

Institut für Physikalische Biologie
(Leitung: Prof. Dr. rer. nat. Dieter Willbold)
Heinrich-Heine-Universität Düsseldorf

**Addressing the multifactorial risk-factors of Schizophrenia using
combination of biophysical, biochemical, and societal approaches**

Habilitationsschrift

(kumulatives Verfahren)

zur Erlangung der Venia Legendi

für das Fach Biophysik

Mathematisch-Naturwissenschaftliche Fakultät

Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Dr. rer. nat. Abhishek Arun Cukkemane

June 2025

Preface

This cumulative thesis is based on my original research contributions (indicated with * as corresponding author) in schizophrenia research. The narrative begins with a general introduction to the subject of schizophrenia and outlines the central research questions addressed in this work. The crux of the dissertation is divided into three main sections, each comprising a thematic overview, selected original articles (mentioned below), followed by a critical perspective and discussion of future directions. The concluding sections summarize the overall research activities and highlight my academic and scientific qualifications. I hope this thesis serves as a valuable resource for readers seeking to deepen their scientific understanding and exploration of schizophrenia and related chronic mental illnesses.

1. **Cukkemane A***, Becker N, Kupreichyk T, Heise H, Willbold D, Weiergräber OH (2025) Snapshots of the aggregation pathway of the scaffold protein DISC1 and the implications for chronic mental illnesses: structure mirrors function. *Journal of Structural Biology*: X. <https://doi.org/10.1016/j.yjsbx.2025.100128>
2. **Cukkemane A***, Dingley AD, Mohrlüder J, Santiago-Schübel B, Weiergräber OH, Willbold D (2025) DISC1 mutants are functionally active but structurally unstable to choreograph their tasks. *European Journal of Pharmaceutical Sciences*. <https://doi.org/10.1016/j.ejps.2025.107148>
3. **Cukkemane A***, Hurelemann R, Sanchez, VG (2025) Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. *Journal of Affective Disorders Reports*. <https://doi.org/10.1016/j.jadr.2025.100893>
4. **Cukkemane, A**, Elmenhorst D and Elmenhorst E-M (2023) Biomarkers for the diagnosis of sleep dis-orders. Reference number – PCT/EP2024/073111 https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2025040590&_cid=P11-MBI9SA-80557-1
5. **Cukkemane, A** and Willbold, D (2022) DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia and other chronic mental disorders (CMDs). Reference number - PCT/EP2023/066796 https://patentscope.wipo.int/search/en/detail.jsf?docId=EP418639163&_fid=WO2023247625
6. **Cukkemane A***, Becker N, Zielinski M, Frieg B, Lakomek NA, Heise H, Schröder GF, Willbold D, Weiergräber OH. (2021) Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl Psychiatry* 11, 639. <https://doi.org/10.1038/s41398-021-01765-1>

And one in review

Cukkemane A*, Sanchez, VG (IN REVISION) Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. *Journal of Affective Disorders Reports*

And two scientific articles in preparation (*corresponding author)

1. **Cukkemane A***, Bochinsky, K, Lange D, Hennecke E, Foerges AL, Beer S, Bauer A, Aeschbach D, Willbold D, Kastenmüller G, Drzezga A, Jordan J, Elmenhorst D, Elmenhorst EM (IN PREPARATION) Toward precision diagnostics for circadian disruption: qNMR metabolomics and metabolic dysregulation
2. **Cukkemane A***, Postin D, Bochinsky K, Scheele D, Willbold D, Hurlemann R (IN PREPARATION) Reconstructing schizophrenia's biochemistry through hybrid diagnostics: clinical and metabolic profiling

Contents

1 Introduction - Schizophrenia: A complex multi-factorial psychiatric disorder	1
1.1 Population Survey: Societal outlook and awareness	1
1.2 Structure-function relationship of DISC1 C-region in context of CMIs	4
1.3 Biochemical and metabolic pathways in Schizophrenia	6
1.4 Loneliness and being alone, the psychosocial elements	9
 2. The big question: Is a comprehensive, multi-disciplinary strategy feasible for managing schizophrenia?	 13
 3 Understanding the viewpoint of various stakeholders on the use of digital technology for the diagnosis of schizophrenia and related CMIs	 15
3.1. Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. <i>Journal of Affective Disorders Reports</i> . https://doi.org/10.1016/j.jadr.2025.100893	18
3.2 Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. <i>Journal of Affective Disorders Reports</i>	30
3.3 Perspective: The challenges, responsibilities, and benefits for society in addressing schizophrenia	66
 4 DISC1 (dys)function and its implications for chronic mental illnesses	 70
4.1. Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. <i>Transl Psychiatry</i> 11, 639. https://doi.org/10.1038/s41398-021-01765-1	76
4.2. Snapshots of the aggregation pathway of the scaffold protein DISC1 and the implications for chronic mental illnesses: structure mirrors function. <i>Journal of Structural Biology: X</i> . https://doi.org/10.1016/j.yjsbx.2025.100128	98
4.3 DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia and other chronic mental disorders (CMDs). Reference number - PCT/EP2023/066796 https://patentscope.wipo.int/search/en/detail.jsf?docId=EP418639163&_fid=WO2023247625	116

4.4	DISC1 mutants are functionally active but structurally unstable to choreograph their tasks. <i>European Journal of Pharmaceutical Sciences</i> . https://doi.org/10.1016/j.ejps.2025.107148	166
4.5	Perspective: Targeting DISC1 for the treatment of psychiatric disorders	187
5	Biochemical pathways and metabolomics in understanding schizophrenia	191
5.1.	Reconstructing schizophrenia's biochemistry through hybrid diagnostics: clinical and metabolic profiling	194
5.2.	Biomarkers for the diagnosis of sleep disorders. Reference number – PCT/EP2024/073111 https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2025040590&_cid=P11-MBI9SA-80557 -	236
5.3	Toward precision diagnostics for circadian disruption: qNMR metabolomics and metabolic dysregulation	
5.4	Perspective: Developing diagnostic tools through insights into aberrant biochemical mechanisms	305
6.	Summary	307
7.	Publication and Achievements	308
8.	Acknowledgements	313
9.	Curriculum Vitae	315

Abbreviations

ϵ	Absorption coefficient
λ	wavelength
Å	Angstrom
°C	Degrees Celsius
ADHD	Attention deficit hyperactivity disorder
AFM	Atomic Force Microscopy
BD	Bipolar disorder
BLI	Bio-layer Interferometry
CD	circular dichroism
CMI	Chronic mental illnesses
Da	Dalton
DISC1	Disrupted in schizophrenia 1
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EM	Electron Microscopy
e.g.	for example
FPLC	fast protein liquid chromatography
GF	Gel-Filtration
h	hours
His	Histidine
HMW	high molecular weight
i.e.	that is
IPTG	isopropyl- β -D-thiogalactoside
IDP	Intrinsically Disordered Protein
IDR	Intrinsically Disordered Region
ITC	Isothermal Titration Calorimetry
MAS	Magic Angle Spinning
NLP	Natural Language Processing
NMR	Nuclear Magnetic Resonance spectroscopy (solution-state)
kDa	kilodalton

Abbreviations

l	Litre
LB	Luria-Bertani
LMW	low molecular weight
m	Metre
M	Molar
mAU	milli absorbance units
MDD	Major depressive disorder
min	minute
MS	mass spectrometry
MW	molecular weight
NRC	National Research Council Canada
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PD	Psychiatric disorder
QCA	Qualitative content analysis
s	seconds
SDS	sodium dodecylsulfate
SPR	Surface Plasmon Resonance
SEC	size exclusion chromatography
ssNMR	Solid-state Nuclear Magnetic Resonance spectroscopy
TE	Tris-EDTA
ThT	Thiflavin T
Tris	Tris(hydroxymethyl)-aminomethane
TBS	Tris buffered saline
UV	Ultraviolet
v/v	volume by volume
w/v	weight by volume
w.r.t	with regards to

Abbreviations

Abbreviation of Amino acid residues

A	Alanine	Ala
R	Arginine	Arg
N	Asparagine	Asn
D	Aspartate	Asp
C	Cysteine	Cys
E	Glutamate	Glu
Q	Glutamine	Gln
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
L	Leucine	Leu
K	Lysine	Lys
M	Methionine	Met
F	Phenylalanine	Phe
P	Proline	Pro
S	Serine	Ser
T	Threonine	Thr
W	Tryptophan	Trp
Y	Tyrosine	Tyr
V	Valine	Val

1. Introduction - Schizophrenia: A complex multi-factorial psychiatric disorder

Chronic mental illnesses (CMIs), such as schizophrenia and related affective disorders are poorly understood psychiatric disorders because of their complex and multifaceted aetiology, which involve a dynamic interaction of biological, environmental, and social factors. As a researcher, I am invested in understanding the various faces of schizophrenia, and have always believed that scientific progress cannot exist in isolation from society. Schizophrenia is not just a clinical diagnosis but is a societal illness (Organization, 2022b, 2023; van Os & Kapur, 2009). Its implications effect far beyond the affected individual, but also challenges living in families, communities, health systems, and public policy. A recent meta-analysis ((NIMH), 2022)conducted by the National Institute of Mental Health (NIMH) and the National Institutes of Health (NIH) helped frame the global context of the disease, and the findings were both sobering and motivating. The analysis revealed that schizophrenia ranks among the top 15 leading causes of disability worldwide. Individuals diagnosed with the condition are not only prone to premature mortality with an average of 28.5 years of life lost per patient in the U.S. with approximately 5% die by suicide. The economic toll is overwhelming with financial costs spanning both direct medical expenditures and indirect losses due to reduced productivity, unemployment, and societal disengagement.

This prompted me to explore how these findings translate to the local context of Germany and/or for the EU. I was particularly curious to investigate the societal perspective, the burden of diagnosis, and awareness among various stakeholders, especially given the country's strong healthcare infrastructure and scientific community. I wanted to understand the landscape from the ground up how research, diagnosis, and regulation intersect with public perception, emotional toll, and the feasibility of implementing new approaches for medical care.

1.1. Population Survey: Societal outlook and awareness

I designed the study to be immersive and dialog-driven, focusing on both direct and emotional stakeholders. My interviewees (Figure 1) included clinicians, researchers, personnel from pharmaceutical and diagnostic companies, social workers, and family members of patients to name a few. I also created an online questionnaire available in both German and English for voluntary participation, where I have been collecting data for the last 4 years.

https://docs.google.com/forms/d/1k2Y9E6kRGk1eUgno8uCxh5kl9nIb0L3UX0TY_vynv4w/

https://docs.google.com/forms/d/1AH_E7iS4XLbfv4nGTIdZk2LsPhifDBDLCberqa4N8Vs/

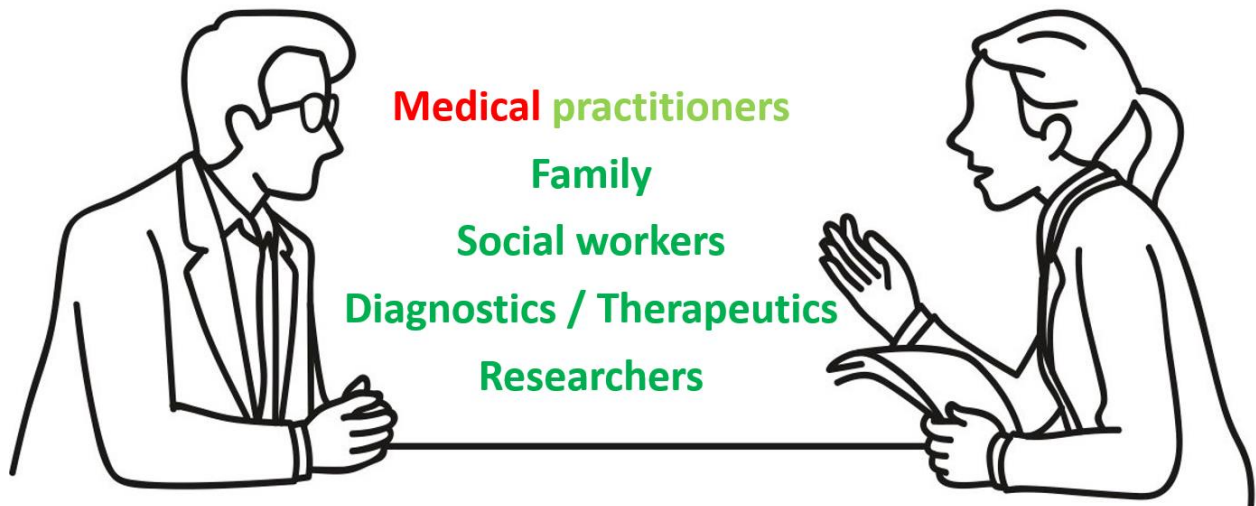


Figure 1: Among the interviewed stake holders, the people who were keen to embrace new technology and were keen on gathering more knowledge and had overall positive outlook in modern-scientific outlook towards schizophrenia research are denoted in green. The medical practitioners could be differentiated into encouraging as well as reluctant (hence in red fonts) on the possibility of biochemical implications to schizophrenia pathology.

The interviews themselves were highly informative, yet emotionally resonant, where stakeholders with a progressive attitude toward technology and a curiosity to understand schizophrenia's scientific basis are marked in green. Some of the most prominent insights that emerged from these discussions are worth highlighting:

- There was near-unanimous agreement on the inadequacy of current diagnostic procedures, which are largely empirical and based on psychiatric interviews and symptom checklists like those in ICD (Organization, 2022a) or DSM-V (Association, 2013). This lack of a non-invasive, objective biomarker left many clinicians and researchers feeling frustrated and constrained.
- Family members expressed that the diagnostic journey was exhausting, not just emotionally, but also financially and logistically. Many spoke about the multiple appointments, travel, lost workdays, and the emotional toll of navigating an often-ambiguous diagnostic landscape.
- Among social workers, burnout and emotional fatigue were common themes. Several confessed that the emotional heaviness of the work led them to seek different career paths. They felt under-resourced and inadequately supported in the current system.
- Representatives from diagnostic and pharmaceutical companies shared their frustration with the complex bureaucratic and regulatory approvals in the EU (IVDR). These procedures are overwhelming and constraints innovation, which not only poses hurdles but are particularly discouraging for start-ups and young innovators, who found the compliance requirements too burdensome to pursue novel diagnostics in schizophrenia.
- An interesting observation was the clear generational divide among medical professionals. Younger clinicians were much more receptive to the idea that schizophrenia has a biochemical

basis and were enthusiastic about integrating molecular and metabolic insights into diagnostic frameworks. In contrast, many older practitioners were sceptical or even dismissive of this direction, holding firmly to established diagnostic protocols.

One of the key lessons from this experience was the critical importance of scientific communication and stakeholder engagement. Except for one reluctant group, i.e., the older clinicians, most of the other stakeholders I spoke with were eager to gain deeper insights into schizophrenia. Their motivations ranged from scientific curiosity and professional growth to the desire for better patient care and societal well-being. It reinforced my belief

- I need to continue engaging with stakeholders on a regular basis to understand how the societal illness is evolving and how may I/we scientist(s) contribute to the ailment. Details of societal and health sciences research based on stakeholder interviews are provided in Section 3 of the thesis
- I have to rationalize and develop newer drug development strategies as there has been a declining patent portfolio since the last 20 years (Figure 2), especially in small-molecule drugs (Section 4), for which deeper understanding of schizophrenia pathophysiology is a must.

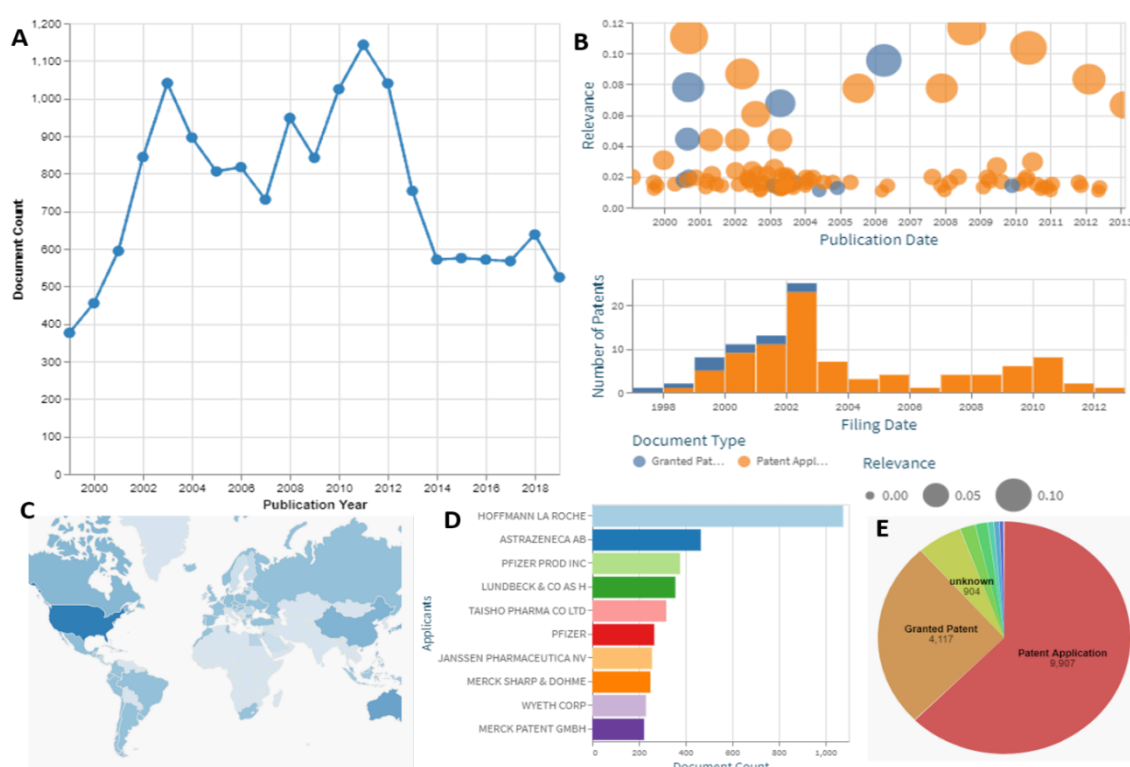


Figure 2: Pictorial representation of the state of therapeutics for schizophrenia and related patents in terms of yearly publications (A), relevance for therapy (B), jurisdiction of the patents (C), Applicants during period A (D) and a pie chart highlighting the success of the application (E); Image was generated using Lens.org using the fooling Boolean operators and filters (Title: schizophrenia OR (Abstract: schizophrenia OR Claims: schizophrenia); Publication Date = (Jan 1, 1999 - Dec 31, 2019): IPCR Classifications = (A61p25/18 (Antipsychotics, I.e. Neuroleptics; Drugs For Mania Or Schizophrenia), A61p25/28 (For Treating Neurodegenerative Disorders Of The Central Nervous System, E.g. Nootropic

Agents, Cognition Enhancers, Drugs For Treating Alzheimer's Disease), A61p25/16(Anti-Parkinson Drugs), A61p25/14(For Treating Abnormal Movements, E.g. Chorea, Dyskinesia)}

- I believe nuclear magnetic resonance (NMR) spectroscopy-based metabolomics experiments (section 5) represents a wonderful approach to address most of the queries of the various stakeholders because the technique offers a powerful approach to study day-to-day physiological changes by analysing the small molecule metabolites in biological samples providing insights into metabolic pathways and how they respond to various stimuli or conditions, and the environment, thereby addressing the unmet needs of various stakeholders

In conclusion, schizophrenia's burden is immense, multifaceted, and deeply felt across various strata of society. The interviews and surveys, I have reported above were conducted in Germany (later across EU stakeholders (Cukkemane, 2025 IN REVIEW; Abhishek Cukkemane, Rene Hurlmann, et al., 2025), see section 3) reaffirmed the findings of the NIMH-NIH meta-analysis. I hope that through my continued dialogue, fieldwork, and lab research, we can move closer to a future where schizophrenia is not just better understood, but more compassionately and accurately treated.

1.2 Structure-function relationship of DISC1 C-region in context of CMIs

The next aspect of my research activity (Section 4) revolves around molecular underpinnings for pathophysiology schizophrenia and related CMIs, I have become particularly interested in the protein Disrupted in schizophrenia 1 (DISC1). This protein has been repeatedly implicated as a major genetic risk factor in not only schizophrenia but also in bipolar disorder and other CMIs (Blackwood et al., 2001; Millar et al., 2000). What initially drew my attention to DISC1 was the extent of its interaction network—it interfaces with more than 300 known proteins (Bradshaw & Korth, 2019; Tropea et al., 2018; Yerabham et al., 2013) involved in diverse physiological processes, from cell cycle regulation to synaptic development. These characteristics make DISC1 a potential central player in the molecular aetiology of CMIs.

What is rather fascinating is that DISC1 functions as a scaffold protein, orchestrating the assembly of multi-protein complexes with clinical and therapeutic relevance. Despite its central role in various cellular pathways, our understanding of its structure-function relationship remains limited due to the lack of a high-resolution 3D structure of DISC1.

Briefly, DISC1 (Soares et al., 2011) is an 854-amino-acid-long protein (Figure 3) that has two principal domains in the N-terminal and the C-terminal tail regions. and sequence-based bioinformatic analyses have provided important insights into its structural organization. These analyses suggest that the N-terminal portion (~1–350 aa) is predominantly disordered, while the C-terminal region (~350–854 aa) contains helical coiled-coil motifs, which are likely responsible for its tendency to oligomerize and, in

some pathological contexts, to aggregate (Yerabham et al., 2017). This structural disorder is predicted to confer flexibility, potentially enabling simultaneous or sequential interactions with multiple protein targets (Soares et al., 2011). I found this dual nature of structural disorder and order particularly compelling, as it underlies DISC1's multifunctionality and dynamic interaction capacity, and is detailed extensively in section 4.

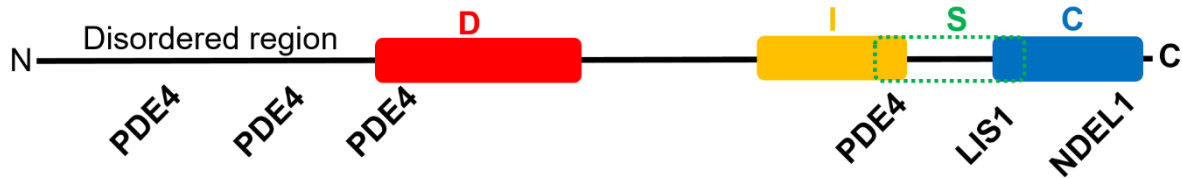


Figure 3: Schematic representation of the DISC1 protein; the D, I, S and C regions and the proposed binding sites for selected interaction partners have been annotated.

The tail domain, in contrast shows some degree of ordered structured elements, which are mnemonically named D, I, S and C regions and comprises of residue 257-383, 539-655, 635-738 and 691-836 (Yerabham et al., 2017). It comprises a series of α -helical regions with strong coiled-coil propensities, i.e., is characterized by the presence of heptad repeats (commonly denoted *abcdefg*, with positions *a* and *d* carrying hydrophobic residues and others being either polar or charged). I find this structured nature particularly intriguing because it allows the different domains of the DISC1 to self-associate. Previous experimental studies have demonstrated that the region on the potential of DISC1's ability to form multimers (dimers, tetramers, octamers) and aggregates (Kamiya et al., 2005; Leliveld et al., 2008; Leliveld et al., 2009; Narayanan et al., 2011; Yerabham et al., 2017). These studies, literally and metaphorically served as modular building blocks for my research activities w.r.t DISC1, particularly the C-region to explore further the structure, and architecture in context to its dys(function) in section 4.

One recurring theme in my exploration of DISC1 is its characterization as a scaffold hub within interaction networks (Soares et al., 2011; Tropea et al., 2018; Yerabham et al., 2013). Among DISC1's extensive list of interactors are several established psychiatric risk proteins (Bradshaw & Hayashi, 2017; Carlyle et al., 2011; Tropea et al., 2018), including isoforms of Phosphodiesterase 4 (PDE4B1 and PDE4D3), Platelet-Activating Factor Acetylhydrolase IB Subunit Alpha (PAFAH1B, formerly LIS1), Activating Transcription Factor 4 (ATF4), Growth factor receptor-bound protein 2 (Grb2), Glycogen Synthase Kinase (GSK3 β), Nuclear Distribution Element 1 (NDE1), and Nuclear Distribution Element-like 1 (NDEL1). Several

lines of evidences (Leliveld et al., 2008; Leliveld et al., 2009; Narayanan et al., 2011) suggests that C-region of the DISC1 is central interaction with proteins from the mitotic spindle complex such as NDEL1. However, whether the physiological unit represents a monomer or multimer remained elusive. This contradiction necessitated the need for detailed structural and functional studies, which was a major part of my research endeavour (see section 4, (Cukkemane et al., 2021)).

Another recurring theme revolves around the tendency of DISC1 aggregation, also referred to as DISC1opathies (Korth, 2012). Aggregated DISC1 has been identified in post-mortem brain samples from patients with schizophrenia and mood disorders (Leliveld et al., 2008), suggesting that misfolding and aggregation may be directly linked to disease pathology. I am particularly interested in characterizing the biophysical properties that govern this aggregation and understanding how it might be pharmacologically modulated (Section 4, (Abhishek Cukkemane, Nina Becker, et al., 2025; A. Cukkemane et al., 2025; Cukkemane, 2022)).

Given the central role of DISC1 in multiple regulatory networks and its implications in psychiatric disorders, I intend to systematically explore its structure-function relationship using of the C-region combination of biophysical and structural biology techniques that includes (detailed in section 4)

- Map interaction motifs along DISC1 C-regions length,
- Understand how multimerization influences binding specificity,
- Characterize the biochemical conditions that promote aggregation,
- Structure-rationale in drug development.

1.3 Biochemical and metabolic pathways in Schizophrenia

As mentioned above, schizophrenia is a complex, multi-aetiological CMI that affects around 1% of the global population (van Os & Kapur, 2009). What boggles me the most about schizophrenia is not only its impact on individuals and families but also the profound gaps that still exist (also see section 3, (Abhishek Cukkemane, Rene Hurlemann, et al., 2025)) in our ability to diagnose it early and accurately. For these reasons, along with development of therapeutics, I pursued the biochemical underpinnings of schizophrenia with a particular focus on metabolomics (Figure 4, (Vignoli et al., 2019; Wishart, 2019)) for both diagnostics as well as a read-out for drug-development strategies. The primary objective of my current research is to understand the pathophysiological mechanisms of this disorder at a molecular level. I am especially interested in the metabolic signatures that may offer insights into the physiological status of individuals affected by schizophrenia. By applying metabolomics, and more specifically quantitative nuclear magnetic resonance (qNMR) spectroscopy, I aim to identify biomarkers that could support diagnostic frameworks for psychiatric disorders.

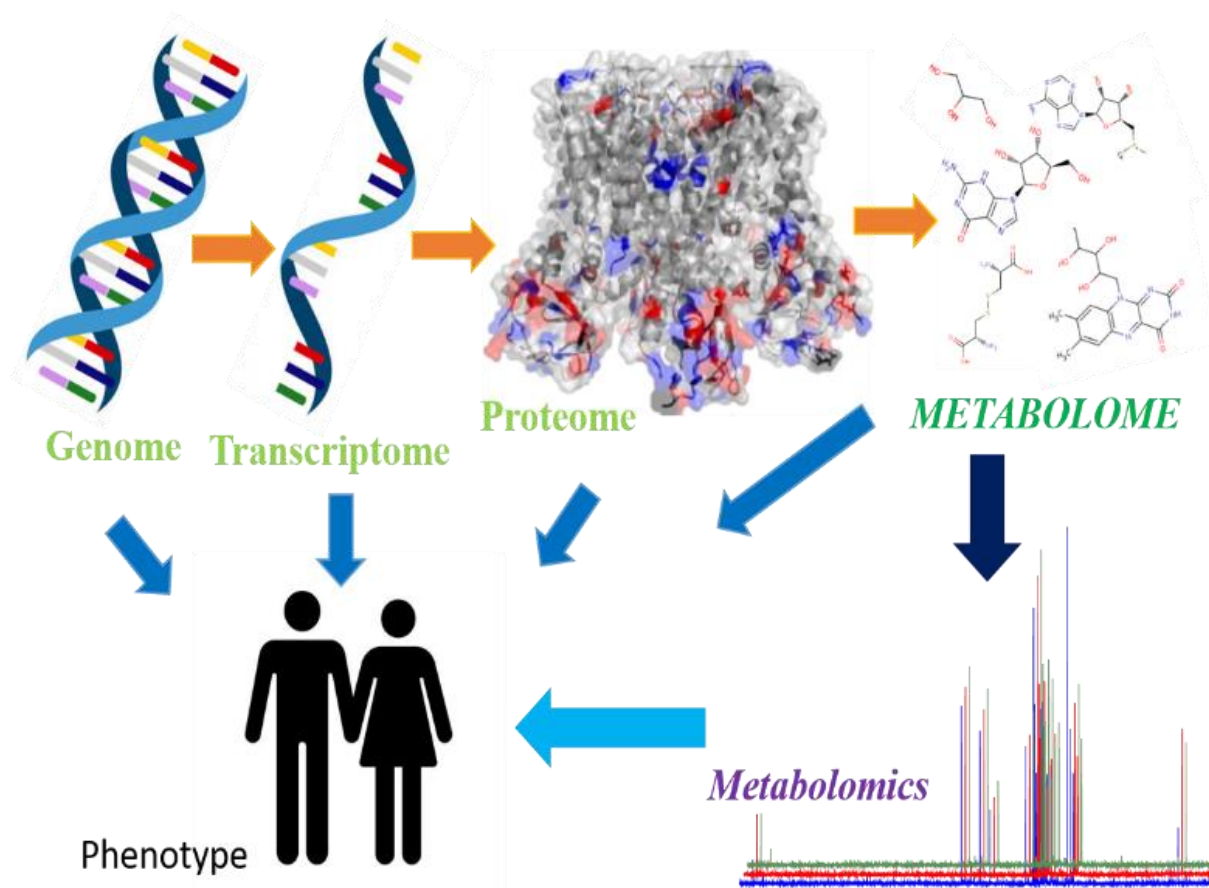


Figure 4: Different forms of functional genomics techniques including metabolomics investigation, which involves study of small molecules ($>1.5\text{kDa}$) in a biological sample

One of the most appealing aspects of qNMR metabolomics is its ability to provide a comprehensive, reproducible, and quantitative snapshot of the metabolic state of a biological system. At the current stage, I am continually generating metabolomics datasets (library) using qNMR, focusing primarily on biological fluids such as blood and urine. These datasets are being analysed using with advanced data science techniques, including chemometrics, statistical modelling (uni- and multi- variate), and machine learning approaches. This integrative approach enables my group and collaborators to explore subtle biochemical changes that may be associated with the onset or progression of schizophrenia.

The reason this endeavour is particularly relevant is the current lack of objective laboratory diagnostics for schizophrenia. Clinicians today still rely heavily on symptom-based assessment tools such as the DSM (Association, 2013) and ICD (Organization, 2022a) which, although valuable, often lead to delayed and mis- diagnoses (Korth & Fangerau, 2020). This delay, which can span several months complicates treatment strategies and may even contribute to disease progression. There is a clear need to develop adjunct diagnostic methods that are rooted in objective molecular data to aid psychiatrists in early identification and intervention. Toward this end, we are developing diagnostic methodologies for schizophrenia (Section 5) and associated symptoms such as sleep disorder (Cukkemane, 2023) through high-throughput qNMR-based metabolomic profiling. The goal is not just to identify biomarkers but to

validate them within clinically relevant settings, accounting for the diversity of populations and the variability in environmental factors.

Why Metabolomics? Metabolomics has long been acknowledged for its potential for diagnostics, especially for systemic diseases (Wishart, 2019; Wishart et al., 2007). However, the technology has been commercially exploited for over a decade for food (honey (Pu et al., 2025), wine (Amargianitaki & Spyros, 2017), juice (Lv et al., 2025), dairy products (Cukkemane et al., 2020) etc) but its potential in medical care is only recently being realized. At the moment, metabolomics is a rapidly emerging technology, which has seen clinical applications by Bruker Inc. following EU-IVDR approval for chemometric platforms. On similar lines, UK-biobank in collaboration with Finnish start-up ‘Nightingale health’ employs NMR metabolomics on blood samples (Julkunen et al., 2023) on cohort of 500,000 participants.

As a short diversion, among the techniques available for metabolomic analysis, mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy are the most widely used. MS is renowned for its sensitivity, being 100 to 1000 times more sensitive than NMR (Vignoli et al., 2019). However, this sensitivity comes at a cost because the method is qualitative (semi-quantitative at best), and requires labour-intensive sample preparation. In contrast, NMR spectroscopy, particularly in its quantitative form (qNMR), offers several advantages in minimal sample preparation, high reproducibility, and, crucially, the ability to deliver absolute quantification. These features make qNMR especially well-suited for clinical applications where consistency and standardization are paramount. Another compelling reason for focusing on NMR technology is its compatibility with existing hospital infrastructure. Many, if not most of the hospitals are already equipped with magnetic resonance imaging (MRI) systems and have staff that are familiar with magnetic resonance techniques. Since NMR shares a similar technological foundation with MRI, its integration into clinical workflows can be less disruptive, relatively seamless, requiring minimal additional training or investment.

Reverting back to the study described in this thesis, understanding the metabolome of a schizophrenic patient offers us a lot more than mere diagnostic insight because it opens the door to personalized medicine. By comparing patient-specific profiles with control data, we hope to not only detect schizophrenia early but also stratify patients based on the severity of metabolic disruptions (Section 4). Such stratification could be invaluable for tailoring treatment plans or monitoring therapeutic responses. Based on the feedback I have received from the various stakeholders (Section 3), such as family members of patients, social workers, clinicians, and researchers, who have expressed their desire in tools that could reduce diagnostic delays and improve care delivery. I hope that this work is not just as a scientific pursuit, but results in a real-world application with potential to improve lives. To this extent, I have a patent application filed for diagnosing one of the symptoms in sleep disorder (Cukkemane, 2023) and am underway to filing two more by the end of this year.

As we move forward, we will continue monitoring the evolving landscape of mental health diagnostics based on stake-holder interviews, evolution of society, and adapt our methodologies accordingly. We are also refining our statistical models using large-scale datasets and integrating multi-omics layers to enhance the robustness of our findings. In short, one of the missions here is to bridge the gap between molecular understanding and clinical reality in psychiatric medicine.

1.4 Loneliness and being alone, the psychosocial elements

As repeatedly mentioned above, schizophrenia is a debilitating psychotic disorder that affects 1% of the global population [1, 2]. The aetiology of this complex disease remains challenging and is affected by an interplay of various factors including biological, social and psychosocial / environmental conditions. Given this background, it is not surprising that loneliness is especially severe in schizophrenia. Individuals with schizophrenia are subject to stigma and have greater clinical (especially negative and cognitive symptoms, etc.) challenges. In addition, on average, individuals with schizophrenia experience greater socio-environmental difficulties (e.g., poverty, low rates of employment, low rates of marriage) and are objectively less integrated within their communities, although their perceived sense of belonging within the community may not differ from individuals without schizophrenia living in the same community. Recent surveys indicate that self-reported annual rates of loneliness among individuals with schizophrenia and other psychotic disorders (76 to 80%) are approximately 2.3 times higher than those in the general population (35%). Highlighting its clinical importance, loneliness has been reported as a significant contributor to reduced quality of life in schizophrenia, and individuals with psychotic disorders cite loneliness as one of the most important challenges in their life, second only to financial concerns (Eglit et al., 2018). In addition, loneliness has been shown to potentially culminate into depression and suicidal ideation, especially in individuals with schizophrenia (Chen et al., 2023; Pitman et al., 2020). The lifetime prevalence of suicidal thoughts in the world population is between 7-10% and rises to 35% in those affected from schizophrenia (Chen et al., 2023; Pitman et al., 2020), underscoring the devastating impact of loneliness.

While loneliness is generally associated with increased morbidity and mortality, it is particularly severe in schizophrenia. Therefore, there is an urgent need to improve treatment options for patients with schizophrenia, focusing on their loneliness as a novel, innovative and promising approach to foster recovery on the one hand and prevent disease chronification and suicide on the other. As this project is at its infancy and we are recruiting candidates, I do not have any data to share but would like to highlight the holistic approach I am implementing in understanding the multi-factorial complexity of the disorder. In this case, in nutshell, the scientific inkling involves performing qNMR metabolomics on blood and urine samples and tally them with UCLA Loneliness Scale (Eglit et al., 2018) to compare clinical diagnosis with biomarker analytics.

References

- (NIMH), N. I. o. M. H. (2022). *Schizophrenia*. U.S. Department of Health and Human Services. https://www.nimh.nih.gov/health/statistics/schizophrenia#part_2544
- Amargianitaki, M., & Spyros, A. (2017). NMR-based metabolomics in wine quality control and authentication. *Chemical and Biological Technologies in Agriculture*, 4(1), 9. <https://doi.org/10.1186/s40538-017-0092-x>
- Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*, Arlington VA, American Psychiatric Association.
- Blackwood, D. H. R., Fordyce, A., Walker, M. T., St Clair, D. M., Porteous, D. J., & Muir, W. J. (2001). Schizophrenia and affective disorders - Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: Clinical and P300 findings in a family. *American Journal of Human Genetics*, 69(2), 428-433. <https://doi.org/Doi 10.1086/321969>
- Bradshaw, N. J., & Hayashi, M. A. (2017). NDE1 and NDEL1 from genes to (mal)functions: parallel but distinct roles impacting on neurodevelopmental disorders and psychiatric illness. *Cell Mol Life Sci*, 74(7), 1191-1210. <https://doi.org/10.1007/s00018-016-2395-7>
- Bradshaw, N. J., & Korth, C. (2019). Protein misassembly and aggregation as potential convergence points for non-genetic causes of chronic mental illness. *Mol Psychiatry*, 24(7), 936-951. <https://doi.org/10.1038/s41380-018-0133-2>
- Carlyle, B. C., Mackie, S., Christie, S., Millar, J. K., & Porteous, D. J. (2011). Co-ordinated action of DISC1, PDE4B and GSK3beta in modulation of cAMP signalling. *Mol Psychiatry*, 16(7), 693-694. <https://doi.org/10.1038/mp.2011.17>
- Chen, Y. L., Jian, C. R., Chang, Y. P., Chao, S. R., & Yen, C. F. (2023). Association of loneliness with suicide risk and depression in individuals with schizophrenia: moderating effects of self-esteem and perceived support from families and friends. *Schizophrenia (Heidelb)*, 9(1), 41. <https://doi.org/10.1038/s41537-023-00368-7>
- Cukkemane, A., & Sanchez, V. G. (2025 IN REVIEW). Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. *Journal of Affective Disorders Reports*.
- Cukkemane, A., Becker, N., Kupreichyk, T., Heise, H., Willbold, D., & Weiergräber, O. H. (2025). Tracing the aggregation pathway of the scaffold protein DISC1: Structural implications for chronic mental illnesses. *Journal of Structural Biology: X*, 11, 100128. <https://doi.org/https://doi.org/10.1016/j.yjsbx.2025.100128>
- Cukkemane, A., Becker, N., Zielinski, M., Frieg, B., Lakomek, N. A., Heise, H., Schroder, G. F., Willbold, D., & Weiergräber, O. H. (2021). Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl Psychiatry*, 11(1), 639. <https://doi.org/10.1038/s41398-021-01765-1>

- Cukkemane, A., Dingley, A. J., Mohrluder, J., Santiago-Schubel, B., Weiergraber, O. H., & Willbold, D. (2025). A peptide mimetic therapeutic strategy targeting dysfunction of the scaffold protein DISC-1 in psychiatric disorders. *Eur J Pharm Sci*, 107148. <https://doi.org/10.1016/j.ejps.2025.107148>
- Cukkemane, A., Elmenhorst, D., Elmenhorst, E-M (2023). *Biomarkers for the diagnosis of sleep disorders* (Germany Patent No. EPO).
- Cukkemane, A., Hurlmann, R., & Sanchez, V. G. (2025). Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. *Journal of Affective Disorders Reports*, 20, 100893. <https://doi.org/https://doi.org/10.1016/j.jadr.2025.100893>
- Cukkemane, A., Kumar, P., & Sathyamoorthy, B. (2020). A metabolomics footprint approach to understanding the benefits of synbiotics in functional foods and dietary therapeutics for health, communicable and non-communicable diseases. *Food Research International*, 128. <https://doi.org/ARTN10867910.1016/j.foodres.2019.108679>
- Cukkemane, A., Willbold, D. (2022). *DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDs)* (EPO Patent No).
- Eglit, G. M. L., Palmer, B. W., Martin, A. S., Tu, X., & Jeste, D. V. (2018). Loneliness in schizophrenia: Construct clarification, measurement, and clinical relevance. *PLoS One*, 13(3), e0194021. <https://doi.org/10.1371/journal.pone.0194021>
- Julkunen, H., Cichonska, A., Tiainen, M., Koskela, H., Nybo, K., Makela, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C., & Wurtz, P. (2023). Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*, 14(1), 604. <https://doi.org/10.1038/s41467-023-36231-7>
- Kamiya, A., Kubo, K., Tomoda, T., Takaki, M., Youn, R., Ozeki, Y., Sawamura, N., Park, U., Kudo, C., Okawa, M., Ross, C. A., Hatten, M. E., Nakajima, K., & Sawa, A. (2005). A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nat Cell Biol*, 7(12), 1167-1178. <https://doi.org/10.1038/ncb1328>
- Korth, C. (2012). Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion*, 6(2), 134-141. <https://doi.org/10.4161/pri.18989>
- Korth, C., & Fangerau, H. (2020). Blood tests to diagnose schizophrenia: self-imposed limits in psychiatry. *Lancet Psychiatry*, 7(10), 911-914. [https://doi.org/10.1016/S2215-0366\(20\)30058-4](https://doi.org/10.1016/S2215-0366(20)30058-4)
- Leliveld, S. R., Bader, V., Hendriks, P., Priekulis, I., Sajnani, G., Requena, J. R., & Korth, C. (2008). Insolubility of disrupted-in-schizophrenia 1 disrupts oligomer-dependent interactions with

- nuclear distribution element 1 and is associated with sporadic mental disease. *Journal of Neuroscience*, 28(15), 3839-3845. <https://doi.org/10.1523/JNEUROSCI.5389-07.2008>
- Leliveld, S. R., Hendriks, P., Michel, M., Sajjani, G., Bader, V., Trossbach, S., Prikulis, I., Hartmann, R., Jonas, E., Willbold, D., Requena, J. R., & Korth, C. (2009). Oligomer assembly of the C-terminal DISC1 domain (640-854) is controlled by self-association motifs and disease-associated polymorphism S704C. *Biochemistry*, 48(32), 7746-7755. <https://doi.org/10.1021/bi900901e>
- Lv, M., Liu, X., Liu, R., Aihaiti, A., Hong, J., Zheng, L., Xing, J., Cui, Y., & Wang, L. (2025). Untargeted metabolomics reveals flavor and metabolic changes in mixed *Lactobacillus*-fermented black mulberry juice. *Food Chem X*, 27, 102367. <https://doi.org/10.1016/j.fochx.2025.102367>
- Millar, J. K., Wilson-Annan, J. C., Anderson, S., Christie, S., Taylor, M. S., Semple, C. A. M., Devon, R. S., St Clair, D. M., Muir, W. J., Blackwood, D. H. R., & Porteous, D. J. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Human Molecular Genetics*, 9(9), 1415-1423. <https://doi.org/DOI> 10.1093/hmg/9.9.1415
- Narayanan, S., Arthanari, H., Wolfe, M. S., & Wagner, G. (2011). Molecular characterization of disrupted in schizophrenia-1 risk variant S704C reveals the formation of altered oligomeric assembly. *J Biol Chem*, 286(51), 44266-44276. <https://doi.org/10.1074/jbc.M111.271593>
- Organization, W. H. (2022a). *ICD-11: International classification of diseases (11th revision)*. <https://icd.who.int/>
- Organization, W. H. (2022b). *Schizophrenia*.
- Organization, W. H. (2023). *Mental Health Gap Action Programme (mhGAP) guideline for mental, neurological and substance use disorders*. Geneva: World Health Organization. <https://www.who.int/publications/book-orders>
- Pitman, A. L., King, M. B., Marston, L., & Osborn, D. P. J. (2020). The association of loneliness after sudden bereavement with risk of suicide attempt: a nationwide survey of bereaved adults. *Soc Psychiatry Psychiatr Epidemiol*, 55(8), 1081-1092. <https://doi.org/10.1007/s00127-020-01921-w>
- Pu, J., Liu, Y., Wu, H., Liu, C., Chen, Y., Tang, W., Yu, Y., Gui, S., Zhong, X., Wang, D., Chen, X., Chen, Y., Chen, X., Qiao, R., Jiang, Y., Zhang, H., Ren, Y., Fan, L., Wang, H., & Xie, P. (2025). Characterizing metabolomic and proteomic changes in depression: a systematic analysis. *Mol Psychiatry*, 30(7), 3120-3128. <https://doi.org/10.1038/s41380-025-02919-z>
- Soares, D. C., Carlyle, B. C., Bradshaw, N. J., & Porteous, D. J. (2011). DISC1: Structure, Function, and Therapeutic Potential for Major Mental Illness. *ACS Chem Neurosci*, 2(11), 609-632. <https://doi.org/10.1021/cn200062k>

- Tropea, D., Hardingham, N., Millar, K., & Fox, K. (2018). Mechanisms underlying the role of DISC1 in synaptic plasticity. *Journal of Physiology-London*, 596(14), 2747-2771. <https://doi.org/10.1113/Jp274330>
- van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet*, 374(9690), 635-645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8)
- Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P., & Luchinat, C. (2019). High-Throughput Metabolomics by 1D NMR. *Angew Chem Int Ed Engl*, 58(4), 968-994. <https://doi.org/10.1002/anie.201804736>
- Wishart, D. S. (2019). Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*, 99(4), 1819-1875. <https://doi.org/10.1152/physrev.00035.2018>
- Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., . . . Querengesser, L. (2007). HMDB: the human metabolome database. *Nucleic Acids Research*, 35, D521-D526. <https://doi.org/10.1093/nar/gkl923>
- Yerabham, A. S., Weiergraber, O. H., Bradshaw, N. J., & Korth, C. (2013). Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol Chem*, 394(11), 1425-1437. <https://doi.org/10.1515/hsz-2013-0178>
- Yerabham, A. S. K., Mas, P. J., Decker, C., Soares, D. C., Weiergraber, O. H., Nagel-Steger, L., Willbold, D., Hart, D. J., Bradshaw, N. J., & Korth, C. (2017). A structural organization for the Disrupted in Schizophrenia 1 protein, identified by high-throughput screening, reveals distinctly folded regions, which are bisected by mental illness-related mutations. *J Biol Chem*, 292(16), 6468-6477. <https://doi.org/10.1074/jbc.M116.773903>

2. The big question: Is a comprehensive, multi-disciplinary strategy feasible for managing schizophrenia?

I believe, yes.

Very challenging indeed but I certainly am advocating for a comprehensive, multi-disciplinary strategy for managing schizophrenia given the complexity of the disorder and multifactorial nature. Schizophrenia is not merely some psychiatric disorder, rather it is a societal illness shaped by the intersection of biological vulnerability, environmental stressors, and systemic healthcare challenges. As highlighted in this introduction, no single domain in clinical psychiatry, natural sciences, or social work can independently address the burden of this condition. Therefore, an integrative approach involving experts from all disciplines is vital.

My fieldwork directly involves engagement with clinicians, patients' families, researchers, pharmaceutical enterprises, and diagnostic firms revealed widespread consensus on the limitations of medicare. The frustration expressed by all stakeholders from emotional burnout among social workers to regulatory bottlenecks faced by innovators exemplifies the urgent need for novel, evidence-based, and scalable diagnostic tools and treatment strategies.

This thesis, which has been prepared as part of my habilitation and my aspiration to lead and educate at the highest levels of academia proposes a dual strategy. On one hand, it seeks to investigate the molecular (dys)function of the Disrupted-in-Schizophrenia 1 (DISC1) protein, a central actor in the aetiology of schizophrenia. On the other hand, I explore the use of nuclear magnetic resonance (NMR) spectroscopy-based metabolomics as a promising translational tool that can bridge molecular insights with clinical applications.

By integrating structure-function analyses of schizophrenia-associated risk proteins such as DISC1 with metabolomic profiling, this research aims to map the biochemical pathways underlying schizophrenia and related chronic mental illnesses (CMIs). Accordingly, the crux of this thesis is structured as follows:

Section 3 - Understanding the viewpoint of various stakeholders on the use of digital technology for the diagnosis of schizophrenia and related CMIs

Section 4- DISC1 (dys)function and its implications for chronic mental illnesses

Section 5- Biochemical pathways and metabolomics in understanding schizophrenia

3. Understanding the viewpoint of various stakeholders on the use of digital technology for the diagnosis of schizophrenia and related CMIs.

“Meine Augen zur Zeit der Erscheinungen (My Eyes in the Time of Apparition)” is a famous painting (Figure 1) in 1913 by a world-renowned artist August Natterer, who was a trained engineer turned artist diagnosed with psychotic disorder early in his life. In this case, Natterer believed that he had visions of divine revelation, and omniscience of the soul but clinically the same image represented hallucination, a positive (ly developed) symptom of schizophrenia. The painting strongly embodies how a subject suffering from psychiatric disorder perceives the world.

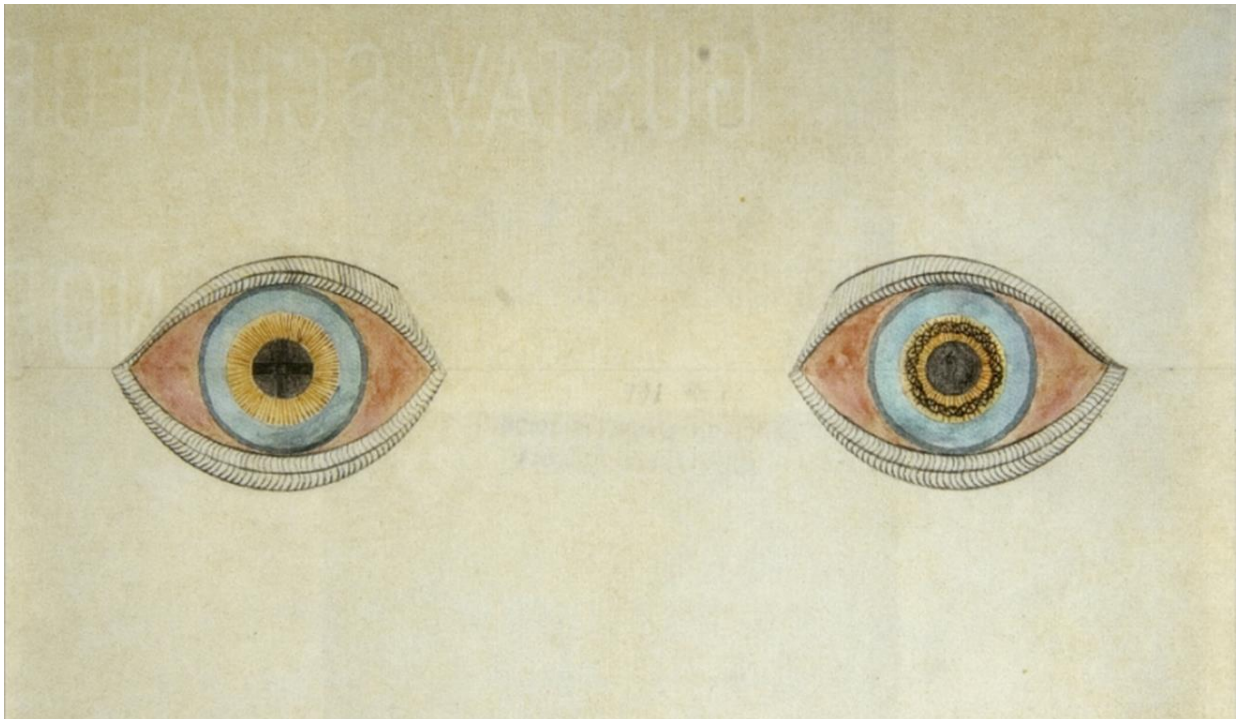


Figure 1: The painting, “My Eyes in the Time of Apparition” reveals the dichotomy reflecting his conflict between vision and delusion, and clarity and chaos. These opposing forces reflect the psychological landscape of a mind oscillating between a search for meaning and the collapse of rational thought.

Psychiatric disorders constitute a significant global health burden. Among them, schizophrenia and other chronic mental illnesses (CMIs) including BD and MDD affect approximately 3–5% of the global population (van Os & Kapur, 2009), yet multitude of challenges persist in medical care. Recently, at the 36th European college of neuropsychopharmacology (ECNP), Barcelona (Neuropsychopharmacology)) 2023), I came across a session devoted to patient opinion and welfare, entitled “Patient session – It takes two to tango: pharmacogenetics in psychiatry (PSY-PGx. Horizon 2020)” that was encouraged by European Psychiatry Association ((association))Global Alliance of Mental Illness Advocacy Network

((GAMIAN-Europe) and European Federation of Associations of Families of People with Mental Illness ((Illness)). The session highlighted the travails patients face during questionnaire-based diagnosis and their opinion is suggestive of feeling of negative emotions, like stress, self-pity, sadness, anger to name a few. While this merely represents the tip of the iceberg, major bottlenecks in current method of diagnosis are in fact demanding and includes:

1. A major facet of the diagnosis, treatment and prognosis is based on a person's attitude and behaviour with symptoms including hallucinations and delusions which also may be caused by substance abuse or another medical condition.
2. The process is complemented with endophenotypic testing in electrophysiology, PET, CT, MRI scans etc, which are expensive as supplementary along with routine blood and urine test for hormonal imbalance, glucose and few other biomarkers
3. The process involves several rounds of sessions to the psychiatrist, who diagnose the mental condition. It can take as low as 1 month for positive symptoms and up to 6 months (or even more) for negative symptoms (DSM-5 Diagnosis: Schizophrenia (Quality, 2016)).

While the angst reflected by the patients is comprehensible, a silver lining in the cloud was highlighted in the form of a collaborative venture between GAMIAN-Europe and the PSY-PGx project, Amsterdam (PSY-PGx) on pharmacogenetics and biomarkers in psychiatry was discussed at length. The patients, the family members and related emotional stakeholders greatly appreciated the value information pertaining their health and recovery.

Inspired by this session, I was curious to understand the perspectives of various stakeholders on schizophrenia and related CMIs, with the aim of exploring the potential for implementing a patient-centric approach to healthcare support. For which, I sought to address three important questions:

1. What are the tasks related to Diagnosis?
2. What are the difficulties in the current procedure?
3. What are the expectations of newer clinical approaches?

These questions are particularly relevant given the current limitations in the health-care environment related to diagnosis of CMIs. Unlike other medical disciplines, psychiatry lacks objective laboratory tests for specific diagnoses such as schizophrenia. Diagnosis is empirical and still (since 1913) relies primarily on descriptive, symptomatic evaluation based on questionnaires that was developed by the renowned psychiatrist Karl Jaspers (Hafner, 2015), which nowadays is outlined in manuals like the ICD-11 (Organization, 2022) and DSM-5 (Bredstrom, 2019). This reliance highlights a critical gap: the absence of reliable and personalized diagnostic tools.

In recent years, however, there has been growing adoption of biomarker-based diagnostics across various branches of medicine. Alongside this, advancements in data analytics, machine learning, and artificial intelligence are enabling more accurate and rapid assessments of individual health states. The identification of biomarkers, both qualitatively and quantitatively (latter sections of this thesis), could offer a foundational basis for more precise diagnosis and treatment of psychiatric conditions. With this in mind, I conducted discussions with a diverse group of stakeholders (excluding patients due to stringent ethical requirements), including clinicians, researchers, and industry professionals. This study was notified with the Data Protection Services of Sikt – Norwegian Agency for Shared Services in education and Research (reference number: 648482). The legal basis for data processing was the GDPR Article 6(1)(a), ensuring compliance with ethical and data protection standards. All participant data was anonymized and written consent was obtained as outlined in Article 5(1)(d) (accuracy) and Article 5(1)(f) (integrity and confidentiality), and Article 32 (security) of the GDPR.

The scientific inkling from our methodology was to gather detailed insights that could inform the development of innovative, individualized approaches to managing CMIs as described articles in the following sub-sections, and to assess their views on the need for biomarker-driven diagnostics and personalized medicine/ therapeutics within a patient-centred framework.

References

- association), E. E. p. <https://www.europsy.net>.
- Bredstrom, A. (2019). Culture and Context in Mental Health Diagnosing: Scrutinizing the DSM-5 Revision. *J Med Humanit*, 40(3), 347-363. <https://doi.org/10.1007/s10912-017-9501-1>
- GAMIAN-Europe. <https://www.gamian.eu>
- Hafner, H. (2015). Descriptive psychopathology, phenomenology, and the legacy of Karl Jaspers. *Dialogues Clin Neurosci*, 17(1), 19-29. <https://doi.org/10.31887/DCNS.2015.17.1/hhaefner>
- Illness), E. E. F. o. A. o. F. o. P. w. M. <https://eufami.org/en>.
- Neuropsychopharmacology), E. E. C. o. <https://www.ecnp.eu/Congress2023/ECNPcongress>.
- Organization, W. H. (2022). *ICD-11: International classification of diseases (11th revision)*. <https://icd.who.int/>
- PSY-PGx. www.psy-pgx.org.
- Quality, C. f. B. H. S. a. (2016). *Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health*. Substance Abuse and Mental Health Services Administration.
- van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet*, 374(9690), 635-645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8)



Research Paper

Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments

Abhishek Cukkemane^{a,b,c,*}, Rene Hurlemann^{d,e}, Veralia Gabriela Sanchez^f

^a Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany

^b Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany

^c Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

^d Department of Psychiatry, School of Medicine and Health Sciences, University of Oldenburg, Oldenburg, Germany

^e Karl-Jaspers-Klinik, Bad Zwischenahn, Germany

^f Faculty of Technology, Natural Sciences and Maritime Sciences, University of South-Eastern Norway, Vestfold, Norway

ARTICLE INFO

Keywords:

Schizophrenia
Chronic mental illnesses
Policy making
Social support
Health care management
Personalized care

ABSTRACT

Chronic mental illnesses (CMIs) such as schizophrenia, schizoaffective disorder, bipolar disorder, major depressive disorders, to name a few are psychiatric disorders poses significant burden to both patients and healthcare systems. The public policy and management of such complex set of spectral disorders often includes procedures and regulations, pharmacological treatments, psychotherapy, and social support. However, these approaches frequently fail to fully address patient requirements, leading to inadequate care and poor outcomes. Therefore, by conducting interviews and synthesizing insights from various stakeholders, including clinicians, researchers, and healthcare professionals, the study investigates the multi-dimensional challenges in CMI management. This report of qualitative analysis explores a patient-centric approach towards three main themes in managing CMIs such as, tasks related to diagnosis, challenges with current procedures, and expectations from newer diagnostic methodologies. Key themes emerged revolving around requirements for emotional care, trust, complexity, and the integration of technology in treatment. Our findings highlight the need for personalized care strategies that encompass biological, psychological, and social dimensions with emphasis on effective communication and resource management. The study underscores the potential of a holistic, patient-centred framework, which serendipitously parallels George Engel's biopsychosocial model to improve outcomes and address the multifaceted nature of spectral psychiatric disorders. In the light of newer emerging technologies and the demands for psychometric intervention, perhaps, we suggest revisiting the model for a comprehensive approach in policy-making, health care management and patient-centric support to address these challenges effectively

1. Introduction

Chronic mental illness (CMI) such as schizophrenia has a complex multi-factorial aetiology that affects approximately 20 million individuals worldwide, with a lifetime prevalence of around 1 %. Along with other related CMIs, such as bipolar disorders (BD), and major depressive disorders (MDD), the number balloons to 3–5 % of the afflicted in the global population (van Os and Kapur, 2009). The management of these are complex, requiring a multifaceted approach involving policy makers, reimbursement policies, health care management and support, and societal integration of the afflicted. Albeit of

several professions participating in the preventive and curative measures for CMIs, current healthcare systems often fall short in providing holistic and personalized care for individuals, leading to suboptimal outcomes and a significant burden on patients, their families, and society.

One of the major associated problems begins at the onset of the entire process, i.e., w.r.t to diagnosis. CMIs are extremely challenging to diagnose albeit of DSM/ICD (Association, 2013; Organization, 2022) guidelines for over 60 years- This is because of overlapping symptomatology between disorders, it is not uncommon to misdiagnose (Smith and Craddock, 2011) bipolar disorder for major depression, and

* Corresponding author.

E-mail address: a.cukkemane@fz-juelich.de (A. Cukkemane).

<https://doi.org/10.1016/j.jadr.2025.100893>

Received 8 November 2024; Received in revised form 15 February 2025; Accepted 18 February 2025

Available online 19 February 2025

2666-9153/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

depression between uni- and bi- polar disorders. In a similar vein, from 1990 to 2020, in 30 years the diagnosis for CMIs improved significantly by 18 % to 34 % for schizophrenia and from 27 % to 46 % for MDD (Paul and Potter, 2024) in Germany yet considerably falls short of effective medical care. The occurrence of misdiagnosis at such high rates is not unsurprising given the similarities in symptomatology of CMIs such as schizophrenia, schizoaffective disorder and BD (van Os and Kapur, 2009) but they can have dire consequences in reducing cases involving young adults with depression who are at high risk of developing BD and influence the choice of treatment approach. For instance, one may employ a more cautious and judicious use of medications, and place a greater emphasis on psychological and non-medical interventions for treatment and recovery.

This has posed a significant challenge on the health care system and the society (Canada, 2012; OECD/EU, 2018). A recent finding of public health insurance and psychiatry hospitals in Germany for patient care with depression have identified deviations and not rigorous implementation of the statutory guidelines resulting in a whopping 21 % cases of remission (Wiegand et al., 2020), indicating the urgent requirement of a newer and less cumbersome outlook to patient care and welfare. Therefore, to understand these limitation, using a combination of semi-structured interview (Kallio et al., 2016) and qualitative research methodology (Malterud, 2001), we probed the requirements of tasks related to diagnosis, challenges with current procedures, and expectations from newer diagnostic methods for a patient-centric approach in mental health care management. The reason we opted for qualitative content analysis (QCA) is because it is proven as a powerful tool to deepen our understanding in medicine for diverse subjects such as nursing (Graneheim and Lundman, 2004; McCormack and McCance, 2006), mental health in veterans (Zinn, 2010), mental health among old people (Anker-Hansen et al., 2019, 2020), attitude towards welfare system and smart houses (Sanchez et al., 2019) to reflect on few. Moreover, qualitative methodology provides deeper complementary insights at microscopic levels that quantitative data alone cannot capture.

Is the approach fruitful in itself? It certainly is worth trying; for instance, in a joint effort by several different stakeholders in clinical, research, and societal backgrounds, experts together proposed a panel of biomarkers from their respective fields for positive patient outcome (Frisoni et al., 2024). A similar collective agreement will prove highly beneficial for management of psychiatric disorder. A recent survey conducted by the Patient-Centered Outcomes Research (PCOR) Institute (Kim et al., 2018) on several stake holders for mental disorders that included patients, caretakers, next-of-kin, advocacy organizations, policy makers to name a few, identified three main aspects: (a) logistics, implementation strategy and the choice of outcome; (b) permitting flexibility in approaches to cater the requirements of health care and subjects; (c) identifying reimbursement protocols and adopting timelines for long-term planning recovery. In a similar vein, we are keen to comprehend the outlook of different stakeholders such as clinicians, researchers, social workers in NGOs, nurses, social scientists, entrepreneurs, pharmaceutical heads to name few who are associated with related CMIs that encompass schizophrenia, schizoaffective, BD, MDD etc. We are keen to understand the view, and identify the prospects for application of new approaches for diagnosis, the unmet needs in patient health care, and the possibilities of risk- management strategies.

The findings of this study highlight the importance of adopting a comprehensive, patient-centric approach to management of schizophrenia and related CMIs that parallels with Engel's biopsychosocial model (Engel, 1978, 1979). Although the model proposed 45 years ago, it represents an important theoretical model that is a part of the curriculum of medical graduates (Roberts, 2023), there has been no experimental evidence suggesting its utility. By interviewing stakeholders, we for the first-time address experimentally that factors of biological/medical, psychosocial/ psychological, and societal/ social factors influence the lives and experiences of individuals involved with

schizophrenia and depression welfare.

2. Experimental procedure

2.1. Aim and research question

Patients and caregivers face increasing challenges due to the complexity of medical support in the management of mental disorders (Anker-Hansen et al., 2019, 2020; Kim et al., 2018). Compounding the challenges include spiralling costs (6, 7), lengthy procedural frameworks, and shortage of medical support and psychotherapists (Hanf et al., 2021) that exacerbates the problem. At the recently concluded 36th European College of Neuropsychopharmacology, Barcelona (Neuropsychopharmacology), a session devoted to patient opinion and welfare, entitled "Patient session – It takes two to tango: pharmacogenetics in psychiatry " highlighted the travails patients face during questionnaire-based diagnosis and their opinion is suggestive of feeling of negative emotions, like stress, self-pity, sadness, anger to name a few. While the angst reflected by the patients is comprehensible, a silver lining in the cloud was highlighted in the form of a collaborative effort by the European Psychiatry Association (EPA); Global Alliance of Mental Illness Advocacy Network (GAMIAN-Europe) and European Federation of Associations of Families of People with Mental Illness (EUFAMI) with PSY-PGx (2024), Amsterdam (PSY-PGx) on pharmacogenetics and biomarkers in diagnosis of psychiatric disorders. The patients, the family members and related emotional stakeholders greatly appreciated the value information pertaining their health and recovery. This narrative reinforced our thinking to understand the benefits, challenges and the possibility to adopt newer technology and approaches for the benefit of the patients using QCA.

2.2. Ethical considerations

An informed consent as prescribed in Article 5(1)(d) accuracy and 5 (1)(f) integrity and confidentiality, as well as Article 32 security was obtained from all participants following the approval from the data protection services of Sikt – Norwegian Agency for Shared Services in Education and Research (reference number - 648482) on a legal framework of General Data Protection Regulation art. 6 nr. 1a.

2.3. Participant recruitment

We recruited individuals involved in health management, support and care of schizophrenia and the related CMIs, which includes their caregivers, and healthcare professionals, including psychiatrists, researchers, social workers, and nurses. Participants were recruited through strategies, that included recruit from the author's network (convenience sampling). At the same time, we requested participation of experts based on their scientific publications using email communications. However, the responses using both the approaches were poor, therefore, snowball sampling was then implemented, which involved word-of-mouth publicity asking knowledgeable people about whom could participate. Altogether, we contacted 57 participants (23 females and 34 males) from countries having social welfare policies that included nations within the EU, Norway, the USA and Canada. From which, a total of 7 participants (Table 1) based in Germany, UK, Norway and Denmark cooperated in the study. Hence, our findings herein will be representative for European countries only. The eligibility criteria included the involvement in the care of individuals with CMIs (for professional caregiver participants), and experience in the management of CMIs (for healthcare professional participants) that includes disorders like schizophrenia and major depression. Also, it is noteworthy that due to the high prevalence of CMIs, which stands at 3-5 % of the global population (van Os and Kapur, 2009), few participants were also caregivers to their family members who were suffering from schizophrenia or depression

Table 1
Participant demographics.

Participant	Sex	Education	Specialization	Current profession	Previous profession	Experience
1	male	MD, PhD	Psychiatry	Scientist	Clinician	>20
2	male	MD, PhD	Neurosciences	Scientist	Clinician*	>15
3	male	PhD, MBA	Biochemistry (Neurosciences)	Entrepreneur (CEO)	Scientist*	>15
4	male	PhD	Medicinal Chemistry	Entrepreneur (COO)	Scientist	>10
5	male	PhD	Psychiatry	BDM Pharmaceutical	Scientist	>20
6	female	MS	Informatics	Social worker (CEO)	IT*	>3
7	male	BSN	Nursing	Nurse	–	>1

* Participants with friends and family members afflicted with CMI; The abbreviations stands for, CEO, chief executive officer; COO, chief operating officer; BDM, business Development manager; and IT, information technology.

2.4. Data collection and analysis

Out of the 7 interviews in English, six were held online and one was held in-person, which lasted 35-50 min. The design of the qualitative study (Malterud, 2001) was prepared with semi-structured interviews (see Table S1 for the questionnaire (Kallio et al., 2016)) with three related frameworks that included tasks related to diagnosis, challenges with current procedures, and expectations from newer diagnostic methods. The interviews were audio-recorded and the transcripts were analysed thematically (Braun and Clarke, 2006). The data was coded using deductive inferences and emerging subcategories, subthemes', and theme were discussed and refined through an iterative process. (Elo and Kyngas, 2008; Graneheim and Lundman, 2004; Krippendorff, 1989) and categorized to multiple sub-themes and a unique theme. The unique theme that aligned with the tasks related to diagnosis, difficulties with current procedures and expectations of newer technologies that would provide vital clues in the pros and cons of the current medical setup for the welfare of patients suffering from CMIs.

The content of the interviews was analysed (Graneheim and Lundman, 2004; Krippendorff, 1989) by individually identifying 'meaning units' that were concise description of the message conveyed in the response. In the subsequent step, the meaning unit was further condensed, coded and categorized into codes and subcategories (see Table 2 for an example analysis methodology). The steps were iteratively analysed 4-6 times for the response of each interviewee until a consensus on the codes was reached. The datasets of the participants and the annotations have been provided in supplementary section as separate Excel sheet. A total of 27 codes (Fig. 1) were identified that were refined into subcategories categories that shared similar sub-themes

(Fig. 2).

3. Results

For the presentation of the analysis, based on the interviews (Table S1), the study investigates the multi-dimensional challenges in management of CMIs and synthesizing insights from various stakeholders. This approach permits for a comprehensive investigation of aspects related in diagnosing, treating, patient-care and managing CMIs as a whole. Our purpose was integrating these diverse viewpoints, one may identify not only the clinical and operational hurdles but also the emotional, social, and economic challenges faced by patients and their families. Additionally, understanding these multi-faceted challenges may enable the formulation of policy recommendations that address the needs of all stakeholders, fostering a more collaborative, and an integrated approach to mental health care.

3.1. Tasks related to diagnosis

Firstly, the narrations on "tasks related to diagnosis", encompasses the process of diagnosing mental conditions involving several critical steps such as diagnosis, prognosis and the various situations present unique challenges, such as differing symptom manifestations in depression, requiring tailored responses and therapeutic goals. The process begins from the moment a mental health professional receives a patient or sample, the structured work-flow and the flexibility in addressing the diverse needs of patients effectively.

The data show a broad variety of replies across the different tasks both professionally as well as in ancillary roles. As an example, P2 said,

Table 2
An example analysis of the data grouped into sub-categories, sub-themes, and the main theme.

Participant (specialization)	Condensed meaning unit	Code	Sub-category	Sub-theme	Theme
01. Psychiatry Researcher	Freedom of participation is important, if unwilling then the participant is free to leave	guidelines	guidelines represent the scope of activity	Biological / Medical	Biopsychosocial model
05. Psychiatry Pharmaceuticals	The resistance to medication stems from the fact that the patient is unclear on the purpose and the outcome of the medication	complexity in treatment	patients resist treatment to medication and recovery program		
03. Neurosciences Entrepreneur	Yes, patients do quite often resist treatment provided they are not mobile		patients resist treatment		
02. Neurosciences Researcher	There is unspoken friction in the very first interaction between patient and doctor. The doctor needs to cautiously win over the patient's trust.	lack of trust	loss in trust due to unclear info results in resistance to medical support	Psychological / Psychosocial	
06. Social worker - IT	Patient journey is very complex and challenging due to mistrust, and uncertainties to every aspect of life	patient journey	patient journey is arduous and filled with several challenges like history, trust, resources. Understanding these are critical for successful intervention		
04. Medicinal chemistry entrepreneur	Challenges comes in several different flavours, in doctors, patients, resources, regulations and insurances	legality and bureaucracy	medical care is severely time restrained and highly over regulated by legislative protocols	Social / Societal	
07. Nurse	Legal issues are very restrictive in medical support and the medical staff can't enforce support unless life-threatening				

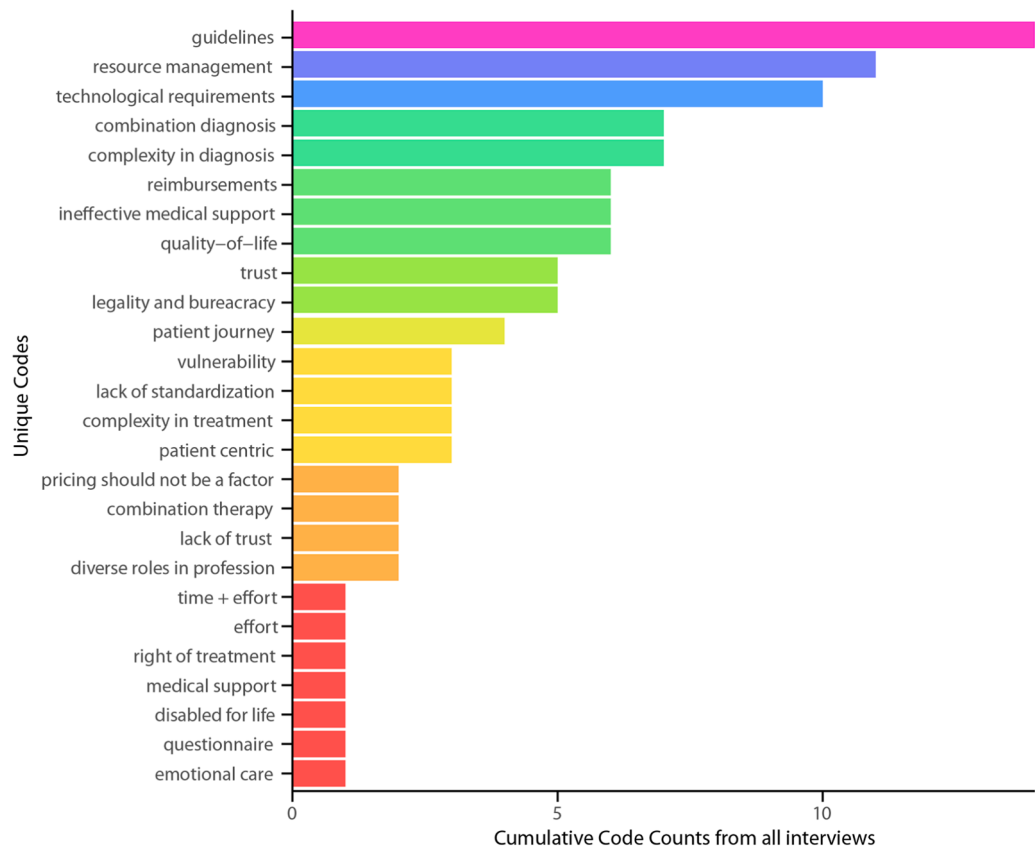


Fig. 1. Illustration of the occurrence (in ascending order) of various codes that were annotated from the interview.

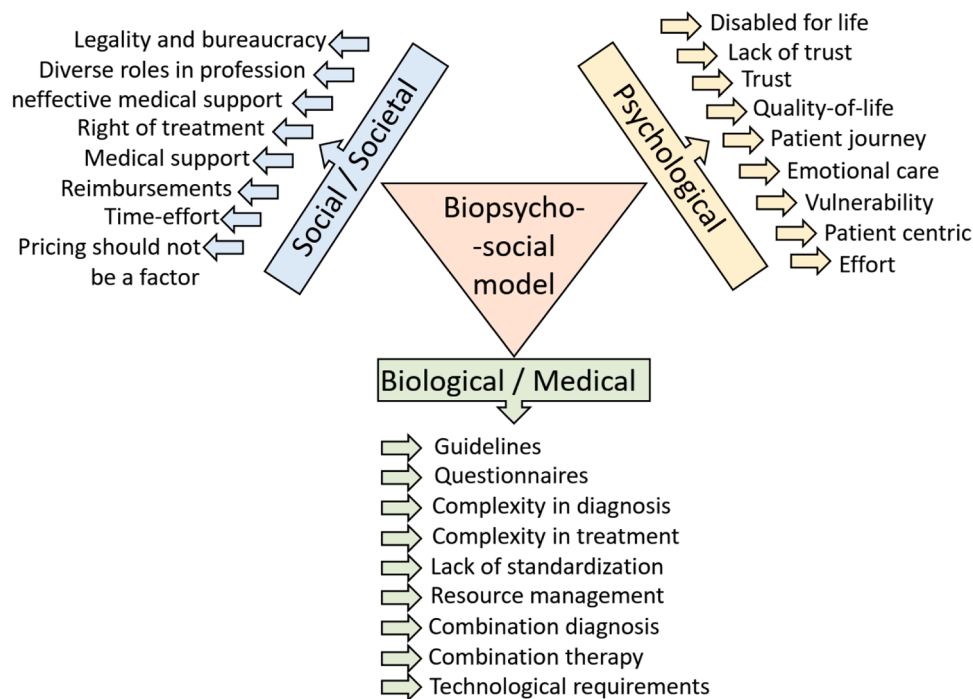


Fig. 2. Schematic representation of the identified codes and its association with the sub-themes, Biological/ Medical, Social/ Societal and Psychological/ Psycho-social factors in the theme biopsychosocial model.

“Neural disorder is quite different from degeneracy because by inspecting the brain you can see some pathology or some theoretical changes in the brain, and you can act accordingly and treated accordingly. Yeah, so by

surgery, our medicines are schematic treatment. But I think in neural disorder disease, I think as a physician in disorder management, we should pay more attention to the emotional support” (P2)

On the other hand, P4's response is

"For us in the new field, it's a multi-step process involving all the stakeholders. So of course, it's getting an idea and evaluating market needs in terms of the laboratory offering, our tests, the physicians who do diagnosis, who decide which therapy to go, who accompanied the therapy along the patient journey and of course the patient. And that's quite complex for us and a component to complete development process then. So, it's not just we're asking some questions, have understood it fully since asking three physicians, getting four answers is more normal than." (P4)

Taken together, the process is challenging not only for diagnosis but also management of the individual and the system. On a related note, P7 relished the ancillary role of the job as worthwhile, something he is proud of,

"I had a patient who was sort of put into my care. Just as she was being discharged, she gave me chocolate and said I saved her life and that was like very meaningful from for like negative, what keeps me awake at night?" (P7).

3.2. Challenges with current procedures

Secondly, the questions on the "challenges with current procedures", that include complexity in diagnosing mental health conditions and potential errors. Furthermore, we discussed on challenges that may arise when patients resist the diagnostic process, how success and failure are gauge for performances, and the budgetary. In this context, we were also keen to comprehend patients and their families seek savings in terms of time, money, and effort for more streamlined and accessible mental health services.

Among several factors, money is a very important criteria for patient welfare, science and development of new technology. One of the participants (P4) who is involved in all three aspects said,

"Most of the expenses are taken by the society since it's not so in terms of indirect effect. Patient journey's inefficiency of the system sector of course, if you are in concrete steps within the system. It's why other insurance companies quite good covered as far as we have experienced and have learned, but to get things new things improvements are need in this system again. It's a very long and much too long journey and we experience every day at the moment that the healthcare system is much too slow and inefficient, which means reimbursement issues take too long time." (P4).

The view point of a care-giver (P6) relates to the above yet markedly differs,

"Fair share of mental health, because listening to the stories day-in and day-out causes us problems as well. So those are the hidden costs that we can't see and there are costs as providers that we have to incur. In fact, and I think that's a higher price that we have to pay in terms of being in that position or being in that service" (P6).

3.3. Expectations from newer diagnostic methods

And lastly, on subject related to "expectations from newer diagnostic methods", we were keen to identify current value propositions in mental health care and their limitations. Also, it was important to understand performance issues related to diagnostic accuracy and treatment efficacy that can directly impact their work, causing frustration and impacting patient outcomes. While existing quality standards are in place, we sought to gather information on a hypothetical "silver bullet" that could simplify and enhance the diagnostic process, aspects of current procedures that professionals appreciate, such as specific features that streamline assessments or improve patient interactions. Lastly, in order to adopt a new value proposition, such as a blood-based biomarker assay, what factors would professionals prioritize better quality, lower

risk, low investments, and reduced costs.

Coming from a person scientifically skilled in the subject but also is a caregiver to family member, P3 says it is alarming to note the opinion of the current diagnostic that,

"So, the diagnostic guidelines, they are just not visible to the patient or the family, it's a black box. When it comes to cancer, or so, at least he has something to kind of hold on to. Here there are just probably no set guidelines that are available to the family members or the caregivers or the kids and kids of the patients, right? I mean that is the most frustrating."

So again, if we can get something to stage, OK, there is a major moment or there is a major, which is quantitative, that's going to drive the decisions both on the diagnostic side and on the treatment side that will be helpful."

On the other end of the spectrum P5 says,

"In the clinical development, I think there's no issue, okay, because it's quite straightforward approach, okay, and because we are always talking about a smaller group. Let's say in the development of a drug, you might need, let's say in 3000 patients. That's a relatively small group. And you have you use clear diagnosis criteria you have this exclusion inclusion criteria. And so, it's for the bigger indication it's not a problem" (P5).

These messages strongly point out the pole-apart differences in opinion between different stake holders and little congruence on the subject related to rating the value of the current diagnosis. However, regardless of whom we interviewed, most of the participants were suggestive of the need for better quality and lower risk technology as the priority for health care and management of the afflicted individual.

4. Discussion

During the course of the interview, we set out with frameworks related to tasks related to diagnosis, challenges with current procedures, and expectations from newer diagnostic methods in healthcare of CMIs. Also, in recent years, there has been an increase in use of biomarker-based analysis for various disease across a spectrum of medical disciplines in addition to increased usage of data-analytics, machine learning and artificial intelligence approaches for personalized medicine. Identifying biomarkers in a qualitative and quantitative fashion can offer strong basis for diagnostic and treatment follow-on for psychiatric disorders. Therefore, we were keen on understanding the viewpoint and gather as much as information of the different stakeholders in CMIs, which revealed several key factors that seem to origin from aspects associated with biological/medical, social/societal and psychological/psychosocial to align with George Engel's biopsychosocial model ((Engel, 1979), Figs. 2 and 3).

Participants highlighted the importance of understanding the technology, access to more information, better resource management, limited availability of newer technologies and the impact on their quality-of-life (Fig. 3). The study highlights the need for a shift in the management of CMIs, moving away from a narrow, biomedical model towards a more integrated approach that is more illustrative of the biopsychosocial model for patient-centred outcome. This involves not only optimizing pharmacological interventions but also incorporating psychological and social support strategies to address the multifaceted needs of individuals with schizophrenia.

Another set of important findings were combination-diagnostics and combination-treatment were vouched by most of the participants. They strongly were suggestive that while the current set of guidelines and therapy is constructive and have benefitted in patient care, it however lacks transparency and clearly is insufficient. In addition, it is necessary to astutely listen to the patients, and where possible, personalized medicine needs to take into account the personal history and cultural settings of the patient (Gomez-Carrillo and Kirmayer, 2023; Gomez-Carrillo et al., 2023). In a similar vein, the response to the treatment needs sincere communications between the clinics and hospital for the

Biological / medical factors 1. Guidelines ➤ guidelines represent the scope of activity ➤ current guidelines are good but at times can be challenging with complexity of disorder 2. Questionnaires ➤ questionnaire for mental disorder defines the strength and specificity of clinical and research practices 3. Complexity in diagnosis ➤ mistakes are possible during diagnosis ➤ due to complexity of the diseases, the diagnosis is often difficult 4. Complexity in treatment ➤ patients resist treatment to medication and recovery program ➤ further diagnostic classification of spectral disorders is required for better drug development 5. Lack of standardization ➤ the procedure for diagnosis and prognosis is subjective, unclear, and has insufficient criteria for validation 6. Resource management ➤ multi-layered challenges are required for clinical diagnostics in the current scenario, ➤ medical outcome requires interaction between physician, society, patient and pharma 7. Combination diagnosis ➤ multiple methodologies using clinical and newer technologies can improve patient outcome 8. Combination therapy ➤ medication, non-medical treatment and patient communication is required quality standards for successful outcome 9. Technological requirements ➤ unmet needs for psychiatric disorders require newer ideas and input from medical society ➤ diagnostic conditions for complex psychiatric disorders with broad symptom requires more than one biomarker ➤ newer technology should be encouraged and modern state-of-art technologies required ➤ drug development requires proper molecule profiling with respect to efficiency, safety and expected follow-ups using a biomarker assay	Social / societal factors 1. Legality and bureaucracy ➤ medical care is severely time restrained and highly over regulated by legislative protocols 2. Diverse roles in profession ➤ one has to assume several roles to gather maximal information for successful outcome 3. Pricing should not be a factor ➤ best possible patient care and quality trumps cost factor 4. Ineffective medical support ➤ clinical setup is outdated, while new technologies are available, they remain unused poor diagnostic solution are not providing adequate support ➤ many difficulties arise due to exhaustive patient journey due to mismanagement and inefficiency 5. Right of treatment ➤ many stake holders need to support the afflicted, which should be their fundamental rights 6. Medical support ➤ prognosis is challenging in neurological field and is critical in-patient recovery 7. Reimbursements ➤ insurance coverage decides on diagnostic frame work ➤ budgetary is critical for research, health insurance and family spending with insurance being a very critical deciding factor ➤ hidden costs go unnoticed and is a heavy societal burden 8. Time-effort ➤ money is not an important factor but time and effort are required for patient care	Psychological / psychosocial factors 1. Disabled for life ➤ due to difficulty in diagnosing, the afflicted individual is affected for life 2. Lack of trust ➤ loss in trust due to unclear information results in resistance to medical support 3. Trust ➤ trust building is vital for proper diagnosis 4. Quality-of-life ➤ outcomes should be pursued towards improving patients' quality-of-life ➤ reasonable diagnostics are available for medical care but not sufficient for stigmatization 5. Patient journey ➤ patient journey is arduous and filled with several challenges like history, trust, resources. Understanding these are critical for successful intervention 6. Emotional care ➤ it is important to recognize and tend to the personal space of the patient 7. Vulnerability ➤ physician needs to build trust and address patients' concerns and apprehension 8. Patient centric ➤ success must be defined on a patient-centric manner 9. Effort ➤ care taking is effort intensive
--	--	---

Fig. 3. The subcategories within sub-themes exemplifying the variety of aspects experienced by professionals in CMI health management.

greater benefit of the patient (Paul and Potter, 2024).

Listening to the patient, gaining their trust, addressing their vulnerability and emotional care emerges as one of the fundamental findings from our investigations from the interviews with clinician, nurse, social worker and pharma. In fact, such an approach has been suggested and laid in the concept of narrative medicine (Gomez-Carrillo and Kirmayer, 2023; Kirmayer et al., 2023). To this extent, George Engel's biopsychosocial approach (Engel, 1978) recognizes the significance of this two-way interaction.

Furthermore, participants shared their experiences of navigating the healthcare system, role of financial support, the ancillary help by care givers and the role of legislative and bureaucracy in the patient recovery (Fig. 3). They expressed frustration with the fragmented and often complex nature of the healthcare system, which posed significant barriers to accessing comprehensive and coordinated care.

This study highlights the need for a new and revamped outlook in management of schizophrenia and related psychotic disorders like bipolar disorder and, major depression towards a more patient-centric. Currently, the most prevalent methodology in patient care involves descriptive psychopathology pioneered by the Karl Jaspers (Hafner, 2015), which takes into account "objective description of abnormal states of the mind" or a phenomenological and later a physiological stance to psychotherapy. However, considering the challenges and limitations that the health care and the society is facing, one needs to evolve with the provisions of medical support in a time relevant manner.

On those lines, George Engel's biopsychosocial model (Engel, 1978, 1979), which is 45 years old is younger than the other models and serves a valuable framework, for instance in oncology for patient-centred (Lauriola and Tomai, 2019) and clinical trials (Chartogne et al., 2021), in cardiovascular diseases (CVD) like myocardial infarction (Tomas et al., 2023), rehabilitation of CVD (Anttila et al., 2021) to name few. It is also a vital component of medical education (Jaini and Lee, 2015; Roberts, 2023) for CMIs and continues to be a vital part of psychiatric and psychotherapy training worldwide. However, the subject matter has received a fair share of criticism in the world of psychiatry (Ghaemi, 2009; Roberts, 2023), which in itself is not totally unfounded.

This is because the utility of a model is best described by the unmet needs and success, but there is no experimental evidence (Williamson, 2022) supporting its requirements in psychiatry, thereby rightfully questioning an otherwise popular model, which is the prerequisite of a critical scientific process.

Interestingly, our findings (Figs. 2 and 3) for the requirements of the challenges and the unmet needs in the world of psychiatry is strongly indicative of applying the model for greater success similar to CVD and oncology treatments mentioned above. A schematic representation (Fig. 2) of our findings shows segregation of the different codes into 3 different themes, namely biological/ medical, societal and psychosocial/ psychological categories. Deeper insights into these codes reveals several different subcategories (Fig. 3) within. For instance, among the biological code, "technological requirement" reveals 5 subcategories that can be summarized as, "the unmet needs in psychiatric disorders necessitate innovative approaches and input from the medical community. Diagnosing complex psychiatric conditions with diverse symptoms demands multiple biomarkers. Advancing state-of-the-art technologies is essential, and drug development should prioritize thorough molecule profiling for efficacy, safety, and biomarker-based monitoring.". This aspect is complemented by social/ societal factors (Fig. 3) such as "ineffective medical support" and "right of treatment". These codes jointly are suggestive of "outdated clinical practises and underutilized new technologies lead to inadequate diagnostic solutions and ineffective medical support resulting in hampering patient journey. It is crucial that all of us, stakeholders ensure the fundamental right to proper treatment for those affected.". Clearly a direct link to the codes from the above two factors points to psychological/ psychosocial factors in "quality-of-life" and "patient journey", where "efforts should focus on enhancing patients' quality-of-life. Although reasonable diagnostic procedures are in place, they don't adequately address stigmatization. Altogether, the patient journey is challenging, involving complexities like medical history, trust, and resource limitations, which are crucial to understand for effective intervention."

Based on these subcategories within the codes, enables us to build a framework that utilizes cross- and trans- disciplinary approaches for psychiatric disorder management. To this extent, several models in

patient welfare have already been mentioned and also applied, some of which are from a social and cultural perspective (Gomez-Carrillo and Kirmayer, 2023; Kirmayer et al., 2023), others are psychological and psychosocial interventions like cognitive behaviour therapy (CBT) (Salahuddin et al., 2024), yet others that are medical in origin (Association, 2013; Organization, 2022), and few others encompass integration of medical and research (Clark et al., 2017), or medical, societal and psychological (Engel, 1978; Gomez-Carrillo et al., 2023; Organization, 2023). In one such major effort of “reinventing schizophrenia” and “updating the construct”, several leading experts provided a diversely differing opinion on schizophrenia health care in an editorial report (Tandon et al., 2023), which matches aptly with our identified codes of “lack of standardization” and “complexity in diagnosis”. Taking into account the stake holder views from our experimental findings in psychiatric health care management unequivocally identify combination approaches to diagnosis and treatment as the best plausible route. Although, combination of medical and research have steadily gained grounds in clinical applications (Clark et al., 2017), clearly the medical fraternity in clinicians seems divided on the best possible strategy (Tandon et al., 2023) due to various valid yet conflicting views. Nevertheless, assuming all these models, thoughts and connotations are true then it clearly reflects a lack in communication between experts of various disciplines tending to the healthcare and welfare of psychiatric ailments like schizophrenia and related CMIs (Moren, 2024). Not surprisingly, the outcomes are indeed reflected in the response to the interview questions, so much so that the interviewees themselves have acknowledged the need of cross- and/or trans- disciplinary approaches.

Interestingly, summarizing the coded factors from the interviews fell into 3 major categories that encompass elements from the perspectives of scientific fields such as biological/medical (e.g., medication, research and management), psychological/psychosocial (e.g., emotions, coping strategies, stigma), and societal/social (e.g., patient, healthcare system navigation) factors (Fig. 3). The strength of this study lies in its qualitative, patient-centric approach, which provides in-depth insights into the experiences and perspectives of professionals associated with support to CMIs like schizophrenia, and depression from scientific, medical, and care-giving. The study’s finding corroborates with Engel’s biopsychosocial model (Engel, 1979) as a guiding framework ensures a comprehensive reductionist yet holistic exploration of the factors influencing management of psychiatric disorders.

Like any other form of qualitative research (Malterud, 2001), the study is limited by small sample size and the potential for selection bias, however, as participants who engaged were from Europe, the findings presented herein is more specific for the geographic region. If any bias, then we are more of the opinion that may arise due to absence of professionals in CMIs who did not engage in the interviews such as practising physicians, social scientists in mental health and nursing, actuarial mathematicians specializing in health insurance policies, and policy makers themselves. In addition, we have many more male participants during the course of the event. It is not uncommon to observe that in many such studies it is challenging to interview female participants (Gubrium and Holstein, 2001). Qualitative research indeed can be challenging (Dickson-Swift et al., 2007) to conduct for both the interviewer and the interviewee due to the sensitive nature of the research. Drawing strength from our work, however we hope that in a similar attempt, “missing” experts would be willing to share their experience. Lastly, few of the experts whom we had contacted declined to participate because of the sheer stress that they had to endure previously caretaking their loved ones afflicted with CMIs. Lastly, it is important to note that qualitative research sampling does not have a set number of participants but is dependent on the richness of the information (Fossey et al., 2002; Marshall, 1996; Sandelowski, 1995) that illustrates the microdetails of the subject under study. Equally important the data obtained using the qualitative study herein will serve as a springboard for further qualitative and quantitative studies for health management for CMIs.

By addressing the multifaceted needs of individuals with

schizophrenia and CMIs, policy makers and healthcare providers can develop general as well as personalized regimes that improve clinical outcomes, reduce the burden on patients and their families, and contribute to the overall well-being of this population. The findings of this study have important implications for the development of more comprehensive and effective management strategies of CMIs. While all of these criteria exist in terms of insurance, health care management, diagnostic criteria, etc., a major aspect goes neglected and is rather confined to literature is the patient perspective that has remained unchanged. In an anonymously drafted report in 1988 (“Where next with psychiatric illness?”, 1988) titled, “Where next with psychiatric illnesses?”, the writer describes schizophrenia as the worst disease and the person disabled for life. We found similar codes, “disabled for life”, “patient journey” and “quality-of-life” from our study. Unfortunately, very little has changed and communication between different stake holders remains awfully poor (Moren, 2024) dramatically affecting medical services and positive patient outcome.

Therefore, it is necessary that the future research focus on implementing and evaluating the impact of a more encompassing model such as biopsychosocial interventions on the clinical, functional, and quality-of-life outcomes of individuals with schizophrenia, depression and related CMIs. A perfect example for such an initiative can be drawn from the European intersocietal recommendations for the biomarker-based diagnosis of neurocognitive disorders (Frisoni et al., 2024). As an example, we have recently observed an interesting shift in DSM IV to DSM 5 (Bredstrom, 2019; Substance Abuse and Mental Health Services Administration, 2016), although inadequate it represents an important shift by incorporating the views of social sciences researchers.

While this manuscript was under revision, a narrative review (Leucht et al., 2024) was published, which emphasizes on the importance of focusing on patient’s experiences in psychiatric diagnosis and treatment. Furthermore, it criticizes modern approaches that abstract patient experiences into higher-order constructs as described in the manuals DSM and ICD, which are necessary for reimbursement for medical claims but have the potential of obscuring actual symptoms and the sufferings the patients experience. This perspective aligns with the findings from our interviews, which highlight the necessity of understanding patients’ personal histories and cultural contexts to provide effective, personalized care. Both sources advocate for a shift from a narrow biomedical model to a more integrated approach, addressing the complex interplay of biological, psychological, and social factors in mental health care.

A silver lining, however in the current psychiatric cloud is the report commissioned by NHS England South, where the psychometric scale such as PANSS (Kay et al., 1987) was recommended by the Trust of NHS-UK (Ltd, 2021) to improve patient outcome at all stages of medical support. On a macroscale, PANSS draws its strengths from its objective assessment for clinical practises, and drug-development by incorporating 30 different scaled-biomarkers in symptoms, social and psychopathological factors. In an essence, PANSS is a conceptual derivative of the biopsychosocial model, and has potentials in providing personalized care for patient outcome, hence is strongly advocated in several clinical practises in a near-ubiquitous manner.

In our opinion, based on the findings, and the recent reports on the patient-centred outcome research (Kim et al., 2018; Largent et al., 2018; Moreno-Poyato et al., 2021) and strategies that involve cross- and transdisciplinary approaches such as collaborating with genomics groups like Schizophrenia Spectrum Biomarkers Consortium (SSBC) ((SSBC), 2024)), biomarker identification (Korth and Fangerau, 2020), antibody-antigen assays (Pils et al., 2023) and phenomics studies using proteomics (Sialana et al., 2018) and metabolomics (Jansen et al., 2023), in a combined fashion holds the key areas of engagement. Therefore, additional efforts are required to integrate several different ideas for a patient-centric approach into the broader healthcare system, ensuring that individuals with schizophrenia have access to the support and resources they need to manage their condition effectively.

Funding

A.C. acknowledges from the Hanse-Wissenschaftskolleg (HWK), Germany and Helmholtz Information & Data Science Academy - Norwegian Artificial Intelligence Research Consortium (HIDA NORA) to conduct the research activity.

CRediT authorship contribution statement

Abhishek Cukkemane: Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Rene Hurlemann:** Writing – review & editing, Supervision, Conceptualization. **Veralia Gabriela Sanchez:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

All authors declare no conflict of interest

Acknowledgements

A.C. thanks Ms. Diana Immel and Mr. Danilo Postin, from the NEMO group in the Karl-Jaspers-Klinik, Germany, Prof. Dr. Dieter Willbold from the Forschungszentrum Jülich, Germany for helpful comments and deeper insights into clinical and social psychiatry. We are extremely grateful to all participants who have shared their professional experience and also sensitive content for the success of this study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jadr.2025.100893](https://doi.org/10.1016/j.jadr.2025.100893).

References

- Anker-Hansen, C., Skovdahl, K., McCormack, B., Tonnessen, S., 2019. Invisible cornerstones. A hermeneutic study of the experience of care partners of older people with mental health problems in home care services. *Int. J. Older. People Nurs.* 14 (1), e12214. <https://doi.org/10.1111/opn.12214>.
- Anker-Hansen, C., Skovdahl, K., McCormack, B., Tonnessen, S., 2020. Collaboration between home care staff, leaders and care partners of older people with mental health problems: a focus on personhood. *Scand. J. Caring Sci.* 34 (1), 128–138. <https://doi.org/10.1111/scs.12714>.
- Anttila, M.R., Soderlund, A., Paajanen, T., Kivisto, H., Kokko, K., Sjogren, T., 2021. Biopsychosocial profiles of patients with cardiac disease in remote rehabilitation processes: mixed methods grounded theory approach. *JMIR Rehabil. Assist. Technol.* 8 (4), e16864. <https://doi.org/10.2196/16864>.
- Association, A.P., 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. American Psychiatric Association, Arlington VA.
- Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. *Qual. Res. Psychol.* 3 (2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- Bredstrom, A., 2019. Culture and context in Mental health diagnosing: scrutinizing the DSM-5 revision. *J. Med. Humanit.* 40 (3), 347–363. <https://doi.org/10.1007/s10912-017-9501-1>.
- Canada, M.H.C.o. (2012). *Making the case for investing in mental health in Canada*. https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf (last visited 24 Oct 2024).
- Chartogne, M., Leclercq, A., Beaune, B., Boyas, S., Forestier, C., Martin, T., Thomas-Ollivier, V., Landry, S., Bourgeois, H., Cojocarasu, O., Pialoux, V., Zanna, O., Messonnier, L.A., Rahmani, A., Morel, B., 2021. Building a biopsychosocial model of cancer-related fatigue: the BIOCARE Factory cohort study protocol. *BMC Cancer* 21 (1), 1140. <https://doi.org/10.1186/s12885-021-08831-3>.
- Clark, L.A., Cuthbert, B., Lewis-Fernandez, R., Narrow, W.E., Reed, G.M., 2017. Three approaches to understanding and classifying Mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychol. Sci. Public Interest* 18 (2), 72–145. <https://doi.org/10.1177/1529100617727266>.
- Dickson-Swift, V., James, E.L., Kippen, S., Liamputtong, P., 2007. Doing sensitive research: what challenges do qualitative researchers face? *Qual. Res.* 7 (3), 327–353. <https://doi.org/10.1177/1468794107078515>.
- Elo, S., Kyngas, H., 2008. The qualitative content analysis process. *J. Adv. Nurs.* 62 (1), 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>.
- Engel, G.L., 1978. The biopsychosocial model and the education of health professionals. *Ann. N. Y. Acad. Sci.* 310, 169–187. <https://doi.org/10.1111/j.1749-6632.1978.tb22070.x>.
- Engel, G.L., 1979. The biopsychosocial model and the education of health professionals. *Gen. Hosp. Psychiatry* 1 (2), 156–165. [https://doi.org/10.1016/0163-8343\(79\)90062-8](https://doi.org/10.1016/0163-8343(79)90062-8).
- Fossey, E., Harvey, C., McDermott, F., Davidson, L., 2002. Understanding and evaluating qualitative research. *Aust. N. Z. J. Psychiatry* 36 (6), 717–732. <https://doi.org/10.1046/j.1440-1614.2002.01100.x>.
- Frisoni, G.B., Festari, C., Massa, F., Cotta Ramusino, M., Orini, S., Aarsland, D., Agosta, F., Babiloni, C., Borroni, B., Cappa, S.F., Frederiksen, K.S., Froelich, L., Garibotto, V., Haliassos, A., Jessen, F., Kamondi, A., Kessels, R.P., Morbelli, S.D., O'Brien, J.T., Nobili, F., 2024. European intersocietal recommendations for the biomarker-based diagnosis of neurocognitive disorders. *Lancet Neurol.* 23 (3), 302–312. [https://doi.org/10.1016/S1474-4422\(23\)00447-7](https://doi.org/10.1016/S1474-4422(23)00447-7).
- Global Alliance of Mental Illness Advocacy Network (GAMIAN-Europe). <https://www.gamian.eu> (last visited 24 Oct 2024).
- Ghaemi, S.N., 2009. The rise and fall of the biopsychosocial model. *Br. J. Psychiatry* 195 (1), 3–4. <https://doi.org/10.1192/bjp.bp.109.063859>.
- Gomez-Carrillo, A., Kirmayer, L.J., 2023. A cultural-ecosocial systems view for psychiatry. *Front. Psychiatry* 14, 1031390. <https://doi.org/10.3389/fpsyt.2023.1031390>.
- Gomez-Carrillo, A., Paquin, V., Dumas, G., Kirmayer, L.J., 2023. Restoring the missing person to personalized medicine and precision psychiatry. *Front. Neurosci.* 17, 1041433. <https://doi.org/10.3389/fnins.2023.1041433>.
- Graneheim, U.H., Lundman, B., 2004. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ. Today* 24 (2), 105–112. <https://doi.org/10.1016/j.nedt.2003.10.001>.
- Gubrium, J., Holstein, J., 2001. Handbook of Interview Research. <https://doi.org/10.4135/9781412973588>.
- Hafner, H., 2015. Descriptive psychopathology, phenomenology, and the legacy of Karl Jaspers. *Dialog. Clin. Neurosci.* 17 (1), 19–29. <https://doi.org/10.31887/DCNS.2015.17.1/hhaefner>.
- Hanf, M., Hirt, J., van den Akker, M., 2021. Primary care professionals' attitudes towards digital health interventions for common mental disorders: study protocol for a mixed methods systematic review. *BMJ Open* 11 (6), e045657. <https://doi.org/10.1136/bmjopen-2020-045657>.
- European Federation of Associations of Families of People with Mental Illness (EUFAMI). <https://eufami.org/en>. (last visited 24 Oct 2024).
- Jaini, P.A., Lee, J.S., 2015. A review of 21st century utility of a biopsychosocial model in United States Medical School education. *J. Lifestyle Med.* 5 (2), 49–59. <https://doi.org/10.15280/jlm.2015.5.2.49>.
- Jansen, R., Milaneschi, Y., Schraner, D., Kastenmuller, G., Arnold, M., Han, X., Dunlop, B.W., Mood Disorder Precision Medicine, C., Rush, A.J., Kaddurah-Daouk, R., Penninx, B.W., 2023. The metabolome-wide signature of major depressive disorder. *Res Sq.* <https://doi.org/10.21203/rs.3.rs-3127544/v1>.
- Kallio, H., Pietila, A.M., Johnson, M., Kangasniemi, M., 2016. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *J. Adv. Nurs.* 72 (12), 2954–2965. <https://doi.org/10.1111/jan.13031>.
- Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13 (2), 261–276. <https://doi.org/10.1093/schbul/13.2.261>.
- Kim, K.K., Khodyakov, D., Marie, K., Taras, H., Meeker, D., Campos, H.O., Ohno-Machado, L., 2018. A novel stakeholder engagement approach for patient-centered outcomes research. *Med. Care* 56 (10 Suppl 1), S41–S47. <https://doi.org/10.1097/MLR.0000000000000790>. *Suppl 10 Suppl 1*.
- Kirmayer, L.J., Gómez-Carrillo, A., Sukhanova, E., Garrido, E., 2023. Narrative medicine. In: Mezzich, J.E., Appleyard, W.J., Glare, P., Snaedal, J., Wilson, C.R. (Eds.), *Person Centered Medicine*. Springer International Publishing, pp. 235–255. https://doi.org/10.1007/978-3-031-17650-0_14.
- Korth, C., Fangerau, H., 2020. Blood tests to diagnose schizophrenia: self-imposed limits in psychiatry. *Lancet Psychiatry* 7 (10), 911–914. [https://doi.org/10.1016/S2215-0366\(20\)30058-4](https://doi.org/10.1016/S2215-0366(20)30058-4).
- Krippendorff, K., 1989. Content Analysis. Oxford University Press, New York, NY. Vol. 1. http://repository.upenn.edu/asc_papers/226.
- Largent, E.A., Weissman, J.S., Gupta, A., Abraham, M., Rozenblum, R., Lynch, H.F., Cohen, I.G., 2018. Patient-centered outcomes research: stakeholder perspectives and ethical and regulatory oversight issues. *IRB* 40 (1), 7–17. <https://www.ncbi.nlm.nih.gov/pubmed/30631218>.
- Lauriola, M., Tomai, M., 2019. Biopsychosocial correlates of adjustment to cancer during chemotherapy: the key role of health-related quality of life. *ScientificWorldJournal*. 2019, 9750940. <https://doi.org/10.1155/2019/9750940>.
- Leucht, S., van Os, J., Jager, M., Davis, J.M., 2024. Prioritization of psychopathological symptoms and clinical characterization in psychiatric diagnoses: a narrative review. *JAMA Psychiatry* 81 (11), 1149–1158. <https://doi.org/10.1001/jamapsychiatry.2024.2652>.
- N. H. S. C. Ltd, 2021. An Independent Investigation into the Care and Treatment of a Mental Health Service User Mr J in Kent. <https://www.england.nhs.uk/south-east/wp-content/uploads/sites/45/2022/06/Final-report-2018-23654-v3.4.pdf>.
- Malterud, K., 2001. Qualitative research: standards, challenges, and guidelines. *Lancet* 358 (9280), 483–488. [https://doi.org/10.1016/S0140-6736\(01\)05627-6](https://doi.org/10.1016/S0140-6736(01)05627-6).
- Marshall, M.N., 1996. Sampling for qualitative research. *Fam. Pract.* 13 (6), 522–525. <https://doi.org/10.1093/fampra/13.6.522>.
- McCormack, B., McCance, T.V., 2006. Development of a framework for person-centred nursing. *J. Adv. Nurs.* 56 (5), 472–479. <https://doi.org/10.1111/j.1365-2648.2006.04042.x>.

- Moren, C., 2024. Put people at the heart of schizophrenia research. *Nature* 630, 531.
- Moreno-Poyato, A.R., El Abidi, K., Rodriguez-Nogueira, O., Lluch-Canut, T., Puig-Llobet, M., 2021. A qualitative study exploring the patients' perspective from the 'Reserved Therapeutic Space' nursing intervention in acute mental health units. *Int. J. Ment. Health Nurs.* 30 (3), 783–797. <https://doi.org/10.1111/inm.12848>.
- European College of Neuropsychopharmacology (ECNP). <https://www.ecnp.eu/Congress2023/ECNPcongress> (last visited 24 Oct 2024).
- OECD/EU, 2018. Health at a Glance: Europe 2018: State of Health in the EU Cycle. OECD Publishing, Paris. https://doi.org/10.1787/health_glance_eur-2018-en last visited 24 Oct 2024.
- World Health Organization, 2022. ICD-11: International Classification of Diseases (11th revision). <https://icd.who.int/>. last visited 24 Oct 2024.
- World Health Organization, 2023. Mental Health Gap Action Programme (mhGAP) Guideline For mental, Neurological and Substance Use Disorders. World Health Organization, Geneva. <https://www.who.int/publications/book-orders>. last visited 24 Oct 2024.
- Paul, S.M., Potter, W.Z., 2024. Finding new and better treatments for psychiatric disorders. *Neuropsychopharmacology* 49 (1), 3–9. <https://doi.org/10.1038/s41386-023-01690-5>.
- Pils, M., Rutsch, J., Eren, F., Engberg, G., Piehl, F., Cervenka, S., Sellgren, C., Trossbach, S., Willbold, D., Erhardt, S., Bannach, O., Korth, C., 2023. Disrupted-in-schizophrenia 1 protein aggregates in cerebrospinal fluid are elevated in patients with first-episode psychosis. *Psychiatry Clin. Neurosci.* 77 (12), 665–671. <https://doi.org/10.1111/pcn.13594>.
- Pharmacogenetics in Psychiatry (PSY-PGx) - A New Intervention for Implementation of Pharmacogenetics in Psychiatry. www.psy-pgx.org (last visited 24 Oct 2024).
- Substance Abuse and Mental Health Services Administration. (2016). *Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health*. Substance Abuse and Mental Health Services Administration (US).
- Roberts, A., 2023. The biopsychosocial model: its use and abuse. *Med. Health Care Philos.* 26 (3), 367–384. <https://doi.org/10.1007/s11019-023-10150-2>.
- Salahuddin, N.H., Schutz, A., Pitschel-Walz, G., Mayer, S.F., Chaimani, A., Sifas, S., Priller, J., Leucht, S., Bighelli, I., 2024. Psychological and psychosocial interventions for treatment-resistant schizophrenia: a systematic review and network meta-analysis. *Lancet Psychiatry* 11 (7), 545–553. [https://doi.org/10.1016/S2215-0366\(24\)00136-6](https://doi.org/10.1016/S2215-0366(24)00136-6).
- Sanchez, V.G., Anker-Hansen, C., Taylor, I., Eilertsen, G., 2019. Older people's attitudes and perspectives of welfare technology In Norway. *J. Multidiscip. Healthc.* 12, 841–853. <https://doi.org/10.2147/JMDH.S219458>.
- Sandelowski, M., 1995. Sample size in qualitative research. *Res. Nurs. Health* 18 (2), 179–183. <https://doi.org/10.1002/nur.4770180211>.
- Sialana, F.J., Wang, A.L., Fazari, B., Kristofova, M., Smidak, R., Trossbach, S.V., Korth, C., Huston, J.P., de Souza Silva, M.A., Lubec, G., 2018. Quantitative proteomics of synaptosomal fractions in a rat overexpressing Human DISC1 gene indicates profound synaptic dysregulation in the dorsal striatum. *Front. Mol. Neurosci.* 11, 26. <https://doi.org/10.3389/fnmol.2018.00026>.
- Smith, D.J., Craddock, N., 2011. Unipolar and bipolar depression: different of the same? *Br. J. Psychiatry* 199 (4), 272–274. <https://doi.org/10.1192/bjp.bp.111.092726>.
- Schizophrenia Spectrum Biomarkers Consortium 2024 (SSBC) <https://ssbcbio.org> (last visited 24 Oct 2024).
- Tandon, R., Keshavan, M., Nasrallah, H., 2023. "Reinventing schizophrenia: updating the construct". Project update and next steps. *Schizophr. Res.* 252, 345–347. <https://doi.org/10.1016/j.schres.2023.01.025>.
- Tomas, J.M., Oliver, A., Torres, Z., Parker, J., Marques-Sule, E., Sentandreu-Mano, T., 2023. A biopsychosocial model predicting myocardial infarction. *J. Clin. Med.* 12 (17). <https://doi.org/10.3390/jcm12175715>.
- van Os, J., Kapur, S., 2009. Schizophrenia. *Lancet* 374 (9690), 635–645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8).
- Where next with psychiatric illness?, 1988. *Nature* 336 (6195), 95–96. <https://doi.org/10.1038/336095a0>.
- Wiegand, H.F., Saam, J., Marschall, U., Chmitorz, A., Kriston, L., Berger, M., Lieb, K., Holzel, L.P., 2020. Challenges in the transition from In-patient to out-patient treatment in depression. *Dtsch. Arztebl. Int.* 117 (27–28), 472–479. <https://doi.org/10.3238/arztebl.2020.0472>.
- Williamson, S., 2022. The biopsychosocial model: not dead, but in need of revival. *BJPsych. Bull.* 46 (4), 1–3. <https://doi.org/10.1192/bjb.2022.29>.
- Zinn, J.O., 2010. The biographical management of risk and uncertainty—british veterans. *Qual Soc Res* 11 (1). <https://doi.org/10.17169/fqs-11.1.1457>.

Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments

Abhishek Cukkemane^{1,2,3*}, Rene Hurlemann^{4,5}, and Veralia Gabriela Sanchez⁶

¹Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

³Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

⁴Department of Psychiatry, School of Medicine and Health Sciences, University of Oldenburg, Oldenburg, Germany.

⁵Karl-Jaspers-Klinik, Bad Zwischenahn, Germany

⁶Faculty of Technology, Natural Sciences and Maritime Sciences, University of South-Eastern Norway, Vestfold, Norway

*Address for correspondence - a.cukkemane@fz-juelich.de

Semi-structured interview

Semi-structured interviews were conducted in a narrative manner. As mentioned in the methodology in the main text, interviews were audio-recorded, and transcribed verbatim prior to analysis.

Table S1: Semi-structured interview guide

I. Tasks related to Diagnosis	
1.	In your profession, what steps are required for tending to mental condition like schizophrenia depression or bipolar disorders?
2a.	How does one prognose the mental condition based on the patient/sample/information?
b.	Do you adopt different standards and guidelines in comparison to previous point?

3. What are the different contexts/situations that you face? For instance, depression encompasses several broad symptoms, how does your response and goals change depend on these different contexts?
- 4a. Can you please walk me through the entire process from the moment you receive the patient/sample/information?
- b. If a patient the how does the interaction look like during the appointment and until they leave?
- 5a. Do you also have to take up supporting jobs and/or switch roles throughout this life cycle; for instance, as a physician, then a teacher or as a mentor?
- b. Why?
6. As we understand borderline cases can take 1-6 months for diagnosis What are the consequences of a misdiagnosis?
 - a. for the patient
 - b. for your reputation?

II. Difficulties with current procedures

7. Are mistakes common during diagnosis/evaluation? Specifically, is it possible that one may evaluate the condition/disorder in the wrong way?
8. Related to the previous questions, do you face situations where it is difficult to getting certain things done, for instance patients who resist the process, answering questions and/or treatment?
9. Success and failure are subjective terms. However, scientific findings describe on the strong benefits and positive attributes of counselling, however the measure for success is ill-defined. How do you gauge performance and measure success or failure from a session and long-term?
10. Chronic mental illnesses are a heavy financial burden to the society, as high as 2.8 – 4 % (1, 2) GDP in the western world;
 - a. Do you know who bears these expenses?
 - b. Do you know who makes the budget decisions?
 - c. How does this affect your work?
11. While public health insurance foots a large amount of the bill pertaining treatment of mental health, a large sum is also borne by the patient and their family members, which savings would make your customers happy in terms of time, money and effort would they value?

III. Expectations of newer diagnostic methods

12. What keeps you up awake at night, i.e., What are the main difficulties and challenges that you encounter?
- 13a. How do you rate current value propositions?
- b. Which features in your opinion are missing?
- c. Do these performance issues affect your job? if yes, in which ways?
- 14a. What are the current quality standards?
- b. Are they satisfactory? or would you wish for something different?
- c. Let say, if there is a possibility to have a silver bullet to aid your work, what would that be?
- 15a. How do current diagnostic procedures delight you?
- b. Which specific features do you enjoy?
- c. What performance and quality do you expect?
16. What would increase your likelihood of accepting and adopting a value proposition, for instance a blood-based biomarker assay? Would you desire lower cost, less investment, lower risk, or better quality? Can you please prioritize starting from the most important one?

References:

1. OECD/EU. Health at a Glance: Europe 2018: State of Health in the EU Cycle: OECD Publishing, Paris; 2018.
2. Canada MHCo. Making the Case for Investing in Mental Health in Canada 2012.

Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics

Abhishek Cukkemane^{1,2,3*} and Veralia Gabriela Sanchez⁴

¹Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

³Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

⁴Faculty of Technology, Natural Sciences and Maritime Sciences, University of South-Eastern Norway, Vestfold, Norway

*Address for correspondence - a.cukkemane@fz-juelich.de

Abstract

Chronic mental illnesses (CMI), such as schizophrenia, schizoaffective disorder, bipolar disorder (BD), and major depressive disorder (MDD), place a substantial burden on individuals, family, society and the healthcare infrastructure. Existing treatment methods falls short in addressing the needs of patients thereby leading to inadequate care and less-than-optimal health outcomes. To address this gap, our study explores a patient-centric approach through leveraging text mining and natural language processing (NLP) techniques by analysing transcribed interviews from various stakeholders, including clinicians, researchers, and healthcare professionals. Using, sentiment analysis, we examine the emotions and sentiments expressed in CMI-related discourse by automated identification and categorization of opinions. This study explores the application of emotional and sentiment analysis using the four different lexicons in the syuzhet package in R to analyse open-ended responses in management of CMIs within academic, social and medical frameworks. The findings indicate that NRC lexicon provided text analysis methods with valuable insights into participants' emotional and attentional focus, thereby deepening our understanding of patient experiences and their reactions to interventions. Additionally, we compare sentiment analysis with outcomes from qualitative content analysis

to evaluate their effectiveness in routine scientific applications and policy making. Integrating sentiment analysis into CMI management has the potential to enhance patient-centred care, ultimately leading to improved treatment outcomes. This research emphasizes the importance of leveraging innovative, data-driven methodologies to supplement conventional psychiatric care and policy development, fostering a more holistic comprehension of CMIs.

Keywords

Schizophrenia; text analysis; emotion; sentiments, natural language processing; patient-centric approach; health-care management

Introduction

Chronic mental illnesses (CMIs), such as schizophrenia, bipolar disorder (BD), and major depressive disorder (MDD), affect millions worldwide and present a significant burden on individuals, healthcare systems, and society (van Os & Kapur, 2009). Effective management of these disorders requires a multifaceted approach, involving healthcare professionals, policymakers, and social support systems. However, despite advancements in medical guidelines (Association, 2013) such as (Diagnostic and Statistical Manual of Mental Disorders) DSM and International Classification of Diseases (ICD), CMIs remain difficult to diagnose due to overlapping symptoms, leading to frequent misdiagnoses. For example, differentiating between various CMIs is often challenging (Smith & Craddock, 2011; van Os & Kapur, 2009), resulting in inappropriate treatment strategies that may hinder recovery. Although diagnostic accuracy has improved in recent decades (Paul & Potter, 2023), healthcare systems still struggle to provide holistic and patient-centred care because substantial gaps in care remain, with deviations from statutory guidelines contributing to poor treatment outcomes.

Given these limitations, there is a growing need for alternative methods to enhance psychiatric care. Understanding the diverse requirements of stakeholders involved in psychiatric welfare is crucial. In a recent study, we conducted thematic qualitative analysis (Cukkeman et al., 2025) based on interviews with professionals, including clinical researchers, scientific researchers, pharmaceutical company managers, doctors, nurses, entrepreneurs, and social workers. The findings hinted for a holistic, patient-centred management approach aligned with the biopsychosocial model (Engel, 1978, 1979) proposed by Dr. George Engel. Qualitative content analysis (QCA) (Anker-Hansen et al., 2020; Graneheim & Lundman, 2004; McCormack &

McCance, 2006) and stakeholder interviews offer a microscopic view of psychiatric health welfare, capturing experiences, challenges, solutions, and daily activities that quantitative approaches often overlook. However, developing more reliable models and policies, one will require analysing a large number of interviews, a process that can be daunting when done manually in the traditional QCA. To address this challenge, we explored the potential of text mining (Arnold, 2017; Rutkowski et al., 2022; Yu, 2011) as a powerful tool in qualitative analytics. Text mining enables researchers to extract meaningful insights from large volumes of textual data. This strength of text mining analytics has found applications in politics (Hoffmann, 2018), social media analytics (Pv et al., 2024; Shehu et al., 2020), arts and humanities (Hoyeol, 2022) to name a few. Particularly in this case, sentiment analysis (Isasi, 2023; Yu, 2011) can computationally assess emotions expressed in stakeholder opinions in management of CMIs. This technique offers deeper insights into patient experiences, emotional states, and unmet needs in mental health care, ultimately enhancing psychiatric welfare through data-driven decision-making.

Herein, we employed sentiment analysis using the NRC Emotion Lexicon, a comprehensive tool developed by the National Research Council Canada (Mohammad & Turney, 2013). Unlike other lexicons, NRC categorizes words not only into positive or negative sentiment but also into eight basic emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. With over 14,000 annotated words, it allows for a nuanced understanding of emotional expression in text. The lexicon is popular among several literary scholars because it is available in over 100 languages, making it accessible for cross-cultural research. By applying NRC-based sentiment analysis to open-ended responses from key stakeholders—including clinicians, researchers, social workers, and patients—we aim to identify critical gaps in current CMI management approaches and compare the findings to the previously studied material.

This study highlights the importance of integrating qualitative sentiment analysis into psychiatric care to complement traditional diagnostic and treatment methods. By capturing emotional and attentional focus in textual data, sentiment analysis can help refine treatment plans, improve patient communication, and guide policy changes. Our findings align with Engel's biopsychosocial model, emphasizing the interconnectedness of biological, psychological, and social factors in mental health care. This research provides experimental evidence supporting a comprehensive, patient-centric approach to CMI management, offering a novel framework for improving psychiatric care through data-driven methodologies.

Methodology

Aim and research question

This study investigates whether text analysis techniques from interviews, particularly sentiment analysis, can provide valuable insights into patient needs, ultimately guiding more effective, patient-centred mental health interventions for the management of CMIs such as schizophrenia, BD, and MDD.

Ethical considerations

In accordance with Article 5(1)(d) (accuracy) and Article 5(1)(f) (integrity and confidentiality), as well as Article 32 (security) of the General Data Protection Regulation (GDPR), informed consent was obtained from all participants and the data was anonymized. This process was notified with the Data Protection Services of Sikt – Norwegian Agency for Shared Services in Education and Research (reference number: 648482). The legal basis for data processing was established under GDPR Article 6(1)(a), ensuring compliance with ethical and data protection standards.

Participant recruitment

As described in reference (Cukkemane et al., 2025), we recruited individuals engaged in the management, support, and care of schizophrenia and related chronic mental illnesses (CMIs). The eligibility criteria for the participants directly involved in CMIs healthcare covering disorders such as schizophrenia, BD and MDD. Also, because of the high global prevalence of CMIs (3–5%), few participants provided personal inputs due to family members affected by CMIs. The professional participants included caregivers, psychiatrists, researchers, social workers, and nurses. Briefly, participants were identified through multiple strategies, such as convenience sampling via the authors' professional network, email invited experts based on their scientific and snowball sampling, where existing contacts recommended other potential participants. Altogether, 7 participants (Table 1) from Germany, the UK, Norway, and Denmark ultimately agreed to participate from a total of 57 (23 females and 34 males) reflecting healthcare systems in the region of Europe.

Data analysis

To explore the emotional undertones embedded in the interview data, sentiment analysis was employed using the *syuzhet* package in R (Jockers, 2020; Naldi, 2019) that includes lexicons such as *syuzhet* (Jockers, 2020), NRC Emotion Lexicon (Mohammad & Turney, 2013), *afinn* (Nielsen, 2011), and *bing* (Bing, 2012). We first tested all the lexicons for the dispersity of the scores across the spectrum of positive to negative sentiments. Next, we expanded on the

sentiment using NRC Emotion Lexicon, based on the dispersity scores and also because unlike traditional sentiment analysis lexicons, the NRC Lexicon categorizes words into eight emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Examination of emotional patterns and the ten most prominent words from the stakeholder narratives provides us with a variety of categorical variables, i.e., the central codes that can be compared with themes from qualitative analytics (Cukkemane et al., 2025).

The sentiment analysis workflow involved the following steps: loading the *syuzhet* package, tokenization of sentences and words from the recorded interview transcripts. This was followed by extraction of sentences using the NRC Emotion Lexicon, emotional and sentiment-related data were extracted from the textual responses. The extracted data were visualized and summarized using tables, bar charts, word clouds, and trajectory plots. The trajectory plot describes normalized sentiment values across time for the participants and can roughly be divided into three key frameworks as described in the questionnaire (Cukkemane et al., 2025), i.e., “tasks related to diagnosis”, “challenges with current procedures”, and “expectations from newer diagnostic methods”. These visualizations illustrate the distribution of sentiment and emotional categories across the participant’s narratives. The R-scripts employed for all the analysis with the different lexicons have been provided in the supplementary section.

Results

Sentiment variations across different lexicons

The sentiment analysis results for seven participants (Table 1) with diverse professional backgrounds reveal notable trends in emotional expression across their transcripts. Overall, the analysis reveals that participants primarily use neutral language, with positive sentiment generally outweighing negative sentiment, with proportions ranging from 62.3% to 96%, indicating that either the majority of the language used in their transcripts or the lexicons themselves are biased to emotionally neutral tones. Positive sentiment varies between 2.7% and 28.7%, with P4 displaying the highest positivity, while P3 and P5 show lower positive sentiment scores. Negative sentiment is generally low, ranging from 0.8% to 13.9%, with P7 exhibiting the highest negativity, possibly reflecting early career stress or challenges.

Among the sentiment analysis methods, NRC consistently assigns higher positive sentiment than the other methods, often exceeding 20%. At the same time, it also assigns a notable amount of negative sentiment but not as much as the positive scores thereby providing a wider range of sentiments from positive and negative scores. AFINN’s sentiment scores tend to show less

extremes in values for both positive and negative sentiments but leans towards neutral sentiment. The difference between positive and negative scores is however smaller than NRC. Bing and Syuzhet show the lowest negative sentiment scores, suggesting they are far more conservative and seem more neutral-heavy.

As NRC based analytics distribute sentiment more widely, we believe it is better suited for qualitative analytics. In addition, the data is also suggestive of variations in sentiment reflecting across professional roles, with entrepreneurial and scientific roles generally showing higher positivity, while clinical and early-career roles exhibit slightly higher negative sentiment.

Emotions Across Stakeholders

Emotional analysis of Participant 1's discourse (**Figure 1, P1**) revealed trust (20%) as the dominant emotion, followed by anticipation (16%) and fear (14%). The prominence of trust aligns with this participant's extensive clinical background (>20 years) and medical training. Anger (12%) ranked higher than in most other participants, potentially reflecting frustration with system limitations encountered during clinical practice.

Analysis from the word cloud (**Figure 1, P1**) highlighted "money", "kind", "disease", "diagnosis", "good", "medical", "interested", "treat", "feeling", and "patient" as ten most prominent words in this participant's discourse. These words can broadly reflect themes "medical diagnosis and treatment", "financial concerns", "healthcare", and "emotional well-being". The sentiment valence trajectory (**Figure 3, P1**) demonstrated significant fluctuations, with notable negative dips when discussing diagnostic challenges and healthcare system fragmentation. Positive peaks occurred during discussions of evidence-based treatments and potential system improvements. The trajectory showed an overall slight positive trend, suggesting cautious optimism despite acknowledged limitations.

Participant 2's emotional profile showed the highest levels of trust (22%) among all participants, followed by anticipation (15%) and fear (13%). This participant also expressed higher levels of sadness (11%) compared to most others (**Figure 1, P2**). The word cloud emphasized "disease", "patient", "success", "suffering", "problem", "tumour", "friendship", "truth", "communication", and "management" as the prominent words and the sentiment valence trajectory (**Figure 3, P2**) showed stability in the positive range when discussing research advances and evidence-based approaches, with negative dips when addressing implementation barriers and research-to-practice gaps. The trajectory demonstrated a pronounced positive surge when discussing integrated and possibilities of newer diagnostics approaches, corresponding

with this participant's emphasis on holistic research frameworks. Altogether, the suggestive central themes fall into “medical diagnosis and treatment”, “healthcare management”, “emotional well-being”, “communication and support”, “problem solving and achievement”.

Emotional analysis of Participant 3 (**Figure 1, P3**) revealed anticipation (18%) as Participant 3's dominant emotion, closely followed by trust (16%) and fear (15%). This participant showed higher levels of joy (10%) than most others, reflecting an entrepreneurial optimism about innovation potential in CMI management. The word cloud (**Figure 1, P3**) highlighted “medical”, “patient”, “disease”; “cancer”; “problem”; “schizophrenia”; “suffering”; “depression”; “career”, and “reward” with central themes being “health and illness”, “emotional well-being and support”, “problem solving and achievement”. This was reflected in the sentiment valence trajectory (**Figure 3, P3**) showed an overall positive trend with significant peaks when discussing technological solutions and innovation opportunities. Negative valleys appeared when addressing regulatory barriers and implementation challenges. The trajectory demonstrated the most pronounced positive-negative oscillations among all participants, potentially reflecting the tension between entrepreneurial optimism and recognition of practical constraints.

Participant 4 displayed the lowest positive sentiment (3.0%) among all participants while maintaining similar negative sentiment levels to others. The emotional analysis revealed trust (18%) as dominant, followed by fear (16%) and anticipation (14%) as highlighted by predominant words (**Figure 1, P4**) in the cloud “medical”, “patient”, “success”, “disease”, “cancer”, “resources”, “happy”, “prognostic”, “treat”, and “risk” reflecting central themes such as “health and illness”, “treatment and care”, “resources and support”, “problem solving and achievement”. The sentiment valence trajectory showed the most pronounced negative trends when discussing regulatory challenges and trial limitations for psychiatric medications. Positive peaks corresponded with discussions of new measurement approaches and potential breakthroughs in targeted treatments. The overall trajectory (**Figure 3, P4**) demonstrated more negative segments than most other participants, potentially reflecting industry-specific challenges in psychiatric drug development. This participant showed the highest levels of disgust (10%) among all participants, potentially reflecting frustration with pharmaceutical industry constraints.

Participant 5 demonstrated a high degree of positive sentiments with emotional analysis (**Figure 2, P5**) showing anticipation (17%) and trust (17%) as dominant factors, followed by joy (11%) and fear (12%). This emotional profile aligns with an innovation-focused perspective

grounded in scientific training as emphasized in the word cloud (**Figure 2, P5**) comprising of “disease”, “patient”, “cancer”, “medical”, “suffering”, “treat”, “schizophrenia”, “risk”, “happy”, and “problem”. The sentiment valence trajectory (**Figure 3, P5**) maintained a consistent positive trend, with minimal negative excursions compared to other participants. The most significant positive peaks occurred during discussions of precision medicine approaches and biomarker development, while slight negative dips accompanied discussions of regulatory challenges and development timelines reflecting expertise in pharmaceutical approaches to mental health. The narrative emphasized themes central to “treatment and care”, “emotional well-being”, “risk management”, and “healthcare”.

Participant 6 demonstrated a completely different emotional profile (**Figure 2, P6**) with higher levels of sadness (12%) and disgust (11%) than other participants, while still maintaining trust (15%) and anticipation (15%) as dominant emotions. The word cloud (**Figure 2, P6**) emphasized “patient”, “medical”, “diagnosis”, “success”, “treat”, “risk”, “time”, “problem”, “illness”, and “improve”, highlighting themes in “communication and support”, “problem solving and improvement”, “treatment and care”, and “time and risk management”. This pattern likely reflects direct experience with the human impact of system limitations as both a social worker and family member of someone with CMI greatly reflecting the focus on practical challenges faced by individuals and families navigating mental healthcare systems. The sentiment valence trajectory (**Figure 3, P6**) showed significant negative excursions when discussing system navigation challenges and caregiver burden. Positive peaks appeared during discussions of community support and advocacy efforts. The trajectory demonstrated rapid oscillations between positive and negative sentiment, potentially reflecting the complex emotional reality of supporting individuals with CMIs through fragmented systems.

As the participant with the least professional experience (>1 year), Participant 7 showed a distinctive emotional profile (**Figure 2, P7**) with higher levels of surprise (11%) than other participants, alongside trust (18%), anticipation (16%), and fear (15%). The predominant words (**Figure 2, P7**) in the cloud highlighted “patient”, “success”, “diagnosis”, “happiness”, “disease”, “time”, “suffering”, “problem”, “trust”, and “resources” reflecting front-line clinical nursing perspectives. The sentiment valence trajectory (**Figure 3, P7**) showed consistent fluctuations with negative dips corresponding to discussions of staffing challenges and care coordination issues. Positive peaks occurred when discussing direct patient care and therapeutic relationships aligning with central themes as “support and relationship”, “emotional well-being”, “time management”, and “problem solving and achievement”. The trajectory

demonstrated an overall neutral trend with balanced positive and negative segments, reflecting the practical, day-to-day reality of psychiatric nursing practice.

Dominant emotions and trajectories across stakeholders

Trust emerged as the most prevalent emotion (**Figure 1 and 2**) across all stakeholder groups, comprising 15-22% of the emotional content. This finding suggests that despite challenges in CMI management, stakeholders maintain confidence in certain aspects of care systems or professional relationships. Following trust, anticipation (14-18%) and fear (12-16%) were the next most prominent emotions, reflecting the tension between hopeful expectations for improved care and concerns about current limitations.

Among healthcare professionals (Participants 1, 2, 5, and 7), trust and anticipation were particularly pronounced, suggesting their professional training may instill greater confidence in established protocols despite recognizing limitations. For participants with personal experiences with CMIs (Participants 2, 3, and 6), fear and sadness showed higher prominence compared to other participants, potentially reflecting the emotional toll of witnessing loved ones navigate complex healthcare systems.

Clinicians (Participants 1, 2, and 5) demonstrated higher levels of trust (18-22%) compared to non-clinical professionals (15-17%), possibly indicating greater faith in medical approaches. However, they also expressed more anger (10-12%) than other groups (7-9%), suggesting frustration with systemic barriers to optimal care.

Entrepreneurs with scientific backgrounds (Participants 3 and 4) showed the highest levels of anticipation (17-18%) and joy (9-10%), potentially reflecting their solution-oriented perspectives and investment in innovative approaches to mental healthcare. The social worker (Participant 6) demonstrated the highest levels of disgust (11%) and sadness (12%), perhaps reflecting direct exposure to the human consequences of healthcare system failures.

The nurse (Participant 7), with the least professional experience among participants, showed the highest level of surprise (11%), possibly indicating gaps between theoretical training and clinical realities in CMI management.

Figure 3 illustrates the emotional trajectories across interview narratives, revealing dynamic shifts in emotional tone as discussions progressed through different topics. For all participants, discussions related to current diagnostic procedures and challenges triggered increases in negative emotions (fear, anger, and sadness), while conversations about potential improvements and future directions stimulated positive emotions (trust, anticipation, and joy).

Discussion

Qualitative research methods (Anker-Hansen et al., 2020; Graneheim & Lundman, 2004; McCormack & McCance, 2006) are a challenging affair as they often require manual analysis that is processed iteratively, hence time-consuming. Due to which data analytics is rather restricted to a handful of samples. We were curious to comprehend whether text mining-based approaches and sentiment analysis (Arnold, 2017; Rutkowski et al., 2022; Yu, 2011) can support large-scale qualitative data analysis. This will help identify hurdles faced by patients and professional stakeholders in management of CMIs from a patient-centric perspective. Therefore, we are testing the utility of different lexicons, and then using the best performing lexicon to compare with QCA results published previously.

Comparative analysis across sentiment lexicons

As outlined earlier, we compared sentiment analysis outcomes across different lexicons to evaluate their range and distribution in analysing sentiments related to CMI discourse. Among these, NRC (Mohammad & Turney, 2013) consistently assigned higher positive sentiment scores exceeding 20% in few instances. At the same time, it also captured a notable amount of negative sentiment, though less than the positive. This indicates NRC's ability to reflect a broader range of emotional complexities beyond simple positive/negative classifications. In contrast, AFINN's (Nielsen, 2011) outputs were less disperse with both positive and negative sentiments and leaned towards neutrality. The other two lexicons, Bing (Bing, 2012) and Syuzhet (Jockers, 2020), appeared less sensitive to sentiment variations and predominantly produced neutral scores.

Our findings of low negative sentiment scores agree with previous findings (Naldi, 2019), although the lexicons containing more negative than positive words. This discrepancy may be attributed to the Syuzhet package's limitations in accounting for negators (Hoyeol, 2022; Naldi, 2019) thereby leading to significant differences in sentiment scores across various lexicons alike our observation (**Table 1**). Thus, the choice of sentiment analysis packages should be made carefully, depending on the specific requirements of the discipline. In our case, as we seek to understand the emotional responses of professionals managing CMIs, it was essential to capture a wide range of emotions. NRC's broader sentiment distribution made it particularly well-suited for our qualitative analysis. Additionally, the data revealed variations in sentiment across professional roles: entrepreneurial and scientific roles exhibited higher positivity, while clinical and early-career roles reflected slightly higher negativity. As shown in **Figure 1 and 2**,

the distribution of words categorized into various emotions across professional groups provides a nuanced understanding beyond mere positive, negative, and neutral scores.

Qualitative analytics through sentiment analytics

In a previous report from our team, we conducted qualitative analytics with data obtained from interviews using a semi-structured questionnaire that had three parts in it viz. “tasks related to diagnosis”, “challenges with current procedures”, and “expectations on newer diagnostics”. Such a framework was necessary to investigate the multi-dimensional challenges that the stakeholders face on a day-to-day routine in management of CMIs.

Tasks related to diagnosis

In the early part of the narratives (**Figure 3**) from the trajectory plot, participants P3, and P6 exhibit negative sentiments, likely reflecting frustration, dissatisfaction, or challenges related to the diagnostic tasks. This could stem from procedural inefficiencies, delays, or inaccuracies often associated with the diagnosis of chronic mental illnesses (CMIs). In contrast, P1 demonstrated positive sentiment during this phase, possibly indicating satisfaction or optimism regarding their experiences with diagnostic procedures. The remaining participants displayed a divergent set of emotions suggests that participants had differing levels of exposure to diagnostic systems or varying perspectives based on individual circumstances.

Challenges with current procedures

In the middle phase of the trajectory plot (**Figure 3**), the narratives included responses for questions concerning challenges procedures, there is a significant fluctuation in negative sentiment scores. Participants such as P3, P4, and P5 show pronounced negative sentiment, highlighting dissatisfaction with the limitations or inefficiencies of existing procedures in managing CMIs. This could include issues such as lack of accuracy, accessibility, reimbursements or patient-centred care in the current systems. Notably, P3 and P4, both of who are in the entrepreneurial space appears to experience tonically negative sentiment, suggesting the deep dissatisfaction during this phase. This corroborates strongly with call of lack of infrastructure and development in the innovative space as highlighted in the Draghi report on EU competitiveness (Draghi, 2024) . On the other hand, the narrative from the remaining participants illustrates that these participants while not completely satisfied with the healthcare system perceive fewer challenges and/ or have adapted to the current systems regardless. Their

perspectives may reflect experience, optimism or resilience in navigating the procedural difficulties.

Expectations from newer diagnostic methods

In the final phase of the trajectory plot (**Figure 3**), which focused on narrations that described the expectations from newer diagnostic methods. At this end, there is a noticeable shift toward positivity for all participants but P1, which show a significant increase in positive sentiment, suggesting optimism, or enthusiasm for advancements in diagnostic tools and approaches. This phase may reflect participants' belief in the potential for newer methods to address the shortcomings of current systems and improve the management of CMIs.

However, not all participants experience a positive shift. P3 and P6 exhibit lingering negative sentiments toward the end of the narrative. This could reflect scepticism about the feasibility or effectiveness of newer diagnostic methods, or perhaps a lack of trust in their ability to resolve the challenges highlighted earlier in the narrative. These outliers illustrate that while many participants are optimistic about future developments, some remain cautious or unconvinced.

Qualitative analytics vs sentiment analysis

One of the main objectives of this study is to assess the feasibility of applying text mining techniques to qualitative content analysis. Previous research (Isasi, 2023; Yu, 2011) suggests that text mining and qualitative methods are epistemologically similar, as both share methodological foundations in content analysis, particularly in terms of reliability and validation. Traditional qualitative approaches often rely on labour-intensive, iterative manual analysis, making it challenging to process large text-based datasets. Therefore, text mining provides an interesting alternative, enabling the extraction of meaningful patterns and insights. So, we were curious to validate whether the sentiment analysis conducted in this study is congruent with findings from a prior report on the same dataset.

Firstly, results obtained from qualitative analytics described approximately 30 different themes in patient-centred management approaches that aligns with George Engel's biopsychosocial model (Engel, 1978, 1979) encompassing biological, psychological, and social dimensions. In comparison, key themes that emerged from the sentiment analysis were fewer. We attribute this lower frequency in themes to the limited association with tokenization of negators and overall negative text in the sentiment analysis. However, when comparing the obtained themes (**Figure 4**), for instance, among the biological /medical factors, we observed "combination diagnosis",

"combination therapy", "complexity in diagnosis", "complexity in treatment", "technological requirements" that were previously reported; herein we noticed parallels in themes "medical diagnosis", "treatment and care", and "health and illness". In a similar vein, when comparing social/ societal factors, we observe similarities in management and money-matters such as "legality and bureaucracy", "medical support", "time-effort", "reimbursements", "pricing should not be a factor" in QCA against "time and risk management", "financial concerns" and "healthcare management". The emphasis on patient-centred elements appears prominently in both. On one hand, we observed themes like "quality-of-life," "patient journey," "emotional care," and "vulnerability" under psychological factors in QCA, which directly corresponds to the "emotional well-being" and "relationship and support" emerges as crucial findings from the interviews.

The challenges related to healthcare system navigation are consistently and reliably identified using sentiment analysis.

Strengths and limitations

The application of sentiment analysis to CMI stakeholder interviews demonstrated several notable strengths. First, the methodology provided a systematic framework for detailing emotional dimensions of qualitative data, enabling more objective comparison across participants and topics. The nuanced approach reveals complexities that otherwise might be missed by traditional content analysis alone, for instance, in this case, we noticed the recurrence of the word "cancer" and "tumour". In fact, among chronic diseases and medical care, cancer serves as the gold standard in healthcare practice that integrates multidisciplinary approaches combining specialized diagnostics, personalized treatment protocols, and comprehensive patient support systems. Other chronic illnesses, such as schizophrenia, BD and MDD, will benefit substantially from adopting similar comprehensive frameworks.

Secondly, sentiment analysis identified emotional patterns across professional backgrounds and experience levels, highlighting how different stakeholder positions influence perspectives on CMI management. These insights can inform more targeted approaches to stakeholder engagement and system improvement. For instance, during the upgrade of DSM IV to DSM 5 (Administration, 2016; Bredstrom, 2019) views of social sciences researchers in context to societal and cultural references were incorporated in diagnostic guidelines.

Third, the longitudinal emotional trajectory analysis revealed dynamic shifts in emotional tone as discussions progressed through different topics, providing context for understanding how various aspects of CMI management elicit. This temporal dimension offers valuable insights for structuring interventions and communications in ways that acknowledge and address emotional concerns.

Although, there are several positive factors, sentiment analysis has few shortcomings that need to be considered. First, sentiment analysis relies on pre-defined lexicons with a finite number of words that have been factorized. Hence, then analysis may not fully capture the contextual nuances of specialized healthcare discourse. Negative words are not properly represented and technical terms common in professional practises may carry emotional connotations that vary from day-to-day language, potentially skewing analysis results.

Second, our small sample size ($n=7$) limits the generalizability of findings, particularly given the European focus of participants. As noted in our methodology, the absence of certain professional perspectives (practicing physicians, social scientists in mental health, health insurance specialists, and policy makers) creates potential gaps in our understanding of the emotional landscape surrounding CMI management.

Third, sentiment analysis captures linguistic expressions of emotion but cannot access unexpressed feelings or nonverbal cues that might provide deeper insight into stakeholder perspectives. The methodology also assumes that emotional content in text reflects genuine emotional states, which may not always be accurate, particularly in professional discourse where emotional expression may be modulated by professional norms.

Conclusion and Perspectives

Sentiment analysis of stakeholder interviews revealed critical insights into the emotional landscape surrounding chronic mental illness (CMI) management. Trust emerged as the dominant emotion across stakeholder groups, suggesting a foundation of confidence in mental healthcare fundamentals despite acknowledged limitations. This coexisted with prominent fear, particularly regarding diagnostic accuracy and care continuity, creating a tension between confidence in established frameworks and concern about their practical implementation.

Professional backgrounds significantly influenced emotional responses: clinicians demonstrated higher trust but also more anger, likely reflecting direct involvement in constrained care systems, while those with entrepreneurial backgrounds expressed greater

anticipation and joy when discussing innovation possibilities. Discussions of current challenges consistently elicited negative sentiments, while conversations about innovation stimulated positive emotional responses. Our findings corroborate with a recently published narrative review (Leucht et al., 2024), which is critical of guidelines in psychiatric manuals like DSM and ICD by focusing more on insurance payments rather than patient welfare.

The integration of sentiment analysis into CMI research offers promising opportunities for more patient-centred approaches that address not only clinical needs but also emotional experiences of patients, caregivers, and healthcare providers. Future research and analytics should include broader stakeholder perspectives, including those of a large number of patients and doctors, whose experiences are central to developing truly patient-centred approaches. Such an approach would bridge the gap between qualitative and quantitative analytics, where opinions from several experts at micromanagement levels will be available for policy making and patient-centric approach. With current computation capabilities, it is now possible to handle voluminous data, for instance 13,000 tweets (Shehu et al., 2020) using sentiment analysis and AI-models like random forest, support vector machine, maximum entropy, and decision tree to comprehend public viewpoint. We are firmly of the opinion that including sentiment analysis with other methodologies, including biomarker research, could create comprehensive frameworks addressing both objective and subjective dimensions of CMI experience.

A recent development in line includes incorporating the views of social sciences researchers going from DSM-IV to DSM 5. By acknowledging and addressing the complex emotional landscape surrounding mental healthcare, we can develop more holistic, responsive approaches that better serve individuals living with chronic mental illnesses and those who support them. Sentiment analysis reveals both challenges and opportunities in CMI management, providing valuable emotional dimensions that complement and enhance traditional qualitative analysis approaches.

Acknowledgements

AC acknowledges funding support from the Hanse-Wissenschaftskolleg (HWK), Germany and Helmholtz Information & Data Science Academy - Norwegian Artificial Intelligence Research Consortium (HIDA-NORA) to conduct the research activity. We are extremely grateful to all participants who have shared their professional experience and also sensitive content for the success of this study.

Author contribution

AC, VGS conceptualized the idea. AC conducted the interviews and performed the programming. AC, and VGS analysed and discussed the data. AC drafted the manuscript and VGS reviewed and edited. All authors critically reviewed the manuscript.

References:

- Administration, S. A. a. M. H. S. (2016). Impact of the DSM-IV to DSM-5 Changes on the National Survey on Drug Use and Health [Internet]. Rockville (MD): Substance Abuse and Mental Health Services Administration (US); Table 3.22, DSM-IV to DSM-5 Schizophrenia Comparison. In. <https://www.ncbi.nlm.nih.gov/books/NBK519704/table/ch3.t22/>
- Anker-Hansen, C., Skovdahl, K., McCormack, B., & Tonnessen, S. (2020). Collaboration between home care staff, leaders and care partners of older people with mental health problems: a focus on personhood. *Scand J Caring Sci*, 34(1), 128-138. <https://doi.org/10.1111/scs.12714>
- Arnold, T. a. L. T., L. (2017). Basic Text Processing in R. *Programming Historian*, 6. <https://doi.org/https://doi.org/10.46430/phen0061>
- Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*, Arlington VA, American Psychiatric Association.
- Bing, L. (2012). *Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies, Band 16)*. Morgan & Claypool Publishers (Springer Cham). <https://doi.org/https://doi.org/10.1007/978-3-031-02145-9>
- Bredstrom, A. (2019). Culture and Context in Mental Health Diagnosing: Scrutinizing the DSM-5 Revision. *J Med Humanit*, 40(3), 347-363. <https://doi.org/10.1007/s10912-017-9501-1>
- Cukkemane, A., Hurlemann, R., & Sanchez, V. G. (2025). Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. *Journal of Affective Disorders Reports*, 20, 100893. <https://doi.org/https://doi.org/10.1016/j.jadr.2025.100893>
- Draghi, M. (2024). *The future of European competitiveness: Report by Mario Draghi*. EU Commission. https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en#paragraph_47059

- Engel, G. L. (1978). The biopsychosocial model and the education of health professionals. *Ann N Y Acad Sci*, 310, 169-187. <https://doi.org/10.1111/j.1749-6632.1978.tb22070.x>
- Engel, G. L. (1979). The biopsychosocial model and the education of health professionals. *Gen Hosp Psychiatry*, 1(2), 156-165. [https://doi.org/10.1016/0163-8343\(79\)90062-8](https://doi.org/10.1016/0163-8343(79)90062-8)
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*, 24(2), 105-112. <https://doi.org/10.1016/j.nedt.2003.10.001>
- Hoffmann, T. (2018). “Too many Americans are trapped in fear, violence and poverty”: a psychology-informed sentiment analysis of campaign speeches from the 2016 US Presidential Election. 4(1). <https://doi.org/doi:10.1515/lingvan-2017-0008> (Linguistics Vanguard)
- Hoyeol, K. (2022). Sentiment Analysis: Limits and Progress of the Syuzhet Package and Its Lexicons. *Digital Humans Quarterly*, 16. <https://www.digitalhumanities.org/dhq/vol/16/2/000612/000612.html>
- Isasi, J. (2023). Sentiment Analysis with 'syuzhet' using R. *Programming Historian*, 12. <https://doi.org/https://doi.org/10.46430/phen0110>
- Jockers, M. (2020). *Introduction to the Syuzhet Package (CRAN)*. <https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html>.
- Leucht, S., van Os, J., Jager, M., & Davis, J. M. (2024). Prioritization of Psychopathological Symptoms and Clinical Characterization in Psychiatric Diagnoses: A Narrative Review. *JAMA Psychiatry*, 81(11), 1149-1158. <https://doi.org/10.1001/jamapsychiatry.2024.2652>
- McCormack, B., & McCance, T. V. (2006). Development of a framework for person-centred nursing. *J Adv Nurs*, 56(5), 472-479. <https://doi.org/10.1111/j.1365-2648.2006.04042.x>
- Mohammad, S., & Turney, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*, 29(3), 436-465. <http://dblp.uni-trier.de/db/journals/ci/ci29.html#MohammadT13>
- Naldi, M. (2019). A review of sentiment computation methods with R packages. *arXiv e-prints*, arXiv:1901.08319. <https://doi.org/10.48550/arXiv.1901.08319>
- Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, 93-98. <https://doi.org/https://doi.org/10.48550/arXiv.1103.2903>

- Paul, S. M., & Potter, W. Z. (2023). Finding new and better treatments for psychiatric disorders. *Neuropsychopharmacology*. <https://doi.org/10.1038/s41386-023-01690-5>
- Pv, M., Karmegam, D., & Saravanan, S. (2024). People's Attitude and Perception of the Pandemic on Twitter: A Case Study of COVID-19 in India. *Cureus*, 16(2168-8184 (Print)). <https://doi.org/doi:10.7759/cureus.76320>
- Rutkowski, R. A., Lee, J. D., Collier, R. J., & Werner, N. E. (2022). How can text mining support qualitative data analysis? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1), 2319-2323. <https://doi.org/10.1177/1071181322661535>
- Shehu, H. A., Haidar Sharif, M., Uyaver, S., Tokat, S., & Ramadan, R. A. (2020, 2020//). Sentiment Analysis of Turkish Twitter Data Using Polarity Lexicon and Artificial Intelligence. *Emerging Technologies in Computing*, Cham.
- Smith, D. J., & Craddock, N. (2011). Unipolar and bipolar depression: different of the same? *Br J Psychiatry*, 199(4), 272-274. <https://doi.org/10.1192/bjp.bp.111.092726>
- van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet*, 374(9690), 635-645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8)
- Yu, C. H., Jannasch-Pennell, A., and DiGangi, S. (2011). Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *The Qualitative Report*, 16(3), 730-744. <https://doi.org/https://doi.org/10.46743/2160-3715/2011.1085>

Table 1: Summary of sentiment analysis across professional stakeholders (*Participants with friends and family members affected by CMIs)

	Positive	Negative	Neutral
P1 (Sex = Male; Education = MD, Ph.D; Specialization = Psychiatry; Current Role = Scientist; Previous Role = Clinician; Experience (in year) > 20)			
NRC	0.274648	-0.08451	0.640845
afinn	0.053883	-0.02324	0.922874
bing	0.029054	-0.01321	0.957739
syuzhet	0.039937	-0.01442	0.945642
P2* (Sex = Male; Education = MD, Ph.D; Specialization = Neurosciences; Current Role = Scientist; Previous Role = Clinician; Experience (in year) > 15)			
NRC	0.263576	-0.09801	0.638411
afinn	0.096886	-0.0346	0.868512
bing	0.052595	-0.01765	0.929758
syuzhet	0.042353	-0.01952	0.938131
P3* (Sex = Male; Education = Ph.D, MBA; Specialization = Biochemistry; Current Role = CEO - Entrepreneurship; Previous Role = Scientist; Experience (in year) > 15)			
NRC	0.229752	-0.12397	0.646281
afinn	0.077807	-0.03808	0.88411
bing	0.041366	-0.02397	0.934668
syuzhet	0.037886	-0.01814	0.943976
P4 (Sex = Male; Education = Ph.D; Specialization = Medicinal Chemistry; Current Role = COO - Entrepreneurship; Previous Role = Scientist; Experience (in year) > 10)			
NRC	0.287719	-0.07485	0.637427
afinn	0.064182	-0.01851	0.917304
bing	0.034806	-0.01333	0.951864
syuzhet	0.044495	-0.00823	0.947272
P5 (Sex = Male; Education = Ph.D; Specialization = Psychiatry; Current Role = MD Pharmaceutics; Previous Role = Scientist; Experience (in year) > 20)			
NRC	0.215479	-0.12401	0.66051
afinn	0.051554	-0.0404	0.908044
bing	0.027239	-0.0192	0.953565
syuzhet	0.030146	-0.02042	0.949433
P6* (Sex = Female; Education = MS; Specialization = Informatics; Current Role = CEO - Social entrepreneurship; Previous Role = IT specialist; Experience (in year) > 3)			
NRC	0.24237	-0.10413	0.653501
afinn	0.064504	-0.02212	0.91338
bing	0.039071	-0.01511	0.945816
syuzhet	0.040509	-0.01426	0.945227
P7 (Sex = Male; Education = BSN; Specialization = Nursing; Current Role = Nurse; Previous Role = None; Experience (in year) < 1)			
NRC	0.237288	-0.13877	0.623941
afinn	0.101779	-0.03874	0.859486
bing	0.053755	-0.0249	0.921344
syuzhet	0.042401	-0.0218	0.9358

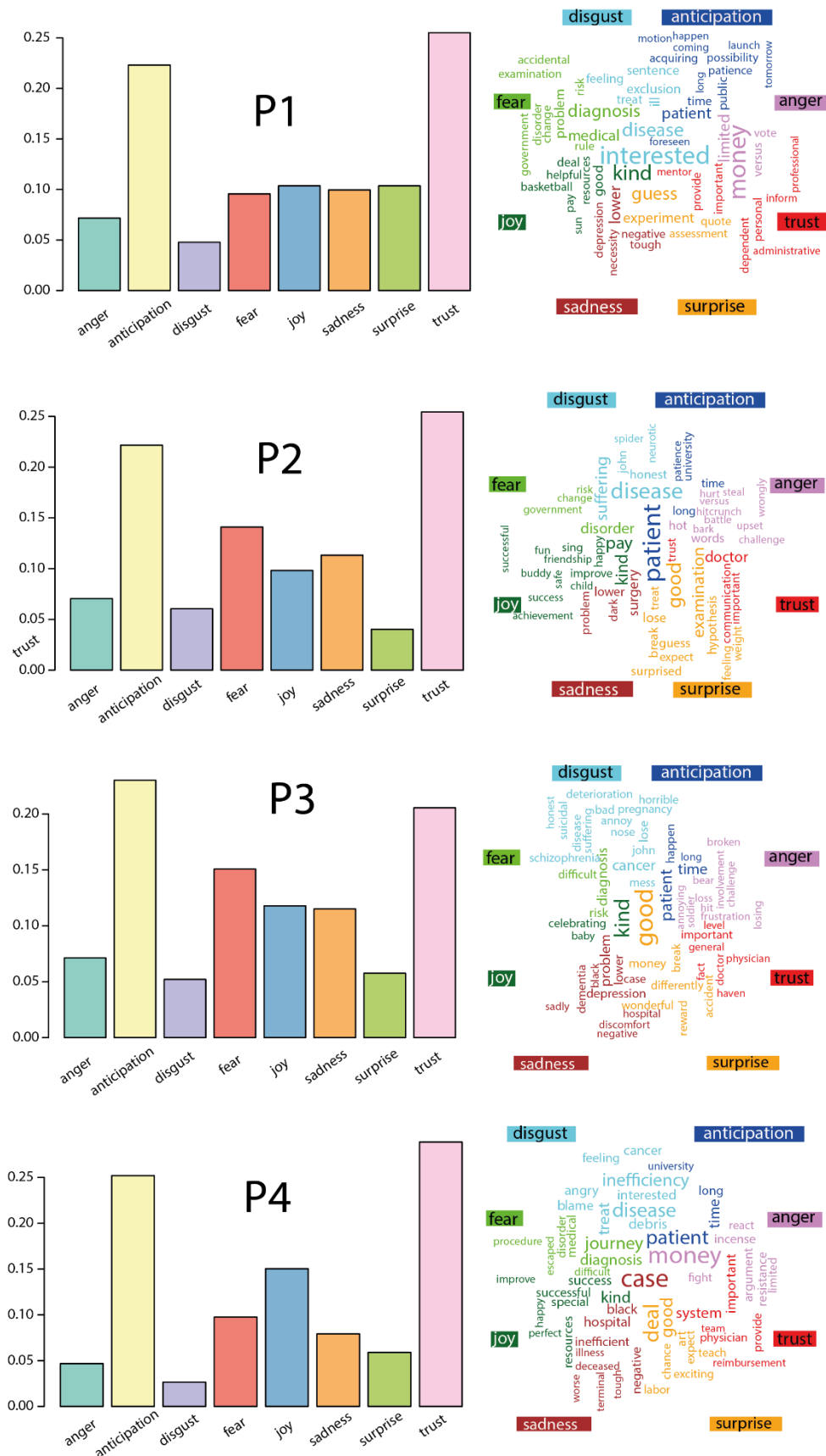


Figure 1: Bar chart showing the calculated scores of the eight emotions from the NRC lexicon and word cloud for participants 1-4 of the most frequent words corresponding to the same emotion set.

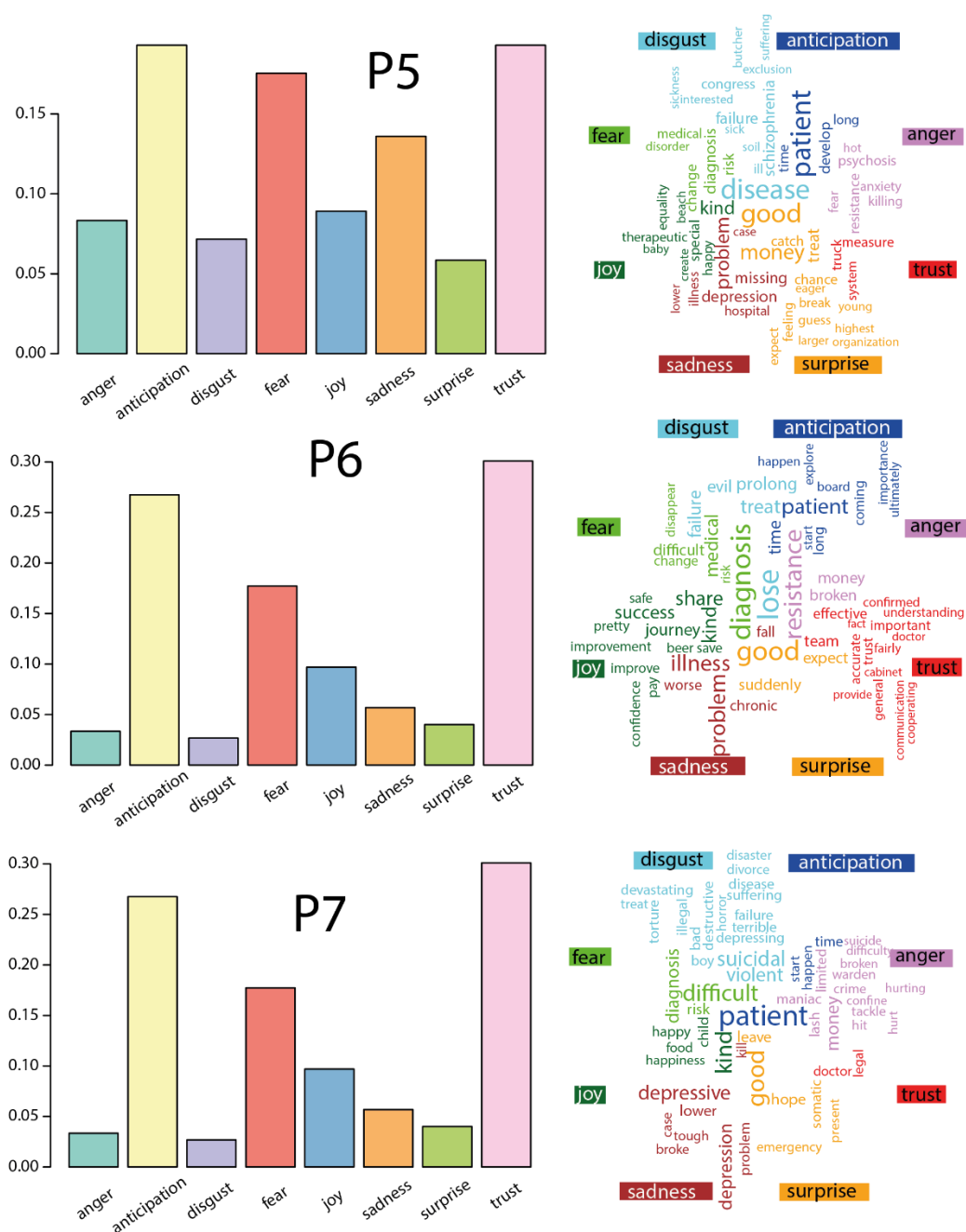


Figure 2: Bar chart displaying the computed scores of eight emotions from the NRC lexicon, alongside a word cloud for participants 5-7 highlighting the most frequently used words associated with these emotions.

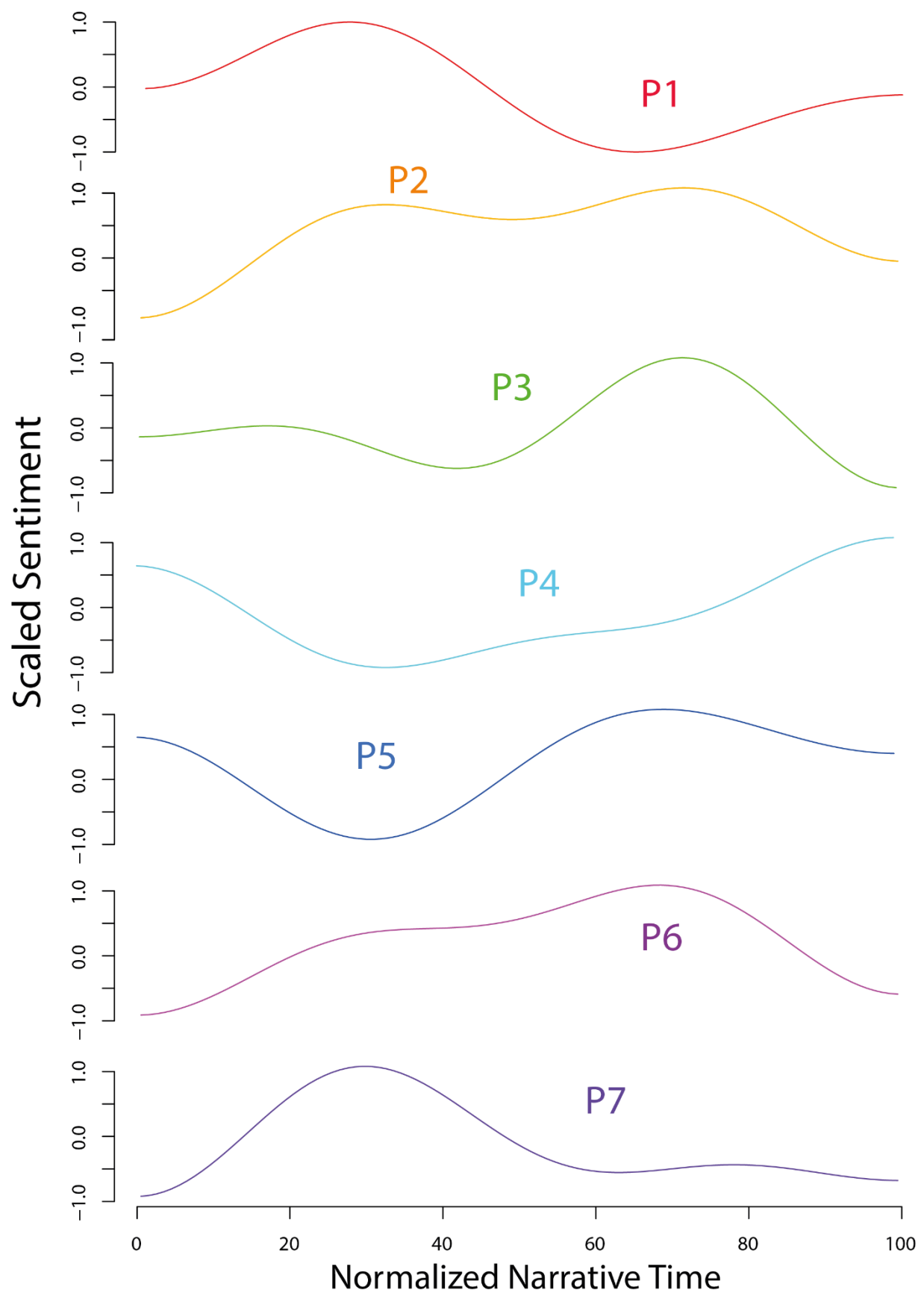


Figure 3: Evolution of sentiments (positive, negative and neutral) the during the course of the interview for the seven participants , which was scored on the basis of the NRC lexicon.

Qualitative Analytics – Biopsychosocial model *		
<u>Biological / medical factors</u> - Guidelines - Questionnaires - Complexity in diagnosis - Complexity in treatment - Lack of standardization - Resource management - Combination diagnosis - Combination therapy - Technological requirements	<u>Social / societal factors</u> - Legality and bureaucracy - Diverse roles in profession - Pricing should not be a factor - Ineffective medical support - Right of treatment - Medical support - Reimbursements - Time-effort	<u>Psychological / psychosocial factors</u> - Disabled for life - Lack of trust - Trust - quality-of-life - patient journey - Emotional care - Vulnerability - Patient centric - Effort

Emotion and Sentiment Analysis		
<u>Biological / medical factors</u> - Medical diagnosis and treatment - Healthcare - Health and illness - Treatment and care - Resources and support	<u>Social / societal factors</u> - Financial concerns - Healthcare management - Communication and support - Problem solving and achievement - Time and risk management - Problem solving and improvement	<u>Psychological / psychosocial factors</u> - Emotional well-being - Support and relationship

Figure 4: Comparison of themes derived from expert-based qualitative analytics (Cukkemane et al., 2025) versus sentiment analytics.

Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics

Abhishek Cukkemane^{1,2,3*} and Veralia Gabriela Sanchez⁴

¹Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

³Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

⁴Faculty of Technology, Natural Sciences and Maritime Sciences, University of South-Eastern Norway, Vestfold, Norway

*Address for correspondence - a.cukkemane@fz-juelich.de

R-script for syuzhet_sentiment_analysis

```
# Load Required Libraries
library(syuzhet)
library(RColorBrewer)
library(wordcloud)
library(tm)

# 1. Data Input -----
setwd("filepath")
text.strings <- list()
for (i in 1:7) {
  file_name <- paste0("Participant", sprintf("%02d", i), "_transcript_sentence.txt")
  text.strings[[i]] <- scan(file = file_name, what = character(), sep = "\n", allowEscapes = TRUE)
}

# Looping through each participant's text string for tokenizing words
text.words <- list()
for (i in 1:7) {
  text.words[[i]] <- get_tokens(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(text.words[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(text.words[[i]], 100), "\n\n")
}

# 2. Sentiment Analysis with Syuzhet -----
# Initialize a summary dataframe
```

```

sentiment.summary.syuzhet <- data.frame(Participant = character(),
                                         Positive = numeric(),
                                         Negative = numeric(),
                                         Neutral = numeric(),
                                         Total_Sentiment = numeric(),
                                         stringsAsFactors = FALSE)

# Loop through each participant's text strings
for (i in 1:length(text.words)) {
  # Get the current participant's texts
  texts <- text.words[[i]]

  # Skip empty transcripts
  if (length(texts) == 0) next

  # Get sentiment scores using syuzhet lexicon
  sentiment_scores <- get_sentiment(texts, method = "syuzhet")

  # Calculate proportions
  positive_count <- sum(sentiment_scores[sentiment_scores > 0])
  negative_count <- sum(sentiment_scores[sentiment_scores < 0])
  total_count <- length(sentiment_scores)

  neutral_count <- total_count - (positive_count + abs(negative_count))

  # Store results in the summary dataframe
  sentiment.summary.syuzhet <- rbind(sentiment.summary.syuzhet, data.frame(
    Participant = paste("Participant", i),
    Positive = positive_count / total_count,
    Negative = negative_count / total_count,
    Neutral = neutral_count / total_count,
    Total_Sentiment = sum(sentiment_scores),
    stringsAsFactors = FALSE
  ))

  # Save plots as PNG and PDF
  png(filename = paste0("Emotional_Arc_Syuzhet_Participant_", i, ".png"), width = 800, height = 600)
  plot(sentiment_scores, type = "l", main = paste("Emotional Arc (Syuzhet) for Participant", i),
       xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
  dev.off()

  pdf(file = paste0("Emotional_Arc_Syuzhet_Participant_", i, ".pdf"), width = 8, height = 6)
  plot(sentiment_scores, type = "l", main = paste("Emotional Arc (Syuzhet) for Participant", i),
       xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
  dev.off()

  # Additional plots
  png(filename = paste0("Emotional_Arc_Syuzhet_Histogram_Participant_", i, ".png"), width = 800,
      height = 600)
  plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (Syuzhet) for
Participant", i),
       xlab = "Sentences", ylab = "Sentiment Score", col = "red")
  dev.off()

```

```

pdf(file = paste0("Emotional_Arc_Syuzhet_Histogram_Participant_", i, ".pdf"), width = 8, height = 6)
plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (Syuzhet) for
Participant", i),
      xlab = "Sentences", ylab = "Sentiment Score", col = "red")
dev.off()

ft_values <- get_transformed_values(sentiment_scores, low_pass_size = 3,
                                   x_reverse_len = 100, padding_factor = 2, scale_vals = TRUE, scale_range =
FALSE)

png(filename = paste0("Emotional_Arc_Syuzhet_Transformed_Participant_", i, ".png"), width = 800,
height = 600)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (Syuzhet) for Participant", i),
      xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()

pdf(file = paste0("Emotional_Arc_Syuzhet_Transformed_Participant_", i, ".pdf"), width = 8, height =
6)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (Syuzhet) for Participant", i),
      xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()
}

# View the sentiment summary
write.csv(sentiment.summary.syuzhet, file = "sentiment.summary.syuzhet.csv", row.names = T)

```

R-script for NRC_sentiment_analysis

```
# Load the Packages
library(syuzhet)
library(RColorBrewer)
library(wordcloud)
library(tm)

# 1. Data Input -----
setwd("filepath")
text.strings <- list()
for (i in 1:7) {
  file_name <- paste0("Participant", sprintf("%02d", i), "_transcript_sentence.txt")
  text.strings[[i]] <- scan(file = file_name, what = character(), sep = "\n", allowEscapes = TRUE)
}

# Looping through each participant's text string for tokenizing words
text.words <- list()
for (i in 1:7) {
  text.words[[i]] <- get_tokens(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(text.words[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(text.words[[i]], 100), "\n\n")
}

#####
#####
#####
#Sentiment Analysis for Positive and Negative Valence -----
# Sentiment Analysis for Positive, Negative, and Neutral Valence -----
sentiment.valance.summary <- data.frame(Participant=character(),
                                         Positive=numeric(),
                                         Negative=numeric(),
                                         Neutral=numeric(),
                                         Total=numeric(),
                                         Positive_Proportion=numeric(),
                                         Negative_Proportion=numeric(),
                                         Neutral_Proportion=numeric(),
                                         stringsAsFactors=FALSE)

for (i in 1:7) {
  # Get sentiment scores using NRC lexicon
  sentiment.words <- get_nrc_sentiment(text.words[[i]], lang = "english")

  # Calculate positive, negative, and neutral counts
  positive_count <- sum(sentiment.words[, "positive"])
  negative_count <- sum(sentiment.words[, "negative"])
  total_count <- sum(sentiment.words)
  neutral_count <- total_count - (positive_count + negative_count)

  # Calculate proportions
  positive_proportion <- ifelse(total_count > 0, positive_count / total_count, 0)
  negative_proportion <- ifelse(total_count > 0, negative_count / total_count, 0)
  neutral_proportion <- ifelse(total_count > 0, neutral_count / total_count, 0)
```

```

# Store results in the summary dataframe
sentiment.valance.summary <- rbind(sentiment.valance.summary,
                                   data.frame(Participant=sprintf("%02d", i),
                                             Positive=positive_count,
                                             Negative=negative_count,
                                             Neutral=neutral_count,
                                             Total=total_count,
                                             Positive_Proportion=positive_proportion,
                                             Negative_Proportion=negative_proportion,
                                             Neutral_Proportion=neutral_proportion,
                                             stringsAsFactors=FALSE))
}

sentiment.valance.summary
write.csv(sentiment.valance.summary, file = "sentiment_summary_positive_negative_neutral.csv",
row.names = FALSE)

# Looping through each participant's tokens and calculate sentiment scores
# Visualizing sentiment scores with bar plots
sentiment.words <- list()
sentiment.summary <- data.frame(Participant=character(), Anger=numeric(), Anticipation=numeric(),
                                Disgust=numeric(), Fear=numeric(), Joy=numeric(),
                                Sadness=numeric(), Surprise=numeric(), Trust=numeric(),
                                stringsAsFactors=FALSE)

for (i in 1:7) {
  sentiment.words[[i]] <- get_nrc_sentiment(text.words[[i]], lang = "english")
  sentiment.data.word <- colSums(prop.table(sentiment.words[[i]][, 1:8]))
  # Store the sentiment summary for the current participant
  sentiment.summary <- rbind(sentiment.summary,
                             data.frame(Participant=sprintf("%02d", i),
                                         t(sentiment.data.word),
                                         stringsAsFactors=FALSE))

  # Create bar plot
  barplot(sentiment.data.word, space = 0.2, horiz = FALSE,
          las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
          main = paste("Transcript Participant (Words)", sprintf("%02d", i)),
          sub = "Qualitative Analytics - Schizophrenia",
          xlab = "Emotions", ylab = NULL)
  # File names for saving *.png and *.pdf
  png_filename <- paste("Word.Sentiment.Bar.Plot.Participant", sprintf("%02d", i), ".png", sep = "")
  pdf_filename <- paste("Word.Sentiment.Bar.Plot.Participant", sprintf("%02d", i), ".pdf", sep = "")

  # Saving as PNG
  png(filename = png_filename, width = 800, height = 600)
  barplot(sentiment.data.word, space = 0.2, horiz = FALSE,
          las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
          main = paste("Transcript Participant (Words)", sprintf("%02d", i)),
          sub = "Qualitative Analytics - Schizophrenia",
          xlab = "Emotions", ylab = NULL)
  dev.off()
  # Saving as PDF

```



```

pdf(file = pdf_filename, width = 8, height = 6)
barplot(sentiment.data.word, space = 0.2, horiz = FALSE,
        las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
        main = paste("Transcript Participant (Words)", sprintf("%02d", i)),
        sub = "Qualitative Analytics - Schizophrenia",
        xlab = "Emotions", ylab = NULL)
dev.off()
}

write.csv(sentiment.summary, file = "sentiment.summary.nrc.csv", row.names = T)

# Looping through each participant's tokenized words and sentiment scores
# Visualizing sentiment scores with word-cloud (comparison cloud)
cloud.emotions.data <- list()
cloud.tdm <- list()

for (i in 1:7) {
  cloud.emotions.data[[i]] <- c(
    paste(text.words[[i]][sentiment.words[[i]]$anger > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$anticipation > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$disgust > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$fear > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$joy > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$sadness > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$surprise > 0], collapse = " "),
    paste(text.words[[i]][sentiment.words[[i]]$trust > 0], collapse = " ")
  )
  cloud.emotions.data[[i]] <- iconv(cloud.emotions.data[[i]])
  cloud.corpus <- Corpus(VectorSource(cloud.emotions.data[[i]]))
  cloud.tdm[[i]] <- TermDocumentMatrix(cloud.corpus)
  cloud.tdm[[i]] <- as.matrix(cloud.tdm[[i]])
  colnames(cloud.tdm[[i]]) <- c('anger', 'anticipation', 'disgust', 'fear', 'joy', 'sadness', 'surprise', 'trust')

  # Create a data frame for saving emotions data
  emotions_df <- data.frame(t(cloud.tdm[[i]]), stringsAsFactors = FALSE)
  emotions_df$Participant <- sprintf("%02d", i) # Add participant identifier

  # Save the emotions data to a CSV file
  csv_filename <- paste("Emotions_Data_Participant", sprintf("%02d", i), ".csv", sep = "")
  write.csv(emotions_df, file = csv_filename, row.names = FALSE)
}

write.csv(emotions_df, "emotions_df.csv")

# Set seed for reproducibility & loop through each participant to create comparison clouds
set.seed(99)
for (i in 1:7) {
  comparison.cloud(cloud.tdm[[i]], random.order = FALSE,
    colors = c("violet", "blue", "cyan", "green", "darkgreen", "brown", "orange", "red"),
    title.size = 1, max.words = 60, scale = c(2.5, 1), rot.per = 0.4,
    main = paste("Comparison Cloud for Participant", sprintf("%02d", i)))

  # File names for saving in .pdf and .png formats

```

```

png_filename <- paste("Comparison_Cloud_Participant", sprintf("%02d", i), ".png", sep = "")
pdf_filename <- paste("Comparison_Cloud_Participant", sprintf("%02d", i), ".pdf", sep = "")

# Save as PNG
png(filename = png_filename, width = 800, height = 600)
comparison.cloud(cloud.tdm[[i]], random.order = FALSE,
                 colors = c("violet", "blue", "cyan", "green", "darkgreen", "brown", "orange", "red"),
                 title.size = 1, max.words = 60, scale = c(2.5, 1), rot.per = 0.4,
                 main = paste("Comparison Cloud for Participant", sprintf("%02d", i)))
dev.off()

# Save as PDF
pdf(file = pdf_filename, width = 8, height = 6)
comparison.cloud(cloud.tdm[[i]], random.order = FALSE,
                 colors = c("violet", "blue", "cyan", "green", "darkgreen", "brown", "orange", "red"),
                 title.size = 1, max.words = 60, scale = c(2.5, 1), rot.per = 0.4,
                 main = paste("Comparison Cloud for Participant", sprintf("%02d", i)))
dev.off()
}

# Set margin parameters for the plots & loop through each participant to create sentiment valence
plots
#Visualizing sentiment valence plots in raw and normalized scales
par(mar = c(4, 4, 2, 1)) # Adjust bottom, left, top, right margins
for (i in 1:7) {
  # Calculate sentiment valence for the current participant
  sentiment.valence.word <- (sentiment.words[[i]]$negative * -1) + sentiment.words[[i]]$positive
  simple_plot(sentiment.valence.word,
              title = paste("NRC Plot (Words) - Participant", sprintf("%02d", i)),
              legend_pos = "top",
              lps = 10,
              window = 0.1)
  # Specify file names for saving as png and pdf
  png_filename <- paste("NRC.Plot.Words.Participant", sprintf("%02d", i), ".png", sep = "")
  pdf_filename <- paste("NRC.Plot.Words.Participant", sprintf("%02d", i), ".pdf", sep = "")

  # Save as PNG
  png(filename = png_filename, width = 800, height = 600)
  simple_plot(sentiment.valence.word,
              title = paste("NRC Plot (Sentences) - Participant", sprintf("%02d", i)),
              legend_pos = "top",
              lps = 10,
              window = 0.1)
  dev.off()
  # Save as PDF
  pdf(file = pdf_filename, width = 8, height = 6)
  simple_plot(sentiment.valence.word,
              title = paste("NRC Plot (Sentences) - Participant", sprintf("%02d", i)),
              legend_pos = "top",
              lps = 10,
              window = 0.1)
  dev.off()
}

```

```
#####
#####
#####
#####
```

```
# Loop through each participant's text string for tokenized sentence
sentence.vectors <- list()
for (i in 1:7) {
  sentence.vectors[[i]] <- get_sentences(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(sentence.vectors[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(sentence.vectors[[i]], 100), "\n\n")
}
```

```
# Looping through each participant's text string for tokenized sentences
text.sentences <- list()
for (i in 1:7) {
  text.sentences[[i]] <- get_sentences(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(text.sentences[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(text.sentences[[i]], 100), "\n\n")
}
```

```
# Loop through each participant's tokens and calculate sentiment scores
# Visualizing the output of sentiment scores using bar plots
sentiment.sentences <- list()
for (i in 1:7) {
  sentiment.sentences[[i]] <- get_nrc_sentiment(text.sentences[[i]], lang = "english")
  sentiment.data.sentence <- colSums(prop.table(sentiment.sentences[[i]][, 1:8]))
}
```

```
# Create the bar plot
barplot(sentiment.data.sentence, space = 0.2, horiz = FALSE,
        las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
        main = paste("Transcript Participant (sentences)", sprintf("%02d", i)),
        sub = "Qualitative Analytics - Schizophrenia",
        xlab = "Emotions", ylab = NULL)
```

```
# Specify file names for saving as png and pdf
png_filename <- paste("Sentence.Sentiment.Bar.Plot.Participant", sprintf("%02d", i), ".png", sep =
"")
pdf_filename <- paste("Sentence.Sentiment.Bar.Plot.Participant", sprintf("%02d", i), ".pdf", sep = "")
```

```
# Save as PNG
png(filename = png_filename, width = 800, height = 600)
barplot(sentiment.data.sentence, space = 0.2, horiz = FALSE,
        las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
        main = paste("Transcript Participant (Sentences)", sprintf("%02d", i)),
        sub = "Qualitative Analytics - Schizophrenia",
        xlab = "Emotions", ylab = NULL)
dev.off()
# Save as PDF
pdf(file = pdf_filename, width = 8, height = 6)
barplot(sentiment.data.sentence, space = 0.2, horiz = FALSE,
```

```

        las = 1, cex.names = 0.7, col = brewer.pal(n = 8, name = "Set3"),
        main = paste("Transcript Participant (Sentences)", sprintf("%02d", i)),
        sub = "Qualitative Analytics - Schizophrenia",
        xlab = "Emotions", ylab = NULL)
    dev.off()
}

# Set margin parameters for the plots & loop through each participant to create sentiment valence
plots
# Visualizing sentiment valence plots
par(mar = c(4, 4, 2, 1)) # Adjust bottom, left, top, right margins
for (i in 1:7) {
  # Calculate sentiment valence for the current participant
  sentiment.valence.sentence <- (sentiment.sentences[[i]]$negative * -1) +
    sentiment.sentences[[i]]$positive

  #Valence plot
  simple_plot(sentiment.valence.sentence,
    title = paste("NRC Plot (Sentences) - Participant", sprintf("%02d", i)),
    legend_pos = "top",
    lps = 10,
    window = 0.1)

  # Specify file names as png and pdf
  png_filename <- paste("NRC.Plot.Sentences.Participant", sprintf("%02d", i), ".png", sep = "")
  pdf_filename <- paste("NRC.Plot.Sentences.Participant", sprintf("%02d", i), ".pdf", sep = "")

  # Save as PNG
  png(filename = png_filename, width = 800, height = 600)
  simple_plot(sentiment.valence.sentence,
    title = paste("NRC Plot (Sentences) - Participant", sprintf("%02d", i)),
    legend_pos = "top",
    lps = 10,
    window = 0.1)
  dev.off()
  # Save as PDF
  pdf(file = pdf_filename, width = 8, height = 6)
  simple_plot(sentiment.valence.sentence,
    title = paste("NRC Plot (Sentences) - Participant", sprintf("%02d", i)),
    legend_pos = "top",
    lps = 10,
    window = 0.1)
  dev.off()
}

```

R-script for afinn_sentiment_analysis

```
# Load Required Libraries
library(syuzhet)
library(RColorBrewer)
library(wordcloud)
library(tm)

# 1. Data Input -----
setwd("filepath")
text.strings <- list()
for (i in 1:7) {
  file_name <- paste0("Participant", sprintf("%02d", i), "_transcript_sentence.txt")
  text.strings[[i]] <- scan(file = file_name, what = character(), sep = "\n", allowEscapes = TRUE)
}

# Looping through each participant's text string for tokenizing words
text.words <- list()
for (i in 1:7) {
  text.words[[i]] <- get_tokens(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(text.words[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(text.words[[i]], 100), "\n\n")
}

# 2. Sentiment Analysis with afinn -----
# Initialize a summary dataframe
sentiment.summary.afinn <- data.frame(Participant = character(),
                                     Positive = numeric(),
                                     Negative = numeric(),
                                     Neutral = numeric(),
                                     Total_Sentiment = numeric(),
                                     stringsAsFactors = FALSE)

# Loop through each participant's text strings
for (i in 1:length(text.words)) {
  # Get the current participant's texts
  texts <- text.words[[i]]

  # Skip empty transcripts
  if (length(texts) == 0) next

  # Get sentiment scores using afinn lexicon
  sentiment_scores <- get_sentiment(texts, method = "afinn")

  # Calculate proportions
  positive_count <- sum(sentiment_scores[sentiment_scores > 0])
  negative_count <- sum(sentiment_scores[sentiment_scores < 0])
  total_count <- length(sentiment_scores)

  neutral_count <- total_count - (positive_count + abs(negative_count))

  # Store results in the summary dataframe
```

```

sentiment.summary.afinn <- rbind(sentiment.summary.afinn, data.frame(
  Participant = paste("Participant", i),
  Positive = positive_count / total_count,
  Negative = negative_count / total_count,
  Neutral = neutral_count / total_count,
  Total_Sentiment = sum(sentiment_scores),
  stringsAsFactors = FALSE
))

# Save plots as PNG and PDF
png(filename = paste0("Emotional_Arc_afinn_Participant_", i, ".png"), width = 800, height = 600)
plot(sentiment_scores, type = "l", main = paste("Emotional Arc (afinn) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
dev.off()

pdf(file = paste0("Emotional_Arc_afinn_Participant_", i, ".pdf"), width = 8, height = 6)
plot(sentiment_scores, type = "l", main = paste("Emotional Arc (afinn) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
dev.off()

# Additional plots
png(filename = paste0("Emotional_Arc_afinn_Histogram_Participant_", i, ".png"), width = 800,
height = 600)
plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (afinn) for Participant",
i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "red")
dev.off()

pdf(file = paste0("Emotional_Arc_afinn_Histogram_Participant_", i, ".pdf"), width = 8, height = 6)
plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (afinn) for Participant",
i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "red")
dev.off()

ft_values <- get_transformed_values(sentiment_scores, low_pass_size = 3,
                                   x_reverse_len = 100, padding_factor = 2, scale_vals = TRUE, scale_range =
FALSE)

png(filename = paste0("Emotional_Arc_afinn_Transformed_Participant_", i, ".png"), width = 800,
height = 600)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (afinn) for Participant", i),
     xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()

pdf(file = paste0("Emotional_Arc_afinn_Transformed_Participant_", i, ".pdf"), width = 8, height = 6)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (afinn) for Participant", i),
     xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()
}

# View the sentiment summary
write.csv(sentiment.summary.afinn, file = "sentiment.summary.afinn.csv", row.names = T)

```

R-script for bing_sentiment_analysis

```
# Load Required Libraries
library(syuzhet)
library(RColorBrewer)
library(wordcloud)
library(tm)

# 1. Data Input -----
setwd("filepath")
text.strings <- list()
for (i in 1:7) {
  file_name <- paste0("Participant", sprintf("%02d", i), "_transcript_sentence.txt")
  text.strings[[i]] <- scan(file = file_name, what = character(), sep = "\n", allowEscapes = TRUE)
}

# Looping through each participant's text string for tokenizing words
text.words <- list()
for (i in 1:7) {
  text.words[[i]] <- get_tokens(text.strings[[i]])
  cat("Participant", i, "Token Count:", length(text.words[[i]]), "\n")
  cat("First 100 tokens of Participant", i, "Transcript:\n", head(text.words[[i]], 100), "\n\n")
}

# 2. Sentiment Analysis with bing -----
# Initialize a summary dataframe
sentiment.summary.bing <- data.frame(Participant = character(),
                                     Positive = numeric(),
                                     Negative = numeric(),
                                     Neutral = numeric(),
                                     Total_Sentiment = numeric(),
                                     stringsAsFactors = FALSE)

# Loop through each participant's text strings
for (i in 1:length(text.words)) {
  # Get the current participant's texts
  texts <- text.words[[i]]

  # Skip empty transcripts
  if (length(texts) == 0) next

  # Get sentiment scores using bing lexicon
  sentiment_scores <- get_sentiment(texts, method = "bing")

  # Calculate proportions
  positive_count <- sum(sentiment_scores[sentiment_scores > 0])
  negative_count <- sum(sentiment_scores[sentiment_scores < 0])
  total_count <- length(sentiment_scores)

  neutral_count <- total_count - (positive_count + abs(negative_count))

  # Store results in the summary dataframe
```

```

sentiment.summary.bing <- rbind(sentiment.summary.bing, data.frame(
  Participant = paste("Participant", i),
  Positive = positive_count / total_count,
  Negative = negative_count / total_count,
  Neutral = neutral_count / total_count,
  Total_Sentiment = sum(sentiment_scores),
  stringsAsFactors = FALSE
))

# Save plots as PNG and PDF
png(filename = paste0("Emotional_Arc_bing_Participant_", i, ".png"), width = 800, height = 600)
plot(sentiment_scores, type = "l", main = paste("Emotional Arc (bing) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
dev.off()

pdf(file = paste0("Emotional_Arc_bing_Participant_", i, ".pdf"), width = 8, height = 6)
plot(sentiment_scores, type = "l", main = paste("Emotional Arc (bing) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "blue")
dev.off()

# Additional plots
png(filename = paste0("Emotional_Arc_bing_Histogram_Participant_", i, ".png"), width = 800,
height = 600)
plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (bing) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "red")
dev.off()

pdf(file = paste0("Emotional_Arc_bing_Histogram_Participant_", i, ".pdf"), width = 8, height = 6)
plot(sentiment_scores, type = "h", main = paste("Emotional Arc Histogram (bing) for Participant", i),
     xlab = "Sentences", ylab = "Sentiment Score", col = "red")
dev.off()

ft_values <- get_transformed_values(sentiment_scores, low_pass_size = 3,
                                   x_reverse_len = 100, padding_factor = 2, scale_vals = TRUE, scale_range =
FALSE)

png(filename = paste0("Emotional_Arc_bing_Transformed_Participant_", i, ".png"), width = 800,
height = 600)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (bing) for Participant", i),
     xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()

pdf(file = paste0("Emotional_Arc_bing_Transformed_Participant_", i, ".pdf"), width = 8, height = 6)
plot(ft_values, type = "l", main = paste("Transformed Emotional Arc (bing) for Participant", i),
     xlab = "Normalized Narrative Time", ylab = "Sentiment Score", col = "green")
dev.off()
}

# View the sentiment summary
write.csv(sentiment.summary.bing, file = "sentiment.summary.bing.csv", row.names = T)

```


3.3 Perspective: The challenges, responsibilities, and benefits for society in addressing schizophrenia

Despite their high prevalence and profound impact on patient well-being and health systems (OECD/EU, 2018; Wiegand et al., 2020), CMIs not only possess a considerable global burden and continues to rely on empirical symptom-based diagnostics due to a lack of objective biomarkers. This is in stark contrast with other medical disciplines (Chartogne et al., 2021; Lauriola & Tomai, 2019; Tomas et al., 2023) where biomarker-based diagnosis and personalized treatments are increasingly mainstream. W.r.t, my approach sought to understand how may we bridge this gap?

By investigating the view-points of key stakeholders in clinicians, academic researchers, diagnostic entrepreneurs, and industrial scientists on adopting a patient-centric model for managing CMIs. Hence, I utilized integrative strategy by combining qualitative analytics (Elo & Kyngas, 2008; Fossey et al., 2002; Graneheim & Lundman, 2004) with advanced computational tools such as natural language processing (NLP (Isasi, 2023; Mohammad & Turney, 2013)) through structured interviews. Many responses were open-ended, as were the questions; the coding and categorization process involved identifying nodes i.e., recurrent themes, which were subcategorized and conceptually linked. By employing such as strategy, using QCA, I was able to systematically classify the textual data into themes matching social /biological /psychosocial constructs (Fig 1) that aligned with the biopsychosocial model (Engel, 1978, 1979) proposed by Dr. George Engel 40 years ago.

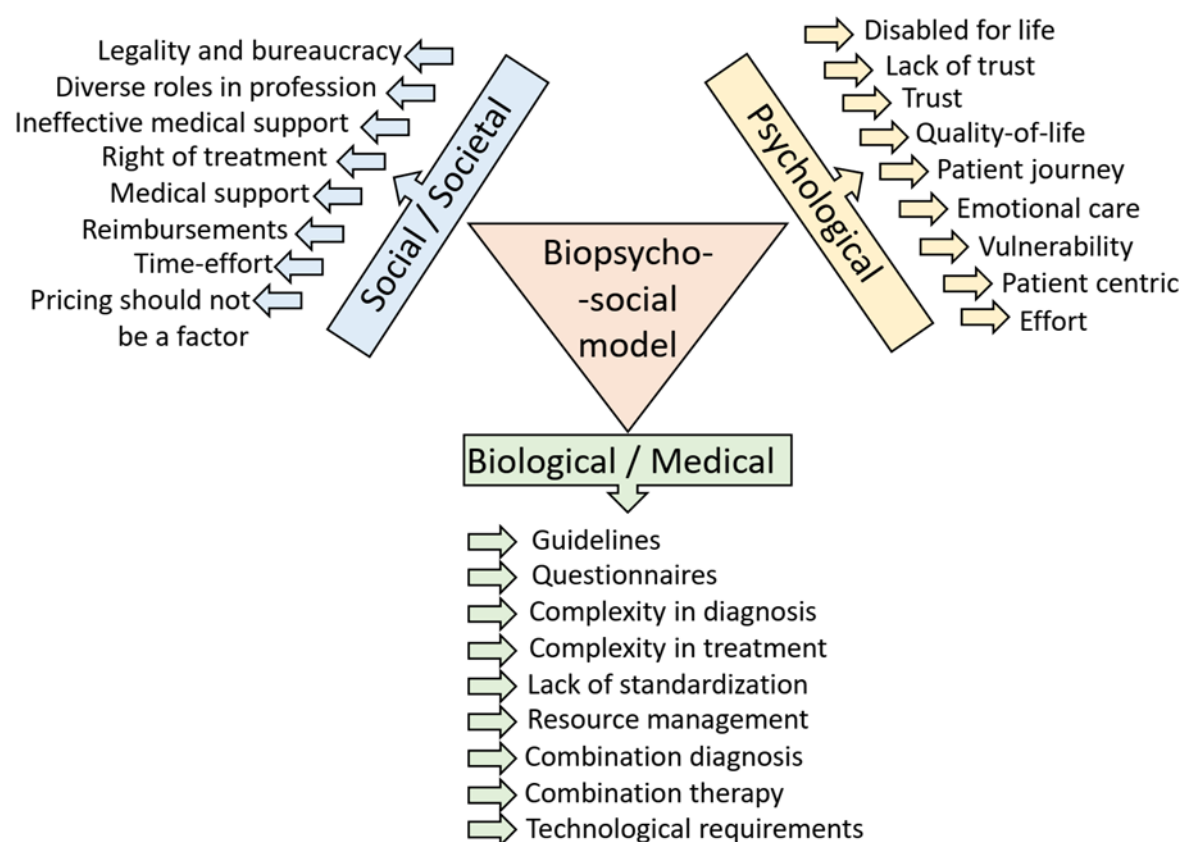


Fig. 1. Schematic representation of the identified codes in the theme of George Engel's biopsychosocial model.

Some of the emerging themes and insights that were particularly critical to us, the scientists in natural sciences were:

- Lack of Diagnostic Objectivity
- Emotional Trust and Therapeutic Alliance
- Technological Integration
- Expectations from Personalized Medicine
- Advantages of the Qualitative Approach

In a nutshell, qualitative analytics provided us (Cukkemane et al., 2025) with gap analytics in psychiatric care, opportunities and potential recommendation to address them. Its integration with emerging computational methods such as natural language processing (NLP) has the potential to further amplify similar qualitative analytics and also enable statistical generalizations. Such hybrid models have the potential of bridging the qualitative-quantitative divide offering a more comprehensive lens to assess interventions and policy strategies.

As the obvious next step from my scientific narrative presented above, I performed NLP (Cukkemane, 2025 IN REVIEW) because I believe that it can serve as an invaluable tool to analyse large volumes of transcribed interview data from stakeholders engaged in the management of CMI. We employed the syuzhet package (Isasi, 2023) in R, a widely used tool for sentiment analysis, by incorporating four primary lexicons: National Research Council Canada (NRC(Mohammad & Turney, 2013)), Bing (Bing, 2012), AFINN (Nielsen, 2011), and Syuzhet (Jockers, 2020). Among these, the NRC lexicon was particularly effective in detecting nuanced emotional states such as trust, fear, anger, anticipation, and joy across stakeholder narratives. Altogether, we could navigate the sentiment trajectory (Fig 2) through the course of discussion revealing several emotional patterns.

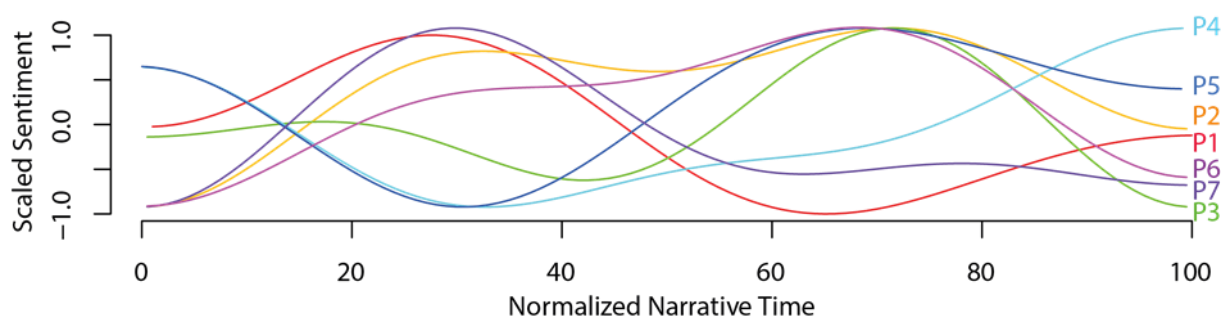


Figure 2: Evolution of sentiments (positive, negative and neutral) the during the course of the interview for the seven participants , which was scored on the basis of the NRC lexicon.

The combination of qualitative analytics and NLP offers a powerful dual approach for understanding the multifaceted challenges in managing CMIs. While qualitative methods provide deeper and detailed thematic exploration, NLP delivers themes from negative perspective slightly short but is scalable, and consistent. Together, these approaches lay the groundwork for a data-driven, patient-centric, and emotionally intelligent framework in psychiatric research, care, and policymaking. From our scientific

expertise, we need to improve and provide better approaches to diagnosis and treatment, and yet at the same time keep other stake-holders regularly informed.

In addition, such studies play a crucial role in reshaping public discourse. Although, I did not touch upon detailed societal requirements, our study also offers the possibility of reducing the social stigmatization surrounding chronic mental illnesses (CMIs). By capturing authentic patient narratives (upon ethics committee approval) and analysing them at scale, we can uncover deeply rooted societal biases and misconceptions. These insights not only foster responsibility among clinicians, policymakers, and the public but also inform targeted anti-stigma interventions and empathy.

References

- Bing, L. (2012). *Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies, Band 16)*. Morgan & Claypool Publishers (Springer Cham).
<https://doi.org/https://doi.org/10.1007/978-3-031-02145-9>
- Chartogne, M., Leclercq, A., Beaune, B., Boyas, S., Forestier, C., Martin, T., Thomas-Ollivier, V., Landry, S., Bourgeois, H., Cojocarasu, O., Pialoux, V., Zanna, O., Messonnier, L. A., Rahmani, A., & Morel, B. (2021). Building a biopsychosocial model of cancer-related fatigue: the BIO CARE Factory cohort study protocol. *BMC Cancer*, 21(1), 1140.
<https://doi.org/10.1186/s12885-021-08831-3>
- Cukkemane, A., & Sanchez, V. G. (2025 IN REVIEW). Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. *Journal of Affective Disorders Reports*.
- Cukkemane, A., Hurlermann, R., & Sanchez, V. G. (2025). Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. *Journal of Affective Disorders Reports*, 20, 100893.
<https://doi.org/https://doi.org/10.1016/j.jadr.2025.100893>
- Elo, S., & Kyngas, H. (2008). The qualitative content analysis process. *J Adv Nurs*, 62(1), 107-115.
<https://doi.org/10.1111/j.1365-2648.2007.04569.x>
- Engel, G. L. (1978). The biopsychosocial model and the education of health professionals. *Ann N Y Acad Sci*, 310, 169-187. <https://doi.org/10.1111/j.1749-6632.1978.tb22070.x>
- Engel, G. L. (1979). The biopsychosocial model and the education of health professionals. *Gen Hosp Psychiatry*, 1(2), 156-165. [https://doi.org/10.1016/0163-8343\(79\)90062-8](https://doi.org/10.1016/0163-8343(79)90062-8)
- Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and evaluating qualitative research. *Aust N Z J Psychiatry*, 36(6), 717-732. <https://doi.org/10.1046/j.1440-1614.2002.01100.x>
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today*, 24(2), 105-112.
<https://doi.org/10.1016/j.nedt.2003.10.001>

- Isasi, J. (2023). Sentiment Analysis with 'syuzhet' using R. *Programming Historian*, 12. <https://doi.org/https://doi.org/10.46430/phen0110>
- Jockers, M. (2020). *Introduction to the Syuzhet Package (CRAN)*. <https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html>.
- Lauriola, M., & Tomai, M. (2019). Biopsychosocial Correlates of Adjustment to Cancer during Chemotherapy: The Key Role of Health-Related Quality of Life. *ScientificWorldJournal*, 2019, 9750940. <https://doi.org/10.1155/2019/9750940>
- Mohammad, S., & Turney, P. D. (2013). Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*, 29(3), 436-465. <http://dblp.uni-trier.de/db/journals/ci/ci29.html#MohammadT13>
- Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*, 93-98. <https://doi.org/https://doi.org/10.48550/arXiv.1103.2903>
- OECD/EU. (2018). *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. OECD Publishing, Paris. https://doi.org/10.1787/health_glance_eur-2018-en
- Tomas, J. M., Oliver, A., Torres, Z., Parker, J., Marques-Sule, E., & Sentandreu-Mano, T. (2023). A Biopsychosocial Model Predicting Myocardial Infarction. *J Clin Med*, 12(17). <https://doi.org/10.3390/jcm12175715>
- Wiegand, H. F., Saam, J., Marschall, U., Chmitorz, A., Kriston, L., Berger, M., Lieb, K., & Holzelt, L. P. (2020). Challenges in the Transition from In-Patient to Out-Patient Treatment in Depression. *Dtsch Arztebl Int*, 117(27-28), 472-479. <https://doi.org/10.3238/arztebl.2020.0472>

4. DISC1 (dys)function and its implications for chronic mental illnesses

Disrupted in Schizophrenia 1 (DISC1) is a multifunctional scaffold protein implicated in various chronic mental illnesses (CMIs), including schizophrenia, bipolar disorder, and autism spectrum disorders. The interactome of DISC1 comprises over 300 proteins (Bradshaw & Korth, 2019; Soares et al., 2012; Yerabham et al., 2013) that are crucial in a variety of physiological processes including cellular replication, neuronal and synaptic development. It has been shown that the protein plays the role of a scaffold modulating the activity of various proteins of clinical and therapeutic relevance. While its role in neurodevelopment and cell division is well-documented, the molecular underpinnings of its dysfunction remain elusive. DISC1 is an 854 amino acid (aa) residues long protein. A high-resolution 3D structure of the protein has remained elusive, which is also the major objective of this proposal. However, bioinformatics analysis (Soares et al., 2012; Yerabham et al., 2013) based on sequence reveals that the major isoform is composed of a disordered N terminal segment (~350 aa), followed by helical regions of coiled-coil nature that are prone to oligomerization and perhaps also central to aggregation in CMI patients.

Recent studies in our group demonstrated that one of the aggregation pathways of the DISC1 involves the C-terminal region the (C-region (residues 691–836)) as a potential link to disease pathology (Cukkemane et al., 2021). This region forms amyloid-like fibrils, and we believe that β -core (residues 717–761) segment that overlaps with binding sites for critical mitotic spindle protein such as nuclear distribution element like protein (NDEL1). Using a combination of structural and biophysical techniques, we demonstrated that the C-terminal region of the DISC1 protein exhibits significant polymorphism, which bears critical implications for its physiological function. Data derived from solid-state NMR (ssNMR) spectroscopy and electron microscopy (EM) reveal that this region not only assembles into symmetric oligomeric species but also forms fibrillar structures that closely resemble amyloids similar to proteinopathies observed in neurodegenerative diseases. Thermodynamic analysis via isothermal titration calorimetry (ITC) confirms that DISC1 aggregation is an exergonic process, characterized by a negative enthalpy change that energetically favors both oligomer (likely tetramer) formation and β -sheet-rich fibrillization.

Through sequence-specific structural investigations, we hypothesized that the core fibrillogenic segment to residues 716–761 (β -core) of the full-length human DISC1 protein. Notably, this segment is missing in the DISC1 ^{Δ 22aa} splice variant (Kamiya et al., 2006), which shows impaired interaction with key dynein motor complex components, namely NDE-like 1 (NDEL1) and platelet activating factor acetylhydrolase 1b regulatory subunit 1 (PAFAH1B1, formerly LIS1), both of which play essential roles in mitotic spindle dynamics. Surface plasmon resonance (SPR) experiments reveal that the oligomeric form of the DISC1 C-region binds to LIS1 and NDEL1 with enhanced affinity and cooperative binding behavior, unlike its monomeric counterpart, which exhibits noncooperative interaction.

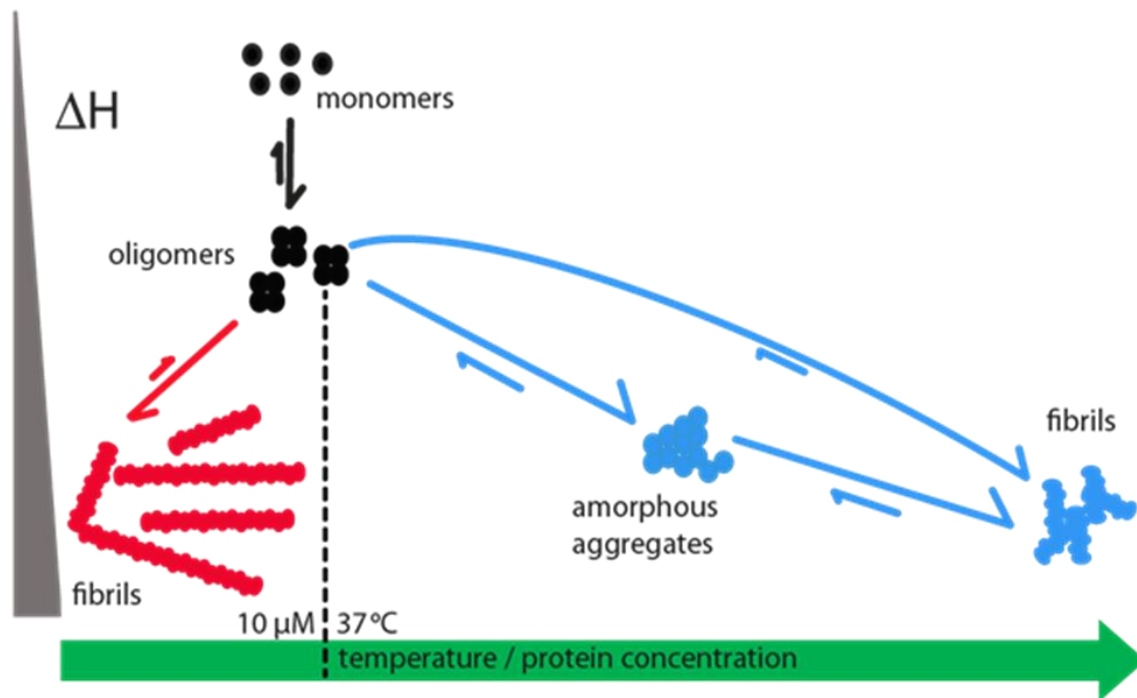


Figure 1: Pictorial depiction of the energy and conformational landscape governing DISC1 C-region self-association. Monomeric and/or oligomeric forms of the DISC1 C-region display a tendency to form extended fibrils at temperatures below 37 °C. However, elevated temperature and protein concentration enhance nucleation rates, resulting in a larger number of shorter fibrils and the concurrent formation of amorphous aggregates

These findings underscore the dual nature of the DISC1 C-terminal region as a functional scaffold in neurodevelopmental signaling complexes and as a driver of pathological aggregation with potential relevance to psychiatric disease (Figure 1).

Structural Plasticity and Aggregation of DISC1 Mutants

Further building on our findings, we were keen to understand whether a thin line separates the delicate balance between structural flexibility and aggregation for the DISC1 C-region protein. While the functional tetrameric form enhances binding to mitotic spindle partners, clinically relevant mutants carrying point mutations, interestingly in our studies bound to NDEL1, PAFAH1B1, and NDE1 with nanomolar affinities (Abhishek Cukkemane et al., 2025). This indicated that the mutations themselves do not abolish the binding interface *per se* but other factors contributed to the pathogenesis. Further probing the mutants using DLS and AFM, we noticed the influence of the mutation in its behaviour towards pathological aggregates. All mutants displayed amorphous aggregates at higher concentrations. The S713E behaved comparably to wild-type, The S704C, however showed increased instability, and the L807-FS was highly unstable at all tested concentrations.

Thermodynamic profiling using ITC reveals that fibrillization was driven by exothermic reactions, with variant-specific enthalpy changes reflecting structural heterogeneity in the aggregates. Water-edited

ssNMR confirmed the presence of tightly bound water, typical of structured fibrils, but with variations suggesting differing internal architectures between the wild-type, S704C, and S713E when compared to L807-FS. Taking all these results together, we were able to draw the aggregation pathway (Figure 2) of the DISC1 C-region protein.

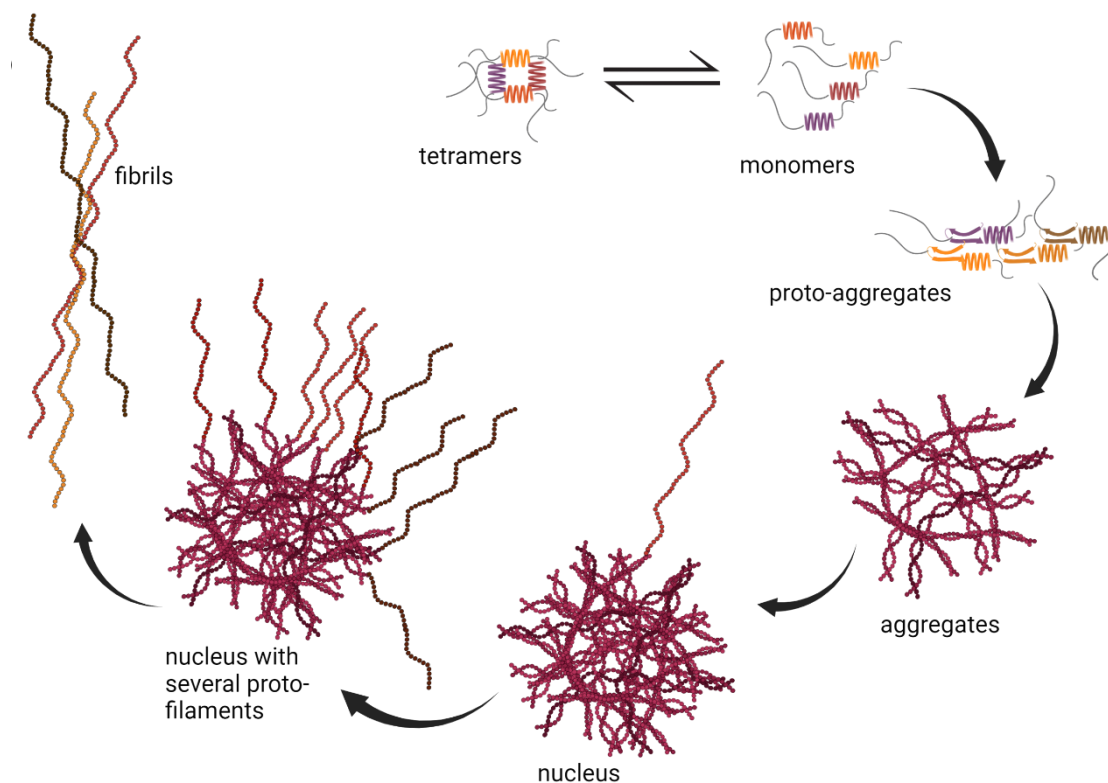


Figure 2: Snapshots describing the aggregation and fibrillar pathway of the DISC1 C-region; the mutant variants are suggested to follow a similar pathway.

As an extension to the aggregation narrative, we also demonstrated the importance of the β -core as a plastic site for function and fibrillization. The β -core also binds to NDEL1, PAFAH1B1, and NDE1 with nanomolar affinities but alone displays inconsistent aggregation behaviour and the thermodynamic profiling using ITC reveals exothermic reactions, with variant-specific enthalpy changes reflecting structural heterogeneity in the aggregates.

Structure Rational Peptide Therapeutic Design

To pharmacologically address aggregation, we employed peptide-based drug development strategy. Using phage display (A. Cukkemane et al., 2025; Cukkemane, 2022), two peptides, Flm01 and Flo01 were selected against monomeric and oligomeric C-region constructs. Both the peptides bound to wild-type and mutant C-region proteins with high affinity in the nM range. Using DLS studies, we demonstrated that peptide treatment reduced the kinetics of aggregation significantly. Flo01 was especially effective in maintaining the DISC1 C-region tetramers at high concentrations of the peptide.

This is crucial because the functional DISC1 C-region tetramer binds to the proteins of mitotic complex with high affinity.

By combining structural biology and pharmacological strategies, we present a cohesive understanding of DISC1 aggregation and its potential therapeutic targeting.

The integrated approach highlighted herein shows the critical role of the DISC1 C-region in both normal neurodevelopmental processes and the pathogenesis of CMIs. Structural and biophysical analyses reveal that the C-region mutants retain their binding capacity to the physiological partners. The potential loss-of-function however arise due to DISC1 aggregation, which maybe the lynchpin of dysfunction in psychiatric disorders. Furthermore, we have been able to convincingly demonstrate that the β -core region contains the essential elements for binding mitotic partners but also serves as the focal point for amyloidogenesis.

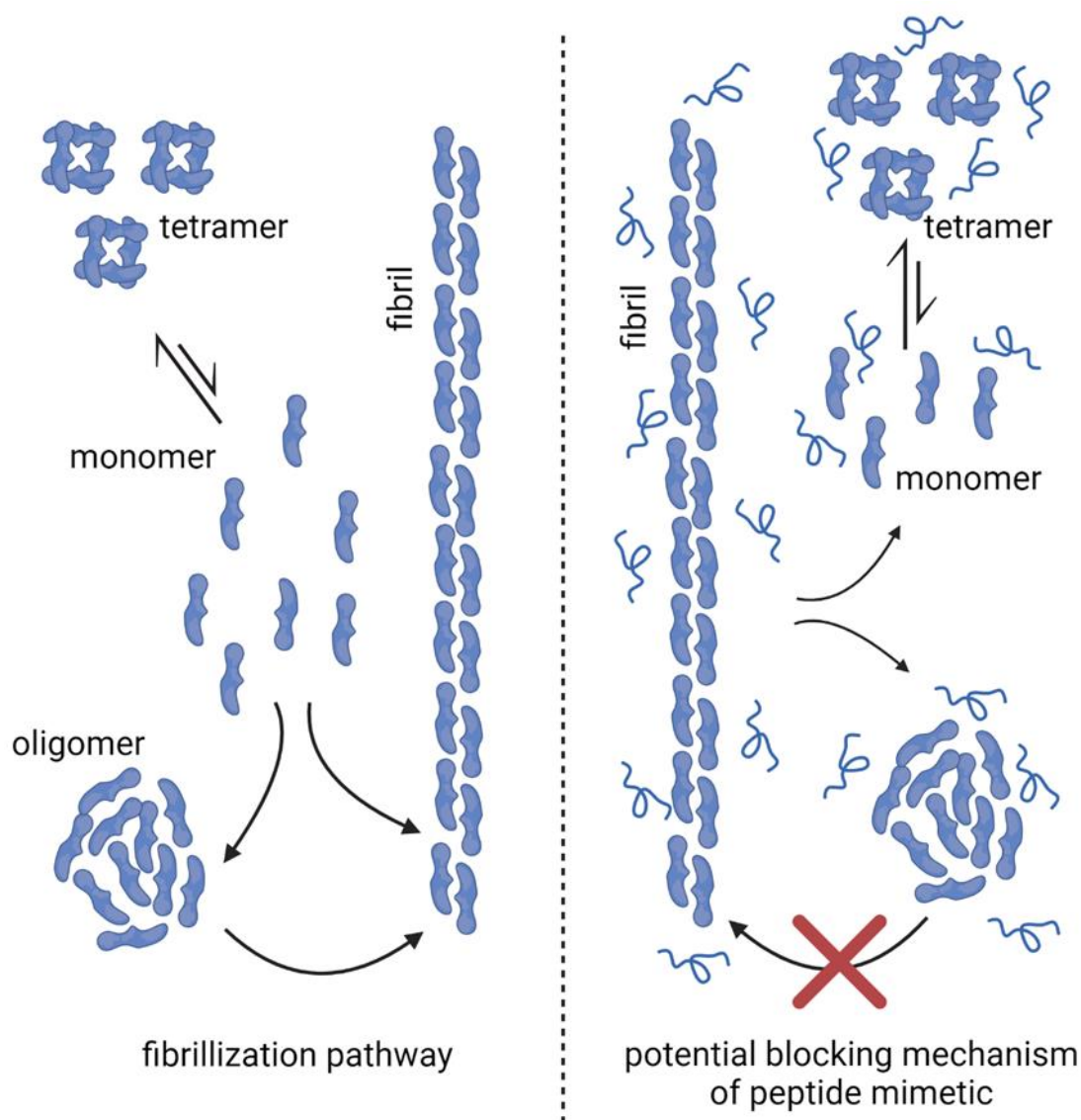


Figure 3: Schematic illustration of the potential activity of the peptide mimetic. In the absence of the peptide (left panel) the monomer is in equilibrium with the tetramer under physiological conditions.

Based on this structural information, using the monomeric and the oligomeric versions of the wild-type C-region as baits, we were able to identify peptide mimetics. The successful validation of two of the peptide mimetics in curtailing DISC1 C-regions aggregation offered a promising start to drug development strategies (Figure 3). These peptides at higher concentrations populated the DISC1 C-region in its functional tetrameric form, thus making the protein available for its physiological roles in mitotic complex formation and neuronal development.

Taken together, the findings present a framework linking DISC1 structure to its (dys)function and drug discovery strategy. This dual approach serves as a template for targeting other regions within the DISC1 protein as well as other intrinsically disordered proteins implicated in neuropsychiatric and neurodegenerative diseases. Thereby, opening new avenues for therapeutic innovations for neglected diseases and disorders, and restoring functional protein interactions in complex cellular environments.

References:

- Bradshaw, N. J., & Korth, C. (2019). Protein misassembly and aggregation as potential convergence points for non-genetic causes of chronic mental illness. *Mol Psychiatry*, 24(7), 936-951. <https://doi.org/10.1038/s41380-018-0133-2>
- Cukkemane, A., Becker, N., Kupreichyk, T., Heise, H., Willbold, D., & Weiergräber, O. H. (2025). Tracing the aggregation pathway of the scaffold protein DISC1: Structural implications for chronic mental illnesses. *Journal of Structural Biology: X*, 11, 100128. <https://doi.org/https://doi.org/10.1016/j.yjsbx.2025.100128>
- Cukkemane, A., Becker, N., Zielinski, M., Friege, B., Lakomek, N. A., Heise, H., Schroder, G. F., Willbold, D., & Weiergraber, O. H. (2021). Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl Psychiatry*, 11(1), 639. <https://doi.org/10.1038/s41398-021-01765-1>
- Cukkemane, A., Dingley, A. J., Mohrluder, J., Santiago-Schubel, B., Weiergraber, O. H., & Willbold, D. (2025). A peptide mimetic therapeutic strategy targeting dysfunction of the scaffold protein DISC-1 in psychiatric disorders. *Eur J Pharm Sci*, 107148. <https://doi.org/10.1016/j.ejps.2025.107148>
- Cukkemane, A., Willbold, D. (2022). *DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDs)* (EPO Patent No.
- Kamiya, A., Tomoda, T., Chang, J., Takaki, M., Zhan, C., Morita, M., Cascio, M. B., Elashvili, S., Koizumi, H., Takanezawa, Y., Dickerson, F., Yolken, R., Arai, H., & Sawa, A. (2006). DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. *Human Molecular Genetics*, 15(22), 3313-3323. <https://doi.org/10.1093/hmg/ddl407>

- Soares, D. C., Bradshaw, N. J., Zou, J., Kennaway, C. K., Hamilton, R. S., Chen, Z. A., Wear, M. A., Blackburn, E. A., Bramham, J., Bottcher, B., Millar, J. K., Barlow, P. N., Walkinshaw, M. D., Rappsilber, J., & Porteous, D. J. (2012). The Mitosis and Neurodevelopment Proteins NDE1 and NDEL1 Form Dimers, Tetramers, and Polymers with a Folded Back Structure in Solution. *Journal of Biological Chemistry*, 287(39), 32381-32393. <https://doi.org/10.1074/jbc.M112.393439>
- Yerabham, A. S., Weiergraber, O. H., Bradshaw, N. J., & Korth, C. (2013). Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol Chem*, 394(11), 1425-1437. <https://doi.org/10.1515/hsz-2013-0178>

ARTICLE OPEN

Conformational heterogeneity coupled with β -fibril formation of a scaffold protein involved in chronic mental illnesses

Abhishek Cukkemane^{1,2} , Nina Becker^{1,2,3} , Mara Zielinski¹ , Benedikt Frieg¹, Nils-Alexander Lakomek^{1,2,3}, Henrike Heise^{1,2,3} , Gunnar F. Schröder^{1,3,4} , Dieter Willbold^{1,2,3} and Oliver H. Weiergräber^{1,3}

© The Author(s) 2021

Chronic mental illnesses (CMIs) pose a significant challenge to global health due to their complex and poorly understood etiologies and hence, absence of causal therapies. Research of the past two decades has revealed dysfunction of the disrupted in schizophrenia 1 (DISC1) protein as a predisposing factor involved in several psychiatric disorders. DISC1 is a multifaceted protein that serves myriads of functions in mammalian cells, for instance, influencing neuronal development and synapse maintenance. It serves as a scaffold hub forming complexes with a variety (~300) of partners that constitute its interactome. Herein, using combinations of structural and biophysical tools, we demonstrate that the C-region of the DISC1 protein is highly polymorphic, with important consequences for its physiological role. Results from solid-state NMR spectroscopy and electron microscopy indicate that the protein not only forms symmetric oligomers but also gives rise to fibrils closely resembling those found in certain established amyloid proteinopathies. Furthermore, its aggregation as studied by isothermal titration calorimetry (ITC) is an exergonic process, involving a negative enthalpy change that drives the formation of oligomeric (presumably tetrameric) species as well as β -fibrils. We have been able to narrow down the β -core region participating in fibrillization to residues 716–761 of full-length human DISC1. This region is absent in the DISC1 ^{Δ 22aa} splice variant, resulting in reduced association with proteins from the dynein motor complex, viz., NDE-like 1 (NDEL1) and lissencephaly 1 (LIS1), which are crucial during mitosis. By employing surface plasmon resonance, we show that the oligomeric DISC1 C-region has an increased affinity and shows cooperativity in binding to LIS1 and NDEL1, in contrast to the noncooperative binding mode exhibited by the monomeric version. Based on the derived structural models, we propose that the association between the binding partners involves two neighboring subunits of DISC1 C-region oligomers. Altogether, our findings highlight the significance of the DISC1 C-region as a crucial factor governing the balance between its physiological role as a multifunctional scaffold protein and aggregation-related aberrations with potential significance for disease.

Translational Psychiatry (2021)11:639; <https://doi.org/10.1038/s41398-021-01765-1>

INTRODUCTION

Chronic mental illnesses (CMIs), e.g., schizophrenia and recurrent affective disorders, remain enigmatic due to their multifactorial etiology that involves an interplay of various factors including biological, environmental, and social conditions. One of the major biological risk factors that was identified about 20 years ago in a Scottish pedigree with severe psychiatric disorders is disrupted in schizophrenia 1 (DISC1) isoform 1 [1, 2]. A growing wealth of information on the physiological role of DISC1 describes its importance in cellular functions such as proliferation, neuronal development, and synaptogenesis [3–5]. Despite the lack of hits for DISC1 in genome-wide association study (GWAS) screens [6], there is evidence of the role of DISC1 mutations that have been shown to be heritable in the pathology of schizophrenia and related CMIs [1, 2, 7]. While the GWAS studies have been designed to identify genetic factors associated with phenotypic traits of a disease, it has underperformed for schizophrenia studies at the population level mainly due to identifying common variants. It does not account for mutations that are rare, such as those

observed in DISC1 [8, 9], and the ones that are associated with altered post-translational modifications, misassembly, and aggregation [10–12]. Like aberrant species found in other proteinopathies among the neurodegenerative and prion amyloid diseases, DISC1 aggregates have been demonstrated to spread between cells in in vitro experiments [13], albeit without causing cell death. The protein plays the pivotal role of a scaffold by regulating the activity of various enzymes and other proteins, many of which are of clinical and therapeutic relevance; in fact, its interactome includes more than 300 different partners [14, 15]. Therefore, high-resolution 3D structures of the DISC1 protein and its complexes with relevant risk factors are pertinent in comprehending its physiological significance.

From a structural perspective (Fig. 1A), evaluation of the amino acid sequence using bioinformatics analysis indicates that the DISC1 N-terminal region of approximately 350 residues is intrinsically disordered, harboring four distinct phosphodiesterase 4 (PDE4) binding sites [16, 17] and two nuclear-localization signals [18]. The C-terminal part is predicted to be largely helical,

¹Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany. ²Institut für Physikalische Biologie, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ³Jülich Centre for Structural Biology (JuStruct), Forschungszentrum Jülich, Jülich, Germany. ⁴Physics Department, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ✉email: a.cukkemane@fz-juelich.de; d.willbold@fz-juelich.de; o.h.weiergraeber@fz-juelich.de

Received: 3 June 2021 Revised: 23 November 2021 Accepted: 7 December 2021

Published online: 17 December 2021

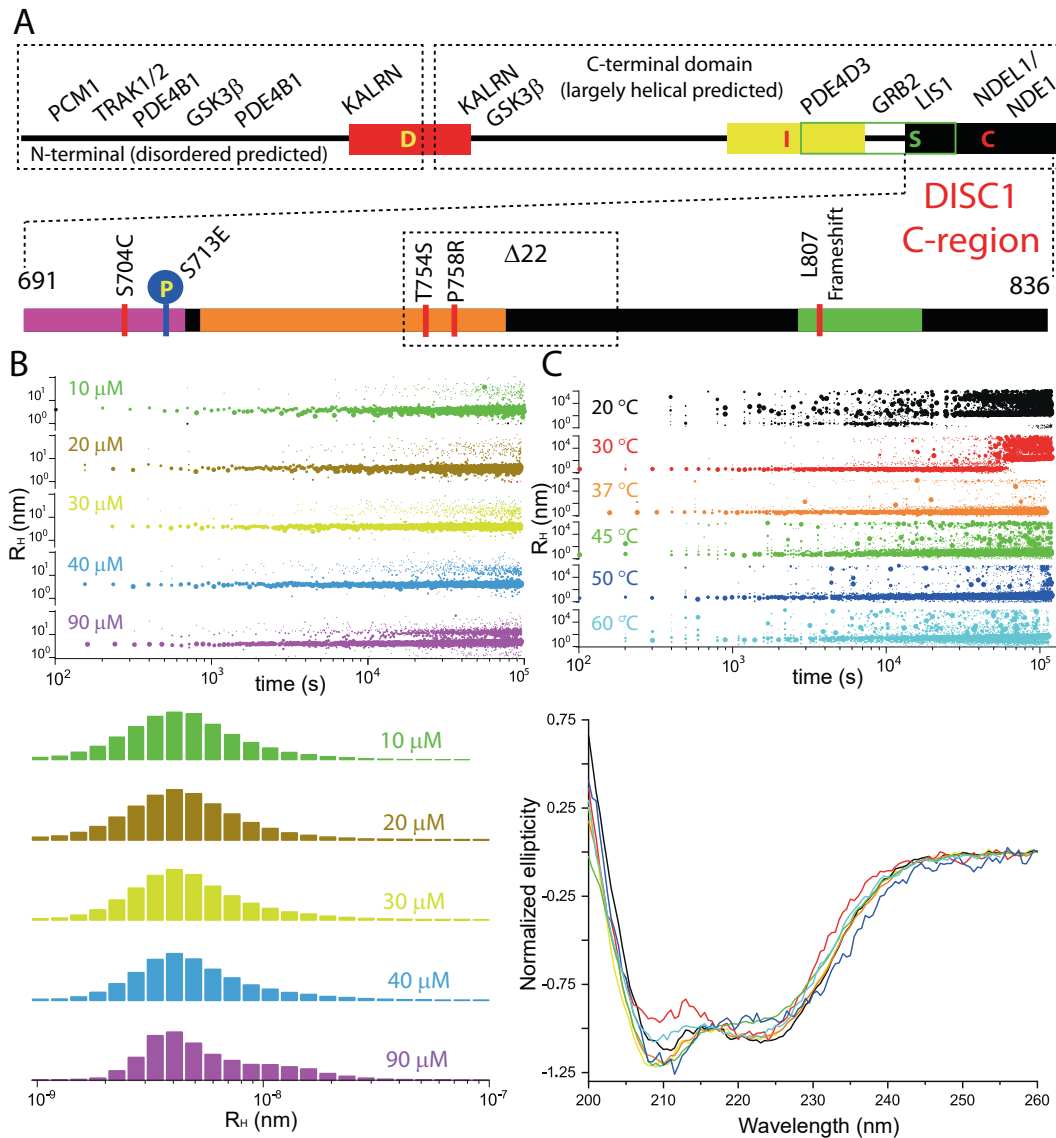


Fig. 1 Size distribution and secondary structure composition of DISC1 C-region oligomers. **A** Pictorial representation of the primary structure of the DISC1; the D, I, S and C regions and the potential binding-partner interaction sites are annotated. For more details, the reader is referred to reviews in references [3–5, 8, 9, and 14]. The lower panel shows the C-region with the V_HH B5 nanobody-binding site (magenta), the (pseudo)repeat sequence (orange), and the C-terminal NDEL1-binding site (green); the region missing in the Δ22 variant is indicated by a dotted box. The same color-coding scheme is also used in the structural model in Fig. 4D. Post-translational modification and mutation sites have been annotated for the C-region. **B** DLS measurements on the His₆-tagged version at 37 °C, using various concentrations: 10 μM (green, apparent hydrodynamic radius (R_H): 4.1 nm ± 0.7 nm, calculated molecular mass (M): 81.5 kDa), 20 μM (brown, R_H : 4 ± 0.6 nm, M: 81.3 kDa), 30 μM (yellow, R_H : 4.1 ± 0.5 nm, M: 84 kDa), 40 μM (sky blue, R_H : 4.4 ± 0.5 nm, M: 87 kDa; R_H : 17.2 ± 3.6 nm, M: 2.3 MDa; R_H : 18.9 ± 4.4 nm, M: 2.8 MDa), and 90 μM (magenta, R_H : 4.1 ± 0.6 nm, M: 85.7 kDa; R_H : 13 ± 2.5 nm, M: 1.2 MDa). The estimated molecular mass of the smallest species is suggestive of a tetramer. While the top panel illustrates the time-dependent evolution of the radius distribution, the bottom panel displays histograms derived from the entire data set. **C** In order to further characterize the oligomerization process of the C-region and the accompanying structural changes, we performed DLS measurements (top panel) along with CD spectroscopy (bottom panel) at various temperatures, 20 °C (black), 30 °C (red), 37 °C (orange), 45 °C (green), 50 °C (blue), and 60 °C (cyan). One additional spectrum (yellow) represents the CD profile of the freshly prepared sample.

including several coiled-coil regions [8, 19] characterized by the presence of heptad repeats (commonly denoted *abcdefg*, with positions *a* and *d* carrying hydrophobic residues and others being either polar or charged). Additionally, this segment has been proposed to contain antiparallel helical hairpin structures known as UVR domains in regions 343–394, 574–625, and potentially also 778–828 [19]. The latter has been confirmed by solution NMR spectroscopy (see below). A systematic search for stably folded and soluble regions using the high-throughput Expression of Soluble Proteins by Random Incremental Truncation (ESPRIT)

technique [20] revealed four putative domains that were mnemonically named D, I, S and C and comprise residues 257–383, 539–655, 635–738, and 691–836 [21], respectively. Among these presumably structured regions of DISC1, the C-region has been characterized most thoroughly, including its mode of binding to a camelid nanobody [22]. Solution NMR structures of a truncated version of the murine DISC1 C-region comprising the distal 73 residues have been determined in the presence of peptide fragments of ATF4 [23] and NDE1 [24], respectively. The interaction between the C-region and its binding

partner is similar in both systems and involves formation of a three-helix coiled-coil structure, with the helical fragment of ATF4/NDE1 clamped by hydrophobic interactions of the heptad repeats with the two antiparallel helices contributed by DISC1.

From a pathophysiological perspective, the C-region of DISC1 deserves particular attention since it is affected by alterations implicated in mental illnesses. Specifically, this segment is deleted in the Scottish variant [2] lacking residues 599–854 as a consequence of a balanced translocation, and it harbors the sites of a mental illness-associated polymorphism (S704C [25]) and of a frameshift mutation (affecting residues downstream of L807) that results in aggregated complexes [7]. A phosphorylation site observed at S713 [11] serves the role of a switch that promotes the transit of neuronal progenitor cells from proliferation to migration during corticogenesis. Moreover, other disease-associated proteins such as NDE-like 1 (NDEL1) and lissencephaly 1 (LIS1), whose impaired function results in microcephaly and lissencephaly, respectively, interact with these regions [26] alongside the motor protein dynein, all of these being crucial for mitosis. Despite increasing insight into the cellular processes that DISC1 is involved in, its precise molecular functions and their regulation remain poorly understood. Contributing to this is the propensity of the DISC1 protein to aggregate [10], which is the major bottleneck impeding biophysical characterization of the native form and, in particular, determination of its 3D structure. This has led to several unanswered questions: Is it possible to structurally characterize DISC1 aggregates? Why does the protein aggregate and how can this process be prevented? Is the molecular nature of aggregation unique or does it share common features with other proteinopathies? How does aggregation of DISC1 affect its interaction with binding partners and the respective cellular functions?

Herein, we address these queries. To comprehend the impact of DISC1 on the pathophysiology of schizophrenia and related CMLs, we focus our efforts on the C-region of the protein. Employing a combination of biophysical techniques probing structure and dynamics, we demonstrate that the C-region gives rise to cross β -fibril-like aggregates and report kinetic and thermodynamic parameters of their formation. Moreover, we show that the higher oligomeric structures differ from the monomeric species regarding protein–protein interaction properties, suggesting severe consequences for cellular signaling. Altogether, our approach provides us with a biophysical basis to understand the relevance of the C-region of the DISC1 protein with respect to both its physiological role as a scaffold protein and its pathophysiological significance in fibril assembly, in relation to other amyloidogenic diseases.

METHODOLOGY

Protein expression and purification

The human DISC1 C-region (residues 691–836) was expressed alternatively as N-terminal His₆- or MBP-His₆-fusion construct. The His₆-tag construct was expressed as described previously [22] using the vector pESPRIT002 [20] for transforming *E. coli* BL21 (DE3) pLysE T1R cells, while the MBP-His₆ fusion construct was cloned in pET-15b and expressed in *E. coli* (DE3) C43 cells. The culture was grown either in Luria-Bertani (LB) broth or M9 minimal medium. Protein expression was induced at an OD₆₀₀ of 0.6 by the addition of 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) for 16 h at 18 °C. Harvested cells were lysed in Tris-buffered saline (TBS) buffer containing 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, and Complete EDTA-free protease-inhibitor cocktail (Roche) using an ice-chilled microfluidizer M100P (Microfluidics MPT) at 15,000 PSI. The insoluble fraction was spun down by centrifugation at 50,000 $\times g$ for 25 min at 4 °C. The soluble fraction was purified on NTA resin (Qiagen) loaded with Co²⁺ and was eluted using TBS containing 500 mM imidazole. In the subsequent step, the His₆-C-region protein was further purified on a HiLoad 16/60 Superdex

200 size-exclusion chromatography column (GE Healthcare). The MBP fusion protein was purified via an amylose column (New England Biolabs) and eluted using TBS containing 25 mM maltose.

Proteins V_HH B5 and NDEL1 were prepared in a similar fashion as described previously [22, 27], which was also adopted for LIS1 expression and purification. Briefly, V_HH B5, LIS1, and NDEL1 proteins were expressed in *E. coli* cells of the strain types Rosetta (DE3) pLysS, C43 (DE3), and BL21 (DE3) pLysE-T1R, respectively. The cells were cultivated at 37 °C until an OD₆₀₀ of 0.6 was reached, following which expression was induced using 1 mM IPTG and maintained for 16 h at 18 °C. The harvested cells were lysed as described above, and the proteins were purified on Ni²⁺-NTA resin (Qiagen), eluted using TBS containing 500 mM imidazole, and purified further on a HiLoad 16/60 Superdex 200 SEC column (GE Healthcare).

Dynamic light scattering (DLS)

Measurements were performed using a SpectroSize 300 (XtalConcepts) instrument and a sample volume of 500 μ l. The temperature dependence (20–60 °C) was investigated at a protein concentration of 10 μ M, while concentration-dependent aggregation (10–90 μ M) was studied at a temperature of 37 °C. Prior to measurements, all samples were centrifuged at 21,000 $\times g$ for 30 min at 4 °C. Diffusion coefficients were obtained from analysis of the decay of the scattered-intensity autocorrelation function and were used to determine apparent hydrodynamic radii via the Stokes–Einstein equation.

In order to assess the thermodynamics (also see below) of fibrillization, we employed a simple 1D “crystallization” model of amyloid formation [28–30]. The model assumes that upon completion of the process, the solution has converted into a suspension of highly ordered fibrils that are at equilibrium with residual monomers/oligomers. The basic reaction describing the association/dissociation of a single unit to/from an existing fibril can be represented as



with equilibrium constant

$$K_F = \frac{[F_i]}{[C][F_{i-1}]} = \frac{k_{\text{ass}}}{k_{\text{diss}}} \quad (2)$$

where C represents a single element (monomer or small oligomer) of the DISC1 C-region, F_{i-1} and F_i stand for fibrils (more precisely, docking sites on fibrils) composed of $i-1$ and i units, respectively, and k_{ass} and k_{diss} are the corresponding rate constants. In the later stages of the reaction, the number of growing fibrils will be largely constant because the depletion of soluble material will limit further nucleation; given that fibrils are unbranched in our case (see below), the number of association/dissociation sites, which are presumably located at the fibril ends, will likewise be constant. In this situation, the concentration of docking sites effectively becomes independent of fibril length i , and eq. [2] simplifies to

$$K_F = \frac{1}{[C]} \quad (3)$$

Determination of $[C]$, the equilibrium concentration of C-region monomers/oligomers, which is popularly referred to as the “critical concentration”, is not straightforward owing to the extensive heterogeneity of the system. Considering that radii corresponding to these building blocks were fitted to the DLS data of fibril suspensions only intermittently, we reasoned that their concentration should be on the order of the detection limit of the instrument, which was determined experimentally. From the resulting equilibrium constant K_F , we can compute the apparent free-energy change of fibrillization $\Delta G_{\text{app}} = -RT \ln K_F = RT \ln [C]$, where R and T are the gas constant and temperature, respectively.

Isothermal titration calorimetry (ITC)

Measurements were performed using an iTC-200 (MicroCal) calorimeter with a 200 μ l sample cell. All protein samples were degassed prior to measurements. Changes in the heat flow were monitored in real time with the reaction cell stirred at 300 rpm and reference power of the cell set to 5 μ cal·s⁻¹.

Thermodynamic parameters were computed on the basis of the Goto-Kardos scheme [29, 30]. Assuming that the observed heat exchange represents the enthalpy change of association of the DISC1 C-region, ΔH at 37 °C was calculated by integrating the respective peak area, followed by normalization against protein amount. The entropy change ΔS was calculated using the values of ΔH and ΔG_{app} using the Gibbs-Helmholtz equation ($\Delta G_{app} = \Delta H - T\Delta S$, assuming ΔH is independent of temperature to a first approximation). Likewise, the heat-capacity change at constant pressure (ΔC_p) was obtained by linear extrapolation of the values from $(\delta\Delta H/\delta T)_p$, based on experiments that were performed for the spontaneous fibrillization of 10 μ M protein at temperatures ranging from 20 to 60 °C. The reference cell contained filtered and degassed water.

Thioflavin-T (ThT) fluorescence assay

The assay was performed on black nonbinding 96-well plates (Sigma-Aldrich) using a total volume of 100 μ l containing 10 μ M ThT per well, and protein concentrations ranging from 5 to 90 μ M. Different temperatures (25 °C, 37 °C, and 45 °C) were evaluated with a protein concentration of 10 μ M. Each reaction was background-corrected by subtraction of the buffer control. Fluorescence was monitored at regular intervals of 3–5 min using a Fluostar microplate reader (BMG Labtech, Offenburg, Germany) with 440 nm excitation and 492 nm emission filters, respectively, in bottom-read mode.

Circular dichroism (CD) spectroscopy

Far-UV CD spectra were recorded on a Jasco J-1100 spectropolarimeter at 20 °C (0.1 cm path-length cuvette) in TBS at a ramp rate of 50 nm min⁻¹. Prior to the measurements, protein samples were centrifuged at 20,000 $\times g$ for 20 min at 4 °C to remove large aggregates and fibrils. Spectra were individually normalized to the ellipticity at wave length of 217 nm.

Surface plasmon resonance (SPR) spectroscopy

Binding of LIS1 and NDEL1 to the C-region constructs was investigated using a Biacore T200 instrument (GE Healthcare). Prior to measurements, all protein samples were dialyzed against phosphate-buffered saline (PBS) containing 20 mM Na-phosphate, pH 7.4, and 150 mM NaCl. For coupling of the ligand, the flow cells on a CM5 sensor chip (GE Healthcare) were activated with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/N-hydroxysuccinimide (NHS) (200 mM/50 mM); the DISC1 C-region fragment was diluted to 30 μ g/mL in 10 mM sodium acetate, pH 4.0, immobilized to a final level of 250 RU, and the flow cell was deactivated with 1 M ethanolamine-HCl. A reference flow cell was treated in a similar fashion. Titrations of the above-mentioned interaction partners were performed at 25 °C with PBS, pH 7.4, as running buffer, using protein concentrations ranging from 0.75 to 10 μ M without intermittent regeneration. The sensorgrams were double-referenced using the reference flow cell and a buffer cycle, while evaluation was performed by plotting the respective equilibrium-response level against the applied analyte concentration. The curves were individually fitted using Hill's equation (OriginPro 2020).

NMR experiments

Solution NMR experiments were performed at a static magnetic field strength of 17.6 T, corresponding to a ¹H Larmor frequency of 750 MHz, using a Bruker Avance III spectrometer equipped with a 5 mm cryogenic triple resonance probe. Measurements were conducted at 298 K. The protein sample was provided in PBS buffer (20 mM Na-phosphate, pH 7.4, 150 mM NaCl) containing

10% (v/v) D₂O for field-frequency locking. 2D ¹H–¹⁵N HSQC spectra were recorded to obtain a fingerprint of the protein and assess the highly dynamic protein regions visible to solution NMR. Decoupling of ¹⁵N during proton acquisition was performed by application of the GARP scheme. Solvent suppression was achieved using the conventional WATERGATE scheme. ¹H chemical shifts (CS) were referenced directly to sodium trimethylsilylpropanesulfonate (DSS) at 0 ppm and ¹⁵N shifts were referenced indirectly to DSS, using the absolute-frequency ratios. The HSQC spectra were processed using a squared sine bell function of 2. Additionally, 2D ¹H–¹⁵N HSQC was recorded to follow the aggregation of the DISC1 protein in solution (three different concentrations) at various time points. Furthermore, to analyze the concentration dependence, HSQC spectra of proteins were recorded at three different concentrations of 10, 60, and 400 μ M, respectively, in a time-dependent manner.

Solid-state NMR (ssNMR) experiments were performed on a uniformly ¹⁵N/¹³C-labeled sample of the DISC1 C-region, dissolved in TBS buffer containing 1 mM Na₂-EDTA and 0.01% (w/v) NaN₃, which was concentrated in a centrifugal device with 3 kDa cutoff membrane (Amicon, Millipore). The concentrated sample was bath-sonicated for 5 min and incubated at 20 °C for 2 weeks. At this point, it contained noticeable particulate matter, had a consistency similar to hydrogel, and was packed into 1.9- and 3.2-mm zirconia magic angle spinning (MAS) rotors (Bruker Biospin). Measurements were conducted at static magnetic field strengths of 16.4 and 18.8 T (corresponding to ¹H Larmor frequencies of 600 and 800 MHz, respectively) and an experimental temperature of 256 K, using 1.9- and 3.2-mm triple-resonance (¹H, ¹³C, and ¹⁵N) Bruker MAS ssNMR probes (Bruker Biospin). Cross-polarization (CP)-based experiments, sensitive to more rigid regions of the protein, were performed at 600 MHz using the 3.2 mm probe. 2D ¹³C–¹³C proton-driven spin-diffusion (PDSD) experiments were recorded at 12.5 kHz MAS frequency with 20 ms mixing time and under weak coupling conditions [31], with mixing time of 20 and 200 ms at MAS frequencies of 8.5 kHz, respectively. ¹H decoupling was applied during evolution and detection periods, using the SPINAL64 scheme at a radio-frequency between 66 and 83 kHz.

To obtain information on the highly dynamic regions of the protein, 2D ¹H–¹⁵N as well as ¹H–¹³C correlation experiments, employing refocused insensitive nuclei enhanced by polarization transfer (refocused INEPT)—with detection on the ¹⁵N and ¹³C nuclei, respectively, were recorded. This was complemented by the proton-detected version of the 2D ¹H–¹⁵N as well as ¹H–¹³C correlation experiments [32], employing refocused INEPT transfers as well. Experiments were conducted at 800 MHz and a MAS frequency of 30 kHz (using the 1.9 mm probe). During acquisition of the ¹H dimension, a 10-kHz weak-power decoupling was applied on the ¹⁵N and ¹³C nuclei using the WALTZ64 scheme. During evolution of ¹⁵N (or ¹³C) nuclei, a 10 kHz weak power proton decoupling was applied using the frequency-swept TPPM scheme. Solvent suppression was achieved by the MISSISSIPPI scheme.

Further, ¹³C–¹³C total through-bond correlation spectroscopy (TOBSY [33]) with 6 ms mixing at 16 kHz MAS was recorded at 800 MHz. Low-power ¹H decoupling using the sequence TPPM at 10 kHz was applied during the detection period. ¹³C and ¹⁵N CS were referenced using the most downfield signal of Ca and N in the tripeptide MLF [34], which is 55 and 100 ppm, respectively.

Water-edited 1D build-up experiments were recorded at 600 MHz and 12.5 kHz MAS, the temperature was maintained at 256 K. A ¹H T₂ filter of 2.5 ms was used to suppress signals from the rigid regions of the protein; for the ¹H–¹H spin diffusion, mixing times (t_m) of 2–500 ms were used to permit spin diffusion from water to the protein. The spectral region of 50–75 ppm was integrated and normalized to the maximum signal intensity (at 100 ms mixing time). A linear fit to the initial build-up rate (t_m^0) was used to determine the water accessibility of the sample. The slope describes the time required to reach 100% magnetization transfer

in the absence of any saturation effects [35] and is inversely proportional to the volume-to-surface area ratio, which is described by $\frac{V}{S} = \sqrt{\frac{D_{\text{eff}} t_m}{\pi}}$ where V represents the volume of the protein, S is the surface area of the protein accessible to water, and D_{eff} is an effective magnetization diffusion coefficient corresponding to 0.2 nm²/ms [36, 37]. Assuming that the fibril represents an elongated cylinder, where fibrillary length is significantly larger than the diameter (d) of the fibril, the ratio of V/S equals $d/4$. Such an approach provides a semiquantitative estimate to the molecular dimensions of water accessible areas of membrane proteins [36] and amyloid fibrils [37] in comparison with their low-resolution structural models. To determine amino acid specific hydration, we employed water-edited 2D ¹³C–¹³C experiments with mixing times of 20 ms to encode intraresidue correlations with the conditions mentioned above. Intensities for regions of interest corresponding to CS of residues in different secondary structural elements (SSE) were integrated. The resulting signals were normalized to the number of scans and signal intensity of the particular amino acid in the respective SSE for a comparison between different ¹H–¹H mixing time.

All data were processed by Topspin 4.0.6 (Bruker Inc.), CS was predicted using the FANDAS [38] software and analyzed using NMRFAM-SPARKY [39].

Transmission electron microscopy (TEM) measurements

The sample solution (4 µl) containing 10 µM protein was applied onto a glow-discharged carbon-coated copper grid (S160-4, Plano). After 2 min, the solution on the grid was blotted off by filter paper. Then the grid was washed with 4 µl of 1% (w/v) uranyl acetate (UrAc) and blotted off immediately, and another 4 µl of 1% (w/v) UrAc was applied onto the grid for 1 min. Then the solution was blotted off and the grid air-dried. TEM images were obtained using a TFS Talos 120 C (Thermo Scientific) with a voltage of 120 kV. Processing of the negatively stained images was performed using RELION 3.1.0 [40] and the contrast transfer function (CTF) was fitted using CTFFIND4 [41]. For the tetramer, 40 particles were selected from ten images (pixel size of 2.5 Å) and extracted with a box size of 80 pixels (200 Å). For the fibrils, 14 fibrils were selected from three images (pixel size of 1.55 Å) and extracted with a box size of 130 pix (201.5 Å), resulting in 393 fibril segments. Subsequently, 2D classification of the tetramers and the fibrils was performed with a mask diameter of 180 Å.

Structural models

The template for the protomer of the DISC1 C-region was obtained from a model published previously, which is consistent with SAXS data [22]. The proposed β-structure was introduced into the model using MODELLER 9.23 [42] after adding restraints defining antiparallel β-strands in segments 716–737 and 739–761, respectively, via the *model-addrsr.py* script. The newly derived structure was subjected to in silico docking simulations using CLUSPRO, targeting tetramers and oligomers in the multimer docking mode [43]. The resulting models were examined for consistency with the results obtained from ssNMR, EM, and SPR measurements and the best candidate selected for further interpretation.

For comparison of the EM-based 3D reconstructions to the structural models, a 3D-density map was simulated for each model using the corresponding PDB file with EMAN2 [44]. Correlations between the reconstructed 3D-density maps and the simulated-model maps were calculated using UCSF Chimera [45].

RESULTS

The DISC1 C-region self-associates into oligomers and aggregates

A DISC1 construct comprising the C-terminal region 691–836 and carrying an N-terminal hexa-histidine tag was studied. While

investigating freshly prepared protein using spectroscopic methods, we noticed formation of large particles during the course of the measurements. As a first approach to characterize this particle formation, we performed DLS experiments at different temperatures and protein concentrations. We observed that the protein, which initially (i.e., freshly prepared off-column) displays an apparent hydrodynamic radius (R_H) in the range of 2.12–3.11 nm (corresponding to a molecular mass (M) of approximately 14–41 kDa for a globular protein), tends to associate into larger particles in the range of 100 nm within a few hours at 20 °C (Fig. S1). Given a theoretical mass of the His₆-tagged protein of 19.4 kDa, the major species would seem to represent a monomer or a dimer (or both), but this assignment may be confounded by the elongated shape of this fragment [22]. Intriguingly, DLS measurements at 37 °C, using protein concentrations ranging from 10 to 90 µM, revealed a predominant fraction with an apparent R_H of 4–4.5 nm (81.3–99.8 kDa), which may correspond to a tetramer, and another stable population at approximately 13–19 nm (1.2–2.8 MDa) that became more apparent with increasing concentration (Fig. 1B). A systematic investigation of temperature dependence of the aggregation profile of the DISC1 C-region (Fig. 1C) showed a smooth transition of the presumed tetramer into several relatively well-defined species with radii above 100 nm, especially at temperatures below 37 °C. At temperatures higher than 37 °C, we do not observe any consistent aggregate sizes but radii scattered across the nm–µm range. The profile at 37 °C, however, was remarkably stable over the measurement duration with tiny-to-negligible scattering fluctuations. In combination with DLS recordings, we performed CD measurements to investigate potential structural transformations at various temperatures (Fig. 1C). Notably, at temperatures lower than 37 °C, we observe a robust change in spectral shape affecting the ratio of the minima at 208 and 222 nm. This transition may reflect loss in helicity and increase in the β-content of the system. In order to ascertain this assumption, we performed ssNMR spectroscopy experiments (*vide infra*) to confirm on the presence of β-strands. At higher temperatures though, the pattern is inconsistent at the characteristic minima of 208 and 222 nm and difficult to rationalize. Perhaps, the spectra represent a mixture of populations with varying secondary-structure element (SSE) composition. Interestingly, alike DLS measurements, the CD spectra at 37 °C agreed well with the freshly prepared sample. In order to ascertain that the tetrameric state represents an intrinsic property of the DISC1 C-region and not an artifact related to the His₆-tag, we generated an alternative construct containing MBP as fusion partner, which should also render the protein more soluble and may thus (at least to some extent) counteract self-association promoted by nonphysiological protein concentrations. Similar to our observations with the His₆ construct, the MBP-tagged version of the DISC1 C-region predominantly existed as a tetramer in solution at different temperatures (Fig. S1). Based on these observations, we hypothesized that the tetramer constitutes a physiologically significant oligomeric state, whereas the larger aggregates represent aberrant forms that may relate to pathology.

To further dissect the relevance of the aggregation phenomenon that we observe at temperatures lower than 37 °C, we proceeded to identify the critical concentration of the DISC1 C-region assuming the 1D crystallization model of amyloid formation [28–30]. The inkling here is that the concentration of fibrils remains stable as the oligomers and monomers are constantly consumed during elongation. As a result, the equilibrium constant (K_F) for fibril formation can be approximated by the inverse of the concentration of protein in equilibrium with the fibril suspension, i.e., the critical concentration (also see “Methodology” section pertaining to DLS measurements), which was observed as 750 nM (Fig. S1). Using this value for the critical concentration, we estimated the ΔG_{app} of fibril formation as $-34.36 \text{ kJ mol}^{-1}$. We proceeded to perform a thorough

biophysical characterization of the different oligomeric states observable via DLS, in order to obtain insight into their architecture and functional properties.

Architecture of the C-region of DISC1

As mentioned above, bioinformatics [8, 19] analysis highlights the predominance of coiled-coil arrangement in the DISC1 C-region, which is experimentally supported by structural information [23, 24] obtained on a truncated version of the murine homolog (73 residues ranging from 764–836); however, the entire C-region comprising 160 amino acid residues has been intractable for structural characterization. In our first attempts using 2D NMR spectroscopy (^1H – ^{15}N HSQC), we noticed that well-defined resonances were grossly under-represented. Signal dispersion is poor and the characteristics of folded protein are not visible. Interestingly, for the few resolved peaks (Fig. 2A), the observed chemical shifts (CS) agree well with the solution NMR data published previously [24] on the truncated murine version mentioned above, and largely localize to the segment spanning residues 790–830 (Fig. S2). More importantly, the broadened signals may indicate oligomerization or the presence of disordered regions or a combination thereof. Regardless of the protein concentration (Fig. S2), we noticed broadened resonances

in all tested conditions and for the few resolved peaks, we observed CS perturbation in a time-dependent manner. Altogether, these findings are suggesting a tendency of the protein to aggrandise, consistent with our observations from the DLS measurements mentioned above. Hence, we believe that the solution NMR data reflect the presence of tetramer that is steadily fibrillizing.

At the same time, we performed ssNMR spectroscopy measurements to separately probe dipolar transfer and scalar couplings (Fig. 2B) on the gel-like sample that is predominantly fibrillar in equilibrium with the tetramer. We observed an overwhelming signal in the INEPT transfer in comparison with the CP-based spectrum, indicating that the C-region comprises elements that are highly mobile on the ssNMR timescale. A close inspection of dipolar-based 2D ^{13}C – ^{13}C transfer data (Fig. 2B), which encode intrasite correlations from rigid parts of the protein, highlights select residues, notably Ala, Gly, Leu, Pro, Ser, Thr, and Val, with CS reminiscent of residues in β -strands. On the other hand, 2D spectra using ^{13}C – ^{13}C -INEPT-TOBSY (Fig. 2B) and ^1H – ^{13}C -INEPT (Fig. S3) correlations exhibit signals exclusively due to amino acids in helices and coils, and are completely devoid of resonances arising from residues in β -strands. To complement the INEPT-transfer experiments, we performed ^1H detected ssNMR

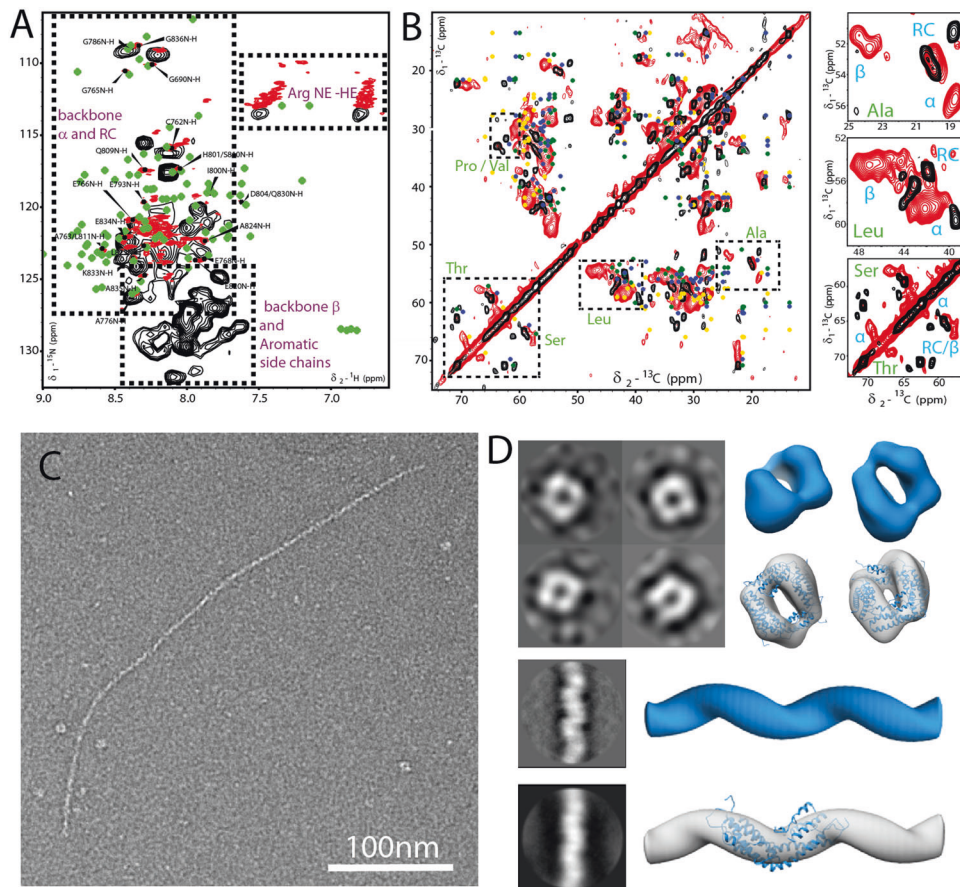


Fig. 2 Structural characterization of the DISC1 C-region using a combination of NMR spectroscopy and EM. In all spectra displayed in the figures, the protein was uniformly ^{13}C - and ^{15}N -labeled. **A** 2D (^1H – ^{15}N)-HSQC spectra of the C-region in solution (red) and solid state (black) at a MAS rate of 30 kHz. Green diamonds represent solution NMR CS of the truncated version of the DISC1 C-region (PDB ID 5YIH) and black annotations are likely assignments of C-region resonances based on proximity to the solution NMR signals. **B** 2D PDSD (red) and INEPT-TOBSY (black) ^{13}C – ^{13}C spectra, diamond-shaped marks in yellow, blue, and green represent average CS (ref. [47]) of residues in helices, coiled regions (RC), and strands, respectively. The box depicts the region where the presence of β -strand-specific CS is expected in dipolar coupling-based experiments for residues such as Ala, Leu, and Ser. **C** Negatively stained EM image of the DISC1 C-region showing the presence of fibrillar protein along with oligomers. **D** 2D class averages of the tetramer and the fibril. For better visualization of the overall fibrillar structure, one class (top) was calculated by ignoring the CTF until the first peak. The 3D reconstructions of the tetramer and the fibril are displayed as surface representations alongside the best-fitting structural model docked within the density map (also see Figs. 4 and S8).

experiments [32, 46] alike solution NMR $^1\text{H}^{15}\text{N}$ - and $^1\text{H}^{13}\text{C}$ -HSQC but at medium MAS rates of 30 kHz. We observed a high degree of overlap (Fig. 2A) between the solution and solid-state $^1\text{H}^{15}\text{N}$ correlations, strongly supporting the notion that the flexible component of the C-region is located in the region comprising residues 764–836. The $^1\text{H}^{15}\text{N}$ correlations at 30-kHz MAS also displayed back-folded signals due to lysine side chains. These sets of experiments are well poised to probe various ranges of protein dynamics that are relevant for protein function and structure (*vide infra* for water-edited spectra). Also, similar to the ^{13}C - ^{13}C correlation spectrum, the ssNMR-HSQC experiments (Fig. S3) are dominated by signals arising from residues in helices and coils, while a CP-based $^1\text{H}^{15}\text{N}$ transfer (Fig. S3) at 30-kHz MAS, albeit inhomogeneously broadened, also shows the presence of CS that is predicted for residues in β -strands [47].

To investigate the molecular architecture of the aggregates, we complemented NMR experiments with EM studies. The monomeric DISC1 C-region has a calculated molecular mass of 19.4 kDa, which roughly corresponds to a hydrodynamic radius of 2.28 nm. In negative-contrast electron micrographs, however, we observed two major populations representing higher oligomers and fibrils, respectively (Fig. 2C), and also noted the presence of amorphous aggregates (Fig. S4). Most of the oligomeric particles in the preparations displayed apparent cyclic symmetry, the predominant species being C4 with an average diameter of 10–12 nm, while minor populations suggesting C5 and C6 were also observed. The 3D reconstruction of the tetramer (Fig. 2D, upper left) is suggestive of few contacts between the subunits and a pore in the center. Fibrils on the other hand appeared as either short segments measuring 100–200 nm in length or longer filaments that extended up to tens of micrometers. The density map of the fibril (Fig. 2D, lower left) is indicative of a dimeric repeat across the length of the structure. The fibrils generally displayed a helical twist with a pitch and average diameter of approximately 9–12 and 5–7 nm, respectively.

Thermodynamics of fibrillization

In ITC measurements (Fig. 3A, B), at concentrations of 10 μM and at 37 $^{\circ}\text{C}$, we observed a weak-intensity exothermic burst, and ThT-fluorescence studies revealed a steady rise in the signal with a plateau that was reached after 8 h (inset of Fig. 3C, black trace). Under this condition, we had previously detected few long fibrils as well as oligomers in electron micrographs (Fig. S4), which is also supported by DLS experiments. At 25 $^{\circ}\text{C}$, we observed an increase in the ThT fluorescence signal over a period of 8–40 h (Fig. S5) and a subsequent gradual decline below the baseline (inset of Fig. 3C, red trace). Indeed, EM studies highlight that the length of fibrils is temperature-dependent and the ones grown at lower temperatures (such as 20 $^{\circ}\text{C}$) are considerably longer than those cultivated at higher temperatures. We hypothesize that the time-dependent reduction in fluorescence signal accompanying the formation of longer fibrils may arise due to increased precipitation and/or adhesion to the plate walls, compared with conditions where numerous smaller fibrils are formed. At temperatures above 37 $^{\circ}\text{C}$, in all conditions tested using ITC, we observed a pronounced exothermic reaction. Alongside, the ThT assay at 45 $^{\circ}\text{C}$ revealed a slow tonic increase in fluorescence like the one observed at 37 $^{\circ}\text{C}$, but with far weaker intensity (Fig. S5); this observation is in agreement with both EM and CD measurements revealing amorphous aggregates and formation of several short fibrils respectively, at this temperature (Figs. 1C and S4). Furthermore, in DLS experiments conducted at 60 $^{\circ}\text{C}$ we did not observe a distinct population but larger particles in nm- μm dimensions, while electron micrographs revealed (Fig. S4) the presence of bundles and clusters of short fibrils. Taken together, this would imply that at higher temperature, the DISC1 C-region undergoes nucleation at an enhanced rate, resulting in fibrils that are more numerous, heterogeneous, and shorter in length because of a rapidly

decreasing concentration of free protein, while at lower temperatures, one observes a slow steady nucleation and longer homogeneous fibrillar growth.

Next, we investigated the effect of protein concentration on the heat exchanges during fibrillization at 37 $^{\circ}\text{C}$. Higher concentrations resulted in isotherms with a broad exothermic burst similar to those for higher temperatures (Fig. 3B). We calculated an enthalpy exchange of $-297.74 \text{ kJ mol}^{-1}$ (Fig. 3B and Table 1) at 37 $^{\circ}\text{C}$, while values ranging from -1.5 to -150 kJ mol^{-1} have been observed for other amyloid systems [29, 30, 48, 49]. Next, we calculated the heat-capacity change for the formation of fibrils and observed steep ΔCp (Fig. 3A) trends of 113.84 and $-373.51 \text{ kJ mol}^{-1} \text{ K}^{-1}$ (Table 1). In almost all amyloids [29, 30, 48, 49] investigated, one observes a negative tendency of ΔCp during fibrillization; besides tight packing of residues in newly formed β -core regions, this effect may arise from solvent-related phenomena such as desolvation of hydrophobic surface areas and trapping of water molecules at protein-protein interfaces.

Another notable feature in all cases of cross- β -fibrillar proteins is a negative entropy change^{24,25,51}, consistent with assembly of highly ordered structures with decreased bulk-solvent access and presence of ordered water molecules. Our calculations of ΔS indeed yield negative values of $-0.9 \text{ kJ mol}^{-1} \text{ K}^{-1}$ (Table 1). Therefore, in order to probe the presence of water accessible at the surface of the fibril, we performed water-edited ssNMR experiments (Fig. 3D, E). Previous studies on A β [50] have shown that ^1H T_2 relaxation times of approximately 250, 150 and 90 ms can describe the presence of water as bulk, interfibrillar, and fiber bound, respectively. Other reports on classical amyloids have shown from build-up experiments that ordered water molecules associated with the core region of β -fibrils [37, 50, 51] in the range of 17–110 ms. Altogether, the interaction between water and the core region occurs in the tens of ms range. While analyzing region 50–75 ppm that is predominated by signals arising from backbone C α in the 1D water build-up experiment (Fig. S6), we observed that the DISC1 C-region contains tightly bound water and requires initial polarization transfer times t_m^s of 45 ms (Fig. 3D). This would represent the presence of water in close proximity to the surface area. In turn, this would correspond to a fibrillary core with a diameter of 6.8 nm, agreeing well with our observations from the EM images of the fibril (approximately 5–7 nm). We next performed 2D water edited ^{13}C - ^{13}C spin diffusion experiments at various ^1H - ^1H mixing times ranging from 5 to 100 ms (Fig. S6). A close analysis in the amino acid-specific regions, viz., Ala, Leu, and Ser (Fig. 3E) reveals the following: for the residue Leu, in all tested conditions, we noticed an approximate equidistribution of population intensities for all SSE with an t_m^s of 19 ms for sheets and coils, and 13 ms for helices. In the case of Ser, we noticed signals arising from β -strands and coils that display build-up times of 15 and 11 ms, respectively. Only in the short mixing time of 5 ms, we observe a weak signal for Ser in α -helix (Fig. S6). The steeper build-up profile for Ser is unsurprising considering that the amino acid side chain is polar and favors interaction with water. Resonances from Ala, however, are weakly observed at lower mixing times and increase starkly from 20 ms onward. The steep increase in the signal profile for all three SSE corresponds to build-up times of 13.5–14 ms.

Altogether, our findings highlight that there is no preferential transfer of magnetization from water to the different SSE and residues of the protein. Due to the unavailability of a high-resolution 3D structure of the DISC1 C-region β -core, we are rather limited with the interpretation of amino acid-specific interaction with water. We can, however, readily deduce: first, the water-edited signals arise due to the fibrils and few residues within, which illustrates that the DISC1 C-region exhibits the presence of ordered water molecules. Second, the association of DISC1 C-region into oligomers/fibrils is enthalpy driven alike other amyloid systems [29, 30, 48, 49].

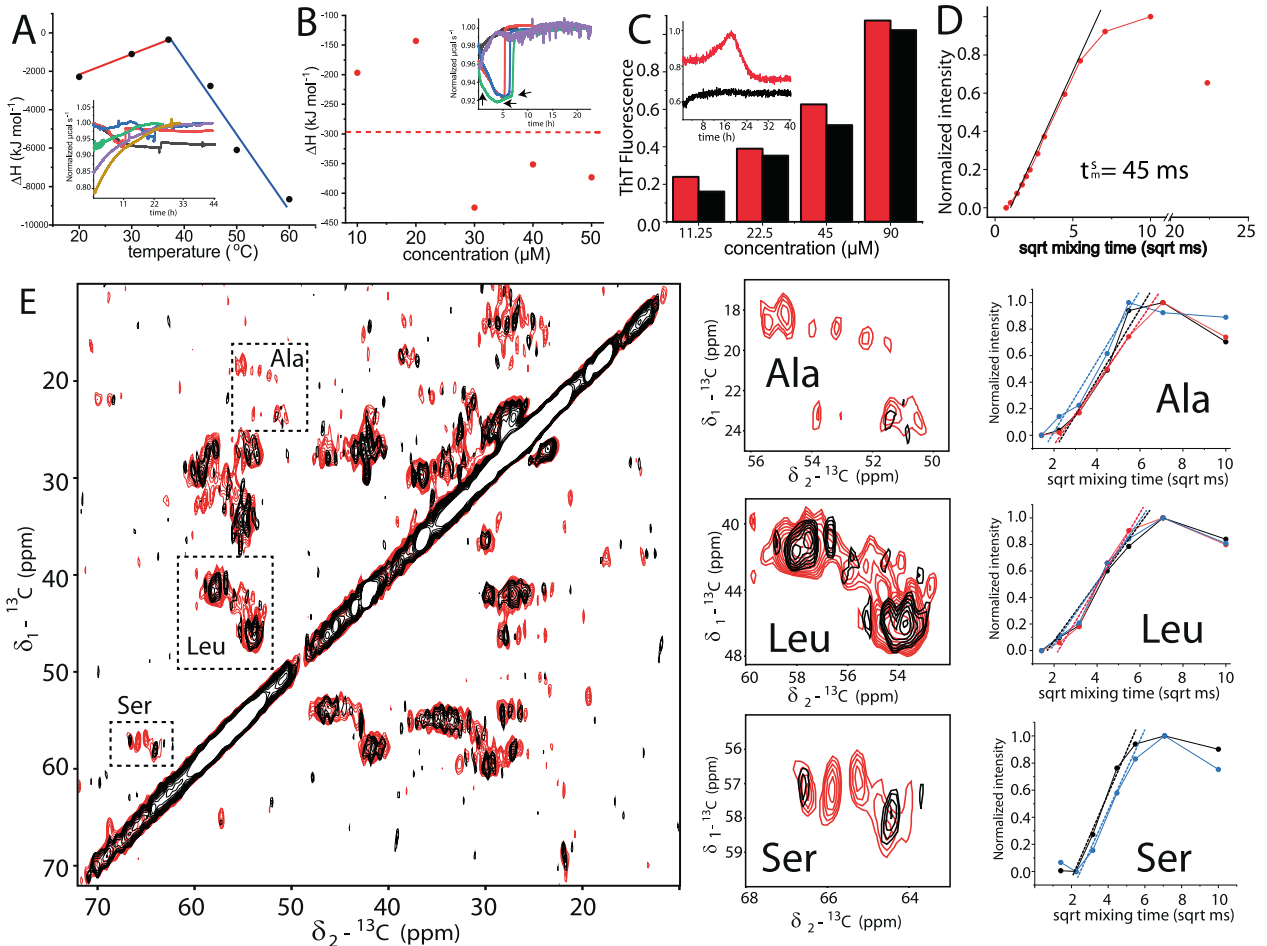


Fig. 3 Mechanistic and thermodynamic characterisation of C-region aggregation. **A** Thermodynamic characterization of the DISC1 C-region via ITC. The enthalpy changes observed for a 10 μM protein solution are plotted as a function of temperature; the inset shows the corresponding thermograms (20 $^{\circ}\text{C}$, black; 30 $^{\circ}\text{C}$, red; 37 $^{\circ}\text{C}$, blue; 45 $^{\circ}\text{C}$, green; 50 $^{\circ}\text{C}$, purple; and 60 $^{\circ}\text{C}$, golden). Heat capacities (ΔC_p) of 113.84 (red) and -373.51 (blue) $\text{kJ mol}^{-1} \text{K}^{-1}$ were derived from the slopes of the two phases using a linear fit. **B** The enthalpy change ($\Delta H = -297.74 \text{ kJ mol}^{-1}$) was calculated by normalizing the observed heat changes in the thermograms by protein amount (in inset; 10 μM (black), 20 μM (purple), 30 μM (red), 40 μM (blue), and 50 μM (green) protein solutions). **C** ThT assay highlighting the fluorescence intensity (in arbitrary units) of the DISC1 C-region at 25 $^{\circ}\text{C}$ (red) and 37 $^{\circ}\text{C}$ (black) as a function of protein concentration. The inset figure shows the raw data obtained at 45 μM protein concentration (also see Fig. S5). **D** 1D water build-up curves of CP intensity in predominantly $\text{C}\alpha$ region corresponding to signals in the region 50–75 ppm (Fig. S6). Dotted line represents the fitted slope that was used for estimation of t_m^s . **E** Left panel displays 2D ^{13}C – ^{13}C correlation spectra with ^1H – ^1H mixing times of 10 ms (black) and 100 ms (red), the right panels depict close-up views for three spectral regions corresponding to the signals from residues Ala, Leu and Ser and their respective water build-up profiles at different ^1H – ^1H mixing times (refer to Fig. S6 for individual spectrum). Corresponding signals arising from secondary-structure elements (SSE) like helices, strands, and coils are color-coded in red, blue, and black, respectively. The corresponding colored dotted lines represent the slopes arising due to individual SSE and were used further for estimation of t_m^s .

Table 1. Thermodynamic properties of DISC1 C-region association.

ΔH [kJ/mol]	ΔG_{app} [kJ/mol]	ΔS [kJ/mol/K]	ΔC_p [kJ/mol/K]
–297.74	–34.36	–0.9	113.84 & –373.51

Oligomerization/fibrillization increases affinity and promotes cooperativity of C-region binding to other physiological partners

The interaction of LIS1 and NDEL1 with DISC1 was studied using SPR, where the C-region protein was immobilized to the sensor chip. We chose to study three different preparations of the DISC1 C-region, in order to assess the impact of oligomerization. As the His₆-tagged version of the protein readily aggregated, we used a freshly prepared sample of MBP-fused C-region to represent the monomer. In fact, ITC and DLS measurements (Fig. S1) showed that the MBP-C region is more stable than the His₆-C region, yet with time displays similar tendencies to oligomerize/fibrillize.

Samples of His-C-region and MBP-C region incubated at 4 $^{\circ}\text{C}$ for two weeks and two months, respectively, served as representatives of the aggranzed fraction.

At first, to see if there were any qualitative differences between the monomeric and oligomeric/fibrillary species, we studied binding of a camelid nanobody (V_HH B5), which we have extensively characterized previously [22] (Figs. 4A and S7). The isotherm was well described using a simple hyperbolic binding model in all three cases, however, both oligomeric fractions showed slightly altered affinities in comparison with the monomeric MBP-C-region. The K_D values (Table 2) calculated were 46, 66, and 70 nM, respectively. Similar to our observations with V_HH B5, the monomeric C-region displayed hyperbolic binding isotherms with K_D values of 71 nM and 28 μM to LIS1 and NDEL1 (Fig. 4B, C), respectively. It is important to note, however, that the higher-order oligomers displayed a significant sigmoidal trend, indicating cooperative binding. In the case of LIS1, we determined Hill coefficients of 2.4 and 2.8 along with K_D values of 31 nM for

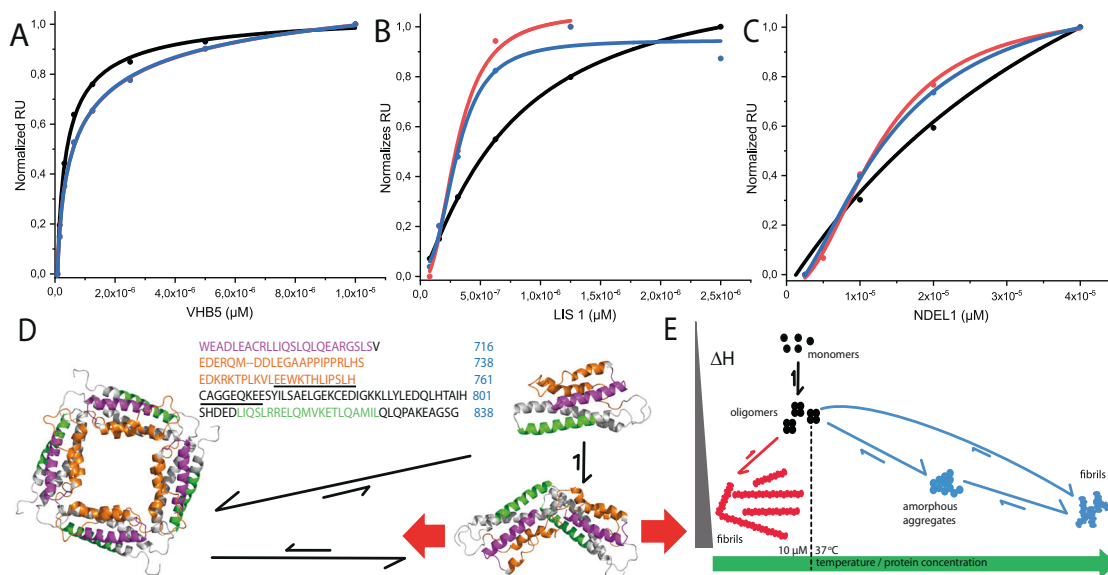


Fig. 4 Impact of C-region self-association on physiological DISC1 interactions. Binding isotherms measured using SPR (Fig. S7 and Table 2) representing protein–protein interaction of (A) V_HB5, (B) LIS1, and (C) NDEL1 with monomeric MBP-C-region (black), oligomeric/fibrillar His₆-C region (red), and oligomeric/fibrillar MBP-C region (blue). **D** Representation of structural changes in the DISC1 C-region transitioning from monomer to oligomers/fibrils. Protein sequence is numbered in blue in the context of the full-length DISC1 protein; V_HB5-binding site is displayed in magenta, (pseudo)repeat sequence in orange, the underlined part describes residues in Δ22 and the green region represents the C-terminal NDEL1-binding site. The same color coding is used in ribbon representations of protein models and the red arrows indicate directions of filament extension (also see Figs. 2D and S8). **E** Schematic representation of the energy and conformational landscape of DISC1 C-region self-association. The monomer and/or oligomer versions of the DISC1 C-region show a propensity to fibrillize into longer filaments at temperatures of 37 °C. With increasing temperature and protein concentration, accelerated nucleation yields more numerous but shorter fibrils along with amorphous aggregates. The fibrillar pathways are color-coded according to the temperature domains depicted in Fig. 3 A.

Table 2. Summary of binding constants.

	V _H B5 K _D [nM] (n)	NDEL1 K _D [μM] (n)	LIS1 K _D [nM] (n)
Monomer (M)	46 (1.4)	20 (1.4)	71 (1.2)
Oligomer/Fibril (H)	66 (1.1)	13 (2.1)	31 (2.4)
Oligomer/Fibril (M)	70 (1.1)	14 (2.5)	31 (2.8)

Legend: n – cooperativity index, M – MBP fusion construct, H – His₆ fusion construct.

both the His and MBP versions, respectively. In a similar manner, NDEL1 displayed cooperative binding (Hill coefficients of 2.1–2.5) together with higher apparent affinities (K_D of 13 and 14 μM) for the oligomeric versions. These observations indicate that oligomerization and/or fibrillization of the DISC1 C-region not necessarily lead to sequestration of inactive material, but—at least for certain partners—tend to enhance protein–protein interactions. Potential implications of this behavior are discussed below.

DISCUSSION

The seminal issue addressed by the current work concerns the question if and to what extent it is possible to structurally characterize DISC1 aggregates. Indeed, our findings reported herein shed light onto several structural and physiological tenets regarding the human DISC1 protein. Using a combination of ssNMR spectroscopy and EM measurements, we are able to show that the C-terminal segment forms oligomers and fibrils, the latter displaying a regular pitch along their axis as well as characteristics of amyloid structure. This piqued our curiosity whether the C-region of the DISC1 protein (perhaps also full-length human DISC1) may behave in ways comparable to other proteinopathies.

In fact, one report has described ThT-positive [52] full-length recombinant human DISC1 fibrils, suggesting a potential amyloidogenic nature of DISC1, even though, to this end, DISC1 amyloid has not been reported either in human post-mortem brains or in animal models [12]. Moreover, for nearly a decade, it was believed that DISC1 may resemble other protein risk factors in various neurological disorders [9], however, structural data at atomistic resolution have been severely lacking.

Amyloid systems, including PrP [53], Tau [54], and RIPK1-K3 [55] heteromers to name a few, have a small fraction of their sequences embedded in the β-core, while the remaining larger parts of the proteins often comprise dynamic regions. One exception may be the in vitro-obtained Aβ amyloid-fibril structure [56], where all 42 residues are part of the fibril core. Our findings from solution NMR and INEPT-based ssNMR experiments highlight a dynamic and highly mobile C-terminal end of the protein. On another note, a protomer–fibril equilibrium [57] has been shown to exist in α-synuclein where monomer-dominated signals were observed in INEPT and fibril-derived resonances populated the CP-based spectra. Likewise, it appears conceivable that our observations from the NMR-spectroscopy data may not only represent differential mobility within fibers, but also point to a dynamic equilibrium between oligomers and fibrils, which is indeed well supported by our findings from DLS, ITC, and EM recordings.

This raises several important questions: why does the protein aggregate and which regions of the DISC1 protein contribute to fibrillization? While the ultimate proof is still lacking, our observations together with published data do allow several important conclusions. First, notable signature motifs of amyloid fibrillization require the presence of hydrophobic residues [58] within short peptide sequences [59], repetitions (single amino acid repeats [37] like polyQ; tandem repeats [60]; prion-like repeats [60, 61]; and pseudorepeats [54, 62]) with an intrinsic propensity to form β-strands. While a few of these features have been incorporated in amyloid prediction softwares, these programs

were all unsuccessful in determining the β -core interface of DISC1. A manual search, however, for (pseudo)repeats using the program RADAR [63] identified one such long tandem motif covering residues 717–737 and 739–761. This region is abundantly rich in amino acids like Ala, Gly, Leu, Pro, and Ser, which we show have CS reminiscent of residues in strands. Also, the sequence contains hydrophobic residues that may support formation of a β -core.

Next, we sought to address the physiological consequences of aggregation as well as the question whether the oligomer is crucial for DISC1 functioning or represents a pathological species. From a physiological perspective, the pseudorepeat is located in the vicinity of crucial residues such as S704 [25] and S713 [11], on one side, it comprises the $\Delta 22$ region [64, 65] (see below) and is flanked by the NDEL1-binding region on the other side. The absence of residues 748–769 as a result of alternative splicing leads to a variant that is referred to as DISC1 $\Delta 22aa$ and displays a marked decrease in binding to NDEL1 [65]. Similarly, it has been convincingly shown by physiological [65, 66] and structural [24] means that region 807–828 is necessary for NDEL1 binding. We speculate that the plasticity surrounding the $\Delta 22$ /(pseudo)repeat region supports the formation of a β -structure from coil, which may be crucial for different types of self-association, including oligomerization and fibrillization. Notably, our structural models suggest that the resulting conformational change leaves both the N- and the C-terminal portions intact, thus explaining the marginal alteration in the affinity of the nanobody V_HH B5 [22], which binds in the region 691–715 in a noncooperative manner. Presumptively, oligomerization promoted due to the (pseudo)repeat region may facilitate cooperative binding of physiological partners wherein the binding partner interacts with 2 protomers of the C-region. In a pathological scenario, the same segment could possibly promote fibrillization (Figs. 4D and S8), resulting in a loss-of-function phenotype.

Our findings settle a long-debated question, i.e., whether NDEL1 interacts with monomeric or oligomeric DISC1. Previously, it has been shown that monomeric DISC1 and the S704C-associated oligomers bind to NDEL1 with similar affinities [67]. However, other findings indicated that higher oligomers [10, 68], presumably octamers based on gel-filtration profiles, displayed a greater affinity for NDEL1. We clearly show that, in fact, both results are valid: NDEL1 does indeed bind to both the monomer and the oligomers, but with varying affinities and degrees of cooperativity. That said, it is important to note that the sample representing oligomers may/will also contain fibrils. Does NDEL1 equally bind oligomers and fibers? Perhaps yes, this would indeed be supportive of the loss-of-function phenotype affecting neurite formation [69, 70] and neuronal development [10, 68], which has been observed in conjunction with the precipitation of DISC1. On the other hand, if binding is exclusively associated with the oligomers [10, 68], then they may represent either the physiologically relevant state as present in the dynein complex or a pathogenic species analogous to other amyloid systems, where the fibrils and the aggresomes thereof occur as a nontoxic end product.

Notably, our attempts to study the DISC1 C-region interaction with NDEL1 were not straightforward, perhaps due to the propensity of the latter to aggregate into fibrils [27]. It will be interesting to see if the fibrils from the NUDEL system, viz., NDEL1 and NDEL1, also contain elements of β -strands. This raises the very pertinent question whether, in addition to the similarities in structural architecture to amyloid diseases, DISC1 aggregation bears any semblance to other known proteinopathies. Akin to aggresomes observed for Alzheimer's and Parkinson's diseases, where a multitude of hallmark fibrillar proteins co-aggregate, DISC1opathies [9], i.e., proteinopathies with aggregation or misassembly of the DISC1 protein, may feature a similar tendency for DISC1 and some of its partner proteins. Indeed, our findings on the DISC1 C-region showed the presence of oligomers and fibrils

using EM and DLS; biophysical characterization revealed broad enthalpy changes and a negative entropy value using ITC, inhomogeneous line broadening, and mobile and rigid portions of the protein using ssNMR spectroscopy, altogether pointing to heterogeneity and plasticity associated with the protein (Fig. 4E). Interestingly though, amyloid fibrils come in various shapes and architectures. Recently, with the advance in state-of-the-art structural biology techniques and equipment, several new fibril structure details have been reported at atomic and near-atomic level (for review see ref. [71]). It is quite bewildering to note the differences in handedness, fibrillar length, helical axis position, and pitch length. Amyloid fibrils can have widths ranging from 5 to 30 nm and in some cases even hundreds of nm, can be twisted with helical pitches that are extremely short and few may also appear as flat ribbons, with diverse physiological roles (for excellent reviews, readers are referred to [72, 73]). Along these lines, DISC1 C-region fibrils fit well into the grand picture. As far as the driving concentrations for nucleation and aggregation are concerned, in most cases, monomeric concentrations of the amyloid fibril proteins are in the nM to low μ M range under physiological conditions, and are tightly regulated by protein synthesis and degradation. Similarly, expression of the DISC1 protein is controlled by the cellular homeostasis machinery. However, under pathological circumstances, an increase of the monomer concentration is likely to occur. This in particular is very interesting because, unlike other amyloid systems where the fibrillary deposits are extensive and readily visualizable by staining techniques, sarkosyl-resistant DISC1 aggresomes obtained from post-mortem brain tissues are highly heterogeneous and usually require enrichment [9, 10]. Contributing to this is the postnatal decrease of expression levels of DISC1 protein [74]. On a side note, due to the poor solubility of the full-length protein, which rendered the system untenable for biophysical characterization, the EPRIT technique was successfully employed to identify the four more soluble regions of the DISC1 protein [21]. Therefore, based on all the earlier findings and in comparison to the classical amyloids from neurodegenerative diseases, we estimate that the DISC1 monomer may be present in pM to nM concentrations within the cell.

Are the oligomers a functional entity and/or do they perhaps act as seeds for fibril formation? This is an unresolved question, but being a scaffold protein, DISC1 has been shown to influence the activity of a variety of enzymes and other proteins [75]. For myriads of physiological roles, a coiled-coil structure is a very convenient feature offering plenty of conformational flexibility. In the context of our work, DISC1, LIS1, and NDEL1 form a functional triad as part of the dynein complex [26]. It is believed that these proteins mediate opposing interactions wherein LIS1 locks dynein movement, while NDEL1 binding releases it, providing the primary stroke. For these interactions to occur with dynein, a dimeric or tetrameric assembly of LIS1 and NDEL1 is a prerequisite [26]. Also, it has been shown that DISC1 monomers bind NDEL1 [67], likewise, we observe a similar interaction in a noncooperative manner, but that may not be sufficient to promote dynein activity. The C-region of DISC1 has been previously shown to harbor a "multimerization/octamerization" and a "dimerization" domain in regions 668–747 and 765–854 [10, 68]. It has been suggested that oligomerization, presumably octamerization, preferentially promotes NDEL1 binding. This region inherently has a tendency to form sarkosyl-resistant precipitates similar to the ones from brain [10], and the propensity to aggregate is further enhanced by the S704C mutant [68], suggesting a possible disruption of the quaternary structure that may lead to a loss-of-function phenotype. Our findings point in the direction that the structural changes in the region comprising the (pseudo)repeat and the $\Delta 22$ (748–769) region likely promote (homo)oligomerization (Fig. 4D). This, in turn, may culminate into cooperative association of DISC1 C-region with NDEL1 and/or LIS1 in promoting dynein activity

during the mitotic cycle. However, at this juncture, we cautiously tread on the interplay between the different binding partners because we cannot rule out the possibility of the presence of oligomeric NDEL1 or LIS1 in the analyte solution. It would be very interesting to further dissect if cooperativity arises with monomeric LIS1/NDEL1 or involves (homo)oligomers interacting with the DISC1 C-region (homo)oligomers.

Given that structure and function are two sides of the same coin, our findings help to reconcile divergent results obtained across various experiments and platforms, including inconsistencies with respect to the binding of DISC1 and NDEL1, the mode of interaction with different oligomeric species to name a few. Moreover, the C-region controls proper DISC1 self-association; truncated versions lacking the C-region [1, 2] resemble the disease-associated variant that was observed in the Scottish family with severe chronic mental illnesses. The truncated DISC1 protein associates with the full-length human DISC1, dissociating it from the dynein complex and redistributing it [69], thereby hampering neuronal development [70]. In light of the above, we believe that the results of this study provide a strong foundation for understanding the (patho)physiology of the DISC1 protein in quantitative biophysical terms. Is it possible to prevent the aggregation of DISC1 protein or reverse the process? While our research findings here provide some insight into the potential pathological mechanism, it is too early to answer this question. As one can envision, the findings presented herein have raised a lot more queries than have been addressed. For instance, how can we translate the structural and biophysical information to other mutants of the DISC1 C-region relevant in schizophrenia pathophysiology, such as the S704C, S713E, and L807 (frameshift) to name a few? From the strategy presented herein, the S704C and frameshift mutants represent genetic variants that can be readily targeted for structural characterization, screening for therapeutic applications due to their propensity to misassemble and aggregate. As highlighted above, the S704C mutant has provided a wealth of information from numerous reports on aggregation propensity of DISC1 and its oligomeric specific association with NDEL1 [67, 68]. From the clinical perspective, this mutation has been identified with increased presence of aggregates in brain tissues and is known to be associated with major depressive disorder (MDD), schizophrenia, increase in brain volume, and decreased cognitive function with age [9]. Analogously, another variant that is of clinical significance that was identified due to familial schizophrenia is the frameshift mutant at position L807, which was first identified in the American family [7]. This mutation results in the replacement of amino acids 809–854 with nine residues at the C terminus, resulting in severe aggregation as well as loss of the binding site for the NDEL1 protein. For both the S704C and the L807-frameshift mutants, we speculate that these residues are critical elements of the heptad repeats within the coiled coil region under native conditions. This conserved motif is populated by charged and polar residues such as Ser, and hydrophobic residues such as Leu in positions *a* and *d*; any mutation that destabilizes this structural core may alter the architecture of the $\Delta 22$ and adjacent regions, thereby impacting on its binding to other protein partners and eventually resulting in aggregation. Interestingly and in similar vein, other mutants within the $\Delta 22$ region such as T754S and P758R result in loss of binding to several key neuronal modulators such as DIX domain-containing 1 (Dixdc1), LIS1, and pericentriolar material 1 (PCM1) to name a few [8]. But given the fact that DISC1 is a scaffold protein, which interacts with over 300 different proteins, the coiled-coil region provides a facile conformational flexibility, especially for post-translational modification as a convenient switch to specify interaction partners. For example, phosphorylation of one crucial site (S713) by protein kinase A (PKA) results in recruitment of Bardet–Biedl syndrome (BBS) proteins to the centrosome, while the dephosphorylated protein interacts with GSK3 β in the Wnt

signaling pathway [11]. Under such conditions, it is plausible that phosphorylation might disrupt the coiled-coil interactions and facilitate the formation of disordered or β -strand conformations to promote interactions with the BBS proteins; this hypothesis will be subject of our future efforts using biophysical approaches to further comprehend the role of the C-region in DISC1 function. It is important to study the individual regions, i.e., the N-terminal disordered region and the four structured regions, which harbor several sites for post-translational modifications and well-characterized mutations promoting misassembly and aggregation, in a reductionist manner to comprehend their significance in the full-length protein. That said, due to the physiological and psychiatric relevance of DISC1, we are of the firm gestalt opinion that the whole is indeed larger than the sum of individual parts; atomistic detail characterization of the full-length DISC1 protein poses a significant challenge that we intend to tackle in the near future.

REFERENCES

- Blackwood DHR, Fordyce A, Walker MT, St Clair DM, Porteous DJ, Muir WJ. Schizophrenia and affective disorders - Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. *Am J Hum Genet.* 2001;69:428–33.
- Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, Semple CAM, et al. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum Mol Genet.* 2000;9:1415–23.
- Bradshaw NJ, Korth C. Protein misassembly and aggregation as potential convergence points for non-genetic causes of chronic mental illness. *Mol Psychiatry.* 2019;24:936–51.
- Tropea D, Hardingham N, Millar K, Fox K. Mechanisms underlying the role of DISC1 in synaptic plasticity. *J Physiol.* 2018;596:2747–71.
- Hikida T, Gamo NJ, Sawa A. DISC1 as a therapeutic target for mental illnesses. *Expert Opin Ther Targets.* 2012;16:1151–60.
- Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet.* 2013;45:1150–9.
- Sachs NA, Sawa A, Holmes SE, Ross CA, DeLisi LE, Margolis RL. A frameshift mutation in disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Mol Psychiatry.* 2005;10:758–64.
- Soares DC, Carlyle BC, Bradshaw NJ, Porteous DJ. DISC1: structure, function, and therapeutic potential for major mental illness. *ACS Chem Neurosci.* 2011;2:609–32.
- Korth C. Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion* 2012;6:134–41.
- Leliveld SR, Bader V, Hendriks P, Prikulis I, Sajjani G, Requena JR, et al. Insolubility of disrupted-in-schizophrenia 1 disrupts oligomer-dependent interactions with nuclear distribution element 1 and is associated with sporadic mental disease. *J Neurosci.* 2008;28:3839–45.
- Ishizuka K, Kamiya A, Oh EC, Kanki H, Seshadri S, Robinson JF, et al. DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature* 2011;473:92–6.
- Trossbach SV, Bader V, Hecher L, Pum ME, Masoud ST, Prikulis I, et al. Misassembly of full-length Disrupted-in-Schizophrenia 1 protein is linked to altered dopamine homeostasis and behavioral deficits. *Mol Psychiatry.* 2016;21:1561–72.
- Zhu S, Abounit S, Korth C, Zurzolo C. Transfer of disrupted-in-schizophrenia 1 aggregates between neuronal-like cells occurs in tunnelling nanotubes and is promoted by dopamine. *Open Biol.* 2017;7:160328.
- Yerabham AS, Weiergraber OH, Bradshaw NJ, Korth C. Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol Chem.* 2013;394:1425–37.
- Millar JK, Christie S, Porteous DJ. Yeast two-hybrid screens implicate DISC1 in brain development and function. *Biochem Biophys Res Commun.* 2003;311:1019–25.
- Millar JK, Pickard BS, Mackie S, James R, Christie S, Buchanan SR, et al. DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science* 2005;310:1187–91.
- Murdoch H, Mackie S, Collins DM, Hill EV, Bolger GB, Klusmann E, et al. Isoform-selective susceptibility of DISC1/phosphodiesterase-4 complexes to dissociation by elevated intracellular cAMP levels. *J Neurosci.* 2007;27:9513–24.
- Sawamura N, Ando T, Maruyama Y, Fujimuro M, Mochizuki H, Honjo K, et al. Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. *Mol Psychiatry.* 2008;13:1138–48.
- Sanchez-Pulido L, Ponting CP. Structure and evolutionary history of DISC1. *Hum Mol Genet.* 2011;20:R175–81.

20. Yumerefendi H, Tarendeau F, Mas PJ, Hart DJ. ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J Struct Biol.* 2010;172:66–74.
21. Yerabham ASK, Mas PJ, Decker C, Soares DC, Weiergraber OH, Nagel-Steger L, et al. A structural organization for the Disrupted in Schizophrenia 1 protein, identified by high-throughput screening, reveals distinctly folded regions, which are bisected by mental illness-related mutations. *J Biol Chem.* 2017;292:6468–77.
22. Yerabham ASK, Muller-Schiffmann A, Ziehm T, Stadler A, Kober S, Indurkha X, et al. Biophysical insights from a single chain camelid antibody directed against the Disrupted-in-Schizophrenia 1 protein. *PLoS ONE.* 2018;13:e0191162.
23. Wang X, Ye F, Wen Z, Guo Z, Yu C, Huang WK, et al. Structural interaction between DISC1 and ATF4 underlying transcriptional and synaptic dysregulation in an iPSC model of mental disorders. *Mol Psychiatry.* 2019;26:1346–60.
24. Ye F, Kang E, Yu C, Qian X, Jacob F, Yu C, et al. DISC1 regulates neurogenesis via modulating kinetochore attachment of Ndel1/Nde1 during mitosis. *Neuron* 2017;96:1204.
25. Hashimoto R, Numakawa T, Ohnishi T, Kumamaru E, Yagasaki Y, Ishimoto T, et al. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Hum Mol Genet.* 2006;15:3024–33.
26. Bradshaw NJ, Hayashi MA. NDE1 and NDEL1 from genes to (mal)functions: parallel but distinct roles impacting on neurodevelopmental disorders and psychiatric illness. *Cell Mol Life Sci.* 2017;74:1191–210.
27. Soares DC, Bradshaw NJ, Zou J, Kennaway CK, Hamilton RS, Chen ZA, et al. The mitosis and neurodevelopment proteins NDE1 and NDEL1 form dimers, tetramers, and polymers with a folded back structure in solution. *J Biol Chem.* 2012;287:32381–93.
28. Jarrett JT, Lansbury PT Jr. Seeding “one-dimensional crystallization” of amyloid: a pathogenic mechanism in Alzheimer’s disease and scrapie? *Cell* 1993;73:1055–8.
29. Ikenoue T, Lee YH, Kardos J, Yagi H, Ikegami T, Naiki H, et al. Heat of supersaturation-limited amyloid burst directly monitored by isothermal titration calorimetry. *Proc Natl Acad Sci USA.* 2014;111:6654–9.
30. Kardos J, Yamamoto K, Hasegawa K, Naiki H, Goto Y. Direct measurement of the thermodynamic parameters of amyloid formation by isothermal titration calorimetry. *J Biol Chem.* 2004;279:55308–14.
31. Seidel K, Lange A, Becker S, Hughes CE, Heise H, Baldus M. Protein solid-state NMR resonance assignments from $(^{13}\text{C}, ^{13}\text{C})$ correlation spectroscopy. *Phys Chem Chem Phys.* 2004;6:5090–3.
32. Chevelkov V, Rehbein K, Diehl A, Reif B. Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration. *Angew Chem Int Ed Engl.* 2006;45:3878–81.
33. Hardy EH, Verel R, Meier BH. Fast MAS total through-bond correlation spectroscopy. *J Magn Reson.* 2001;148:459–64.
34. Hong M, Griffin RG. Resonance assignments for solid peptides by dipolar-mediated $^{13}\text{C}/^{15}\text{N}$ correlation solid-state NMR. *J Am Chem Soc.* 1998;120:7113–4.
35. Schmidt-Rohr K, Spiess HW CHAPTER ONE - Introduction. In: Schmidt-Rohr K, Spiess HW, editors. *Multidimensional solid-state NMR and polymers.* San Diego: Academic Press; 1994. 1–12.
36. Ader C, Schneider R, Seidel K, Etkorn M, Becker S, Baldus M. Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy. *J Am Chem Soc.* 2009;131:170–6.
37. Schneider R, Schumacher MC, Mueller H, Nand D, Klaukien V, Heise H, et al. Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J Mol Biol.* 2011;412:121–36.
38. Gradmann S, Ader C, Heinrich I, Nand D, Dittmann M, Cukkemane A, et al. Rapid prediction of multi-dimensional NMR data sets. *J Biomol NMR.* 2012;54:377–87.
39. Lee W, Tonelli M, Markley JL. NMRFA-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* 2015;31:1325–7.
40. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife.* 2018;7:e42166.
41. Rohou A, Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol.* 2015;192:216–21.
42. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993;234:779–815.
43. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. *Nat Protoc.* 2017;12:255–78.
44. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, et al. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol.* 2007;157:38–46.
45. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25:1605–12.
46. Zhou DH, Shah G, Cormos M, Mullen C, Sandoz D, Rienstra CM. Proton-detected solid-state NMR spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J Am Chem Soc.* 2007;129:11791–801.
47. Wang Y, Jardetzky O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* 2002;11:852–61.
48. Jeppesen MD, Westh P, Otzen DE. The role of protonation in protein fibrillation. *FEBS Lett.* 2010;584:780–4.
49. Ikenoue T, Lee YH, Kardos J, Saiki M, Yagi H, Kawata Y, et al. Cold denaturation of alpha-synuclein amyloid fibrils. *Angew Chem Int Ed Engl.* 2014;53:7799–804.
50. Wang T, Jo H, DeGrado WF, Hong M. Water distribution, dynamics, and interactions with Alzheimer’s beta-amyloid fibrils investigated by solid-state NMR. *J Am Chem Soc.* 2017;139:6242–52.
51. Dregni AJ, Duan P, Hong M. Hydration and dynamics of full-length tau amyloid fibrils investigated by solid-state nuclear magnetic resonance. *Biochemistry* 2020;59:2237–48.
52. Tanaka M, Ishizuka K, Nekooki-Machida Y, Endo R, Takashima N, Sasaki H, et al. Aggregation of scaffolding protein DISC1 dysregulates phosphodiesterase 4 in Huntington’s disease. *J Clin Invest.* 2017;127:1438–50.
53. Helmus JJ, Surewicz K, Nadaud PS, Surewicz WK, Jaroniec CP. Molecular conformation and dynamics of the Y145Stop variant of human prion protein in amyloid fibrils. *Proc Natl Acad Sci USA.* 2008;105:6284–9.
54. Fitzpatrick AWP, Falcon B, He S, Murzin AG, Murshudov G, Garringer HJ, et al. Cryo-EM structures of tau filaments from Alzheimer’s disease. *Nature* 2017;547:185–90.
55. Mompean M, Li W, Li J, Laage S, Siemer AB, Bozkurt G, et al. The structure of the necrosome RIPK1-RIPK3 Core, a human hetero-amyloid signaling complex. *Cell* 2018;173:1244–53 e10.
56. Gremer L, Scholzel D, Schenk C, Reinartz E, Labahn J, Ravelli RBG, et al. Fibril structure of amyloid-beta(1–42) by cryo-electron microscopy. *Science* 2017;358:116–9.
57. Heise H, Hoyer W, Becker S, Andronesi OC, Riedel D, Baldus M. Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc Natl Acad Sci USA.* 2005;102:15871–6.
58. Gazit E. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J.* 2002;16:77–83.
59. Gazit E. Mechanisms of amyloid fibril self-assembly and inhibition. Model short Peptides as a key research tool. *FEBS J.* 2005;272:5971–8.
60. Gazit E. Global analysis of tandem aromatic octapeptide repeats: the significance of the aromatic-glycine motif. *Bioinformatics* 2002;18:880–3.
61. Cherny I, Rockah L, Levy-Nissenbaum O, Gophna U, Ron EZ, Gazit E. The formation of Escherichia coli curli amyloid fibrils is mediated by prion-like peptide repeats. *J Mol Biol.* 2005;352:245–52.
62. Wasmer C, Lange A, Van Melckebeke H, Siemer AB, Riek R, Meier BH. Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 2008;319:1523–6.
63. Heger A, Holm L. Rapid automatic detection and alignment of repeats in protein sequences. *Proteins* 2000;41:224–37.
64. Taylor MS, Devon RS, Millar JK, Porteous DJ. Evolutionary constraints on the disrupted in Schizophrenia locus. *Genomics* 2003;81:67–77.
65. Kamiya A, Tomoda T, Chang J, Takaki M, Zhan C, Morita M, et al. DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. *Hum Mol Genet.* 2006;15:3313–23.
66. Brandon NJ, Handford EJ, Schurov I, Rain JC, Pelling M, Duran-Jimeniz B, et al. Disrupted in Schizophrenia 1 and Nudel form a neurodevelopmentally regulated protein complex: implications for schizophrenia and other major neurological disorders. *Mol Cell Neurosci.* 2004;25:42–55.
67. Narayanan S, Arthanari H, Wolfe MS, Wagner G. Molecular characterization of disrupted in schizophrenia-1 risk variant S704C reveals the formation of altered oligomeric assembly. *J Biol Chem.* 2011;286:44266–76.
68. Lelieveld SR, Hendriks P, Michel M, Sajani G, Bader V, Trossbach S, et al. Oligomer assembly of the C-terminal DISC1 domain (640–854) is controlled by self-association motifs and disease-associated polymorphism S704C. *Biochemistry.* 2009;48:7746–55.
69. Kamiya A, Kubo K, Tomoda T, Takaki M, Youn R, Ozeki Y, et al. A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nat Cell Biol.* 2005;7:1167–78.
70. Ozeki Y, Tomoda T, Kleiderlein J, Kamiya A, Bord L, Fujii K, et al. Disrupted-in-Schizophrenia-1 (DISC-1): mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. *Proc Natl Acad Sci USA.* 2003;100:289–94.
71. Willbold D, Strodel B, Schroder GF, Hoyer W, Heise H. Amyloid-type protein aggregation and prion-like properties of amyloids. *Chem Rev.* 2021;121:8285–307.
72. Iadanza MG, Jackson MP, Hewitt EW, Ranson NA, Radford SE. A new era for understanding amyloid structures and disease. *Nat Rev Mol Cell Biol.* 2018;19:755–73.
73. Ke PC, Zhou R, Serpell LC, Riek R, Knowles TPJ, Lashuel HA, et al. Half a century of amyloids: past, present and future. *Chem Soc Rev.* 2020;49:5473–509.

74. Schurov IL, Handford EJ, Brandon NJ, Whiting PJ. Expression of disrupted in schizophrenia 1 (DISC1) protein in the adult and developing mouse brain indicates its role in neurodevelopment. *Mol Psychiatry*. 2004;9:1100–10.
75. Teng S, Thomson PA, McCarthy S, Kramer M, Muller S, Lihm J, et al. Rare disruptive variants in the DISC1 interactome and regulome: association with cognitive ability and schizophrenia. *Mol Psychiatry*. 2018;23:1270–7.

ACKNOWLEDGEMENTS

We thank Dr. Bernd König and Mr. Ian Gering for discussions pertaining to ITC measurements; Dr. Rudolf Hartmann and Mr. Kevin Bochinsky for assisting with the solution NMR experiments; Dr. Antony Yerabham (NIH, USA) for cloning the MBP fusion construct of the DISC1 C-region; Dr. Nicholas Bradshaw (Rijeka University, Croatia) and Prof. Dr. Carsten Korth (Heinrich Heine University Düsseldorf, Germany) for plasmids containing the genes for DISC1 and its variants, NDEL1 and LIS1. We extend our sincere thanks to Prof. Dr. Carsten Korth and Dr. Nicholas Bradshaw for critical reading of the paper.

AUTHOR CONTRIBUTIONS

AC, DW and OHW designed the experiments and drafted the paper. AC prepared all the samples and performed experiments using DLS, CD, SPR, ITC, and NMR measurements. ssNMR experiments were performed by AC alongside NB, NL, and HH. EM measurements were performed by MZ, BF, and GS. All authors critically reviewed the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-021-01765-1>.

Correspondence and requests for materials should be addressed to Abhishek Cukkemane, Dieter Willbold or Oliver H. Weiergräber.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

SUPPLEMENTARY INFORMATION

Conformational heterogeneity coupled with β -fibril formation of a scaffold protein involved in chronic mental illnesses

Abhishek Cukkemane^{1,3*}, Nina Becker^{1,2,3}, Mara Zielinski¹, Benedikt Frieg¹, Nils-Alexander Lakomek^{1,2,3}, Henrike Heise^{1,2,3}, Gunnar F. Schröder^{1,2,4}, Dieter Willbold^{1,2,3*} and Oliver H. Weiergräber^{1,2*}

¹Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany.

²Jülich Centre for Structural Biology (JuStruct), Forschungszentrum Jülich, Jülich, Germany.

³Institut für Physikalische Biologie, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

⁴Physics Department, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

*email – a.cukkemane@fz-juelich.de; d.willbold@fz-juelich.de; o.h.weiergraeber@fz-juelich.de

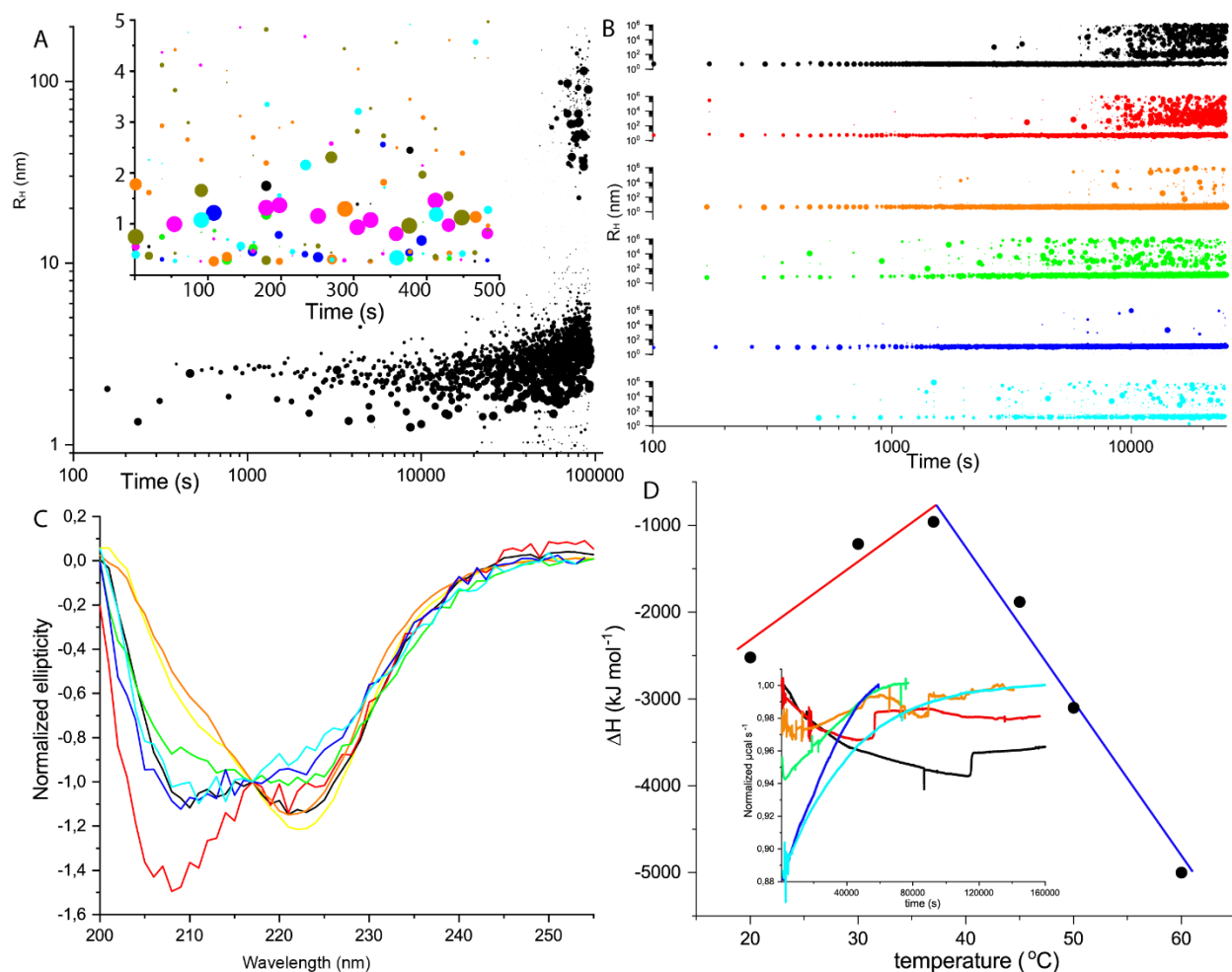
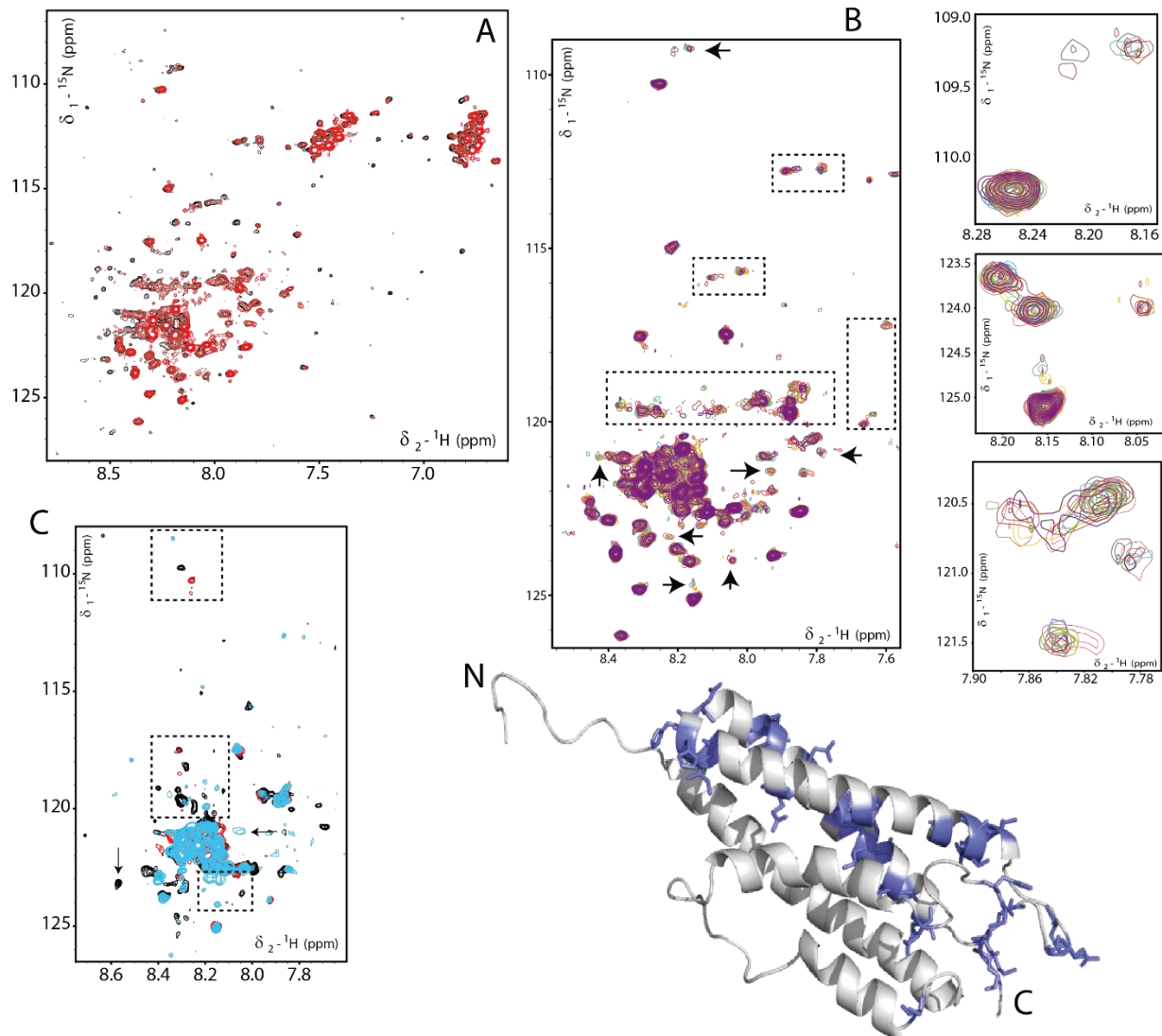


Figure S1: (A) DLS measurement of a freshly prepared His₆-tagged DISC1 C-region sample at 20 °C. During the initial phase of the measurement, the sample has a R_H of 2.94 nm and aggrandizes to ~100 nm towards the end of the measurement. The inset figure illustrates determination of the minimum detectable concentration of the His₆-tagged version at 20 °C, using various concentrations: 15 μ M (olive green), 10 μ M (orange), 7.5 μ M (pink), 5 μ M (cyan), 2.5 μ M (blue), 1.25 μ M (green) and 0.75 μ M (black). To study the oligomerization propensity of the MBP fusion construct of the DISC1 C-region, we performed (B) DLS and (C) CD measurements to follow changes in SSE at various temperatures, 20 °C (black), 30 °C (red), 37 °C (orange), 45 °C (green), 50 °C (blue) and 60 °C (cyan). One additional spectrum (yellow) in the CD profile depicts the freshly prepared sample. (D) Thermograms (inset) of 10 μ M protein solutions recorded at 20 °C (black), 30 °C (red), 37 °C (orange), 45 °C (green), 50 °C (blue) and 60 °C (cyan), with corresponding enthalpy changes plotted in the main graph. ΔC_p values of 94.3 and -179.8 kJ mol⁻¹ K⁻¹ were derived from the slopes of the two segments using a linear fit.



29

30 **Figure S2:** (A) Solution NMR spectroscopy measurements of C-region DISC1 at protein
 31 concentrations, 400 μM (black) and 60 μM (red) describes spectra of oligomerized and/or unfolded
 32 protein; we expect to observe 166 backbone resonances for a monomeric folded protein. Spectra
 33 was processed with squared sine bell function of 2 (B) and (C) describes time dependent
 34 measurements at protein concentrations of 60 μM (sine bell function of 2) and 10 μM (processed
 35 using exponential line broadening factor of 10 Hz), respectively. The colour coding in the spectra
 36 depicts various time points in minutes; red (0-300), grey (300-600), green (600-900), yellow (900-
 37 1200), cyan (1200-1500), khaki (1500-1800), bronze (1800-2100) and purple (2100-2400).

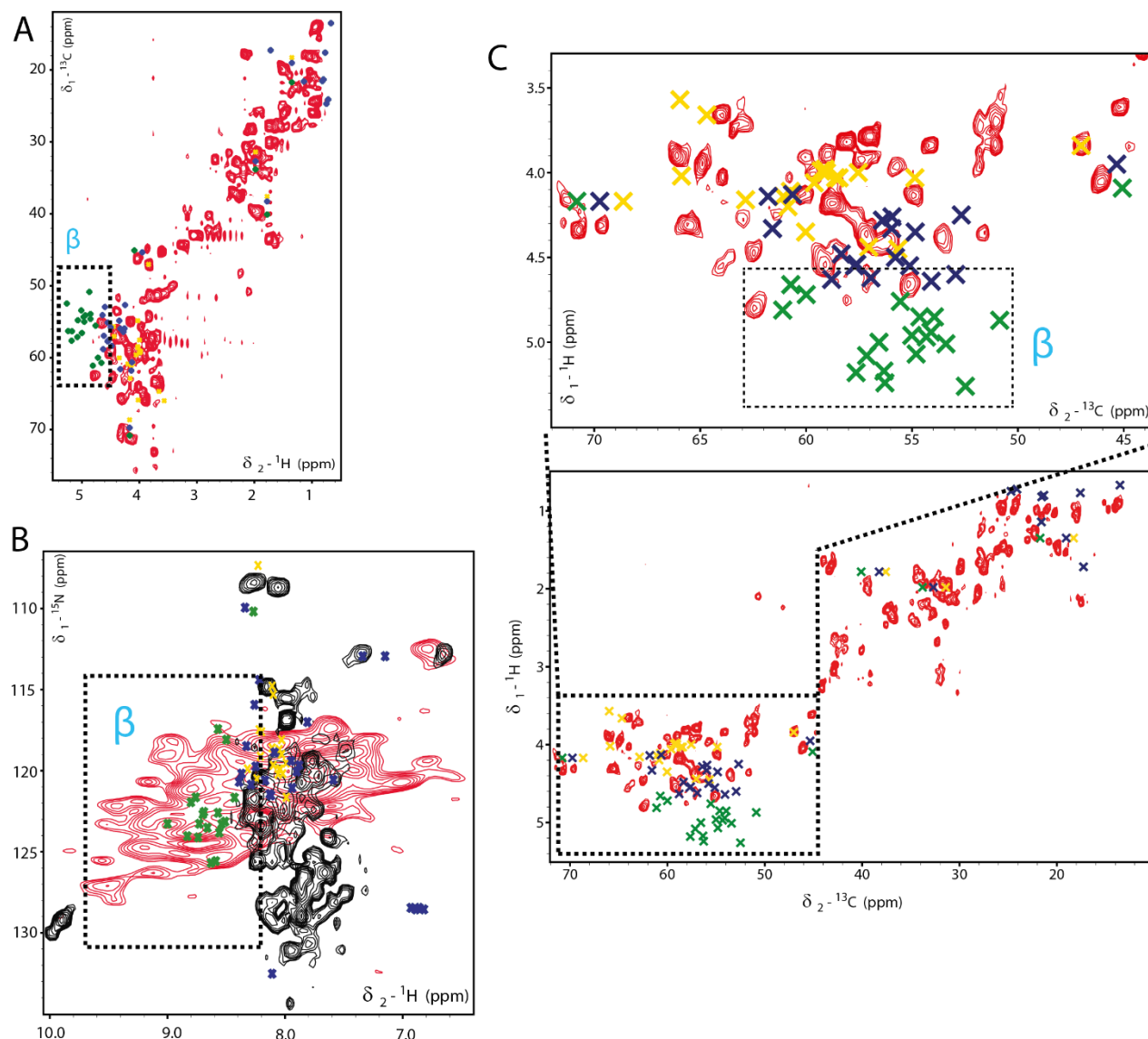
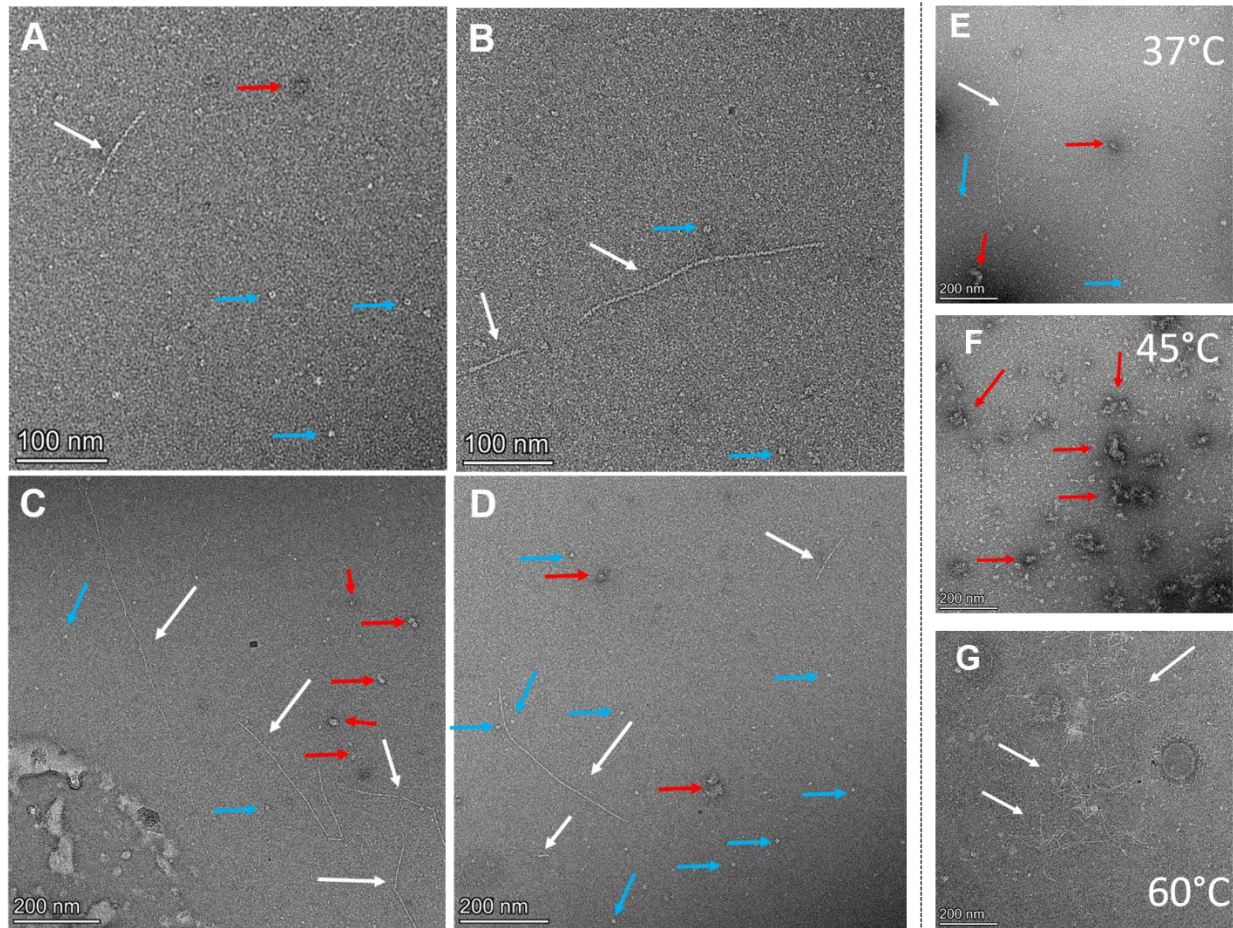


Figure S3: (A) 2D ^1H - ^{13}C -HSQC-MAS NMR spectrum of the solid sample; as in panels (B) and (C), annotations in yellow, blue and green represent average CS (Wang and Jardetzky, 2001 <https://doi.org/10.1110/ps.3180102>) of residues in helices, coiled regions (RC) and strands, respectively. The box depicts the region where CS of residues in β -strands are expected. (B) 2D ^1H - ^{15}N correlations on scalar (black) and dipolar (red) couplings, respectively at 30 kHz MAS displaying the predominance of CS arising from residues in β -strands in the rigid part of the protein. (C) 2D ^1H - ^{13}C -INEPT based correlation spectrum was applied to the DISC1 C-region protein. The top panel displays a zoom-in of the backbone ^1H - ^{13}C correlation; the boxed region with green cross-peaks are suggestive of expected signals from residues if present in β -sheets.



48

49 **Figure S4:** Electron micrographs (A-D) of the DISC1 C-region predominantly shows 2 major
 50 populations comprising of the fibrils (white arrows) and oligomers (blue arrows). One can observe
 51 fibrils at various lengths ranging from 100 nm up to 800-1000 nm. We also observed amorphous
 52 aggregates (red arrows) in low abundance at temperatures of 37 °C and lower. (D) Zoom-out image
 53 shown in Figure 2C. (E-G) Negatively stained electron micrographs of the DISC1 C-region at
 54 various temperatures. Arrows coloured in white, red and blue point towards fibrils, oligomers and
 55 amorphous aggregates, respectively.

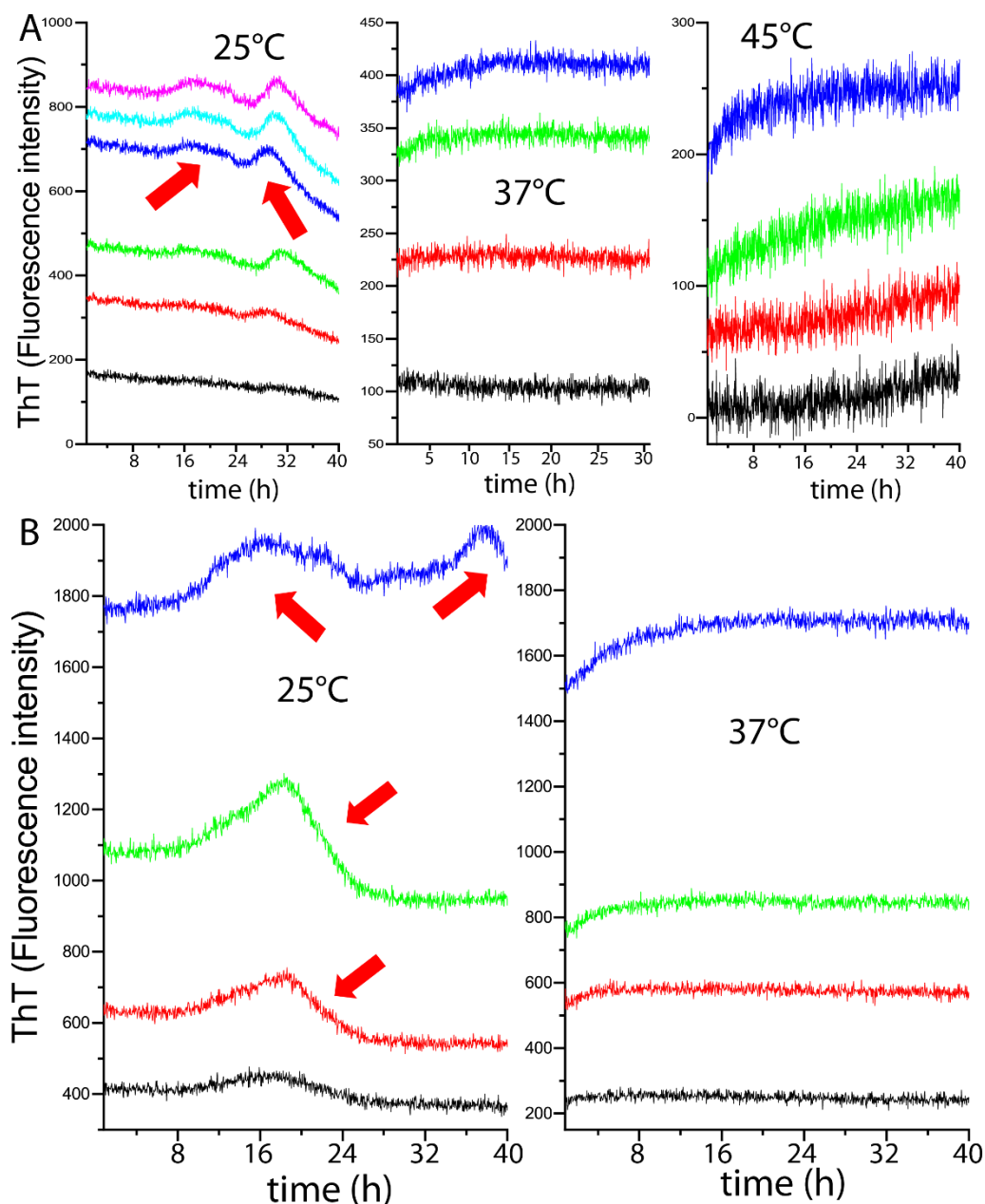
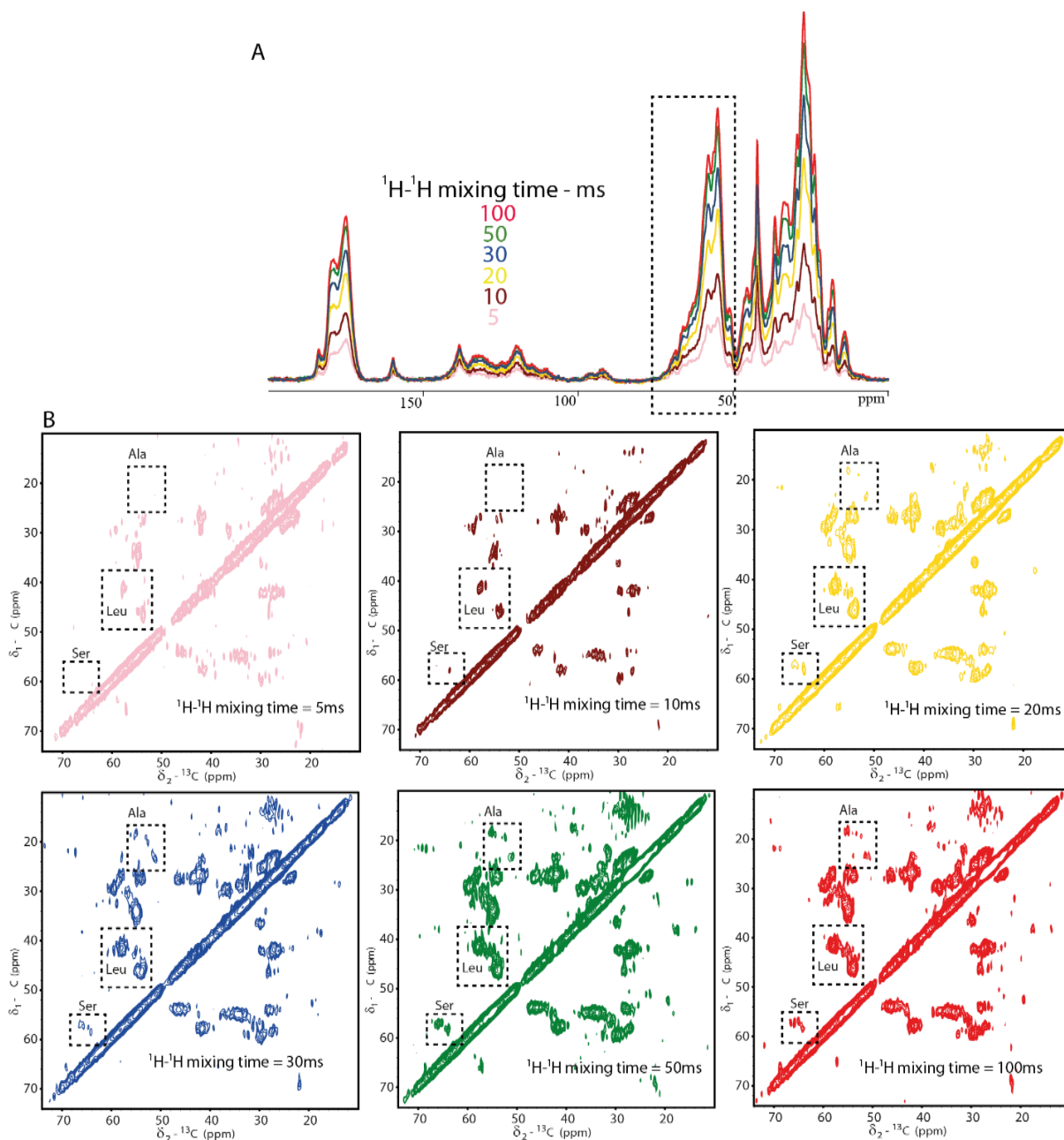


Figure S5: (A) Temperature dependent effect of ThT fluorescence intensity at various DISC1 C-region protein concentrations, 5 μM (black), 10 μM (red), 15 μM (green) and 20 μM (blue); 2 additional protein concentrations were tested at 25 μM (cyan) and 30 μM (pink).

(B) ThT fluorescence assay performed at 90 μM (blue), 45 μM (green), 22.5 μM (red) and 11.25 μM (black) at 25 $^{\circ}\text{C}$ and 37 $^{\circ}\text{C}$. Oscillations between high and low intensities is most likely due to aggregation of the larger fibrils sticking to the sides of the plate during the measurement window followed by growth of another fibrillar segment(s).



65

Figure S6: (A) 1D ^{13}C water build-up spectra with several with ^1H - ^1H mixing times. The signal intensity in the $\text{C}\alpha$ region (boxed) was plotted for the build-up curve depicted in Figure 3E. (B) 2D ^{13}C - ^{13}C correlation spectra with ^1H - ^1H mixing times of 5 ms (pink), 10 ms (maroon), 20 ms (yellow), 30 ms (blue), 50 ms (green) and 100 ms (red), the boxed regions correspond to the signals from residues Ala, Leu and Ser that were used for the amino acid specific water build-up analysis depicted in Figure 3E.

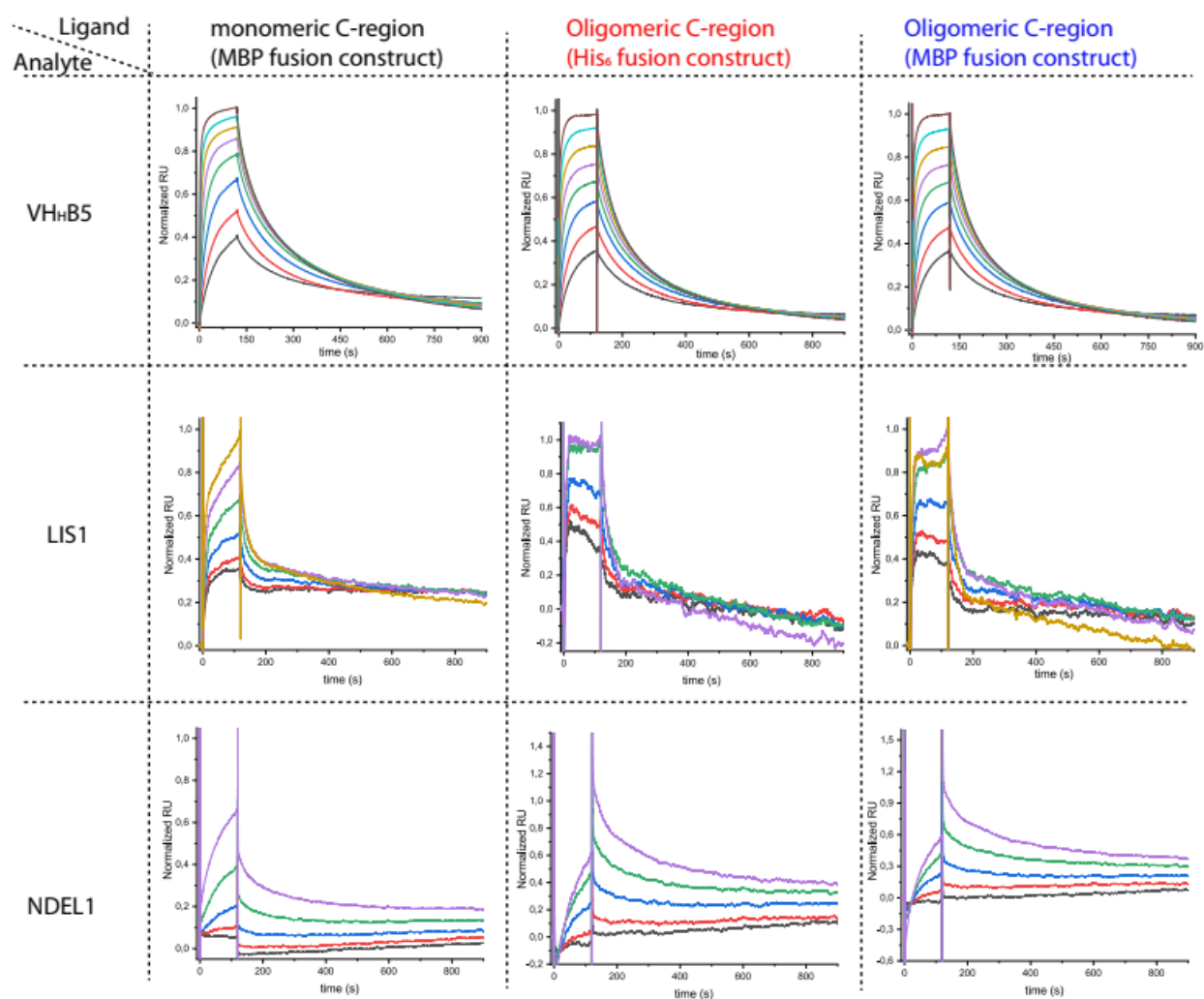
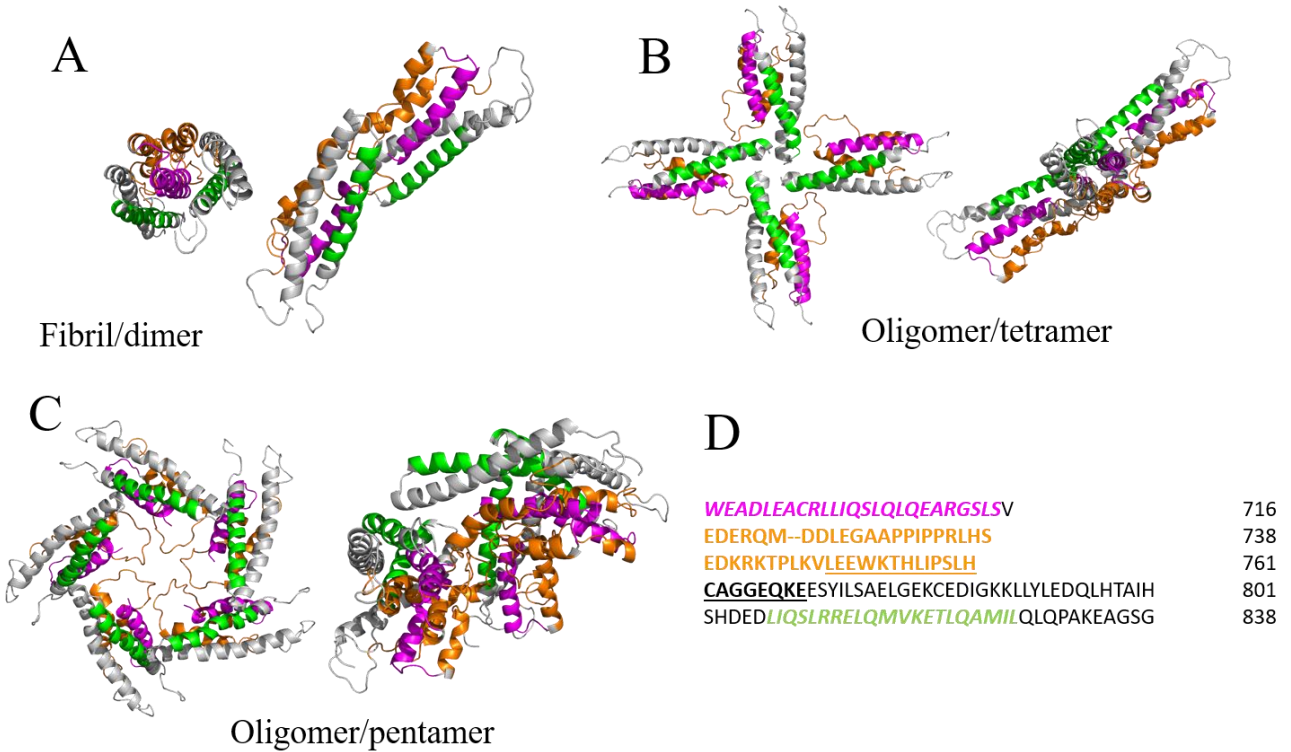


Figure S7: (A) SPR sensorgrams featuring binding of V_HH B5, LIS1 and NDEL1 with monomeric MBP-C-region, oligomeric His₆-C-region and oligomeric MBP-C-region.



76

77 **Figure S8:** Additional structural models (also see figure 4) of the DISC1 C-region, which were
 78 generated using CLUSPRO2 and agree well with the experimental data. Fibrils (A) and oligomers
 79 (B and C) are colour coded as shown in the protein sequence in (D). Colours represent V_HH B5
 80 binding in magenta; (pseudo)repeat sequence in orange; underlined part describes residues in Δ22
 81 and the green region represents the NDEL1 binding site.



Tracing the aggregation pathway of the scaffold protein DISC1: Structural implications for chronic mental illnesses

Abhishek Cukkemane^{a,b,*}, Nina Becker^{a,b}, Tatsiana Kupreichyk^{a,b}, Henrike Heise^{a,b}, Dieter Willbold^{a,b,*}, Oliver H. Weiergräber^{a,*}

^a Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany

^b Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany

ARTICLE INFO

Keywords:

Schizophrenia
Disrupted in schizophrenia 1
Aggregation pathway
Proteinopathy
Biophysical analysis
Drug development
Protein-protein interaction

ABSTRACT

Disrupted in schizophrenia 1 (DISC1) is a pleiotropic scaffold protein that is postulated to comprise large disordered regions and four distinct structured segments with a high proportion of helical or coiled-coil fold. DISC1 associates with over 300 proteins and is associated with several physiological roles ranging from mitosis to cellular differentiation. Yet, the structural features of the protein are poorly characterized. The C-terminal region (C-region, res. 691–836) forms a tetramer and can also aggregate into amyloid-like fibers, potentially linked to schizophrenia and other chronic mental illnesses. Using a combination of biophysical and structural biology applications, we investigate the structural heterogeneity of three mutants of the C-region, viz., the S713E, S704C and L807-frameshift mutants. We provide evidence for the plasticity of the C region; a thin border separates the conformational flexibility of DISC1 required for interaction with a myriad of partners from disruptive aggregation. Snapshots of aggregates and fibrils growing from a nucleus are presented, along with data supporting the role of the minimal fibrillizing element in the C-region, the β -core. This segment also houses a stretch of residues that is critical for the binding of NDEL1 proteins in the mitotic spindle complex and is absent in the non-binding splice variant DISC1 Δ 22aa. Physiologically, both the splice variant and the fibers represent loss-of-function states that disrupt cellular division. Our findings highlight the need to decipher the structural elements within the DISC1 C-region to comprehend its physiological role and aggregation-related anomalies, and to establish a rationale for drug development.

Introduction

Schizophrenia, major depressive disorder (MDD), bipolar disorder (BD) and autism spectrum disorder represent closely associated neurodevelopmental and chronic mental illnesses (CMI). The aetiological factors that contribute to these disorders remain poorly understood as they involve an interplay of multiple factors, including biological, environmental, and social conditions. A major biological risk factor that was identified about two decades ago in a Scottish family with several severe psychiatric disorders was disrupted in schizophrenia 1 (DISC1) (Blackwood et al., 2001; Millar et al., 2000). DISC1 is a multi-functional hub that regulates the activities of over 300 different enzymes and proteins, including molecules of clinical and therapeutic relevance (Yerabham et al., 2013; Millar et al., 2003). Considering its regulatory significance, DISC1 is involved in a myriad of physiological roles across various cellular functions such as mitosis, proliferation, neuronal

development, and synaptogenesis (Bradshaw and Korth, 2019; Tropea et al., 2018; Hikida et al., 2012).

In a recent study, the C-region (Cukkemane et al., 2021) was demonstrated to assemble into a tetramer that represents the functional unit, which associates with Nuclear Distribution Element 1 (NDEL1, formerly known as NudE), Nuclear Distribution Element Like 1 (NDEL1, formerly known as Nudel) and platelet activating factor acetylhydrolase 1b regulatory subunit 1 (PAFAH1B1, formerly LIS1) in a cooperative manner in the mitotic spindle complex. Dysfunctional DISC1 aggregates into amyloid-like fibrils that are stained by the thioflavin T (ThT) dye. Using bioinformatics analysis, we identified a pseudo-repeat sequence (res. 717–761) that may represent the scaffold, the β -core of the fibril (Cukkemane et al., 2021). This region encompasses a stretch of amino acids that will herein be referred to as the Δ 22 region (748–769). The Δ 22 region is absent in the splice variant DISC1 Δ 22aa (Taylor et al., 2003; Kamiya et al., 2006) (also known as Lv-DISC1 (Nakata et al., 2009),

* Corresponding authors at: Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany.

E-mail addresses: a.cukkemane@fz-juelich.de (A. Cukkemane), d.willbold@fz-juelich.de (D. Willbold), o.h.weiergraeber@fz-juelich.de (O.H. Weiergräber).

<https://doi.org/10.1016/j.jysbx.2025.100128>

Received 15 April 2025; Received in revised form 23 May 2025; Accepted 23 May 2025

Available online 24 May 2025

2590-1524/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

which renders it unable to bind to NDEL1 and PAFAH1B1 proteins, resulting in defective neurite outgrowth in PC12 cells.

Further progress towards understanding the role of the C-region of DISC1 (Fig. 1A and 1B) and the significance of the $\Delta 22$ region requires delineating the structure–function relationships of the protein from a pathological perspective. Therefore, it is important to understand how mutations within the C-region contribute to the pathophysiology of the various disorders associated with DISC1 dysfunction. In this context, three different C-region variants were selected, *i.e.*, S704C, S713E and the L807-frameshift (L807-FS). The S704C (Leliveld et al., 2009; Leliveld et al., 2008; Narayanan et al., 2011; Hashimoto et al., 2006) mutant forms aggregate deposits in 25 % of post-mortem brain samples from subjects suffering from schizophrenia, BD and MDD. The S713E mutation (Ishizuka et al., 2011) is responsible for Bardet–Biedl syndrome; the Ser residue serves the role of a phosphorylation switch that promotes the transition of neuronal progenitor cells from proliferation to migration during corticogenesis. Finally, the frameshift mutation affecting residues downstream of L807 was first identified in an American family suffering from schizophrenia and schizoaffective disorders and gives rise to aggregated complexes (Sachs et al., 2005).

To comprehend the roles of these residues in DISC1 pathophysiology, it is important to view the function of the protein from a structural perspective. NMR structures of an N-terminally truncated version of the murine C-region (Ye et al., 2017; Wang et al., 2019) highlighted that this region was largely unstructured but adopted helical motifs containing coiled-coil (CC) structure in the presence of a ligand. A typical CC domain is characterized by the presence of heptad repeats (*abcdefg*) in which positions *a* and *d* are occupied by hydrophobic residues and the remaining sites by either polar or charged ones. In these structures, residue L807 is part of the coiled-coil domain. An ab-initio model (Yerabham et al., 2018) of the C-region that is consistent with SAXS data positioned S704 as polar entity in the coiled-coil domain, which also flanks the β -core and S713 present in between two helical coils. These serine residues do not contribute to the oligomeric interface (Cukkemane et al., 2021) and both are accessible to interact with physiological partners. This naturally poses several interesting

questions. Firstly, how exactly do mutations within a CC heptad sequence that replace a polar side chain with a small apolar or a charged one, specifically S704C and S713E, affect protein stability? Secondly, does the β -core that houses part of the $\Delta 22$ region represent the scaffold of the DISC1 C-region fibrils?

To address these questions and comprehend the impact of point mutations in DISC1 on the pathophysiology of schizophrenia and related CMI, we employed a combination of biophysical techniques and structural biology applications. Here, we demonstrate that the β -core contains determinants for the aggregation of the C-region. The isolated β -core peptide and the DISC1 C-region mutants form aggregates that may serve as nuclei that are oily droplet-like. These findings are complemented with ThT fluorescence data and thermodynamic parameters of the aggregates formed by the C-region variants. To conclude, we show the individual steps involved in the complexation of the DISC1 C-region into supramolecular entities. These findings extend our previous report (Cukkemane et al., 2021) on the significance of the aggrandization of the DISC1 C-region into fibrillar assemblies from a clinical and pathophysiological perspective and provide insights into therapeutic strategies.

Experimental procedures

Protein expression and purification

The DISC1 C-region (WT), C-region fused to maltose-binding protein (MBP-C-region), S704C, S713E, L807-FS and β -core (717–761) proteins were expressed as His₆ fusion constructs, as described previously (Cukkemane et al., 2021; Yerabham et al., 2018), using the vector pESPRIT002 (Yumerefendi et al., 2010) for transforming *E. coli* BL21 (DE3) pLysE T1R cells. The residue sequences for the constructs are detailed in Fig. S1. The culture was grown either in Luria-Bertani (LB) broth or M9 minimal medium. Protein expression was induced at an OD₆₀₀ of 0.6 by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and culturing was continued for 16 h at 18 °C. Harvested cells were lysed in Tris-buffered saline (TBS) containing 10 mM Tris-HCl pH 7.4, 150 mM NaCl and Complete EDTA-

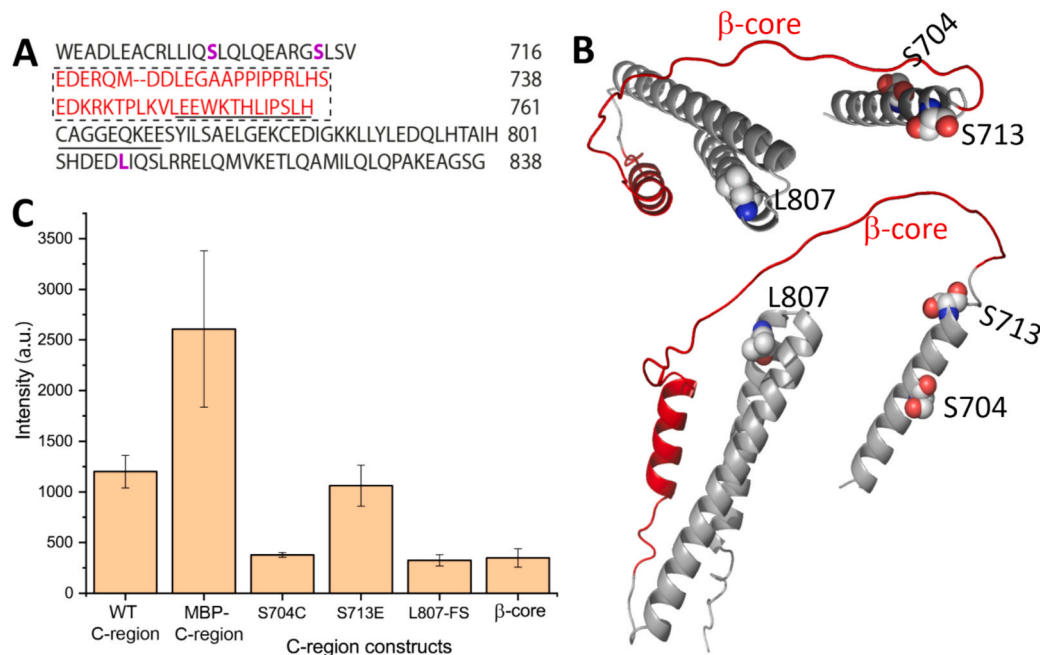


Fig. 1. (A) The protein sequence of the DISC1 C-region numbered in the context of the full-length protein, with the pseudo-repeat sequence constituting the β -core (717–761) shown in the dotted box in red; the underlined part corresponds to residues in the $\Delta 22$ region (748–769); residues S704, S713, and L807 highlighted in bold pink. (B) Representation of the DISC1 monomeric C-region model generated using AlphaFold2, featuring two different views. (C) ThT fluorescence assay of the DISC1 C-region, its mutant variants, and the β -core at 25 °C. Experiments for the different proteins (*n* = 3) were corrected for background by subtracting the buffer control. Bars illustrate mean $\pm$ SD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

free protease inhibitor cocktail (Roche) using an ice-chilled microfluidizer M100P (Microfluidics MPT) at 15,000 PSI. The insoluble fraction was removed by centrifugation at $50,000 \times g$ for 25 min at 4 °C. The soluble fraction was loaded onto Co²⁺-loaded NTA resin (Qiagen) and the target protein eluted using TBS containing 500 mM imidazole. The His₆-C-region protein was further purified using a HiLoad 16/60 Superdex 200 size exclusion chromatography column (GE Healthcare Bio-Sciences).

Thioflavin-T (ThT) fluorescence assay

The Thioflavin T (ThT) assay was conducted by preparing the reaction mixture immediately before measurement 25 °C on black non-binding 96-well plates (Sigma-Aldrich) with the reaction mixture containing 10 μ M ThT and 10 μ M protein in a total volume of 100 μ L. Experiments were performed in triplicate and each reaction was background-corrected by subtraction of the buffer control. Fluorescence was monitored at regular intervals of 3–5 min using a Fluostar microplate reader (BMG Labtech, Offenburg, Germany) with 440 nm excitation and 492 nm emission filters in bottom-read mode. The signal intensity is depicted as an average value in the form of mean $\pm$ SD.

Dynamic light scattering (DLS)

Measurements of the protein samples were performed using a SpectroSize 300 (XtalConcepts GmbH) instrument with a sample volume of 500 μ L at 20 °C. Prior to measurements, all samples were centrifuged at $21,000 \times g$ for 30 min at 4 °C. Diffusion coefficients were obtained from analysis of the decay of the scattered intensity autocorrelation function and were used to determine apparent hydrodynamic radii (R_H) via the Stokes-Einstein equation, as implemented in the instrument software (SpectroCrystal). Moreover, we determined the detection limit of the DLS instrument as a proxy to the minimal (“critical”) concentration of building blocks supporting fibril growth (see below).

Isothermal titration calorimetry (ITC)

Measurements were performed using an iTC-200 (MicroCal) calorimeter with 10 μ M protein in a 200 μ L sample cell. Changes in the heat flow were monitored in real time with the reaction cell stirred at 300 rpm and reference power of the cell set to 5 μ cal/s. To compare the thermograms for the different protein, the data set was normalized by dividing each value by the maximum value in the dataset. Furthermore, thermodynamic parameters were computed on the basis of the Goto-Kardos scheme (Kardos et al., 2004; Ikenoue et al., 2014). Assuming that the observed heat exchange represents the enthalpy change of self-association of the DISC1 C-region, ΔH at 20 °C was calculated by integrating the respective peak area. As a framework for describing the thermodynamics of fibrillization, we have previously employed (Cukkemane et al., 2021) a simple 1D “crystallization” model of amyloid formation (Kardos et al., 2004; Ikenoue et al., 2014; Jarrett and Lansbury, 1993); we assume that, upon completion of the process, the protein suspension contains fibrils at equilibrium with residual monomers/oligomers as represented by the “critical concentration” $[C]$, which, as a first approximation, equals the reciprocal of the equilibrium association constant K_F . Based on the critical concentration estimated via DLS (Cukkemane et al., 2021; Kardos et al., 2004; Ikenoue et al., 2014; Jarrett and Lansbury, 1993), we obtained the apparent free energy change ($\Delta G_{app} = -RT \ln K_F = RT \ln [C]$, where R and T are the gas constant and the absolute temperature, respectively). The entropy change ΔS was calculated using the values of ΔH and ΔG_{app} using the Gibbs-Helmholtz equation ($\Delta G_{app} = \Delta H - T\Delta S$).

Solid-state NMR experiments (ssNMR)

SsNMR experiments were performed on uniformly ¹³C/¹⁵N- double-

labelled samples of the DISC1 C-region and its variants, which were dissolved in TBS buffer containing 1 mM Na₂-EDTA and 0.01 % (w/v) NaN₃, and concentrated in a centrifugal device with a 3 kDa cut-off membrane (Amicon, Millipore). The concentrated sample was bath-sonicated for 5 min and incubated at 20 °C for 2 weeks to allow for aggregation. The sample was packed into 3.2 mm zirconia magic angle spinning (MAS) rotors (Bruker Biospin). Measurements were conducted at a static magnetic field strength of 14.1 T (corresponding to a ¹H Larmor frequency of 600 MHz) using 3.2 mm triple-resonance (¹H, ¹³C, ¹⁵N) Bruker MAS ssNMR probes (Bruker Biospin) at a 12.5 kHz MAS rate. The temperature of the VT gas was set to 256 K, resulting in a sample temperature of 266 ± 5 K. Water-edited 1D build-up experiments were recorded using a ¹H T_2 filter of 2.5 ms to suppress signals from the rigid regions of the protein. For ¹H-¹H spin diffusion, mixing times (t_m) of 2–500 ms were used to permit spin diffusion from water to the protein. ¹H decoupling was applied during evolution and detection periods using the SPINAL64 (Fung et al., 2000) scheme at a radio-frequency of 83 kHz.

While semi-quantitative, water-edited ssNMR studies provide a reasonable estimate of the molecular dimensions of water accessible areas of membrane proteins (Ader et al., 2009; Luo and Hong, 2010) and amyloid fibrils (Schneider et al., 2011). We implemented a similar approach to calculate and compare the dimensions of WT-C-region fibrils (Cukkemane et al., 2021). Briefly (details in SI methods), the spectral region of 50–75 ppm was integrated and normalized to the maximum signal intensity, which was plotted against the mixing time. In the absence of any saturation effects, the slope describes the time required to reach 100 % magnetization transfer (Ader et al., 2009; Schneider et al., 2011; Schmidt-Rohr et al., 1994). A linear fit to the initial build-up rate (t_m^s) was used to determine the water accessibility of the sample and calculate the molecular dimensions of the fibril. All data were processed and analyzed using Topspin 4.0.6 (Bruker Inc).

Atomic force microscopy (AFM)

For sample preparation, 5 μ L of DISC1 WT C-region, the mutants, or the β -core at concentrations ranging from 1 to 50 μ M was applied onto a freshly cleaved muscovite mica surface. Samples were incubated for 10 min under a humid atmosphere, followed by washing thrice with Milli-Q water (100 μ L) and finally drying under a stream of N₂ gas. Images were recorded in intermittent contact mode using a JPK NanoWizard 3 atomic force microscope equipped with a silicon cantilever and tip (OMCL-AC160TS-R3, Olympus). The tip has an average radius of 9 ± 2 nm, a force constant of 26 /Nm and a resonant frequency of approx. 300 kHz. The images were processed using JPK Data Processing Software (version spm-5.0.84). The height profiles in the figures were obtained from measured heights using a polynomial fit that was subtracted from each scan line first independently and then using limited data range.

Results

Sensitivity of ThT dye to the C-region mutants

The enhanced and red-shifted fluorescence of ThT upon binding to structures enriched in cross β -sheet is used in standard assays to study amyloidogenic proteins. In the case of human DISC1, one study investigating the recombinant full-length protein noted increases ThT fluorescence, indicating formation of cross β -fibrils (Tanaka et al., 2017). We reported previously (Cukkemane et al., 2021) that fibrils formed by the isolated C-region of the DISC1 protein bind ThT and share structural similarities with other well-characterized neurological proteinopathy markers. However, the aggregates isolated from brain tissue of deceased subjects suffering from schizophrenia, BD and MDD do not stimulate ThT fluorescence (Leliveld et al., 2009; Korth, 2012). To resolve these apparent discrepancies, we investigated the effects of clinically identified alterations S704C, S713E, and L807-FS in ThT fluorescence assays

of the DISC1 C-region (Fig. 1C). The S713E protein displays ThT fluorescence similar to the WT protein. However, the S704C and L807-FS mutants yielded lower signal intensities. According to our hypothesis, the β -core region represents the lynchpin motif for amyloid-like fibrillization of the DISC1 C-region (Cukkemane et al., 2021). Nonetheless, ThT-reactivity of the β -core construct was minimal with a fluorescence intensity resembling that observed for S704C and L807-FS.

This raises the important question, which structural features within the protein provide the interface for ThT binding and to what extent these traits are modified in the variants investigated. To address these questions, we used a combination of biophysical and structural biology techniques, including DLS, and AFM to comprehend the behavior of the proteins in solution.

DISC1 C-region aggregates seed fibrillar growth

As reported previously (Cukkemane et al., 2021), the WT DISC1 C-region (Fig. 2A) self-associates into fibrils at 20 °C. Notably, the protein adopts additional multimeric states such as amorphous aggregates at 45 °C and shorter clumped fibrils at 50–60 °C, which do not display ThT fluorescence. In the DLS experiments with C-region variants, at a protein concentration of 10 μ M, the major populations of S704C and S713E (Fig. 2B and 2C) showed apparent R_H values of 2.8 ± 0.4 nm and 2.9 ± 0.6 nm, respectively, similar to the WT C-region monomer (2.6 ± 0.5 nm) (Cukkemane et al., 2021). In a couple of situations, the S713E variant did form larger species ($R_H = 6.3 \pm 1.2$ nm, magenta trace in Fig. 2C). The third C-region variant, L807-FS, on average displayed a particle radius of 8.9 ± 1.9 nm (Fig. 2B). This is in stark contrast with the behavior of the WT C-region protein, which initially presented as a tetramer ($R_H = 4.25 \pm 0.25$ nm), but quantitatively transformed into μ m range aggregates in the course of the experiment (Fig. 2A), in accordance with previous observations (Cukkemane et al., 2021). Unlike the

WT C-region, the mutant variants were present in different states of oligomerization/aggregation but did not show an obvious tendency to convert into larger species in a sigmoidal fashion.

To complement the above findings, we performed a series of AFM (Fig. 3, S2 and S3) measurements to characterize the architecture of the different species observed in the DLS experiments. At protein concentrations ≥ 10 μ M (Fig. S2 and S3), we consistently observed the presence of numerous amorphous aggregates. Serendipitously, for the 3 μ M MBP-C-region protein, we captured an AFM image of a seemingly amorphous aggregate from which fibrils of various lengths were radiating out (Fig. 3A). This raises an interesting prospect: Are the amorphous aggregates the seeds for fibrillar growth? Does a higher concentration of the protein (≥ 10 μ M) promote amorphous aggregate formation to an extent that does not leave sufficient soluble protein for fibrillar growth? To systematically investigate this effect, we repeated the experiment with a protein concentration of 1 μ M (Fig. 3B–E). However, unlike the WT-C-region, in this case we did not observe any fibril formation. Interestingly, in these cases, we observed fewer but larger aggregates than the ones detected with protein concentrations above 10 μ M (Fig. S2). This indeed indicates that at higher protein concentration, the C-region variants undergo nucleation at an enhanced rate, resulting in the formation of a large number of smaller aggregates because of a rapidly decreasing concentration of free protein, while at lower protein concentrations, one observes slow nucleation yielding larger aggregates. While we do not have enough evidence at the moment, we would like to speculate that these larger aggregates promote the growth of longer fibrillar structures (Cukkemane et al., 2021); Fig. 3A) either by directly feeding the growing fibrils or due to the higher concentration of free protein. Our attempts to perform DLS experiments with protein concentrations < 3 μ M (Fig. S4) did not yield reproducible results because of finite instrument sensitivity.

Next, reverting to one of our major hypotheses, does the β -core

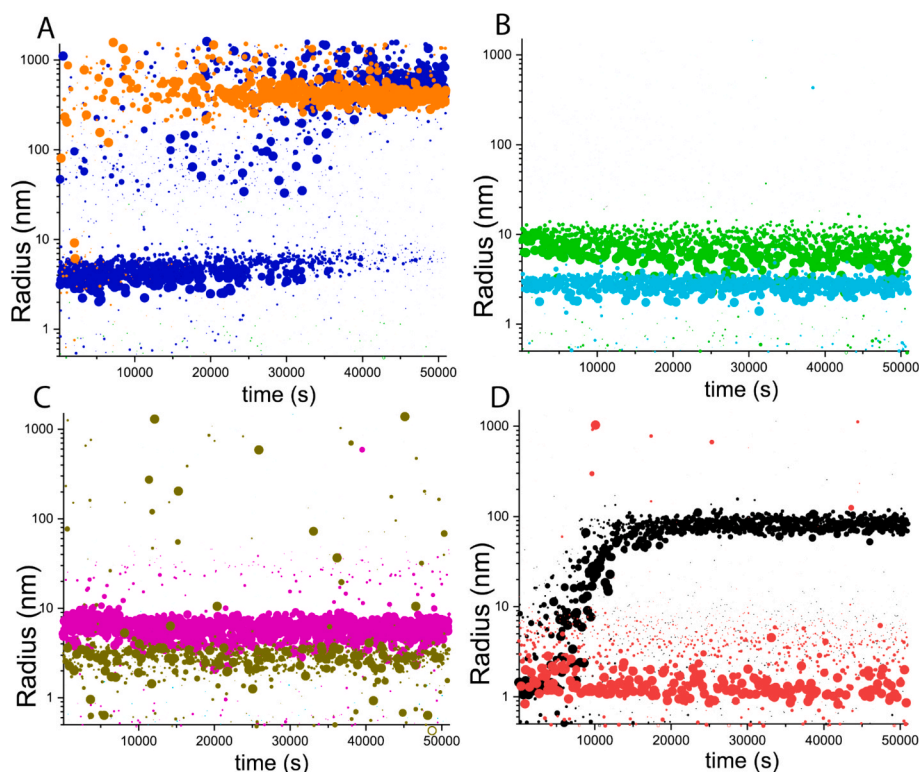


Fig. 2. Intensity-weighted DLS isotherms of the DISC1 C-region, mutant variants, and the β -core at 20 °C. (A) MBP-DISC1-C-region (orange) and the remaining constructs carried the his-tag namely WT-C-region (blue); (B) S704C (cyan) and L807-FS (green); (C) two preparations of S713E (magenta and olive green); (D) two samples of the β -core under identical conditions but from two different batches. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

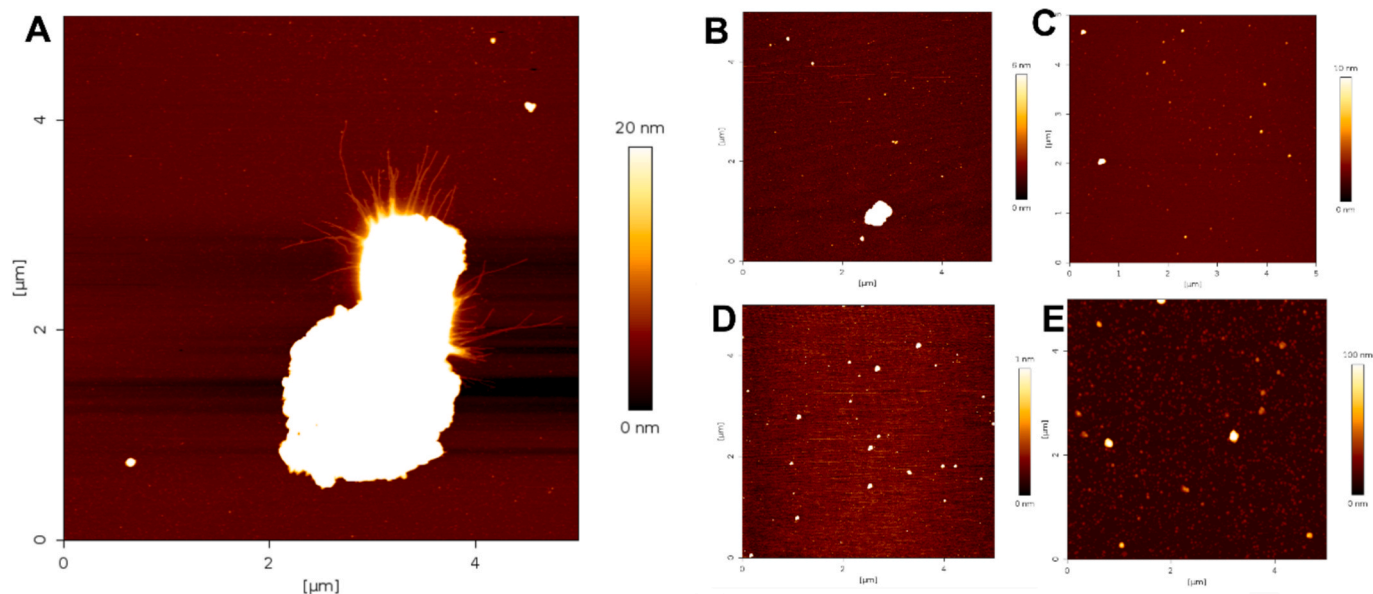


Fig. 3. Structural characterization of the DISC1 C-region using AFM (also see Fig. S2 and S3). (A) AFM image of fibrillary outgrowth from an aggregate of MBP-C-region. AFM images of the His-tagged DISC1 C-region variants S713E (B), S704C (C), β -core (D) and L807-FS (E) show the presence of aggregates. Protein concentrations were 3 μ M (panel A) and 1 μ M (panels B-E).

represent the scaffold of DISC1 fibrils? In DLS recordings, we observe a sigmoidal aggrandization of the motif. However, unlike the C-region protein that consistently formed larger species at 20 °C, β -core aggregated in $\sim$ 25 % of trials only (Fig. 2D), and the observed aggregates resembled those of the other C-region variants in AFM (Fig. 3D) as opposed to the μ m long fibrils observed for the WT- and MBP-C-region (Cukkemane et al., 2021; Fig. 3A). Together, these findings suggest that the β -core has a tendency to form larger aggregates at lower protein concentrations than the WT-C-region. Based on the DLS experiments we can surmise that it also contains the minimal building block, and that factors like protein concentration and temperature are critical for nucleation and fibril growth. Finally, to sum up our findings on the self-association of the DISC1 C-region variants, we observed the presence of aggregates that were rather heterogeneous in size.

Thermodynamic underpinning of C-region and variants fibrillization

We were interested in understanding why the mutant proteins and the β -core at 20 °C displayed characteristics that were apparently different from those of the WT-C-region. Thus, we investigated the thermodynamics of aggregation and related our findings to the structure of the assembly. For the WT-C-region, we reported previously (Cukkemane et al., 2021) that fibrils formed at lower temperatures (20–37 °C) were considerably longer than those cultivated at higher temperatures (45–60 °C), and that fibril morphology was related to reaction enthalpy. Intriguingly, the C-region variants investigated at a single temperature were found to divide into two categories simulating the two thermal regimes of the WT protein. Specifically, the S713E and S704C mutants showed an exothermic burst ΔH of -2144 kJ/mol and -1004 kJ/mol, respectively, in the course of self-association at 20 °C (Table 1, Fig. 4A, and 4B), similar to the WT-C-region at lower temperatures (ΔH of -2289 kJ/mol at 20 °C and -1103 kJ/mol at 30 °C; Table 2, and Fig. 4B). In contrast, the β -core and L807-FS revealed a more profound exothermic aggregation reaction with ΔH of -3528 kJ/mol and -4444 kJ/mol, respectively (Table 1), resembling those observed for the WT-C-region at elevated temperatures (ΔH of -2761 kJ/mol, -6092 kJ/mol and -8657 kJ/mol at 45, 50 and 60 °C; Fig. 4B, and 4C, Table 2 (Cukkemane et al., 2021)). However, comparing the enthalpic changes and structural features observed during WT-C-region aggregation at different temperatures with the data obtained for the

Table 1

Thermodynamic properties of the self-association process of the DISC1 C-region mutants and the β -core. The value of ΔG_{app} at 37 °C represents an estimate determined for the WT protein.

C-region variant	ΔH [kJ/mol]	ΔG_{app} [kJ/mol]
S704C	-1004	-34
S713E	-2144	
L807-FS	-4444	
β -core	-3528	
S704C*	-3635	

*One week old sample.

mutant variants, we observed intriguing similarities: pronounced exothermic reactions correlate with the formation of aggregates but not with the development of long fibrils.

Based on the estimated critical concentration of 0.75 μ M at 20 °C (Cukkemane et al., 2021), we calculated the ΔG_{app} for fibrillization of the WT-C-region to be -34 kJ/mol. Assuming the overall free energy change is not altered dramatically in the variants investigated in this study, the widely varying reaction enthalpies determined experimentally would still translate into negative entropy changes (between -3.5 and -15 kJ/mol/K) indicating the assembly of highly ordered structures possibly incorporating ordered water molecules. We performed water-edited ssNMR experiments to investigate the mechanism of fibrillization in a qualitative fashion and to understand the molecular structure of the aggregates. In such a study, polarization from water is transferred from the interfacial region to the protein, where a short transfer time (t_m^s) is indicative of high degree of water accessibility for a large part of the protein.

On performing the 1D water build-up experiment (Fig. S5 and Fig. 4D) and analyzing signals in the region of 50–75 ppm that represent primarily backbone α nuclei, we observed tightly bound water, with average initial t_m^s values of 31.14, 39.21, 62.50 and 34.36 ms (Table S1) for S704C, S713E, L807-FS, and β -core variants, respectively. In comparison the WT-C-region has t_m^s of 45 ms (Cukkemane et al., 2021), and previous studies on A β (Dregni et al., 2020) have shown that water which associated with fibers, presented in interfibrillar spaces and in the bulk have relaxation times of 90, 150, and 250, respectively. This observation supports the notion that water molecules are in proximity to

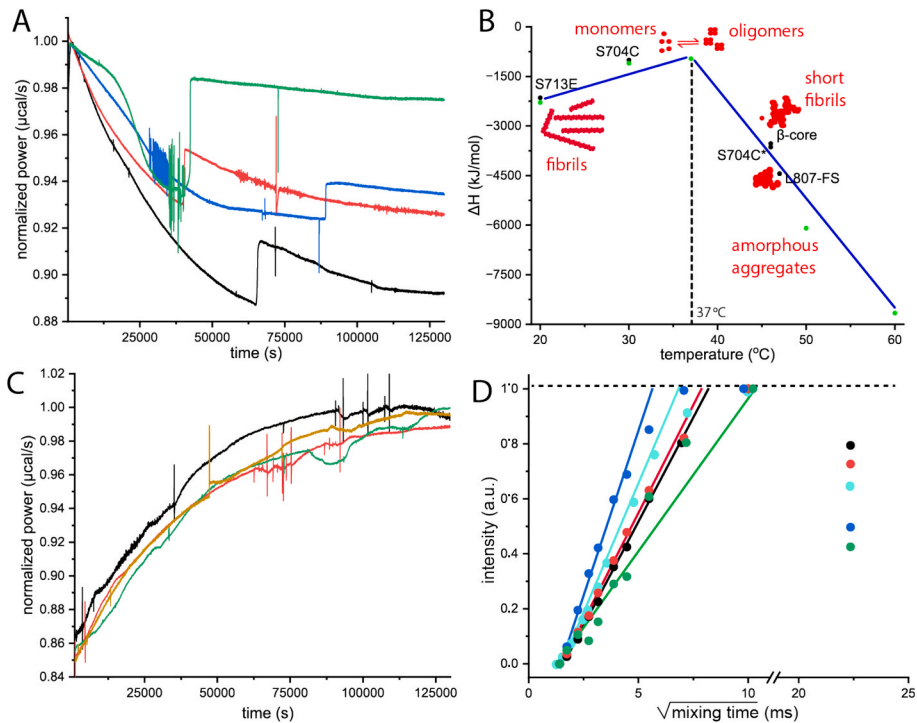


Fig. 4. Thermodynamic characterization of the self-association process of the WT C-region, its mutants, and the β -core. (A) The enthalpy changes observed for a 10 μ M his-tagged protein solution of WT-C-region at 20 $^{\circ}$ C (blue), WT-C-region at 30 $^{\circ}$ C (green), S713E (black), S704C (red). (B) Pictorial representation highlighting the similarities of enthalpy changes between the WT-C-region at various temperatures (green circles) on the one hand and the mutant variants as well as the β -core at 20 $^{\circ}$ C (black circles) on the other. (C) L807-FS (red), β -core (green), WT-C-region at 50 $^{\circ}$ C (orange (image adapted from Fig. 3A from Cukkemane et al., 2021 (Cukkemane et al., 2021))), and one-week old S704C (black). (D) Normalized intensities obtained for 1D water-edited (see Fig. S5 for spectra) build-up curves from 13 C spectra for signals corresponding to C α region (50–75 ppm) were plotted against the square root of the 1 H– 1 H mixing time for S713E (blue), S704C (red), L807-FS (green), β -core (black), and the his-tagged WT-C-region (cyan (image adapted from Fig. 3D from Cukkemane et al., 2021 (Cukkemane et al., 2021))). Straight lines represent the fitted slopes that were used for estimation of t_m^s . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

Comparison of enthalpy changes recorded for C-region mutant variants and the β -core with those of the WT-C-region at various temperatures (20–60 $^{\circ}$ C).

C-region variant	ΔH [kJ/mol]	WT-C-region (Temperature $^{\circ}$ C)	ΔH [kJ/mol]
S713E	–2144	20	–2289
S704C	–1004	30	–1103
β -core	–3528	37	–962
S704C *	–3635	45	–2761
L807-FS	–4444	50	–6092
		60	–8658

*One week old sample.

the protein surface for all the constructs.

However, to understand the finer details of the aggregates observed in AFM images (Fig. 3 and Fig. S3), we were curious to understand their molecular organization. We hypothesized that the aggregates are composed of a mesh of filaments that serves as a seed for fibrillar growth. Therefore, we assumed a structural model of an elongated fibril with a circular cross-sectional area (Schmidt-Rohr et al., 1994); based on this, we calculated the diameters (supplementary methodology) of the different fibrils to be 5.73, 5.56, 15.92, and 5.92 nm (Table S1) for S704C, S713E, L807-FS, and β -core, respectively. Barring the L807-FS mutant, these fibril diameters would be comparable to those of the WT construct (5–7 nm (Cukkemane et al., 2021)). But unlike the spectra for other C-region constructs and the β -core, the L807-FS shows a noisy spectrum due to protein instability during preparation and lesser sample in the NMR rotor. However, comparing the buildup time of L807-FS (62.50 ms) with those reported for other amyloid fibrils, which have

t_m^s of 17–110 ms (Schneider et al., 2011; Dregni et al., 2020; Wang et al., 2017), for ordered water molecules interacting with the core region of β -fibrils, the data suggest that the aggregates of L807-FS may contain an underlying fibrillar assembly. The findings suggest that the seemingly amorphous aggregates may constitute a mesh-like network that contains water molecules associated within the structure and some bound on the surface (Fig. 5A). However, with the information we have acquired, we believe that the aggregated material acts as the seed that promotes the assembly of the larger structures observed in the DLS and AFM analyses.

Discussion

The aim of this study has been to further structurally characterize the DISC1 C-region and to understand its significance to the pathophysiology of CMIs such as schizophrenia, MDD and BD. Following this line, we characterized three mutants (Fig. 1A and 1B) of the C-region using a combination of biophysical and structural biology applications. Intuitively, one would expect that the mutation in the S713E C-region variant will not have a dramatic effect on protein stability, while the impact of the Ser-to-Cys exchange in position 704 is more difficult to predict. In line with this expectation, S713E displays ThT fluorescence similar to that of the WT-C-region (Fig. 1C), along with a similar thermogram at 20 $^{\circ}$ C (Fig. 3B). In contrast, S704C shows weak ThT reactivity (Fig. 1C) and its thermogram is reminiscent of the WT-C-region at 30 $^{\circ}$ C. We also tested a one-week-old S704C sample for its thermodynamic properties; in this case, the thermogram at 20 $^{\circ}$ C resembled the profile of the WT-C-region at 50 $^{\circ}$ C (Fig. 4B and Table 2). Our findings on the S704C C-region may therefore hint at a heterogeneous nature of the aggregates in post-mortem brain samples (Leliveld et al., 2009; Tanaka et al., 2017) and their inability to bind to the ThT dye. To summarize our

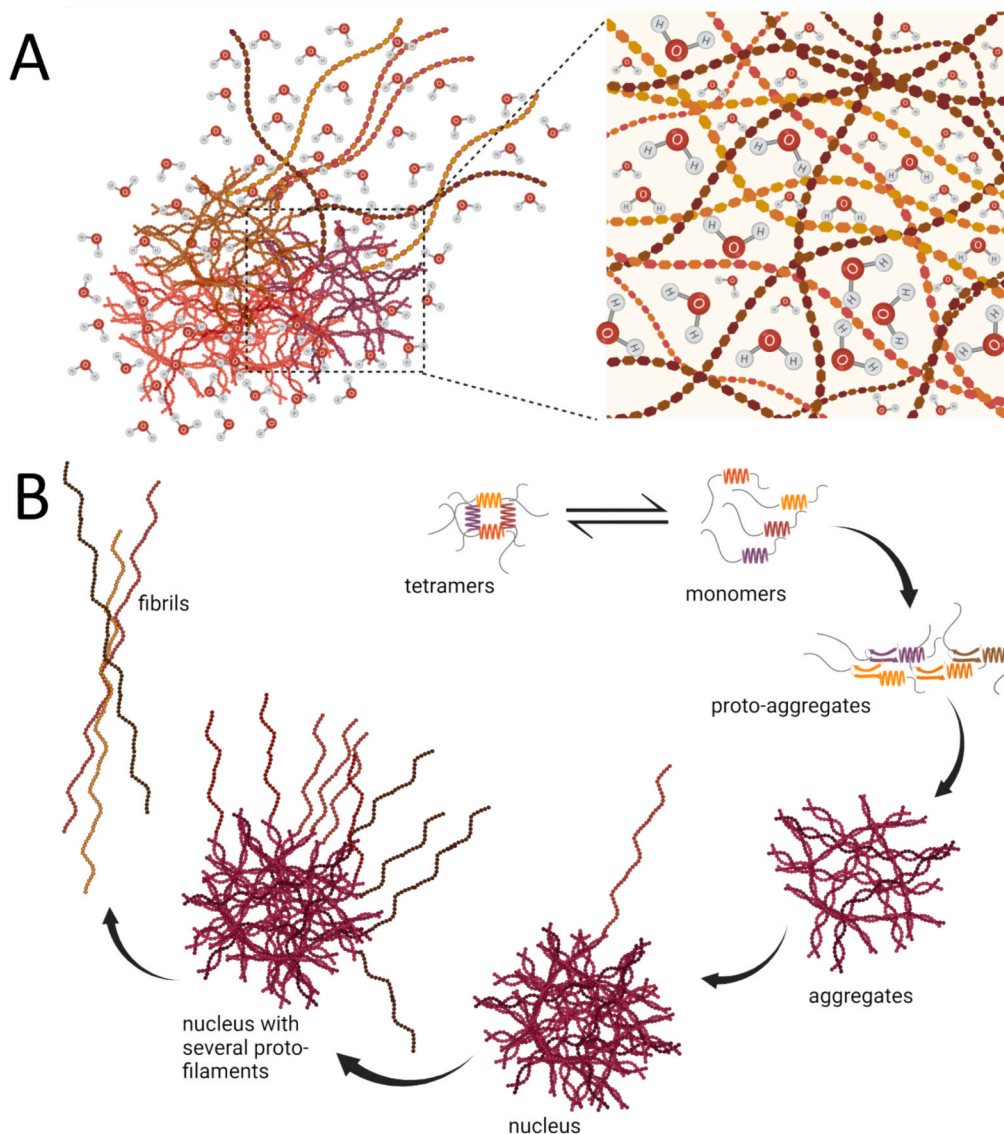


Fig. 5. (A) Presumptive model of a DISC1 C-region aggregate including water molecules filling cavities and associated along the length of the fibrils. (B) Snapshots describing the aggregation and fibrillar pathway of the DISC1 C-region; the mutant variants are suggested to follow a similar pathway.

observations, the mutations where a polar Ser is exchanged to Glu/Cys are tolerated but the behavior and the stability of the resulting proteins are reminiscent of the WT-C-region maintained at different temperatures. In comparison with the mutant variants mentioned above, L807-FS also showed low intensity ThT fluorescence (Fig. 1C), but it rapidly aggregated and was very unstable in solution. The L807-FS mutant displayed a calorimetric footprint that resembled the WT protein at 50 °C.

As an extension of our previous study on the DISC1 C-region fibrilization process (Cukkemane et al., 2021), we wanted to comprehend the role of the pseudo-repeat sequence (717–761), i.e., the β -core. In accordance with our hypothesis, the β -core displayed a sigmoidal trend of increase in particle size like the WT-C-region, with amorphous aggregates being the prevalent species under all tested conditions. These findings are of strong physiological significance because it includes half of the amino acid stretch of the $\Delta 22$ region (748–769, (Kamiya et al., 2006), which is absent in DISC1 $\Delta 22aa$. When expressed in PC12 cells, DISC1 $\Delta 22aa$ is incapable of interacting with NUDEL and LIS1 from the mitotic complex, which prohibits the proliferation of the cell line. In a similar vein, we believe that the participation of the β -core in fibrillar assembly will impair DISC1 function, presumably hampering mitosis

and ensuing processes such as cell division, neuronal proliferation, and migration.

Intriguingly, the ThT fluorescence signal arising from this dye binding to DISC1 has been reported only in a study evaluating DISC1 function in relation to Huntington's disease (Tanaka et al., 2017), where the presence of DISC1 as fibers was demonstrated, and in our previous report (Cukkemane et al., 2021), where we showed that the fibers of the C-region were β -fibrils. This raises the important question why DISC1 reactivity with ThT has been observed only sporadically. At least part of the explanation may be related to the heterogeneous nature of the aggregates, which is corroborated by the coexistence of (seemingly) amorphous material and fibrils, which is concentration and temperature dependent (Cukkemane et al., 2021) (Fig. 3, S2-S3). In this context, it is also important to highlight that binding of fluorescent dyes such as ThT and Congo red occurs in the fibrillar grooves, which are formed by multiple protofilaments (Schutz et al., 2015; Frieg et al., 2020). Such a process involves the intercalation of ThT between two filaments. Until now, however, all fibrils of the DISC1 C-region (Cukkemane et al., 2021) (Fig. 3A), that have been observed appear as mono-filamentous assemblies. This may underlie the observed poor fluorescence of ThT in the presence of the WT-C-region (Cukkemane et al., 2021) as well as the

recombinantly expressed mutants; extrapolating to the full-length protein, these considerations would also explain the absence of fluorescent signals from samples obtained from post-mortem brain tissues (Leliveld et al., 2009; Korth, 2012).

So, what drives the DISC1 C-region into an amyloid-like structure and what similarities does it share with its counterparts from the neurodegenerative world? From a thermodynamic perspective, protein association can often be described by entropy-enthalpy compensation. In such a scenario, formation of aggregates and fibrils is driven by enthalpy loss (Kardos et al., 2004; Ikenoue et al., 2014), while concomitantly the entropy of the system (with protein and solvent contributions S_{prot} and S_{sol} , respectively) is reduced. For all tested samples, a strong exothermic reaction and presence of water was observed. The common feature among amyloids that likely drives aggregation and fibrillar assembly is the accumulation of several weak interactions within the core of the protein involving hydrophobic residues (Gazit, 2002) in repeats (Schneider et al., 2011; Gazit, 2002; Cherny et al., 2005) and pseudo-repeats (Wasmer et al., 2008; Fitzpatrick et al., 2017), which is reflected by enthalpic changes. In several natively disordered and amyloidogenic proteins (Kardos et al., 2004; Ikenoue et al., 2014; Lawrence et al., 2014; Rezaei et al., 2002; Ikenoue et al., 2014; Kinoshita et al., 2018; Jeppesen et al., 2010) a highly exothermic process was reported, potentially highlighting the association of the hydrophobic core. For many amyloids, it has been indicated that water is aligned along the length of the fibril (Schneider et al., 2011; Dregni et al., 2020; Wang et al., 2017). The ordered water may also be arranged in small pockets within the hydrophobic core (Dregni et al., 2020) or, in other examples, be associated as compartments in different pools (Wang et al., 2017) such as surface-bound water (loosely or tightly bound), interfibrillar water and bulk-like matrix water. Altogether, water molecules arranged in a well-ordered fibrillar system will be an indirect indicator of low protein entropy. And a more accurate description of the molecular architecture of the aggregates can be obtained from a combination of high-resolution spectra, D_2O exchange experiments (Soares et al., 2011), and residue-specific water-edited experiments (Wang et al., 2017).

What exactly constitutes the supra-molecular structure of the DISC1 C-region? The DISC1 protein (Fig. 1B) is composed of largely unstructured and four distinct structured regions (Yerabham et al., 2013; Yerabham et al., 2017; Levin et al., 2014). The C-region, which is one of the structured segments is composed of helical coiled-coils (Ye et al., 2017; Wang et al., 2019) and unstructured elements (Cukkemane et al., 2021; Ye et al., 2017), especially in the absence of an interacting partner. The partially disordered and coiled-coil regions provide ample plasticity for integration into various cellular signaling pathways, interacting with homo- or hetero-oligomeric partners. We demonstrated previously (Cukkemane et al., 2021) that the DISC1 C-region has a strong propensity to form tetrameric structures, aggregates, and fibrils. The reaction may largely be driven by hydrophobic interactions (Fig. 1A) among residues in the pseudo-repeat of the β -core. Perhaps, the aggregation mechanism potentially progresses via liquid phase separation (Banach et al., 2019; Patel et al., 2015; Fandrich et al., 2018).

These amorphous-like aggregates are likely composed of a mesh-like protein network with water molecules embedded within the structure (Fig. 5A), which then serves as a seed for fibrillar growth. We tend to favor the model describing the liquid droplets of the RNA-binding protein FUS (Fandrich et al., 2018), which is a risk factor for amyotrophic lateral sclerosis (ALS). Similar to FUS, we observe that the aggregates (Fig. 3) of the WT C-region serve as seeds and that fibrils extend from such supra-molecular structures. Assuming the C-region and its S704C and S713E mutants share the same pathway to fibrillization, we provide a model with pictorial snapshots (Fig. 5B) depicting the growth of the assembly, where the building block is the tetramer, which progressively self-associates to form larger structures; this in turn promotes seeding followed by elongation of the fibrils, which are either short or long depending on the conditions, e.g., temperature.

Taking into account the results presented herein, we presume the conformational flexibility of the C-region and perhaps also that of the full-length DISC1 may be key to enabling interactions with a myriad of partners (Yerabham et al., 2013; Millar et al., 2003) as a scaffold protein. At the same time, this plasticity likely facilitates the formation of heterogeneous aggregates (Fig. 3, S2 and S3) in DISC1opathies (Korth, 2012), which is relevant to the clinical findings from human brain tissue. For this reason, we strongly advocate a structure-based rationale in drug design. This is because the heterogeneous nature of the particles including the sporadic presence of fibrils has posed a significant challenge to the drug development strategy aiming to prevent aggregation in neurodegenerative disorders (Plascencia-Villa and Perry, 2023). As an example of recent relevance, several reasons have been put forward for the relatively poor efficacy of the drug mAb Aducanumab directed against aggregated A β (Wojtunik-Kulesza et al., 2023), including structural implications (Colvin et al., 2016). It has been argued that the mAb binds to an N-terminal region that is unstructured (Gremer et al., 2017) or structured (Cukkemane and Willbold, 2022) under different conditions, thereby limiting the effectiveness of the drug. Taking a leaf out of the lessons with neurodegenerative disease targets and avoiding similar pitfalls, we have implemented a different approach. We are developing peptide therapeutics based on a hit-and-lead strategy using the phage-display technique [57], explicitly targeting either monomers or oligomers, which are generated using the conditions described herein and reported previously (Cukkemane et al., 2021). The aim is to identify peptides which bind individual protomers in the aggregate or fibril, thereby uncoupling them from the large supramolecular assembly and restoring DISC1 function. Clearly, DISC1 is involved in a variety of physiological processes, interacts with many potential partners, and may also have overwhelming structural diversity. Therefore, it is pertinent to further characterize structure and function of the different domains of this large scaffold protein when rationalizing an adequate drug development strategy.

Authors contribution

AC, DW, and OHW designed the experiments and drafted the manuscript. AC prepared all the samples and performed experiments using DLS and ITC. ssNMR experiments were performed by AC alongside NB and HH. AFM was performed by TK. All authors critically reviewed the manuscript.

CRediT authorship contribution statement

Abhishek Cukkemane: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Nina Becker:** Writing – review & editing, Methodology, Investigation. **Tatsiana Kupreichyk:** Writing – review & editing, Methodology, Investigation. **Henrike Heise:** Writing – review & editing, Resources, Methodology, Investigation, Funding acquisition. **Dieter Willbold:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Oliver H. Weiergräber:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: AC and DW are inventors of patents covering the peptide drugs. All other authors declare no competing interests.

Acknowledgments

Access to the Jülich-Düsseldorf Biomolecular NMR Center jointly run by Forschungszentrum Jülich and HHU and support by the DFG (HE 3243/4-1 and INST 208/771-1 FUGG) are acknowledged. We thank the Alexander von Humboldt foundation (Ref 3.4 - 1142747 - HRV – IP) for financial support. We thank the izv. prof. dr. sc. Nicholas J. Bradshaw from the University of Rijeka, Croatia, and Prof. Dr. Andrew Dingley from the Forschungszentrum Jülich, Germany for critically reading the manuscript. We thank the group of Prof. Dr. Gunnar Schröder for helpful discussions on sample preparation for structural characterization.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsbx.2025.100128>.

Data availability

No data was used for the research described in the article.

References

- Ader, C., Schneider, R., Seidel, K., Etzkorn, M., Becker, S., Baldus, M., 2009. Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 131, 170–176.
- Banach, M., Konieczny, L., Roterman, I., 2019. The Amyloid as a Ribbon-Like Micelle in Contrast to Spherical Micelles Represented by Globular Proteins. *Molecules* 24.
- Blackwood, D.H.R., Fordyce, A., Walker, M.T., St Clair, D.M., Porteous, D.J., Muir, W.J., 2001. Schizophrenia and affective disorders - Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: Clinical and P300 findings in a family. *Am. J. Hum. Genet.* 69, 428–433.
- Bradshaw, N.J., Korth, C., 2019. Protein misassembly and aggregation as potential convergence points for non-genetic causes of chronic mental illness. *Mol. Psychiatry* 24, 936–951.
- Cherny, I., Rockah, L., Levy-Nissenbaum, O., Gophna, U., Ron, E.Z., Gazit, E., 2005. The formation of *Escherichia coli* curli amyloid fibrils is mediated by prion-like peptide repeats. *J. Mol. Biol.* 352, 245–252.
- Colvin, M.T., Silvers, R., Ni, Q.Z., Can, T.V., Sergeyev, I., Rosay, M., Donovan, K.J., Michael, B., Wall, J., Linse, S., Griffin, R.G., 2016. Atomic Resolution Structure of Monomeric Abeta42 Amyloid Fibrils. *J. Am. Chem. Soc.* 138, 9663–9674.
- A. Cukkemane, Willbold, D., DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDs), in, vol. PCT/EP2023/066796, Forschungszentrum Juelich GmbH, EPO, 2022.
- Cukkemane, A., Becker, N., Zielinski, M., Frieg, B., Lakomek, N.A., Heise, H., Schroder, G.F., Willbold, D., Weiergraber, O.H., 2021. Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl. Psychiatry* 11, 639.
- Dregni, A.J., Duan, P., Hong, M., 2020. Hydration and Dynamics of Full-Length Tau Amyloid Fibrils Investigated by Solid-State Nuclear Magnetic Resonance. *Biochemistry* 59, 2237–2248.
- Fandrich, M., Nystrom, S., Nilsson, K.P.R., Bockmann, A., LeVine 3rd, H., Hammarstrom, P., 2018. Amyloid fibril polymorphism: a challenge for molecular imaging and therapy. *J. Intern. Med.* 283, 218–237.
- Fitzpatrick, A.W.P., Falcon, B., He, S., Murzin, A.G., Murshudov, G., Garringer, H.J., Crowther, R.A., Ghetti, B., Goedert, M., Scheres, S.H.W., 2017. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* 547, 185–190.
- Frieg, B., Gremer, L., Heise, H., Willbold, D., Gohlke, H., 2020. Binding modes of thioflavin T and Congo red to the fibril structure of amyloid-beta(1-42). *Chem. Commun. (Camb)* 56, 7589–7592.
- Fung, B.M., Khitritin, A.K., Ermolaev, K., 2000. An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* 142, 97–101.
- Gazit, E., 2002. Global analysis of tandem aromatic octapeptide repeats: the significance of the aromatic-glycine motif. *Bioinformatics* 18, 880–883.
- Gazit, E., 2002. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB J.* 16, 77–83.
- Gremer, L., Scholzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R.B.G., Tusche, M., Lopez-Iglesias, C., Hoyer, W., Heise, H., Willbold, D., Schroder, G.F., 2017. Fibril structure of amyloid-beta(1-42) by cryo-electron microscopy. *Science* 358, 116–119.
- Hashimoto, R., Numakawa, T., Ohnishi, T., Kumamaru, E., Yagasaki, Y., Ishimoto, T., Mori, T., Nemoto, K., Adachi, N., Izumi, A., Chiba, S., Noguchi, H., Suzuki, T., Iwata, N., Ozaki, N., Taguchi, T., Kamiya, A., Kosuga, A., Tatsumi, M., Kamijima, K., Weinberger, D.R., Sawa, A., Kunugi, H., 2006. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Hum. Mol. Genet.* 15, 3024–3033.
- Hikida, T., Gamo, N.J., Sawa, A., 2012. DISC1 as a therapeutic target for mental illnesses. *Expert Opin Ther Tar* 16, 1151–1160.
- Ikenoue, T., Lee, Y.H., Kardos, J., Yagi, H., Ikegami, T., Naiki, H., Goto, Y., 2014. Heat of supersaturation-limited amyloid burst directly monitored by isothermal titration calorimetry. *PNAS* 111, 6654–6659.
- Ikenoue, T., Lee, Y.H., Kardos, J., Saiki, M., Yagi, H., Kawata, Y., Goto, Y., 2014. Cold denaturation of alpha-synuclein amyloid fibrils. *Angew. Chem. Int. Ed. Engl.* 53, 7799–7804.
- Ishizuka, K., Kamiya, A., Oh, E.C., Kanki, H., Seshadri, S., Robinson, J.F., Murdoch, H., Dunlop, A.J., Kubo, K., Furukori, K., Huang, B., Zeledon, M., Hayashi-Takagi, A., Okano, H., Nakajima, K., Houslay, M.D., Katsanis, N., Sawa, A., 2011. DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature* 473, 92–96.
- Jarrett, J.T., Lansbury Jr., P.T., 1993. Seeding “one-dimensional crystallization” of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73, 1055–1058.
- Jeppesen, M.D., Westh, P., Otzen, D.E., 2010. The role of protonation in protein fibrillation. *FEBS Lett.* 584, 780–784.
- Kamiya, A., Tomoda, T., Chang, J., Takaki, M., Zhan, C., Morita, M., Cascio, M.B., Elashvili, S., Koizumi, H., Takanezawa, Y., Dickerson, F., Yolken, R., Arai, H., Sawa, A., 2006. DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. *Hum. Mol. Genet.* 15, 3313–3323.
- Kardos, J., Yamamoto, K., Hasegawa, K., Naiki, H., Goto, Y., 2004. Direct measurement of the thermodynamic parameters of amyloid formation by isothermal titration calorimetry. *J. Biol. Chem.* 279, 55308–55314.
- Kinoshita, M., Lin, Y., Dai, I., Okumura, M., Markova, N., Ladbury, J.E., Sterpone, F., Lee, Y.H., 2018. Energy landscape of polymorphic amyloid generation of beta2-microglobulin revealed by calorimetry. *Chem. Commun. (Camb)* 54, 7995–7998.
- Korth, C., 2012. Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion* 6, 134–141.
- Lawrence, C.W., Kumar, S., Noid, W.G., Showalter, S.A., 2014. Role of Ordered Proteins in the Folding-Up-On-Binding of Intrinsically Disordered Proteins. *J. Phys. Chem. Lett.* 5, 833–838.
- Leliveld, S.R., Bader, V., Hendriks, P., Prikulis, I., Sajani, G., Requena, J.R., Korth, C., 2008. Insolubility of disrupted-in-schizophrenia 1 disrupts oligomer-dependent interactions with nuclear distribution element 1 and is associated with sporadic mental disease. *J. Neurosci.* 28, 3839–3845.
- Leliveld, S.R., Hendriks, P., Michel, M., Sajani, G., Bader, V., Trossbach, S., Prikulis, I., Hartmann, R., Jonas, E., Willbold, D., Requena, J.R., Korth, C., 2009. Oligomer assembly of the C-terminal DISC1 domain (640-854) is controlled by self-association motifs and disease-associated polymorphism S704C. *Biochemistry* 48, 7746–7755.
- Levin, A., Mason, T.O., Adler-Abramovich, L., Buell, A.K., Meisl, G., Galvagnoni, C., Bram, Y., Stratford, S.A., Dobson, C.M., Knowles, T.P., Gazit, E., 2014. Ostwald's rule of stages governs structural transitions and morphology of dipeptide supramolecular polymers. *Nat. Commun.* 5, 5219.
- Luo, W., Hong, M., 2010. Conformational changes of an ion channel detected through water-protein interactions using solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 132, 2378–2384.
- Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A.M., Devon, R.S., St Clair, D.M., Muir, W.J., Blackwood, D.H.R., Porteous, D.J., 2000. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* 9, 1415–1423.
- Millar, J.K., Christie, S., Porteous, D.J., 2003. Yeast two-hybrid screens implicate DISC1 in brain development and function. *Biochem Biophys Res Commun* 311, 1019–1025.
- Nakata, K., Lipska, B.K., Hyde, T.M., Ye, T., Newburn, E.N., Morita, Y., Vakkalanka, R., Barenboim, M., Sei, Y., Weinberger, D.R., Kleinman, J.E., 2009. DISC1 splice variants are upregulated in schizophrenia and associated with risk polymorphisms. *PNAS* 106, 15873–15878.
- Narayanan, S., Arthanari, H., Wolfe, M.S., Wagner, G., 2011. Molecular characterization of disrupted in schizophrenia-1 risk variant S704C reveals the formation of altered oligomeric assembly. *J. Biol. Chem.* 286, 44266–44276.
- Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoykov, S., Mahamid, J., Saha, S., Franzmann, T.W., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L.A., Weigert, M., Myers, E.M., Grill, S., Drechsel, D., Hyman, A.A., Alberti, S., 2015. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162, 1066–1077.
- Plascencia-Villa, G., Perry, G., 2023. Lessons from anti-amyloid-beta immunotherapies in Alzheimer's disease. *Handb. Clin. Neurol.* 193, 267–292.
- Rezaei, H., Choiset, Y., Eghiaian, F., Treguer, E., Mentre, P., Debey, P., Grosclaude, J., Haertle, T., 2002. Amyloidogenic unfolding intermediates differentiate sheep prion protein variants. *J. Mol. Biol.* 322, 799–814.
- Sachs, N.A., Sawa, A., Holmes, S.E., Ross, C.A., DeLisi, L.E., Margolis, R.L., 2005. A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Mol. Psychiatry* 10, 758–764.
- Schmidt-Rohr, K., Spiess, H.W., 1994. CHAPTER ONE - Introduction. In: Schmidt-Rohr, K., Spiess, H.W. (Eds.), *Multidimensional Solid-State NMR and Polymers*. Academic Press, San Diego, pp. 1–12.
- Schneider, R., Schumacher, M.C., Mueller, H., Nand, D., Klauk, V., Heise, H., Riedel, D., Wolf, G., Behrmann, E., Rauner, S., Seidel, R., Engelhard, M., Baldus, M., 2011. Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy. *J. Mol. Biol.* 412, 121–136.
- Schutz, A.K., Vagt, T., Huber, M., Ovchinnikova, O.Y., Cadalbert, R., Wall, J., Guntert, P., Bockmann, A., Glockshuber, R., Meier, B.H., 2015. Atomic-resolution three-dimensional structure of amyloid beta fibrils bearing the Osaka mutation. *Angew. Chem. Int. Ed. Engl.* 54, 331–335.

- Soares, D.C., Carlyle, B.C., Bradshaw, N.J., Porteous, D.J., 2011. DISC1: Structure Function, and Therapeutic Potential for Major Mental Illness. *ACS Chem Neurosci* 2, 609–632.
- Tanaka, M., Ishizuka, K., Nekooki-Machida, Y., Endo, R., Takashima, N., Sasaki, H., Komi, Y., Gathercole, A., Huston, E., Ishii, K., Hui, K.K., Kurosawa, M., Kim, S.H., Nukina, N., Takimoto, E., Houslay, M.D., Sawa, A., 2017. Aggregation of scaffolding protein DISC1 dysregulates phosphodiesterase 4 in Huntington's disease. *J. Clin. Invest.* 127, 1438–1450.
- Taylor, M.S., Devon, R.S., Millar, J.K., Porteous, D.J., 2003. Evolutionary constraints on the Disrupted in Schizophrenia locus. *Genomics* 81, 67–77.
- Tropea, D., Hardingham, N., Millar, K., Fox, K., 2018. Mechanisms underlying the role of DISC1 in synaptic plasticity. *J. Physiol-London* 596, 2747–2771.
- Wang, T., Jo, H., DeGrado, W.F., Hong, M., 2017. Water Distribution, Dynamics, and Interactions with Alzheimer's beta-Amyloid Fibrils Investigated by Solid-State NMR. *J. Am. Chem. Soc.* 139, 6242–6252.
- Wang, X., Ye, F., Wen, Z., Guo, Z., Yu, C., Huang, W.K., Rojas Ringeling, F., Su, Y., Zheng, W., Zhou, G., Christian, K.M., Song, H., Zhang, M., Ming, G.L., 2019. Structural interaction between DISC1 and ATF4 underlying transcriptional and synaptic dysregulation in an iPSC model of mental disorders. *Mol. Psychiatry*.
- Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A.B., Riek, R., Meier, B.H., 2008. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319, 1523–1526.
- Wojtunik-Kulesza, K., Rudkowska, M., Orzel-Sajdlowska, A., 2023. Aducanumab-Hope or Disappointment for Alzheimer's Disease. *Int. J. Mol. Sci.* 24.
- Ye, F., Kang, E., Yu, C., Qian, X., Jacob, F., Yu, C., Mao, M., Poon, R.Y.C., Kim, J., Song, H., Ming, G.L., Zhang, M., 2017. DISC1 Regulates Neurogenesis via Modulating Kinetochore Attachment of Ndel1/Nde1 during Mitosis. *Neuron* 96, 1204.
- Yerabham, A.S.K., Mas, P.J., Decker, C., Soares, D.C., Weiergraber, O.H., Nagel-Steger, L., Willbold, D., Hart, D.J., Bradshaw, N.J., Korth, C., 2017. A structural organization for the Disrupted in Schizophrenia 1 protein, identified by high-throughput screening, reveals distinctly folded regions, which are bisected by mental illness-related mutations. *J. Biol. Chem.* 292, 6468–6477.
- Yerabham, A.S.K., Muller-Schiffmann, A., Ziehm, T., Stadler, A., Kober, S., Indurkha, X., Marreiros, R., Trossbach, S.V., Bradshaw, N.J., Prikulis, I., Willbold, D., Weiergraber, O.H., Korth, C., 2018. Biophysical insights from a single chain camelid antibody directed against the Disrupted-in-Schizophrenia 1 protein. *PLoS One* 13, e0191162.
- Yerabham, A.S., Weiergraber, O.H., Bradshaw, N.J., Korth, C., 2013. Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol. Chem.* 394, 1425–1437.
- Yumerefendi, H., Tarendeau, F., Mas, P.J., Hart, D.J., 2010. ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J. Struct. Biol.* 172, 66–74.

SUPPLEMENTARY INFORMATION

Tracing the aggregation pathway of the scaffold protein DISC1: structural implications for chronic mental illnesses

Abhishek Cukkemane^{1,2*}, Nina Becker^{1,2}, Tatsiana Kupreichyk^{1,2}, Henrike Heise^{1,2}, Dieter Willbold^{1,2*} and Oliver H. Weiergräber^{1*}

¹Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

*email – a.cukkemane@fz-juelich.de; d.willbold@fz-juelich.de; o.h.weiergraeber@fz-juelich.de

Supplementary methods

Analysis of the water-edited 1D build-up experiments was performed by integrating the spectral region of 50-75 ppm, followed by normalization to the maximum signal intensity (at 100 ms mixing time). A linear fit to the initial build-up rate $\frac{1}{s} = (t_m^s)$ was used to determine the water accessibility of the sample. The slope (s) describes the time that is required to reach 100% magnetization transfer in the absence of any saturation effects [1], which is inversely proportional to the volume-to-surface area ratio (V/S) described by $\frac{V}{S} = \sqrt{\frac{D_{eff} t_m^s}{\pi}}$ where D_{eff} describes the effective magnetization diffusion coefficient corresponding to 0.2 nm²/ms [2, 3]. Assuming that the fibril represents an elongated cylinder where the fibrillary length greatly exceeds the diameter (d) of the fibril, the ratio of V/S equals d/4. Such an approach provides a semi-quantitative estimate of the molecular dimensions of water-accessible areas of membrane proteins [2, 4] and amyloid fibrils [3] in comparison to their low-resolution structural models.

Using a similar approach previously, we determined the diameter of the fibrils of the WT C-region to be 6.8 nm, that corresponds well with the EM findings, which was in the range of 5-7 nm [5].

Table S1: Water-edited buildup data for the C-region constructs investigated. The initial rate slope was obtained from a linear fit to the data points of the buildup curve. Also see Fig. S3 and Fig. 4

C-region variant	Initial rate slope (1/ms)	t_m^s (ms)	V/S (nm)	d (nm)
S713E	0.033	30.30	1.39	5.56
S704C	0.031	32.26	1.43	5.73
L807-FS	0.016	62.50	3.98	15.92
β -core	0.029	34.48	1.48	5.92

43 SUPPLEMENTARY FIGURES

44 WT-C-region (MBP-fusion construct)

45 MGVHSSHHHHHSSSENLYFQSRTKIEEGKLVINGDKGYNGLAEVGKKFEKDTGIK
46 VTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPAAAFQDKLYP
47 FTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKSALM
48 FNLQEPYFTWPLIAADGGYAFKYAAGKYDIKDVGVNDAGAKAGLTFLVDLIKNNHM
49 NADTDYSIAEHAFNHGETAMTINGPWAWSNIDTSAVNYGVTVLPTFKGQPSKPFVGV
50 LSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELVKDPRVA
51 ATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDAALAAQAQTNHMGGGS
52 GGGSASLEVLFGQPTSSELWEADLEACRLLIQSLQLQEARGSLSVEDERQMDDLEGAA
53 PPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGKK
54 LLYLEDQLHTAIHSHDEDLIQLRRELQMVKETLQAMILQLQPAKEAG

55 WT-C-region

56 MGHHHHHHDYDIPTTENLYFQGWADLEACRLLIQSLQLQEARGSLSVEDERQMDD
57 LEGAAPPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCE
58 DIGKKLLYLEDQLHTAIHSHDEDLIQLRRELQMVKETLQAMILQLQPAKEAGSG

59 S704C

60 MGHHHHHHDYDIPTTENLYFQGWADLEACRLLIQCLQLQEARGSLSVEDERQMDD
61 LEGAAPPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCE
62 DIGKKLLYLEDQLHTAIHSHDEDLIQLRRELQMVKETLQAMILQLQPAKEAGSG

63 S713E

64 MGHHHHHHDYDIPTTENLYFQGWADLEACRLLIQSLQLQEARGELSVEDERQMDD
65 LEGAAPPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCE
66 DIGKKLLYLEDQLHTAIHSHDEDLIQLRRELQMVKETLQAMILQLQPAKEAGSG

67 L807FS

68 MGHHHHHHDYDIPTTENLYFQGRVWEADLEACRLLIQSLQLQEARGSLSVEDERQ
69 MDDLEGAAPPIPPRLHSEDKRKTPLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELG
70 EKCEDIGKKLLYLEDQLHTAIHSHDEDLSSLGGSSRWGS

71 β -core

72 MGHHHHHHDYDIPTTENLYFQGDERQMDDLEGAAPPIPPRLHSEDKRKTPLKVLEE
73 WKTHLIPSLHCAGGEQKEESSG

74

75 Fig S1: The amino acid sequences of the DISC1 WT-C-region protein constructs, S704C,
76 S713E, L807 frame-shift and the β -core

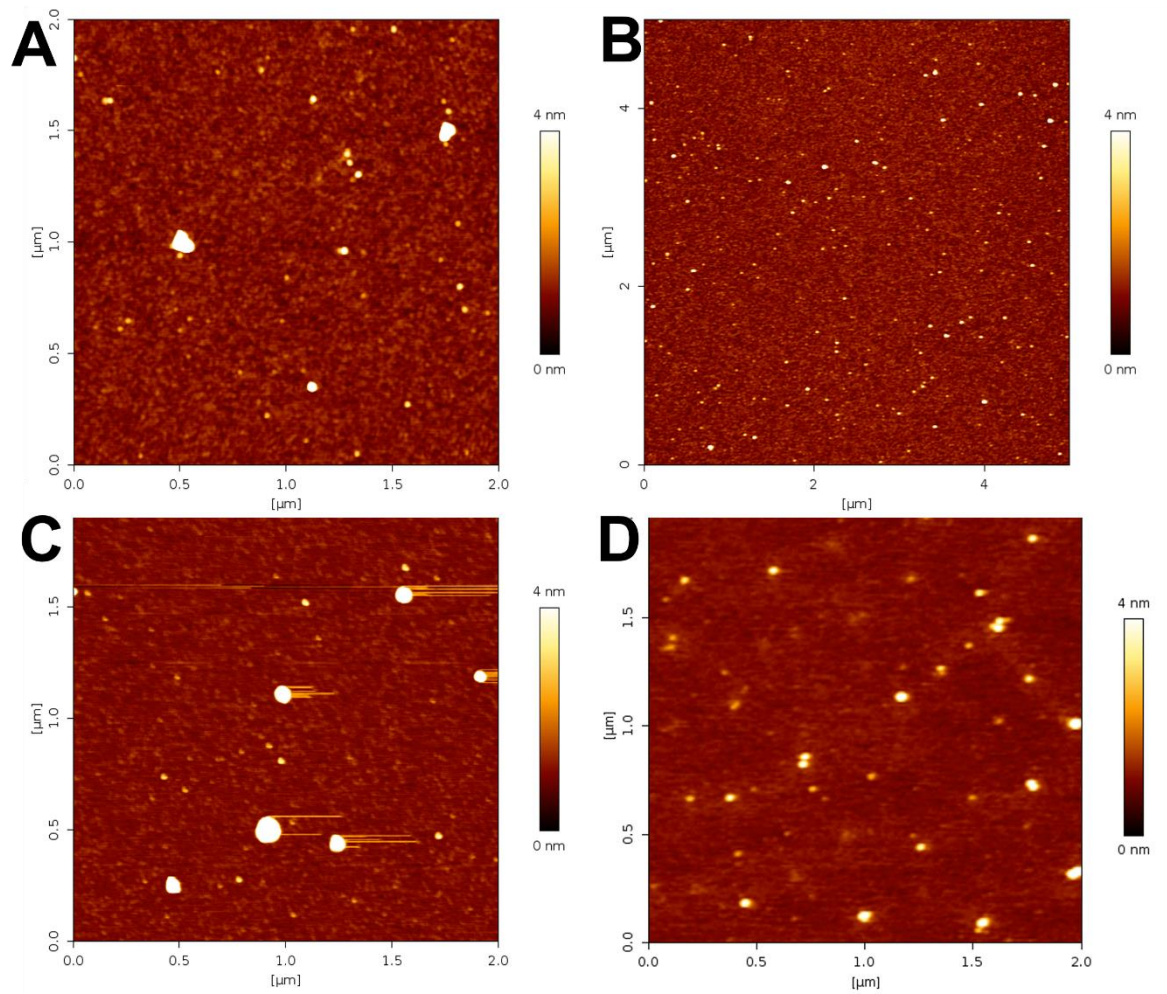


Fig. S2: Images of the His-tagged DISC1 C-region variants and the β -core at concentrations of 10 μ M in panels were recorded using atomic force microscopy (AFM) for (A) S713E, (B) S704C, (C) β -core, and L807-FS.

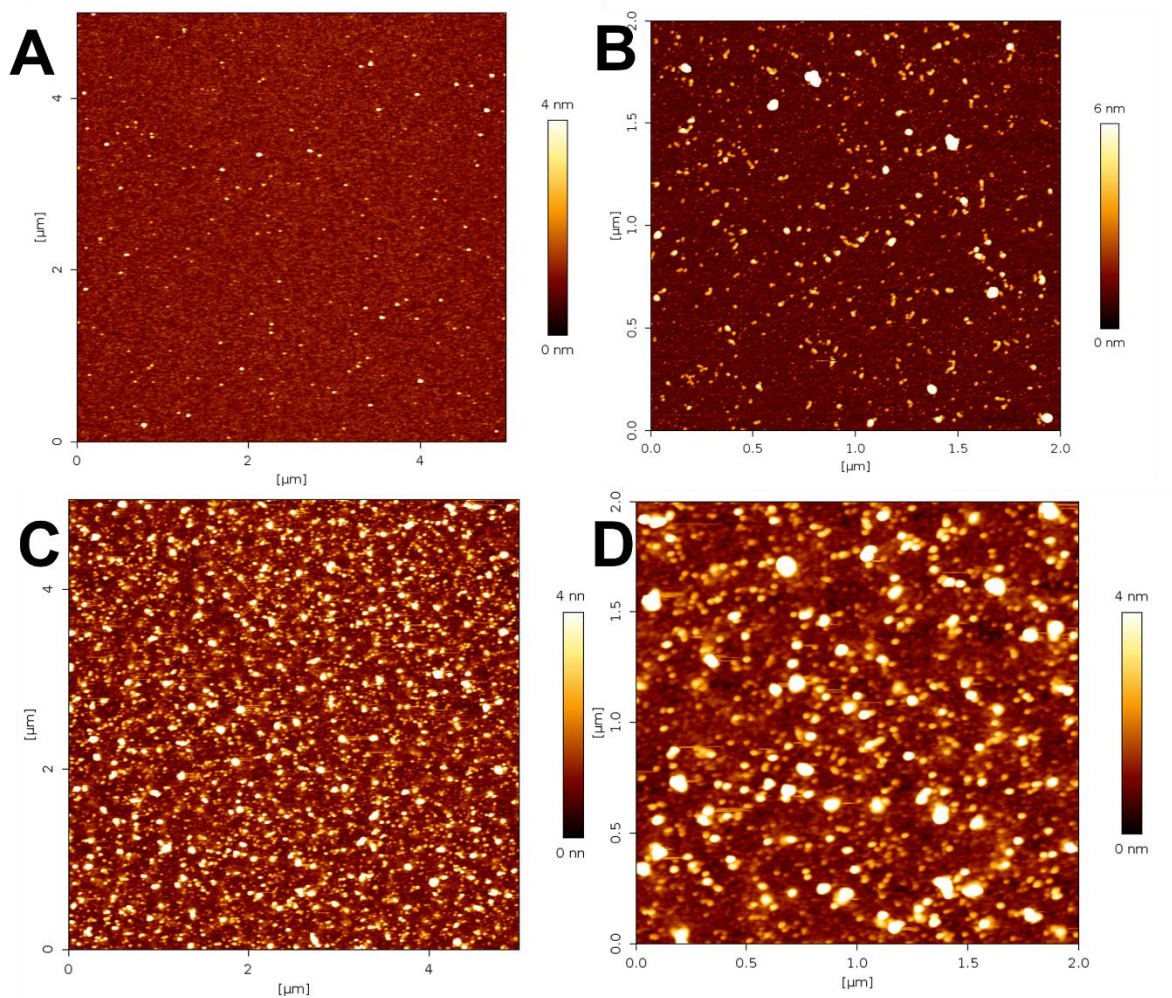


Fig. S3: Images of the His-tagged DISC1 C-region variants and the β -core at concentrations of 50 μ M in panels were recorded using atomic force microscopy (AFM) for (A) S713E, (B) S704C, (C) β -core, and (D) L807-FS.

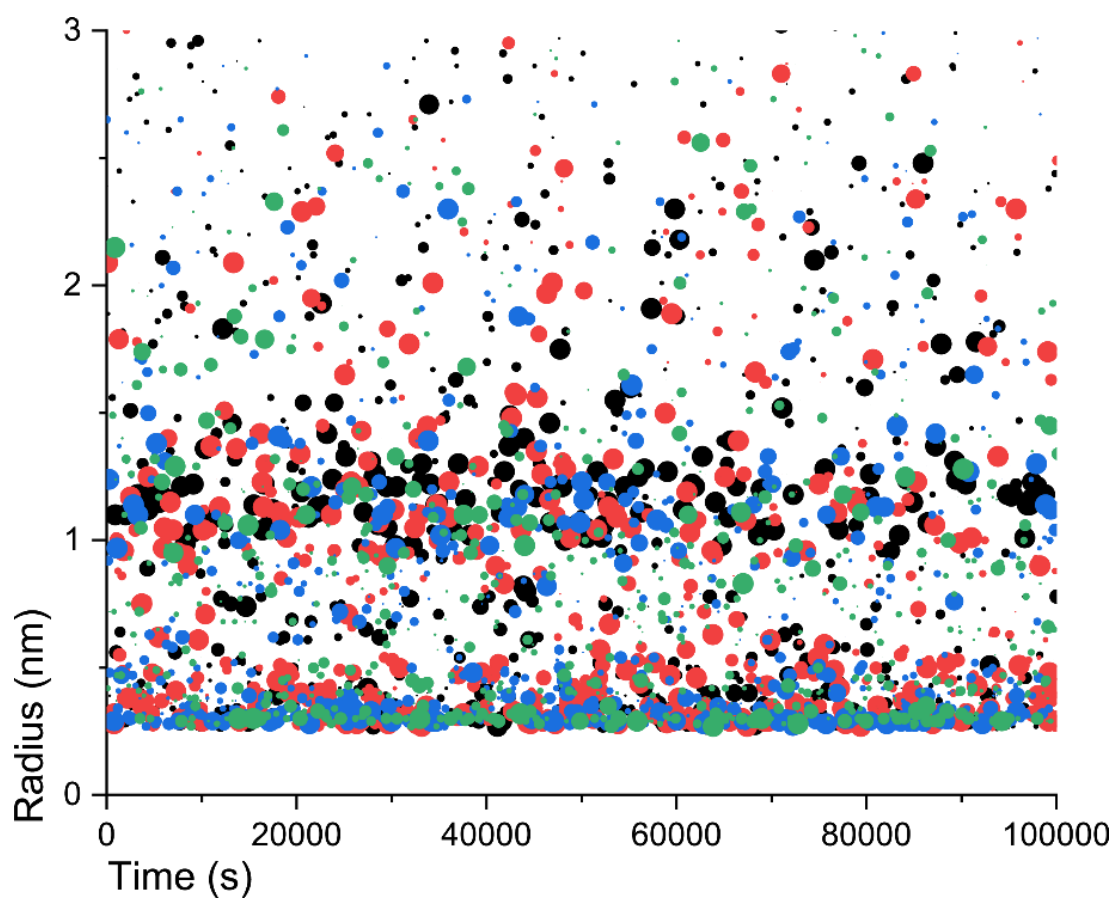


Fig. S4: The figure describes the DLS results of C-region variants at a concentration of 2.5 μ M for S704C (red), S713E (black), L807-FS (blue), and β -core (green) at 20 $^{\circ}$ C.

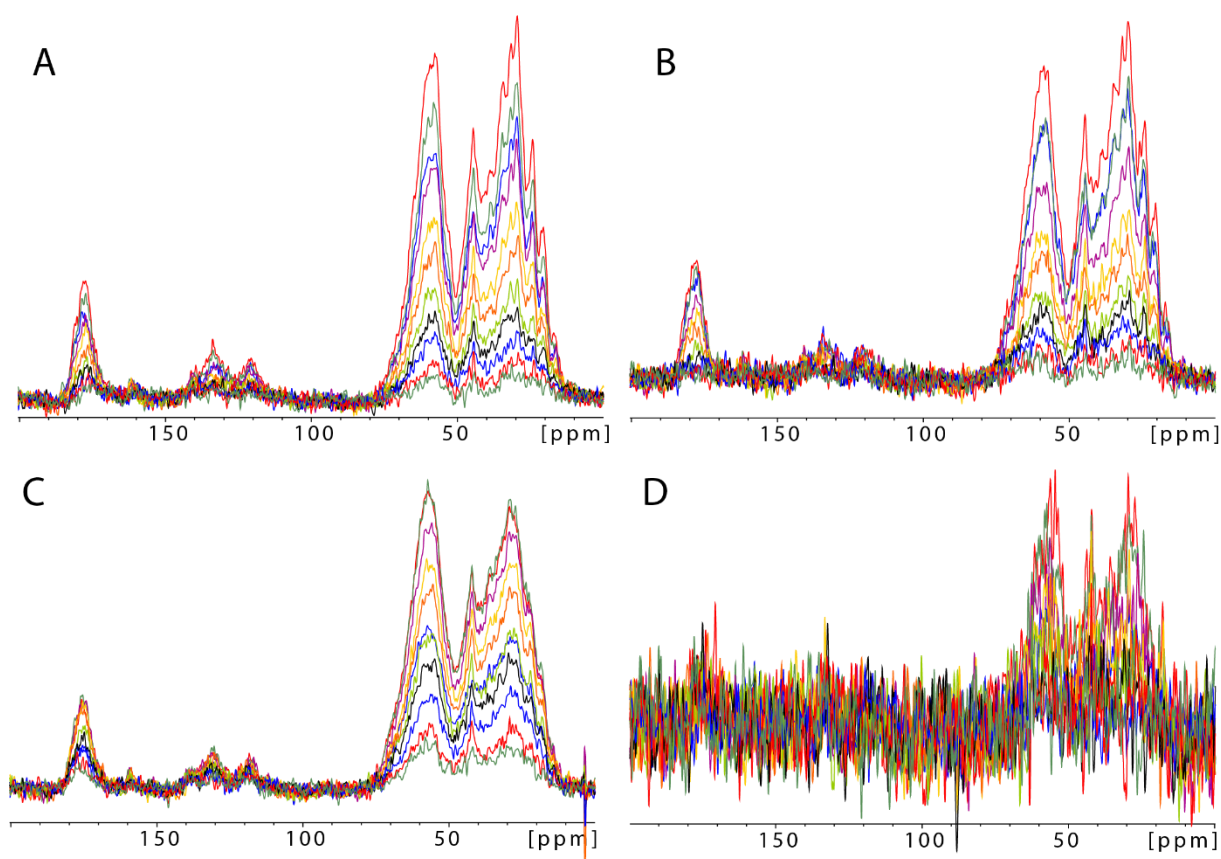


Fig. S5: 1D ^{13}C water build-up spectra of (A) S713E, (B) S704C, (C) β -core, and (D) L807-FS with several ^1H - ^1H mixing times ranging from 500 ms (blue), 100 ms (red), 50 ms (green), 30 ms (magenta), 20 ms (yellow), 15 ms (orange), 10 ms (light-green), 7 ms (black), 5 ms (blue), 3 ms (red), and 2 ms (green), respectively.

References:

1. Schmidt-Rohr K, Spiess HW: **CHAPTER ONE - Introduction**. In: *Multidimensional Solid-State NMR and Polymers*. edn. Edited by Schmidt-Rohr K, Spiess HW. San Diego: Academic Press; 1994: 1-12.
2. Ader C, Schneider R, Seidel K, Etzkorn M, Becker S, Baldus M: **Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy**. *J Am Chem Soc* 2009, **131**(1):170-176.
3. Schneider R, Schumacher MC, Mueller H, Nand D, Klaukien V, Heise H, Riedel D, Wolf G, Behrmann E, Raunser S *et al*: **Structural characterization of polyglutamine fibrils by solid-state NMR spectroscopy**. *J Mol Biol* 2011, **412**(1):121-136.
4. Luo W, Hong M: **Conformational changes of an ion channel detected through water-protein interactions using solid-state NMR spectroscopy**. *J Am Chem Soc* 2010, **132**(7):2378-2384.
5. Cukkemane A, Becker N, Zielinski M, Frieg B, Lakomek NA, Heise H, Schroder GF, Willbold D, Weiergraber OH: **Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses**. *Transl Psychiatry* 2021, **11**(1):639.

(19) World Intellectual Property
Organization

International Bureau

(43) International Publication Date
28 December 2023 (28.12.2023)(10) International Publication Number
WO 2023/247625 A9

(51) International Patent Classification:

C07K 7/08 (2006.01) A61K 38/04 (2006.01)

(21) International Application Number:

PCT/EP2023/066796

(22) International Filing Date:

21 June 2023 (21.06.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

22180453.7 22 June 2022 (22.06.2022) EP

(71) Applicant: FORSCHUNGSZENTRUM JÜLICH
GMBH [DE/DE]; 52425 Jülich (DE).(72) Inventors: CUKKEMANE, Abhishek; Düsseldorf
Straße 17, 52428 Jülich (DE). WILLBOLD, Dieter;
Theodor-Heuss-Straße 179, 52428 Jülich (DE).(74) Agent: RUPPRECHT, LL.M., Kay; Meissner Bolte
Patentanwälte Rechtsanwälte Partnerschaft mbB, Kaiser-
swerther Straße 183, 40474 Düsseldorf (DE).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,
SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ,
RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ,
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,
LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,
SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))
- with information concerning authorization of rectification
of an obvious mistake under Rule 91.3 (b) (Rule 48.2(i))
- in black and white; the international application as filed
contained color or greyscale and is available for download
from PATENTSCOPE

(48) Date of publication of this corrected version:

07 March 2024 (07.03.2024)

(15) Information about Correction:

see Notice of 07 March 2024 (07.03.2024)

(54) Title: DISC1-BINDING PEPTIDES AND THE USE THEREOF FOR THE TREATMENT AND DIAGNOSIS OF SCHIZOPHRENIA, MAJOR DEPRESSIVE DISORDERS (MDD), AND BIPOLAR DISORDERS (BD), AND AUTISM AND OTHER CHRONIC MENTAL DISORDERS (CMDS)

(57) Abstract: The invention relates to a peptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, and homologues, fragments and portions thereof.

DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDs)

DESCRIPTION

The invention relates to peptides comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, homologues, fragments, combinations and portions thereof, and to such a peptide for use in the diagnosis and/or treatment of chronic mental disorders, in particular schizophrenia.

Chronic mental disorders (CMDs) such as schizophrenia represent multi-factorial complex disease that is affected by an interplay of mental, biological, social and environmental conditions. With an incidence of 1%, schizophrenia represents a major social and economic burden. Taken together with related CMDs, the lifetime rate increases to 2.3 to 3.5%. The disorders represent neurodevelopmental impairments that do not manifest until decades usually during the onset of adolescence and are accompanied with significant comorbidity. A significantly rated biological risk factor that has been observed as aggregates in the post mortem brain tissue of patients suffering from schizophrenia is the protein Disrupted in Schizophrenia 1 (DISC1). However, the concept of proteinopathy in schizophrenia has remained elusive due to its microscopic nature. Recently, we demonstrated that pathophysiology of schizophrenia bears remarkable structural resemblance to other protein conformational disorder [1] akin its neurodegenerative cousins also forms β -fibrils.

Most forms of therapy that exist today are empirical and prescribed on a trial-and-error basis. They represent the 2nd and 3rd generation atypical antipsychotics that target the dopamine (D2R) and recently also serotonergic (5HT-R) receptors and primarily serve to treat symptoms (mostly psychotic) and preserve remaining abilities. Thus, it is of utmost urgency to develop an active substance that target other known important receptors, thereby target other symptoms and slow down the course of the disease.

One of the aims of the present invention was thus the selection and development of small peptides which bind with high affinity to DISC1 and/or one of the clinically identified mutant versions of DISC1 and can be used in particular as molecular tracers in the diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders, and/or can be used for the treatment of schizophrenia and other chronic mental disorders.

One object of the present invention was therefore to develop new chemical entities which bind with high affinity and high specificity to DISC1 and/or one of the clinically identified mutant versions of DISC1. Such a high-affinity and highly specific binding to DISC1 and/or one of the clinically identified mutant versions of DISC1 gives the peptides according to the invention a particular suitability for the diagnosis and/or therapy of schizophrenia, MDD, and BD, and autism and other chronic mental disorders.

Another object of the present invention was therefore to develop new chemical entities which bind with a certain affinity to DISC1 and/or one of the clinically identified mutant versions of DISC1. Such a specific affinity to DISC1 and/or one of the clinically identified mutant versions of DISC1 gives the peptides according to the invention a particular suitability for the therapy and/or diagnostics of schizophrenia, MDD, and BD, and autism and other chronic mental disorders.

Another aspect of the present invention was to develop new chemical entities which bind with significantly higher affinity to clinically identified mutant versions of the DISC1 protein associated with chronic mental disorders, than to the wild type DISC1.

The present inventors have characterized 12 peptide ligands originally identified by phage display techniques. The hits were generated against the wild-type

construct of the C-region of DISC1 and the identified leads bind to it as well. More importantly, the peptides bind to 3 clinically identified mutant versions of the DISC1 protein with higher affinity. These mutant versions of the C-region include S704C (observed as aggregate deposits from post-mortem brain), S713E (serves the role of a conformational switch in Bardet–Biedl syndrome) and L807-frameshift (L807-FS; identified in American family suffering from mental disorders). The S704C [2] mutant in particular is very interesting because several reports are available on its aggregation propensity, oligomer specific association with NDEL1, and this mutant has been identified among aggregates in brain tissues of major depressive disorder (MDD) and schizophrenia. The L807-frameshift [3] was first identified in an American family. This mutation results in the replacement of residues 808–854 with nine residues at the C terminus at the binding site for the NDEL1. The mutant protein is prone to severe aggregation and perhaps also may be coupled to loss of the binding to NDEL1 protein. A unique version of the C-region that is susceptible to phosphorylation is S713 [4]. In its phosphorylated state, the full-length DISC1 recruits Bardet-Biedl syndrome (BBS) proteins to the centrosome but in its dephosphorylated state interacts with GSK3 β in the Wnt signalling pathway. This residue acts as switch going from proliferation (dephosphorylated state) to migration (phosphorylated state). It will be interesting to compare the atomistic structural changes between the WT-phospho-C-region against the S713E mutant [4]. This will provide us with valuable clue on the conformational switch associated with DISC1 functioning.

The peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, were found using optimised phage display selection. An example of the phage display selection use is described in the experimental part.

Further preferred embodiments are defined in the dependent claims. In the following, the term “comprise” shall also include “consist of”.

In the conventional phage display technology, a recombinant library of randomised peptide sequences, presented at the gp3 protein of the M13 phage and encoded in its genome, is selected against the L-enantiomeric version of the target protein, i.e. the bait, which in this case is the C-region of DISC1. Thereafter, using the retro-inverso technique, the amino acid sequence of the

identified lead peptides is reversed along with the inversion of chirality from L-enantiomeric amino acid to the D-enantiomer.

In the mirror image phage display, alternatively, the recombinant library of randomised peptide sequences, presented at the gp3 protein of the M13 phage and encoded in its genome, is selected against the exact mirror image (D-enantiomer) of a naturally occurring L-enantiomeric target molecule (e.g. C-region of DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS). The gp3 molecule, also called gene product 3, is a protein that sits on the coat of the phage and is required for contact with the host cell.

In the mirror image phage display, for example, a recombinant library of randomised peptide sequences, presented at the gp3 protein of the M13 phage and encoded in its genome, is selected against the exact mirror image (D-enantiomer) of a naturally occurring L-enantiomeric target molecule (e.g. C-region of DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS). The gp3 molecule, also called gene product 3, is a protein that sits in the phage's coat and is required for contact with the host cell.

The peptide sequence is advantageously presented at the N-terminus of the gp3 protein of the M13 phage and is present encoded in its genome.

The DNA sequence of the p3 gene of a selected phage is linked to the DNA sequence containing the genetic information about the corresponding peptide sequence on the gp3 molecule, allowing it to be sequenced. After sequencing, the genomic sequence can be transcribed into an amino acid sequence and synthesised as a D-enantiomeric peptide that binds to the physiological L-enantiomeric form of the target molecule (e.g. a target molecule). The totality of phages that present different peptides as a fusion protein with gp3 on their surface is referred to below as the phage library. The corresponding peptides represent the biomolecules to be selected in the experiment.

So-called panning rounds (selection rounds) can be carried out, e.g. three rounds. In this process, the phage library is brought into contact with a fixed target molecule, also called bait, and binding phages are isolated from the billion-fold background of other, non-binding phages.

In order to reduce the accumulation of phages with an affinity for plastic, L-arginine and Lysine or streptavidin, different substrate surfaces are preferably used in all panning rounds. In this case, the substrate surface is defined as chemically functionalized or biotinylated substrate used (polystyrene or polypropylene surface) and the blocking or quenching agents used. In the successive selection rounds, the selection pressure is successively increased. For this purpose, while the concentration of the target molecule remains stable, the number of washing steps after phage incubation is continuously increased from the second selection round onwards in order to remove phages that do not have an affinity for a target molecule.

Furthermore, a different substrate surface is selected in each selection round of the phage display by using other agents to block the surface after immobilisation of the target molecule on the substrate (e.g. L-arginine and Lysine, milk powder, no blocking).

By way of example, it is possible to alternate between a BSA-blocked polystyrene surface in round 1, a milk powder-blocked polypropylene surface in round 2 and a BSA-blocked polystyrene surface in round 3.

Switching between different substrate surfaces increases the specificity for the target molecule, or bait, to the surface. There is also a reduction in ligands that bind non-specifically to plastic surfaces, L-arginine and Lysine or other components of the substrate surface other than the bait molecule.

Parallel to the actual phage display selection, exemplary control selections can be carried out, which are identical to the main selection in terms of execution - with the important difference that no bait is used here. A data analysis of the sequences resulting from the control selections enables the identification of peptides that accumulate during the selection even without bait and are therefore irrelevant for all subsequent steps.

The method is thus characterised by the following steps:

- a) Providing an immobilised bait on a substrate surface.
- b) Contacting the immobilised molecule acting as bait with a solution containing a library of molecules to be selected.

- c) Bringing the immobilised baits occupied by the molecules into contact with a washing solution.
- d) Separation and multiplication of the molecules still bound to the bait after
- e) contacting the immobilised bait occupied by the molecules with wash solution.
- f) Repetition of the above steps, using a different substrate in each repetition,
- g) Identification of the sequence of molecules remaining on the baits after repetition.

A different substrate is used, for example, by changing the type of substrate and/or blocking or not blocking it using reagents.

The bait used is a molecule selected from the group consisting of proteins, peptides, RNA, DNA, m-RNA and chemical compounds. In particular, DISC1 C-region (residues 691-836) or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS can be used as bait.

The surface to which the bait is immobilised is, for example, a component from the group consisting of functionalized microtitre plates, magnetic particles, agarose or sepharose beads.

The bait according to point a) is therefore a compound to which the biomolecule to be selected is to be bound. It is fixed to a first surface according to methods known in the prior art. Exemplary, but not limiting, baits may be proteins, peptides, RNA or DNA molecules. Exemplary surfaces that may be used include functionalized microtitre plates, magnetic particles, agarose or sepharose beads. The surface with the immobilised bait can subsequently be quenched, whereby the functional groups of the substrate are inactivated. In addition, blocking of the hydrophobic free areas remaining on the substrate can be carried out with suitable agents.

In the second step b), the immobilised bait is brought into contact with a randomised library of molecules - specifically biomolecules. These biomolecules compete for binding to the bait. The randomised library is a mixture of very many, for example 10^{12} , but also 10^4 or only 100 different molecules in a mixture. Such a library can, for example, consist of peptides, proteins, DNA, RNA or m-RNA, which are bound to certain vehicles and which can bind to bait.

Vehicles may include, for example, phages, polysomes or bacterial surfaces. The library may comprise artificial components or components isolated from nature, or a mixture of both. Artificial in the sense of the invention means, for example, compounds produced from oligonucleotide synthesis.

The immobilised bait with biomolecules can be brought into contact with a washing substance in step c). For this purpose, a washing step is carried out in step c), in which a buffer solution is brought into contact with or rinsed with the immobilised baits. This means that the solution containing the library of biomolecules is preferably repeatedly replaced by a possibly similar or identical solution. Thus, those library molecules are removed which dissociate less rapidly from the immobilised baits than other library molecules. The speed of the dissociation reaction of the binding library molecules is mainly determined by the different dissociation constants (especially the k_{off} values) of the individual molecules. Those with a small k_{off} value statistically remain bound to the immobilised bait the longest and thus have a lower statistical probability of being washed away by the wash buffer. The liquid containing the wash buffer is preferably aqueous and may contain a pH buffer. Optional components of the solution for the specificity wash step may be salts, detergents or reducing agents.

After the specificity washing step in step c), the bound biomolecules are separated from the bait and amplified in step d). The separation or elution can be done, for example, by changing the pH, heating or other changes, especially increasing the salt concentration. Subsequently, the separated or eluted biomolecules are multiplied according to known methods. For this purpose, for example, phage particles obtained and eluted according to steps a) to c), which carry the biomolecules on their surface, can be introduced into cells and multiplied.

In step e), the concentration of the selected biomolecules is increased in the solution added to the bait after step a). Preferably, 3 to 6 selection rounds containing steps a) to e) are performed. However, 1 to 10 or 1 to 20 repetitions may also be carried out. Preferably, the increase in the competitor concentration in step c) with increasing number of cycles also leads to an improved selection. Preferably, an increase in the number of washing steps in step c) with an increasing number of cycles leads to an improved selection.

In this way, the following peptides were developed.

SEQ ID NO 1:	IWAGHIHDWWRRGQGS-NH ₂ (Flm1)
SEQ ID NO 2:	AWQHHELKWKQYQGWR-NH ₂ (Flm2)
SEQ ID NO 3:	GIKSLISNKMHTQWNY-NH ₂ (Flm5)
SEQ ID NO 4:	WNHPMNIRSISWHKDS-NH ₂ (Flm6)
SEQ ID NO 5:	PWYNAPSHSHGFWGSGR-NH ₂ (Flm9)
SEQ ID NO 6:	HKLHHLLKWENQHKVE-NH ₂ (Flm11)
SEQ ID NO 7:	FPHPTLKHREFMRHVS-NH ₂ (Flo1)
SEQ ID NO 8:	HISMQTYKWSQWAKHQ-NH ₂ (Flo3)
SEQ ID NO 9:	QDNPWWHEHKHWKSWP-NH ₂ (Flo5)
SEQ ID NO 10:	SNWHPRHFFMSPSSPQ-NH ₂ (Flo6)
SEQ ID NO 11:	KYTAIHEWYKAHNGGA-NH ₂ (Flo8)
SEQ ID NO 12:	TSGVPWWKFQQQVRRT-NH ₂ (Flo11)

The task according to the invention is also solved by homologues, fragments, combinations, and parts of the amino acid sequence comprising a peptide according to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12.

Sequence processing of the peptides according to the invention (e.g. sequence variation) offers the possibility of developing further substances that can be used therapeutically in the treatment of chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism.

The present invention may also relate to other peptides that can be identified using the method disclosed above.

The peptides according to the invention can be used in diagnostic methods for clinical evaluation, monitoring of disease progression and early differential diagnosis of chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism by specific binding to DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS).

Furthermore, the peptides according to the invention can also be used as therapeutic substances against chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism.

By DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS) is preferably understood here the human DISC1 or the human form of the clinically identified mutant versions S704C, and/or 713E and/or L807-FS).

"Homologous sequences" or "homologues" means, for the purposes of the invention, that an amino acid sequence has an identity with one of the above amino acid sequences of the monomers of at least 50, 55, 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100%. Preferred are 80% and 90%. Instead of the term "identity", the terms "homolog" or "homology" are used synonymously in the present description. The identity between two nucleic acid sequences or 30 polypeptide sequences is determined by comparison using the program BESTFIT based on the algorithm of Smith, T. F. and Waterman, M. S (Adv. Appl. Math. 2: 482-489 (1981)) with the following parameters for amino acids: gap creation penalty: 8 and gap extension penalty: 2; and the following parameters for nucleic acids: gap creation penalty: 50 and gap extension penalty: 3. Preferably, the identity between two nucleic acid sequences or polypeptide sequences is defined by the identity of the nucleic acid sequence/polypeptide sequence over the respective entire sequence length as calculated by comparison using the program GAP based on the algorithm of Needleman, S. B. and Wunsch, C. D. (J. Mol. Biol. 48: 443-453) with the following parameters set for amino acids: Gap creation penalty: 8 and Gap extension penalty: 2; and the following parameters for nucleic acids Gap creation penalty: 50 and Gap extension penalty: 3.

Two amino acid sequences are identical within the meaning of the present invention if they have the same amino acid sequence.

In a further variant, the peptides according to the invention have sequences that differ from the indicated sequences by up to two or three amino acids.

Furthermore, sequences containing the above-mentioned sequences can also be used as peptides.

The peptide according to the invention is further preferably characterised in that an acid amide group (CONH₂ group) or a COH group, COCl group, COBr group,

CONH-alkyl radical or a CONH-alkyl-amine radical is present at the free C-terminus instead of the carboxyl group, or else the peptide is cyclised.

In this way, the task of providing a peptide without a negative charge at the C-terminus is additionally solved in a particularly advantageous way. This has the advantage that it can bind to the target molecule with a higher affinity than a peptide that has a carboxyl group at the free C-terminus. Peptides with a free, unmodified carboxyl group have a negative charge at this end in the physiological state.

In one embodiment of the invention, the peptide according to the invention in the physiological state, in particular at pH 6-8, in particular 6.5-7.5, in particular at pH 6.0, pH 6.1, pH 6.2, pH 6.3, pH 6.4, pH 6.5, pH 6.6, pH 6.7, pH 6.8, pH 6.9, pH 7.0, pH 7.1, pH 7.2, pH 7.3, pH 7.4, pH 7.5, pH 7.6, pH 7.7, pH 7.8, pH 7.9, respectively pH 8.0 are modified in such a way that the C-terminus does not carry a negative charge, but instead is neutral or has one or more positive charges.

In one embodiment, the peptide is characterised in that an acid amide group is present at the free C-terminus instead of the carboxyl group. Instead of the carboxyl group (-COOH group), an acid amide group (-CONH₂ group) is arranged at the C-terminus.

The peptide is thus particularly advantageously amidated at the free C-terminus. The peptide is thus particularly advantageously amidated at the free C-terminus and unmodified at the free N-terminus.

This solves the further problem of providing a peptide without negative excess charge, which can bind more affinely to the target molecule and is easily obtainable.

In a further embodiment, the following further groups are present in place of the carboxyl group: COH, COCl, COBr, CONH-alkyl radical, CONH-alkyl-amine radical (net positive charge), etc, although one is not limited thereto provided that the technical teaching of the main claim is followed.

In a further preferred embodiment of the invention, the affinity of binding of the peptides modified according to the invention without a negative charge at the C-terminus, compared to linear peptides with a negative charge at the C-terminus

but otherwise the same amino acid sequence, is therefore increased by 1%, 2, 3, 4, 5, 6, 7, 8, 9, in particular 10%, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, especially 100%, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, especially 200%, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, in particular 300%, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, in particular 400%, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, advantageously even 500%, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599,

particularly advantageously 600%, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, particularly advantageously 700%, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, also particularly advantageously 800%, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, also particularly advantageously 900%, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, or even increased by 1000%, or even increased by 10000%, or even increased by up to 100000% or 1000000%, any intermediate value being assumed.

This is indicated by a correspondingly lower K_D value. The K_D value as a measure of the affinity of binding of a modified peptide to the target molecule is compared to a linear binding peptide with negative charge at the free C-terminus, by 1%, 2, 99.3, 99.4, 99.5%, 99.6, 99.7, 99.8, 99.9 lowered up to 99.99 or even 99.999%, whereby any intermediate value can be assumed.

The peptide according to the invention is further preferably characterised in that it comprises 2, 3, 4, 5, 6, 7, 8, 9, 10 or more copies of the sequences having SEQ ID NO: 1 or homologues, dimers and portions thereof.

Variants are also conceivable wherein the peptide comprises 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more peptides having SEQ ID NO: 1 or homologues, dimers and portions thereof.

Particularly preferred are dimers of the sequence containing SEQ ID NO: 1 or homologues, dimers and parts thereof.

In one embodiment, the peptide according to the invention is further characterised in that the peptide consists essentially of D-amino acids.

For the purposes of the present invention, the term "consisting essentially of D-enantiomeric amino acids" means that the monomers to be used according to the invention are composed of at least 50%, 55%, 60%, 65%, 70% preferably 75%, 80%, particularly preferably 85%, 90%, 95%, especially 96%, 97%, 98%, 99%, 100% of D-enantiomeric amino acids.

In one embodiment, the peptide according to the invention is further characterised in that the peptide consists essentially of L-amino acids.

For the purposes of the present invention, the term "consisting essentially of L-enantiomeric amino acids" means that the monomers to be used according to the invention are composed of at least 50%, 55%, 60%, 65%, 70% preferably 75%, 80%, particularly preferably 85%, 90%, 95%, especially 96%, 97%, 98%, 99%, 100% of L-enantiomeric amino acids.

The peptide according to the invention is further preferably characterised in that it consists of an amino acid sequence according to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, or homologues, dimers and parts thereof.

The peptide according to the invention is further preferably characterised in that the peptide is linked to a further substance, in particular to a labelling substance,

preferably selected from the group consisting of fluorescent tags, radioactive tags, and isotope labelled groups at various positions in the amino acids.

For the purposes of the invention, the linkage is a chemical bond as defined in Römpp Chemie Lexikon, 9th edition, volume 1, page 650 ff, Georg Thieme Verlag Stuttgart, preferably a principal valence bond, in particular a covalent bond.

In an alternative, the linked substances are medicinal products or active substances, defined according to the Medicinal Products Act §2 or §4 (19). §In an alternative, active substances are therapeutically active substances that are used as medicinally active substances.

In another variant, the substances are compounds that enhance the effect of the peptides.

In another alternative, these are compounds that improve the solubility of the peptides and/or the passage of the blood-brain barrier.

In an alternative, the peptides according to the invention have any combination of at least two or more features of the variants, embodiments and/or alternatives described above.

The peptide according to the invention is further preferably characterised in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, or homologues, dimers and parts thereof are covalently or non-covalently linked to each other.

For the purposes of the invention, a covalent linkage of the peptide units is present if the peptides are linked linearly head to head, tail to tail or head to tail, with or without interposed linkers or linker groups.

The peptide according to the invention is further preferably characterised in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, or homologues, dimers and parts thereof are linked to each other without a linker, i.e. directly to each other, or are linked to each other with a linker group.

A non-covalent linkage in the sense of the invention is present if the peptides are linked to each other, for example, via biotin and streptavidin, in particular streptavidin tetramer.

The peptide according to the invention is further preferably characterised in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, or homologues, dimers and parts thereof are linked together near or branched.

In one variant of the present invention, the peptides may be linked linearly to one another, in particular as described above. In another variant, the peptides are linked together in a branched manner to form the peptide according to the invention.

According to the invention, a branched peptide may be a dendrimer in which the monomers are covalently or non-covalently linked to each other.

Alternatively, the peptides can be linked to a platform molecule (such as PEG or sugar) to form a branched peptide.

Alternatively, combinations of these options are also possible.

In one embodiment of the invention, the affinity of the binding of the peptides is defined by the dissociation constant (K_D value).

The dissociation constant (K_D value) of a peptide according to the invention is thereby advantageously reduced in an advantageous embodiment of the invention. This is associated with improved properties of the peptides according to the invention, such as higher affinity for binding to DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS.

In one variant of the invention, peptides are used which are bound to a DISC1 protein with a dissociation constant (K_D -value) of at most 500 μM , preferably 250, 100, 50 μM , particularly preferably 25, 10, 1 μM , whereby any intermediate value can be assumed.

The peptides according to the invention are further preferably characterised in that they bind to the DISC1 C-region (residues 691-836) protein with a K_D of less than 100 μM , in particular less than 10 μM , in particular 100 μM to 0,01 μM , more preferably 10 μM to 0,05 μM .

In one variant of the invention, peptides are used which are bound to the clinically identified DISC1 C-region (residues 691-836) mutant versions S704C, and/or 713E and/or L807-FS with a dissociation constant (K_D -value) of at most 50 μM , preferably 25, 10, 5 μM , particularly preferably 2,5, 1, 0,1 μM , whereby any intermediate value can be assumed.

The peptides according to the invention are further preferably characterised in that they bind to the clinically identified DISC1 C-region (residues 691-836) mutant versions S704C, and/or 713E and/or L807-FS with a K_D of less than 10 μM , in particular less than 1 μM , in particular 10 μM to 0,001 μM , more preferably 1 μM to 0,005 μM .

The peptides according to the invention are further preferably characterised in that they bind to the clinically identified DISC1 C-region (residues 691-836) mutant versions S704C, and/or 713E and/or L807-FS with a K_D of less than 10 nM, in particular less than 5 nM, in particular 1 nM to 5 pM.

The peptides can be produced, for example, via chemical synthesis or peptide synthesis or by recombinant methods.

The present invention also relates to the peptide described above for use as a probe for the identification, quantitative and/or qualitative determinations of DISC1 or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS.

However, the peptides according to the invention can also be used as therapeutic agents, in particular for the treatment of chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism.

The present invention thus also relates to the peptides described above for use in the treatment of chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism.

Furthermore, the present invention also relates to the use of the above-described peptides as a probe for the identification, quantitative and/or qualitative determination of DISC1 and/or the clinically identified mutant versions S704C, and/or 713E and/or L807-FS, in particular as a tracer in an imaging method for clinical evaluation and/or monitoring of disease progression and/or for the early differential diagnosis of chronic mental disorders, in particular schizophrenia, MDD, and BD, and autism.

Description of the figures

- Fig. 1: A schematic representation of the phage display technique. Two parallel selection rounds were performed to enrich peptide displaying phages during the TS1 and ES1. Direct controls (DC2 and DC3) were utilized as negative controls sequences, which were compared with target selection sequences after NGS sequencing to value which peptides were not target specific.
- Fig. 2: Sequences selected on WT C-region ranked by empty score. The first 10 sequences (Flm, where m = monomer) were further investigated.
- Fig. 3: Sequences selected on WT C-region ranked by empty score. The first 10 sequences (Flo, where o = oligomer) were further investigated.
- Fig. 4: BLI response featuring binding of Flm1 (IWAGHIHDWWRRGQGS-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 5: BLI response featuring binding of Flm2 (AWQHHELKWKQYQGWR-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 6: BLI response featuring binding of Flm5 (GIKSLISNKMHTQWNY-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 7: BLI response featuring binding of Flm6 (WNHPMNIRSISWHKDS-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 8: BLI response featuring binding of Flm9 (PWYNAPHSHGFWGSGR-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.

- Fig. 9: BLI response featuring binding of Flm11 (HKLHLLKWKENQHKVE-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 10: BLI response featuring binding of Flo1 (FPHPTLKHREFMRHVS-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 11: BLI response featuring binding of Flo3 (HISMQTYKWSQWAKHQ-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 12: BLI response featuring binding of Flo5 (QDNPWWHEHKHWKSWP-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 13: BLI response featuring binding of Flo6 (SNWHPRHFFMSPSSPQ-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 14: BLI response featuring binding of Flo8 (KYTAIHEWYKAHNGGA-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 15: BLI response featuring binding of Flo11 (TSGVPWWKFQQQVRRT-NH2) with C-region constructs of DISC1 and clinically identified mutant versions of DISC1.
- Fig. 16: Binding affinities of the 12 different peptide hits with the C-region constructs of DISC1 and clinically identified mutant versions of DISC1.

Material and Methods

Protein expression and purification. The wild-type (WT) human DISC1 C-region (residues 691-836) was expressed alternatively as N-terminal His₆ - fusion construct, which was expressed as previously described [5] using the E. coli BL21 (DE3) pLysE T1R cells. The culture was grown either in Luria-Bertani (LB) broth

and protein expression was induced at an OD₆₀₀ of 0.6 by the addition of 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) for 16 h at 18 °C. Harvested cells were lysed using an ice-chilled microfluidizer M100P (Microfluidics MPT) at 15,000 PSI in Tris-buffered saline (TBS) buffer containing 10 mM Tris-HCl pH 7.4, 150 mM NaCl and Complete EDTA- free protease inhibitor cocktail (Roche). The soluble fraction was purified on Co²⁺-NTA resin (Qiagen) and was eluted using TBS containing 500 mM imidazole. In the next step, the WT C-region protein was further purified on a HiLoad 16/60 Superdex 200 size exclusion chromatography column (GE Healthcare Bio Sciences).

Phage display technique for identification of peptide hits. Freshly prepared 100 μ g/ml of the monomeric and 15 mg/ml of a week-old oligomeric WT proteins were used as the targets for the phage display selection included the monomeric and the oligomeric versions of the above-mentioned protein, which was immobilized on 3D-NHS Matrix plates (PolyAn molecular surface engineering). During the phage display selection experiment multiple control selections were performed to ensure the quality of the selection. From the library (TriCo16, BioLabs) two starting selections were performed. The first, called TS1 for target selection 1, used a surface that contained the target, while the second one, called ES1 for empty selection 1, did not contain the target on the surface and served as a control. The empty selection was performed in parallel to the target selection. In the first round of selection, the wells were blocked with 22 mg/ml of L-Arginine.

The phages derived from TS1 were later applied to the next target selection (TS2) as well as the direct control (DC2) which did not contain the target protein on the surface and served as a negative control for direct comparison. In the next round of selection, the wells were blocked with 15 mg/ml of L-Lysine. A subsequent third selection round was performed analogously and the wells were blocked with 22 mg/ml of L-Arginine. Figure 1 illustrates this course of the selection and at which points a control selection had been performed which subsequently enables the identification of target specific binders via next-generation sequencing (NGS) results. The different selection types were named "target selection" (TS), "direct control" (DC) and "empty selection" (ES). Next generation sequencing (NGS) of the phage DNA was performed by the Genomics & Transcriptomics Laboratory of the Heinrich-Heine-Universität Düsseldorf. For NGS all target selection inputs as well as direct control input three and empty

selection input three were prepared and sequenced, this included a sample of the library.

Data analysis for identification of peptide hits. For the initial analysis and grouping of the NGS data using an in-house developed program TSAT (Target Sequence Analysis Tool). A secondary analysis was performed using the clustering software Hammock [6]. The program TSAT takes raw DNA data in a Fasta or Fastq format as an input and examines this data for the provided framing regions of the randomized peptide region. If the framing regions are present in the given DNA the randomized region coding for the peptide is excised, grouped, translated into amino acids and saved in a database. The saved sequences are further processed by filtering, sequence frequency of the last target selection round is analyzed for their enrichment from target round to target round as well as compared to empty selection and direct control. The filtering is performed in the following order:

TS3 ≥ ES3 & TS3 ≥ DC3 & TS3 ≥ TS2 & TS2 ≥ ES2 & TS2 ≥ DC2 & TS2 ≥ TS1 & TS1 ≥ ES1 & TS1 ≥ Library

Each sequence is assigned an enrichment score and an empty score. These are based on the ratio between the last target selection round, the library and the last empty selection round. The scores are calculated according to the following formulas:

$$\text{Enrichment score} = \frac{(\text{Frequency of a sequence in TS3})}{(\text{Frequency of the same sequence in library})}$$

$$\text{Empty score} = \frac{(\text{Frequency of a sequence in TS3})}{(\text{Frequency of the same sequence in ES3})}$$

Bio-layer interferometry (BLI) for identification of peptide leads. Binding of the peptides with the C-region constructs (WT, S713E, S704C and L807-FS) were investigated using an Octet (ForteBio) instrument. Prior to measurements, all peptides were solubilized in phosphate buffered saline (PBS) containing 20 mM Na-phosphate, pH 7.4, and 150 mM NaCl. For coupling of the peptide (ligand), the A2RG sensor chip (Sartorius GmbH) were activated with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/N hydroxysuccinimide (NHS) (200

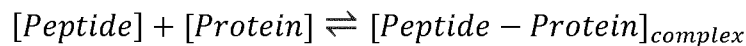
mM/50 mM); the peptides were diluted to 2 μ M in 10 mM sodium acetate, pH 4.0, and immobilized to the sensor chip, and the flow cell was deactivated with 0.1 M ethanolamine-HCl, pH 8.5. Titrations of the above-mentioned C-region and the constructs were performed at 25 °C with TBS, pH 7.4 as the running buffer using protein concentrations of 2 μ M without intermittent regeneration.

Data analysis of peptide-protein interaction. In order to assess the binding isotherms of the protein to the immobilized peptides, I derived the binding affinity (K_D) by computing the on-rate (K_{on}) and the off-rate (K_{off}) by deriving the reciprocals of the exponential fittings of association and dissociation (OriginPro 2020).

$$\begin{aligned} \text{Association time} \rightarrow y \\ = y_0 + A1 * (1 - \exp(-\frac{x}{t1})) + A2 * (1 - \exp(-\frac{x}{t2})) \end{aligned}$$

$$\text{Dissociation time} \rightarrow y = y_0 + A1 * \exp(-\frac{x}{t1}) + A2 * \exp(-\frac{x}{t2})$$

The basic reaction describing the association/dissociation of the peptide-protein interaction can be represented as



$$K_D = \frac{[Peptide - Protein]_{complex}}{[Peptide][Protein]} = \frac{k_{off}}{k_{on}}$$

The binding affinity (K_D) was derived by dividing the off-rate (K_{off}) with the on-rate (K_{on}).

Results

Peptide selection and next generation sequencing. Our activity pertains identifying therapeutics for the protein risk factor associated in schizophrenia *viz.* Disrupted in Schizophrenia 1 (DISC1) using phage display technology (Figure 1). Most forms of therapy that exist today are empirical and prescribed on a trial-and-error basis. They represent the 2nd and 3rd generation atypical antipsychotics that target the Dopamine (D2R) and recently also serotonergic (5HT-R) receptors and primarily serve to treat symptoms (mostly psychotic) and preserve remaining abilities. Therefore, it is essential to find atypical binding partners of the monomeric and oligomeric atypical target proteins. DISC1 mutant proteins have been clinically identified in the families as well as post-mortem tissues of schizophrenic and MDD patients [7]. Recently our group has demonstrated the atomistic details of the architecture of the oligomers and the fibrils, which show remarkable resemblance to other proteinopathies in neurodegenerative diseases.

To this end, we identified several binding partners (16mer peptides) for the monomeric (Figure 2) and the oligomeric (Figure 3) target WT C-region of the DISC1 protein. To identify the peptide hits, NGS was performed, which has the ability to sequence millions of phage genomes coding for peptides presented on the phages. As mentioned above, using TSAT and HAMMOCK several rounds of filtering was performed and compared with the empty selection and direct control. Sequences that were enriched from one selection round to another and were more frequent than ES and DC pass the filter. Finally, the empty and enrichment scores (see methodology, Figures 2 and 3) were calculated for the filtered sequences and the peptides were rated largely on the basis of their empty and enrichment scores as well as the size of the generated clusters. Overall 10 sequences were identified against the monomeric and oligomeric targets, respectively. These sequences exhibited a high enrichment and empty scores and formed large cluster of similar sequences following Hammock analysis. Only sequences with an empty and enrichment score of above 1,000 and 8,000 for the monomeric and oligomeric targets, respectively were considered for further testing using BLI to study peptide-protein interactions.

Protein-peptide interactions. The interaction of WT C-region of DISC1 was studied using BLI, where the identified 16mer peptides were immobilized to the sensor chip. I chose to study three different mutant preparations of the DISC1 C-

region along with *viz.* S713E, S704C and L807-FS, in order to assess potential of the peptides for identifying clinically relevant targets. From this set of experiments, we identified 12 potential peptides that bind to the C-region and the mutant versions of the protein:

Flm1 -> IWAGHIHDWWRRGQGS-NH2
Flm2 -> AWQHHELKWKQYQGWR-NH2
Flm5 -> GIKSLISNKMHTQWNY-NH2
Flm6 -> WNHPMNIRSISWHKDS-NH2
Flm9 -> PWYNAPHSFGFWGSGR-NH2
Flm11 -> HKLHHLLKWENQHKVE-NH2
Flo1 -> FPHPTLKHREFMRHVS-NH2
Flo3 -> HISMQTYKWSQWAKHQ-NH2
Flo5 -> QDNPWWHEHKHWKSWP-NH2
Flo6 -> SNWHPRHFFMSPSSPQ-NH2
Flo8 -> KYTAIHEWYKAHNGGA-NH2
Flo11 -> TSGVPWWKFQQQVRRT-NH2

The binding profiles of the peptides to the C-region constructs are shown in the Figures 4-15 for Flm1 (Figure 4), Flo1 (Figure 5), Flm5 (Figure 6), Flm6 (Figure 7), Flm9 (Figure 8), Flm11 (Figure 9), Flo1 (Figure 10), Flo3 (Figure 11), Flo5 (Figure 12), Flo6 (Figure 13), Flo8 (Figure 14), and Flo11 (Figure 15). Most of the peptides displayed low μM affinity for the WT C-region and nM affinities for the mutant versions (Table 1 and Figure 16). The lowest K_D values were observed for the L807-FS mutant.

	Flm6 (K _D μM)	Flo3 (K _D μM)	Flo1 (K _D μM)	Flm1 (K _D μM)	Flo8 (K _D μM)	Flo11 (K _D μM)
S713E	1.07	0.42	0.07	0.08	0.46	0.89
S704C	2.24	0.27	0.16	0.08	1.04	1.06
L807-FS	0.53	0.40	0.04	0.20	0.17	0.16
WT	2.01	2.62	0.10	0.15	2.62	3.68

	Flo6 (K _D μM)	Flm9 (K _D μM)	Flm5 (K _D μM)	Flm2 (K _D μM)	Flm11 (K _D μM)	Flo5 (K _D μM)
S713E	0.46	1.50	0.08	0.22	0.42	0.21
S704C	0.72	1.49	0.15	0.15	0.72	0.36
L807-FS	0.17	0.39	0.11	0.11	0.54	0.27
WT	2.36	3.20	0.64	0.64	1.27	0.25

Table 1: Summary of binding constants

Efficacy. The results suggest that the identified peptides bind to the WT C-region of the DISC1 protein with an affinity in the low micromolar range. This is critical because the peptides can be used as therapeutic agents against DISC1, which is a scaffold protein that interacts with over 300 different proteins and is required for several physiological process. If the binding affinity is too high, then it may prevent other interaction proteins to bind and lead to undesirable effects. Hence, the peptides can serve as a therapeutic agent against aggregation of the WT DISC1 protein.

On the other hand, most of the peptides display 10-fold and higher affinity for the mutant versions of the protein. This provides a facile diagnostic tool for identification of DISC1 related schizophrenia. As the entire diagnosis and therapy structure in schizophrenia and related chronic mental illnesses is empirical, identifying DISC1 anomaly would be critical for further treatment. The solution presented here is based on a risk factor biomarker i.e. the DISC1 protein, which has been observed as aggregated protein in 20-25% of post-mortem brain tissues [7] from samples involving schizophrenia, MDD and BD. One such application would involve an affinity chromatography procedure wherein a peptide among the 12 is bound to the column and the DISC1 mutant is isolated from biological fluid (blood and/or CSF) sample.

Conclusion

The itinerary of the current invention is two-pronged for the various lead peptides that we have identified. On one hand the high affinity peptides can be used to capture DISC1 proteins from blood and/or CSF as a diagnostic marker.

References

1. Cukkemane A, Becker N, Zielinski M, Frieg B, Lakomek NA, Heise H, Schröder GF, Willbold D, Weiergräber OH (2021) Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl. Psychiatry* 11:639
2. Hashimoto R, Numakawa T, Ohnishi T, Kumamaru E, Yagasaki Y, Ishimoto T, Mori T, Nemoto K, Adachi N, Izumi A, Chiba S, Noguchi H, Suzuki T, Iwata N, Ozaki N, Taguchi T, Kamiya A, Kosuga A, Tatsumi M, Kamijima K, Weinberger DR, Sawa A, Kunugi H (2006) Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Hum Mol Genet.* 15:3024-3033.
3. Sachs NA, Sawa A, Holmes SE, Ross CA, DeLisi LE, Margolis RL (2005) A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Mol Psychiatry.* 10:758-764.
4. Ishizuka K, Kamiya A, Oh EC, Kanki H, Seshadri S, Robinson JF, Murdoch H, Dunlop AJ, Kubo K-I, Furukori K, Huang B, Zeledon M, Hayashi-Takagi A, Okano H, Nakajima K, Houslay MD, Katsanis N, Sawa A (2011) DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature* 473:92–96.
5. Yerabham ASK, Mas PJ, Decker C, Soares DC, Weiergräber OH, Nagel-Steger L, Willbold D, Hart DJ, Bradshaw NJ and Korth C (2017) A structural organization for the Disrupted in Schizophrenia 1 protein, identified by high-throughput screening, reveals distinctly folded regions, which are bisected by mental illness-related mutations. *J. Biol. Chem.* 292:6468-6477.
6. Krejci A, Hupp TR, Lexa M, Vojtesek B and Muller P (2016) Hammock: a hidden Markov model-based peptide clustering algorithm to identify protein-interaction consensus motifs in large datasets. *Bioinformatics.* 32:9-16.
7. Korth C. Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion.* 2012;6(2):134-41.

CLAIMS

1. A peptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, and homologues, combinations, fragments and portions thereof.
2. A peptide according to claim 1, characterized in that the peptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and/or SEQ ID NO: 12, or homologues having an identity of at least 80% thereof.
3. A peptide according to any one of the preceding claims, characterised in that an acid amide group (CONH₂ group) or a COH group, COCl group, COBr group, CONH-alkyl radical or a CONH-alkyl-amine radical is present at the free C-terminus instead of the carboxyl group, or else the peptide is cyclised.
4. A peptide according to any one of the preceding claims, characterized in that it contains 2, 3, 4, 5, 6, 7, 8, 9, 10 or more copies of the sequences having SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12.
5. A peptide according to any one of the preceding claims, characterised in that the peptide consists essentially of D-amino acids.
6. A peptide according to any one of the preceding claims, characterized in that the peptide consists of an amino acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12.

7. A peptide according to any one of the preceding claims, characterized in that the peptide is linked to a further substance, in particular to a labelling substance, preferably selected from the group consisting of fluorescent tags, radioactive tags, and isotope labelled groups at various positions in the amino acids.
8. A peptide according to any one of the preceding claims, characterized in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12 are covalently or non-covalently linked to each other.
9. A peptide according to any one of the preceding claims, characterized in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12 are linked to each other without a linker, i.e. directly to each other, or are linked to each other with a linker group.
10. A peptide according to any one of the preceding claims, characterized in that several peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12 are linked together in a linear or branched manner.
11. A peptide according to any one of the preceding claims, characterized in that it is a dendrimer, wherein peptides of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12 are linked to a platform molecule.
12. A peptide according to any of the preceding claims, characterized in that it binds to the DISC1 C-region (residues 691-836) protein with a K_D of less than 100 μM , in particular less than 10 μM , in particular 100 μM to 0,01 μM , more preferably 10 μM to 0,05 μM .
13. A peptide according to any of the preceding claims, characterized in that it binds to the clinically identified mutant version of the DISC1 C-region (residues 691-836) protein S704C, and/or the clinically identified mutant

version of the DISC1 C-region (residues 691-836) protein L807-FS, and/or the clinically identified DISC1 C-region (residues 691-836) mutant version S704C with an affinity of less than 10 μ M, in particular less than 1 μ M, in particular 10 μ M to 0,001 μ M, more preferably 1 μ M to 0,005 μ M.

14. A peptide according to any one of the preceding claims for use as a probe for the identification, quantitative and/or qualitative determination of DISC1 and/or clinically identified mutant version of the DISC1 protein S704C, and/or the clinically identified mutant version of the DISC1 protein L807-FS, and/or the clinically identified DISC1 mutant version S704C.
15. A peptide according to at least one of claims 1 to 13 for use in the treatment of chronic mental disorders, in particular schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism.
16. Use of a peptide according to at least one of the preceding claims 1 to 13 as a probe for the identification, quantitative and/or qualitative determination of the DISC1 protein S704C, and/or the clinically identified mutant version of the DISC1 protein L807-FS, and/or the clinically identified DISC1 mutant version S704C in particular as a tracer in an imaging method for clinical evaluation and/or monitoring of disease progression and/or for the early differential diagnosis of chronic mental disorders, in particular schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism.

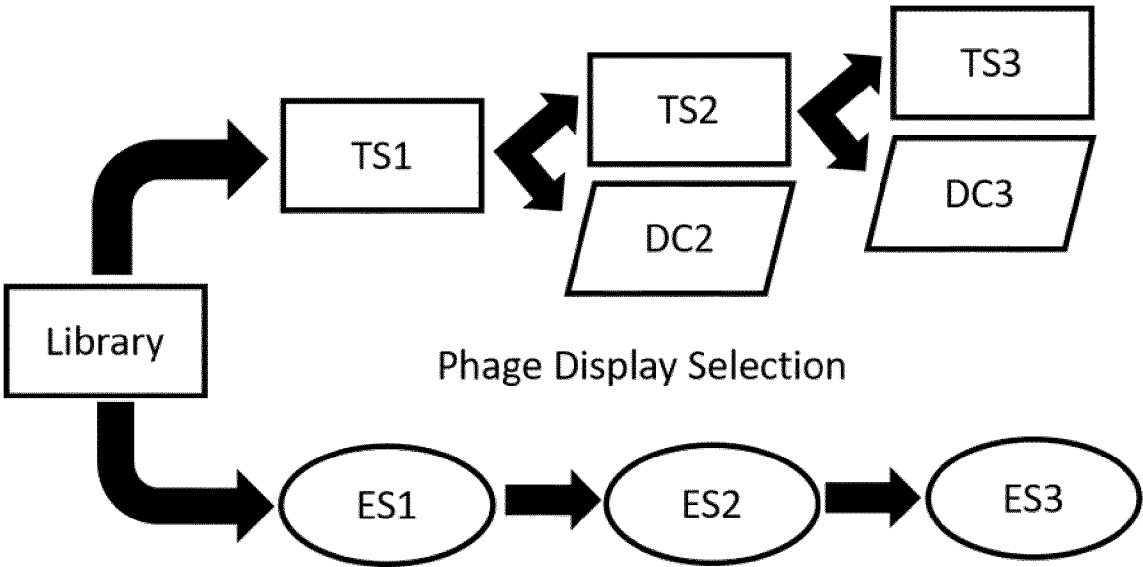


Fig. 1

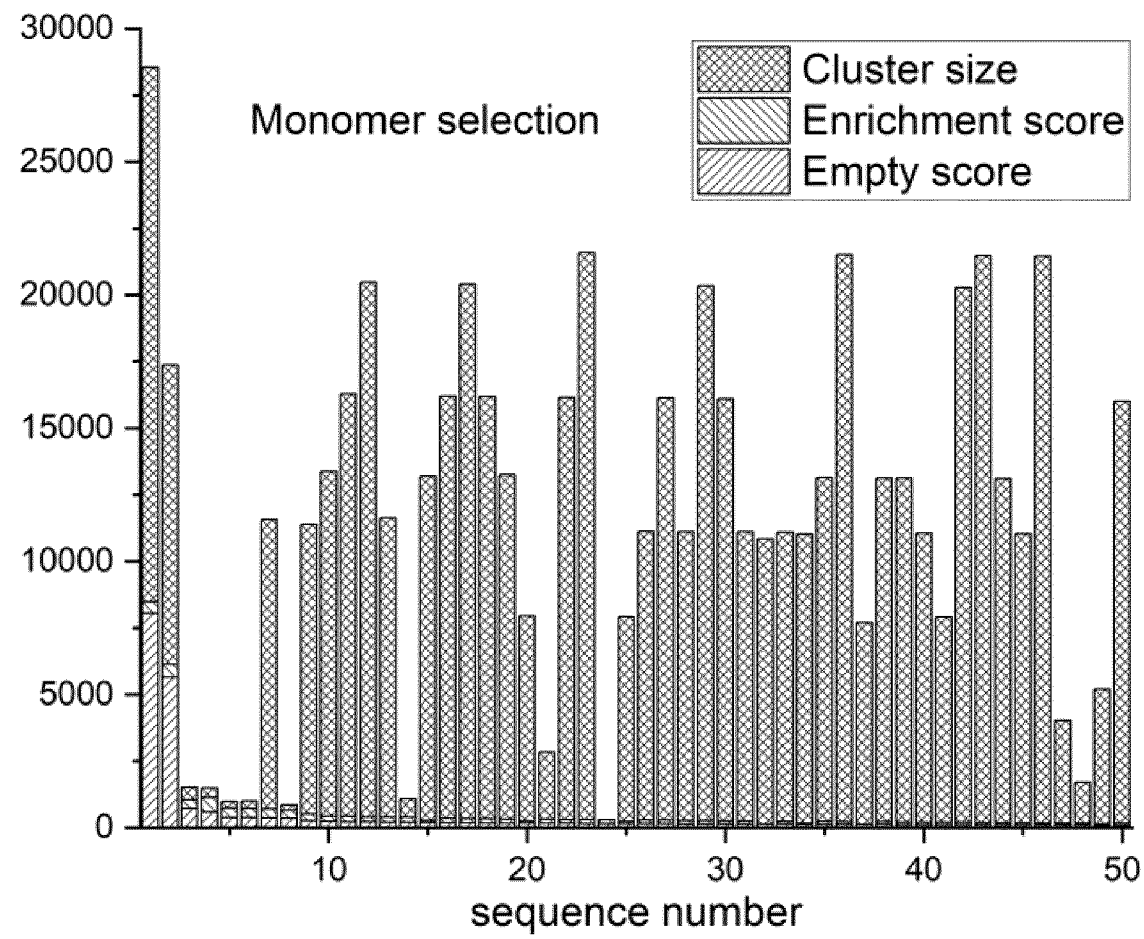


Fig. 2

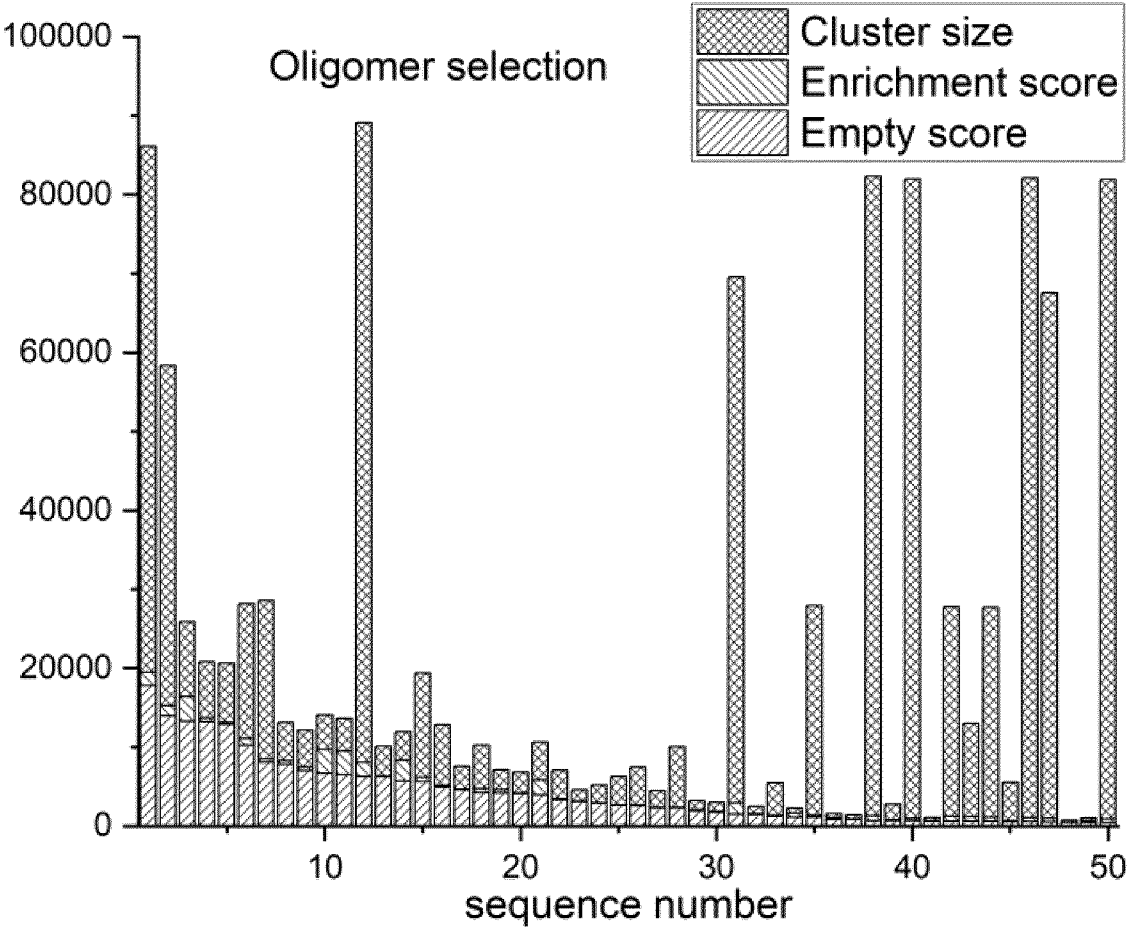


Fig. 3

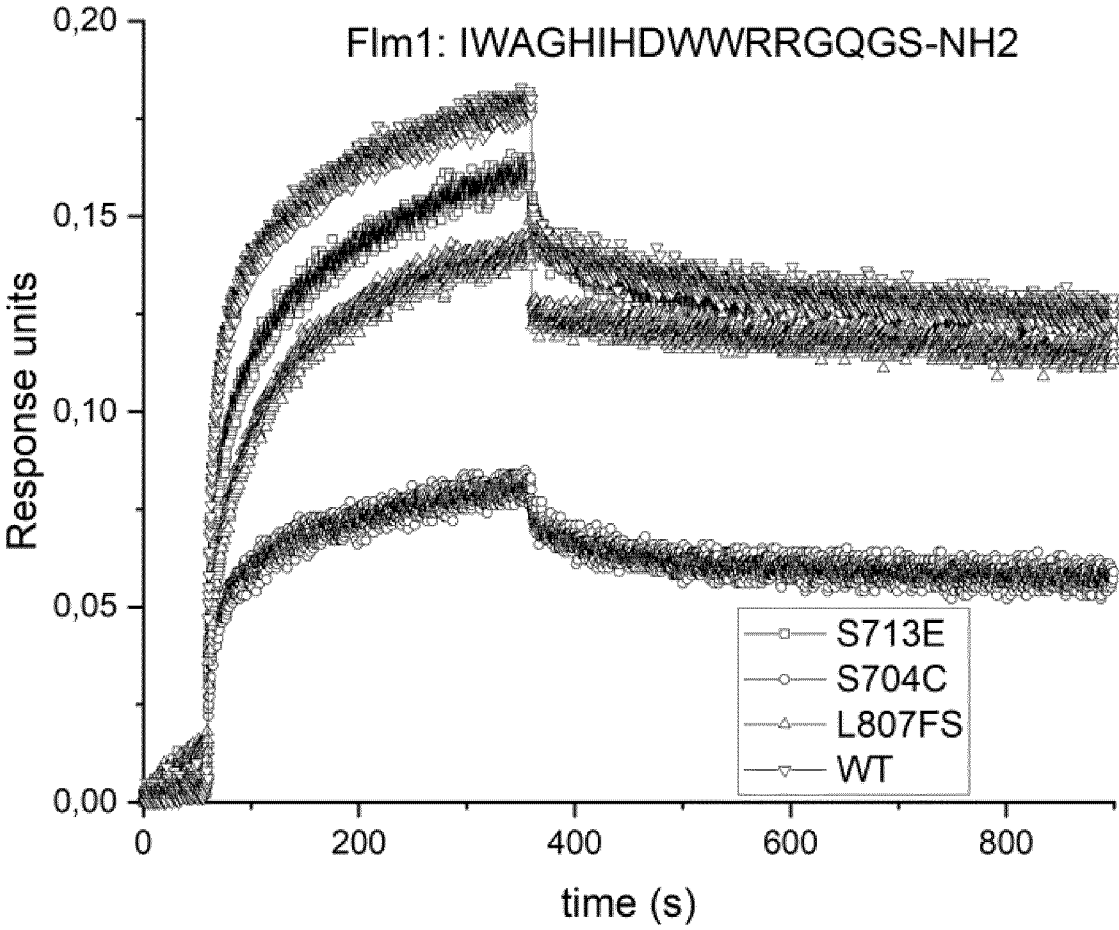


Fig. 4

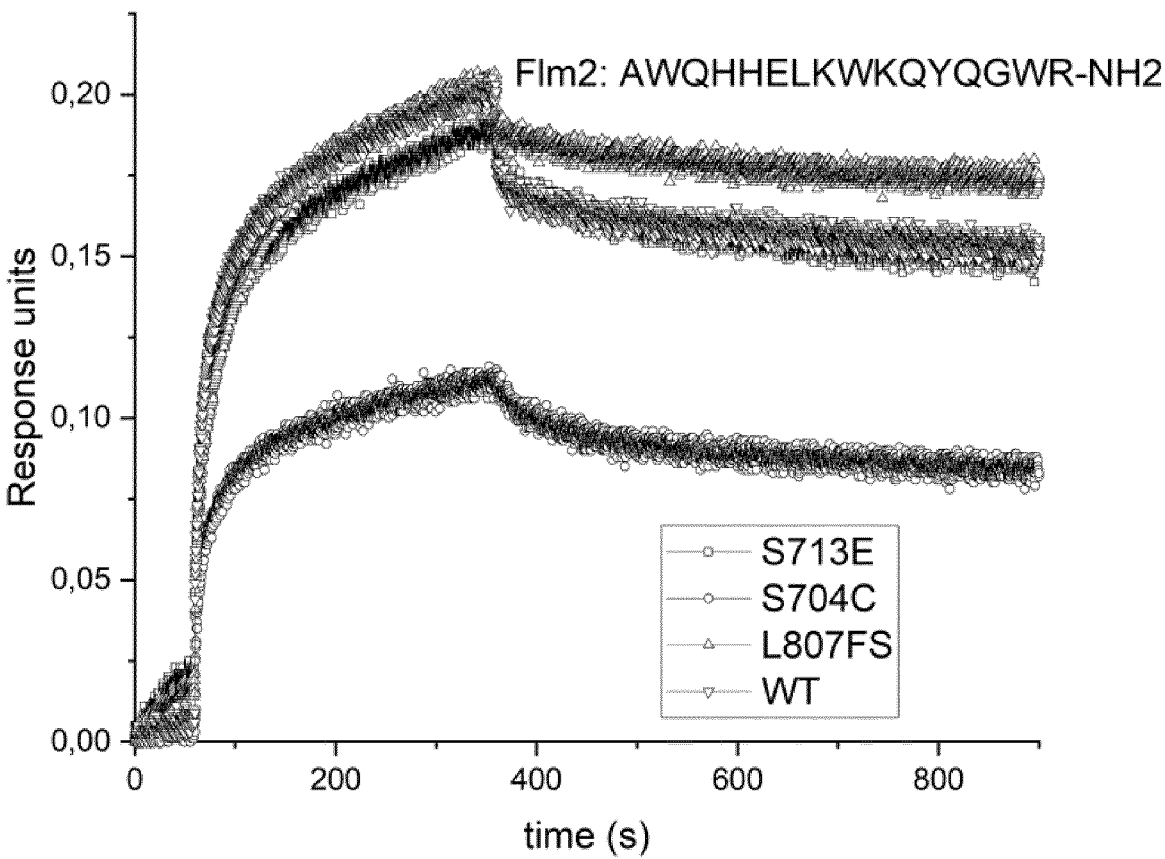


Fig. 5

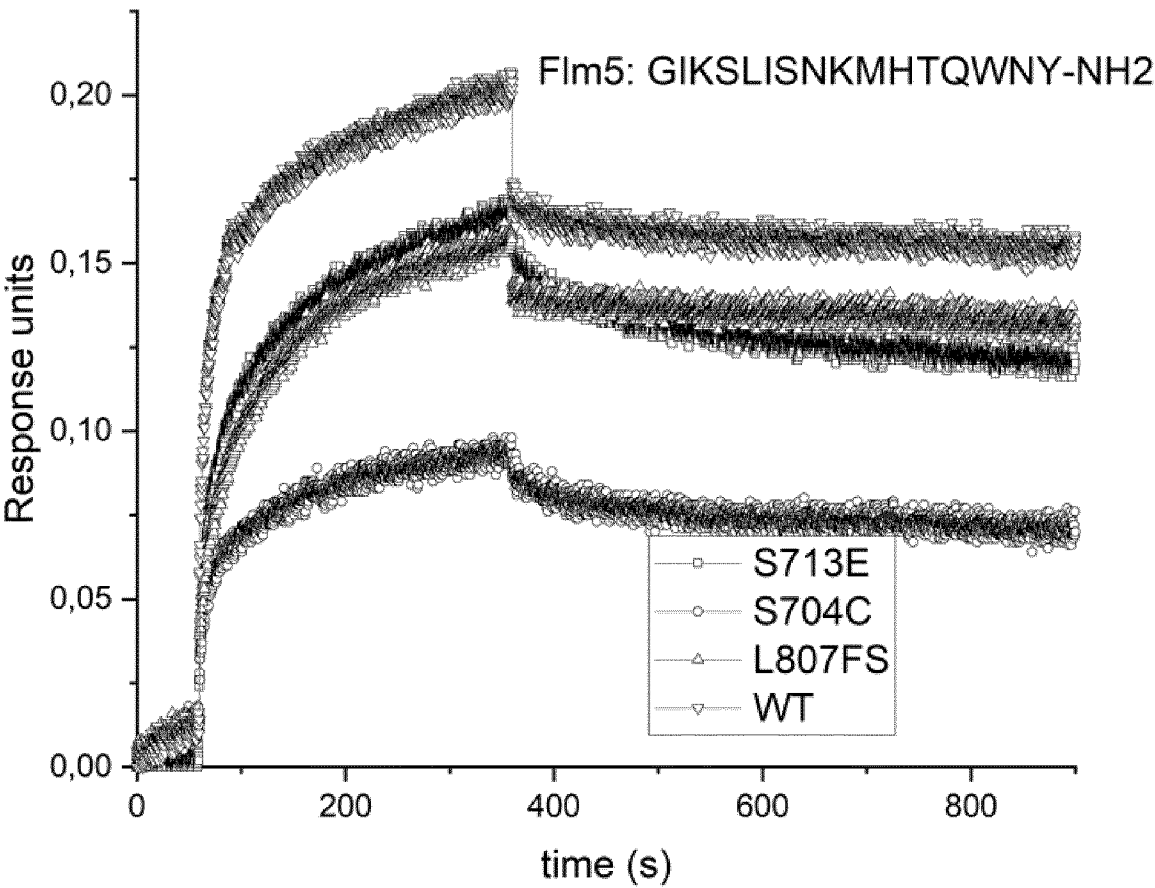


Fig. 6

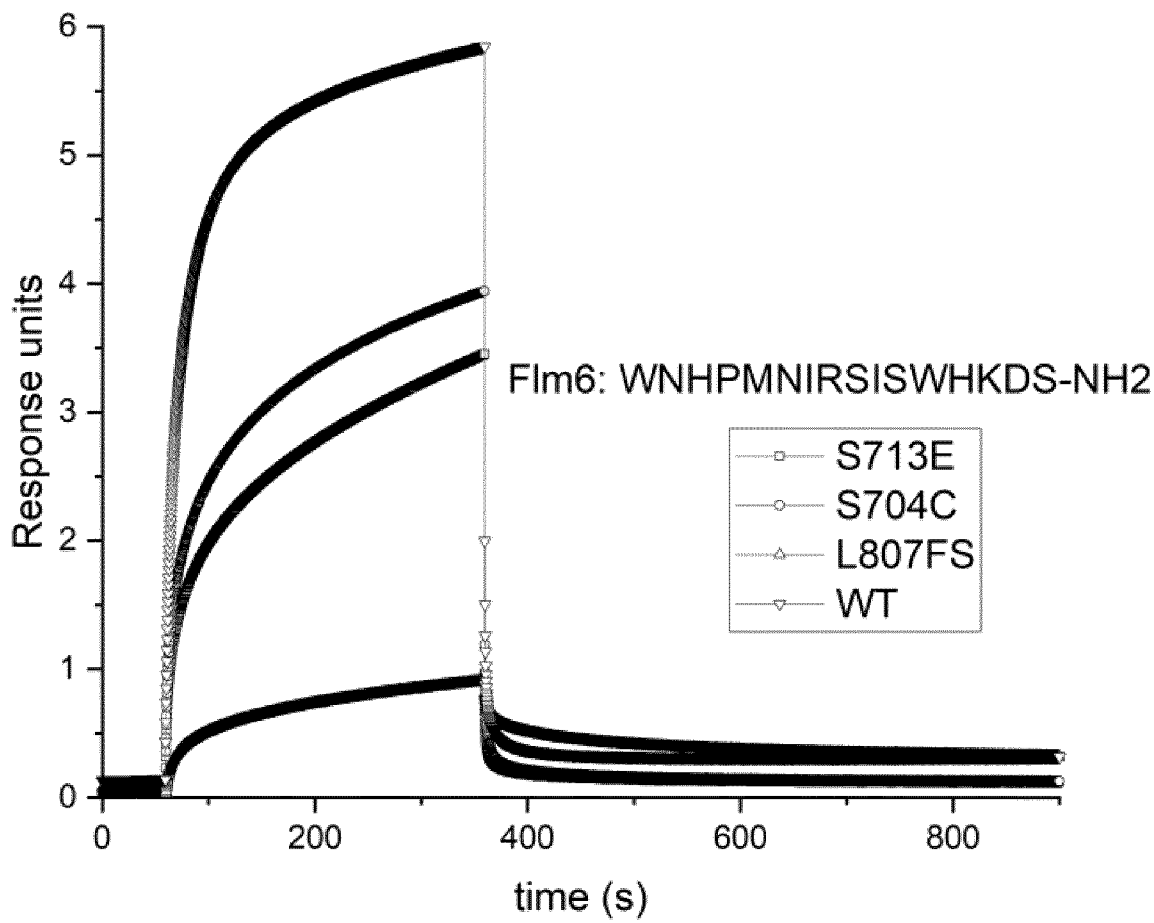


Fig. 7

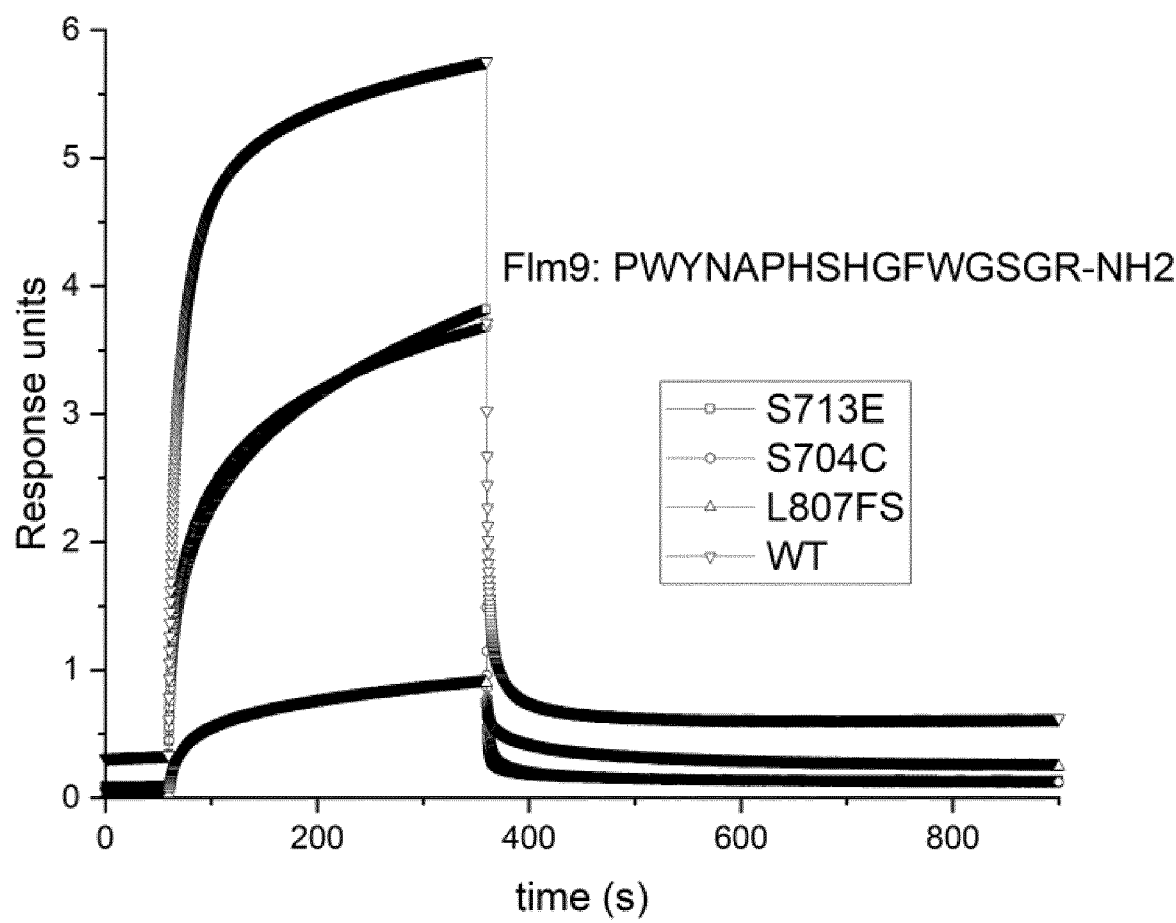


Fig. 8

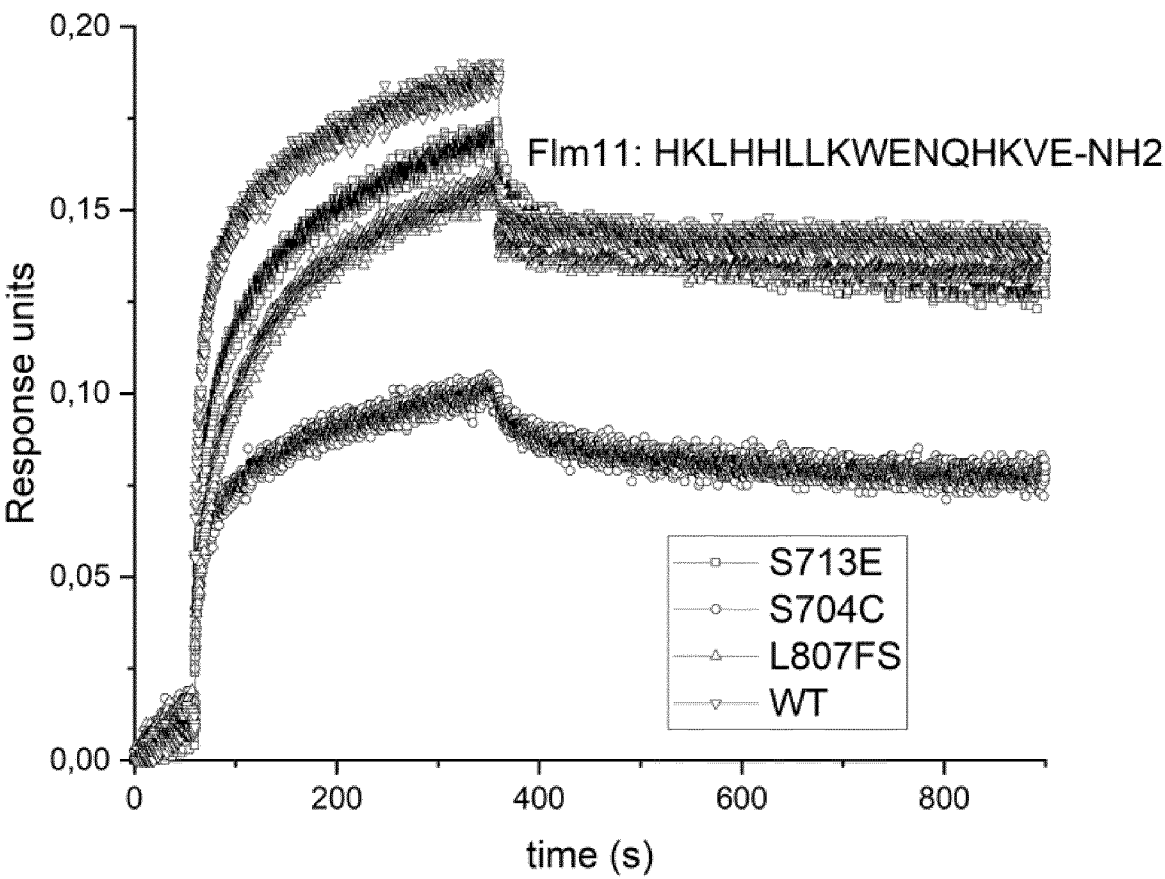


Fig. 9

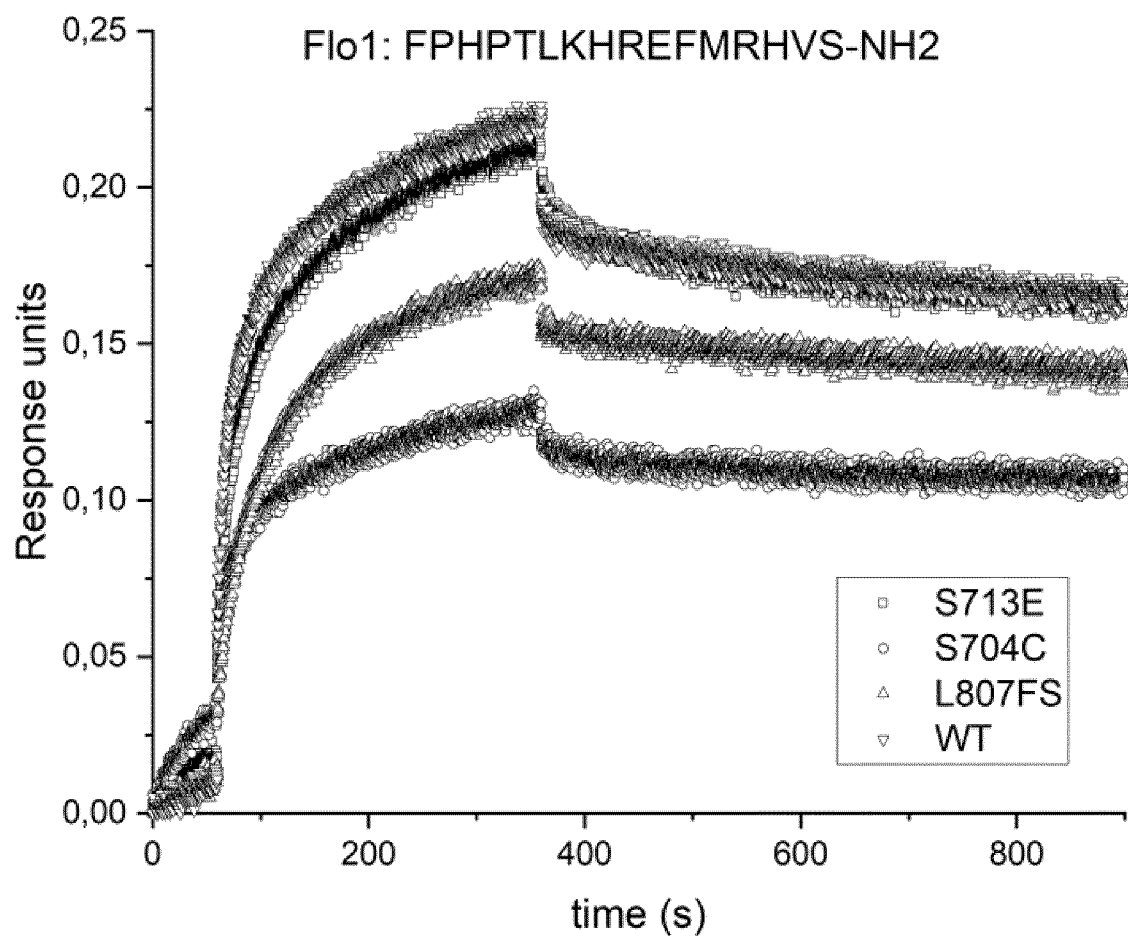
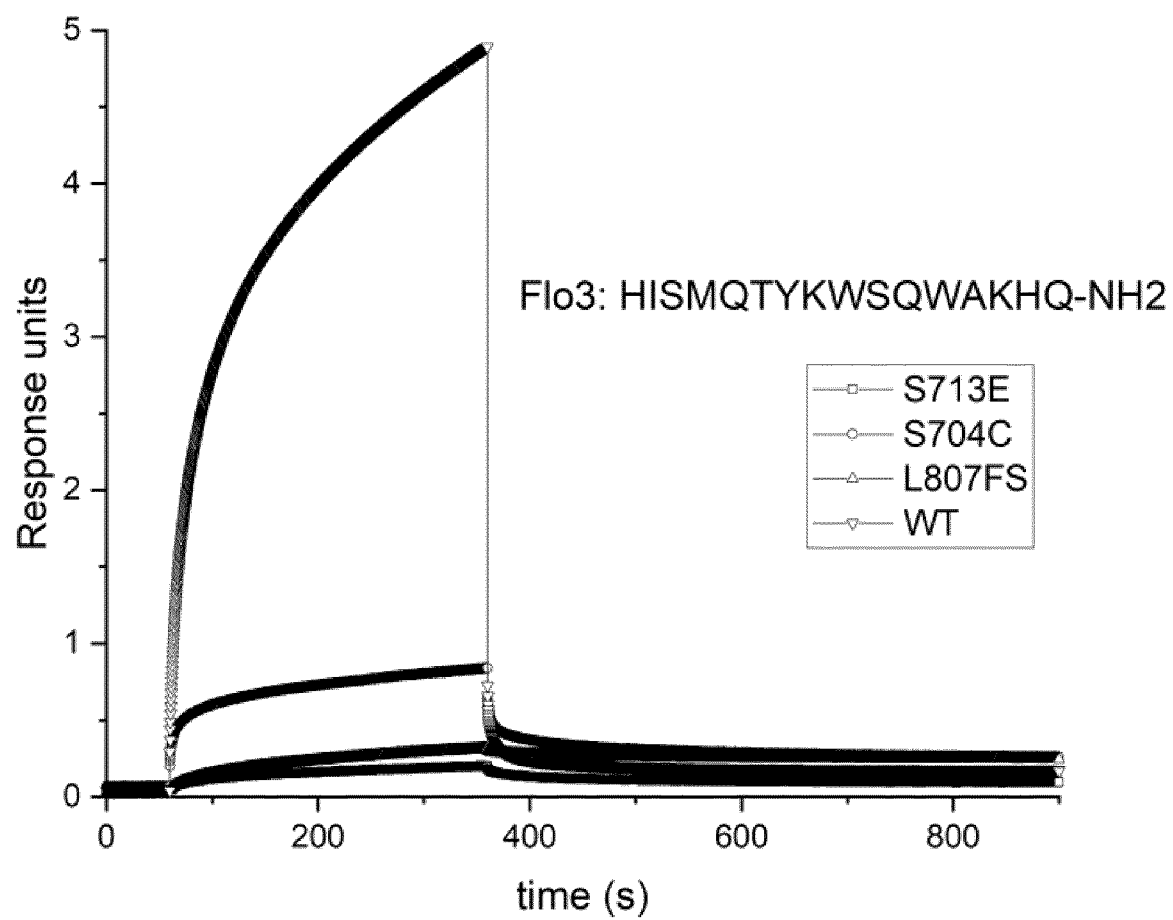


Fig. 10

**Fig. 11**

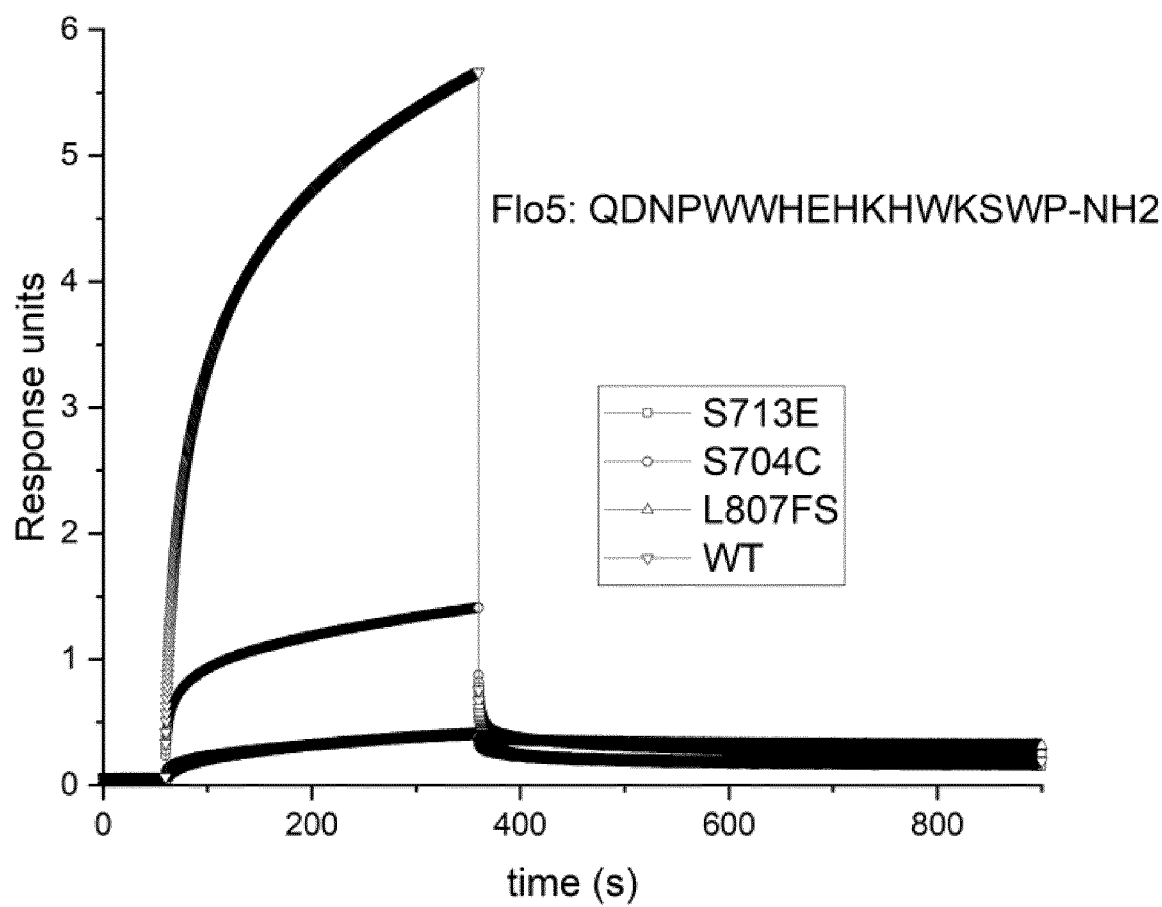


Fig. 12

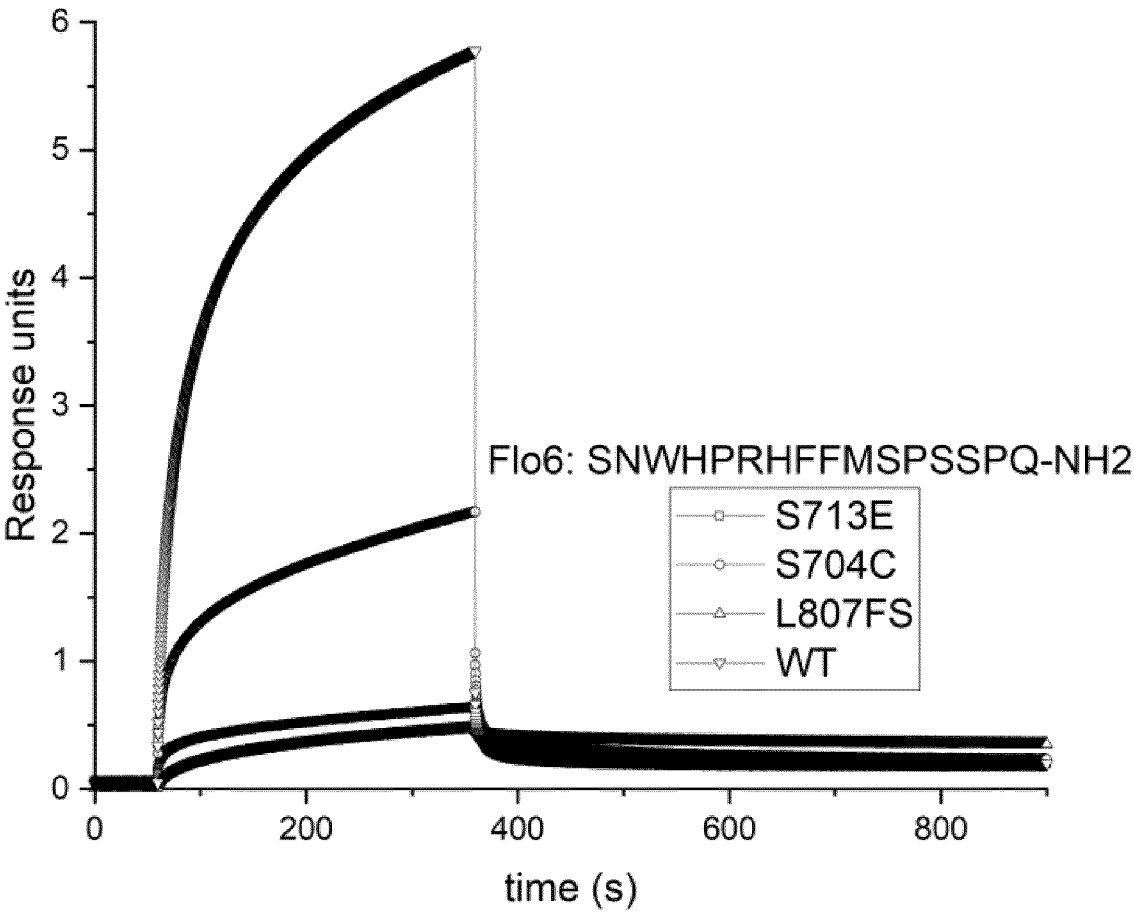


Fig. 13

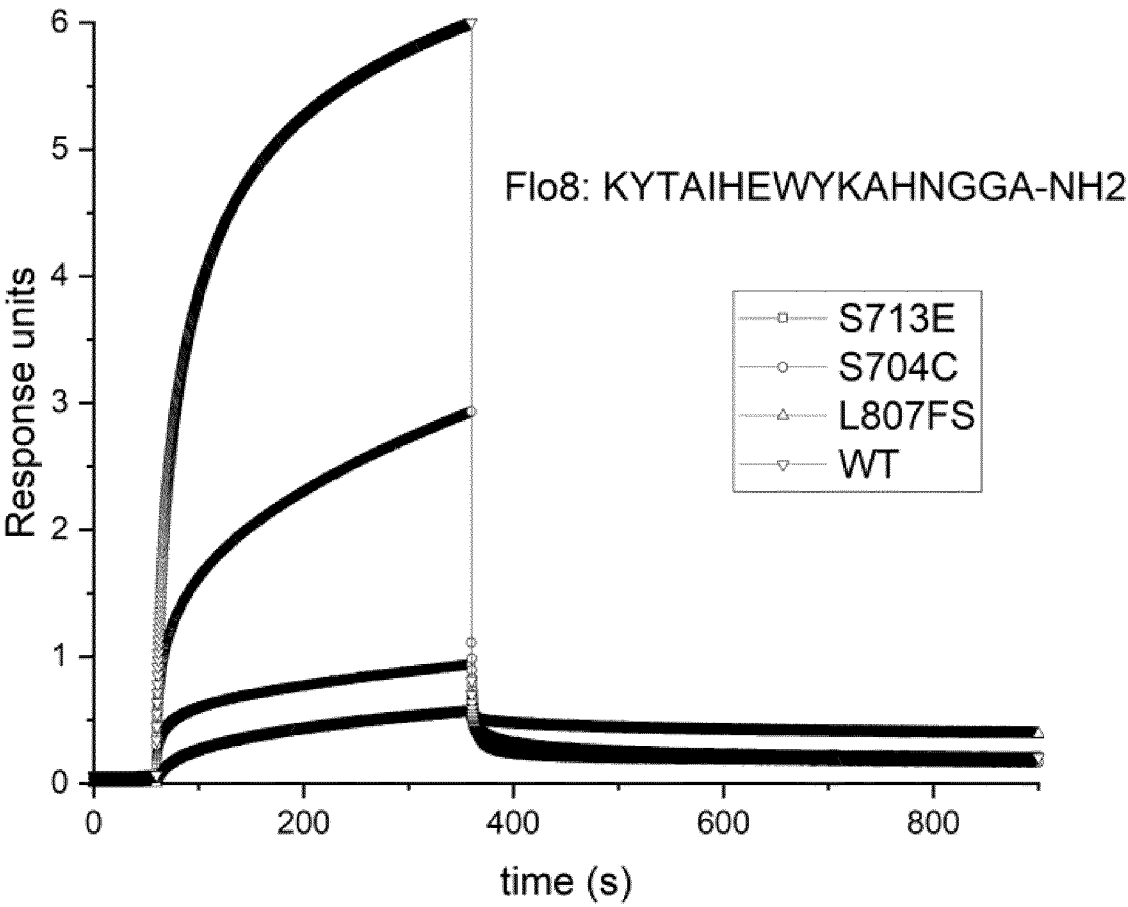


Fig. 14

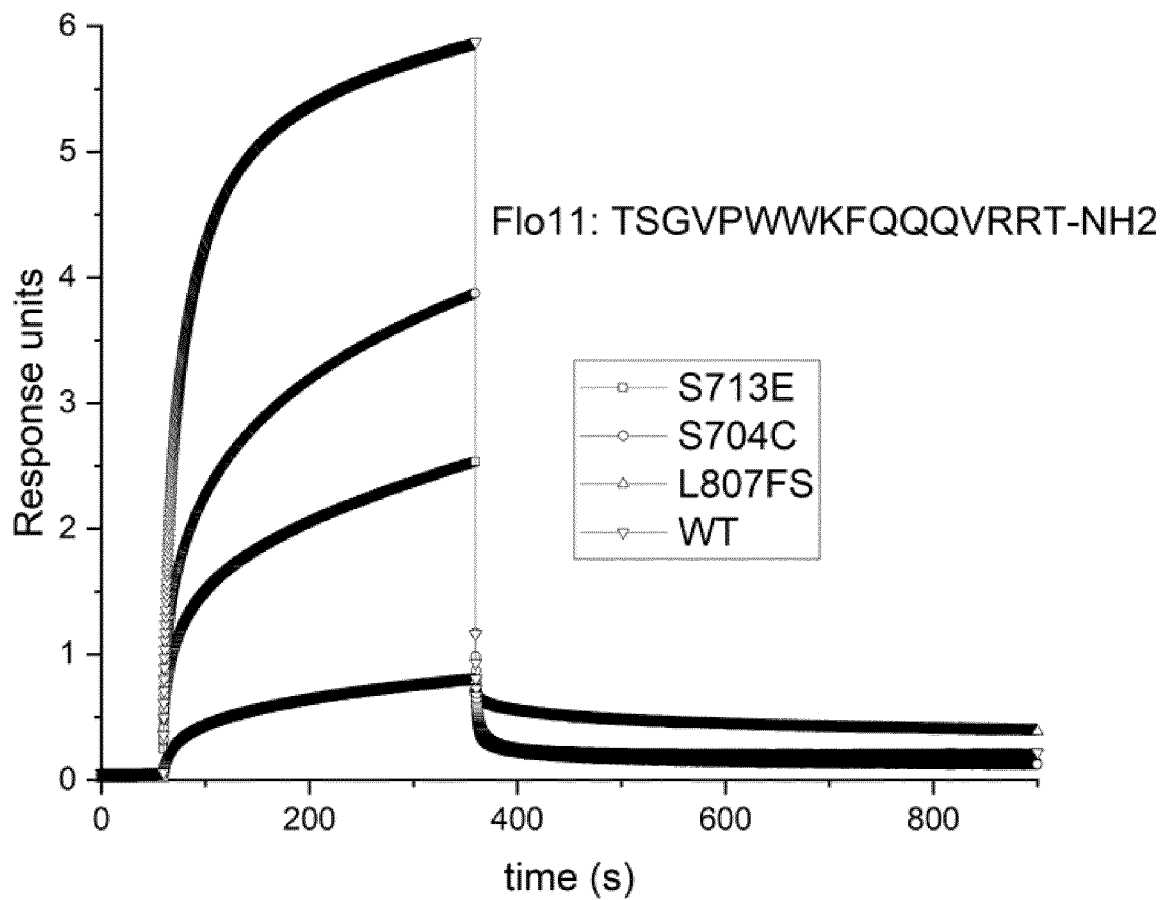


Fig. 15

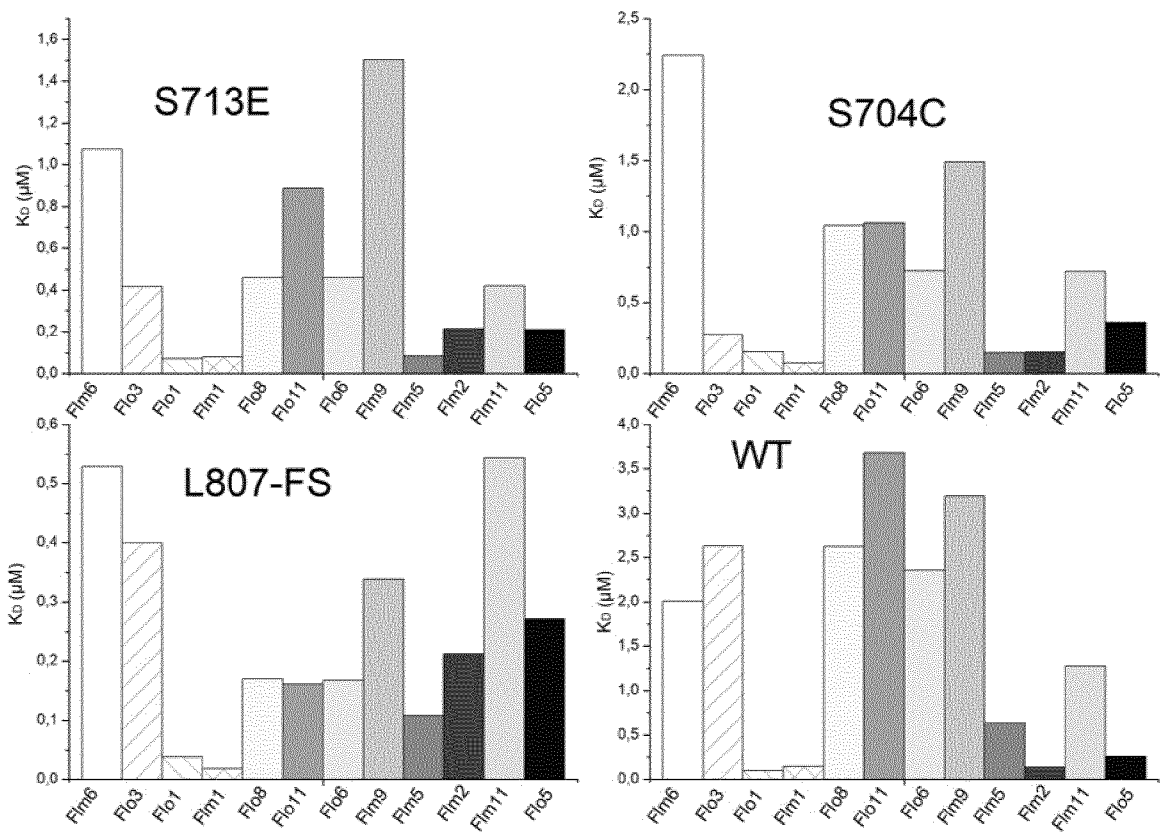


Fig. 16

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2023/066796

A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K7/08 A61K38/04
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data, CHEM ABS Data, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2005/255500 A1 (SAWA AKIRA [US]) 17 November 2005 (2005-11-17) the whole document -----	1-16
Y	CUKKEMANE A ED - FINEBERG N A ET AL: "P.675 Comprehending structural and functional attributes of the C-region of DISC1, a potential therapeutic target for schizophrenia and related mental illnesses", EUROPEAN NEUROPSYCHOPHARMACOLOGY, ELSEVIER SIENCE PUBLISHERS BV , AMSTERDAM, NL, vol. 40, 1 November 2020 (2020-11-01), XP086353769, ISSN: 0924-977X, DOI: 10.1016/J.EURONEURO.2020.09.498 [retrieved on 2020-11-17] the whole document ----- -/--	1-16

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

7 September 2023

Date of mailing of the international search report

18/09/2023

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Pinheiro Vieira, E

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2023/066796

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YERABHAM ANTONY S. K. ET AL: "Biophysical insights from a single chain camelid antibody directed against the Disrupted-in-Schizophrenia 1 protein", PLOS ONE , vol. 13, no. 1 11 January 2018 (2018-01-11), page e0191162, XP055979515, DOI: 10.1371/journal.pone.0191162 Retrieved from the Internet: URL:https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0191162&type=printable the whole document -----	1-16
A	CUKKEMANE ABHISHEK ET AL: "Conformational heterogeneity coupled with [beta]-fibril formation of a scaffold protein involved in chronic mental illnesses", TRANSLATIONAL PSYCHIATRY , vol. 11, no. 1 1 December 2021 (2021-12-01), XP055979520, DOI: 10.1038/s41398-021-01765-1 Retrieved from the Internet: URL:https://www.nature.com/articles/s41398-021-01765-1.pdf the whole document -----	1-16
Y	MURDOCH H. ET AL: "Isoform-Selective Susceptibility of DISC1/Phosphodiesterase-4 Complexes to Dissociation by Elevated Intracellular cAMP Levels", THE JOURNAL OF NEUROSCIENCE, vol. 27, no. 35, 29 August 2007 (2007-08-29), pages 9513-9524, XP055979524, US ISSN: 0270-6474, DOI: 10.1523/JNEUROSCI.1493-07.2007 the whole document -----	1-16

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2023/066796

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005255500 A1	17-11-2005	JP 4948768 B2	06-06-2012
		JP 2005245437 A	15-09-2005
		US 2005255500 A1	17-11-2005



A peptide mimetic therapeutic strategy targeting dysfunction of the scaffold protein DISC-1 in psychiatric disorders

Abhishek Cukkemane^{a,b,*}, Andrew J. Dingley^a, Jeannine Mohrlüder^a,
Beatrix Santiago-Schübel^a, Oliver H. Weiergräber^{a,*}, Dieter Willbold^{a,b,*}

^a Institute of Biological Information Processing, Structural Biochemistry (IBI-7), Forschungszentrum Jülich, Jülich, Germany

^b Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany

ARTICLE INFO

Keywords:

Schizophrenia
Disrupted in schizophrenia 1
Protein aggregation
Proteinopathy
Biophysical analysis
Drug development
Protein-protein interaction

ABSTRACT

Disrupted in schizophrenia 1 (DISC1) is a scaffold protein that regulates several physiological processes ranging from cellular division to neurodevelopment, and its dysfunction contributes to various neurological disorders including schizophrenia, bipolar and mood disorders, and autism. Thus, deciphering its native functions and pathophysiological roles is crucial. In this report, three disease-associated mutants of the C-region of DISC1, i.e., S713E, S704C, and L807-frameshift, were examined to further elucidate the role of DISC1 in cell division. We demonstrate that the mutations do not render the variants functionally inactive; instead, the interaction sites are presumably lost during the aggregation of the DISC1 C-region into amyloid-type fibrils. The minimal fibrillizing element in the C-region is the intrinsically disordered β -core (716–761) that houses a segment absent in the splice variant DISC1 ^{Δ 22aa}, which cannot bind proteins of the mitotic spindle complex and thus hampers cellular proliferation. Based on these structure-function relationships, we present a rational drug development strategy using phage display technology and highlight the role of peptide mimetics in curtailing the agglomeration of fibrils.

1. Background

Schizophrenia, major depressive disorder (MDD), bipolar (BD) and autism spectrum disorders represent closely associated chronic mental illnesses (CMIs) that burden 3 %–5 % of the population (van Os and Kapur, 2009). Additionally, several aetiological factors, such as biological, environmental, social, and psychosocial conditions, are shared by these disorders but are still poorly understood. A critical biological risk factor that was discovered in a Scottish family with a diverse range of psychiatric disorders, including some affective CMIs, was the disrupted in schizophrenia 1 (DISC1) gene (Blackwood et al., 2001; Millar et al., 2000). Subsequently, studies of the cell biology (Brandon et al., 2004; Ishizuka et al., 2011; Kamiya et al., 2008), structural biology and biophysics of DISC1 (Cukkemane et al., 2021; Wang et al., 2019; Ye et al., 2017; Yerabham et al., 2018), as well as the characterization of mutant variants of the C-region (Ishizuka et al., 2011; Leliveld et al., 2008, 2009; Sachs et al., 2005) have added a wealth of information. Such mutations include the S704C, the S713E and L807-frameshift (L807-FS), which are of particular interest from a structural and

functional perspective. For example, mutation of a single residue yielding S704C (Hashimoto et al., 2006; Leliveld et al., 2008, 2009; Narayanan et al., 2011) renders the protein highly insoluble. Consequently, in over 20 % of sixty cases from the biobank at the Stanley Medical Research Institute had the S704C mutant has been observed as aggregates in post-mortem brain tissue (Leliveld et al., 2008, 2009) from subjects suffering from schizophrenia, BD and MDD. The S713E mutant (Ishizuka et al., 2011) mimics a phosphorylation-related signaling cascade, responsible for the Bardet-Biedl syndrome, where neuronal progenitor cells cannot proliferate and migrate during corticogenesis. The L807-FS mutation piques our interest because this mutation causes severe schizophrenia and schizoaffective disorders and culminates in aggregated complexes (Sachs et al., 2005). Unlike the two Ser mutants, the L807 frameshift disrupts the normal reading frame downstream of L807 that comprise one of the binding sites for nuclear distribution element like 1 (NDEL1, formerly known as Nudel) (Brandon et al., 2004; Kamiya et al., 2006; Morris et al., 2003) in DISC1. The disruption of this DISC1–NDEL1 interaction hampers the formation of the mitotic spindle complex and associated cellular processes. These are a few examples

* Corresponding authors.

E-mail addresses: a.cukkemane@fz-juelich.de (A. Cukkemane), o.h.weiergraeber@fz-juelich.de (O.H. Weiergräber), dieter.willbold@hhu.de (D. Willbold).

<https://doi.org/10.1016/j.ejps.2025.107148>

Received 28 April 2025; Received in revised form 16 May 2025; Accepted 28 May 2025

Available online 29 May 2025

0928-0987/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

highlighting the pivotal role of the DISC1 C-region, which provides interaction sites for partners that regulate several key cellular processes, and is therefore of clinical relevance (Bradshaw and Korth, 2019; Hikida et al., 2012; Millar et al., 2003; Tropea et al., 2018; Yerabham et al., 2013).

DISC1 has evolved from being a gene that was initially linked to schizophrenia to a broader molecular candidate for therapeutic intervention in a range of CMIs and neurodevelopmental disorders. This conceptual expansion has led to the term “DISC1opathies”, encompassing a spectrum of DISC1-driven protein conformational disorders rooted in disturbed proteostasis and non-genetic cellular stress (Korth, 2012). This includes a spectrum of DISC1-driven protein conformational disorders underlying disturbed proteostasis and non-genetic cellular stress (Bradshaw and Korth, 2019), which is similar in mechanism to classical neurodegenerative diseases (Cukkemane et al., 2021). These aggregates exhibit prion-like properties (Korth, 2012) and can recruit other proteins, such as dysbindin, enhancing their pathological impact. These findings are supported by postmortem analyses, in vitro assays, and animal models demonstrating that intracellular DISC1 misassembly is present in a significant subset of patients with CMI (Bradshaw and Korth, 2019). Furthermore, because DISC1 serves as a regulatory protein, several post-transcriptional modifications and the stability of its binding partners have also been considered as promising therapeutic targets (Hikida et al., 2012). Altogether, these insights highlight the relevance of DISC1 as a potential to bridge molecular pathology across psychiatric and neurodegenerative spectra.

Recent biophysical and structural biology characterization of the C-region revealed that the functional unit is a tetramer that binds NDEL1 and platelet activating factor acetylhydrolase 1b regulatory subunit 1 (PAFAH1B1, formerly LIS1) in a cooperative manner (Cukkemane et al., 2021). In contrast, aggregated DISC1 and its C-region (691–836) bind thioflavin T (ThT) (Cukkemane et al., 2021; Tanaka et al., 2017), indicating the formation of cross- β amyloidogenic fibrils. Under such conditions, the aggregated protein is most likely dysfunctional and a pseudo-repeat sequence, the β -core (716–761), was hypothesized to form the fibrillar axis (Cukkemane et al., 2021). This region houses part of the 22-residue stretch (748–769) that is absent in the DISC1 ^{Δ 22aa} splice variant (Kamiya et al., 2006; Taylor et al., 2003) and is critical for binding NDEL1 to promote neurite outgrowth. Here, we refer to this stretch as the Δ 22 region.

Studying clinically identified mutants of the C-region of DISC1 should provide a wealth of information about DISC1 (dys)function. Moreover, it is critical to examine the significance of the β -core and the Δ 22 region with respect to the common underlying theme, i.e., their interaction with nuclear distribution element 1 (NDE1, formerly known as NudE), NDEL1 and PAFAH1B1 in the mitotic spindle complex. Understanding the function in relation to protein structure should enable rational development of therapeutic agents targeting the DISC1 C-region. Therefore, to comprehend the effects of the DISC1 C-region mutants on the pathophysiology of schizophrenia and related CMIs, we employed a combination of biophysical and structural biology applications to shed light on DISC1 function.

We demonstrate the functional importance of the β -core for interaction with NDE1, NDEL1 and PAFAH1B1 and extend this analysis to the mutant proteins. In all cases, we observe nM affinity for the mutants to the above-mentioned protein partners. The findings strongly highlight the possibility that the loss of function involved in the pathophysiology of DISC1 and its mutants is not a primary consequence of mutations but rather results from pronounced aggregation. Next, using phage display technology, we identified several peptide hits targeting the wild-type (WT) monomeric and oligomeric versions of the DISC1 C-region. The top candidate peptides interacted with the C-region and mutant variants with high nM affinity. Their potential as drug candidates was studied using a combination of nuclear magnetic resonance (NMR) and dynamic light scattering (DLS) with the C-region, revealing that the peptides attenuate growth of DISC1 fibrils without requiring a specific binding

site. The results provide insight into the poorly defined functional landscape of DISC1 and illustrate a strategy for developing drugs against DISC1 fibrillization.

2. Methodology

2.1. Protein expression and purification

The four DISC1 C-region constructs and the β -core (residue sequence in Fig. S1) in pESPRIT002 (Yumerefendi et al., 2010) were expressed separately as His₆-fusion constructs, as described previously (Cukkemane et al., 2021; Yerabham et al., 2018), by transforming *Escherichia coli* BL21 (DE3) pLysE T1R cells. Cultures were grown primarily in Luria-Bertani (LB) broth. Cells were cultivated in M9 minimal medium supplemented with ¹⁵N-NH₄Cl for preparing NMR samples. Protein expression in minimal and rich media was induced by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM when cell cultures reached an optical density at 600 nm (OD₆₀₀) of ~0.6. Cells were then grown for 16 h at 18 °C. The cells were harvested and lysed in Tris-buffered saline (TBS; 10 mM Tris-HCl, 150 mM NaCl and Complete EDTA-free protease inhibitor cocktail (Roche), pH 7.4) using an ice-chilled microfluidizer M100P (Microfluidics MPT) at 15,000 psi. The insoluble fraction was removed by centrifugation at 50,000 $\times$ g for 25 min at 4 °C. The His₆-fusion construct was purified by loading the soluble fraction onto a column with Co²⁺-charged NTA resin (Qiagen). The bound protein was eluted using TBS containing 500 mM imidazole. The His₆-C-region was further purified on a HiLoad 16/60 Superdex 200 size exclusion chromatography (SEC) column (GE Healthcare).

PAFAH1B1, NDE1 and NDEL1 were prepared using reported protocols (Cukkemane et al., 2021; Soares et al., 2012b; Yerabham et al., 2018). Briefly, PAFAH1B1 was expressed in *E. coli* BL21 C43 (DE3), whereas NDE1 and NDEL1 were expressed in *E. coli* BL21 (DE3) pLysE-T1R. The cells were cultivated at 37 °C until an OD₆₀₀ of 0.6, and expression was induced using 1 mM IPTG and cells grown for a further 16 h at 18 °C. As described above, the harvested cells were lysed and the soluble fraction loaded onto Ni²⁺-NTA resin (Qiagen). After washing with TBS containing 20 mM imidazole, the target protein was eluted using TBS containing 500 mM imidazole and further purified on a HiLoad 16/60 Superdex 200 SEC column (GE Healthcare).

2.2. Circular dichroism (CD) spectroscopy

Far-UV CD spectra for the different C-region proteins were recorded using 20 μ M samples in TBS on a Jasco J-1100 spectropolarimeter at 20 °C (0.2 cm path-length cuvette). The scan speed was 50 nm min⁻¹. Prior to the measurements, protein samples were centrifuged at 20,000 $\times$ g for 20 min at 4 °C to remove any potentially large aggregates and fibrils. The raw data obtained was normalized for the amino acid residues into mean residue ellipticity (MRE) to offer comparison between the different C-region variants and the β -core.

2.3. Dynamic light scattering (DLS)

Measurements of protein samples (10 μ M) were performed using a SpectroSize 300 (XtalConcepts GmbH) instrument at 20 °C with a sample volume of 500 μ L. Peptides were added to the protein sample at different ratios to study their effects on aggregation. Prior to measurements, all samples, without or with peptides, were centrifuged at 21,000 $\times$ g for 30 min at 4 °C. Diffusion coefficients were obtained from analysis of the decay of the scattered intensity autocorrelation function and were used to determine apparent hydrodynamic radii via the Stokes-Einstein equation, as implemented in the manufacturer's software (SpectroCrystal).

2.4. Phage display technique for identification of peptide hits

The phage display selection was performed as described previously (Santur et al., 2021) (Fig. S2) with a few modifications. Briefly, freshly prepared 100 µg/mL monomeric and 15 mg/mL a week-old oligomeric WT C-region DISC1 was used as targets for phage display selection, after immobilization on 3D—NHS matrix plates (PolyAn molecular surface engineering). During the selection rounds, the wells were blocked in alternate selections either with 22 mg/mL l-arginine or 15 mg/mL l-lysine. Multiple control selections were performed to ensure the quality of the selection starting from the library (TriCo16, Creative BioLabs). The different selection types were named “target selection” (TS), “direct control” (DC) and “empty selection” (ES). The ES was performed separately but analogously to the target selection without using the target protein. In the case of DC, the phages of the respective round of target selection were used as input but again no target protein was immobilized. Both control selections (ES and DC) were performed to enable reliable ranking of the resulting peptide sequences. Next-generation sequencing (NGS) of the phage DNA was performed at the Genomics & Transcriptomics Laboratory of the Heinrich-Heine-Universität Düsseldorf, Germany. For NGS, a sample from the library, and all TS, DS and ES inputs were prepared and sequenced.

The phage DNA sequences were processed and filtered using TSAT (target sequence analysis tool (Altendorf et al., 2024; Santur et al., 2021)). NGS data were filtered based on an overall increase in frequency of a sequence in the TS (library < TS1 < TS2 < TS3), enrichment because of the presence of the target in individual rounds (TS2 > DC2; TS3 > DC3) and lastly, enrichment with respect to a mock selection (TS1 > ES1; TS2 > ES2; TS3 > ES3). The resulting hits were ranked by enrichment score (TS3/library) and empty score (TS3/ES3). These filtered sequences were then used as input for HAMMOCK clustering (Krejci et al., 2016). From the list of peptides, we selected those with the highest enrichment scores from TS with the monomeric (Flm1) and oligomeric (Flo1) C-region proteins for further biophysical and structural characterization.

2.5. Biolayer interferometry (BLI)

Binding of physiological partners, i.e., PAFAH1B1, NDE1, and NDEL1 to the different DISC1 C-region constructs at protein concentration of 2 µM was studied using the Octet (ForteBio GmbH) instrument. Protein samples were dialyzed against phosphate buffered saline (PBS). The A2RG sensor chip (Sartorius GmbH) was activated with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/N-hydroxysuccinimide (NHS) (200 mM/50 mM) for coupling the ligand (C-region constructs). The peptides were diluted to 2 µM in 10 mM sodium acetate, pH 4.0, and immobilized to the sensor chip. The flow cell was deactivated with 100 mM ethanolamine-HCl, pH 8.5.

Binding of the peptides to the C-region constructs (WT, S713E, S704C, L807-FS, β-core) was investigated using a similar procedure. Peptides were dissolved in PBS prior to measurements, followed by coupling of the peptide-ligand to the A2RG sensor chip (Sartorius GmbH) using the activation and deactivation processes described above. The dissociation constant (K_D) was determined as the ratio of the off- and on-rate constants, using exponential dissociation and association fittings, respectively (OriginPro 2020, OriginLab, USA).

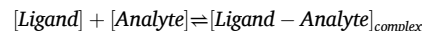
$$\text{Association time} \rightarrow y = y_0 + A1 * \left(1 - \exp\left(-\frac{x}{t1}\right)\right) + A2 * \left(1 - \exp\left(-\frac{x}{t2}\right)\right)$$

$$\text{Dissociation time} \rightarrow y = y_0 + A1 * \exp\left(-\frac{x}{t1}\right) + A2 * \exp\left(-\frac{x}{t2}\right)$$

The rate constants (k) were derived by calculating the reciprocal of the average association/dissociation time (t), which was calculated as

$$\frac{1}{k} = t = \frac{A1 * t1 + A2 * t2}{A1 + A2}$$

The basic reaction describing the association/dissociation of the ligand-analyte interaction was represented as



$$K_D = \frac{[Ligand - Analyte]_{complex}}{[Ligand][Analyte]} = \frac{k_{off}}{k_{on}}$$

The average dissociation constant and standard deviation of the ligand-analyte complex for DISC1-C-region constructs and β-core with the physiological partners is tabulated in Table S1 and the peptide-protein complex was measure once that is tabulated in Table S3. The inclusion criteria were sample dependent and greatly reflected on their stability. For instance, to determine the binding affinity of PAFAH1B 1 to the DISC1 C-region proteins, it was possible to have a sample count greater than 3, while for NDE1, due to its poor stability was limited to $n = 3$. The exact sample size for the measurements is described in Table S1 in supplementary section.

2.6. NMR experiments

Solution NMR experiments were performed at a ^1H Larmor frequency of 750 MHz using a Bruker Avance III spectrometer equipped with a 5 mm cryogenic triple resonance probe. Measurements were performed at 25 °C. DISC1 C-region samples were prepared in PBS containing 10 % (v/v) D_2O and 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). 2D ^1H - ^{15}N HSQC spectra were recorded in the absence and presence of peptide ligands. Proton chemical shifts were referenced to DSS and the ^{15}N chemical shift was referenced indirectly (Wishart et al., 1995). Datasets were processed using Topsin 4.0.6 (Bruker Inc.) and analyzed using NMRFAM-Sparky (Lee et al., 2015).

2.7. Chemical cross-linking reaction and protein digestion

The oligomeric interface of the DISC1 C-region (WT and mutants) was examined by crosslinking experiments of C-region proteins (WT and mutants) coupled with mass spectrometry (MS) analysis. Preliminary tests were performed with DSBU (disuccinimidyl dibutyric urea), 1,1-carbonyldiimidazole (CDI) and disuccinimidyl sulfoxide (DSSO). We identified DSSO as suitable cross-linker. DISC1 C-region-protein (10 µg) samples were incubated for 1 h at 25 °C in the presence of excess of cross-linker (>1 mM) DSSO, which has a spacer arm length of 10.1 Å. The reactions were quenched by the addition of 500 mM Tris-Cl, pH 8.0 (final concentration). The samples were denatured in the presence of 1 % SDS and 10 mM DTT by heating to 95 °C for 10 min. After cooling the samples, cysteines were alkylated by incubation with 50 mM iodoacetamide in darkness at 20 °C for 30 min. Next, 1 M DTT was added to a final concentration of 50 mM and the sample further incubated at 20 °C for 30 min to quench the reaction. For buffer exchange, single-pot solid phase (SP3) paramagnetic beads were used (Hughes et al., 2019). Proteins were bound to the beads by adding nine sample volumes acetonitrile (ACN, 90 % final concentration), washed with 90 % ACN, and released by reconstitution in digestion buffer containing 100 mM HEPES pH 7.5, 2.5 mM CaCl_2 . Trypsin was added in a protein:protease ratio of 100:1 and incubated for 18 h at 37 °C. The samples were acidified to pH < 3.0 by adding 1 % formic acid and desalted by using self-packed double-layer C18-STAGE (STop And Go Extraction, 3 M USA)(Rappsilber et al., 2003) tips.

2.8. LC-MS/MS analysis

An estimated 1 µg of desalted peptides were loaded onto an Ultimate 3000 RSLC nano chromatography system (ThermoFischer) operating

with a two-column setup, a μ PAC reversed-phase trap column (PharmaFluidics) and a 50 cm μ PAC reversed-phase analytical column (PharmaFluidics), both equilibrated with 0.1 % formic acid in water, at a flow rate of 600 nL min⁻¹ and a column temperature of 40 °C. Bound peptides were recovered using a binary gradient from 2 % to 30 % of eluent (0.1 % formic acid in ACN), with 90 min effective separation time and a total runtime of 2 h per sample. Separated peptides were ionized using a CaptiveSpray nano-ESI source (Bruker) with the nitrogen gas saturated with ACN using a Nanobooster (Bruker) and introduced to an Impact II ultra-high resolution QqTOF MS (Bruker) as described in (Beck et al., 2015). Briefly, MS data was acquired with the Bruker HyStar Software (v5.1, Bruker) in line-mode with a mass range of 200–1750 m/z at an acquisition rate of 5 Hz. The 14 most intense ions were selected for fragmentation, with fragment spectra automatically acquired between 5 and 20 Hz. Selected precursors were excluded for the next 0.4 min unless the signal-to-noise ratio improved at least three-fold.

Cross-linked peptide fragments were subsequently identified using the software MaxLynx (Yilmaz et al., 2022) v.2.2.0.0 based on the amino acid sequences of the C-region proteins. The searches were performed using 10 ppm as the precursor, 20 ppm as the fragment ion precision and a false discovery rate of 1 %. Carbamidomethylation of cysteine side chains was assumed and variable degrees of N-terminus acetylation and methionine oxidation considered. Up to three missed cleavages of trypsin were allowed.

2.9. Structural models

The model of the DISC1 C-region protomer was obtained using AlphaFold2 (Jumper et al., 2021), and subjected to *in-silico* docking using CLUSPRO (Kozakov et al., 2017) in the multimer docking mode to generate models of the tetramer. The quality of the structural models was assessed using two sets of restraints, (i) the inter-subunit cross-link of K741-K741 from the MS data and (ii) the hydrodynamic radius (R_H) of the tetramer determined using DLS.

3. Results

3.1. DISC1 C-region mutants bind to their physiological partners

PAFAH1B1 and NDEL1 were previously shown by surface-plasmon resonance (SPR) to bind the native DISC1 C-region with nanomolar affinity (Cukkemane et al., 2021). However, the mutant proteins tended to aggregate readily and block the valves and capillaries of the SPR chip. NDEL1 and PAFAH1B1 were also susceptible to rapid precipitation in solution (Cukkemane et al., 2021; Soares et al., 2012a). These issues prohibited the use of SPR and thus the binding of the C-region variants with these physiological partners were studied by bio-layer interferometry (BLI). The C-region variants were coupled to the sensor tip, and the physiological partners served as analytes. Remarkably, all tested

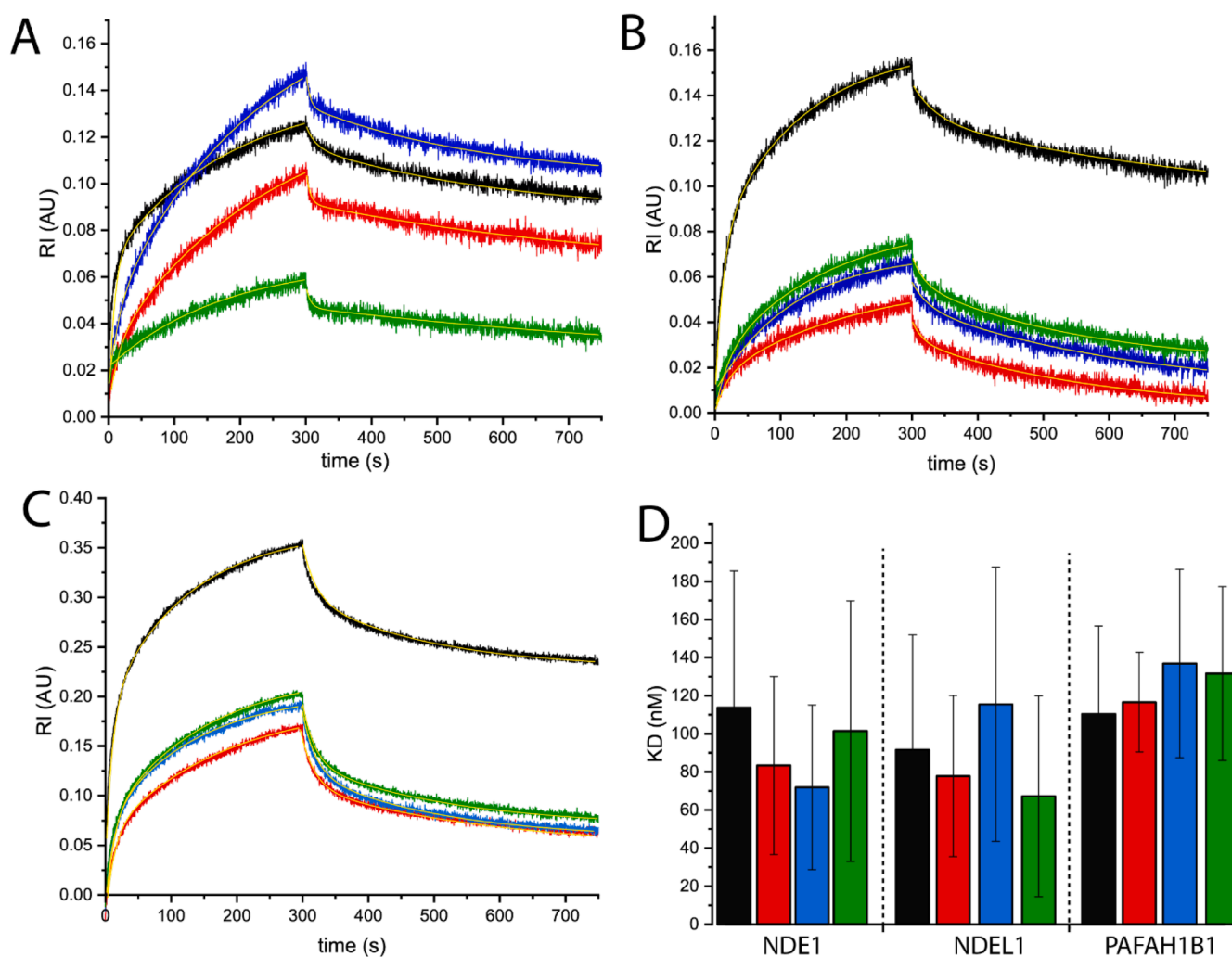


Fig. 1. BLI sensorgrams of (A) NDEL1 (B) NDEL1, (C) PAFAH1B1 and (D) the binding affinity summary in bar plot (mean $\pm$ standard deviation) with WT C-region (black), S704C (red), S713E (blue) and L807-FS (green). The yellow solid line represents the exponential fittings to the kinetic data. The detailed analysis is provided in Tables 1 and S1.

C-region proteins interacted with the three proteins, NDE1, NDEL1 and PAFAH1B1. The kinetics of association and dissociation of the physiological partners to/from the C-region variants were similar (Fig. 1, Tables 1 and S1) under our experimental conditions, although the signal amplitudes differed markedly. Our findings thus provide a first indication that the clinically identified mutants bind to the physiological partners with similar sub- μ M affinities as the WT C-region. This observation implies that the active sites for the interactions with the physiological partners in the mitotic spindle complex are not affected by and possibly located remote from the mutation sites. Thus, we hypothesize that the lynchpin motif facilitating the interaction is likely the β -core (717–761) that carries a pseudo-repeat sequence (Cukkemane et al., 2021). Additionally, it was not always possible to fit mono-exponential association/dissociation functions and hence we had to calculate the binding affinity using an average rate function (see Methodology). This suggests some heterogeneity within and between samples. This potential heterogeneity may have affected the immobilization efficiencies of the ligands. Thus, we resorted to CD spectropolarimetry to assess the state of the proteins in solution.

3.2. Stability of the C-region mutants

Freshly prepared samples were centrifuged to remove any large scattering particles prior to the CD experiments. Nevertheless, the presence of such aggregates was indirectly reflected in the CD spectra as a reduction in the signal intensities for the mutant variant samples (Fig. 2A). The L807-FS mutant showed a very weak ellipticity reading when compared with the ellipticity observed for the other C-region variants. This is in agreement with a previous report (Yerabham et al., 2017) where a decrease in the helical content of the frame shift mutant was observed over a period of 5 days. The S704C and the S713E mutant samples were relatively stable and yielded signals that were half the intensity of the WT C-region. Our attempts to deconvolute the spectra were not successful because the signal intensity in the far UV region in the 190–200 nm was very noisy. However, a qualitative evaluation of the spectra of the WT, S704C and S713E suggests the presence of helical and coiled elements. Taking into account of the normalized amplitudes of the different constructs in the CD spectra (Fig. 2A) indicates poor stability of the mutants in comparison to the WT-C-region that rapidly aggregate and thereby reduce the soluble protein concentration. Translating these observations to the full-length DISC1 protein, we speculate that reduced stability of the mutants may be the primary cause of DISC1-associated pathogenesis, with the mutations exacerbating aggregation.

3.3. The β -core is the minimal functional element that associates with physiological partners of the mitotic spindle complex

To address the second aspect of the hypothesis, i.e., whether the β -core is the functional lynchpin of the C-region, we performed BLI (Fig. 2B and 2C) and CD (Fig. 2A) measurements. Initially, the β -core was titrated against PAFAH1B1, NDEL1 or NDE1. As shown above for the WT C-region and its mutants, high-affinity binding in the sub- μ M range to these physiological partners was observed. This finding strongly

Table 1

Binding affinities (K_D) of the physiological partners viz. NDEL1, NDE1 and PAFAH1B1 w.r.t. the different C-region variants and the β -core.

	WT (nM)	S713E (nM)	S704C (nM)	L807-FS (nM)	β -core (nM)
PAFAH1B1	110.35 $\pm$ 46.18	136.86 $\pm$ 49.43	116.56 $\pm$ 26.19	131.60 $\pm$ 45.57	132.06 $\pm$ 38.07
NDEL1	91.53 $\pm$ 60.34	115.43 $\pm$ 71.98	77.81 $\pm$ 42.23	67.19 $\pm$ 52.73	77.63 $\pm$ 42.61
NDE1	113.71 $\pm$ 71.72	71.92 $\pm$ 43.17	83.35 $\pm$ 46.73	101.39 $\pm$ 68.39	105.04 $\pm$ 16.26

Table 2

Binding affinities (K_D) of the peptide therapeutics w.r.t. the different C-region variants.

	Flo1 (nM)	Fm1 (nM)
WT	97.6	146.3
S704C	156.0	78.0
S713E	73.4	81.0
L807FS	37.2	17.9

suggests that the pseudo-repeat sequence is an important functional entity, which corroborates cell biology experiments where the DISC1 Δ 22aa isoform lacks cell proliferative capability because of weak association with NDEL1 (Kamiya et al., 2006). Circular dichroism experiment on the normalized intensity of β -core (Fig. 2A) suggested that this region of DISC1 is highly unstable under in-vitro conditions and aggregates rapidly. Also, the spectra of β -core shows the presence of structured elements predominantly coiled but also presence of some helical content. However, CD measurements reported previously (Yerabham et al., 2017) indicated the β -core is exclusively unstructured. In the same report, the size-exclusion profile on two different samples of β -core resulted in profiles containing exclusively monomers and another populated with oligomers. Taken together, the results illustrate that the β -core is prone to oligomerization/aggregation and adopts several conformations, some of which are relevant for the function of the C-region. In this study, as we had centrifuged the sample prior to measurements, it is likely that we have removed the large aggregating disordered β -core and have a small population in solution.

Overall, these findings indicate that the mutations do not lead to inactive proteins when the β -core is preserved for association with physiological partners. Rather, the pathophysiology associated with the mutants arises because of their poor stability in solution.

3.4. Structural model of the tetrameric DISC1 C-region

We performed chemical crosslinking-coupled MS of proteolytic fragments to identify sites of self-association and hence characterize the architecture of the C-region tetramer. However, in this case, we encountered serious challenges in identifying proteolytic fragments. We surmised that this problem was caused by the poor stability of the C-region proteins under in-vitro conditions, which may further be compounded by chemical modification during crosslinking. We then tested the WT C-region, the various C-region mutants, and the β -core in the presence of DSSO to identify potential interfacial regions. Under experimental conditions using DSSO as the cross-linker, we observed intra-subunit loop-links among the Lys residues, i.e., K741–K743, K743–K747, K747–K753, K753–K768, K787–K788 and K819–K833 (Table S3). Serendipitously, in a unique instance, we observed an inter-subunit crosslink (K741–K741) in the S704C sample (Fig. S3).

Starting from the AlphaFold2 model of the hDISC1 C-region protomer, we next generated tetramer models using ClusPro (Kozakov et al., 2017). We analyzed candidate structures for the plausibility of the K741–K741 crosslink within a distance of 10.1 Å, which represents the length of the chemical linker. Furthermore, we considered the longest axis of the model as a rough estimate relating to the R_H of $\sim$ 5 nm observed from DLS experiments. Only one model fulfilled the criteria (Fig. 3), whereas two models had acceptable solutions (Fig. S4) but were missing the inter-subunit cross-link possibility. Careful inspection of the model that fulfilled the restraints revealed that the β -core is an essential element for oligomerization, which is mediated by a short α -helix (745–755). Moreover, the disordered region in the β -core is exposed to solvent, thereby potentially offering conformational flexibility to associate with physiological partners. Lastly, in addition to K741–K741 that was obtained as an unambiguous inter-subunit restraint, the model suggests that inter- and intra-subunit cross-links for K741–K743 should

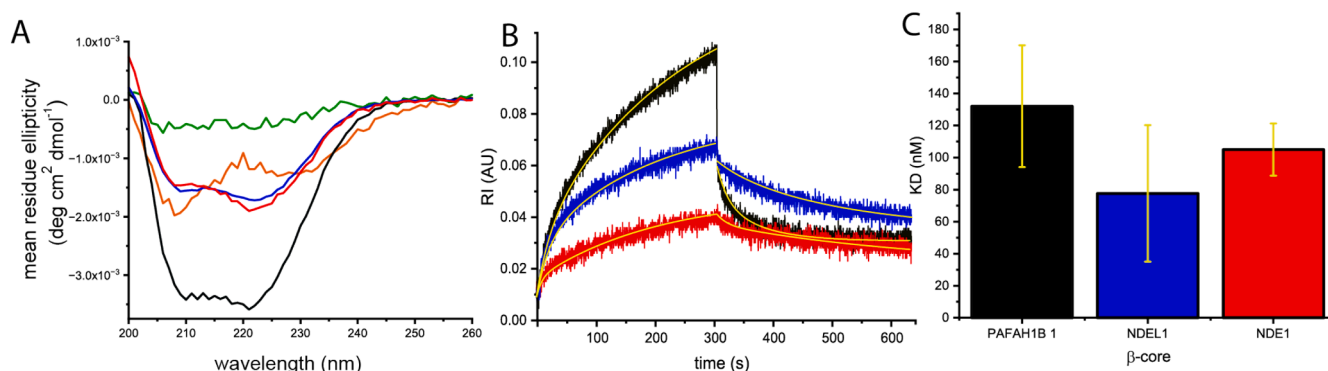


Fig. 2. (A) CD measurements of the various C-region constructs. β -core (orange), WT C-region (black), S713E (blue), S704C (red), and L807-FS (green). (B) BLI sensorgrams of the β -core binding to PAFAH1B1 (black) NDEL1 (blue) and NDE1 (red) with the fitted data shown in yellow. Results of the analysis are given in [Tables 1](#) and [S1](#) and as a bar plot mean $\pm$ standard deviation in (C).

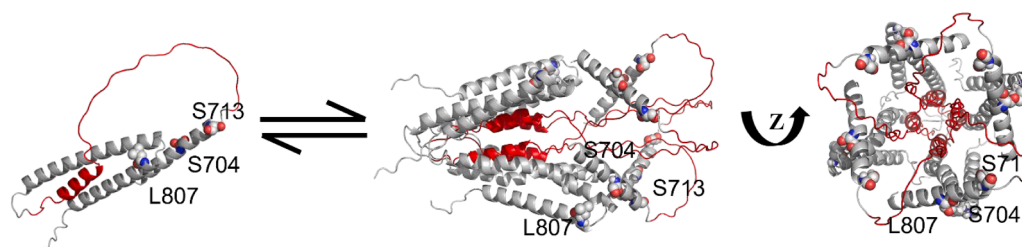


Fig. 3. Structural model of the DISC1 C-region. Left, model of the monomeric protein, as generated by AlphaFold2. Middle and right, tetrameric structure obtained by CLUSPRO docking. Based on the intra- and inter-subunit ([Table S2](#)) restraints that were derived from the MS-crosslinking experiments, the best-fitting model was selected. The residues that are affected by the mutations studied in this manuscript are shown in cpk (for Corey–Pauling–Koltun) mode, the β -core is drawn in red.

also be possible because these two residues are well within 10 Å.

3.5. Discovery of potential therapeutic peptides that target the C-region of DISC1

Phage display selection targeting the C-region of the protein was used as a strategy for developing drugs against DISC1 aggregation ([Jaroszewicz et al., 2022](#)). In this strategy ([Fig. S2](#)), we used a library containing phages that were decorated with 16-mer peptides. We identified several binding partners for the oligomeric and monomeric preparations of the WT DISC1 C-region. Next generation sequencing (NGS) was performed to identify the peptide hits, followed by several rounds of sequence filtering using TSAT ([Santur et al., 2021](#)) and HAMMOCK ([Krejci et al., 2016](#)). Finally, the empty and enrichment scores ([Fig. S5](#)) were calculated for the filtered sequences, and the peptides were ranked and selected on the basis of their empty and enrichment scores and the size of the generated clusters. From the several hits identified, we selected the highest-ranking peptide for both the monomeric and the oligomeric target (termed Flm01 and Flo01, respectively) for biophysical and structural characterization of the protein-peptide interaction.

3.6. Protein-peptide interactions

Flo01 and Flm01 were tested for their specific interactions with the DISC1 C-region. As a first step, the WT C-region of DISC1 was studied using BLI ([Fig. 4A and 4B](#)), where the identified 16-mer peptides were immobilized to the sensor chip and the C-region served as the analyte. This approach was extended to study the interaction of the potential peptide therapeutics with the three clinically identified mutants. The peptides bound to the WT C-region and the mutant variants with affinities in the nM range ([Fig. 4C](#)) ([Cukkemane and Willbold, 2022](#)). Although the affinities were high, the kinetics for the interactions

differed remarkably ([Table S3](#)). The basis of these differences was examined by solution NMR to probe the binding modes of the peptides, while the effects on DISC1 C-region aggregation were assessed by DLS.

3.7. DISC1 C-region aggregation is attenuated in the presence of potential peptide therapeutics

The selected peptides were further tested for their specific interactions with the WT DISC1 C-region. We tested the WT DISC1 C-region, which aggregated rapidly (< 30 min following centrifugation to remove visible aggregates). When challenged with equimolar concentrations of Flo01 ([Fig. 4D, 4 E and Fig. S6](#)) and Flm01 ([Fig. 4D](#)) the fibrillar growth kinetics were reduced markedly. During the experimental time window, the aggregates of the untreated C-region are larger than 250 nm in radius, whereas in the peptide-treated samples, the largest observable species had an apparent R_H of ~50 nm. We also observed the presence of the tetramer when the protein:peptide ratio was maintained at 1:1 ([Fig. 4D](#)) for Flm01. This indicates that a higher concentration ratio of protein-ligand is required for inhibiting and/or curtailing aggregation. In a similar experimental setup, increasing concentrations of Flm01 yielded no further change in the kinetics of the aggregates. However, Flo01 at protein:peptide ratios of 1:5, 1:10 and 1:25 caused a clear decrease in the aggregate size, and at higher ratios a stable population of the tetramer was observed ([Fig. 4E](#)) over the duration of the DLS experiment. Additionally, a stable population with a R_H of 12–16 nm was observed and we postulate that this population likely represents the oligomeric fraction of the C-region. This is because Flo01 was identified using the oligomeric species as the capture ligand. These promising results raise two interesting questions: how do the two peptides that were generated against two different species of the DISC1 C-region interact with the protein and which residues are critical for binding of the peptide drug?

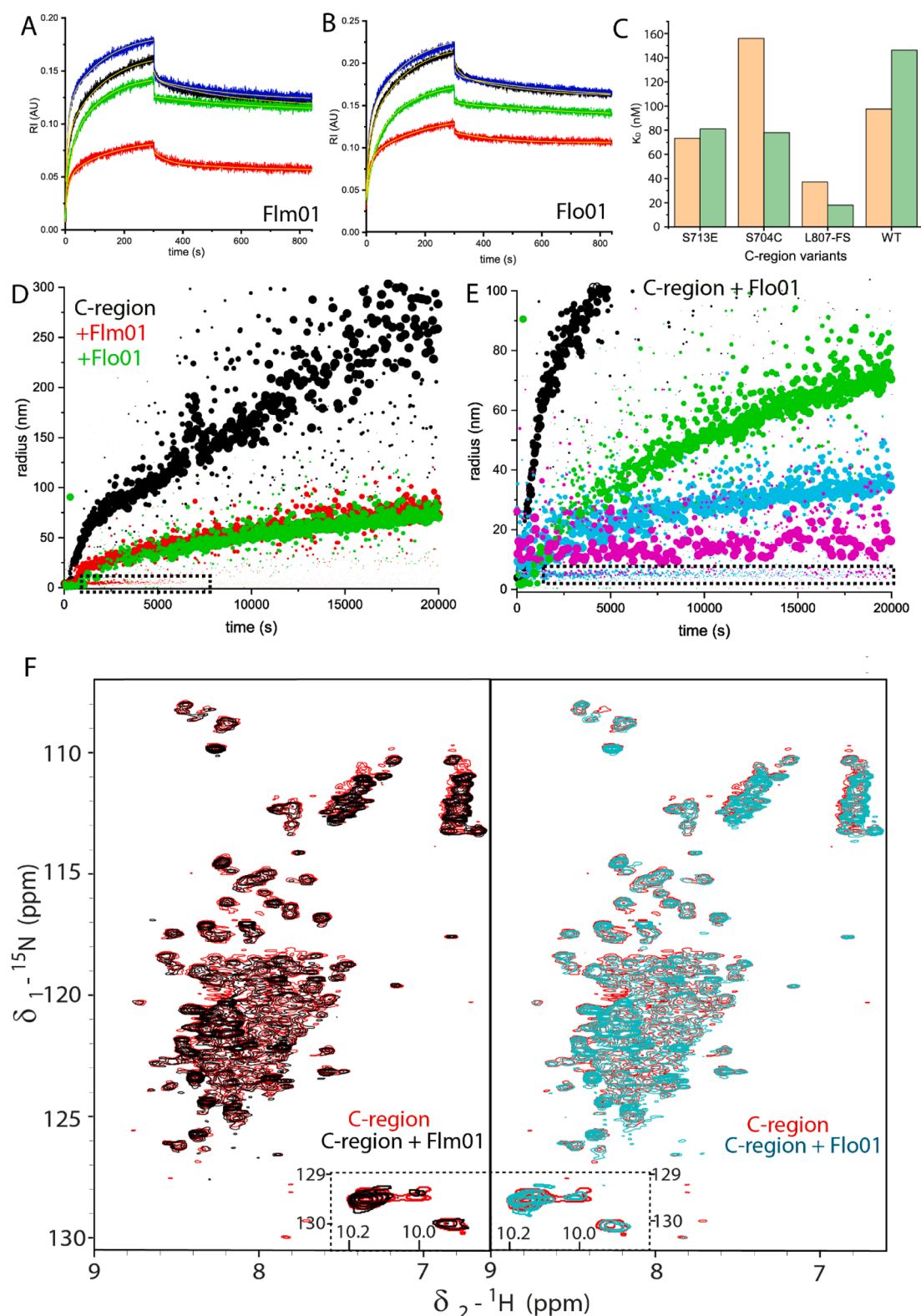


Fig. 4. BLI sensorgrams featuring binding of (A) Flm01 and (B) Flo01 to WT C-region (blue), S704C (red), S713E (black) and L807-FS (green). (C) The binding affinity of the peptides Flo01 (light brown) and Flm01 (green) to the various region constructs is shown in the bar plot, the details of which are provided in [Tables 2](#) and [S3](#). (D) Intensity-weighted DLS isotherms of the WT DISC1 C-region (10 μ M, black) and equimolar mixtures with 16mer peptides Flm01 (red) and Flo01 (green). The dotted box represents the population of the tetrameric species. (E) Close-up (also refer to [Fig. S6](#) for the full isotherm) of a titration of the WT C-region (10 μ M, black) with Flo01 at concentrations of 50 μ M (green), 100 μ M (cyan) and 250 μ M (magenta). Again, the tetrameric population is marked by a dotted box. (F) Overlay of the 2D ^1H - ^{15}N HSQC spectra of unliganded C-region (75 μ M, red) and the C-region bound to Flm01 (125 μ M, black; left panel) and Flo01 (125 μ M, cyan; right panel). Insets (dotted boxes) provide close-up views of the signals arising from the Trp indole side chains.

3.8. DISC1 C-region is intrinsically disordered in the absence of a physiological binding partner

Solution NMR was used to examine the conformation of the DISC1 C-region. The 2D ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectrum of the WT C-region showed low ^1H chemical shift (CS) dispersion of the peaks (Fig. 4F). Similar findings were reported previously (Cukkemane et al., 2021) and also for a truncated mDISC1 C-region (Ye et al., 2017). However, unlike the truncated mDISC1 C-region, where the spectral dispersion improved noticeably on binding a fragment from NDEL1 (Ye et al., 2017) or ATF4 (Wang et al., 2019) in the context of a fusion protein, the overall ^1H CS dispersion of the signals in the presence of either peptide (Fig. 4F) remained essentially unchanged, indicating no pronounced structural change occurred was formed on binding. Nonetheless, minor shifts were observed for a small subset of signals, indicating an interaction between peptide and protein. Increasing the peptide:protein ratio to values similar to those used in the DLS experiments also yielded small changes in the 2D ^1H - ^{15}N HSQC spectrum (Fig. S7). The observed poor ^1H CS dispersion may reflect either intrinsic disorder or oligomerization or both. Considering the findings from the CD spectra (Fig. 2) and the DLS experiments ((Cukkemane et al., 2021), Fig. 4D and 4E), we surmise that the DISC1 C-region can adopt various different conformations and exists as a tetramer and larger species and contains coiled-coil regions, as modeled by AlphaFold2 (Jumper et al., 2021) and reported previously (Wang et al., 2019; Ye et al., 2017), and intrinsically disordered regions (IDRs).

We investigated the CS changes upon binding of Flo01 and Flm01 to the WT C-region. CS perturbations of several distinguishable signals in the 2D ^1H - ^{15}N HSQC spectra of the WT C-region in the presence of Flo01 were observed. Overall, the differences in the CS between the liganded and unliganded states are small. Even smaller CS changes are observed for signals from the WT C-region in the presence of Flm01. In both cases, residues within the IDRs may also interact with the peptides. However, the limited spectral dispersion precludes analysis of these interactions without resorting to various labeling strategies, solid-state NMR spectroscopy, or a combination of both, to obtain residue specific information of the binding interface of the protein-peptide complexes.

4. Discussion

In this report, we sought to understand the function of the DISC1 C-region by studying mutant variants. DISC1 function is associated with the initial stages of neurodevelopment, involving a regulatory role in the mitotic spindle complex (Bradshaw and Korth, 2019; Tropea et al., 2018; Yerabham et al., 2013), but its mechanism of action is still poorly understood. Do the mutations identified in patients actually render the C-region dysfunctional? To this extent, we performed BLI measurements on the different C-region mutant variants. The S704C mutant has previously been identified as aggregates in post-mortem brain tissue of patients suffering from schizophrenia, BD and MDD (Korth, 2012; Leliveld et al., 2009). Similarly, the L807-FS mutant was found to aggregate in subjects with debilitating mental illness (Sachs et al., 2005). In accordance with these data, we observed the S704C and the L807-FS variants to be severely aggregating, but to our surprise, aggregation was also obvious for the S713E protein. Interestingly, all tested mutants, displayed high-affinity binding in the sub- μM range with their physiological partners, NDEL1, NDE1 and PAFAH1B1. The WT C-region binds PAFAH1B1 (Cukkemane et al., 2021) and NDEL1 (Ye et al., 2017) with dissociation constants in the sub- μM range. Additionally, the S704C constitutes a silent mutation that is fully functional and capable of binding to NDEL1 (Kamiya et al., 2006; Leliveld et al., 2009; Narayanan et al., 2011). These findings strongly indicate that the mutations do not directly affect interaction with physiological binding partners, but that loss of function may secondarily arise because of protein aggregation. Studying these interactions is challenging because of protein instability (both ligands and analytes), which is reflected, by the large standard

deviation of the binding constants. Lastly, we note one major difference with respect to our previous findings: We previously observed a μM binding affinity for the C-region and NDEL1 (Cukkemane et al., 2021), in contrast to the nM affinity reported here, which is consistent with the findings of Ye et al. (Ye et al., 2017). We believe the source of discrepancy may be the difficulty in handling the proteins as they are highly unstable in solution.

In the context of protein instability, the L807-FS mutant is a peculiar case among DISC1 mutants, because the region downstream of L807 was postulated to harbor a site critical for binding to NDEL1 (Brandon et al., 2004). The DISC1 segments interacting with NDEL1 have been a matter of intense debate, and regions other than the C-region have been shown to be involved (Kamiya et al., 2006; Morris et al., 2003). One of the strongest indicators that the NDEL1 binding site exists within the C-region is related to the DISC1 $^{\Delta 22\text{aa}}$ splice variant, which binds only weakly to NDEL1. The 3D solution structure (Ye et al., 2017) of the truncated mDISC1-NDEL1 fusion protein is strongly supportive of this notion and highlights an amino acid stretch upstream of L807 that is also critical for the interaction with NDEL1. Herein, we showed that the β -core encompassing residues 716–761, which houses a major portion of the $\Delta 22$ region (748–769), is sufficient to bind to NDEL1, NDE1 and PAFAH1B1 with high affinity. In a previous report (Cukkemane et al., 2021), the β -core was shown to harbor a pseudo-repeat sequence, which is an important feature in amyloidogenic proteins. Indeed, a fibrillar WT C-region with the β -core serving as the fibrillar scaffold that is unavailable for other protein-protein interactions, would resemble the DISC1 $^{\Delta 22\text{aa}}$ splice variant in its inability to bind physiological partners.

To better comprehend how the C-region, particularly the β -core, on one hand is able to bind physiological partners and on the other is susceptible to pathological effects, we derived a structural model. We first generated a template protomer of the C-region using AlphaFold2 (Jumper et al., 2021). The model of the monomer suggests that the β -core is sandwiched between a helix containing S704 and S713 and the coiled-coil domain that was observed experimentally using solution NMR spectroscopy (Wang et al., 2019; Ye et al., 2017). The β -core in the monomeric C-region is likely to display structural plasticity and dynamics to facilitate binding to various physiological partners. In the tetrameric state, the disordered region of the β -core is predicted to be exposed to solvent, enabling multivalent interaction. In analogy to fibril assembly pathways (Willbold et al., 2021) established for amyloidogenic proteins, where IDRs of a monomeric feeder protein transform into the cross- β scaffold of fibrils, we hypothesize monomeric DISC1 to serve as the building block for fibrillization.

Based on our findings, previously reported data (Cukkemane et al., 2021) and the recent success of phase-I clinical trials of the therapeutic d-enantiomeric peptide RD2 (Schemmert et al., 2019) in eliminating A β oligomers for treating Alzheimer's disease piqued our interest in identifying peptide drugs targeting DISC1 using a similar approach. We initiated a selection using monomeric and oligomeric C-region proteins as bait to identify peptide binders. Two lead peptides, one identified against the monomer (Flm1) and the other against the oligomer (Flo1), displayed nanomolar affinities (Fig 4A and 4B; (Cukkemane and Willbold, 2022)) to the DISC1 WT C-region and mutants. Although both peptides did not abolish aggregation of the WT C-region, they remarkably slowed the kinetics of fibril growth and successfully curtailed the formation of large particles (see Fig. 5 for proposed model). At a protein-to-peptide ratio of 1:1, both peptides reduced C-region aggregate sizes throughout the experiment (Fig. 4D). Remarkably, when applied at tenfold excess, Flo1 (Fig 4E) was even able to bring particle sizes down to 12 to 16 nm, possibly representing the oligomeric species that was used for its selection. These findings open a whole new avenue and a wide range of applications. On the one hand, the peptides can be used to capture DISC1 from blood and/or CSF as a diagnostic marker. On the other hand, they might be further modified and improved for application as therapeutic agents, with the aim to limit or prevent formation of DISC1 aggregates and/or fibrils.

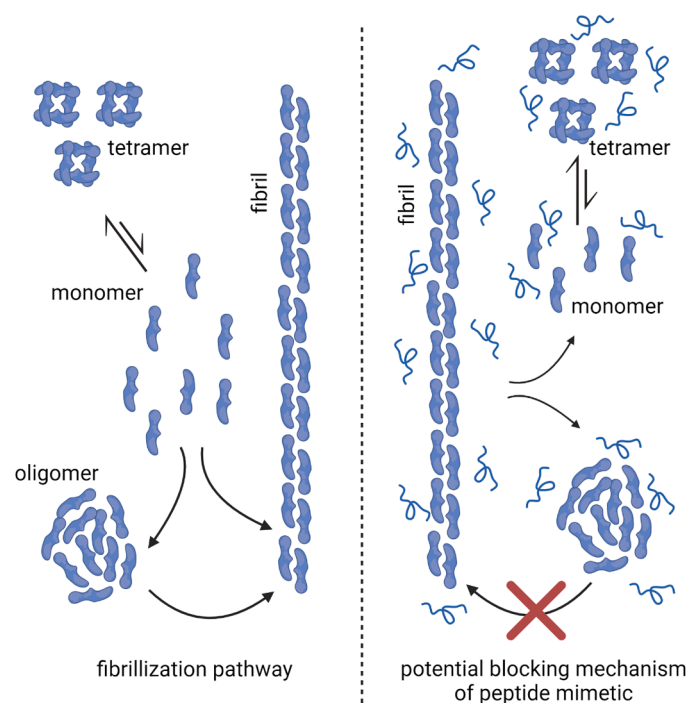


Fig. 5. Schematic illustration of the potential activity of the peptide mimetic. In the absence of the peptide (left panel) the monomer is in equilibrium with the tetramer under physiological conditions. Under atypical conditions, the monomer serves the role of the building block for the formation of oligomers and fibrils. In the presence of the peptide (right panel), we postulate that the peptide interacts with the monomers in a competitive manner such that a competition ensues for the monomer between the peptide and the aggregates. This results in curtailing the kinetics and the size of the aggregates. In the case of Flo01, at sufficiently high concentrations, the oligomeric species is stabilized along with the tetramer that can no longer participate in the formation of fibrils, while in Flm01 the reaction continues to proceed albeit with slow growth kinetics.

Impressed by the activity of the peptides, we next proceeded to characterize residue-specific interactions between the WT C-region and the peptide ligands using solution NMR spectroscopy. The finding of only minor chemical shift changes for a limited number of C-region resonances is intriguing given that the peptides bind to the protein with high affinity and significantly affect its self-association kinetics. Notably, similar minor peak shifts in NMR data have been reported for the ultrahigh affinity complex between the two intrinsically disordered proteins histone H1 and ProTα (Borgia et al., 2018). In this study, complex formation involves a large number of opposite charge interactions without defined interacting residues or sites. We hypothesize that a similar situation exists for the interaction between DISC1 and the 16mer peptides, with long-range electrostatic attraction mediating the high affinity interaction between the two polypeptide chains. Although the affinity between the peptide and the C-region involves high-affinity binding in the nm range, we require higher ratios of the peptide to slow the kinetics of aggregation growth to near stabilization of the population of the oligomers and tetramers (Figs 4D and 4 E). However, the formation of aggregates and fibrils is also highly favored thermodynamically, with a binding affinity of the C-region monomers to the fibrils of ~750 nM (Cukkemane et al., 2021). This results in a competition between the peptides and the aggregates for the free monomers in the solution (Fig 5). Based on the results from the DLS, we observe a slow tonic increase in the size of the aggregates (Figs 4D) at a stoichiometry of 1:1. However at higher, concentrations of the peptide namely, Flo01, we see a total inhibition of the formation of the aggregates and stabilization of the potentially oligomeric fraction and the tetrameric species.

Despite these promising results, admittedly several limitations persist. First, both the C-region protein and its variants exhibited significant instability due to poor solubility, particularly in the case of the L807-FS mutant. This compromised protein handling and likely contributed to the observed variability in binding affinity measurements. More importantly, extrapolating these findings to the full-length DISC1 protein presents an even greater challenge, because the protein contains several intrinsically disordered and aggregation-prone regions that complicate biochemical analysis. Second, the following study is purely biophysical in nature. It limits our ability to infer therapeutic potential of the peptide mimetics from the application point-of-view to assess the biological efficacy, toxicity, or pharmacodynamics of the peptides. Third, the pharmacokinetics and metabolic stability of the peptides have not yet been evaluated. As L-enantiomeric drugs are susceptible to proteolytic degradation, it is necessary to conduct in-vitro and in-vivo tests on the d-enantiomeric versions for therapeutic use as drug candidates. These limitations highlight the need for further validation in cellular and animal models.

5. Conclusion

In summary, folding and aggregation of proteins are two sides of the same coin (Pastore and Temussi, 2012). Proteinopathic agents more often contain large stretches of IDRs (Lapidus, 2013). In the absence of a physiological partner, the unstructured regions likely exist in a dynamic equilibrium. Together with the coiled-coil domain, the IDR structures provide ample freedom for DISC1 to associate with a plethora of potential partners, and we were able to narrow down the segment that binds components of the mitotic spindle complex (NDEL1, NDE1, PAFAH1B1) to the β-core (716-761). We also showed that disease-associated mutations in the C-region do not directly affect protein function, but rather influence protein stability.

Lastly, we describe a potential target for treating psychotic disorders with a strategy that does not involve prospective drugs of low molecular weight (< 900 Da). Most forms of therapy that exist today are empirical. They comprise the 2nd and 3rd generation atypical antipsychotics (Paul and Potter, 2023), which target dopamine receptors (D2R) and, recently, also serotonin receptors (5HTR), and primarily serve to treat symptoms (mostly psychotic) and preserve remaining abilities. Therefore, it is essential to identify molecular structures that are critical for pathogenesis and target them for causal intervention. Using the phage-display strategy, we demonstrate the selection of lead peptides directed at monomeric and oligomeric states of the C-region that are capable of slowing down the fibrillization kinetics. This raises further potential to develop new strategies for drug development targeting risk factors for schizophrenia, bipolar disorder, major depressive disorder, and related chronic mental illnesses.

In order to realize the full potential of the peptide mimetics as therapeutic candidates, it is essential to address the limitations mentioned above that requires several critical steps. This includes optimizing peptide stability through chemical modifications such as cyclization, incorporation of non-natural amino acids, or PEGylation, which may significantly enhance their resistance to degradation, improve their pharmacokinetic properties and bioavailability. Among these one of the most successfully implemented methodologies in our group has been the use of d-enantiomeric peptides (Funke and Willbold, 2012), which were generated using mirror-image phage display technique (Sun et al., 2012). Notably, this method has already yielded the therapeutic peptide PRI-002, which has successfully completed Phase Ib (Kutzsche et al., 2025) is currently undergoing Phase II clinical trials for Alzheimer's disease. As the quintessential next step, we plan to evaluate the efficacy and safety of the peptides in DISC1 mouse models namely Tg (tetO-DISC1*)1001Plet (Pletnikov et al., 2008) and the DISC1-Boymaw (Zhou et al., 2010). In the near future, we aim to complete the pre-clinical trials to assess the peptides' mechanism of action in mice models, identify potential off-target effects, and explore formulations

that support clinical delivery. All these efforts together will be crucial for translating these biophysically validated peptide leads into clinically viable candidates for CMIIs.

CRediT authorship contribution statement

Abhishek Cukkemane: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Andrew J. Dingley:** Writing – review & editing, Validation, Methodology, Investigation. **Jeannine Mohrlüder:** Writing – review & editing, Methodology, Investigation. **Beatrix Santiago-Schübel:** Writing – review & editing, Methodology, Investigation. **Oliver H. Weiergräber:** Writing – review & editing, Writing – original draft, Resources, Project administration, Formal analysis, Conceptualization. **Dieter Willbold:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

A.C. and D.W. are co-inventor of patents covering the peptide drugs. All other authors declare no competing interests.

Acknowledgments

We thank the Biologisch Medizinisches Forschungszentrum of the Heinrich Heine University Düsseldorf for performing the NGS analysis. Access to the Jülich-Düsseldorf Biomolecular NMR Center jointly run by Forschungszentrum Jülich and HHU, and supported by the DFG (HE 3243/4–1 and INST 208/771–1 FUGG) are acknowledged. We also thank the Alexander von Humboldt foundation (Ref 3.4 – 1142747 – HRV – IP) for financial support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ejps.2025.107148](https://doi.org/10.1016/j.ejps.2025.107148).

Data availability

No data was used for the research described in the article.

References

- Altendorf, T., Mohrlüder, J., Willbold, D., 2024. TSAT: efficient evaluation software for NGS data of phage/mirror-image phage display selections. *Biophys Rep (N Y)* 4 (3), 100166. <https://doi.org/10.1016/j.bpr.2024.100166>.
- Beck, S., Michalski, A., Raether, O., Lubeck, M., Kaspar, S., Goedecke, N., Baessmann, C., Hornburg, D., Meier, F., Paron, I., Kulak, N.A., Cox, J., Mann, M., 2015. The Impact II, a very high-resolution quadrupole time-of-flight instrument (QTOF) for deep shotgun proteomics. *Mol Cell Proteomics* 14 (7), 2014–2029. <https://doi.org/10.1074/mcp.M114.047407>.
- Blackwood, D.H.R., Fordyce, A., Walker, M.T., St Clair, D.M., Porteous, D.J., Muir, W.J., 2001. Schizophrenia and affective disorders – cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: clinical and P300 findings in a family. *Am. J. Hum. Genet.* 69 (2), 428–433. <https://doi.org/10.1086/321969>.
- Borgia, A., Borgia, M.B., Bugge, K., Kissling, V.M., Heidarsson, P.O., Fernandes, C.B., Sottini, A., Soranno, A., Buholzer, K.J., Nettels, D., Kragelund, B.B., Best, R.B., Schuler, B., 2018. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* 555 (7694), 61–66. <https://doi.org/10.1038/nature25762>.
- Bradshaw, N.J., Korth, C., 2019. Protein misassembly and aggregation as potential convergence points for non-genetic causes of chronic mental illness. *Mol Psychiatry* 24 (7), 936–951. <https://doi.org/10.1038/s41380-018-0133-2>.
- Brandon, N.J., Handford, E.J., Schurov, I., Rain, J.C., Pelling, M., Duran-Jimeniz, B., Camargo, L.M., Oliver, K.R., Beher, D., Shearman, M.S., Whiting, P.J., 2004. Disrupted in Schizophrenia 1 and Nudel form a neurodevelopmentally regulated protein complex: implications for schizophrenia and other major neurological disorders. *Mol Cell Neurosci* 25 (1), 42–55. <https://doi.org/10.1016/j.mcn.2003.09.009>.
- Cukkemane, A., Becker, N., Zielinski, M., Frieg, B., Lakomek, N.A., Heise, H., Schroder, G.F., Willbold, D., Weiergräber, O.H., 2021. Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl Psychiatry* 11 (1), 639. <https://doi.org/10.1038/s41398-021-01765-1>.
- Cukkemane, A., & Willbold, D. (2022). DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDS). (PCT/EP2023/066796).
- Funke, S.A., Willbold, D., 2012. Peptides for therapy and diagnosis of Alzheimer's disease. *Curr Pharm Des* 18 (6), 755–767. <https://doi.org/10.2174/138161212799277752>.
- Hashimoto, R., Numakawa, T., Ohnishi, T., Kumamaru, E., Yagasaki, Y., Ishimoto, T., Mori, T., Nemoto, K., Adachi, N., Izumi, A., Chiba, S., Noguchi, H., Suzuki, T., Iwata, N., Ozaki, N., Taguchi, T., Kamiya, A., Kosuga, A., Tatsumi, M., Kunugi, H., 2006. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Hum. Mol. Genet.* 15 (20), 3024–3033. <https://doi.org/10.1093/hmg/ddl244>.
- Hikida, T., Gamo, N.J., Sawa, A., 2012. DISC1 as a therapeutic target for mental illnesses. *Expert Opin. Ther. Targets* 16 (12), 1151–1160. <https://doi.org/10.1517/14728222.2012.719879>.
- Hughes, C.S., Moggridge, S., Muller, T., Sorensen, P.H., Morin, G.B., Krijgsvel, J., 2019. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* 14 (1), 68–85. <https://doi.org/10.1038/s41596-018-0082-x>.
- Ishizuka, K., Kamiya, A., Oh, E.C., Kanki, H., Seshadri, S., Robinson, J.F., Murdoch, H., Dunlop, A.J., Kubo, K., Furukori, K., Huang, B., Zeledon, M., Hayashi-Takagi, A., Okano, H., Nakajima, K., Houslay, M.D., Katsanis, N., Sawa, A., 2011. DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature* 473 (7345), 92–96. <https://doi.org/10.1038/nature09859>.
- Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L., Węgrzyn, G., 2022. Phage display and other peptide display technologies. *FEMS Microbiol Rev* 46 (2). <https://doi.org/10.1093/femsre/fuab052>.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596 (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- Kamiya, A., Tan, P.L., Kubo, K., Engelhard, C., Ishizuka, K., Kubo, A., Tsukita, S., Pulver, A.E., Nakajima, K., Cascella, N.G., Katsanis, N., Sawa, A., 2008. Recruitment of PCM1 to the centrosome by the cooperative action of DISC1 and BBS4: a candidate for psychiatric illnesses. *Arch Gen Psychiatry* 65 (9), 996–1006. <https://doi.org/10.1001/archpsyc.65.9.996>.
- Kamiya, A., Tomoda, T., Chang, J., Takaki, M., Zhan, C., Morita, M., Cascio, M.B., Elashvili, S., Koizumi, H., Takanezawa, Y., Dickerson, F., Yolken, R., Arai, H., Sawa, A., 2006. DISC1-NDEL1/NUDEL protein interaction, an essential component for neurite outgrowth, is modulated by genetic variations of DISC1. *Hum. Mol. Genet.* 15 (22), 3313–3323. <https://doi.org/10.1093/hmg/ddl407>.
- Korth, C., 2012. Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion* 6 (2), 134–141. <https://doi.org/10.4161/pri.18989>.
- Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhomy, D., Yueh, C., Beglov, D., Vajda, S., 2017. The ClusPro web server for protein-protein docking. *Nat Protoc* 12 (2), 255–278. <https://doi.org/10.1038/nprot.2016.169>.
- Krejci, A., Hupp, T.R., Lexa, M., Vojtesek, B., Muller, P., 2016. Hammock: a hidden Markov model-based peptide clustering algorithm to identify protein-interaction consensus motifs in large datasets. *Bioinformatics* 32 (1), 9–16. <https://doi.org/10.1093/bioinformatics/btv522>.
- Kutzsche, J., Cosma, N.C., Kauselmann, G., Fenski, F., Bieniek, C., Bujnicki, T., Pils, M., Bannach, O., Willbold, D., Peters, O., 2025. Oral PRI-002 treatment in patients with MCI or mild AD: a randomized, double-blind phase 1b trial. *Nat Commun* 16 (1), 4180. <https://doi.org/10.1038/s41467-025-59295-z>.
- Lapidus, L.J., 2013. Understanding protein aggregation from the view of monomer dynamics. *Mol Biosyst* 9 (1), 29–35. <https://doi.org/10.1039/c2mb25334h>.
- Lee, W., Tonelli, M., Markley, J.L., 2015. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* 31 (8), 1325–1327. <https://doi.org/10.1093/bioinformatics/btu830>.
- Leliveld, S.R., Bader, V., Hendriks, P., Prikulis, I., Sajjani, G., Requena, J.R., Korth, C., 2008. Insolubility of disrupted-in-schizophrenia 1 disrupts oligomer-dependent interactions with nuclear distribution element 1 and is associated with sporadic mental disease. *Journal of Neuroscience* 28 (15), 3839–3845. <https://doi.org/10.1523/JNEUROSCI.5389-07.2008>.
- Leliveld, S.R., Hendriks, P., Michel, M., Sajjani, G., Bader, V., Trossbach, S., Prikulis, I., Hartmann, R., Jonas, E., Willbold, D., Requena, J.R., Korth, C., 2009. Oligomer assembly of the C-terminal DISC1 domain (640–854) is controlled by self-association motifs and disease-associated polymorphism S704C. *Biochemistry* 48 (32), 7746–7755. <https://doi.org/10.1021/bi900901e>.
- Millar, J.K., Christie, S., Porteous, D.J., 2003. Yeast two-hybrid screens implicate DISC1 in brain development and function. *Biochem. Biophys. Res. Commun.* 311 (4), 1019–1025. <https://doi.org/10.1016/j.bbrc.2003.10.101>.
- Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A.M., Devon, R.S., St Clair, D.M., Muir, W.J., Blackwood, D.H.R., Porteous, D.J., 2000. Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* 9 (9), 1415–1423. <https://doi.org/10.1093/hmg/9.9.1415>.
- Morris, J.A., Kandpal, G., Ma, L., Austin, C.P., 2003. DISC1 (Disrupted-In-Schizophrenia 1) is a centrosome-associated protein that interacts with MAP1A, MIPT3, ATF4/5

- and NUDEL: regulation and loss of interaction with mutation. *Hum. Mol. Genet.* 12 (13), 1591–1608. <https://doi.org/10.1093/hmg/ddg162>.
- Narayanan, S., Arthanari, H., Wolfe, M.S., Wagner, G., 2011. Molecular characterization of disrupted in schizophrenia-1 risk variant S704C reveals the formation of altered oligomeric assembly. *J Biol Chem* 286 (51), 44266–44276. <https://doi.org/10.1074/jbc.M111.271593>.
- Pastore, A., Temussi, P.A., 2012. The two faces of Janus: functional interactions and protein aggregation. *Curr Opin Struct Biol* 22 (1), 30–37. <https://doi.org/10.1016/j.sbi.2011.11.007>.
- Paul, S.M., Potter, W.Z., 2023. Finding new and better treatments for psychiatric disorders. *Neuropsychopharmacology*. <https://doi.org/10.1038/s41386-023-01690-5>.
- Pletnikov, M.V., Ayhan, Y., Nikolskaia, O., Xu, Y., Ovanesov, M.V., Huang, H., Mori, S., Moran, T.H., Ross, C.A., 2008. Inducible expression of mutant human DISC1 in mice is associated with brain and behavioral abnormalities reminiscent of schizophrenia. *Mol Psychiatry* 13 (2), 173–186. <https://doi.org/10.1038/sj.mp.4002079>, 115.
- Rappsilber, J., Ishihama, Y., Mann, M., 2003. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75 (3), 663–670. <https://doi.org/10.1021/ac026117i>.
- Sachs, N.A., Sawa, A., Holmes, S.E., Ross, C.A., DeLisi, L.E., Margolis, R.L., 2005. A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Mol Psychiatry* 10 (8), 758–764. <https://doi.org/10.1038/sj.mp.4001667>.
- Santur, K., Reinartz, E., Lien, Y., Tusche, M., Altendorf, T., Sevenich, M., Tamguney, G., Mohrluder, J., Willbold, D., 2021. Ligand-induced stabilization of the native Human superoxide dismutase 1. *ACS Chem Neurosci* 12 (13), 2520–2528. <https://doi.org/10.1021/acschemneuro.1c00253>.
- Schemmert, S., Schartmann, E., Honold, D., Zafiu, C., Ziehml, T., Langen, K.J., Shah, N.J., Kutzsche, J., Willuweit, A., Willbold, D., 2019. Deceleration of the neurodegenerative phenotype in pyroglutamate-Abeta accumulating transgenic mice by oral treatment with the A-beta oligomer eliminating compound RD2. *Neurobiol Dis* 124, 36–45. <https://doi.org/10.1016/j.nbd.2018.10.021>.
- Soares, D.C., Bradshaw, N.J., Zou, J., Kennaway, C.K., Hamilton, R.S., Chen, Z.A., Wear, M.A., Blackburn, E.A., Bramham, J., Bottcher, B., Millar, J.K., Barlow, P.N., Walkinshaw, M.D., Rappsilber, J., Porteous, D.J., 2012a. The mitosis and neurodevelopment proteins NDE1 and NDEL1 form dimers, tetramers, and polymers with a folded back structure in solution. *J Biol Chem* 287 (39), 32381–32393. <https://doi.org/10.1074/jbc.M112.393439>.
- Soares, D.C., Bradshaw, N.J., Zou, J., Kennaway, C.K., Hamilton, R.S., Chen, Z.A., Wear, M.A., Blackburn, E.A., Bramham, J., Bottcher, B., Millar, J.K., Barlow, P.N., Walkinshaw, M.D., Rappsilber, J., Porteous, D.J., 2012b. The mitosis and neurodevelopment proteins NDE1 and NDEL1 form dimers, tetramers, and polymers with a folded back structure in solution. *Journal of Biological Chemistry* 287 (39), 32381–32393. <https://doi.org/10.1074/jbc.M112.393439>.
- Sun, N., Funke, S.A., Willbold, D., 2012. Mirror image phage display-generating stable therapeutically and diagnostically active peptides with biotechnological means. *J Biotechnol* 161 (2), 121–125. <https://doi.org/10.1016/j.jbiotec.2012.05.019>.
- Tanaka, M., Ishizuka, K., Nekooki-Machida, Y., Endo, R., Takashima, N., Sasaki, H., Komi, Y., Gathercole, A., Huston, E., Ishii, K., Hui, K.K., Kurosawa, M., Kim, S.H., Nukina, N., Takimoto, E., Houslay, M.D., Sawa, A., 2017. Aggregation of scaffolding protein DISC1 dysregulates phosphodiesterase 4 in Huntington's disease. *J Clin Invest* 127 (4), 1438–1450. <https://doi.org/10.1172/JCI85594>.
- Taylor, M.S., Devon, R.S., Millar, J.K., Porteous, D.J., 2003. Evolutionary constraints on the disrupted in Schizophrenia locus. *Genomics* 81 (1), 67–77. [https://doi.org/10.1016/s0888-7543\(02\)00026-5](https://doi.org/10.1016/s0888-7543(02)00026-5).
- Tropea, D., Hardingham, N., Millar, K., Fox, K., 2018. Mechanisms underlying the role of DISC1 in synaptic plasticity. *Journal of Physiology-London* 596 (14), 2747–2771. <https://doi.org/10.1113/Jp274330>.
- van Os, J., Kapur, S., 2009. Schizophrenia. *Lancet* 374 (9690), 635–645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8).
- Wang, X., Ye, F., Wen, Z., Guo, Z., Yu, C., Huang, W.K., Rojas Ringeling, F., Su, Y., Zheng, W., Zhou, G., Christian, K.M., Song, H., Zhang, M., Ming, G.L., 2019. Structural interaction between DISC1 and ATF4 underlying transcriptional and synaptic dysregulation in an iPSC model of mental disorders. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-019-0485-2>.
- Willbold, D., Strodel, B., Schroder, G.F., Hoyer, W., Heise, H., 2021. Amyloid-type protein aggregation and prion-like properties of amyloids. *Chem Rev* 121 (13), 8285–8307. <https://doi.org/10.1021/acs.chemrev.1c00196>.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J. L., Sykes, B.D., 1995. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* 6 (2), 135–140. <https://doi.org/10.1007/BF00211777>.
- Ye, F., Kang, E., Yu, C., Qian, X., Jacob, F., Yu, C., Mao, M., Poon, R.Y.C., Kim, J., Song, H., Ming, G.L., Zhang, M., 2017. DISC1 Regulates neurogenesis via modulating kinetochore attachment of Ndel1/Nde1 during mitosis. *Neuron* 96 (5), 1204. <https://doi.org/10.1016/j.neuron.2017.11.034>.
- Yerabham, A.S., Weiergraber, O.H., Bradshaw, N.J., Korth, C., 2013. Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol Chem* 394 (11), 1425–1437. <https://doi.org/10.1515/hsz-2013-0178>.
- Yerabham, A.S.K., Mas, P.J., Decker, C., Soares, D.C., Weiergraber, O.H., Nagel-Steger, L., Willbold, D., Hart, D.J., Bradshaw, N.J., Korth, C., 2017. A structural organization for the Disrupted in Schizophrenia 1 protein, identified by high-throughput screening, reveals distinctly folded regions, which are bisected by mental illness-related mutations. *J Biol Chem* 292 (16), 6468–6477. <https://doi.org/10.1074/jbc.M116.773903>.
- Yerabham, A.S.K., Muller-Schiffmann, A., Ziehml, T., Stadler, A., Kober, S., Indurkha, X., Marreiros, R., Trossbach, S.V., Bradshaw, N.J., Prikulis, I., Willbold, D., Weiergraber, O.H., Korth, C., 2018. Biophysical insights from a single chain camelid antibody directed against the disrupted-in-Schizophrenia 1 protein. *PLoS One* 13 (1), e0191162. <https://doi.org/10.1371/journal.pone.0191162>.
- Yilmaz, S., Busch, F., Nagaraj, N., Cox, J., 2022. Accurate and automated high-coverage identification of chemically cross-linked peptides with MaxLynx. *Anal Chem* 94 (3), 1608–1617. <https://doi.org/10.1021/acs.analchem.1c03688>.
- Yumerefendi, H., Tarendeau, F., Mas, P.J., Hart, D.J., 2010. ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *J Struct Biol* 172 (1), 66–74. <https://doi.org/10.1016/j.jsb.2010.02.021>.
- Zhou, X., Chen, Q., Schaukowitz, K., Kelsoe, J.R., Geyer, M.A., 2010. Insoluble DISC1-Boymaw fusion proteins generated by DISC1 translocation. *Mol Psychiatry* 15 (7), 669–672. <https://doi.org/10.1038/mp.2009.127>.

A peptide mimetic therapeutic strategy targeting the scaffold protein DISC-1 dysfunction in psychiatric disorders

Abhishek Cukkemane^{1,2*}, Andrew J. Dingley¹, Jeannine Mohrlüder¹, Beatrix Santiago-Schübel¹,
Oliver H. Weiergräber^{1*} and Dieter Willbold^{1,2*}

¹Institute of Biological Information Processing (IBI-7: Structural Biochemistry),
Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

*email – a.cukkemane@fz-juelich.de; o.h.weiergraeber@fz-juelich.de; d.willbold@fz-juelich.de

Table S1: Summary of rate constants for the binding of physiological partners viz. NDEL1, NDE1 and PAFAH1B1 to the 2 μ M of the different C-region variants and the β -core

C-region	Partner	n	K _D (nM)	K _{on} ($\times 10^4$ M ⁻¹ .s ⁻¹)	K _{off} (10e^{-2} s ⁻¹)
WT	PAFAH1B 1	14	110.35 $\pm$ 46.18	25.48 $\pm$ 17.46	2.67 $\pm$ 1.77
	NDEL1	10	91.53 $\pm$ 60.34	8.17 $\pm$ 11.07	0.90 $\pm$ 1.64
	NDE1	4	113.71 $\pm$ 71.72	11.61 $\pm$ 16.55	2.12 $\pm$ 3.69
S704C	PAFAH1B 1	7	116.56 $\pm$ 26.19	21.94 $\pm$ 14.19	2.56 $\pm$ 1.77
	NDEL1	7	77.81 $\pm$ 42.23	5.65 $\pm$ 8.73	0.20 $\pm$ 0.05
	NDE1	3	83.35 $\pm$ 46.73	6.91 $\pm$ 6.68	0.42 $\pm$ 0.22
S713E	PAFAH1B 1	7	136.86 $\pm$ 49.43	13.69 $\pm$ 4.94	2.56 $\pm$ 1.98
	NDEL1	7	115.43 $\pm$ 71.98	5.41 $\pm$ 7.53	1.08 $\pm$ 2.26
	NDE1	3	71.92 $\pm$ 43.17	11.69 $\pm$ 15.20	2.22 $\pm$ 3.36
L807FS	PAFAH1B 1	6	131.60 $\pm$ 45.57	20.65 $\pm$ 16.08	0.42 $\pm$ 0.25
	NDEL1	6	67.19 $\pm$ 52.73	3.06 $\pm$ 1.23	0.16 $\pm$ 0.08
	NDE1	3	101.39 $\pm$ 68.39	3.15 $\pm$ 0.71	0.31 $\pm$ 0.16
β -core	PAFAH1B 1	7	132.06 $\pm$ 38.07	19.61 $\pm$ 12.45	2.71 $\pm$ 1.70
	NDEL1	6	77.63 $\pm$ 42.61	6.90 $\pm$ 7.28	0.78 $\pm$ 1.33

	NDE1	4	105.04 ±16.26	2.33 ±2.54	3.23 ±3.34
--	------	---	---------------	------------	------------

Table S2: Loop-linked cross links for the C-region constructs that were identified using the crosslinker disuccinimidyl sulfoxide (DSSO) coupled with LCMS/MS. The residues marked in red correspond to the mutations of the C-region protein. The underlined sequence represents the β -core. Amino acids marked in bold blue are the Lys residues that were identified cross-linked based mass-spectrometry

WEADLEACRL	LIQ SL QLQEA	RG SL S VEDER	<u>QMDDLEGAAP</u>	730
<u>PIPPRLHSED</u>	<u>KRKTPLKVLE</u>	<u>EWKTHLIPSL</u>	<u>HCAGGEQKEE</u>	770
SYILSAELGE	K CEDIG KKLL	YLEDQLHTAI	HSHDED LIQS	810
LRRELQMV K E	TLQAMILQLQ	PA K EAG		836

C-region construct	loop-linked (intra-subunit)
WT-C-region	K743-K747
	K819-K833
S704C	K741-K743
	K743-K747
	K747-K753
	K753-K768
	K787-K788
	K819-K833
S713E	K743-K747
	K747-K753
	K753-K768
	K781-K787
	K819-K833
L807-FS	K747-K753
	K753-K768

Table S3: Summary of rate constants for the binding of physiological partners viz. NDEL1, NDE1 and PAFAH1B1 to the different C-region variants and the β -core

		K_D (nM)	K_{on} ($\times 10^4 M^{-1}.s^{-1}$)	K_{off} ($10e^{-3} s^{-1}$)
Flo01	WT	97.6	2.9	2.8
	S704C	156.0	4.6	7.2
	S713E	73.4	6.4	4.7
	L807FS	37.2	6.0	2.2
Flm1	WT	146.3	30.7	15.3
	S704C	78.0	61.7	12.3
	S713E	81.0	60.3	12.0
	L807FS	17.9	63.0	12.6

SUPPLEMENTARY FIGURES

WT-C-region

MGHHHHHHHDYDIPTTENLYFQGWEADLEACRLLIQSLQLQEARGSLSVEDERQMDDLE
GAAPPIPPRLHSEDKRKTPCLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGK
KLLYLEDLHTAIHSHDEDLIQSLRRELQMVKETLQAMILQLQPAKEAGSG

S704C

MGHHHHHHHDYDIPTTENLYFQGWEADLEACRLLIQCLQLQEARGSLSVEDERQMDDLE
GAAPPIPPRLHSEDKRKTPCLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGK
KLLYLEDLHTAIHSHDEDLIQSLRRELQMVKETLQAMILQLQPAKEAGSG

S713E

MGHHHHHHHDYDIPTTENLYFQGWEADLEACRLLIQSLQLQEARGELSVEDERQMDDLE
GAAPPIPPRLHSEDKRKTPCLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCEDIGK
KLLYLEDLHTAIHSHDEDLIQSLRRELQMVKETLQAMILQLQPAKEAGSG

L807FS

MGHHHHHHHDYDIPTTENLYFQGRVLWEADLEACRLLIQSLQLQEARGSLSVEDERQMD
DLEGAAPPIPPRLHSEDKRKTPCLKVLEEWKTHLIPSLHCAGGEQKEESYILSAELGEKCED
IGKKLLYLEDLHTAIHSHDEDLSLSGGSSRWSG

β -core

MGHHHHHHHDYDIPTTENLYFQGDERQMDDLEGAAPPIPPRLHSEDKRKTPCLKVLEEWK
THLIPSLHCAGGEQKEESSG

Fig S1: The amino acid sequences of the DISC1 WT-C-region protein constructs, S704C, S713E, L807 frame-shift and the β -core

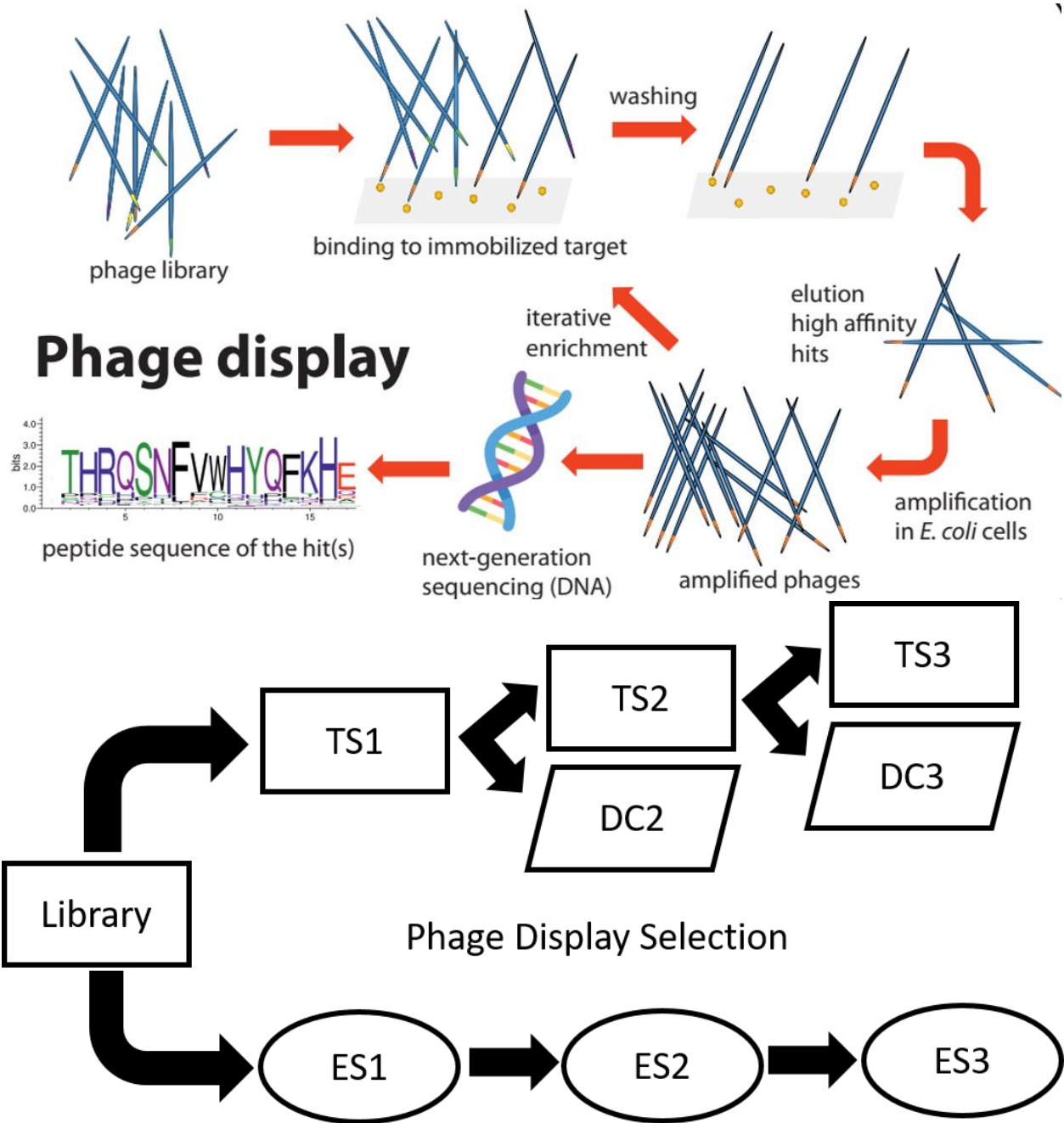


Fig S2: (A) Schematic representation of the phage display technique. (B) Two parallel selection rounds were performed to enrich peptide displaying phages during the TS1 and ES1. Direct controls (DC2 and DC3) were utilized as negative controls sequences, which were compared with target selection sequences after NGS sequencing to value which peptides were not target specific.

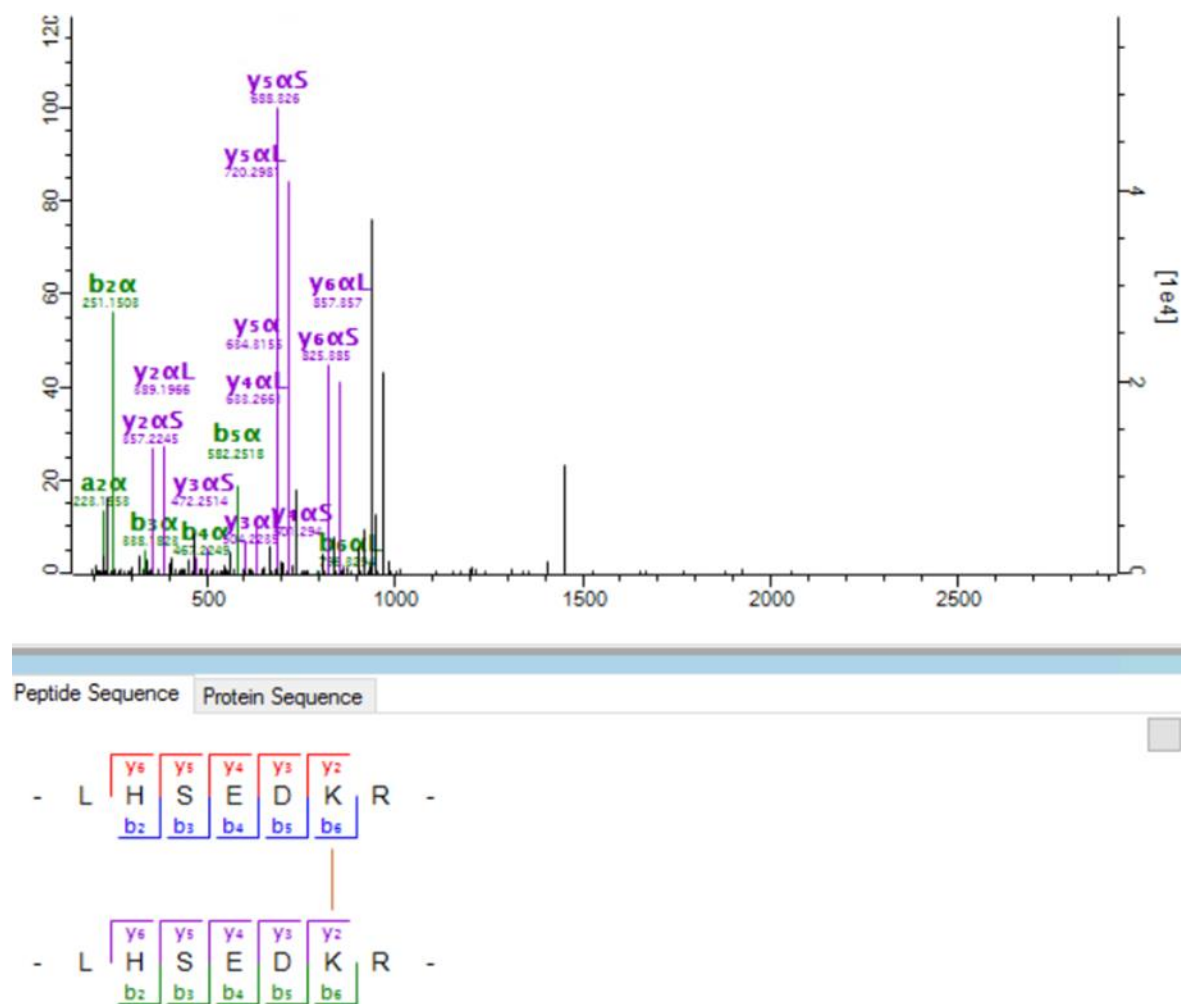


Fig S3: Fragmentation pattern and identified crosslink K741-K741 of the S704C construct

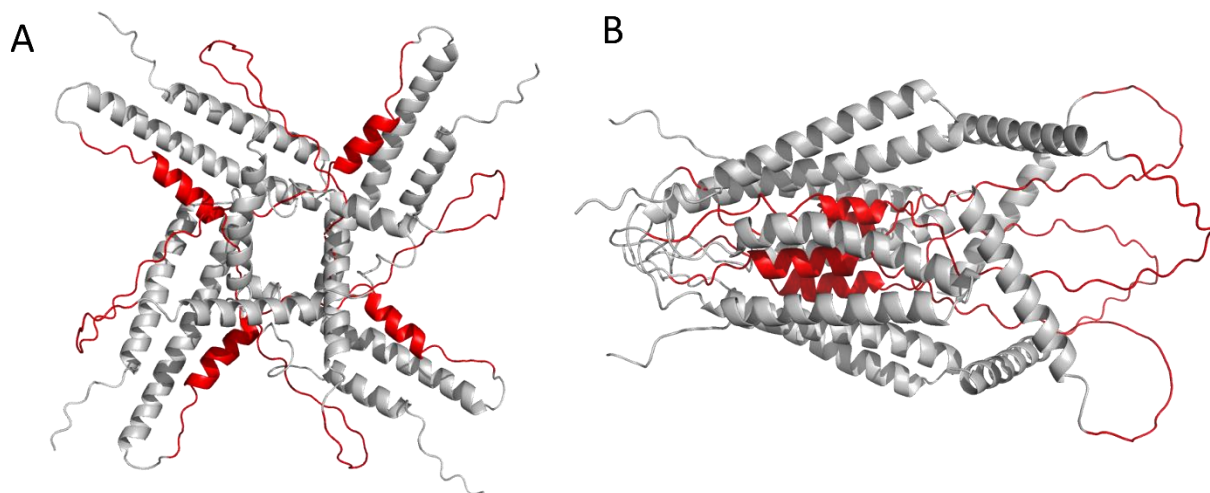


Fig S4: (A and B) Two structural models that satisfy all the restraints obtained from cross-linked MS and DLS studies apart from the intra-subunit cross-link. The region in red represents the β -core. The model in (B) is similar to the model in Fig. 3 except that K741-K741 is 21 Å. We do not, however, rule out the possibility that both models are perhaps identical and the differences may arise due to dynamic nature side chain mobility.

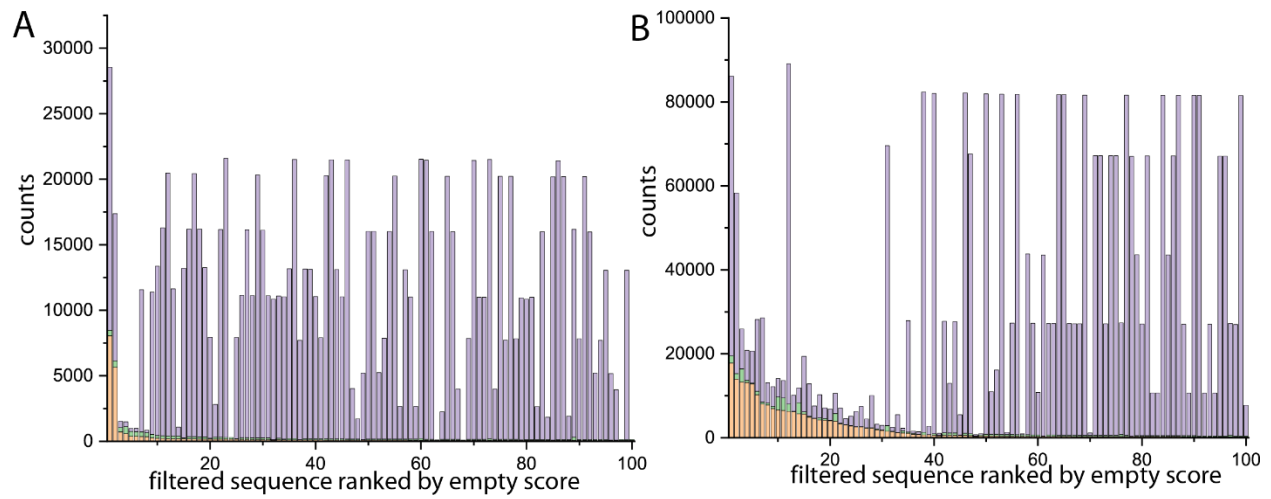


Fig S5: Sequences selected on WT C-region monomer (A) and oligomer (B). The sequences found by NGS sequencing were ranked by their empty score (orange), which is the ratio between the frequency of the respective sequence in the target selection and the empty selection. The 50 highest empty scores are depicted. The enrichment factor (green) and the respective clusters (purple) are visualized within the bar chart.

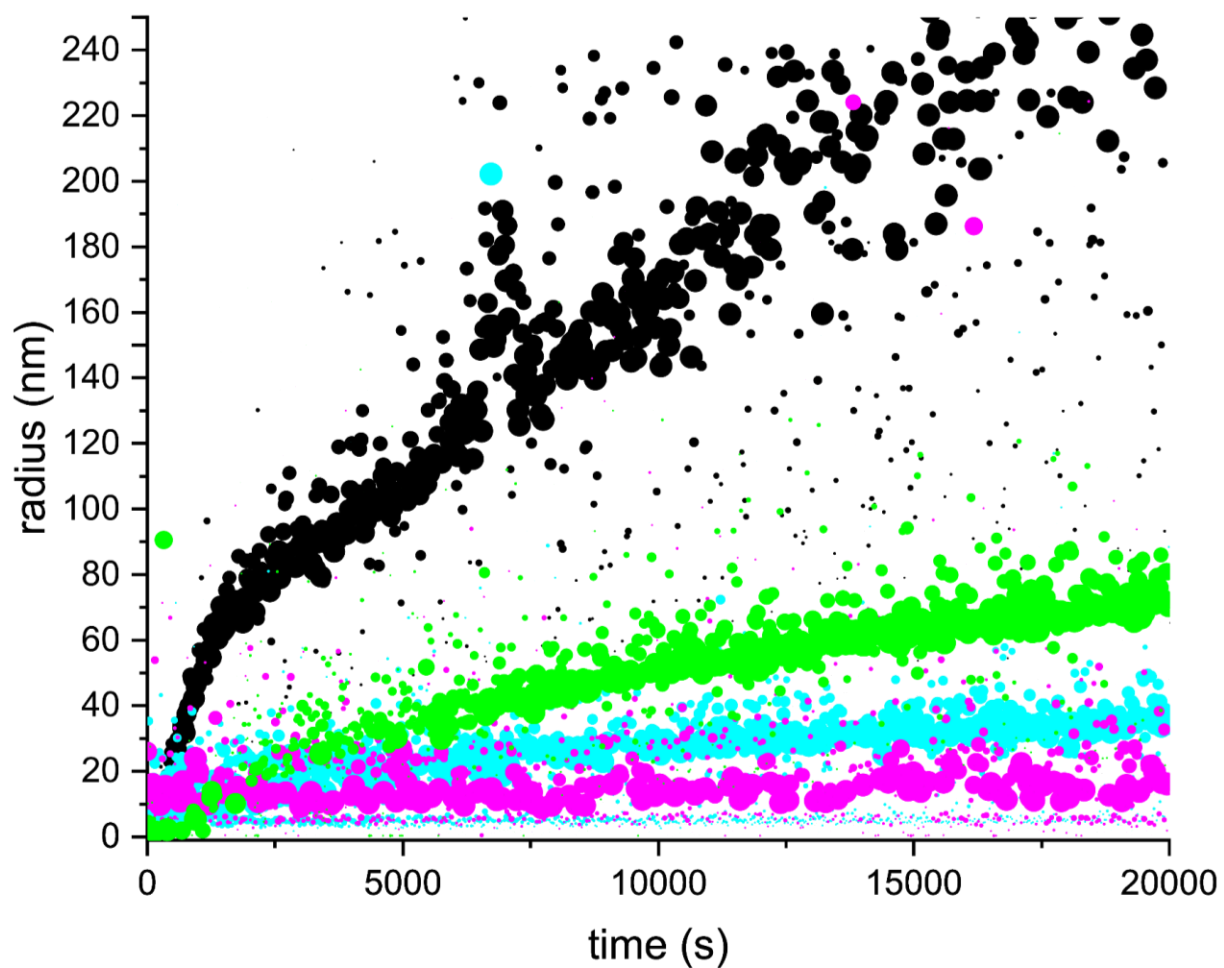


Fig S6: Kinetics of fibrillar growth of DISC1 C-region in the presence of Flo01 at concentrations of untreated (black), 10 μ M (green), 50 μ M (cyan), and 100 μ M (magenta).

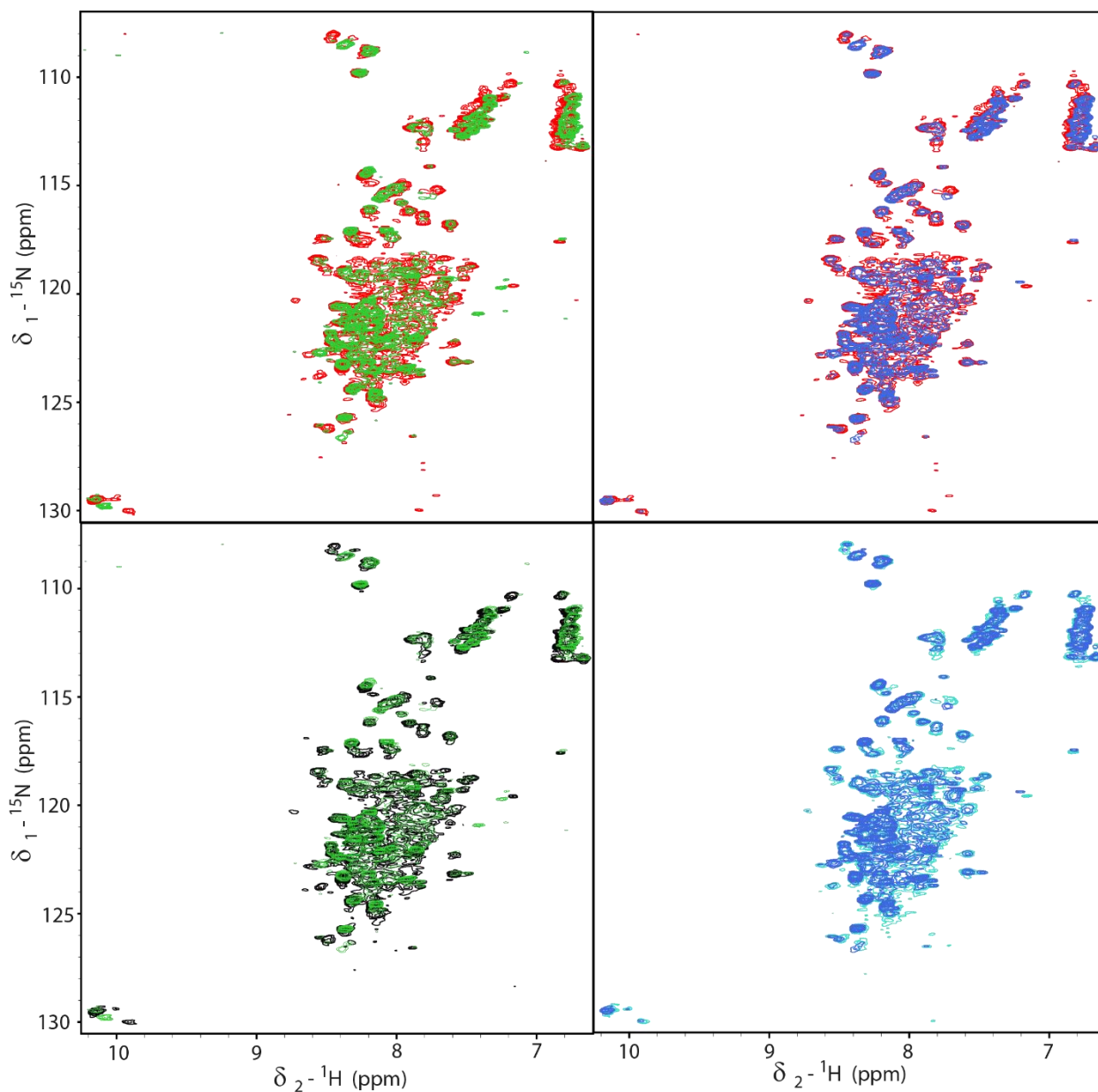


Fig S7: The top panel contains solution NMR spectra of unliganded WT-C-region (red) vs 5 mM ligand; top left panel 5 mM Flm01 (green); top right panel 5 mM Flo01 (royal blue).

The bottom panel contains spectra of the liganded states of WT-C-region offering comparison between 125 μM vs 5 mM ligand; bottom left panel describes the WT-C-region bound to Flm01 (125 μM (black) vs 5 mM (green)); bottom right panel describes the WT-C-region bound to Flo01 (125 μM (sky blue) vs 5 mM (royal blue)).

4.4 Perspective: Targeting DISC1 for the treatment of psychiatric disorders

In the last five and half years, since I began conducting research attempting to understand the pathophysiological behaviour of the DISC1 protein have provided strong evidence for its role in the development of major psychiatric disorders (Cukkemane et al., 2021). DISC1 has previously been associated with diseases such as schizophrenia, bipolar disorder, and major depressive disorder; however, the molecular mechanisms underlying these associations have remained poorly characterized. Through a combination of structural biology and biophysical methods, my group, colleagues, and I have demonstrated that the DISC1 C-region has a pronounced tendency to form amyloid-like fibrils, similar in morphology and molecular architecture to those found in neurodegenerative disorders such as Alzheimer's and Parkinson's disease (Cukkemane et al., 2021).

These findings revealed that the C-region, possesses a high degree of structural flexibility. While this conformational plasticity enables DISC1 to form diverse interactions with itself in the form of tetramer, with other cellular partners, from the mitotic spindle complex like NDE1, NDEL1, and PAFAH1B1, (formerly LIS1), it also renders the protein vulnerable to misfolding events. These misfolded species aggregate into long ordered fibrillar assemblies at low concentrations and temperatures, closely resembling classical amyloid fibrils; and form aggregates and short fibrils at higher concentrations and temperatures (Cukkemane et al., 2021).

To further explore the functional implications of DISC1 aggregation, I was curious to understand how point mutations that are clinically relevant promote mental disorders. We demonstrated that the disease-associated mutations (L807-frame shift and S704C) accelerate the transition to aggregated states (Abhishek Cukkemane et al., 2025). These aggregation-prone mutants displayed no loss-in-function when binding to the physiological partners but reduced solubility *in vitro* rendering them inactive. These findings support the notion that DISC1 C-region misfolding occurs in the wild-type, but can be rendered mutation-sensitive and biophysically tractable. One of the most significant observations from our study was the identification of a fibrillar nucleus, captured using AFM, which revealed a distinct aggregated particle from which short fibrils radiated outward. The presence of such a seed structure provides compelling evidence that DISC1 aggregation follows a nucleation-dependent polymerization pathway, akin to amyloid formation observed in neurodegenerative diseases.

Building on these mechanistic insights, I pursued a structure-rational design (A. Cukkemane et al., 2025; Cukkemane, 2022) to identify potential therapeutic interventions to prevent or reverse DISC1 C-region aggregation. For which, using phage display technology, I identified interacting peptides. These peptide mimetics demonstrated high binding affinity to the C-region and the clinically-relevant mutants. They were effectively curtailing fibril formation by reducing growth kinetics. The treated samples retained C-region in a soluble, functional tetrameric conformation, allowing for the maintenance of its protein-protein interaction capacity. These results suggest that targeted stabilization of DISC1 represents a

viable therapeutic strategy for preventing the structural transformation that underlies its pathological behaviour.

In the next steps, we will perform validation studies. Initial work in rat astrocytes cell culture to study the blood brain barrier permeability (Figure 1) indicated that a couple of peptides were able to diffuse with high-efficiency through the barrier. In the subsequent steps, we will assess the efficacy of these compounds in cellular-human iPSC (Ye et al., 2017), which we will purchase from the NIH and in vivo model like the DISC1 Boymaw mouse (Zhou et al., 2010) model (<https://www.jax.org/strain/030517>) that we have already procured from Jackson Laboratory, USA. The long-term objective of this research is to advance these findings into pre-clinical development and, ultimately, to initiate clinical trials for peptide-based therapeutics addressing the molecular basis of psychiatric disorders.

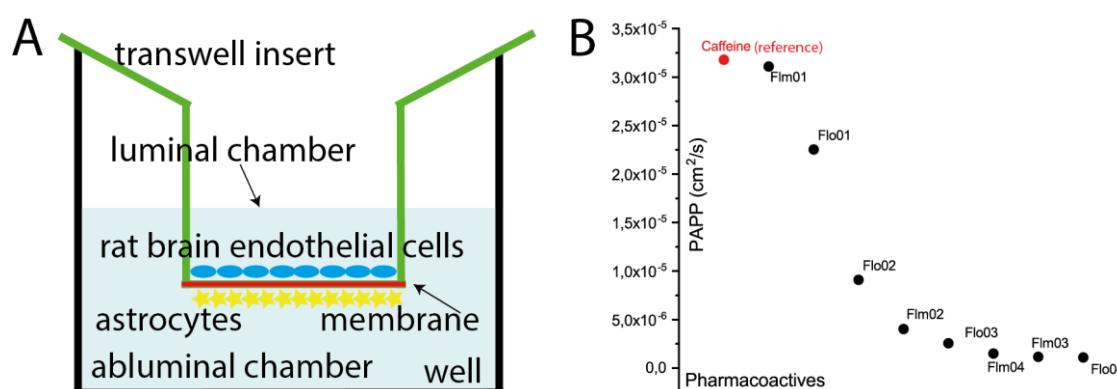


Figure 1: (A) Schematic representation of an in-vitro blood brain barrier testing- (B) The results highlights that few of the peptides (PCT/EP2023/066796) inherently can readily cross-over the blood brain barrier.

In parallel, my colleague Dr. Taras Balandin and I are performing Molecular Dynamics (MD) simulations using GROMACS, which is vital in our understanding of translational drug discovery for bridging the gap between basic molecular research and clinical application. These simulations are being performed at the Jülich Super Computing facility and can offer atomic-level insights into the dynamic behaviour of C-region interaction with peptide drugs (e.g., Flm 01, Fig 02) over time. Our inkling here is that MD simulation results can further feed into structure-based drug design, virtual screening, and fragment-based drug discovery pipelines. Hopefully, integration of GROMACS-based MD into AI/ML

models and pharmacophore development further accelerates the translational outcomes of DISC1 therapeutics.

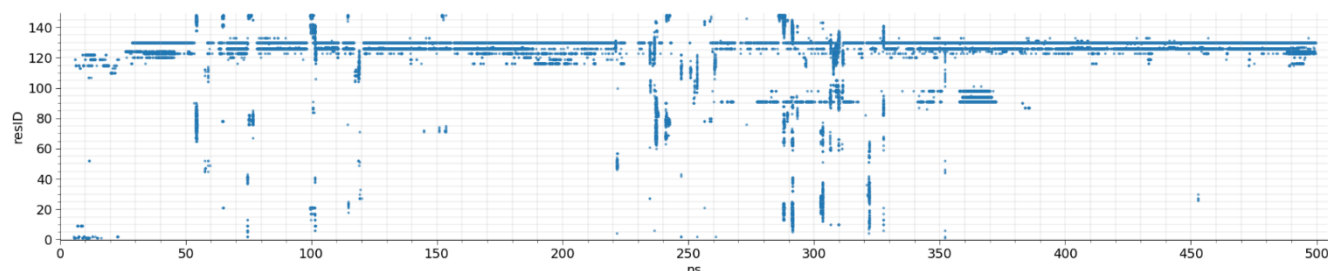


Fig 02: Contact map highlighting the interfacial behaviour between the DISC1 C-region and Flm01 in 0.5 ns simulation. Residues in the 110–140 range maintain persistent contacts within a distance of 3 Å throughout the trajectory suggesting a stable interface of the protein-peptide complex.

Further extending our studies for the DISC1 protein beyond the C-region, a post-graduate student from the HHU, Ms. Sophie Deselaers (currently pursuing her PhD studies at ETH Zürich) under my supervision studied the I-region (Yerabham et al., 2013) of the DISC1 protein. Unlike the C-region, I-region was insoluble and was expressed as inclusion bodies in *E. coli* cells. Using denaturing and refolding techniques, we succeeded in obtaining in protein that permitted biophysical and structural biology studies. The I-region oligomerized as hexamers (Fig 03) in contrast to the tetrameric arrangement of the C-region. However, the I-region was also prone to aggregation and fibrillization.

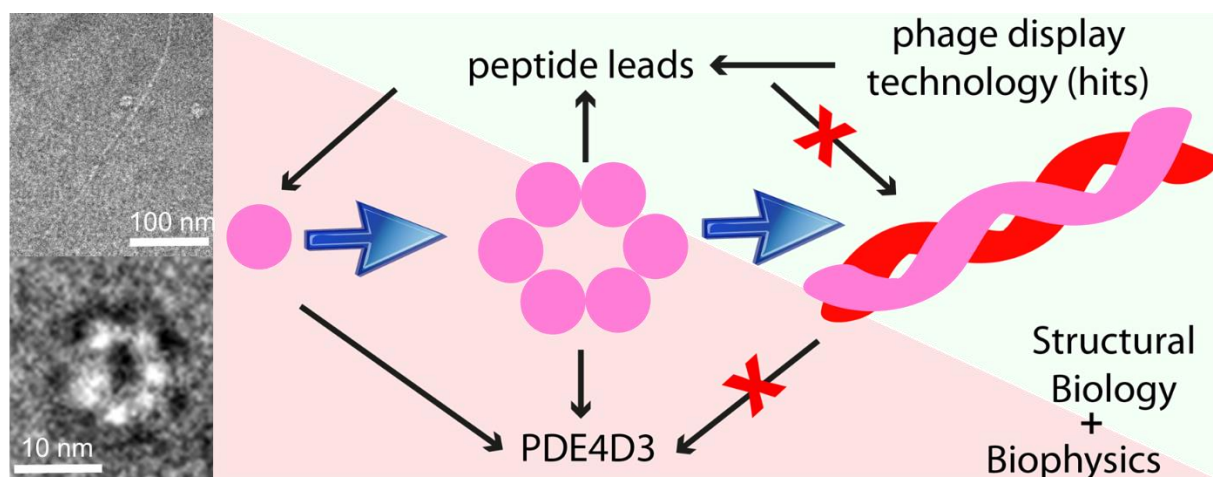


Fig 03: Left Panel contains electron micrographs of the fibrillar and the hexameric assembly of the I-region. The right panel highlights the strategy we wish to undertake to optimize drug leads.

As the next step, we plan to identify peptide leads (Fig 03) using phage display technology and study the impact of the interaction of the I-region with phosphodiesterase 4D3.

To sum all of the above, the integration of structural biology, molecular biophysics, protein biochemistry, animal models and medicinal chemistry in the upcoming future may represent a

significant advancement in the field of neuropsychiatric research from the lens of DISC1. By elucidating the structural vulnerabilities of DISC1 and demonstrating their pathological and therapeutic relevance, the studies I have conducted thus far, and will further pursue along with my colleagues will provide a foundation for potential treatments that target an important risk factor in chronic mental illness.

References

- Cukkemane, A., Becker, N., Kupreichyk, T., Heise, H., Willbold, D., & Weiergräber, O. H. (2025). Tracing the aggregation pathway of the scaffold protein DISC1: Structural implications for chronic mental illnesses. *Journal of Structural Biology: X*, 11, 100128.
<https://doi.org/https://doi.org/10.1016/j.yjsbx.2025.100128>
- Cukkemane, A., Becker, N., Zielinski, M., Frieg, B., Lakomek, N. A., Heise, H., Schroder, G. F., Willbold, D., & Weiergraber, O. H. (2021). Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. *Transl Psychiatry*, 11(1), 639. <https://doi.org/10.1038/s41398-021-01765-1>
- Cukkemane, A., Dingley, A. J., Mohrluder, J., Santiago-Schubel, B., Weiergraber, O. H., & Willbold, D. (2025). A peptide mimetic therapeutic strategy targeting dysfunction of the scaffold protein DISC-1 in psychiatric disorders. *Eur J Pharm Sci*, 107148.
<https://doi.org/10.1016/j.ejps.2025.107148>
- Cukkemane, A., Willbold, D. (2022). *DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia, major depressive disorders (MDD), and bipolar disorders (BD), and autism and other chronic mental disorders (CMDs)* (EPO Patent No.
- Ye, F., Kang, E., Yu, C., Qian, X., Jacob, F., Yu, C., Mao, M., Poon, R. Y. C., Kim, J., Song, H., Ming, G. L., & Zhang, M. (2017). DISC1 Regulates Neurogenesis via Modulating Kinetochore Attachment of Ndel1/Nde1 during Mitosis. *Neuron*, 96(5), 1204.
<https://doi.org/10.1016/j.neuron.2017.11.034>
- Yerabham, A. S., Weiergraber, O. H., Bradshaw, N. J., & Korth, C. (2013). Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biol Chem*, 394(11), 1425-1437.
<https://doi.org/10.1515/hsz-2013-0178>
- Zhou, X., Chen, Q., Schaukowitch, K., Kelsoe, J. R., & Geyer, M. A. (2010). Insoluble DISC1-Boymaw fusion proteins generated by DISC1 translocation. *Mol Psychiatry*, 15(7), 669-672.
<https://doi.org/10.1038/mp.2009.127>

5. Biochemical pathways and metabolomics in understanding schizophrenia

Among the few functional genomics techniques, metabolomics (Vignoli et al., 2019; Wishart, 2019) offers a comprehensive snapshot of the biochemical state of a biological system. Among its analytical platforms, mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy are the most widely employed techniques. While MS has garnered significant attention due to its exceptional sensitivity for detecting low-abundance metabolites, NMR-based metabolomics has emerged as a robust, non-invasive, and highly reproducible alternative. In the context of psychiatric research in our group for clinical applications, NMR provides critical advantages as it requires minimal sample preparation, allows absolute quantification as per the EU-IVDR guidelines, and exhibits high reproducibility across instruments and cohorts. Moreover, due to its no-destructive nature, one can recover the sample for other studies. Hence NMR is increasingly favoured in clinical and population-level studies due to its compatibility with automation, rapid processing, and low inter-laboratory variability.

In this introductory chapter, we will discuss the technical and clinical advantages of NMR metabolomics with a particular focus on bipolar disorder. We will explore (1) sample preparation strategies that preserve metabolic integrity and improves resolution; (2) translational applications for preclinical and clinical models; and (3) the potential of multi-omics integration using NMR.

Together, these topics will illustrate how NMR metabolomics can evolve into a core platform for personalized diagnostics and treatment planning in psychiatry.

Sample preparation

In NMR-based metabolomic studies, plasma and serum are commonly used biofluids due to their accessibility and clinical relevance. However, these fluids contain high levels of proteins and lipids that can obscure the spectral signatures of low-molecular-weight metabolites. Moreover, enzymatic activity in unprocessed samples may alter metabolite concentrations during measurement, compromising data accuracy. In our opinion, protein removal is thus a crucial step in sample preparation. While precipitation methods using organic solvents like methanol or acetonitrile are common (Tiziani et al., 2008), they may lead to the partial loss of hydrophobic or weakly bound metabolites. Comparative studies have shown that ultrafiltration (Tiziani et al., 2008), despite its higher cost, offers superior metabolite recovery and reproducibility. It effectively separates proteins and large lipid vesicles without significantly altering the metabolite composition, making it ideal for applications requiring absolute quantification and longitudinal consistency.

In our work, we have adopted a chloroform-methanol extraction method originally proposed by Bligh and Dyer (Bligh & Dyer, 1959) for food analysis. This technique denatures proteins, removes lipids, and retains polar metabolites, thereby producing high-resolution spectra. It is especially advantageous for processing brain tissue samples from post-mortem biobanks, where controlling enzymatic degradation during thawing and extraction is vital. For large-scale, cohort-based samples, such as those available at the UK Biobank and KORA (Wichmann et al., 2005), standardized operating procedures (SOPs) are essential to ensure consistency across thousands of samples. When optimized for NMR, these SOPs preserve sample integrity and permit for cross-cohort comparison. This consistency is critical for identifying reproducible biomarkers of across spectra of CMIs.

Translational applications and Multi-omics integration

In large human cohorts, plasma NMR profiling has revealed metabolic signatures associated with a range of psychiatric and somatic disorders. For instance, in the UK Biobank and other studies like Nightingale Health's Atlas (Julkunen et al., 2023) <https://research.nightingalehealth.com/>, NMR-

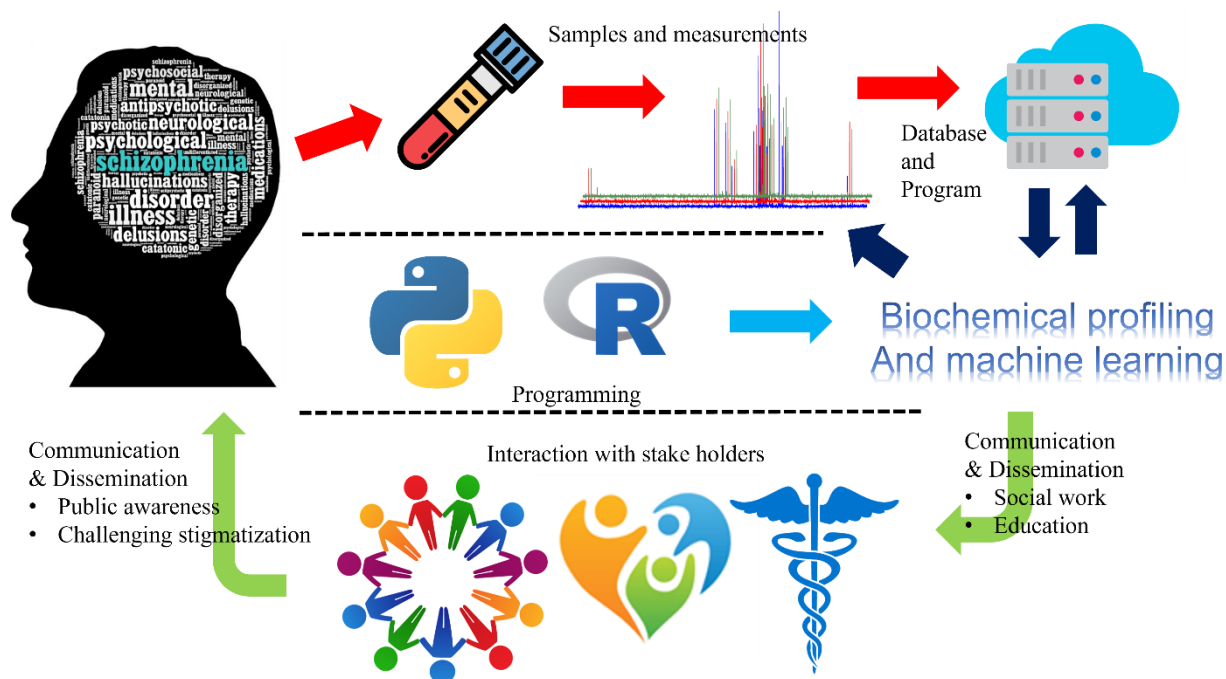
derived biomarkers have been linked a variety of disorders and can be filtered on the basis of their ICD-code including mood disorders, cognitive decline, and neurodegenerative conditions. These datasets highlight the potential of NMR to facilitate population-wide screening, risk stratification, and precision medicine approaches in psychiatry.

Beyond metabolomics, multi-omics integration amplifies the utility of NMR. For instance, metabolome-wide genome-wide association studies (mGWAS) combine NMR spectral data with single nucleotide polymorphism (SNP) genotyping to reveal metabolite-genotype associations (Raffler et al., 2013; Suhre et al., 2011) in the KORA cohort. These findings link genetic polymorphisms to biochemical phenotypes, offering a framework to understand heterogeneity in psychiatric presentations. Such genetic-metabolic insights can inform pharmacogenomic strategies, tailoring medication to a patient's unique biochemical and genetic profile.

Furthermore, the integration of NMR-based data into machine learning and artificial intelligence pipelines holds significant promise. Predictive models can leverage multi-omics data to classify disease subtypes, forecast relapse, or identify treatment responders. The automated, reproducible, and scalable nature of NMR makes it especially compatible with real-time clinical decision-making tools. The future of psychiatry may thus lie in platforms that combine biochemical precision with computational intelligence.

Toward a systems biology framework for CMIs

Given its strengths in accuracy, scalability, and translational relevance, NMR metabolomics provides several distinct advantages. As part of our long-term vision, we have designed a workflow itinerary (Figure 1) that integrates NMR-based metabolomics with clinical diagnostics for conditions like schizophrenia (section 5.1), and sleep disorders (section 5.2 and 5.3).



This roadmap includes extending our studies to other psychiatric conditions, such as DISC1opathies (Korth, 2012), and to other chronic mental illnesses (CMIs) such as bipolar disorders (BD)(Lan et al., 2009) and major depressive disorder (MDD). The goal is to shift from descriptive symptom-based

classifications to biologically anchored subtypes, enabling early intervention and personalized care pathways by adopting a systems biology perspective, supported by high-throughput NMR analytics.

References

- Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol*, 37(8), 911-917. <https://doi.org/10.1139/o59-099>
- Julkunen, H., Cichonska, A., Tiainen, M., Koskela, H., Nybo, K., Makela, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C., & Wurtz, P. (2023). Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*, 14(1), 604. <https://doi.org/10.1038/s41467-023-36231-7>
- Korth, C. (2012). Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion*, 6(2), 134-141. <https://doi.org/10.4161/pri.18989>
- Lan, M. J., McLoughlin, G. A., Griffin, J. L., Tsang, T. M., Huang, J. T., Yuan, P., Manji, H., Holmes, E., & Bahn, S. (2009). Metabonomic analysis identifies molecular changes associated with the pathophysiology and drug treatment of bipolar disorder. *Mol Psychiatry*, 14(3), 269-279. <https://doi.org/10.1038/sj.mp.4002130>
- Raffler, J., Romisch-Margl, W., Petersen, A. K., Pagel, P., Blochl, F., Hengstenberg, C., Illig, T., Meisinger, C., Stark, K., Wichmann, H. E., Adamski, J., Gieger, C., Kastenmuller, G., & Suhre, K. (2013). Identification and MS-assisted interpretation of genetically influenced NMR signals in human plasma. *Genome Medicine*, 5(2), 13. <https://doi.org/10.1186/gm417>
- Suhre, K., Wallaschofski, H., Raffler, J., Friedrich, N., Haring, R., Michael, K., Wasner, C., Krebs, A., Kronenberg, F., Chang, D., Meisinger, C., Wichmann, H. E., Hoffmann, W., Volzke, H., Volker, U., Teumer, A., Biffar, R., Kocher, T., Felix, S. B., . . . Nauck, M. (2011). A genome-wide association study of metabolic traits in human urine. *Nat Genet*, 43(6), 565-569. <https://doi.org/10.1038/ng.837>
- Tiziani, S., Emwas, A. H., Lodi, A., Ludwig, C., Bunce, C. M., Viant, M. R., & Gunther, U. L. (2008). Optimized metabolite extraction from blood serum for ¹H nuclear magnetic resonance spectroscopy. *Anal Biochem*, 377(1), 16-23. <https://doi.org/10.1016/j.ab.2008.01.037>
- Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P., & Luchinat, C. (2019). High-Throughput Metabolomics by 1D NMR. *Angew Chem Int Ed Engl*, 58(4), 968-994. <https://doi.org/10.1002/anie.201804736>
- Wichmann, H. E., Gieger, C., Illig, T., & Group, M. K. S. (2005). KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen*, 67 Suppl 1, S26-30. <https://doi.org/10.1055/s-2005-858226>
- Wishart, D. S. (2019). Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*, 99(4), 1819-1875. <https://doi.org/10.1152/physrev.00035.2018>

Reconstructing schizophrenia's biochemistry through hybrid diagnostics: clinical and metabolic profiling

Abhishek Cukkemane^{1,2,3*}, Danilo Postin^{4,5}, Kevin Bochinsky^{1,6}, Simon Barton^{4,5}, Dirk Scheele^{4,5,7}, Gabi Kastenmüller⁸, Dieter Willbold^{1,2}, and Rene Hurlemann^{4,5}

¹Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

³Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

⁴Department of Psychiatry, School of Medicine and Health Sciences, University of Oldenburg, Oldenburg, Germany.

⁵Karl-Jaspers-Klinik, Bad Zwischenahn, Germany

⁶current address -Bruker Biospin, Ettlingen, Germany

⁷current address - Medizinische Fakultät, Ruhr-Universität Bochum, Bochum, Germany

⁸ Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, Neuherberg, Germany.

*Address for correspondence - a.cukkemane@fz-juelich.de

Abstract: Schizophrenia is a chronic mental illness (CMI), which characterized by a complex interplay of positive, negative, cognitive, and affective symptoms. While genome-wide association studies (GWAS) have identified common genetic variants associated with schizophrenia, their clinical utility remains limited. This study employs high-resolution quantitative nuclear magnetic resonance (qNMR) spectroscopy to investigate metabolomic signatures in plasma and urine samples from individuals with schizophrenia and integrating these findings with symptom-specific clinical scales including PANSS, BNSS, MADRS, and GFS/GFR. Our results reveal significant alterations in several pathways such as amino acid and nitrogen metabolism, impairment in neurotransmitter synthesis and energy regulation, membrane turnover and circadian rhythm. Importantly, several metabolites showed domain-specific correlations with clinical symptoms: for instance, creatinine levels were positively associated with MADRS scores, while glutamate levels aligned with PANSS-positive symptoms among others. These findings demonstrate that schizophrenia is not a unitary ailment a spectral disorder, which is biologically heterogeneous condition with distinct metabolic correlates across symptom dimensions. The integration of metabolomics with clinical phenotyping offers a non-invasive and scalable approach for identifying biomarkers that will guide precision diagnostics and personalized interventions. This study lays a foundational framework for future research aimed at clinical translation, ultimately contributing to more targeted treatment strategies and a deeper understanding of schizophrenia's complex pathophysiology.

Introduction:

Schizophrenia is a debilitating chronic mental illness (CMI) that is heritable and shares genetic vulnerability to bipolar disorders (BD), major depressive disorders (MDD)(Blackwood et al., 2001; International Schizophrenia

et al., 2009; O'Donovan & Owen, 2016; Owen et al., 2007). The disorder is characterized by a wide array of symptoms spanning positive (psychotic such as hallucinations and delusions), negative (such as social withdrawal and anhedonia), cognitive (including impairments in memory and attention), and affective domains (with depressive and manic features, though less pronounced than in bipolar disorder). These clinical features form the basis of the Positive and Negative Syndrome Scale (PANSS)(Kay et al., 1987), which remains the gold standard for the assessment of psychotic disorders. However, distinguishing related CMIs remains complex (Paul & Potter, 2024; Smith & Craddock, 2011) and pose a significant burden to the health management and society (Canada, 2012; OECD/EU, 2018; Wiegand et al., 2020). Therefore, there is a large unmet need for detection and improved early detection strategies.

It is reasonable to assume that due to high heritability of schizophrenia, it encompasses a strong genetic component in its aetiology. Genome-wide association studies (GWAS) have shown that common genetic variants can account for approximately one-third to one-half of the genetic risk, making GWAS a valuable tool in elucidating the biological underpinnings of the disorder (Ripke et al., 2013; Schizophrenia Working Group of the Psychiatric Genomics, 2014). Although, polygenic risk scores (PRS) derived from GWAS have predictive ability, their clinical utility remains limited (Schizophrenia Working Group of the Psychiatric Genomics, 2014). Moreover, this approach has other limitations, for instance, rare genetic mutations are not typically captured in GWAS due to their low population frequency. For instance, Disrupted in Schizophrenia 1 has not been identified among the major GWAS hits, even though compelling evidence from familial studies that mutations in this gene contribute to disease pathology (Porteous et al., 2014; Sullivan, 2013). This mirrors findings in neurodegenerative disorders like Alzheimer's disease (AD), where few key genes involved are rarely highlighted in GWAS screens, despite being central to the disease mechanism (Bertram et al., 2010). Furthermore, a poor correlation between clinical and genomic findings has been repeatedly observed; The assessment of the contribution of risk-factor genes to the statistical variation of the disorders remain challenging (Chen et al., 2019). Beyond genetic influences, non-genetic factors play a critical role in the etiology of schizophrenia and related CMIs. One such factor is disrupted protein homeostasis, which includes processes like protein misfolding, misassembly, and aggregation. These molecular pathologies involving proteins like DISC1 or other neurodegenerative risk factors is not well-captured by GWAS highlighting the importance of exploring alternative biomarker detection strategies to fully understand the complexity of schizophrenia.

On these lines, alternative approaches in proteomics (Martins-de-Souza et al., 2011) and metabolomics (Vignoli et al., 2019; Wishart, 2019) are invaluable because they offer insights into phenotype by analyzing proteins and metabolites, respectively. Proteomics studies using mass-spectrometry in schizophrenia have contributed to our understanding of important biochemical pathways many of which are immunological in origin (Schwarz et al., 2010; Sialana et al., 2018). Despite its strength in high-sensitivity and phenotyping ability, proteomics research has yet to yield a reliable diagnostic tool in schizophrenia and related CMIs (Chan et al., 2015; Korth & Fangerau, 2020; Schwarz et al., 2010); despite identifying numerous proteins with potential relevance, the lack of reproducibility and high associated costs have prevented translation into clinical practice.

Metabolomics offers a promising complementary approach to preventive health care and clinical diagnostics. A notable success story is the Nightingale Health platform (Julkunen et al., 2023), which provides comprehensive information on blood-based biomarkers. Thanks to the simplicity of sample preparation, measurement, and analysis, metabolic biomarker profiling has been widely adopted in both research and diagnostic settings. Inspired by this, we employed quantitative nuclear magnetic resonance (qNMR) spectroscopy to identify and quantify metabolites capable of distinguishing individuals with schizophrenia from healthy controls. Additionally, we compared our metabolic findings with clinically relevant scales, including PANSS (Kay et al., 1987), Brief Negative Symptom Scale (BNSS, (Kirkpatrick et al., 2018)), Montgomery–Åsberg Depression Rating Scale (MADRS, (Montgomery & Åsberg, 1979)), and Global Functioning Scales: Social and Role (GFS /GFR, (Carrion et al., 2019; Piskulic et al., 2011)). This integrative approach offers insights into which combinations of metabolites are associated with specific clinical parameters. Importantly, these biomarkers are non-disruptive and can be readily incorporated into existing clinical workflows.

Methodology

Study design

The study secured approval from the Medizinische Ethikkommission, Carl von Ossietzky Universität Oldenburg (Identification-Nr.:2020-149) and the Medizinische Fakultät Ethikkommission, Heinrich-Heine-Universität Düsseldorf (Identification-Nr.:2022-1813). All participants provided written informed consent and the study was prospectively registered at the Karl-Jaspers Psychiatry clinic, Bad Zwischenahn, Germany.

Participant recruitment: We rigorously screened participants' physical and psychological health using questionnaires and medical examinations. The diagnosis of schizophrenia is made according to DSM-V (Diagnostic Statistical Manual of Mental Disorders-V)(Association, 2013). The diagnosis of high-risk clinical psychosis is made according to the recommendations of the European Psychiatric Association Guidance Project (Campion et al., 2012).

The diagnosis of schizophrenia was established according to standard clinical criteria. To meet the diagnostic threshold, at least two or more characteristic symptoms had to be present for a significant portion of a one-month period, or for a shorter duration if successfully treated. Crucially, at least one of the core symptoms, i.e., delusions, hallucinations, or disorganized speech was required. Additional symptoms that could support the diagnosis included grossly disorganized or catatonic behaviour and negative symptoms such as diminished emotional expression or avolition.

In addition to symptom presentation, a marked decline in functioning in one or more major areas such as work, social relationships, or self-care had to be evident over a significant period of time following the onset of the disorder. The total duration of disturbance was required to be at least six months, with one month of active-phase symptoms included within this timeframe. Diagnoses of schizoaffective disorder or mood disorders with psychotic features were ruled out, and symptoms were not attributable to substance use or medical conditions.

For individuals at clinically high risk for psychosis, diagnosis was based on the presence of specific early indicators, which included at least one attenuated psychotic symptom, such as delusional ideas held with incomplete conviction. Additionally, the presence of at least two basic cognitive disturbances were experienced and reported by the individual as novel since their premorbid state was necessary. These symptoms, such as thought interference, had to occur at least weekly over the past three months and could not be explained by the effects of a substance. Alternatively, a diagnosis could be made if the individual exhibited at least one transient psychotic symptom, classified as a Brief Limited Intermittent Psychotic Symptom (BLIPS), such as spontaneously remitting hallucinations or intense delusions. In all cases, a current or past psychotic episode was ruled out, as were symptoms resulting from physical illness.

Inclusion and exclusion criteria Participants were included in the study provided they met specific eligibility requirements. All participants were required to demonstrate willingness and the ability to provide written informed consent. Individuals diagnosed with schizophrenia or those meeting criteria for a high-risk clinical stage of psychosis were eligible for inclusion. Healthy first-degree relatives of schizophrenia patients could also be included, provided they had no current neurological or psychiatric illnesses and no history of schizophrenia or affective psychosis. Healthy control subjects were required to have no current or past neurological or psychiatric conditions, no history of schizophrenia or affective psychosis, and no first-degree relatives with a known diagnosis of schizophrenia or affective psychosis. All participants had to be between 18 and 65 years of age to control for potential age-related variations in social-emotional functioning. Furthermore, participants were required to have a very good command of the German language to ensure comprehension of study procedures and accurate self-reporting.

Exclusion criteria were applied to ensure participant safety and data quality. Individuals with contraindications to MRI scanning, such as the presence of a pacemaker, pregnancy, or severe claustrophobia, were excluded. Additional exclusion criteria included acute suicidality, substance dependence within the past six months, a history of significant brain trauma or other neurological condition.

Sample Preparation

Plasma metabolites were analysed following extraction using a modified Bligh and Dyer method (Bligh & Dyer, 1959). Each 600 μ L plasma sample was mixed with an equal volume of a 1:1 chloroform:methanol solution and vortexed for one minute at room temperature to promote efficient extraction. The mixture was then centrifuged at 10,000 g to achieve phase separation. The resulting aqueous (upper) layer was carefully collected and extracted again by adding an equal volume of chloroform, followed by another one-minute vortexing at room temperature and centrifugation at 10,000 g. After the second phase separation, the purified aqueous phase was collected and dried under vacuum using a SpeedVac concentrator set at 25 °C. The dried extract was then reconstituted in 660 μ L of an aqueous buffer containing 10 mM sodium phosphate (pH 7.4), 150 mM sodium chloride, 10% (v/v) D₂O, and 200 μ M trimethylsilylpropanoic acid (TSP), which served as an internal reference for chemical shift calibration and quantification in subsequent analyses.

Urine samples were centrifuged at 10,000 g for 5 minutes to remove cellular debris and particulates. Subsequently, 540 μ L of the supernatant was mixed with the following components to a final volume of approximately 660 μ L with 10 mM sodium phosphate buffer (pH 7.4), 150 mM sodium chloride, 10% (v/v) D₂O, and 200 μ M TSP.

NMR spectroscopy, data processing and metabolic profiling

Nuclear magnetic resonance (NMR) spectroscopy was performed using a Bruker Avance 900 MHz spectrometer equipped with a 5 mm probe. Spectral acquisition was conducted at a constant temperature of 293 K (20 °C), employing a standard one-dimensional 1D- ¹H-NMR pulse sequence with water pre-saturation using a square-shaped pulse. D₂O was used to provide the field-frequency lock, and chemical shifts (CS) were referenced to the internal standard TSP, set at 0 ppm. Spectral visualization and batch processing were carried out using TopSpin software (version 4.0.6, Bruker BioSpin Ltd.).

The 1D ¹H-NMR spectra served as metabolic fingerprints, capturing a broad profile of plasma metabolites. Prior to annotation, all spectral data underwent comprehensive pre-processing, including Fourier transformation, manual phasing, baseline correction via Chenomx spline fitting, deletion of the water resonance region, shim correction (adjusted to a line width of 1.2 Hz), apodization using a 1 Hz exponential line broadening function, and final referencing to the TSP peak. Metabolites were then identified and quantified using the batch processing and legacy auto-fitting functions within the Chenomx NMR Suite Professional v10 software (Chenomx Inc., Edmonton, Canada), referencing a curated library of 330 known metabolites to support untargeted metabolomic profiling.

Statistical Analysis

To investigate metabolic features associated with sleep diagnostics, statistical analyses were performed using R software. Prior to modelling, the dataset was carefully examined for outliers and curated to ensure data integrity. To account for metabolite dilution and their specific representation in urine samples, probabilistic quotient normalization (PQN)(Dieterle et al., 2006) was applied to correct for technical variability and sample dilution effects.

Thereafter, metabolite concentrations (both urine and plasma) were log₂-transformed to stabilize variance and normalize the distribution. Mean concentrations were calculated, and linear mixed-effects models were applied to evaluate associations. In these models, the time points of sample collection were treated as fixed effects, while participant age and sex were included as covariates. Individual participants were modelled as random effects to account for within-subject variability. A significance threshold of $p < 0.05$ was used to identify metabolites of potential diagnostic relevance.

Results

Plasma metabolomics

To compare the metabolites that are significantly different between the healthy controls and subjects suffering from schizophrenia, we performed linear modelling where metabolites were tested with condition of the

participant. Age and sex were used as covariates in the model. Figure 1A shows the metabolic differences between schizophrenia and control groups (Figure 1) where significance thresholds were set at an adjusted p-value < 0.05 and fold change > ± 1.5 . Among them, we observed a significant downregulation of metabolites such 2-Aminobutyrate, Phenylalanine, Arginine, Guanidinoacetate 2-Oxoisocaproate and Glutamine indicating disruption in urea cycle, creatine biosynthesis, branched chain amino acid (BCAA) metabolism altogether indicating altered nitrogen metabolism that is reminiscent of depression (De Deyn et al., 2009; Pu et al., 2025). Increased levels of Acetaminophen, Ibuprofen, and Gentisate suggest the presence of pharmaceuticals or metabolites thereof, potentially reflecting medication (See Table S1 for the full list of medication administered to patients) use within the study population. Dimethyl sulfone is commonly associated with sulfur metabolism or dietary intake, and its reduction may indicate differences in exposure or metabolic processing. Neuroactive metabolites like N-Acetylserotonin, 5-Hydroxyindole-3-acetate, and 3-Hydroxykynurenine are components of the serotonin and tryptophan pathways, implying potential links to neurochemical changes and/ or sleep-wake regulation, especially relevant in sleep diagnostics. We also noticed changes in several amino acids and derivatives are differentially regulated, such as Threonine, Arginine, Methionine, Phenylalanine, and Tyrosine derivatives. These findings suggest act<five changes in protein turnover, methylation potential, or neurotransmitter synthesis. A hierarchical clustering heatmap (Figure S1) was generated to visualize patterns of metabolite abundance across all plasma samples, and a correlation matrix (Figure S2) to examine pairwise associations among the quantified plasma metabolites. Using these significantly altered metabolites as input, a biochemical interaction network (Figure 2) was constructed highlighting a visual map of metabolic pathway disruptions.

Urine metabolomics

Similar to studies with plasma metabolomics, urine samples were subjected to NMR-based metabolomic profiling to explore systemic metabolic output differences. Given its role in eliminating metabolic waste and capturing diet, microbiota, and physiological status, urine offers a unique perspective on metabolomic changes. However, prior to performing linear modelling, we corrected for dilution effects using PQN (Dieterle et al., 2006). The metabolites that were upregulated in schizophrenia were Asparagine and O-Phosphoethanolamine (Figure 1B and Figure S3, and S4). Few downregulated metabolites included Glycine, Melatonin, Ethylene glycol, and Ribose reflecting altered neuroactive and energy metabolic pathways.

Metabolomics correlation with clinical scales

One of the biggest challenges in schizophrenia and related CMIs present as a spectrum of disorders with significant heterogeneity in symptom expression and severity. In clinical practice and pharmaceutical research, various standardized scales are employed to assess symptom severity and functional impairment. However, the use of multiple and sometimes overlapping scales such as PANSS (Kay et al., 1987), BNSS, (Kirkpatrick et al., 2018), MADRS, (Montgomery & Åsberg, 1979), and GFS and GFR, (Carrion et al., 2019; Piskulic et al., 2011) that complicates cross-scale comparison and data integration across clinical settings. To address this, we focused not

only on group-level comparisons between controls and patients, but on identifying metabolite-level differences associated with symptom severity as defined by each clinical scale. Linear modelling was performed independently for each scale to identify metabolites in plasma and urine that showed statistically significant associations with symptom severity. The results of these analyses are visualized using volcano plots for plasma (Figure 3A) and urine (Figure 3B), highlighting the most significant metabolites across all scales. These metabolites demonstrate scale-specific and shared patterns of dysregulation. A detailed breakdown of metabolite associations for each scale is presented in Figure 4, with additional data provided in the Supplementary Materials. These findings provide insight into the biological heterogeneity of schizophrenia and suggest that specific metabolic signatures may be linked to different clinical dimensions of the disorder.

To investigate symptom domain-specific metabolic changes in schizophrenia, we analysed individual plasma metabolites in relation to distinct clinical scales. Figure 4 and S5 illustrates selected significant associations. Creatinine levels positively correlated with BNSS scores (Figure 4A), suggesting involvement of energy metabolism in negative symptomatology. Elevated creatine in subjects with more severe negative symptoms may reflect increased neuronal energy demand or disrupted mitochondrial function. Propylene glycol, a metabolite not typically linked with endogenous pathways, showed a moderate positive association with PANSS-negative scores (Figure 4B), potentially indicating environmental exposure or metabolic stress in individuals with more pronounced negative symptoms. Homovanillate, a dopamine metabolite, was positively associated with PANSS-positive scores (Figure 4C), consistent with dopaminergic dysregulation observed in patients exhibiting stronger positive symptoms. Importantly, uracil levels showed an inverse relationship with MADRS scores (Figure 4D), where higher depressive symptoms were associated with lower uracil concentrations. This suggests possible disruptions in pyrimidine metabolism or RNA turnover in individuals with comorbid depressive features, aligning with prior findings of nucleic acid metabolism changes in mood disorders.

To further explore the relationship between clinical symptom severity and metabolic alterations, in urine samples (Figure 5 and S6), we examined associations between selected metabolites and clinical scale scores. As shown in Figure 5A, Glutamate concentration increased with higher PANSS-positive scores, indicating a potential link between excitatory neurotransmission and positive symptom burden. Boxplots confirm elevated glutamate in subjects compared to controls. Similarly, N-Acetylglutamate showed a positive association with Global Functioning: Role scores (Figure 5B), suggesting potential involvement in cognitive or functional domains. Boxplots further support higher N-Acetylglutamate levels among subjects (middle panel, right). Fructose levels also exhibited a positive trend with increasing PANSS- positive and negative scores (Figures 5C and 5D), implicating altered carbohydrate metabolism in negative symptom severity. Notably, subject variability appeared greater than that of controls, consistent with clinical heterogeneity. These findings were consistent across sexes, although no strong sex-specific clustering was observed. These metabolite-clinical relationships support the potential of metabolic markers for stratifying symptom dimensions within schizophrenia, contributing to precision diagnostics and individualized treatment approaches. A more extensive and elaborate data set is provided in the section.

Discussion

Schizophrenia, like many chronic mental illnesses (CMI), presents a formidable challenge to modern psychiatry due to its complex symptomatology, heterogeneous presentation, and elusive pathophysiology (van Os & Kapur, 2009). Despite its high heritability and extensive genome-wide association study (GWAS) efforts, clinically useful genetic biomarkers remain limited, particularly in distinguishing schizophrenia from closely related disorders like bipolar disorder (BD) and major depressive disorder (MDD) (Schizophrenia Working Group of the Psychiatric Genomics, 2014). The work presented here aims to bridge this gap by exploring the metabolomic landscape of schizophrenia using high-resolution quantitative nuclear magnetic resonance (qNMR) spectroscopy. Through integrating metabolite quantification with clinical scale scoring, this study identifies potential biomarkers that may contribute to precision diagnostics and treatment strategies by extending the current medical diagnostic practises. The metabolomic profiling of plasma revealed significant alterations in metabolites associated with nitrogen metabolism, amino acid pathways, and neuroactive compounds. Downregulation of amino acids such as 2-Aminobutyrate, Phenylalanine, Arginine, Guanidinoacetate, and Glutamine points to a consistent dysregulation in nitrogen balance and urea cycle processes. These disturbances are of particular interest given their critical role in neurotransmitter synthesis and neuronal function. Altered levels of Glutamine, for example, can impact both excitatory and inhibitory neurotransmission by modulating glutamate and GABA pathways, both of which are implicated in schizophrenia pathology. Furthermore, multiple neuroactive metabolites within the serotonin and tryptophan pathways, including N-Acetylserotonin, 5-Hydroxyindole-3-acetate, and 3-Hydroxykynurenine, were differentially regulated. These molecules are central to mood regulation, and circadian rhythm, which are frequently disrupted in schizophrenia. Their dysregulation may reflect deeper impairments in monoaminergic transmission and help explain comorbid symptoms such as insomnia, affective flattening, or mood lability.

However, the presence of pharmaceuticals such as Acetaminophen and Ibuprofen in the plasma metabolome requires caution in interpretation of the findings in the absence of patient history as this reflects medication use among participants (see table S1). These exogenous compounds have the potential to influence endogenous metabolic processes (Figure 2), also adding to noise and thereby confounding interpretations of certain metabolic shift (Liu et al., 2020). These findings highlight the need for carefully interpreting false- positive and negative effects when interpreting metabolomics in clinical populations.

While plasma reflects real-time systemic metabolism, urine offers a complementary insight in to metabolic waste profiling. In our study, urine analysis revealed both upregulated and downregulated metabolites, contributing a more comprehensive view of systemic metabolic derangement in schizophrenia. Asparagine and O-Phosphoethanolamine were upregulated in patients, indicating potential shifts in amino acid metabolism and membrane phospholipid turnover, which may suggest membrane remodelling or neuroinflammatory processes, both of which have been implicated in schizophrenia pathophysiology. The observed downregulation of Glycine, Melatonin, Ethylene glycol, and Ribose is suggestive of broad disruptions across neurotransmission, sleep regulation, and energy metabolism (Kawai et al., 2015; Kunz, 2004). Reduced urinary melatonin supports clinical

observations of circadian rhythm disturbances frequently reported in schizophrenia, which are rarely accounted for in traditional diagnostic paradigms.

The big question however persists, can we effectively translate our findings into clinical applications? Is it possible to integrate blood and urine metabolic profiling with clinical phenotyping?

Indeed, linear modelling revealed scale-specific associations with plasma metabolites. For instance, Among the several correlates (Figure 4, 5 and supplementary section), let us take the example of Creatinine's positive correlation with MADRS scores (Figure 4A) indicates that depressive symptoms may be particularly linked with disruptions in energy metabolism, possibly due to mitochondrial dysfunction or heightened cellular stress. This observation resonates with previous research showing impaired bioenergetics in schizophrenia and suggests that peripheral creatinine could serve as a proxy for neuronal metabolic inefficiency. Propylene glycol is not an endogenous metabolite but is used as a solvent or additive in food and several commercial products, was positively associated with PANSS-negative scores (Figure 4B), which might reflect metabolic responses to environmental exposure or dysregulated detoxification systems (Maurer & Kessler, 1988) in more severely affected individuals. In such cases, understanding and exploring patients' history and societal settings can shed further information in to the lifestyle factors, medication, or underlying disease processes of the individual. Association of Homovanillate with PANSS-positive scores (Figure 4C) reinforces the long-standing dopaminergic hyperactivity (Davis et al., 1991) hypothesis in positive symptom domains. Uracil's inverse relationship with MADRS scores (Figure 4D) presents an interesting link; Uridine the breakdown product of uracil is under test for improving between mood dysregulation (Carlezon et al., 2005) indirectly indicating an impaired pyrimidine metabolic pathway.

In similar vein, urine metabolomics supported domain-specific patterns. Increased glutamate levels with PANSS-positive scores (Figure 5A) align with the glutamatergic hypothesis of schizophrenia (Kruse & Bustillo, 2022), where excitatory neurotransmission may become hyperactive or dysregulated. Interestingly, N-acetylglutamate's positive association with global functioning (Figure 5C) may reflect compensatory metabolic activity (Alonso et al., 1991) supporting cognitive resilience or recovery capacity. Meanwhile, fructose's correlation with PANSS-negative scores (Figure 5C and D) highlights altered carbohydrate metabolism, possibly related to insulin resistance and impact on brain function (Spagnuolo et al., 2020). These observations reinforce the concept that schizophrenia is not a single disease entity but a spectrum or syndrome composed of biologically distinguishable symptomatic clusters. By linking specific metabolite profiles to individual symptom dimensions, this study sets the stage for more nuanced patient stratification and opens the door to personalized interventions targeting specific metabolic disruptions.

Strengths, Limitations, and Future Directions

This study benefits from a comprehensive and integrative design, combining plasma and urine metabolomics with high-dimensional clinical data. The use of qNMR spectroscopy ensures reproducibility and quantitative accuracy, while the multi-scalar symptom assessments offer a rich landscape for exploring biological heterogeneity.

However, several limitations prevent a full-proof exploration of our study. First, the influence of medication cannot be fully disentangled (Table S1). Although observed increases in acetaminophen and ibuprofen levels are likely due to treatment regimens, these compounds may also influence endogenous metabolism and contribute to shifts in metabolic changes (Liu et al., 2020). Second, a significant limitation in our approach lies in the sample preparation methodology requires organic phase separation, which removes lipids and cholesterol contained within high-density and low-density lipoproteins (HDL and LDL). These compounds represent important risk factors in numerous neurological. Next, the deconvolution of 1D NMR spectra to identify and quantify metabolites can introduce errors due to overlapping CS between structurally similar molecules. Finally, while sample sizes are adequate for exploratory analyses, larger cohorts are necessary to validate findings and enhance generalizability. Inclusion of longitudinal follow-up could also help identify dynamic biomarkers that change with disease progression or treatment response. Despite these limitations, the current findings provide a template for future translational research.

Conclusion

In this study, we applied qNMR-based metabolomic profiling to plasma and urine samples to uncover metabolic alterations associated with schizophrenia. Our findings reveal disruptions in amino acid metabolism, nitrogen balance, neurotransmitter pathways, and carbohydrate metabolism in patients compared to healthy controls. Furthermore, integration of metabolomics with symptom-specific clinical scales such as PANSS, BNSS, MADRS, and GFS/GFR marks a significant advancement in this study, offering a novel means to stratify symptom domains biologically.

Altogether, our findings herein support the hypothesis that schizophrenia is a biologically heterogeneous disorder, and that metabolomic signatures can reflect this diversity. This approach provides a valuable framework for future biomarker development, therapeutic targeting, personalized medicine initiatives, and information dissemination targeting social stigmatization. While further validation in larger and longitudinal cohorts is necessary, the present study underscores the promise of metabolomics as a non-invasive, scalable, and informative tool in the management of complex mental illnesses such as schizophrenia.

Author Contributions

A.C. and D.S. designed research; A.C., and K.B. performed research; D.P., D.S., and S.B., contributed analytic tools; A.C., and G.K. analysed data; and A.C., wrote the paper; A.C., D.S., D.W., and R.H. secured financial support and ethics approval. All authors revised the manuscript

Acknowledgments and funding

A.C. acknowledges from the Hanse-Wissenschaftskolleg (HWK), Germany and Helmholtz Information & Data Science Academy (HIDA) to conduct the research activity. Access to the Jülich-Düsseldorf Biomolecular NMR

Center jointly run by Forschungszentrum Jülich and HHU and support by the DFG (HE 3243/4-1 and INST 208/771-1 FUGG) are acknowledged.

Conflict of Interest: A.C., is drafting an EP application, which will be submitted to the German Patent office in Munich. All the other authors declare no competing interests.

References

- Alonso, E., Garcia-Perez, M. A., Bueso, J., & Rubio, V. (1991). N-acetyl-L-glutamate in brain: assay, levels, and regional and subcellular distribution. *Neurochem Res*, 16(7), 787-794. <https://doi.org/10.1007/BF00965688>
- Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*, Arlington VA, American Psychiatric Association.
- Bertram, L., Lill, C. M., & Tanzi, R. E. (2010). The genetics of Alzheimer disease: back to the future. *Neuron*, 68(2), 270-281. <https://doi.org/10.1016/j.neuron.2010.10.013>
- Blackwood, D. H. R., Fordyce, A., Walker, M. T., St Clair, D. M., Porteous, D. J., & Muir, W. J. (2001). Schizophrenia and affective disorders - Cosegregation with a translocation at chromosome 1q42 that directly disrupts brain-expressed genes: Clinical and P300 findings in a family. *American Journal of Human Genetics*, 69(2), 428-433. <https://doi.org/10.1086/321969>
- Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol*, 37(8), 911-917. <https://doi.org/10.1139/o59-099>
- Campion, J., Bhui, K., Bhugra, D., & European Psychiatric, A. (2012). European Psychiatric Association (EPA) guidance on prevention of mental disorders. *Eur Psychiatry*, 27(2), 68-80. <https://doi.org/10.1016/j.eurpsy.2011.10.004>
- Canada, M. H. C. o. (2012). *Making the Case for Investing in Mental Health in Canada*. https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2016-06/Investing_in_Mental_Health_FINAL_Version_ENG.pdf
- Carlezon, W. A., Jr., Mague, S. D., Parow, A. M., Stoll, A. L., Cohen, B. M., & Renshaw, P. F. (2005). Antidepressant-like effects of uridine and omega-3 fatty acids are potentiated by combined treatment in rats. *Biol Psychiatry*, 57(4), 343-350. <https://doi.org/10.1016/j.biopsych.2004.11.038>
- Carrion, R. E., Auther, A. M., McLaughlin, D., Olsen, R., Addington, J., Bearden, C. E., Cadenhead, K. S., Cannon, T. D., Mathalon, D. H., McGlashan, T. H., Perkins, D. O., Seidman, L. J., Tsuang, M. T., Walker, E. F., Woods, S. W., & Cornblatt, B. A. (2019). The Global Functioning: Social and Role Scales-Further Validation in a Large Sample of Adolescents and Young Adults at Clinical High Risk for Psychosis. *Schizophr Bull*, 45(4), 763-772. <https://doi.org/10.1093/schbul/sby126>

- Chan, M. K., Cooper, J. D., & Bahn, S. (2015). Commercialisation of Biomarker Tests for Mental Illnesses: Advances and Obstacles. *Trends Biotechnol*, 33(12), 712-723. <https://doi.org/10.1016/j.tibtech.2015.09.010>
- Chen, H. H., Petty, L. E., Bush, W., Naj, A. C., & Below, J. E. (2019). GWAS and Beyond: Using Omics Approaches to Interpret SNP Associations. *Curr Genet Med Rep*, 7(1), 30-40. <https://doi.org/10.1007/s40142-019-0159-z>
- Davis, K. L., Kahn, R. S., Ko, G., & Davidson, M. (1991). Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry*, 148(11), 1474-1486. <https://doi.org/10.1176/ajp.148.11.1474>
- De Deyn, P. P., Vanholder, R., Eloot, S., & Glorieux, G. (2009). PROGRESS IN UREMIC TOXIN RESEARCH: Guanidino Compounds as Uremic (Neuro)Toxins. *Seminars in Dialysis*, 22(4), 340-345. <https://doi.org/https://doi.org/10.1111/j.1525-139X.2009.00577.x>
- Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H. (2006). Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem*, 78(13), 4281-4290. <https://doi.org/10.1021/ac051632c>
- International Schizophrenia, C., Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., & Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460(7256), 748-752. <https://doi.org/10.1038/nature08185>
- Julkunen, H., Cichonska, A., Tiainen, M., Koskela, H., Nybo, K., Makela, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C., & Wurtz, P. (2023). Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*, 14(1), 604. <https://doi.org/10.1038/s41467-023-36231-7>
- Kawai, N., Sakai, N., Okuro, M., Karakawa, S., Tsuneyoshi, Y., Kawasaki, N., Takeda, T., Bannai, M., & Nishino, S. (2015). The sleep-promoting and hypothermic effects of glycine are mediated by NMDA receptors in the suprachiasmatic nucleus. *Neuropsychopharmacology*, 40(6), 1405-1416. <https://doi.org/10.1038/npp.2014.326>
- Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*, 13(2), 261-276. <https://doi.org/10.1093/schbul/13.2.261>
- Kirkpatrick, B., Saoud, J. B., Strauss, G. P., Ahmed, A. O., Tatsumi, K., Opler, M., Luthringer, R., & Davidson, M. (2018). The brief negative symptom scale (BNSS): Sensitivity to treatment effects. *Schizophrenia Research*, 197, 269-273. <https://doi.org/10.1016/j.schres.2017.11.031>
- Korth, C., & Fangerau, H. (2020). Blood tests to diagnose schizophrenia: self-imposed limits in psychiatry. *Lancet Psychiatry*, 7(10), 911-914. [https://doi.org/10.1016/S2215-0366\(20\)30058-4](https://doi.org/10.1016/S2215-0366(20)30058-4)
- Kruse, A. O., & Bustillo, J. R. (2022). Glutamatergic dysfunction in Schizophrenia. *Transl Psychiatry*, 12(1), 500. <https://doi.org/10.1038/s41398-022-02253-w>

- Kunz, D. (2004). Chronobiotic protocol and circadian sleep propensity index: new tools for clinical routine and research on melatonin and sleep. *Pharmacopsychiatry*, 37(4), 139-146. <https://doi.org/10.1055/s-2004-827167>
- Liu, J., Lahousse, L., Nivard, M. G., Bot, M., Chen, L., van Klinken, J. B., Thesing, C. S., Beekman, M., van den Akker, E. B., Sliker, R. C., Waterham, E., van der Kallen, C. J. H., de Boer, I., Li-Gao, R., Vojinovic, D., Amin, N., Radjabzadeh, D., Kraaij, R., Alferink, L. J. M., . . . van Duijn, C. M. (2020). Integration of epidemiologic, pharmacologic, genetic and gut microbiome data in a drug-metabolite atlas. *Nat Med*, 26(1), 110-117. <https://doi.org/10.1038/s41591-019-0722-x>
- Martins-de-Souza, D., Guest, P. C., Vanattou-Saifoudine, N., Harris, L. W., & Bahn, S. (2011). Chapter 4 - Proteomic Technologies for Biomarker Studies in Psychiatry: Advances and needs. In P. C. Guest & S. Bahn (Eds.), *International Review of Neurobiology* (Vol. 101, pp. 65-94). Academic Press. <https://doi.org/10.1016/B978-0-12-387718-5.00004-3>
- Maurer, H., & Kessler, C. (1988). Identification and quantification of ethylene glycol and diethylene glycol in plasma using gas chromatography-mass spectrometry. *Arch Toxicol*, 62(1), 66-69. <https://doi.org/10.1007/BF00316260>
- Montgomery, S. A., & Åsberg, M. (1979). A New Depression Scale Designed to be Sensitive to Change. *British Journal of Psychiatry*, 134(4), 382-389. <https://doi.org/10.1192/bjp.134.4.382>
- O'Donovan, M. C., & Owen, M. J. (2016). The implications of the shared genetics of psychiatric disorders. *Nat Med*, 22(11), 1214-1219. <https://doi.org/10.1038/nm.4196>
- OECD/EU. (2018). *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. OECD Publishing, Paris. https://doi.org/10.1787/health_glance_eur-2018-en
- Owen, M. J., Craddock, N., & Jablensky, A. (2007). The genetic deconstruction of psychosis. *Schizophr Bull*, 33(4), 905-911. <https://doi.org/10.1093/schbul/sbm053>
- Paul, S. M., & Potter, W. Z. (2024). Finding new and better treatments for psychiatric disorders. *Neuropsychopharmacology*, 49(1), 3-9. <https://doi.org/10.1038/s41386-023-01690-5>
- Piskulic, D., Addington, J., Auther, A., & Cornblatt, B. A. (2011). Using the global functioning social and role scales in a first-episode sample. *Early Interv Psychiatry*, 5(3), 219-223. <https://doi.org/10.1111/j.1751-7893.2011.00263.x>
- Porteous, D. J., Thomson, P. A., Millar, J. K., Evans, K. L., Hennah, W., Soares, D. C., McCarthy, S., McCombie, W. R., Clapcote, S. J., Korth, C., Brandon, N. J., Sawa, A., Kamiya, A., Roder, J. C., Lawrie, S. M., McIntosh, A. M., St Clair, D., & Blackwood, D. H. (2014). DISC1 as a genetic risk factor for schizophrenia and related major mental illness: response to Sullivan. *Mol Psychiatry*, 19(2), 141-143. <https://doi.org/10.1038/mp.2013.160>
- Pu, J., Liu, Y., Wu, H., Liu, C., Chen, Y., Tang, W., Yu, Y., Gui, S., Zhong, X., Wang, D., Chen, X., Chen, Y., Chen, X., Qiao, R., Jiang, Y., Zhang, H., Ren, Y., Fan, L., Wang, H., & Xie, P. (2025). Characterizing

- metabolomic and proteomic changes in depression: a systematic analysis. *Mol Psychiatry*, 30(7), 3120-3128. <https://doi.org/10.1038/s41380-025-02919-z>
- Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kahler, A. K., Akterin, S., Bergen, S. E., Collins, A. L., Crowley, J. J., Fromer, M., Kim, Y., Lee, S. H., Magnusson, P. K., Sanchez, N., Stahl, E. A., Williams, S., Wray, N. R., Xia, K., Bettella, F., . . . Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*, 45(10), 1150-1159. <https://doi.org/10.1038/ng.2742>
- Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), 421-427. <https://doi.org/10.1038/nature13595>
- Schwarz, E., Izmailov, R., Spain, M., Barnes, A., Mapes, J. P., Guest, P. C., Rahmoune, H., Pietsch, S., Leweke, F. M., Rothermundt, M., Steiner, J., Koethe, D., Kranaster, L., Ohrmann, P., Suslow, T., Levin, Y., Bogerts, B., van Beveren, N. J., McAllister, G., . . . Bahn, S. (2010). Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. *Biomark Insights*, 5, 39-47. <https://doi.org/10.4137/bmi.s4877>
- Sialana, F. J., Wang, A. L., Fazari, B., Kristofova, M., Smidak, R., Trossbach, S. V., Korth, C., Huston, J. P., de Souza Silva, M. A., & Lubec, G. (2018). Quantitative Proteomics of Synaptosomal Fractions in a Rat Overexpressing Human DISC1 Gene Indicates Profound Synaptic Dysregulation in the Dorsal Striatum. *Front Mol Neurosci*, 11, 26. <https://doi.org/10.3389/fnmol.2018.00026>
- Smith, D. J., & Craddock, N. (2011). Unipolar and bipolar depression: different of the same? *Br J Psychiatry*, 199(4), 272-274. <https://doi.org/10.1192/bjp.bp.111.092726>
- Spagnuolo, M. S., Iossa, S., & Cigliano, L. (2020). Sweet but Bitter: Focus on Fructose Impact on Brain Function in Rodent Models. *Nutrients*, 13(1). <https://doi.org/10.3390/nu13010001>
- Sullivan, P. F. (2013). Questions about DISC1 as a genetic risk factor for schizophrenia. *Mol Psychiatry*, 18(10), 1050-1052. <https://doi.org/10.1038/mp.2012.182>
- van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet*, 374(9690), 635-645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8)
- Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P., & Luchinat, C. (2019). High-Throughput Metabolomics by 1D NMR. *Angew Chem Int Ed Engl*, 58(4), 968-994. <https://doi.org/10.1002/anie.201804736>
- Wiegand, H. F., Saam, J., Marschall, U., Chmitorz, A., Kriston, L., Berger, M., Lieb, K., & Holz, L. P. (2020). Challenges in the Transition from In-Patient to Out-Patient Treatment in Depression. *Dtsch Arztebl Int*, 117(27-28), 472-479. <https://doi.org/10.3238/arztebl.2020.0472>
- Wishart, D. S. (2019). Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*, 99(4), 1819-1875. <https://doi.org/10.1152/physrev.00035.2018>

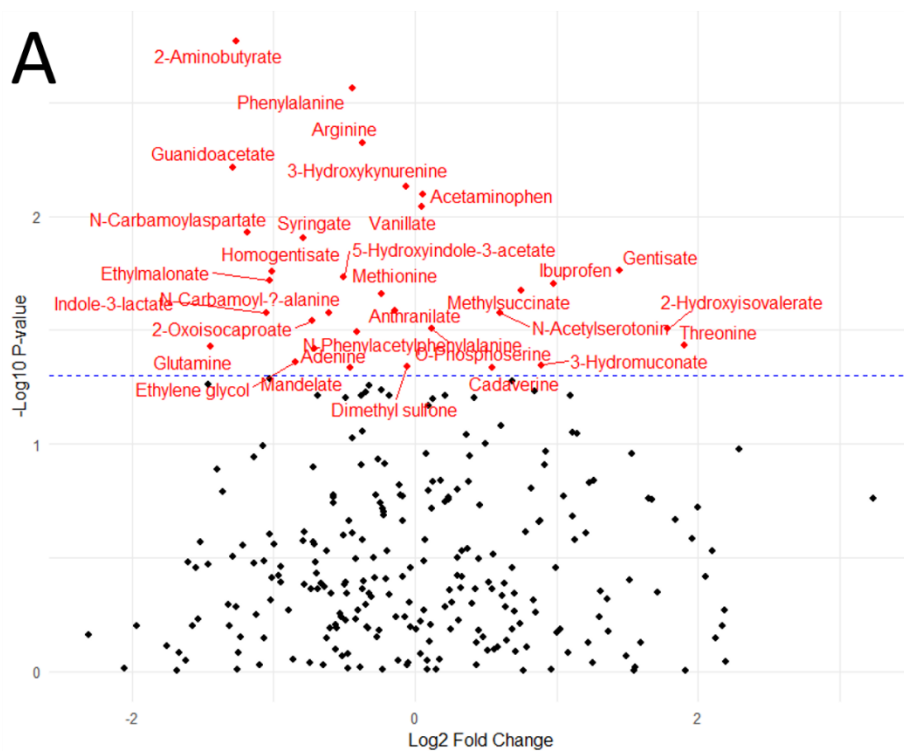
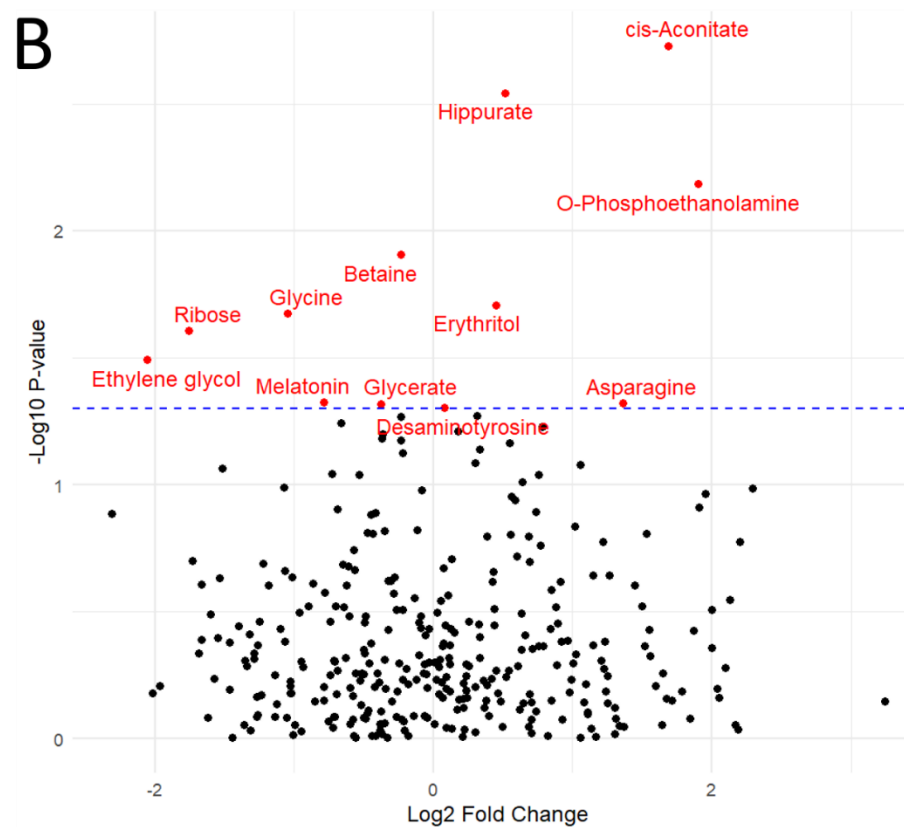


Figure 1. Volcano plots (also see Figures S1 and S3 for heat maps) with the results of the linear model comparing the significant metabolites of the various test groups for (A) plasma-, and (B) urine metabolites



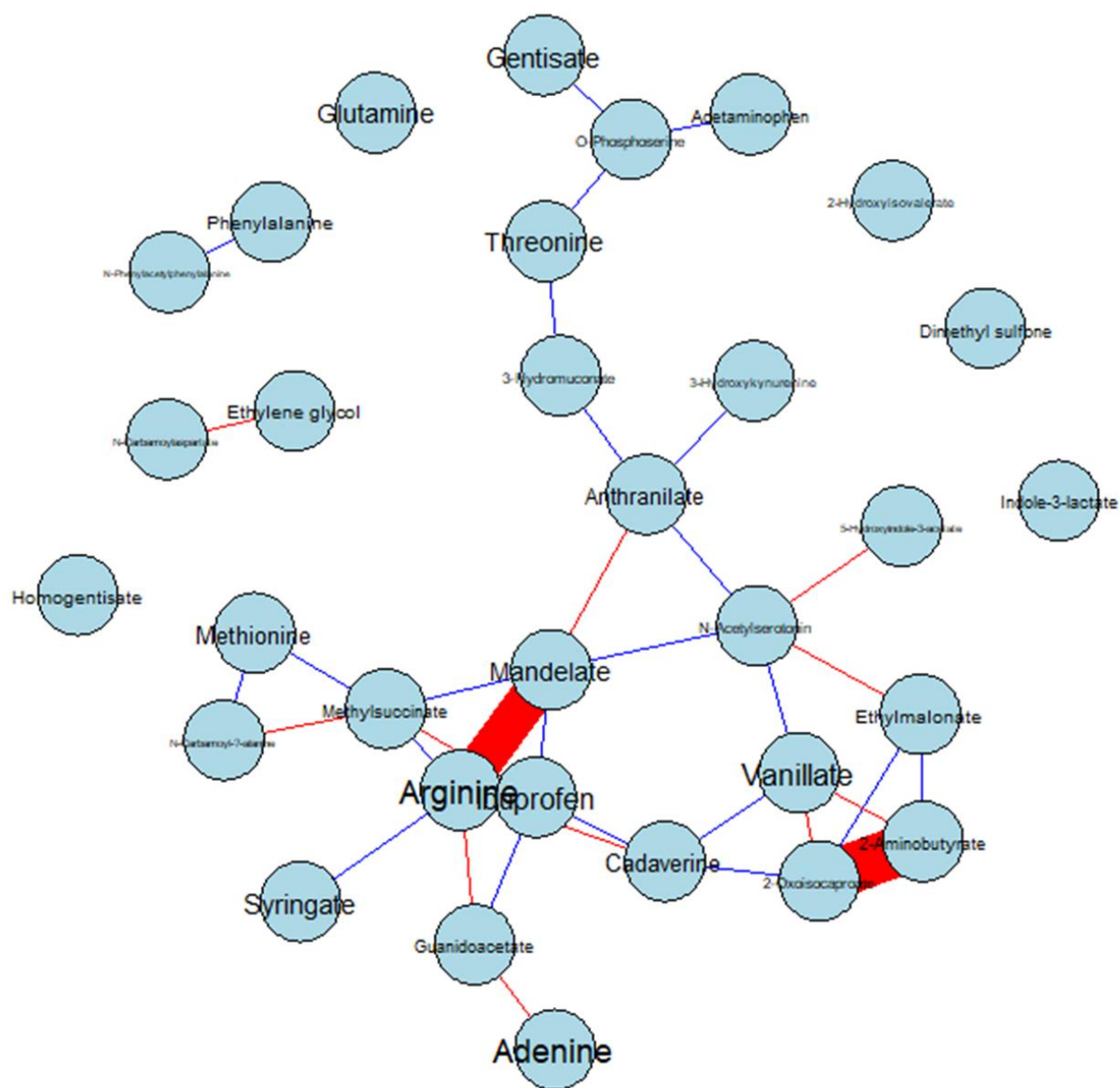


Figure 2. Biochemical interaction network of significant plasma metabolites in healthy controls and individuals with schizophrenia. The network was generated in R using the igraph package, based on Pearson partial correlation coefficients (see also Figure S2). Nodes represent individual metabolites, while edges reflect significant correlations between them—red lines indicate positive correlations, and blue lines indicate negative correlations. This visual map illustrates clusters of co-regulated metabolites and highlights potential metabolic pathway disruptions associated with schizophrenia, offering insight into disease-specific biochemical alterations.

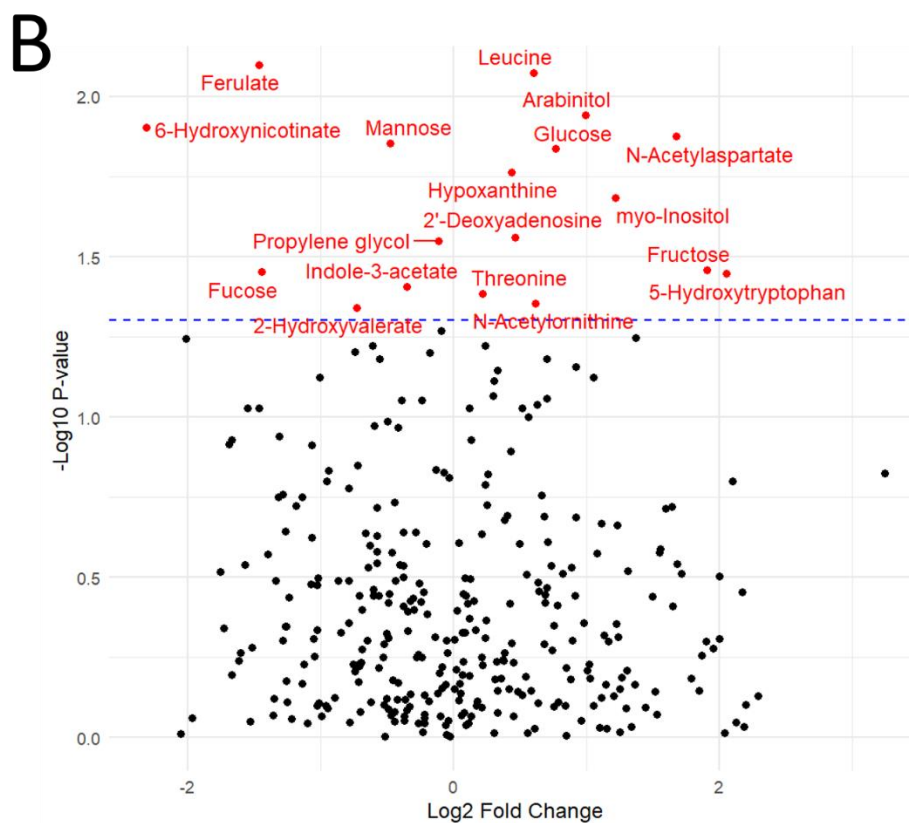
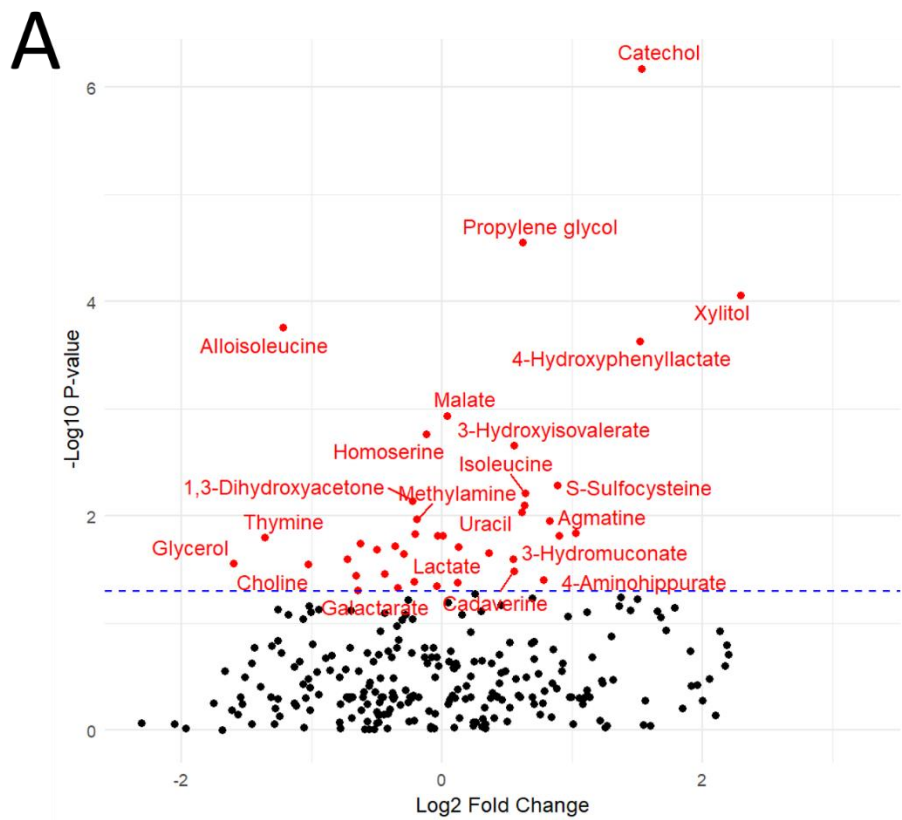


Figure 3: Volcano plots (also see Figures 4, 5 and supplementary section) with the results of the linear model comparing the significant metabolites associated with clinical diagnostic scales for (A) plasma-, and (B) urine metabolites

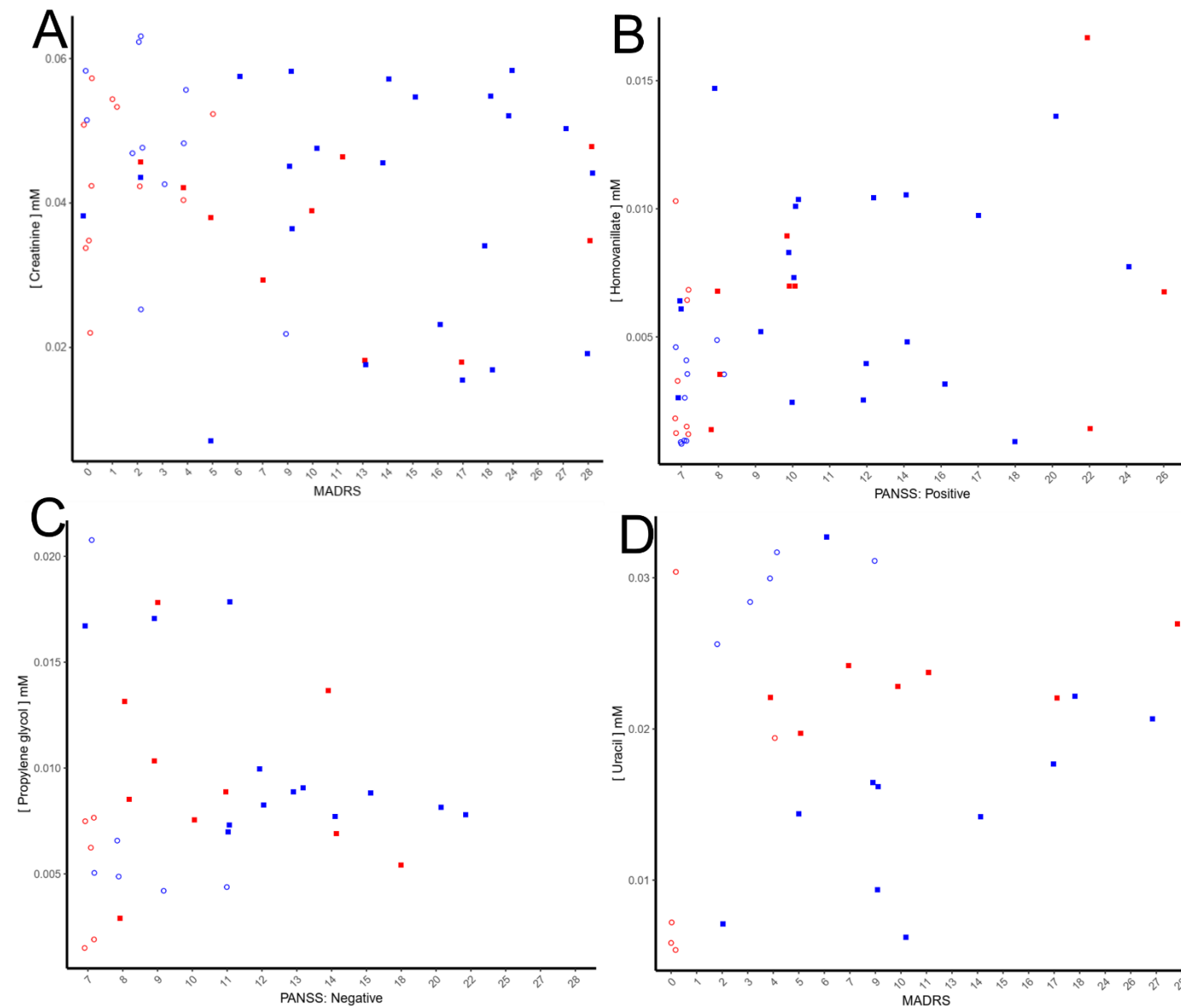


Figure 4. Associations between plasma metabolite concentrations and clinical severity scores. Scatter plots show correlations between: (A) creatinine and MADRS, (B) homovanillate and PANSS-positive, (C) propylene glycol and PANSS-negative, (D) Uracil and MADRS. Data points depicted in red and blue present females, and males respectively: circles represent controls while squares represent individuals with schizophrenia.

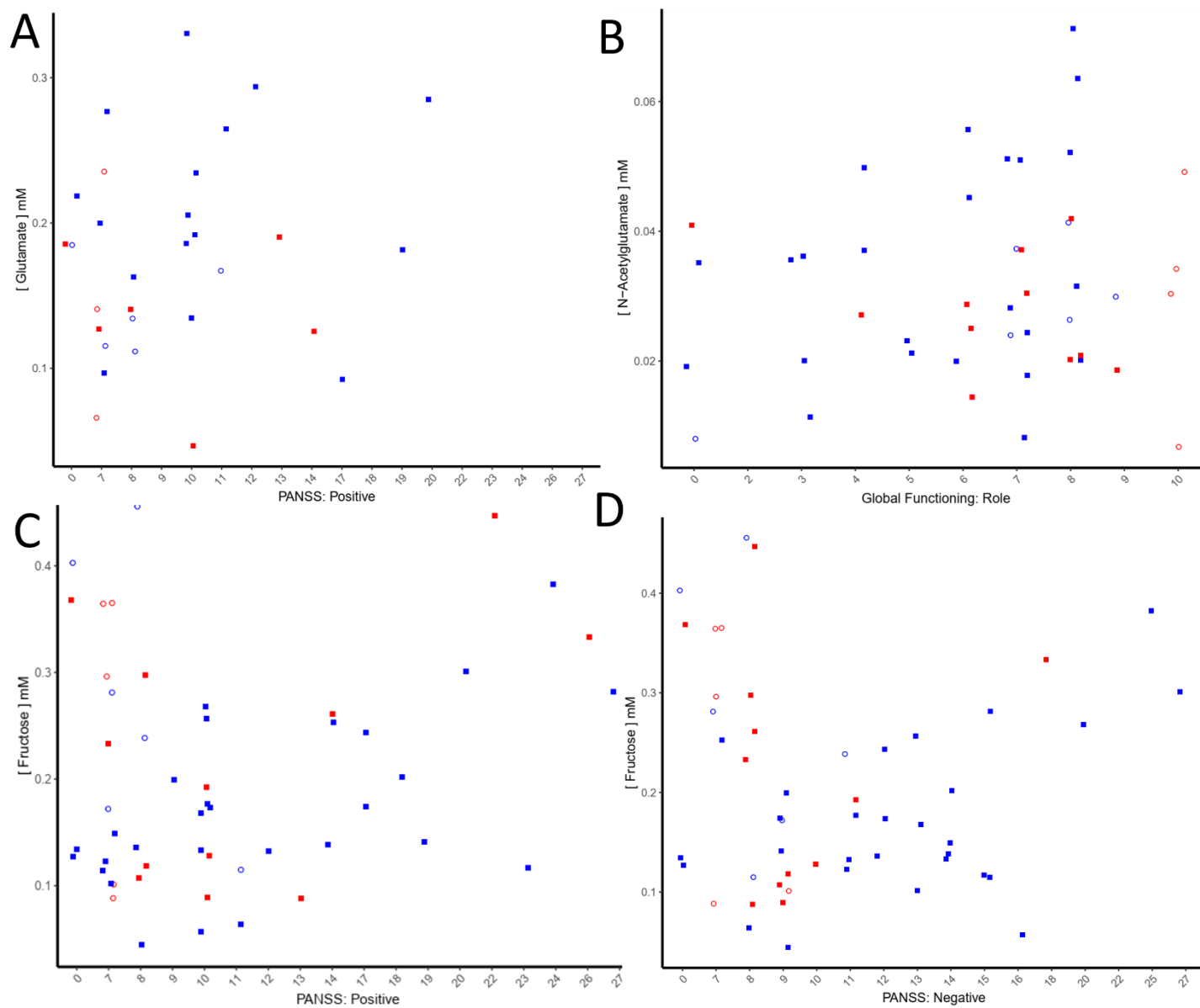


Figure 5: Associations between plasma metabolite concentrations and clinical severity scores. Scatter plots show correlations between: (A) Glutamate and PANSS. Positive, (B) N-Acetylglutamate and GFR, (C) Fructose and PANSS-positive, (D) and PANSS-negative. Data points depicted in red and blue present females, and males respectively: circles represent controls while squares represent individuals with schizophrenia.

Reconstructing schizophrenia's biochemistry through hybrid diagnostics: clinical and metabolic profiling

Abhishek Cukkemane^{1,2,3*}, Danilo Postin^{4,5}, Kevin Bochinsky^{1,6}, Simon Barton^{4,5}, Dirk Scheele^{4,5,7}, Gabi Kastenmüller⁸, Dieter Willbold^{1,2}, and Rene Hurlmann^{4,5}

¹Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, Jülich, Germany.

²Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

³Hanse-Wissenschaftskolleg Institute for Advanced Study, Delmenhorst, Germany

⁴Department of Psychiatry, School of Medicine and Health Sciences, University of Oldenburg, Oldenburg, Germany.

⁵Karl-Jaspers-Klinik, Bad Zwischenahn, Germany

⁶ current address -Bruker Biospin, Ettlingen, Germany

⁷ current address - Medizinische Fakultät, Ruhr-Universität Bochum, Bochum, Germany

⁸ Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, Neuherberg, Germany.

*Address for correspondence - a.cukkemane@fz-juelich.de

Table S1: The different types and classes of medication used to support individuals suffering from Schizophrenia

drug	class	utility	treatment
Amisulpride	antiemetic, atypical antipsychotic, D2ergic	comparable efficacy to olanzapine for schizophrenia	primary
Lisinopril/Amlodipine	Ca ²⁺ channel blocker	hypertension (high blood pressure)	adjunct/com bi
Aripiprazole	atypical antipsychotic (1st line treatment)	schizophrenia, obsessive compulsive disorder (OCD), and bipolar disorder	primary in most cases
ASS	NSAID	headache, pain, nausea	adjunct/com bi
Atorvastatin	cardiovascular disease	anti-cholesterol- inhibiting HMG-CoA reductase	adjunct/com bi
Benperidol	typical antipsychotic, D2ergic	schizophrenia and hypersexuality	primary
Biperiden	anticholinergic	Parkinson disease, certain drug-induced movement disorders and depression	adjunct/com bi
Bisoprolol	beta blocker	Tachyarrhythmias, high blood pressure, and heart failure	
Carbamazepine	anticonvulsant	epilepsy, neuropathic pain and schizophrenia	adjunct/com bi
Cariprazine	atypical antipsychotic	schizophrenia, bipolar disorder, bipolar depression, MDD	primary

Chlorprothixene	typical antipsychotic, Dopaminergic	psychotic disorders (e.g., schizophrenia) and of acute mania occurring as part of bipolar disorders	
Clozapine	first atypical antipsychotic	schizophrenia, parkinsons(adjunct)	primary
Dapagliflozine	antidiabetic (type II)	sodium-glucose co-transporter-2 (SGLT-2) inhibitor (removes sugar through urine)	
Doxazosin	quinazoline(hypertension), recently anti-aging	high blood pressure	adjunct/com bi
Empagliflozin	antidiabetic (type II)	sodium-glucose co-transporter-2 (SGLT-2) inhibitor (removes sugar through urine)	adjunct/com bi
Sacubitril/valsartan	angiotensin receptor blocker	ACE inhibitor or angiotensin II receptor antagonist - heart failure	adjunct/com bi
Escitalopram	SSRI anti-depressant	major depressive disorder and generalized anxiety disorder	primary
Esomeprazole	H+ pump inhibitor	reduce stomach acid	adjunct/com bi
Febuxostat	NSAID	gout due to high uric acid levels	adjunct/com bi
Pirenzepine/Gastrozepin	M1 selective antagonist	peptic ulcers and muscle anti-spasms	
Acetazolamide	1st gen. carbonic anhydrase inhibitor	glaucoma, epilepsy, acute mountain sickness, periodic paralysis, idiopathic intracranial hypertension	
Haloperidol	typical antipsychotic ((1st line treatment)	schizophrenia, tics in Tourette syndrome, BD, delirium, agitation, acute psychosis, and hallucinations from alcohol withdrawal	
Hydrochlorothiazide	diuretic	hypertension (high blood pressure)	
L-Thyroxine	Thyroid functioning hormone for I2	thyroid hormone deficiency (hypothyroidism)	
Lamotrigine	anti acute depression	epilepsy and stabilize mood in bipolar disorder	
Lithium ret.	mood stabilizer	BD; MDD and autism	
Lorazepam	benzodiazepine	anxiety disorders), trouble sleeping, severe agitation, active seizures, alcohol withdrawal	
Melperon	atypical antipsychotic	atypical although structurally related to haloperidol - schizophrenia and psychosis (parkinsonism)	
Metformin	1st line antidiabetic (type II)	overweight - lessen the risk of metabolic syndrome in	adjunct/com bi

		people who take antipsychotics	
Metoprolol	beta blocker	Tachyarrhythmias, high blood pressure, and heart failure	adjunct/combi
Milnacipran	SNRI anti-depressant	antidepressant - MDD	
Mirtazapine	atypical tetracyclic antidepressant	depression complicated by anxiety or insomnia	
Bethanechol	agonist muscarinic receptors	stimulates the parasympathetic nervous system like glaucoma	
Olanzapine	atypical antipsychotic (1st line treatment)	schizophrenia and bipolar disorder	primary
Pantoprazole/Pantozol	H ⁺ pump inhibitor	reduce stomach acid	
Perphenazine	typical antipsychotic drug	agitated depression	adjunct/combi
Pipamperone	typical antipsychotic drug (butyrophenone)	antipsychotic, normalise mood, sleep patterns and anti-depressant	
Pregabalin	anticonvulsant	epilepsy, neuropathic pain, fibromyalgia, restless leg syndrome, opioid withdrawal, and generalized anxiety disorder	
Promethazin + Venlafaxin	1st gen antihistamine/antipsychotic + SNRI	antipsychotic, sedative, and antiemetic + anti-depressant	adjunct/combi
Quetiapine	5-HT _{2A} antagonists (1st line treatment)	schizophrenia, bipolar disorder, and major depressive disorder, parkinsons(adjunct)	primary
Ramipril	ACE inhibitor	high blood pressure, heart failure, and diabetic kidney disease	primary
Risperidone	5-HT _{2A/2C} antagonists (1st line treatment)	schizophrenia and bipolar disorder	primary
Rosuvastatin	cardiovascular disease	Dyslipidemia	adjunct/combi
Sertraline	SSRI anti-depressant	panic disorder, social anxiety disorder, generalized anxiety disorder (GAD), OCD	
Tamsulosin	alpha blocker	symptomatic benign prostatic hyperplasia (BPH) and chronic prostatitis	
Torsemide	diuretic	fluid overload due to heart failure, kidney disease, and liver disease	
Desfesoterodine succinate	antimuscarinic agent diuretic	urinary frequency or incontinence	
Tranlycypromine	MAOI (2nd line treatment)	major depressive disorder, including atypical depression	
Valproate		epilepsy and bipolar disorder and prevent migraine headaches	

Venlafaxin	SNRI anti-depressant	anti-depressant	adjunct/ combi
Vitamin B9 (Folic acid)	vitamin	anaemia	
Vitamin B12 (cyanocobalamine)	vitamin	fatty acid metabolism and amino acid metabolism.	
Vitamin D3 (Choleocalciferol)	vitamin-hormonal	interferon type I activator, familial hypophosphatemia, hypoparathyroidism, and Fanconi syndrome	
Ziprasidone	atypical antipsychotic	schizophrenia and bipolar disorder	
Zopiclone	nonbenzodiazepine	insomnia-zopiclone increases the normal transmission of GABA	
Zuclopentixole	D1 and D2 antagonist, α 1-adrenergic and 5-HT ₂ antagonist	schizophrenia and other psychoses	

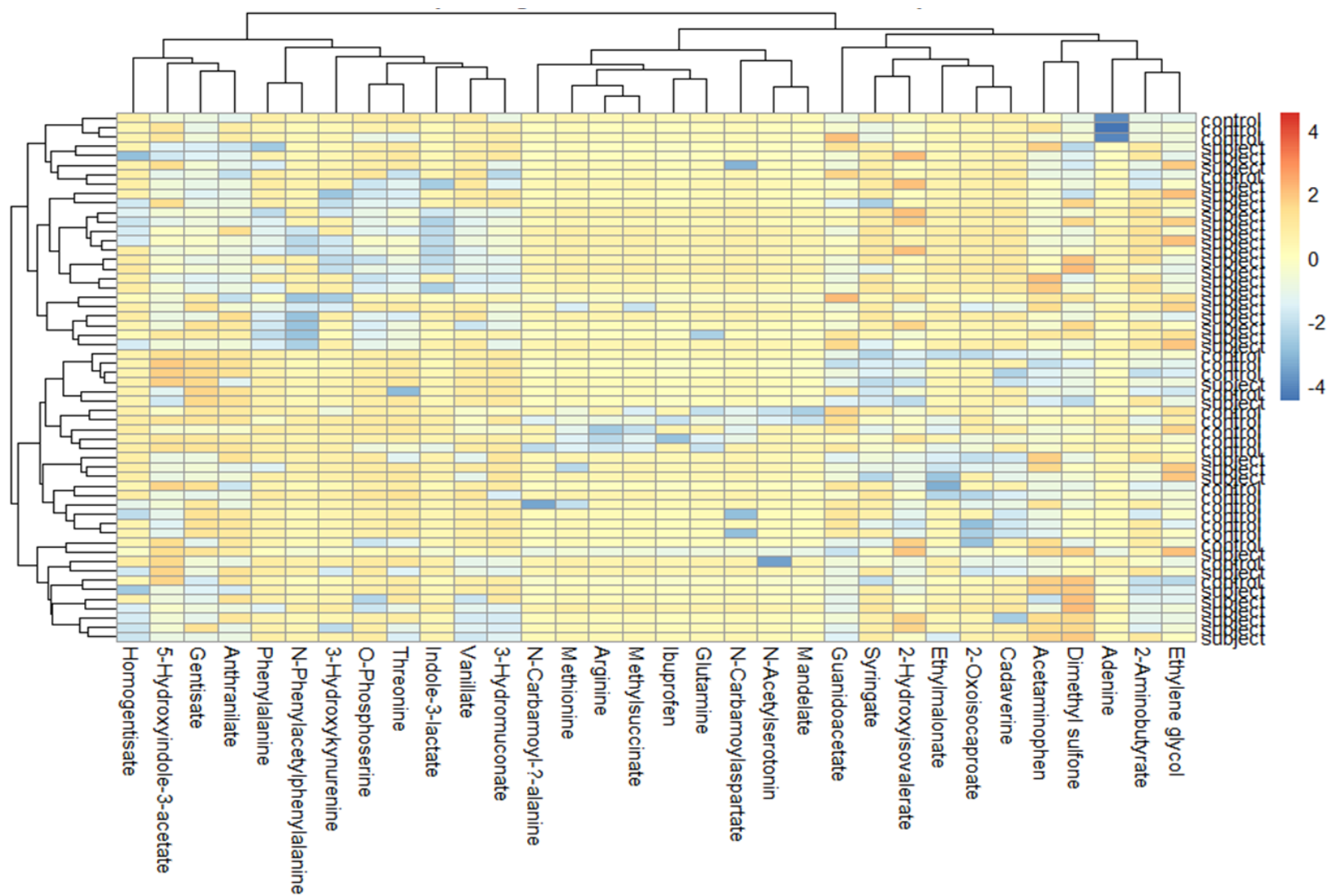


Figure S1. Heat map with hierarchical clustering of significant plasma metabolites identified through linear regression modeling comparing healthy controls to individuals with schizophrenia. Metabolite intensities are color-coded, with red indicating higher levels and blue indicating lower levels.

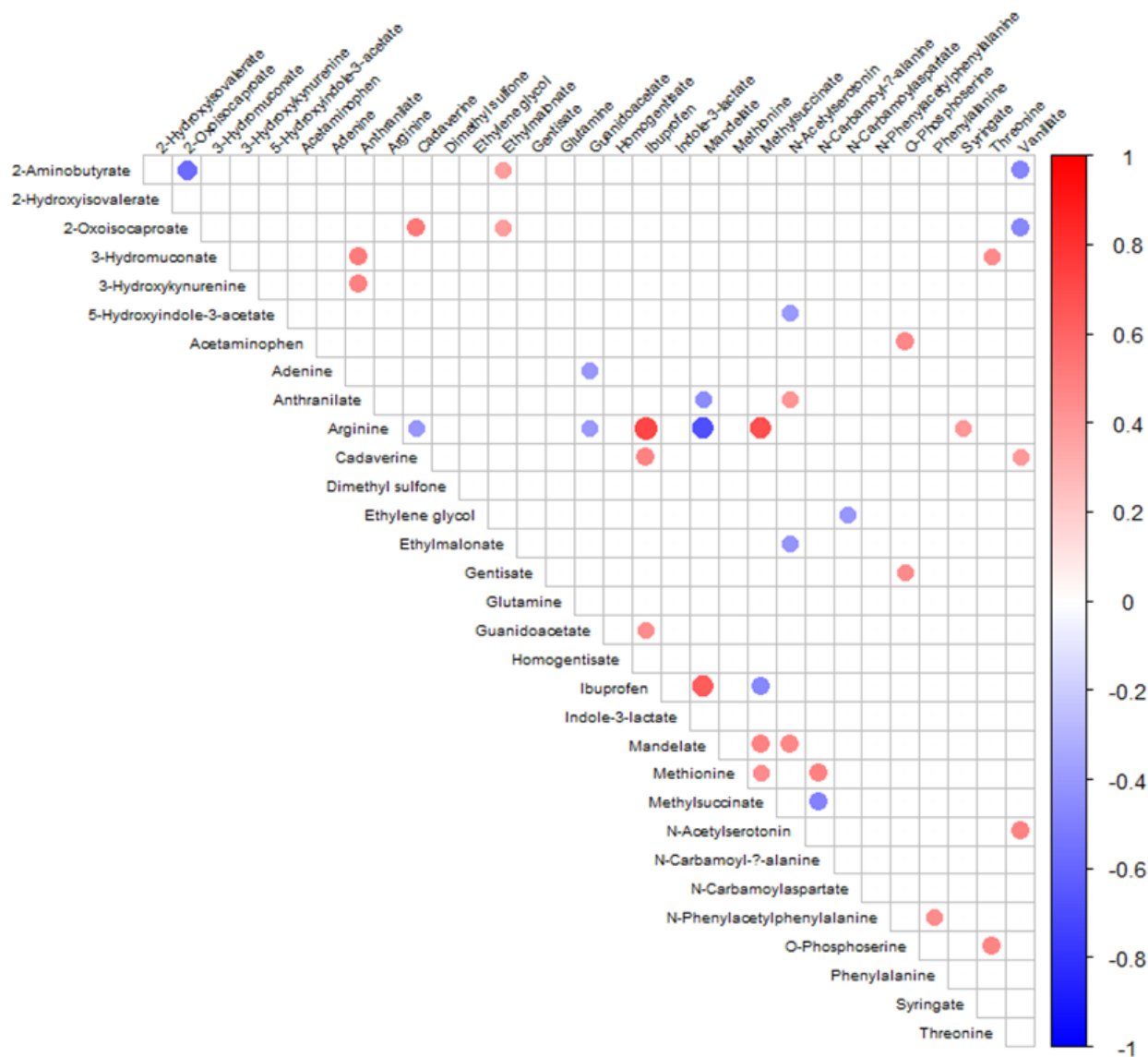


Figure S2: A correlation matrix for the significant plasma metabolites between the controls and individuals with schizophrenia. The matrix was generated using partial correlation coefficients via Pearson's method and plotted as a color-coded matrix, where red indicated positive correlations and blue indicated negative correlations.

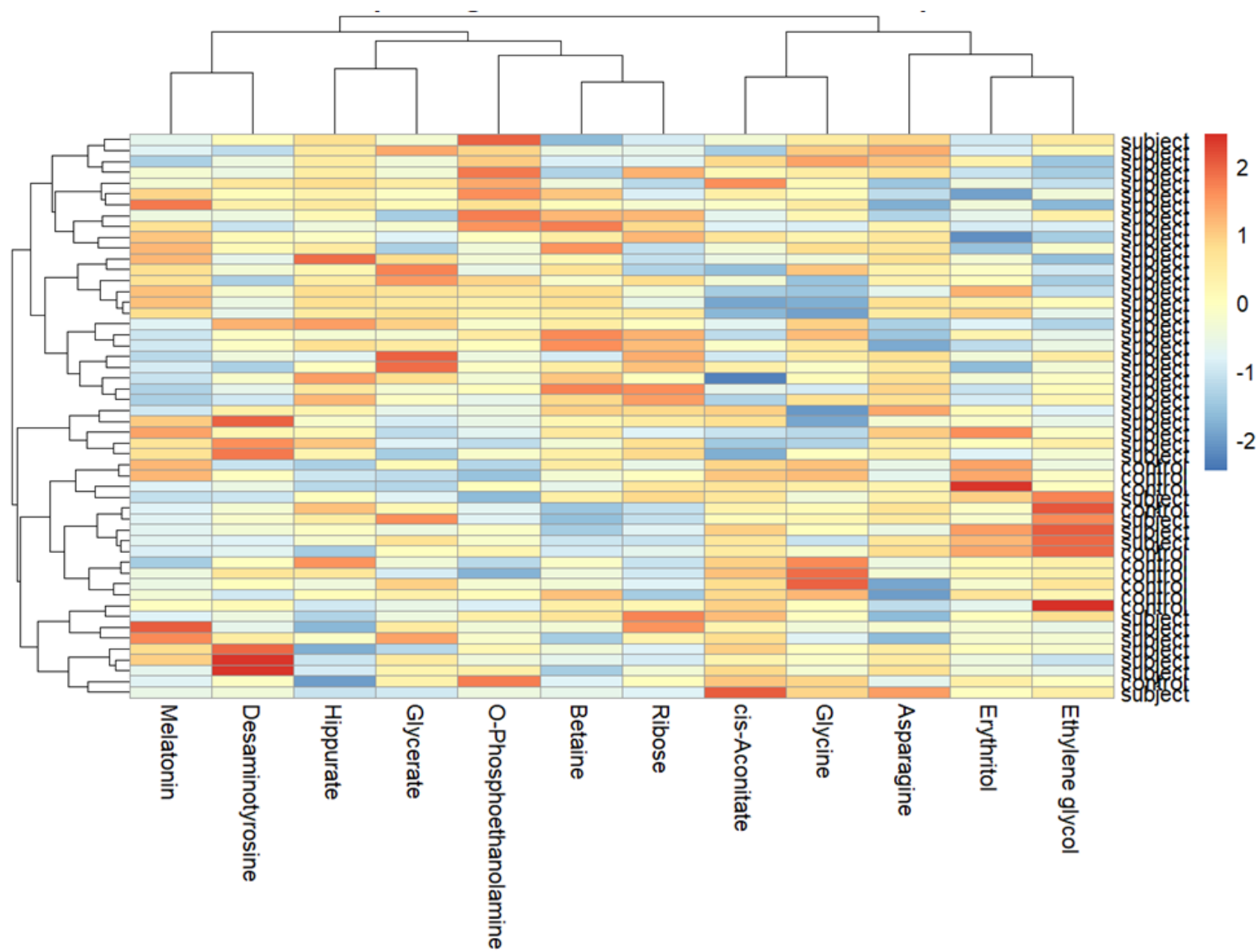


Figure S3. Heat map with hierarchical clustering of significant urine metabolites between healthy controls to individuals with schizophrenia.

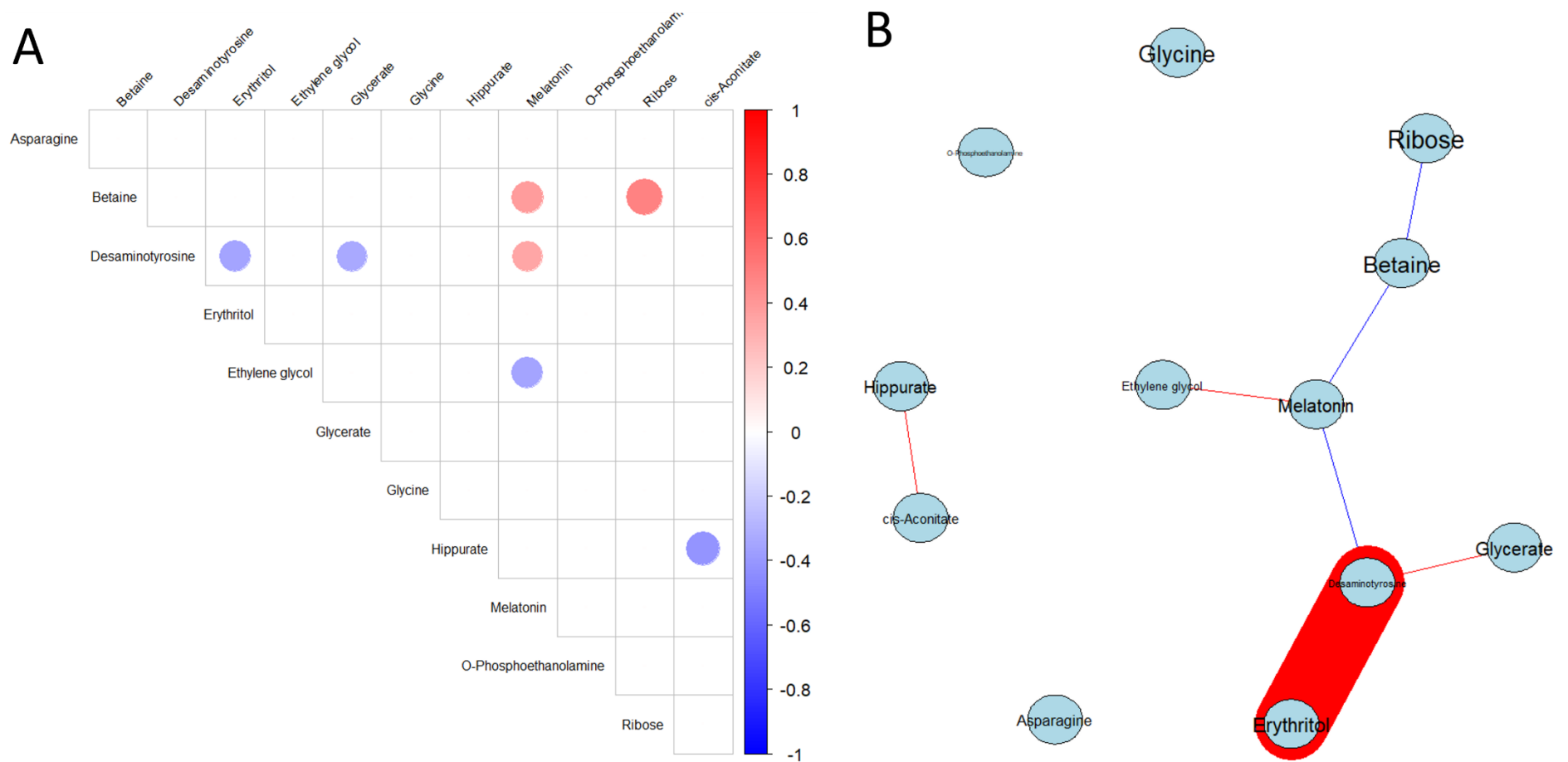


Figure S4: (A) Correlation matrix and (B) biochemical interaction network pathway for the significant urine metabolites between the controls and individuals with schizophrenia. The plots were generated using partial correlation coefficients via Pearson's method and plotted as a color-coded matrix, where red indicated positive correlations and blue indicated negative correlations.

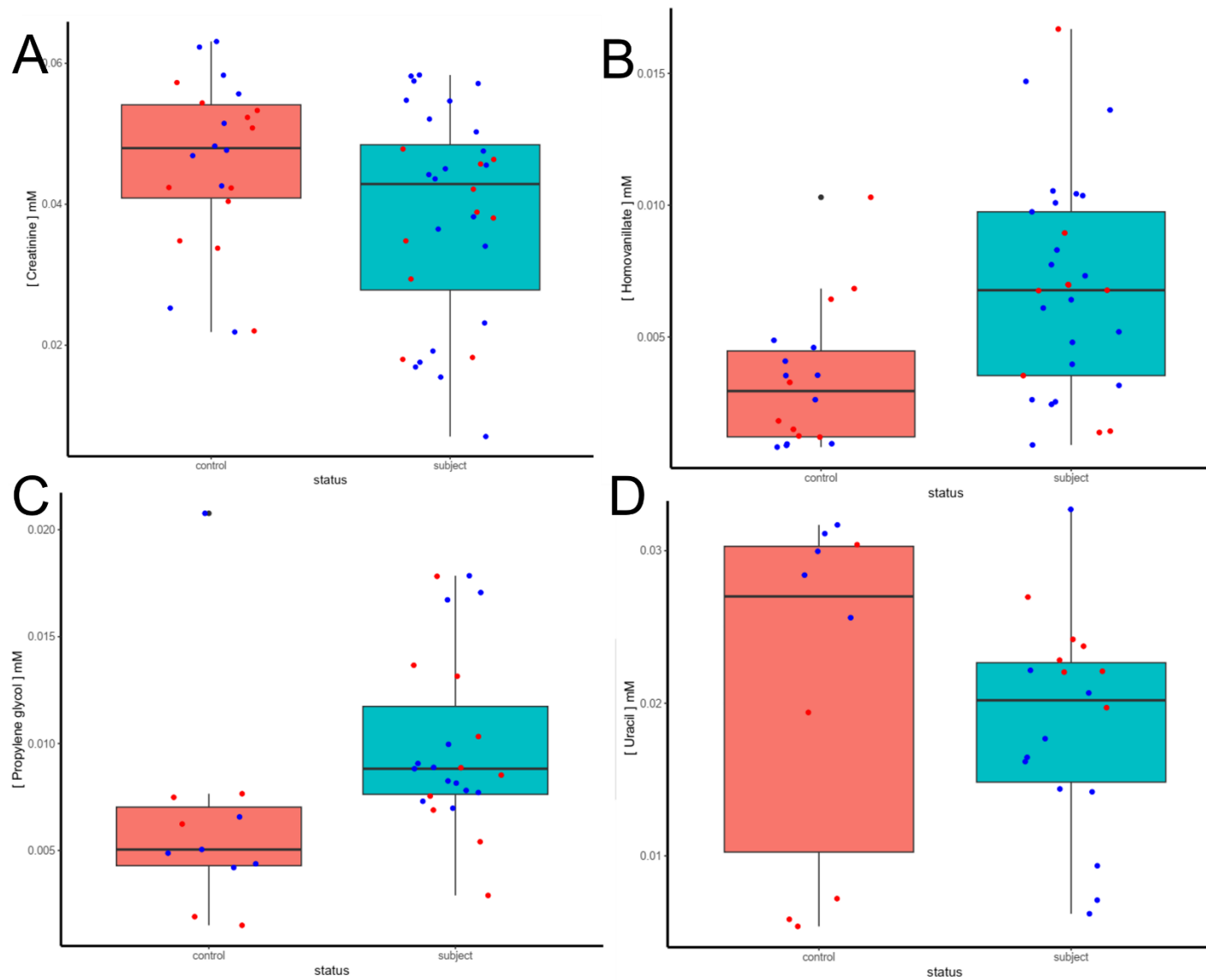


Figure S5: Boxplot comparisons of metabolite levels (also see Figure 4) between controls (dark orange) and subjects (dark cyan) between females (red circles) and males (blue circles) of (A) Creatinine, (B) Homovanillate, (C) Propylene glycol, and (D) Uracil.

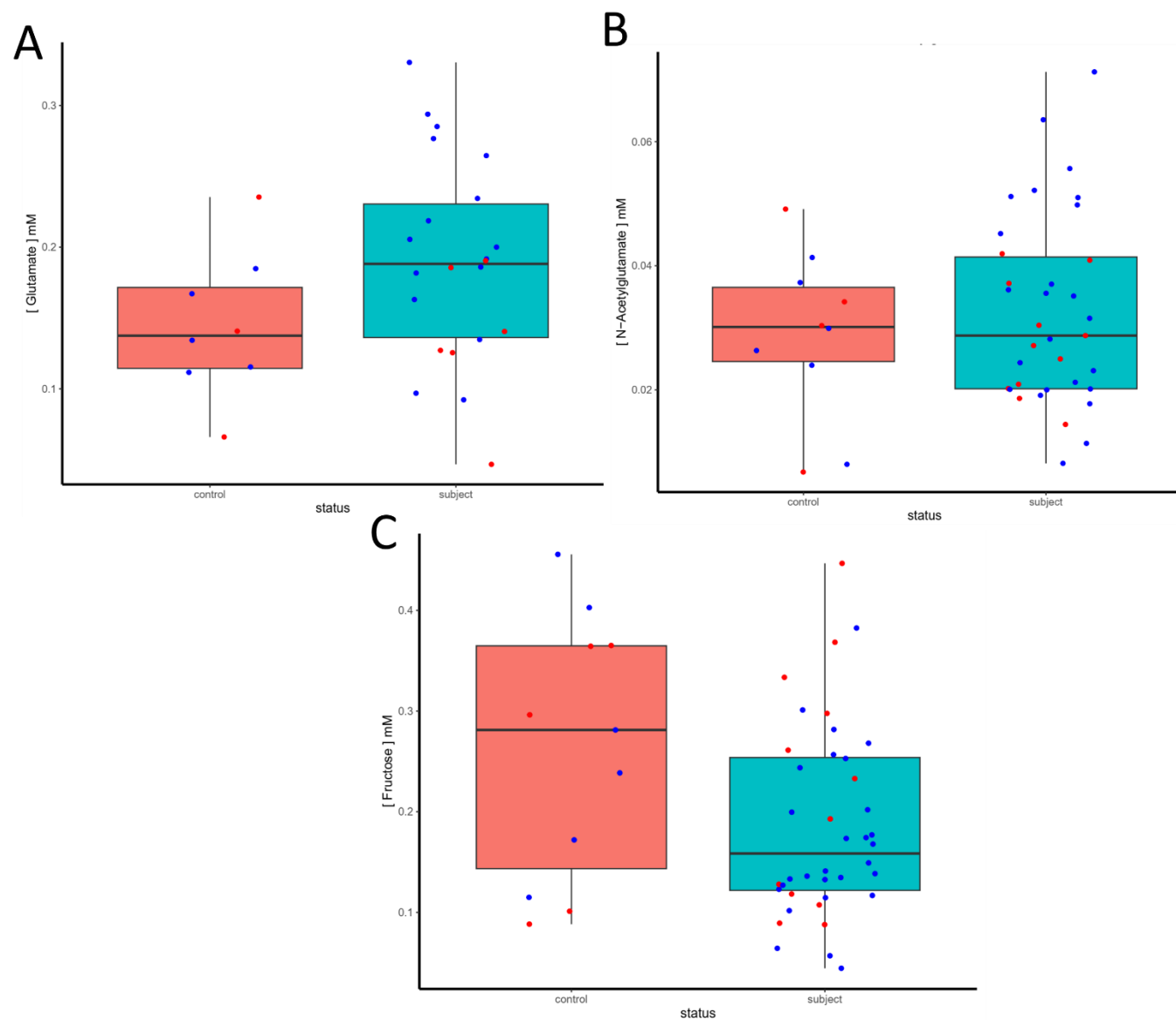
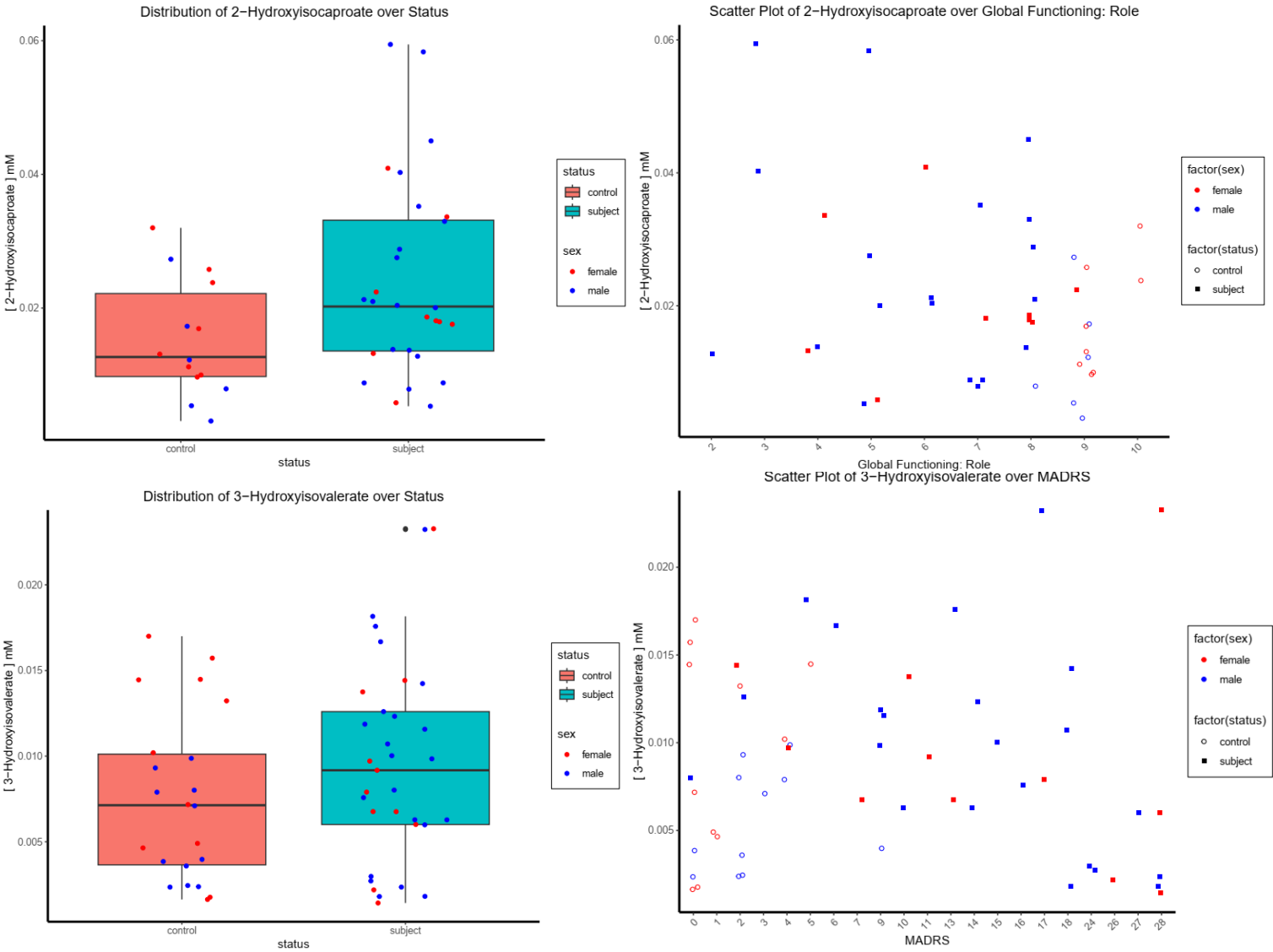
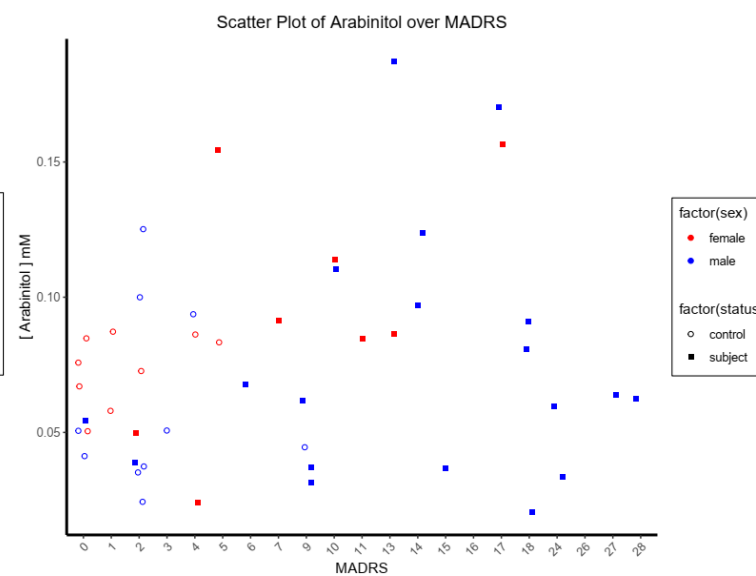
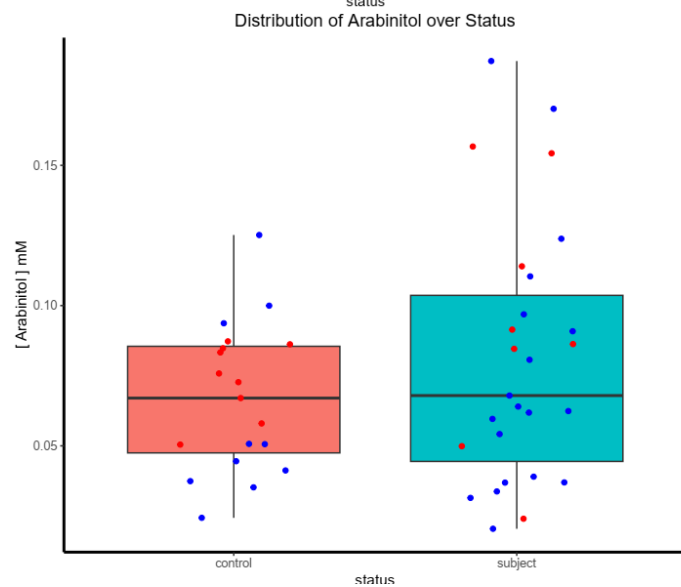
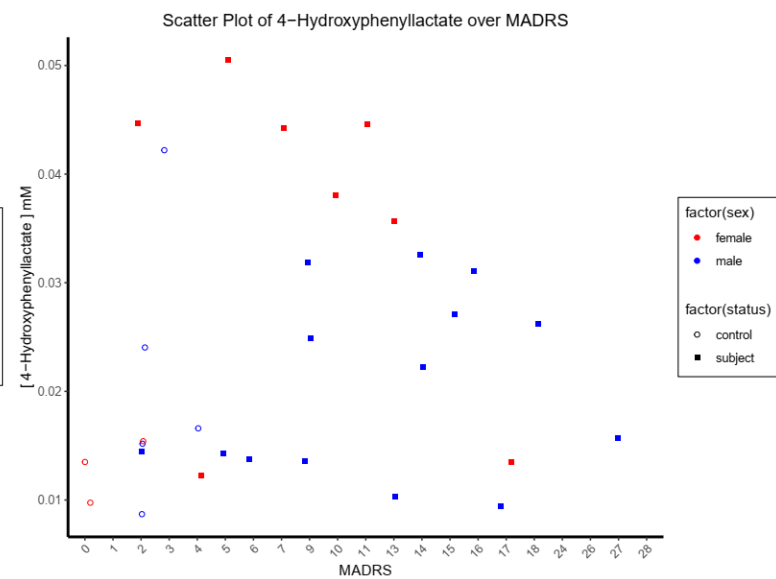
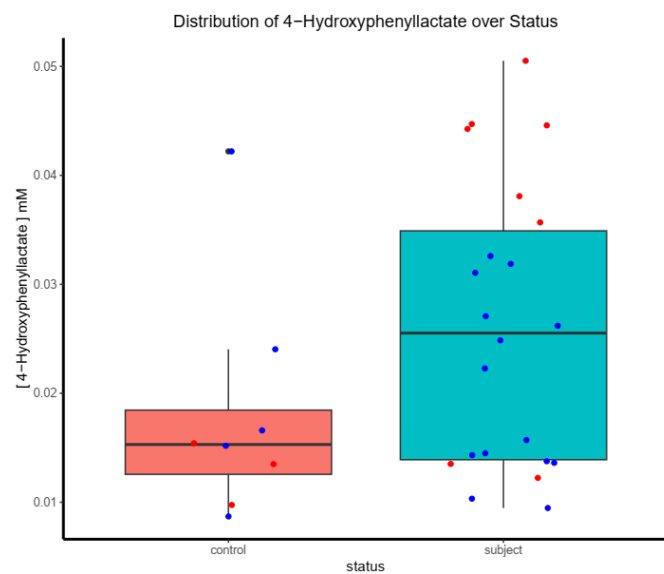
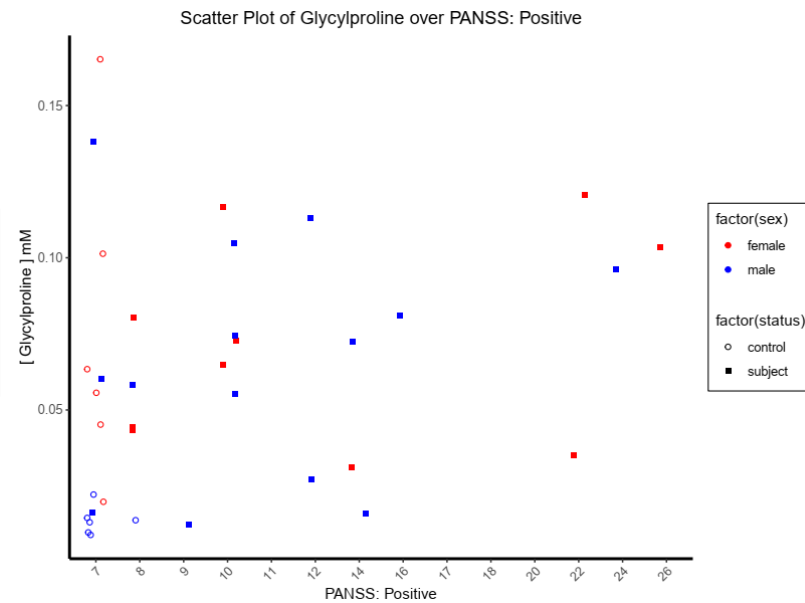
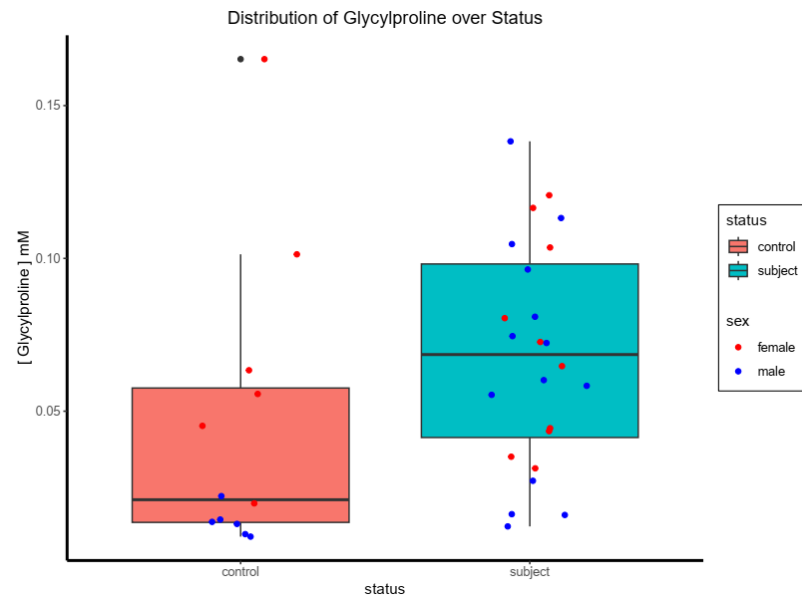
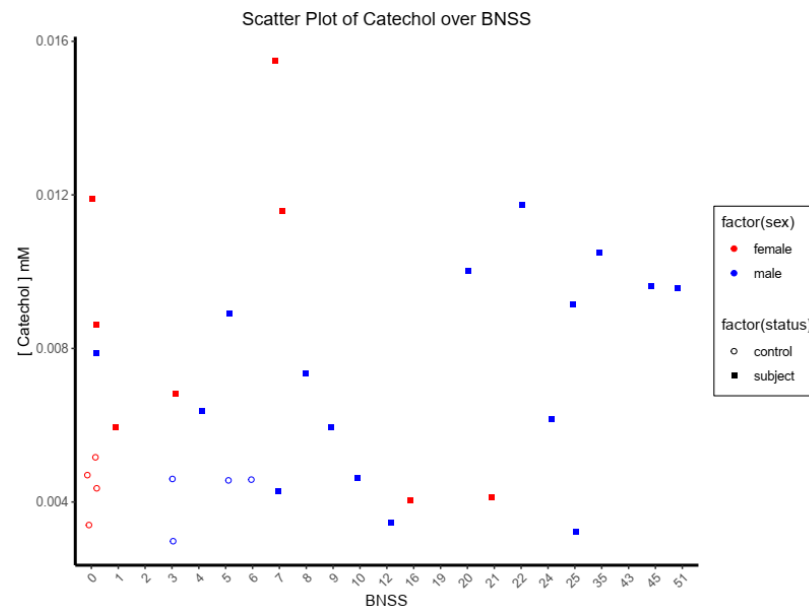
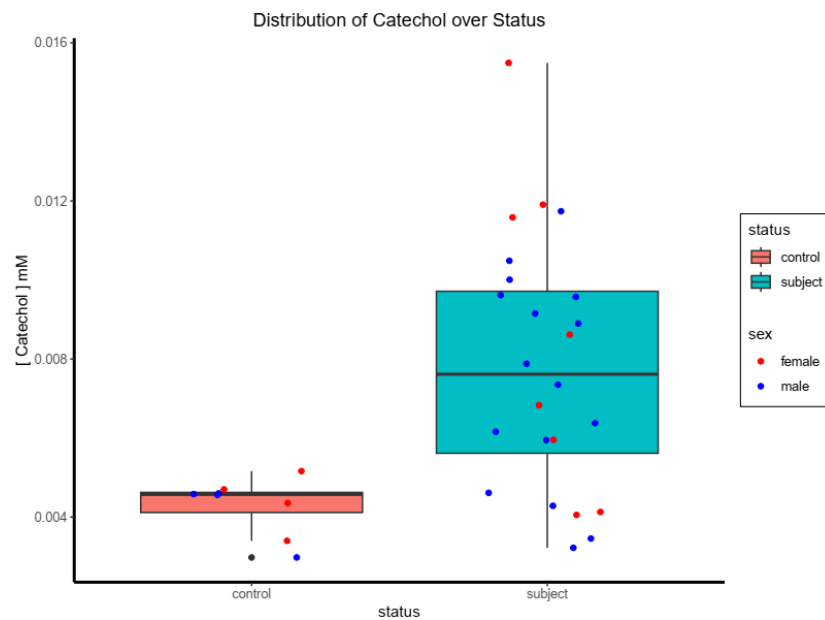


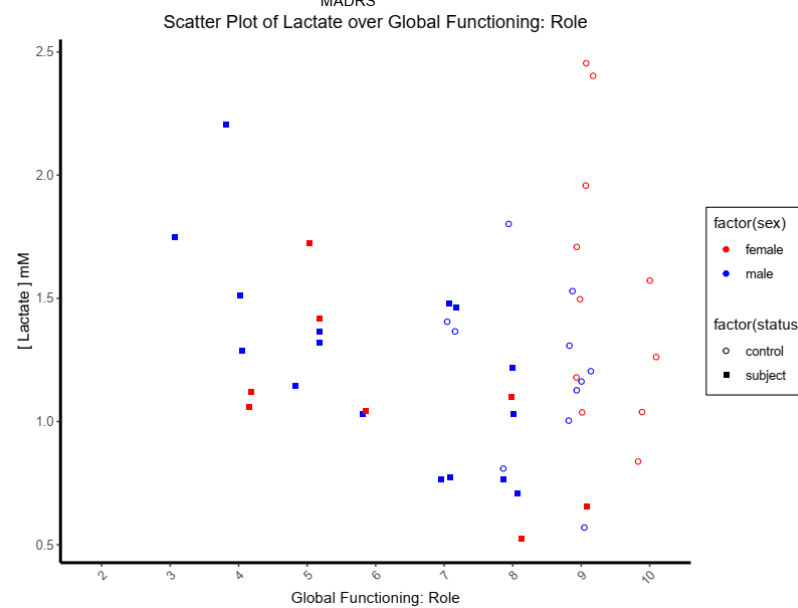
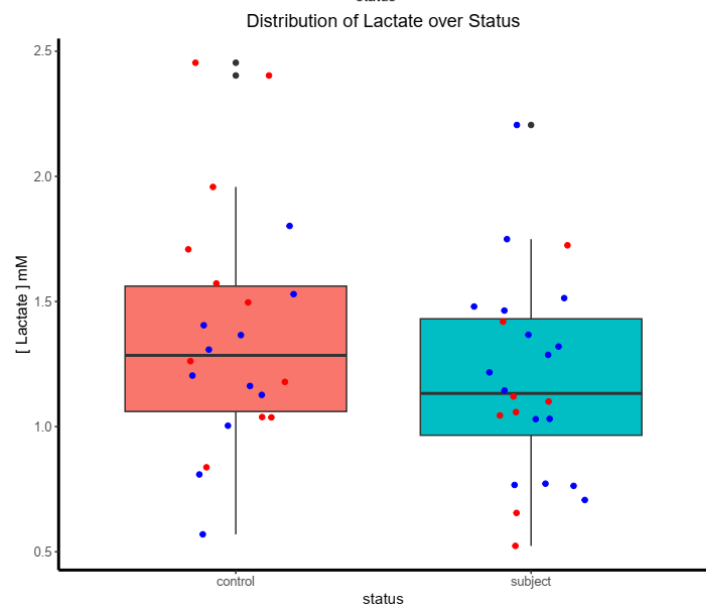
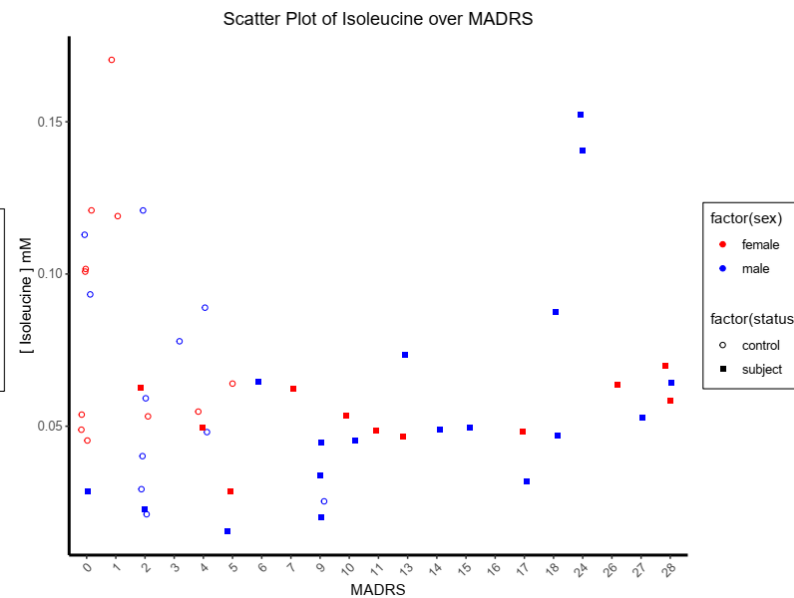
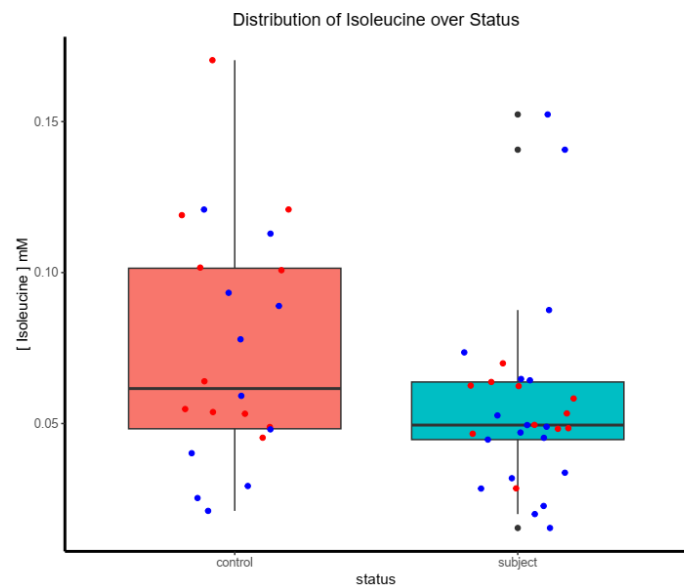
Figure S6: Boxplot comparisons of metabolite levels (also see Figure 5) between controls (dark orange) and subjects (dark cyan) between females (red circles) and males (blue circles) of (A) Glutamate, (B) N-Acetylglutamate, and (C) Fructose.

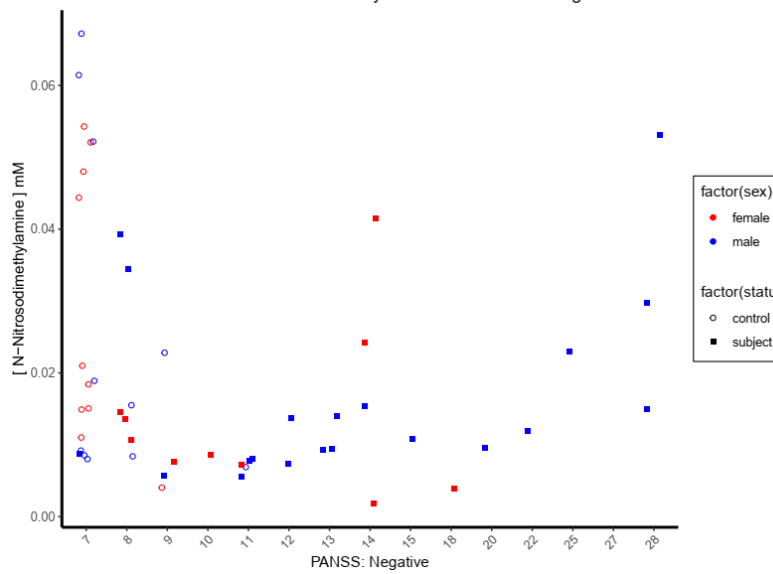
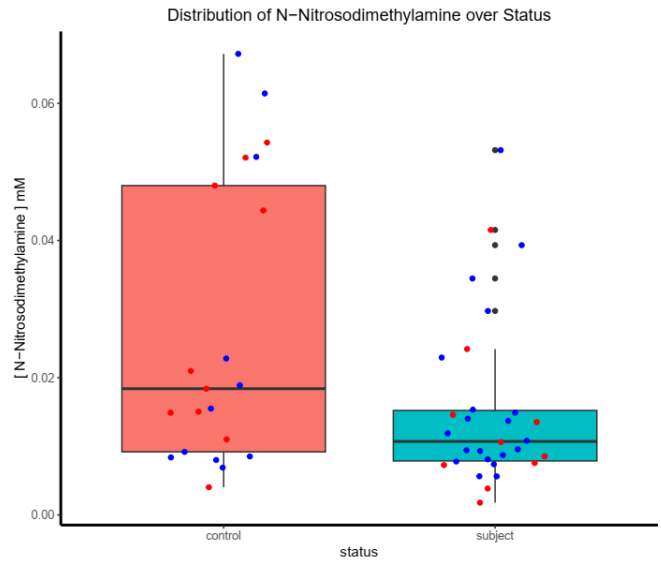
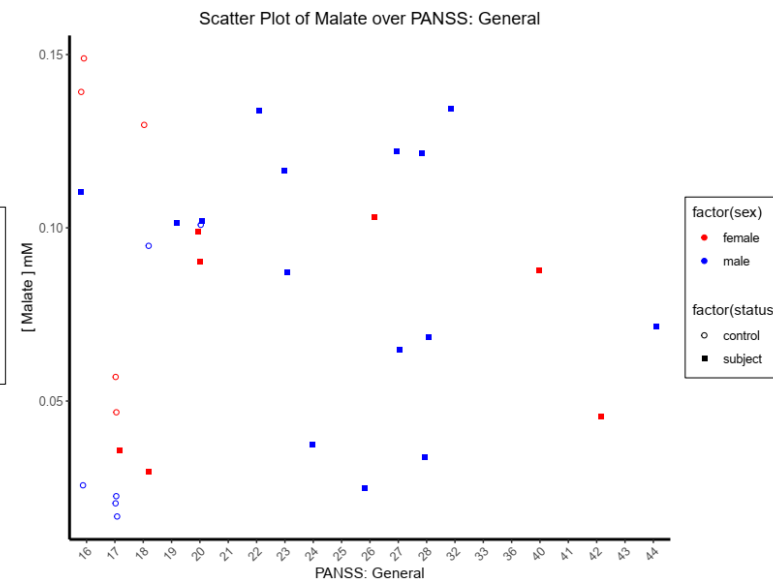
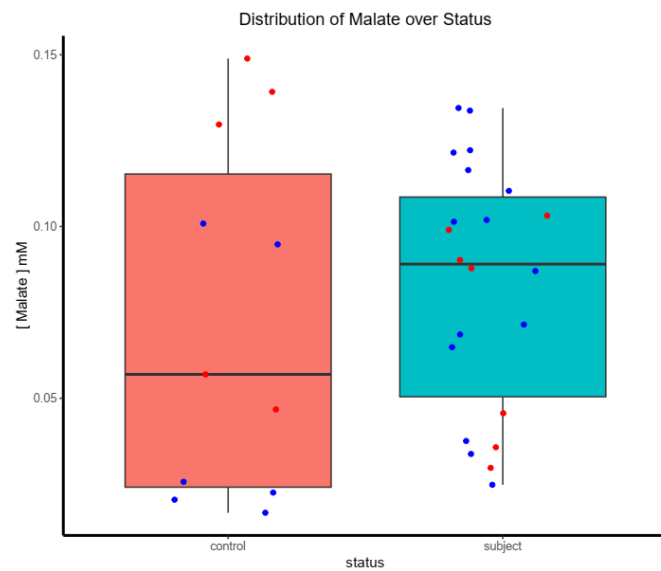
SUPPLEMENTARY FIGURES OF FEW SIGNIFICANT PLASMA METABOLITES (See Figure 3A)

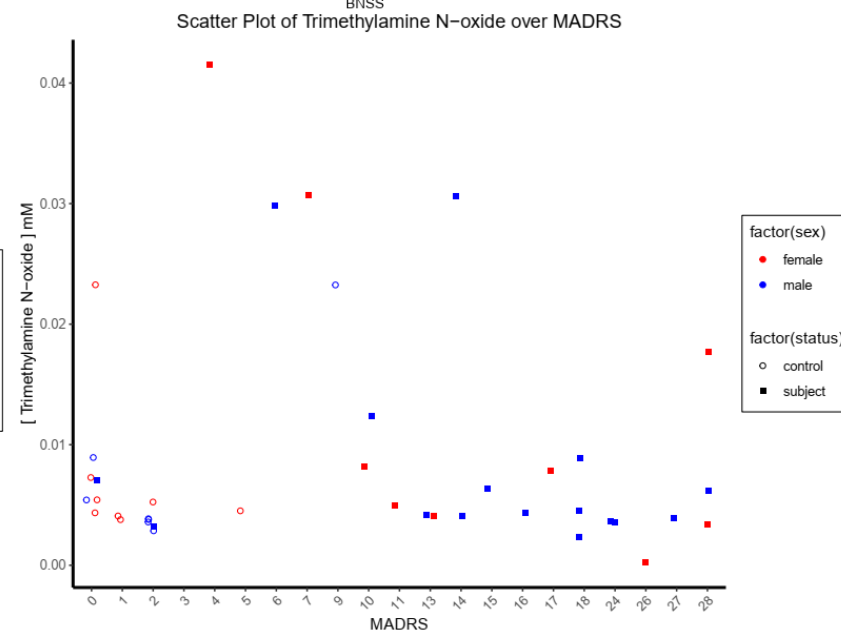
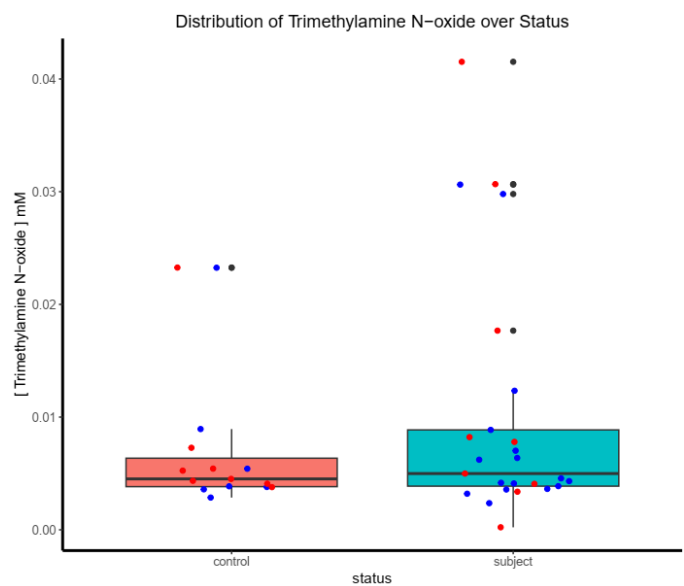
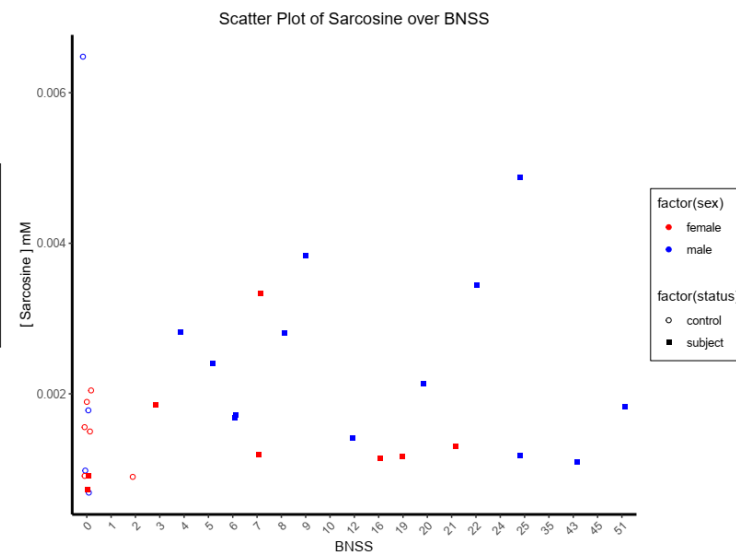
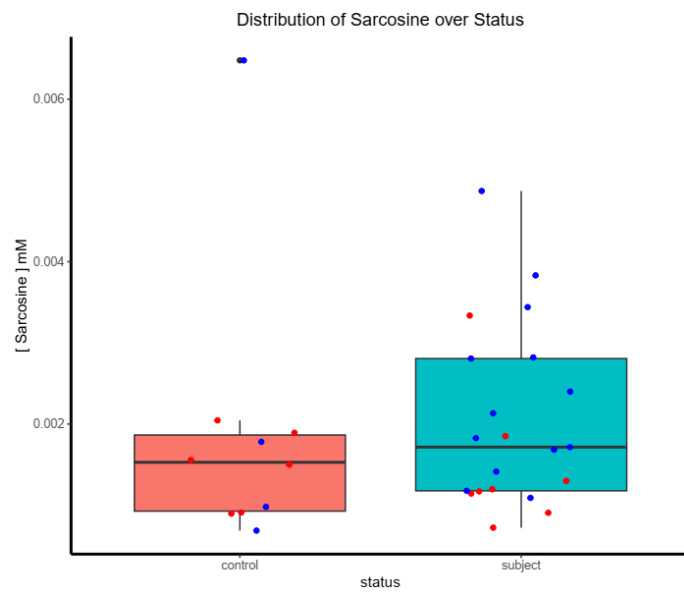




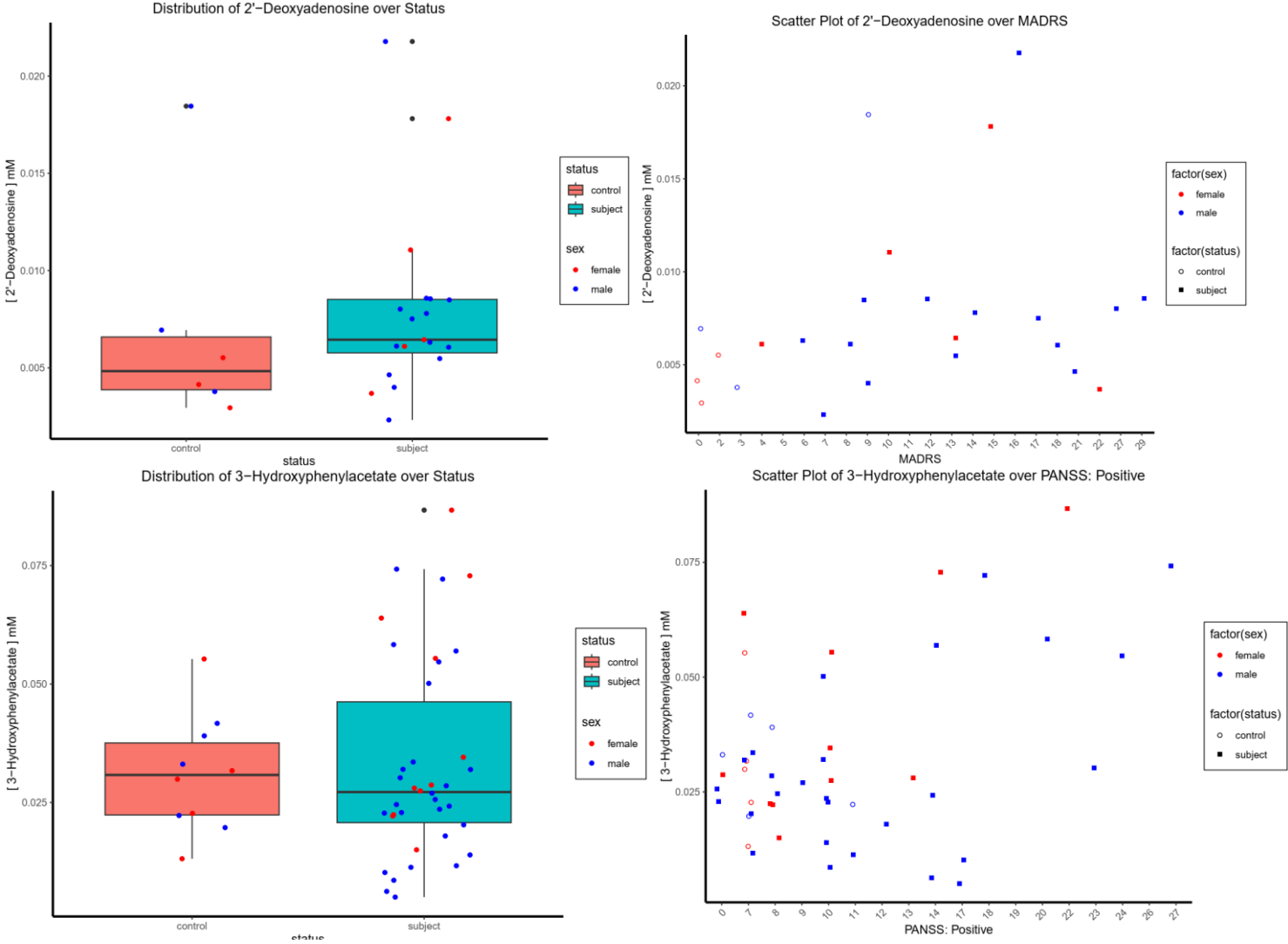


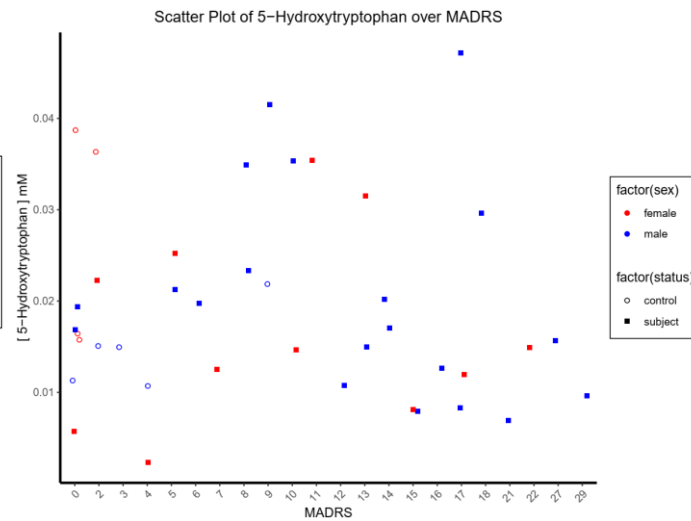
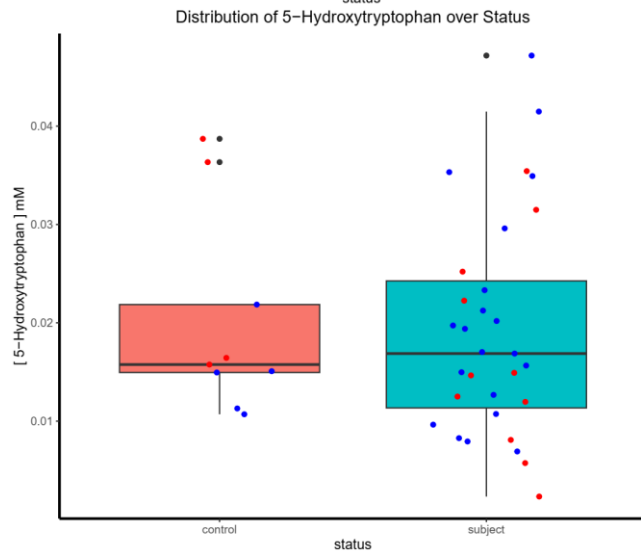
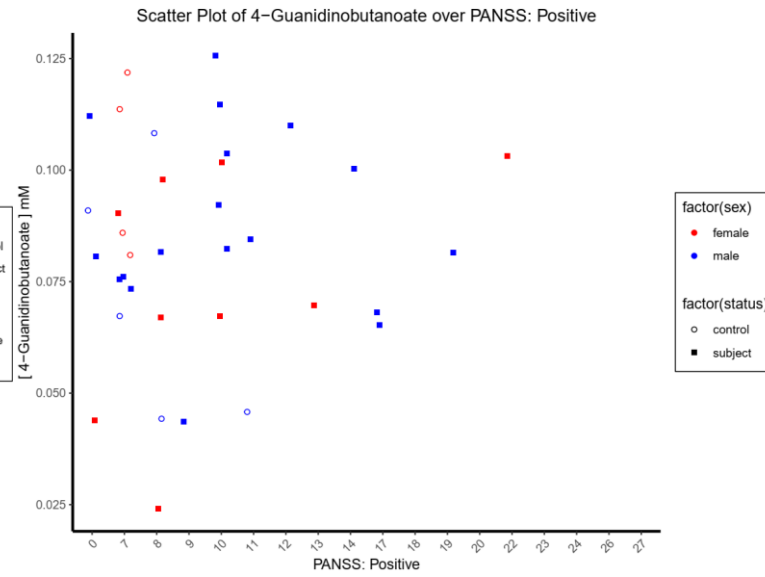
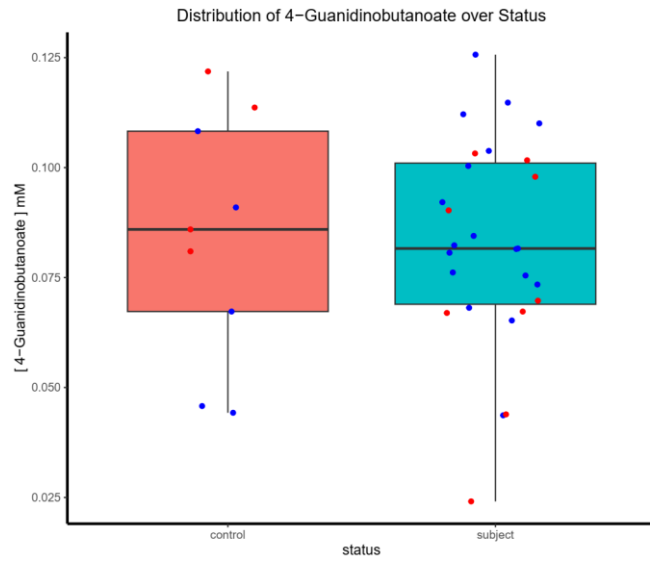


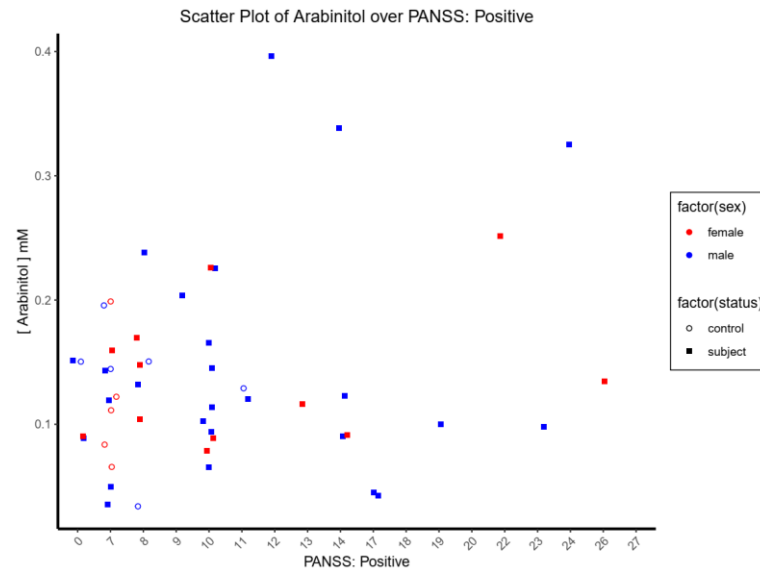
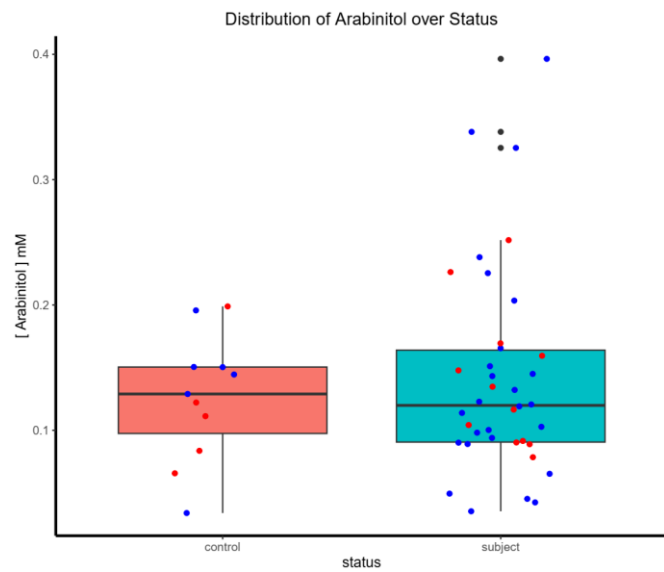
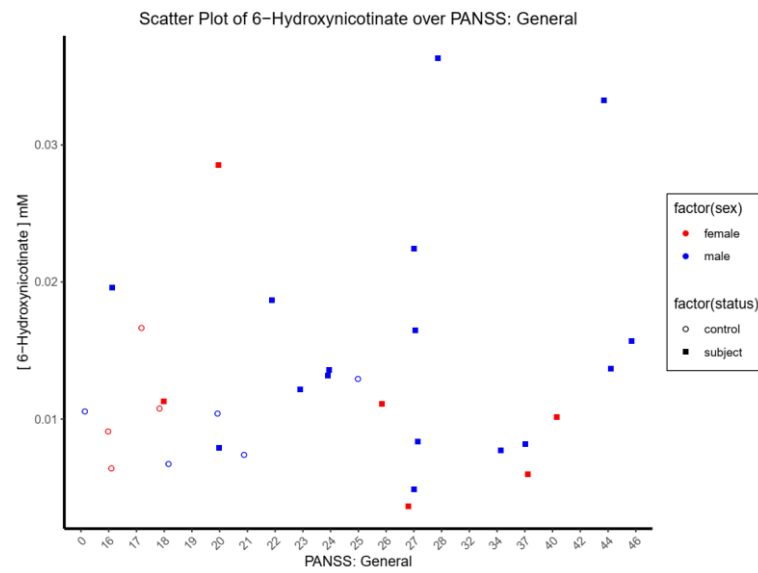
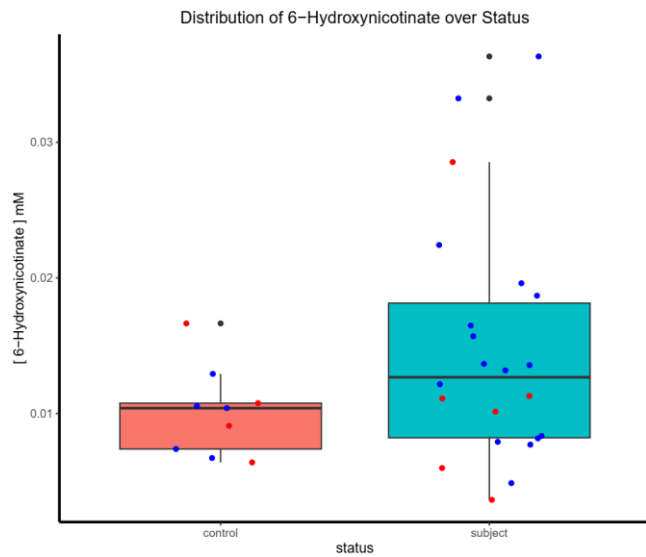


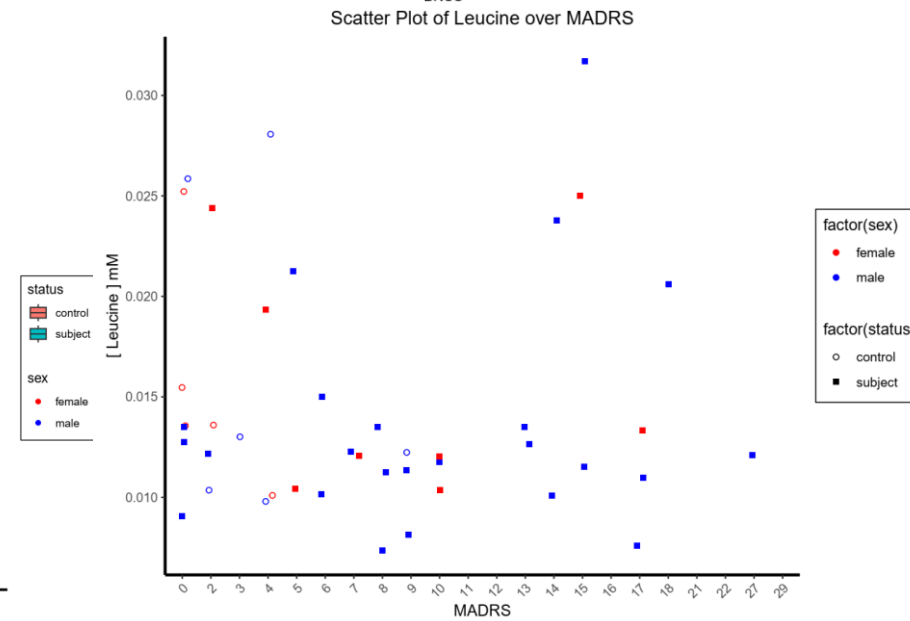
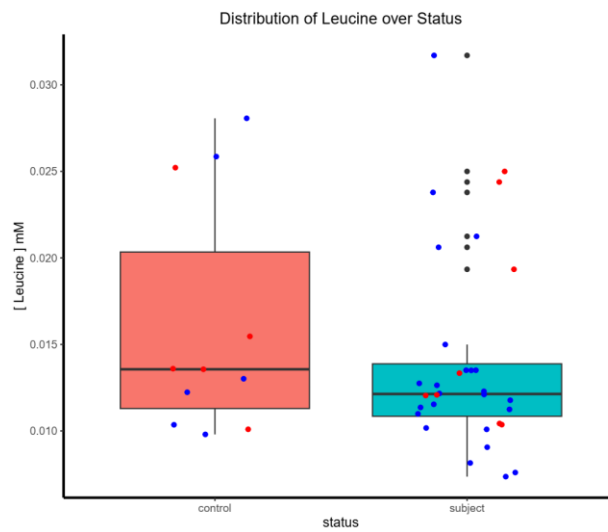
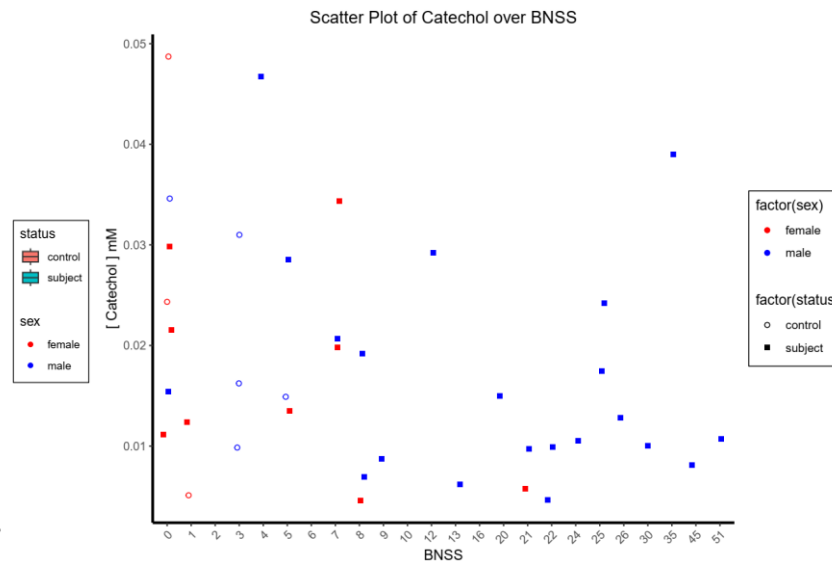
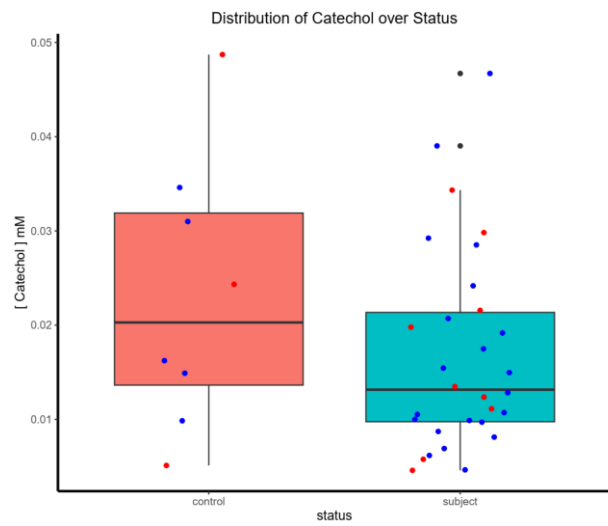


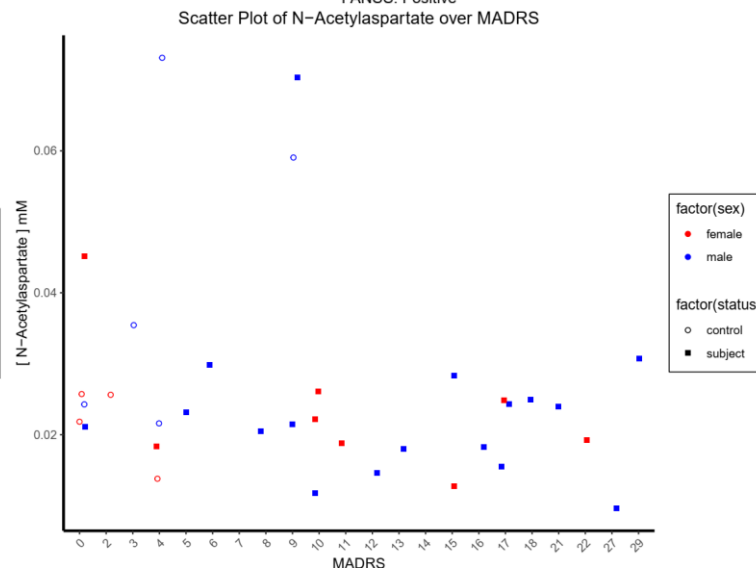
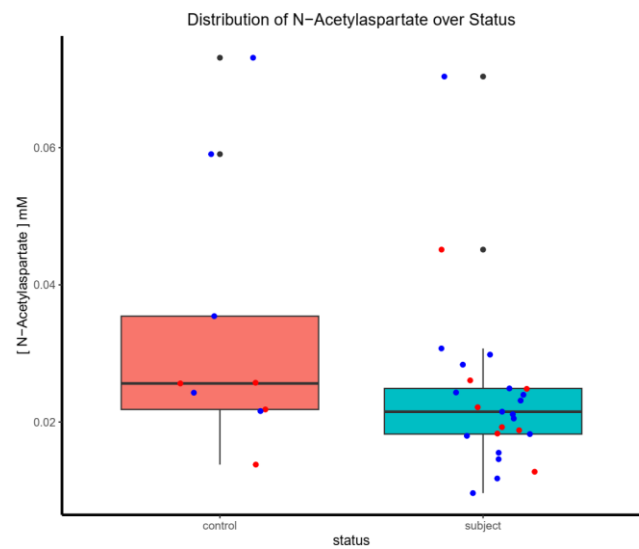
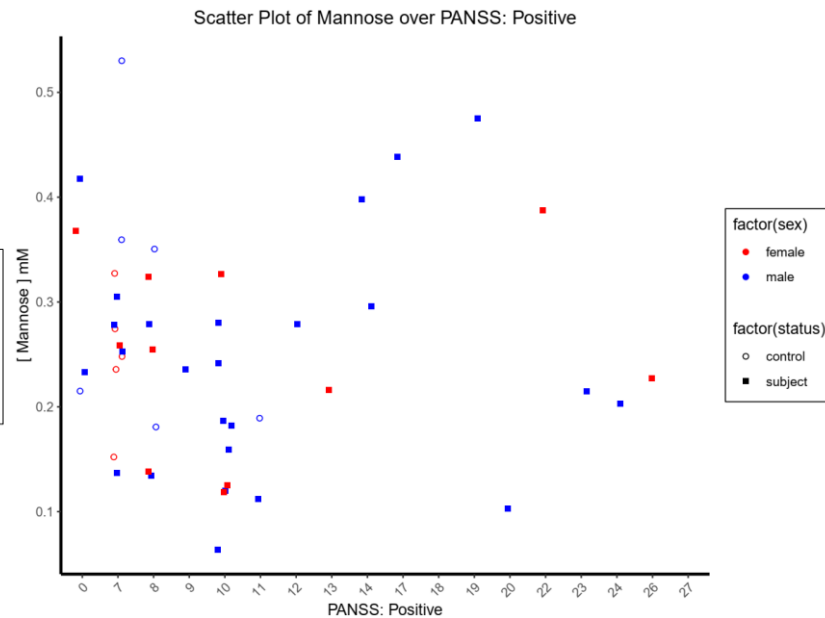
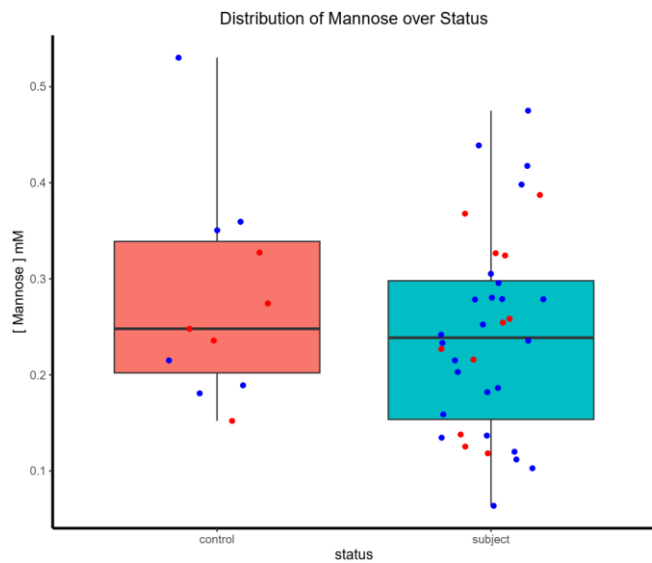
SUPPLEMENTARY FIGURES OF FEW SIGNIFICANT URINE METABOLITES (See Figure 3B)

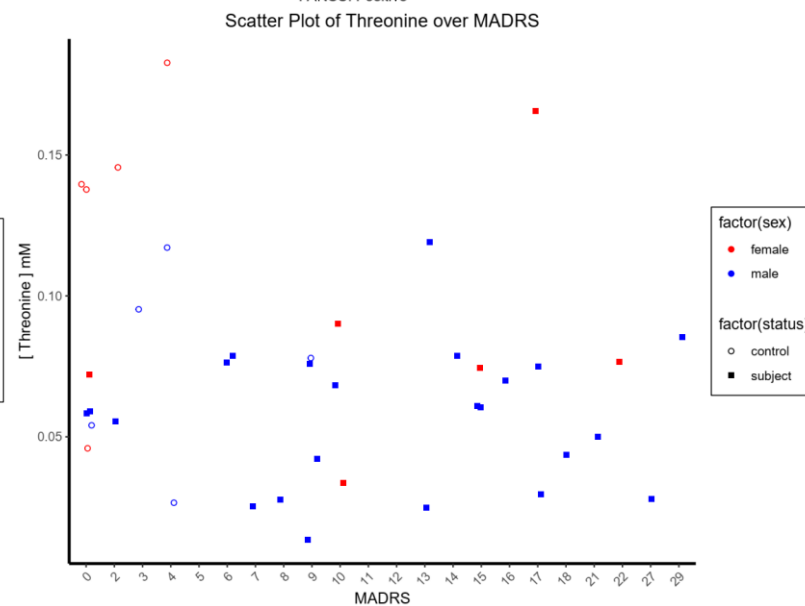
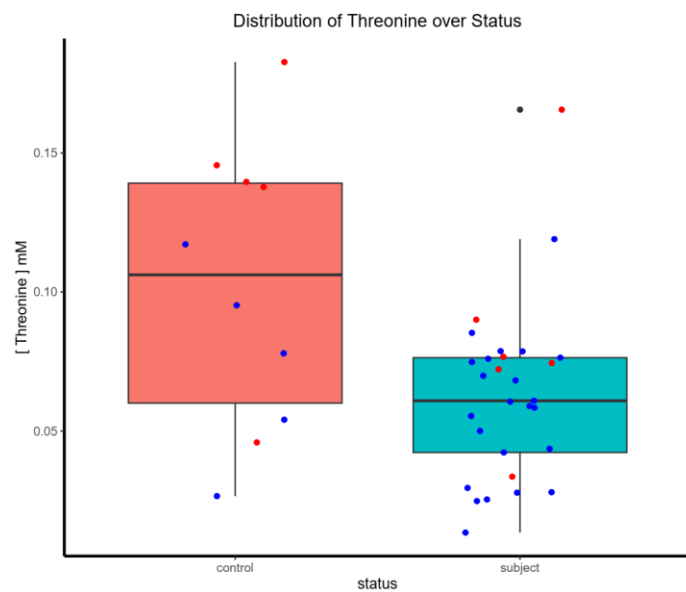
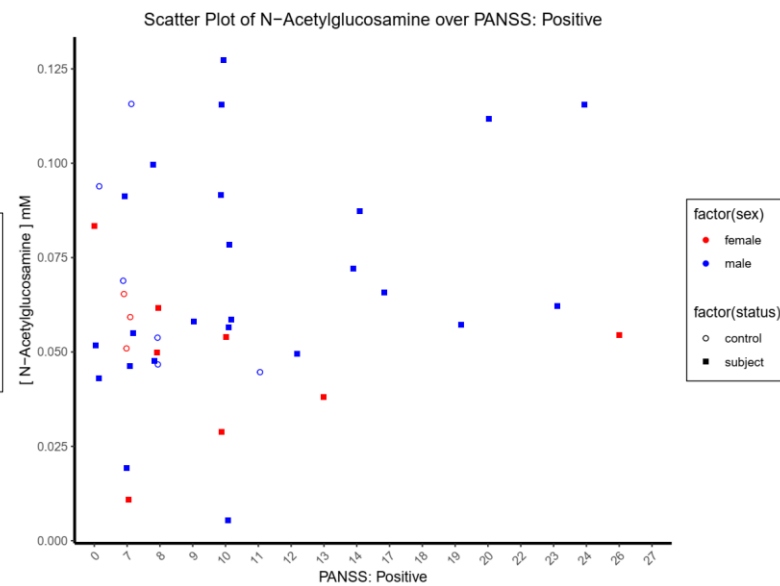
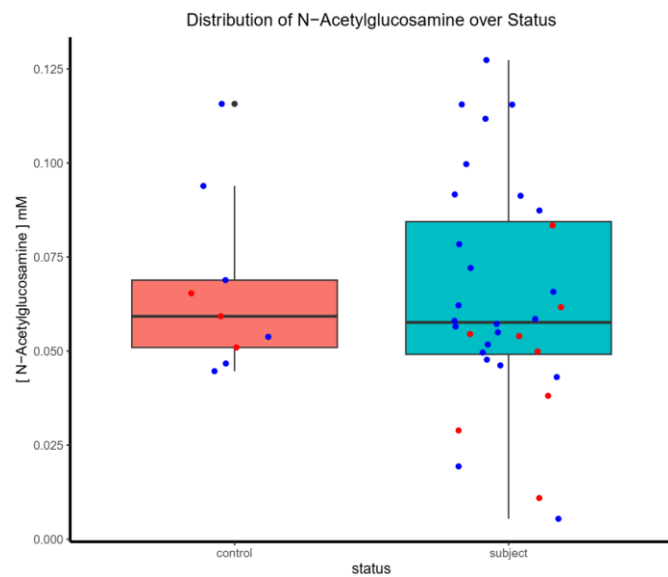


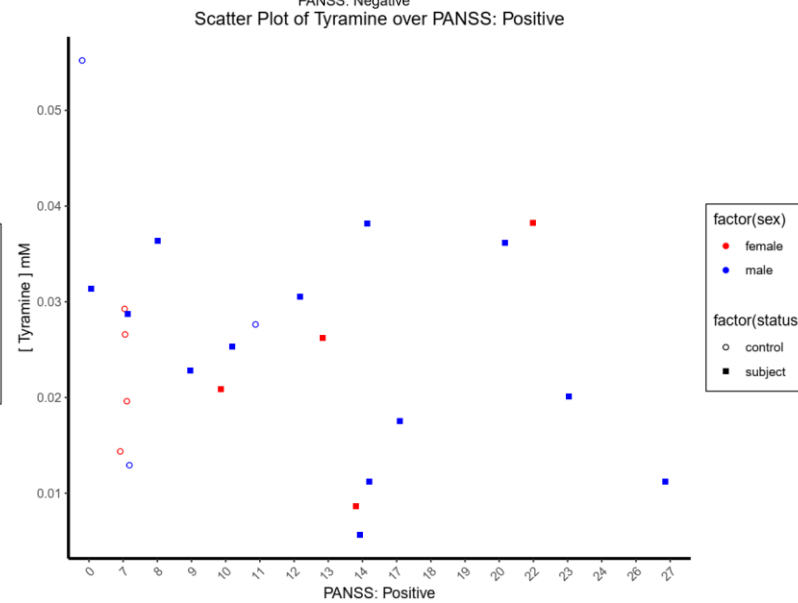
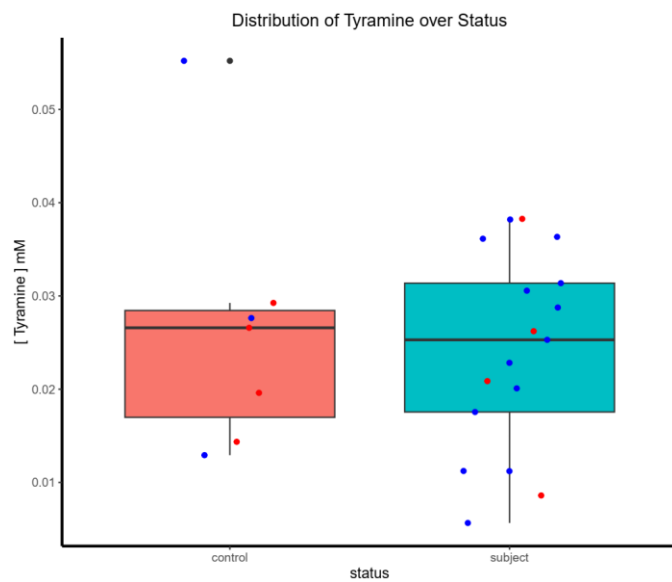
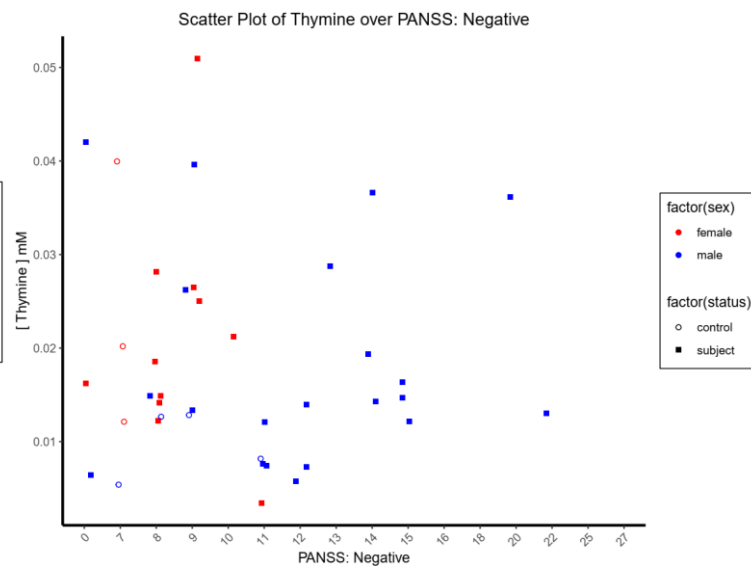
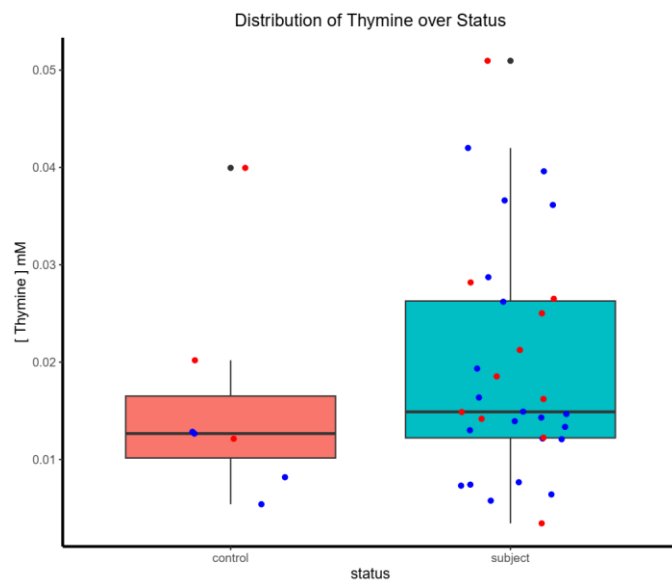












- (51) International Patent Classification:

G01N 33/82 (2006.01) G01N 33/84 (2006.01)
- (21) International Application Number:

PCT/EP2024/073111
- (22) International Filing Date:

16 August 2024 (16.08.2024)
- (25) Filing Language:

English
- (26) Publication Language:

English
- (30) Priority Data:

10 2023 122 636.3

23 August 2023 (23.08.2023) DE
- (71) Applicants:

FORSCHUNGSZENTRUM JÜLICH GMBH [DE/DE]; 52425 Jülich (DE). DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT E.V. [DE/DE]; Königswinterer Str. 522 - 524, 53227 Bonn (DE).
- (72) Inventors:

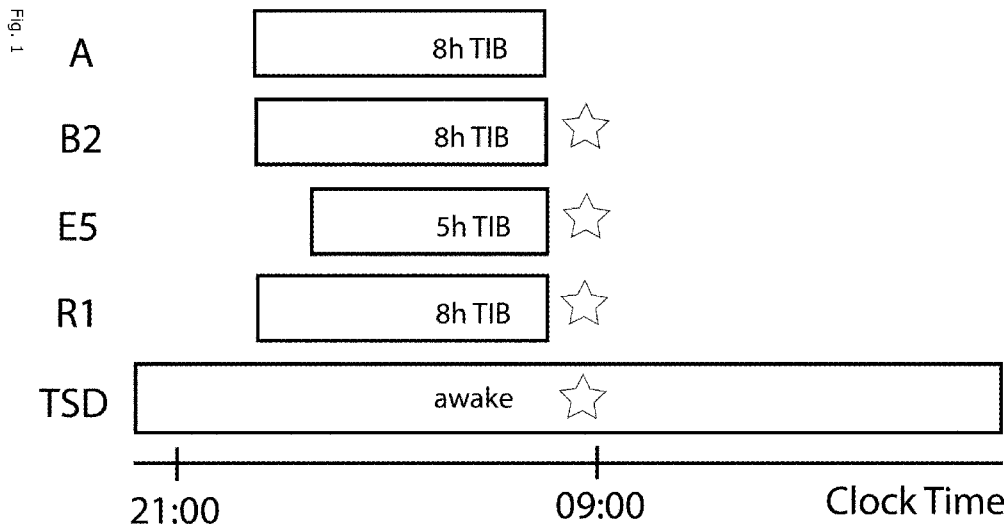
CUKKEMANE, Abhishek; Düsseldorf StraÙe 17, 52428 Jülich (DE). ELMENHORST, David; UhlandstraÙe 6, 52382 Niederzier (DE). ELMENHORST, Eva-Maria; UhlandstraÙe 6, 52382 Niederzier (DE).
- (74) Agent:

MEISSNER BOLTE DÜSSELDORF PATENTANWÄLTE RECHTSANWÄLTE PART-
- (81) Designated States (unless otherwise indicated, for every kind of national protection available):

AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available):

ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: BIOMARKERS FOR THE DIAGNOSIS OF SLEEP DISORDERS



(57) Abstract: The present invention is directed to biomarkers indicative of sleep disorder, processes for diagnosing sleep disorder as well as further systems and computer related implementations in this regard.

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

Biomarkers for the diagnosis of sleep disorders

The present invention is directed to biomarkers indicative of sleep disorder, processes
5 for diagnosing sleep disorder as well as further systems and computer related implementations in this regard.

Any patent, publication, or other disclosure material, in whole or in part, that is said
to be incorporated by reference herein is incorporated herein only to the extent that
10 the incorporated material does not conflict with existing definitions, statements, or other disclosure material set forth in this disclosure. As such, and to the extent necessary, the disclosure as set forth herein supersedes any conflicting material incorporated herein by reference. Any material, or portion thereof, that is said to be
15 incorporated by reference herein, but which conflicts with existing definitions, statements, or other disclosure material set forth herein is only incorporated to the extent that no conflict arises between that incorporated material and the existing disclosure material.

Prior art:

20 Sleep disorders represent a medical condition that is in contrast to an individual's regular sleep pattern and that may affect one's day-to-day function. Modern work environments exacerbate this condition due to over the clock work at the cost of insufficient sleep and increased risk for fatigue-related human error. For instance, CDC-USA (https://www.cdc.gov/sleep/data_statistics.html) performed an extensive
25 set of studies where they identified prevalence of up to 25%, 80% and 40% among children (4 months to 14 years), teens and adults, respectively.

A significant number of people across all ages suffers from sleep disorders. Until recently, sleep disorder was not considered as an ailment in itself because it was associated as a symptom with other forms of psychiatric disorders. However, recently,
30 WHO-ICD-10 and DSM-5 have classified a number of different forms of sleep disorder. Typically, a person is diagnosed for lack/abnormal sleep activity in a sleep lab using polysomnography, which represents a gold standard in studying various sleep ailments. In such a study, several parameters are recorded which include brain wave patterns, the oxygen level in blood, heart rate, breathing, muscle activity, eye and
35 leg movements. Unfortunately, a considerable time lag of a few months follows for

the appointment and in the meanwhile the affected individual may suffer from health deterioration and related complications. Therefore, it is imperative to explore new diagnostic approaches to aid clinicians in the evaluation, to save time and provide for effort saving for all stakeholders.

- 5 As an example, for the requirements of the need for rapid diagnostics, reference could be made to the study pertaining to drowsy driving by the CDC-USA (<https://www.cdc.gov/sleep/features/drowsy-driving.html>). Such a phenomenon is in fact common and prevalent in a number of professions such as health care, transportation, travel services, call centres to name a few. As per the estimates of
- 10 CDC-USA, drowsy driving results in ca. 6,000 fatal accidents per year in the USA alone due to poor health arising from lack of sleep. This further exemplifies the necessity to develop a rapid diagnostic tool to diagnose people suffering from sleep disorder but also for risk-management purposes to identify potentially afflicted individuals, thereby to prevent any accidents.
- 15 Usually, the entire diagnosis and therapy of sleep disorders and related psychiatric diseases is structured around diagnosis based on polysomnography, oxygen levels and related biomarker characterization in clinical settings. Such a system poses significant challenges to and burden on the medical facilities and subjects due to limited number of setups and experts' availability. It thereby, ensues a chain reaction
- 20 where the troubled individual is diagnosed at wrong juncture(s) while her/his health deteriorates and treatment may be a tad late for rapid recovery.
- Therefore, it is important to develop a rapid diagnostic tool to diagnose people suffering from sleep disorder.

25 Object of the invention:

It was an object of the present invention to provide a process of diagnosing sleep disorder in a patient.

It was a further object of the invention to provide a process for the determination whether a therapy against sleep disorder is successful.

- 30 Additionally, it was an object of the present invention to provide a computer-implemented process for diagnosing sleep problems in a patient as well as a computer-readable storage medium having computer-readable instructions stored thereon, particularly useable in or for this process or similar processes.

Further, it was an object of the present invention to provide a system for diagnosing

35 sleep problems in patients.

Not the least, it was an object of the present invention to provide biomarkers indicative of sleep disorder in a patient and/or a set of biomarkers indicative of sleep disorder in a patient.

- 5 Further objects will become apparent to those skilled in the art upon regarding the following disclosure.

Solution:

- 10 These and other objects are solved by the matter outlined in the appended independent claims.

Preferred embodiments are outlined in the dependent claims as well as the following description.

15 Detailed description:

It should be understood that this invention is not limited to the embodiments disclosed in the description that follows, but is intended to cover modifications that are within the spirit and scope of the invention, as defined by the claims.

- 20 In the present application, including the claims, other than in the operating examples or where otherwise indicated, all numbers expressing quantities or characteristics are to be understood as being modified in all instances by the term "about". Accordingly, unless indicated to the contrary, any numerical parameters set forth in the following description may vary depending on the desired properties one seeks to obtain in the compositions and processes according to the present disclosure. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter described in the present description should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques.

- 25 Also, it should be understood that any numerical range recited herein is intended to include all sub-ranges subsumed therein. For example, a range of "1 to 10" is intended to include all sub-ranges between (and including) the recited minimum value of 1 and the recited maximum value of 10, that is, having a minimum value equal to or greater than 1 and a maximum value of equal to or less than 10. The terms "one," "a," or "an" as used herein are intended to include "at least one" or "one or more," unless
35 otherwise indicated.

In the context of the present invention, unless otherwise specified, reactions and operations are done at room temperature, which is understood as being 20°C.

In the context of the present invention, unless otherwise specified, reactions and operations are done at ambient pressure, which is understood as being 101,325 kPa.

5 In the context of the present invention, all indications of quantity are to be understood as indications of weight, unless otherwise indicated.

In the context of the present invent, the term "sleep disorder" preferably means any deviation from the sleep pattern, sleep duration or sleep behaviour, which is considered as normal. These deviations include alterations of the duration of the sleep
10 period or of sleep stages, difficulties in falling asleep, maintaining sleep or early morning awakenings (such as insomnia or hypersomnia), disturbances of sleep due to circadian misalignment of different causes (such as non-24, shift work, jetlag), due to sleep-related disorders (such as sleep apnoea, restless legs syndrome, periodic limb movements, narcolepsy), due to abnormal sleep behaviours (such as nightmares,
15 night terror or sleep walking) or secondary to other diseases (such as cancer or psychiatric disorders) more preferably, "sleep disorder" means any of the six major classes, including subclasses, as defined by WHO-ICD-10 as of August 18, 2023, in particular "sleep disorder" means any of sleep apnoea, restless legs syndrome, periodic limb movements, narcolepsy (or combinations thereof).

20

The present invention in a first major embodiment is directed to a process of diagnosing sleep disorder in a patient comprising or consisting of the following steps:

- i) measuring, in a blood sample from a patient, the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine
25 (vitamin B6), N-acetyl glutamine (NAG), glycocholate, 3-hydroxyphenyl acetate (3-HPA), formate, tyramine, 4-Guanidinobutanoate, trimethyl amine N-oxide (TMAO) and mixtures thereof;
- ii) comparing the measured levels for the detected biomarkers with predetermined threshold levels;
- 30 iii) if at least two of the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, diagnosing the patient
 - iiia) as having sleep disorder, or
 - iiib) as being at least in danger of developing sleep disorder.

35

The discrimination between whether one is diagnosed with sleep disorder or being in danger of developing sleep disorder in the context of the present invention is preferably based on the amount with which the measured levels deviate from the baseline the pre-determined threshold values. The amount necessary to distinguish the one from the other is not always a sharp one and differs, naturally, from individual to individual. However, for the respective biomarkers as well as various combinations thereof tables connecting the values to sleep disorders are empirically prepared and can be used for this. Additionally, such tables can be prepared based on computer models.

It should be noted, that in preferred embodiments of the present invention the diagnosis is that a certain individual "has sleep disorder or is in danger of developing sleep disorder", but not actually a discrimination between these two possibilities. This is due to the fact that individuals (as the word implies) are each different and different values can mean different states for them.

In this sense in some preferred embodiments of the present invention, the biomarker detection of the present invention is the first step in a more encompassing diagnosis; if the biomarker values detected according to the present invention suggest or even prove (if the values deviate sufficiently high for the threshold values) that a patient suffers from sleep disorders, he is further examined, for example with the methods known from the prior art. The advantage here is that the biomarker detections can show that a more specific and encompassing examination is of (some) urgency.

It should be noted that if salts like "acetate", "formate", "glycocholate" or the like are mentioned in the present application, this is intended to encompass those with any counterions as well as the free acids themselves.

In this process, it is beneficial – and thus preferred, if at least three, more preferably at least four, especially all eight, of the biomarkers are selected and analysed. It is to be understood that this goes for all the major embodiments of the present invention, i.e., it is always preferred if at least three, more preferably at least four, especially all eight, of the biomarkers are subject of the respective embodiments. This is because the accuracy of the diagnosis increases with the number of biomarkers.

In preferred embodiments of this process at least the following two biomarkers are measured: N-acetyl glutamine and 3-hydroxyphenyl acetate.

In other preferred embodiments of this process at least the following three biomarkers are measured: pyridoxine, N-acetyl glutamine and 3-hydroxyphenyl acetate are measured.

In still other preferred embodiments of this process at least the following four biomarkers are measured: pyridoxine, N-acetyl glutamine, 3-hydroxyphenyl acetate and trimethyl amine N-oxide.

In preferred embodiments of this process the biomarker levels in the sample are determined by quantitative NMR (qNMR) metabolomics, preferably with ^1H -NMR.

Metabolomics is an important tool among the functional genomics studies. It encompasses studying a collection of metabolites, or small molecule chemicals (with a molecular mass $<1,500$ Da) that can be detected in biological samples. It is because metabolites are sensitive to external stimuli and physiological processes, the metabolome of an individual describes her/his phenotype. Such a metabolic readout of the phenotype is referred to as a "metabotype". To study ones metabotype, the biological samples are, in the context of the present invention, measured using quantitative nuclear magnetic resonance (qNMR) spectroscopy, which is a technique that enables one to accurately and reproducibly quantify signals from metabolites in complex mixtures. In the present invention, it is thus possible to compare the biochemical profile of healthy and diseased subjects and (quantitatively) identify the relevant biomarkers.

For these measurement's, the blood samples are treated, i.e. prepared, as usual and known in the art, preferably by mixing with organic solvent or solvent mixture, in particular equivolume 1:1 mixture of chloroform:methanol, and then separating aqueous layer and organic solvent layer, in particular by centrifugation, and drying the aqueous layer. The dried sample can then be prepared for NMR as usual by re-suspending/re-solving in appropriate solvent, in particular in water or an aqueous mixture, optionally with added salts, and adding an internal standard, such as known in the art, preferably sodium trimethylsilylpropanesulfonat (DSS), tetramethylsilane (TMS) or trimethylsilylpropanoic acid (TSP), in particular trimethylsilylpropanoic acid (TSP).

Throughout this invention, the threshold levels are the mean values determined for the respective biomarkers from a number of individuals, healthy persons during normal life circumstances, preferably at least 18 individual persons.

- 5 In some preferred embodiments the mean values for the concentrations of the respective biomarkers in the aqueous phase from plasma of blood samples [as determined with $^1\text{H-NMR}$] are
- 2.693 $\mu\text{mol/L}$ for pyridoxine;
 - 101.300 $\mu\text{mol/L}$ for 3-hydroxyphenyl acetate;
 - 10 - 79.660 $\mu\text{mol/L}$ for formate;
 - 25.690 $\mu\text{mol/L}$ for N-acetylglutamine;
 - 32.800 $\mu\text{mol/L}$ for tyramine;
 - 11.480 $\mu\text{mol/L}$ for 4-guanidinobutanoate;
 - 16.190 $\mu\text{mol/L}$ for glycocholate;
 - 15 - 17.020 $\mu\text{mol/L}$ for trimethylamine N-oxide.

A second major embodiment of the present invention is a process for the determination whether a therapy against sleep disorder is successful comprising or consisting of the following steps:

- 20 I) diagnosing sleep disorder in a patient, preferably but not necessarily, with a process as outlined above;
- II) measuring levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide
- 25 and mixtures thereof, in a blood sample taken from the patient at the time of diagnosis or at least prior to the start of the therapy;
- III) measuring the levels of the same biomarkers as in step II) in blood samples taken from the patient during the therapy, preferably at predetermined times; and
- 30 IV) comparing each of the measured levels for the detected biomarkers in the respective blood samples with predetermined threshold levels, in particular, if the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%;

- V) determining whether the results of the comparison of the biomarkers change during the therapy, and thereby determining whether the therapy is effective.

It should be noted, that in the context of steps IV) and V) effective can be positive effectiveness meaning that the therapy is successful. But it is also possible to come to negative effectiveness, which would mean that the therapy not only is not effective, but on the contrary is detrimental to the sleep disorder of the respective patient.

As in the process of the first major embodiment, it is preferred if the biomarker levels in the sample are determined by quantitative NMR metabolomics.

A third major embodiment of the present invention is a process for the detection of at least two biomarkers in a blood sample from a patient comprising or consisting of the steps of:

- obtaining a blood sample from a patient;
- detecting the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in the sample by quantitative NMR.

A fourth major embodiment of the present invention is a computer-readable storage medium having computer-readable instructions stored thereon, which, when executed, cause the computer to:

- a) access data attributed to at least one blood sample from a patient, wherein the data comprises the measured levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in the respective blood sample;
- b) compare the data with pre-determined levels for the respective biomarkers, the pre-determined levels being indicative of sleep disorder; and
- c1) if the comparison reveals that the measured levels of at least two of the biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, output of an alert to the computer's user that the patient in question is diagnosed as

having sleep disorder, or is diagnosed as being at least in danger of developing sleep disorder,

or

c2) if the comparison reveals that the measured levels of none or at most one of the biomarkers do not deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, output of an information to the computer's user that the patient in question is not diagnosed as having sleep disorder, or, at the time the respective blood sample was taken, is not in danger of developing sleep disorder.

In further preferred embodiments of the present invention, the processes described above in the context of the first, second and/or third major embodiments employ the computer-readable storage medium to help in the execution of the respective process steps.

A fifth major embodiment of the present invention is a computer-implemented process for diagnosing sleep problems in a patient comprising or consisting of the following steps:

- i) measuring, in a blood sample from a patient, the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof;
- ii) comparing the measured levels for the detected biomarkers with predetermined threshold levels;
- iii) diagnosing the patient with sleep disorder, if at least two of the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a percentage of at least 3%;

wherein the process is carried out on a computer in which a computer-readable storage medium is present having stored thereon instructions that cause the computer to execute the process, preferably a computer-readable storage medium as described above.

In further preferred embodiments of the present invention, the processes described above in the context of the first, second and/or third major embodiments are preferably made to be at computer-implemented process as just described.

A sixth major embodiment of the present invention is a system for diagnosing sleep problems in patients comprising or consisting of:

- 5 A) a detection device configured to measure the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in blood samples by performing quantitative NMR measurements;
- 10 B) an analysing device comprising a computer, and at least one computer readable storage medium as described above;
- C) a database storing pre-determined levels for the respective biomarkers, the pre-determined levels being indicative of sleep disorder.

15 In further preferred embodiments of the present invention, the processes described above in the context of the first, second and/or third major embodiments comprise a system as just described.

A seventh but by no means least major embodiment of the present invention is a set of biomarkers indicative of sleep disorders comprising or consisting of at least two
20 biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, and preferably obtained from a blood sample and especially measured with qNMR.

25 In the context of the present invention biomarkers as well as a set of biomarkers from plasma samples of individuals for diagnosis of (acute) sleep disorder that is prevalent with professions such as those associated with jet-lag and travel, shift-workers (nurses and doctors), tourism and transport sectors etc. were identified and quantified.

30 That is, in the present invention eight biomarkers in the form of metabolites from different biochemical pathways that are affected due to restrictions and deprivation of regular sleep were identified and made the basis for the respective processes, systems and employments according to the present invention.

It was found in the context of the present invention that a single biomarker of those found in the context of the present invention will often not suffice for diagnosis because variations in concentration of each molecule in itself has been implicated in one or several diseases – for that reason the present invention is directed to subject matter always on the basis of at least two of the biomarkers.

Not only is a combination of the biomarkers disclosed above best suited for diagnosis of sleep disorders but their respective concentrations at different time points can well describe the phase or the intensity of the sleep disorder.

Particularly preferred embodiments of the invention are based on all eight biomarkers. However, not always all eight biomarkers are necessary and their combinations can be selected depending on the specific circumstances. For instance, in another embodiment, out of the eight biomarkers, NAG and 3-HPA are chosen as reference biomolecules for either non-transformed or log transformed data sets.

In yet another embodiment of the invention, combinations of the biomarkers can be used in different ratios and proportions, Gaussian models and other mathematical transformations for prediction and diagnosis that can serve for regular data-analytics, machine learning and artificial intelligence-based approaches.

According to the present invention, preferably a metabolomics approach by quantitative nuclear magnetic resonance (qNMR) spectroscopy to identify biomarkers is applied. The major factor in consideration of the diagnosis according to the invention here is based on identification and quantification of the metabolites, i.e., the biomarkers. This is because metabolites (biomarkers) describe an individual's phenotype and are sensitive to various external as well as internal stimuli. Despite the strength of metabolomics for biomarker-based management of sleep-wake disorders, a major limitation that has – prior to the present invention – hampered its use for diagnostics is the lack of normative values of metabolite concentrations. Moreover, metabolic profiles can be considerably affected by various different parameters such as gender, nutrition, medication etc., to name a few. However, in the context of the present invention, the above-described biomarkers (metabolites) were identified and quantified from the mixture of biomolecules that describes the biochemical fingerprint of an individual at a particular time point. The identified, quantifiable biomarkers can thus serve as a reference for diagnostic approaches and the processes and system according to the present invention.

Advantages of the present invention are at least two-fold for the various biomarkers that were found. On one hand, the biomarkers are quantifiable as diagnostic markers and hence are EU-IVDR compliant for use at clinical levels. This can be applied in the context of shift-workers and is a very beneficial strategy to mitigate potential personal health risks. On the other hand, the qNMR metabolomics approach for diagnosis of sleep disorder is quick and cost-effective. Unlike other medical tests that were previously employed, where the subject needs to stay overnight in the clinic, in the context of the present invention, it suffices to measure one's blood sample and provide the results for doctor's analysis in few hours.

The invention is now described in more detail with reference to the following figures and non-limiting examples. The following exemplary, non-limiting examples are provided to further describe the embodiments presented herein. Those having ordinary skill in the art will appreciate that variation of these examples are possible within the scope of the invention.

The individual parts or process steps of the present invention's respective devices or processes are in operative connection with each other in a customary and known manner.

It should be noted that the various methods employed to take samples, prepare them (for further steps), measuring of samples and the like are, if not otherwise specified, usual methods known to those in the art.

It is to be noted that in preferred embodiments of the present invention, each of the methods, systems and other embodiments are not directed to the treatment of mammals of any kind.

The various embodiments of the present invention, e.g. - but not exclusively - those of the various dependent claims or individual embodiments described in the figures, may be combined with one another in any desired manner, provided such combinations do not contradict one another. In particular, the respective major embodiments can be combined. For example, the computer readable-medium according to the present invention can in preferred embodiments of the present

invention be used in the processes of the first and/or second major embodiments in order to effectively carry those out. Similarly, for example the computer-implemented process according to the present invention can in preferred embodiments of the present invention be used in the processes of the first and/or second major
5 embodiments in order to effectively carry those out. Again, similarly, the for example the system according to the present invention can in preferred embodiments of the present invention be used in the processes of the first and/or second major embodiments in order to effectively carry those out.

10 Examples:

Participants and sample collection:

Volunteers underwent a selection procedure on the basis of questionnaires pertaining to their general health status, substance abuse, sleep habits/abnormalities and
15 psychological screening followed by physical examination including blood and urine testing. Plasma samples were collected from a total of 18 healthy, non-smoking women and men with regular sleep times who were aged between 20-39 years. The study was approved by the Ethics Committee of the North Rhine Medical Board, Germany. Participants gave written informed consent prior to participation.

20

Procedure:

At the start of the experiments, during the first week the participants adhered to a sleep-wake protocol with time-in-bed (TIB) from 22:00-07:00 or 23:00-08:00 hours, which was checked with actigraphs and sleep diaries. Participants abstained from
25 consumption of alcohol and caffeine during the entire course of the study. Figure 1 provides an overview of the study design.

Participants spent one adaptation night (A) and 2 baseline nights (B) in the laboratory followed by chronic sleep restriction (CSR) to five hours TIB for 5 nights (E = examination night). Thereafter, the subjects had an eight hour recovery night (23:00-
30 07:00 or 24:00-08:00 o'clock) (R1) followed by 38 hours of acute total sleep deprivation (TSD). Participants were under constant supervision and were not allowed to perform vigorous activities.

Plasma samples were collected at four different time points early in the morning before dietary consumptions, the time points being at baseline (i.e., after the second

baseline night B2), at CSR (i.e., after the fifth examination night – E5), at recovery (i.e., after the recovery night - R1) and finally after the sleep deprivation (TSD).

Sample preparation

5 To characterize the metabolites in the plasma of the study group, 600 μ L of sample was treated with equivolume of 1:1 mixture of chloroform:methanol. Then it was mixed for 1 minute at room temperature (RT). The sample was then spun at 5,000 g for 5 minutes and the aqueous layer was separated from the chloroform phase. The aqueous layer was further treated with 600 μ L of chloroform, mixed for 1 minute at
10 room temperature (RT) followed by centrifugation at 5,000 g for 5 minutes. The two layers were separated. The chloroform layer from the first and the second treatment was pooled together. The aqueous and the apolar layers were then separately dried under vacuum using speed-vac at 25°C. The dried sample was resuspended in 600 μ L of aqueous solution containing 10 mM Na-phosphate, pH 7.4, 150 mM NaCl, 10% v/v
15 D₂O and 30 μ M trimethylsilylpropanoic acid (TSP). TSP served the purpose of an internal calibration reference for the chemical shifts as well as quantifiable marker.

NMR spectroscopy

All samples were measured on a Bruker NMR spectrometer (Bruker Biospin) with
20 proton (¹H)-frequency of 900 MHz equipped with a 5 mm probe at constant temperature of 293K (20°C) applying a standard 1D ¹H-NMR with an excitation sculpted water pre-saturation pulse. The NMR lock signal was obtained by adding D₂O. ¹H-NMR spectra were calibrated with respect to TSP chemical shift (CS) at 0 ppm. The spectra were visualized and batch processed using TopSpin NMR software (version
25 4.0.6, Bruker Biospin Ltd.).

Identification and quantification of metabolites

The ¹H-NMR spectrum contains the biomolecular fingerprint and all the necessary information about the effective hydrophilic metabolites present in the sample, which
30 were identified and quantified using Chenomx V9 (Chenomx Inc., Edmonton, AB, Canada). Prior to the identification, the NMR CS peak lists were binned from -0.04 to 10 ppm with a bin width of 0.04 ppm. Spectral regions corresponding to the water peak from 4.4 to 5.0 were omitted from the analysis.

35 Statistical analysis

To identify the relevant metabolites for sleep diagnostics data-analytics were performed using the statistical programming language R using standard functions and libraries within the CRAN repository namely lme4 and lmerTest for mixed linear models. Two sets of analysis were performed, one of which contained non-transformed data and the other one containing the same data, however, was log 2 transformed.

For the non-transformed data set, mean concentrations of the metabolites from the various participants at different time points were used directly for analysis. In the other set of analysis, the concentrations of the metabolites from the various participants at different time points were log 2 transformed followed by calculating the mean. In both cases, mixed linear model was used for the analytics, in which the mean concentration or the mean concentration from the log-transformed data of the metabolites were analysed against the various time points of sample withdrawal as the fixed effect and the individual participants as the random effect. A p-value <0.05 was used as a statistically significant indicator for the relevance of the potential biomarker.

Results

As described in the methodology above for sample preparation, metabolites present in the aqueous phase from plasma were identified and quantified. The identified biochemicals included several different forms of fatty acids (SCFA and MCFA), such as acetate, butyrate, valerate and pyruvate, citrate, phenyl acetate and many of their variants; branched chained amino acids (BCAA), unbranched chain amino acids and their respective keto-acids such as leucine, valine, isoleucine, glycine, serine, alanine, beta-alanine, serine, threonine, orotate, hippurate, threonine; nucleotides and cellular energy currencies (i.e. compounds/metabolites/biochemicals that provide energy) such as adenosine triphosphate (ATP), uracil triphosphate (UTP) and their derivatives, creatine, creatinine, carnitine, choline, in abundance, in the NMR spectral profiling. Also noticed was identification of several key vitamins and regulators comprising of water-soluble vitamins, anti-oxidants and glutathione. Among them pyridoxine (Vitamin B6) was identified and is of particular interest as it has been implicated in sleep disorders and is also used as a supplement to promote sleep and has been tested in clinical trials. The identification of pyridoxine validates the biomarker analysis. Therefore, robust statistical analysis was performed using mixed-linear model to identify and quantify the concentrations of relevant metabolites under

different phases of sleep (see Figure 1). In such a manner, a panel of biomarkers and their concentrations were identified that can be used for diagnosis of acute sleep disorders.

As a first step, the distribution of the average metabolite concentrations across the various participants was studied. A near normal distribution of the concentrations in most of the identified metabolites was observed. Hence, a linear-mixed model analysis of the data was performed. To identify the relevant metabolites for sleep diagnostics data-analytics using the statistical programming language R was performed. In non-transformed data, mean concentrations of the metabolites from the various participants at different time points were used directly for the mixed linear model analysis where the metabolites were analysed against the various time points of sample withdrawal as the fixed effect and the individual participants as the random effect. In all tested cases, a p-value <0.05 was used as a statistically significant indicator for the relevance of the potential biomarker. Using this approach, five relevant metabolites, namely 3-hydroxyphenyl acetate (3-HPA), formate, N-acetyl glutamine (NAG), pyridoxine (vitamin B6) and tyramine, were identified.

In another instance, the concentrations of the metabolites from the various participants at different time points were log 2 transformed followed by calculating the mean. The transformation was so performed as to reduce the skewness associated with metabolite distribution across various samples and time point yet not dramatically affecting the metabolite concentration from the non-transformed data. Above all, log transformation represents linear and additive models better than a non-transformed version. Alike the non-transformed data, mixed linear analysis of the data was performed and five relevant metabolites namely 3-HPA, NAG, 4-guanidinobutanoate, glycocholate and trimethyl amine N-oxide (TMAO), were identified.

A number of different biomarkers were identified in the context of the present invention. Of particular interest, and thus preferred, are eight different biomarkers that include 3-HPA, formate, NAG, pyridoxine, tyramine, 4-guanidinobutanoate, glycocholate and TMAO. As the data was near normal distributed, a rather mild two-fold log transformation of the data to avoid any potential skewness and add robustness to the linear regression was performed. This is reflected by the presence of NAG and 3-HPA in both sets of analysis. In fact, for the diagnostic purposes preferably both these molecules as reference metabolites for the different transformation-based analysis are used in the present invention.

The respective measured values (in $\mu\text{mol/L}$) for the eight above mentioned compounds are given in the following table:

		B2 [$\mu\text{mol/L}$]	E5 [$\mu\text{mol/L}$]	change vs B2 [%]	R1 [$\mu\text{mol/L}$]	change vs B2 [%]	TSD [$\mu\text{mol/L}$]	change vs B2 [%]
Pyridoxine	1st Qu.	1.95	2.05		2.3		2.7	
	Median	2.4	2.8	16,67%	2.7	12,50%	3.3	37,50%
	Mean	2.693	2.993	11,14%	2.854	5,98%	3.18	18,08%
	3rd Qu.	2.95	3.55		3.2		3.75	
3-HPA	1st Qu.	87.3	84.55		78.7		77.05	
	Median	94.7	93.3	-1,48%	85.5	-9,71%	84.1	-11,19%
	Mean	101.3	98.2	-3,06%	88.05	-13,08%	91.69	-9,49%
	3rd Qu.	121.2	107.4		102.3		105.3	
Formate	1st Qu.	49.9	40.1		59		54.9	
	Median	65.4	56.6	-13,46%	85.8	31,19%	90.2	37,92%
	Mean	79.66	55.17	-30,74%	75.51	-5,21%	79.92	0,33%
	3rd Qu.	89.85	65.7		97.1		101.25	
NAG	1st Qu.	18.85	24.8		25.1		26.45	
	Median	26.05	29.3	12,48%	27.5	5,57%	30.5	17,08%
	Mean	25.69	29.15	13,47%	27.63	7,55%	29.16	13,51%
	3rd Qu.	29.95	32.85		29.7		34.5	
Tyramine	1st Qu.	27.45	25.75		22.8		20.3	
	Median	30.9	29.5	-4,53%	26.4	-14,56%	26.8	-13,27%
	Mean	32.8	32.59	-0,64%	28.43	-13,32%	27.13	-17,29%
	3rd Qu.	38.2	40.3		31.6		31.5	
4-Guanidino- butanoate	1st Qu.	3.65	8.1		3.825		11.6	
	Median	12.6	11.85	-5,95%	10.05	-20,24%	14.1	11,90%
	Mean	11.48	11.13	-3,05%	10.742	-6,43%	15.12	31,71%
	3rd Qu.	16	14.55		17		20.3	
Glycocholate	1st Qu.	13	13.7		10		10.55	
	Median	15.9	14.9	-6,29%	13.1	-17,61%	13.3	-16,35%
	Mean	16.19	15.64	-3,40%	13.44	-16,99%	13.5	-16,62%
	3rd Qu.	20.35	19		17.7		15.35	
TMAO	1st Qu.	13.35	11.3		12.4		11.55	
	Median	15.6	14.7	-5,77%	14.1	-9,62%	14.3	-8,33%
	Mean	17.02	14.29	-16,04%	16.07	-5,58%	15.56	-8,58%

	3rd Qu.	19.65	17.25		17.8		17.95	
--	---------	-------	-------	--	------	--	-------	--

Figures:

Figure 1 shows – very simplified – the study design used in the present invention in order to illustrate when blood samples were taken. The volunteers were subjected to sleep-based experiments. Days are displayed as rows: A, adaptation; B2, baseline; E5, experimental day 5; R1 recovery; TSD is total sleep deprivation; TIB is time in bed. The time points at which the blood samples were collected for the NMR metabolomics experiments are shown as stars (see also the description above).

Figure 2 shows a graphical representation of a mixed linear model on the non-transformed data showing the most significant metabolites from plasma samples collected from various participants at different time points B2, E5, R1 and TSD (see e.g., figure1), wherein the respective compounds (3-HPA, formate, NAG, pyridoxine and tyramine) are plotted against their respective concentrations given in mM (mmol/L) in the aqueous phase from plasma of blood samples.

Figure 3 shows a graphical representation of a mixed linear model on the log 2 transformed data showing the most robust metabolites from plasma samples collected from various participants at different time points B2, E5, R1 and TSD (see e.g., figure1), wherein the respective compounds (3-HPA, 4-guanidinobutanoate, glycocholate, NAG, TMAO) are plotted against their respective concentrations given in mM (mmol/L) in the aqueous phase from plasma of blood samples.

Figures 4 to 7 show some NMR spectra prepared from blood samples of four different individuals. From these spectra the respective changes of the concentrations in relation to the different points in time, B2, E5, R1, TSD, when the respective samples were taken can easily be seen (and quantified, in particular on a connected computer).

Claims:

1. Process of diagnosing sleep disorder in a patient comprising or consisting of the following steps:
- 5
- i) measuring, in a blood sample from a patient, the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof;
- 10
- ii) comparing the measured levels for the detected biomarkers with predetermined threshold levels;
- iii) if at least two of the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, diagnosing the patient
- 15
- iiia) as having sleep disorder, or
- iiib) as being in danger of developing sleep disorder.
2. Process according to claim 1, characterized in that
- 20
- at least the following two biomarkers are measured: N-acetyl glutamine and 3-hydroxyphenyl acetate,
- or
- at least the following three biomarkers are measured: pyridoxine, N-acetyl glutamine and 3-hydroxyphenyl acetate,
- 25
- or
- at least the following four biomarkers are measured: pyridoxine, N-acetyl glutamine, 3-hydroxyphenyl acetate and trimethyl amine N-oxide.
3. Process according to any one of the preceding claims, characterized in that the biomarker levels in the sample are determined by quantitative NMR metabolomics, preferably with ¹H-NMR.
- 30
4. Process according to any one of the preceding claims, characterized in that the threshold levels are the mean values of a number of concentrations of the respective biomarkers in the aqueous phases from plasma of blood samples
- 35

determined from a number of individuals, healthy persons during normal life circumstances, preferably at least 18 individual persons.

5. Process for the determination whether a therapy against sleep disorder is successful comprising or consisting of the following steps:

- I) diagnosing sleep disorder in a patient, preferably with a process according to any one of claim 1 to 4;
- II) measuring levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in a blood sample taken from the patient at the time of diagnosis or at least prior to the start of the therapy;
- III) measuring the levels of the same biomarkers as in step II) in blood samples taken from the patient during the therapy, preferably at predetermined times; and
- IV) comparing each of the measured levels for the detected biomarkers in the respective blood samples with predetermined threshold levels, in particular, if the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%;
- V) determining whether the results of the comparison of the biomarkers change during the therapy, and thereby determining whether the therapy is effective.

6. Process according to claim 5, characterized in that the biomarker levels in the sample are determined by quantitative NMR metabolomics, preferably with ¹H-NMR.

7. Process for the detection of at least two biomarkers in a blood sample from a patient comprising or consisting of the steps of:

- obtaining a blood sample from a patient;
- detecting the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate,

trimethyl amine N-oxide and mixtures thereof, in the sample by quantitative NMR.

8. Computer-readable storage medium having computer-readable instructions stored thereon, which, when executed, cause the computer to:

- a) access data attributed to at least one blood sample from a patient, wherein the data comprises the measured levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in the respective blood sample;
- b) compare the data with pre-determined levels for the respective biomarkers, the pre-determined levels being indicative of sleep disorder; and
- c1) if the comparison reveals that the measured levels of at least two of the biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, output of an alert to the computer's user that the patient in question is diagnosed
 - as having sleep disorder, or
 - as being in danger of developing sleep disorder,or
- c2) if the comparison reveals that the measured levels of none or at most one of the biomarkers do not deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%, output of an information to the computer's user that the patient in question is not diagnosed as having sleep disorder, or, at the time the respective blood sample was taken, is not in danger of developing sleep disorder.

9. Computer-implemented process for diagnosing sleep problems in a patient comprising or consisting of the following steps:

- i) measuring, in a blood sample from a patient, the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate,

tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof;

ii) comparing the measured levels for the detected biomarkers with predetermined threshold levels;

5 iii) diagnosing the patient with sleep disorder, if at least two of the levels of the detected biomarkers deviate from the predetermined threshold values for the respective biomarkers by a pre-selected percentage, in particular by at least 3%;

10 wherein the process is carried out on a computer in which a computer-readable storage medium is present having stored thereon instructions that cause the computer to execute the process, preferably a computer-readable storage medium according to claim 8.

10. System for diagnosing sleep problems in patients comprising or consisting of:

15 A) a detection device configured to measure the levels of at least two biomarkers, preferably selected from the group consisting of pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof, in blood samples by performing quantitative NMR
20 measurements;

 B) an analysing device comprising
 - a computer, and
 - at least one computer readable storage medium according to claim 9;

25 C) a database storing pre-determined levels for the respective biomarkers, the pre-determined levels being indicative of sleep disorder.

11. Set of biomarkers indicative of sleep disorders comprising or consisting of at least two biomarkers, preferably selected from the group consisting of
30 pyridoxine, N-acetyl glutamine, glycocholate, 3-hydroxyphenyl acetate, formate, tyramine, 4-guanidinobutanoate, trimethyl amine N-oxide and mixtures thereof.

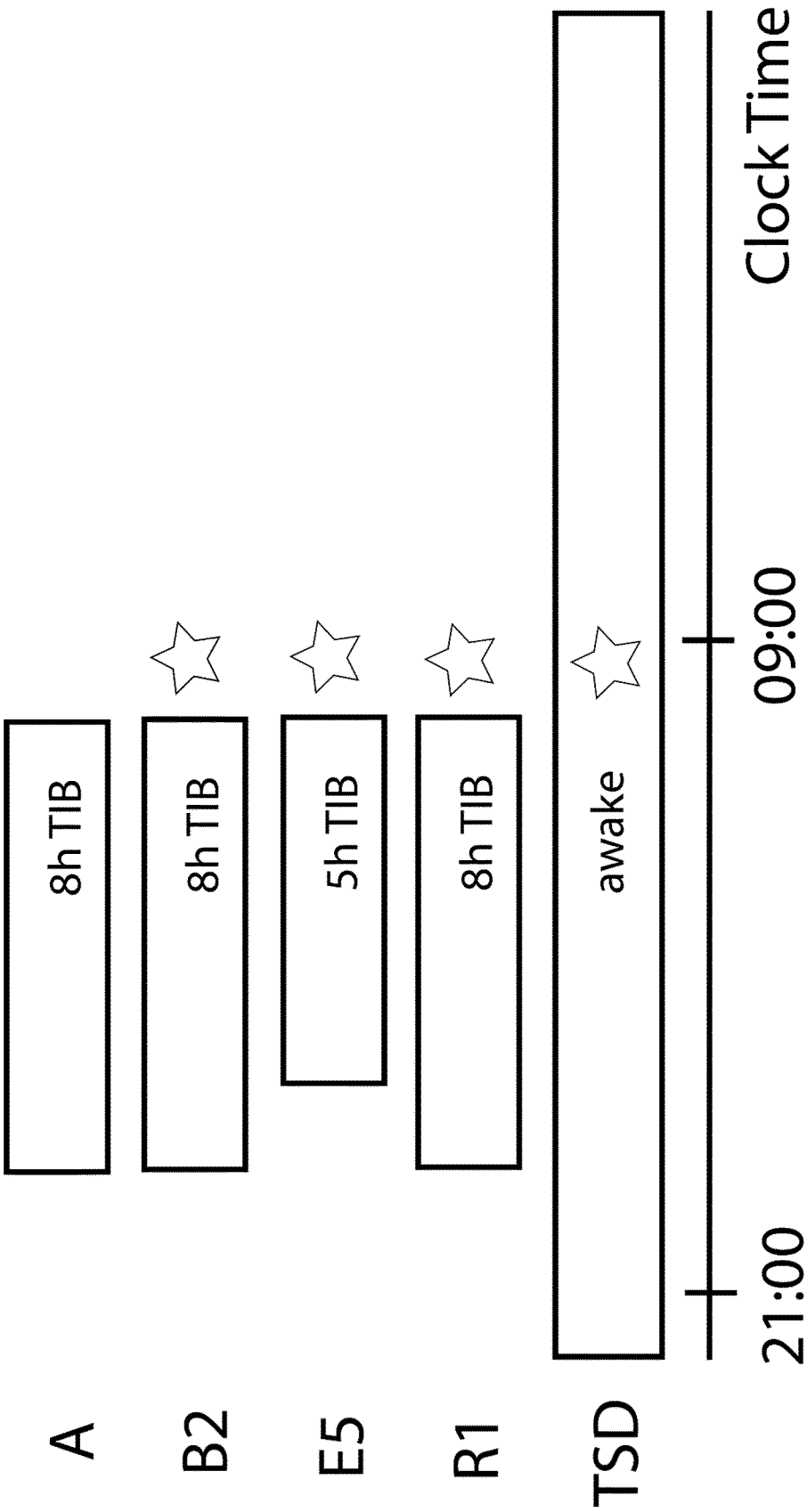


Fig. 1

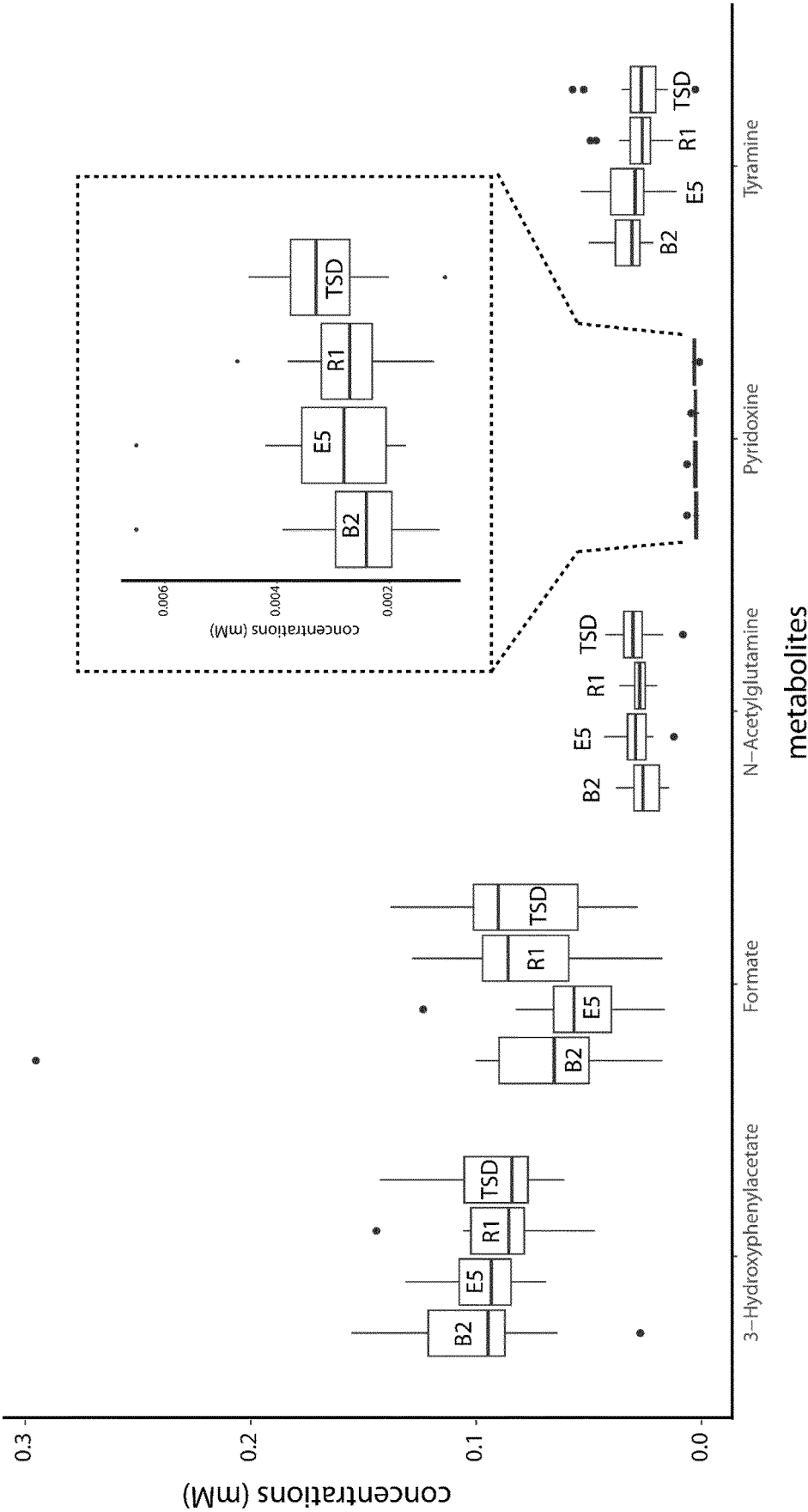


Fig. 2

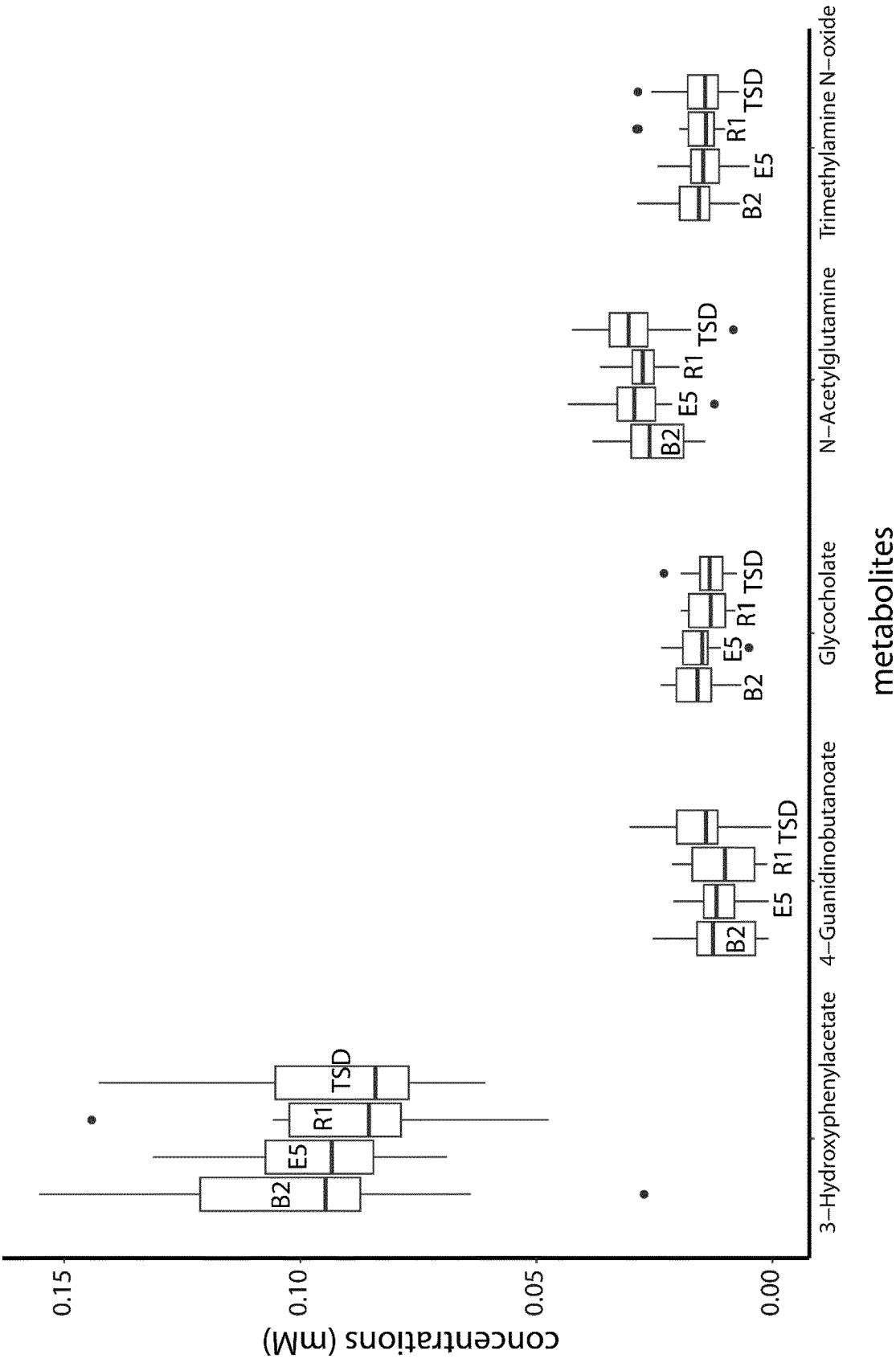


Fig. 3

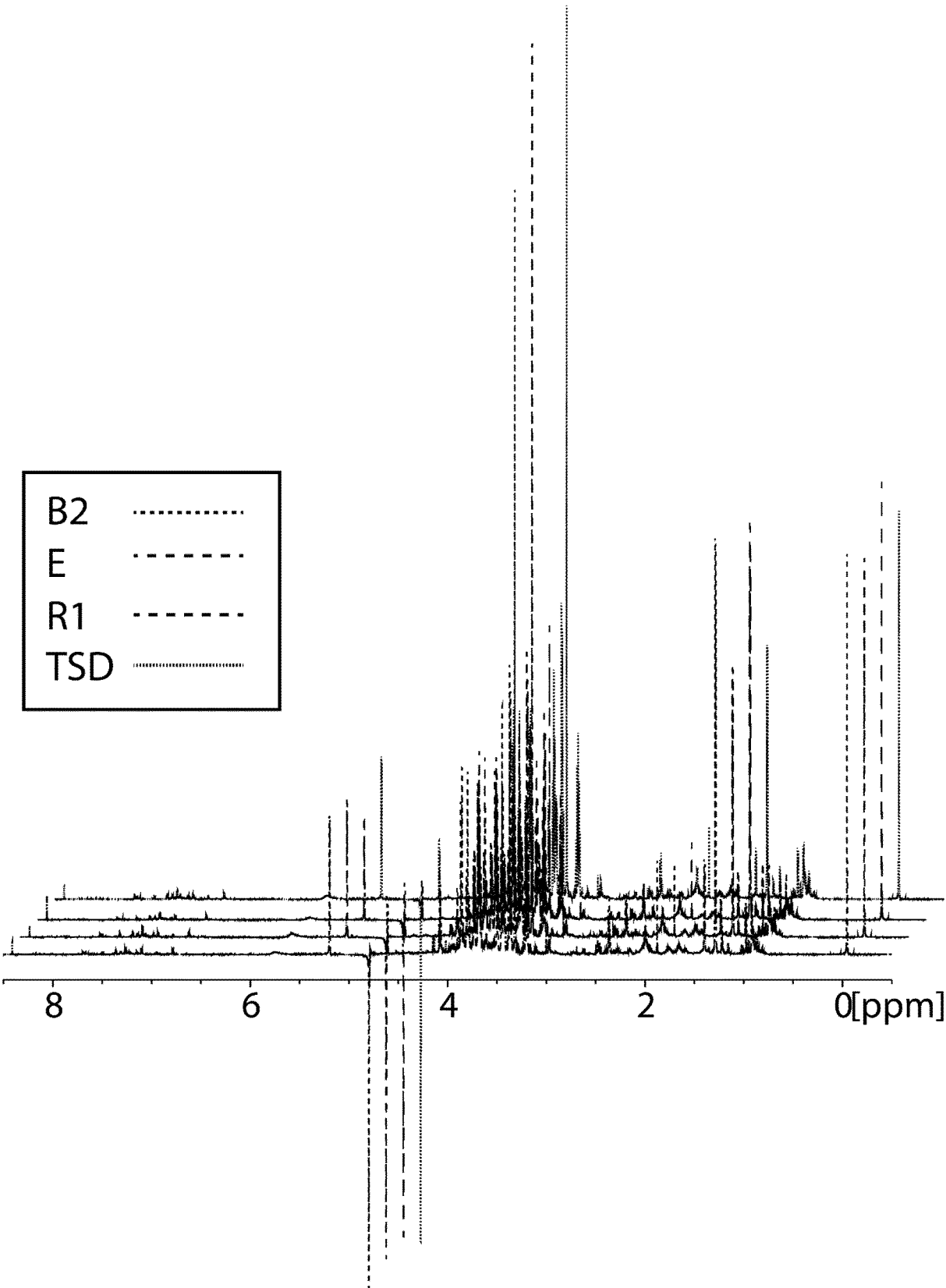


Fig. 4

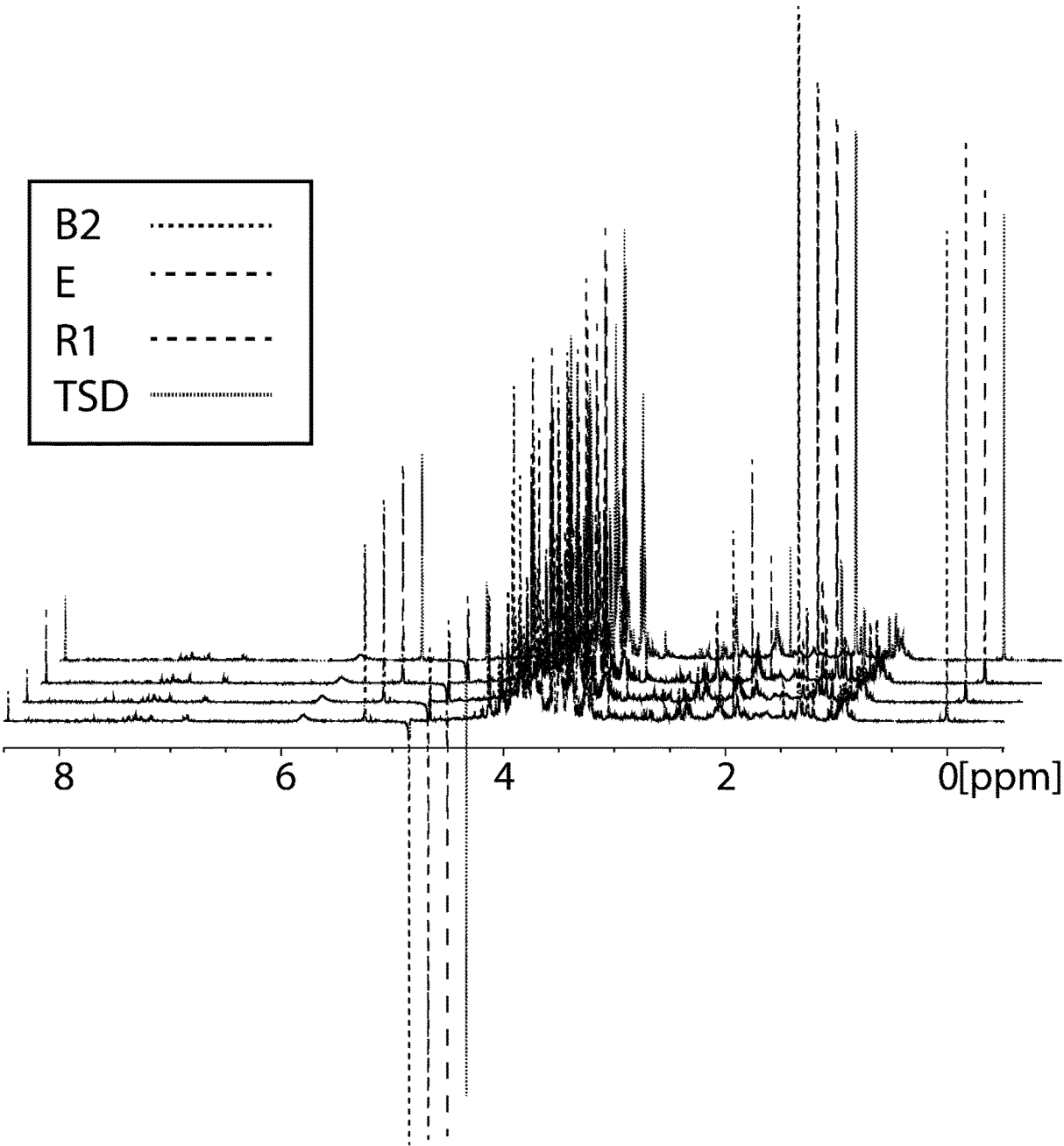


Fig. 5

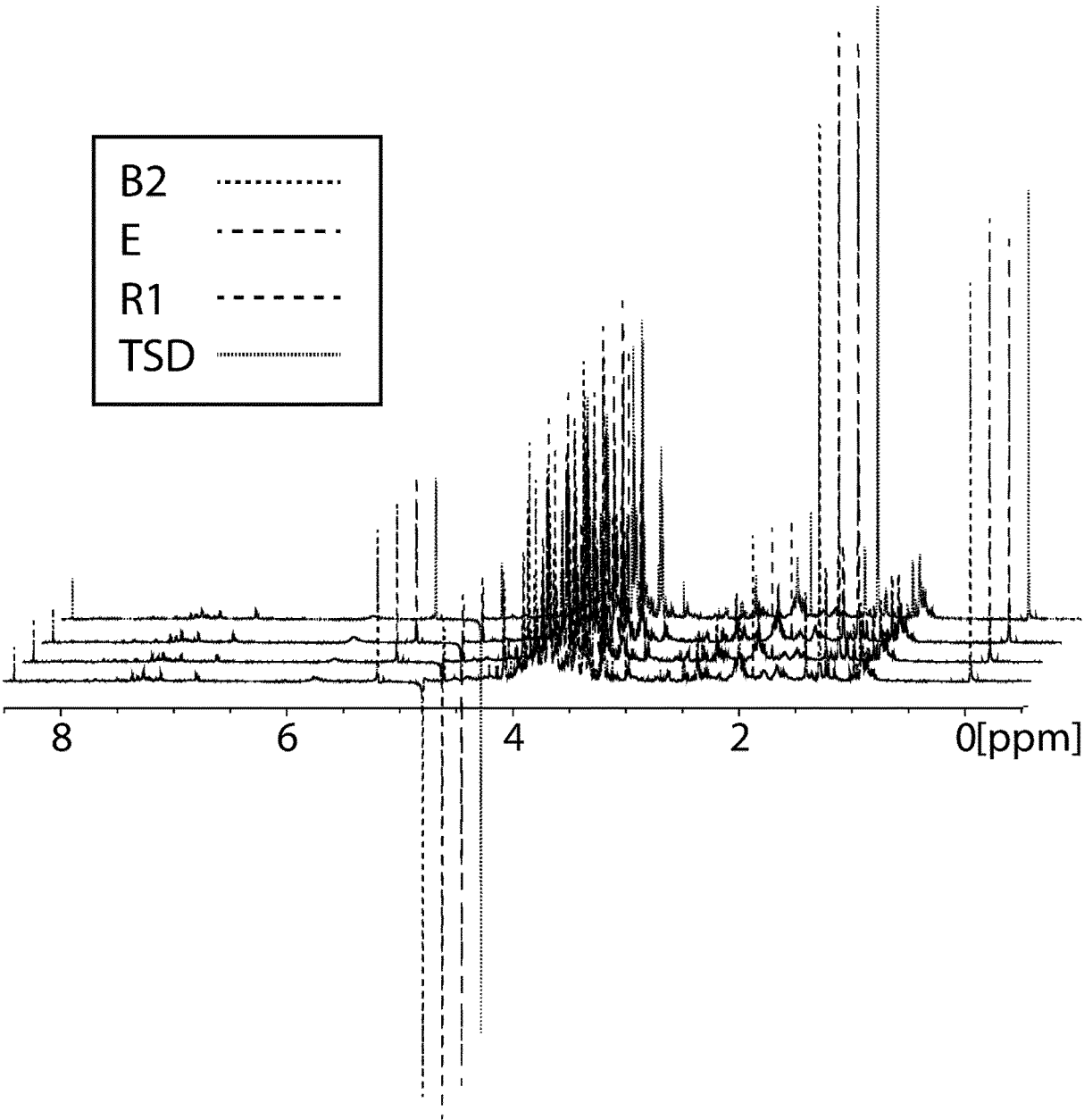


Fig. 6

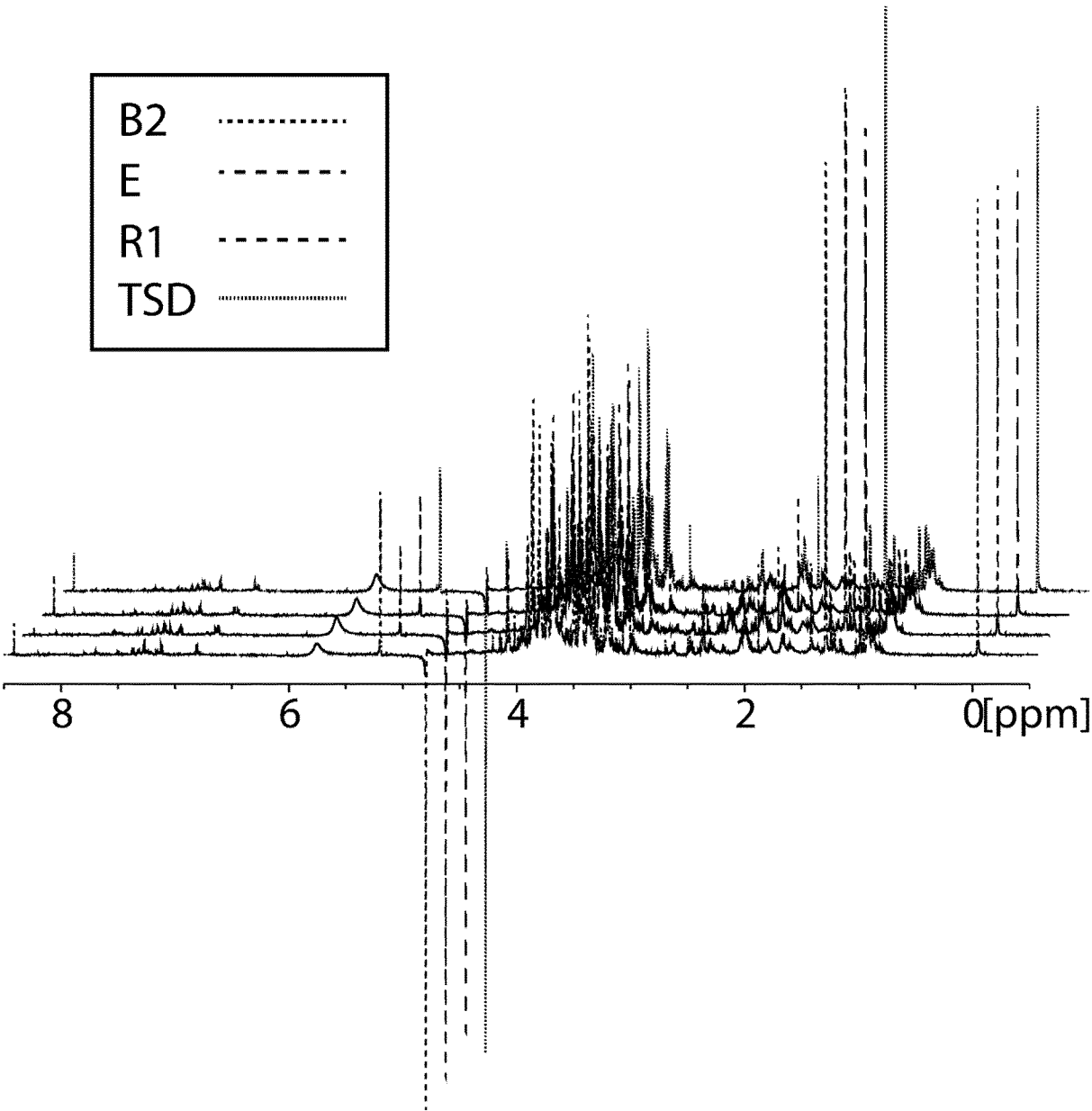


Fig. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/073111

A. CLASSIFICATION OF SUBJECT MATTER INV. G01N33/82 G01N33/84 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2017/059587 A1 (ALTERKI ABDULMOHSEN EBRAHIM [KW]) 2 March 2017 (2017-03-02) relevant for non unity; claims 1-3,6-7 -----	1
A	US 2013/101984 A1 (SHAW PAUL J [US] ET AL) 25 April 2013 (2013-04-25) relevant for non unity; paragraph [0019]; claims 9-20; example 3 -----	1
X	US 2018/156820 A1 (SOUTHWICK PAULA C [US] ET AL) 7 June 2018 (2018-06-07) claims 1-32 ----- -/-	1,3-6, 8-10
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 7 November 2024		Date of mailing of the international search report 23/01/2025
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer van der Kooij, M

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/073111

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ZAMANI ZAHRA ET AL: "A Metabolic Study on Colon Cancer Using 1 H Nuclear Magnetic Resonance Spectroscopy", BIOCHEMISTRY RESEARCH INTERNATIONAL, [Online] vol. 2014, 1 January 2014 (2014-01-01), pages 1-7, XP093009085, US ISSN: 2090-2247, DOI: 10.1155/2014/348712 Retrieved from the Internet: URL:http://downloads.hindawi.com/journals/bri/2014/348712.pdf> [retrieved on 2024-11-06] section 2.1; table 1	7,11
A	----- DEPNER CHRISTOPHER M ET AL: "Developing preliminary blood metabolomics-based biomarkers of insufficient sleep in humans", SLEEP, [Online] vol. 43, no. 7, 2 January 2020 (2020-01-02), pages 1-13, XP093221765, US ISSN: 0161-8105, DOI: 10.1093/sleep/zsz321 Retrieved from the Internet: URL:https://watermark.silverchair.com/zsz321.pdf?token=AQECAHi208BE49Ooan9kKhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAA2owggNmBgkqhkiG9w0BBwagggNXMIIDUwIBADCCA0wGCSqGSib3DQEHATAeBglnghkgBZQMEAS4wEQQMBqQz53n1UGejXLsxAgEQgIIDHZmQRTbFtqdFZzfy1bgmxkqd_i7y8ObXhUz_vowAybrgCekGTVHpnHvxHeWk8ArpMplDeTGWdH8A8Eh7coFfgeNzfVxAs> [retrieved on 2024-11-07] supplementary Table S2 & Depner Christopher M ET AL: "Supplementary Material - Developing preliminary blood metabolomics-based biomarkers of insufficient sleep in humans Authors and Author Affiliations", Sleep, volume 43, Issue 7, 2 January 2020 (2020-01-02), XP093221775, Table S2 -----	1-11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2024/073111

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:
1-11 (partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-11(partially)

Process of diagnosing sleep disorder in a patient based on the measurement of at least 2 biomarkers comprising at least pyridoxine;

Process for the determination whether a therapy against sleep disorder is successful based on said combination, process of detection of said combination; computer-readable storage medium; computer-implemented process; system based on said combination of biomarkers and the set of biomarkers per se..

2-8. claims: 1-11(partially)

Process of diagnosing sleep disorder in a patient based on the measurement of at least 2 biomarkers comprising at least N-acetyl glutamine (invention 2)up to..... trimethyl amine N-oxide (invention 8), in as far as mentioned in the corresponding claims.

Process for the determination whether a therapy against sleep disorder is successful based on said combination, process of detection of said combination; computer-readable storage medium; computer-implemented process; system based on said combination of biomarkers and the set of biomarkers per se..

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2024/073111

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017059587 A1	02-03-2017	NONE	
US 2013101984 A1	25-04-2013	NONE	
US 2018156820 A1	07-06-2018	EP 3194981 A1	26-07-2017
		US 2018156820 A1	07-06-2018
		WO 2016197053 A1	08-12-2016

Toward precision diagnostics for circadian disruption: qNMR metabolomics and metabolic dysregulation

Abhishek Cukkemane^{j,k,*}, Kevin Bochinsky^{j,l}, Denise Lange^{a,b}, Eva Hennecke^a, Anna L. Foerges, Simone Beer, Andreas Bauer^{f,g}, Daniel Aeschbach^{a,h,i}, Dieter Willbold^{j,k}, Gabi Kastenmüller^m, Alexander Drzezga, Jens Jordan, David Elmenhorst^{c,1}, Eva-Maria Elmenhorst^{a,b,1,*}

a. Department of Sleep and Human Factors Research, Institute of Aerospace Medicine, German Aerospace Center, 51170 Cologne, Germany

b. Institute of Occupational, Social, and Environmental Medicine, Medical Faculty, RWTH Aachen University, 52074 Aachen, Germany

c. Institute for Translational Psychiatry, University of Münster, 48149 Münster, Germany

d. Institute of Neuroscience and Medicine (INM-5), Forschungszentrum Jülich, 52425 Jülich, Germany

e. Institute of Radiochemistry and Experimental Molecular Imaging, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Straße 62, 50937 Cologne, Germany

f. Institute of Neuroscience and Medicine (INM-2), Forschungszentrum Jülich, 52425 Jülich, Germany

g. Department of Neurology, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

h. Institute of Experimental Epileptology and Cognition Research, University of Bonn, Medical Center, 53127 Bonn, Germany

i. Division of Sleep Medicine, Harvard Medical School, Sleep Division, Boston 02115, MA, USA

j. Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, 52425 Jülich, Germany.

k. Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

l. current address -Bruker Biospin, Ettlingen, Germany

m. Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, Neuherberg, Germany.

1. A.C, E.-M.E. and D.E. contributed equally to this work

* Corresponding Author: Eva-Maria Elmenhorst, MD
DLR-Institute of Aerospace Medicine,
51170 Cologne, Germany

Phone: +49 2203 601 4735

E-Mail: eva-maria.elmenhorst@dlr.de

Abhishek Cukkemane, PhD
Forschungszentrum Jülich,
52428 Jülich, Germany

Phone: +49 2461 61 9448

E-Mail: a.cukkemane@fz-juelich.de

Keywords

Chronic sleep restriction, human, sleep, quantitation, metabolomics, qNMR spectroscopy

Abstract

Chronic mental illnesses (CMIs), such as depression and schizophrenia, represent a significant global health challenge, with many patients experiencing insufficient symptom relief. Sleep and circadian rhythm disruptions are increasingly recognized as core features, often preceding and perpetuating psychiatric episodes. Modern lifestyles exacerbate these disturbances, leading to widespread sleep deprivation and increased risks for metabolic, cardiovascular, and neurodegenerative diseases. Despite their critical impact, sleep disorders remain underdiagnosed due to limitations of current diagnostic methods.

Metabolomics presents a promising approach to identify the biochemical disruptions linked to both CMIs and sleep disorders, revealing distinct metabolic signatures. Recent advancements in quantitative Nuclear Magnetic Resonance (qNMR) spectroscopy offer a powerful, non-destructive tool for metabolomics research. This study introduces a qNMR-based metabolomic strategy for the detection and monitoring of sleep-wake disorders. We identify and validate a panel of up to 30 small-molecule biomarkers in blood plasma, encompassing essential vitamins, co-factors, energy metabolites, and gut microbiota-derived molecules. This biomarker-driven diagnostic process using qNMR metabolomics offers a robust method for detecting and monitoring sleep disorders. We perceive integrating these findings into machine learning and artificial intelligence approaches can open new possibilities for scalable and automated diagnostics. This research aims to bridge the gap between clinical need and current limitations, offering a more accessible, cost-effective, and efficient alternative to traditional sleep diagnostics.

Significance Statement

In conclusion, we propose a metabolomics-based diagnostic process for circadian sleep disorders involving qNMR metabolomics of specific plasma biomarkers. Deviations in a panel of 20 – 30 small-molecule biomarkers are indicative of the presence or risk of sleep dysfunction. While on one hand this approach offers a promising, non-invasive tool for diagnosis, monitoring, and potentially guiding therapy for sleep-related conditions; on the other hand, it also highlights to the potential of implementing robust predictive protocols using AI and/or ML.

Introduction

Chronic mental illnesses (CMIs) including depression, bipolar disorder, schizophrenia, and anxiety disorders represent a significant global health burden (van Os & Kapur, 2009). These conditions affect hundreds of millions worldwide, causing substantial suffering and reduced quality of life (Cukkemane et al., 2025). Despite extensive research into their genetic, neurochemical, and environmental origins, many patients continue to experience inadequate symptom relief, highlighting the need for more

comprehensive understanding of these conditions. Sleep and circadian rhythm disruptions (Sharma & Kavuru, 2010) stand out as fundamental features of CMIs rather than mere secondary symptoms (Edgar & McClung, 2013; Humer et al., 2020; Li et al., 2022; McClung, 2013). These disturbances range from insomnia to irregular sleep-wake patterns often precede and perpetuate psychiatric episodes. Recent chronobiology research reveals that circadian mechanisms extend beyond the brain's central pacemaker to influence hormone secretion, metabolism, immunity, and cellular repair throughout the body. This positions circadian dysfunction as a major factor in mental illness development rather than just a symptom, creating a concerning feedback loop where psychiatric conditions and circadian disruption mutually reinforce each other.

Meanwhile, sleep disorders themselves including insomnia, sleep apnea, and delayed sleep phase disorder are increasingly recognized as serious health concerns affecting cognitive function, emotional regulation, and physical wellbeing. Modern lifestyle factors such as artificial light exposure (Taillard et al., 2021), shift work (Fink, 2024; Shah et al., 2022; Wickwire et al., 2017), and irregular schedules (Olson et al., 2009; Trammer et al., 2024) further compromise the delicate balance between internal biological rhythms and environmental cues. Epidemiological evidence connects chronic sleep deprivation with increased risks for metabolic, cardiovascular, and neurodegenerative diseases. Alarmingly, average sleep duration has declined dramatically in modern times, with many adults regularly getting less than six hours nightly (Sharma & Kavuru, 2010). Current CDC data (Prevention, 2024) show concerning rates of sleep problems across all age groups, constituting a major public health challenge that intersects with epidemics of mood disorders and metabolic dysfunction.

Despite this critical underlying concern, sleep disorders remain underdiagnosed. Polysomnography (Riemann et al., 2023), while is the golden standard, has limited accessibility due to resource requirements. Encouragingly, recent diagnostic framework revisions now recognize sleep-wake disorders as independent clinical entities requiring dedicated attention (Riemann et al., 2022). Recent advances in Nuclear Magnetic Resonance (NMR) spectroscopy to the forefront of metabolomics research (Wishart, 2019; Wishart et al., 2007), with applications extending across diverse domains, from food biotechnology to clinical diagnostics (Julkunen et al., 2023; Xu et al., 2023). Metabolomics research has emerged as a promising approach to understand connections between various diseases and disorders. By analysing small-molecule metabolites, researchers can identify biochemical disruptions associated with both CMIs and sleep disorders. Among the various modalities of NMR, quantitative NMR (qNMR) spectroscopy (Vignoli et al., 2019) offers distinct advantages that make it particularly suitable for biomedical applications. These advantages include its non-destructive nature, minimal sample preparation, the ability to analyze a wide range of biological fluids such as blood and urine, and high reproducibility. Furthermore, qNMR is highly amenable to automation and integration with digital platforms, enhancing its potential for high-throughput clinical workflows. Studies have found distinct metabolic signatures in various sleep conditions altered branched-chain amino acid (BCAA) metabolism in insomnia and disrupted lipid profiles in sleep apnea (Gehrman et al., 2018; Humer et al., 2020). These

findings suggest sleep disorders have specific biochemical footprints that may explain their overlap with psychiatric conditions.

In this study, we present a qNMR-based metabolomic strategy aimed at the detection and monitoring of sleep-wake disorders, with particular focus on its applicability as a diagnostic tool. The principal strength of this approach lies in its capacity to simultaneously identify and quantify specific metabolic biomarkers directly from minimally processed blood samples. This enables a detailed molecular snapshot of an individual's physiological state, facilitating the recognition of metabolic patterns associated with sleep disruption. This approach may ultimately lead to more precise diagnostics and personalized treatments for those experiencing disorders in sleep, metabolism, and mental health.

Methodology

Study group 1

Participants. The study secured approval from the Ethics Committee of the North Rhine Medical Board (identifier 2015226) and the German Federal Office for Radiation Protection (identifier Z 5-22461/2-2015-006). All participants provided written informed consent and the study was prospectively registered with the German Clinical Trials Register (DRKS 00010194).

We rigorously screened participants' physical and psychological health using questionnaires and medical examinations. Only healthy, non-shift working individuals with normal sleep patterns were included. To ensure this, substance abuse and smoking were exclusionary criteria, verified by urinary tests at the study's onset. During the initial laboratory night, we measured the periodic leg movement index (PLMI) and the apnea-hypopnea index (AHI). Ideally, PLMI shouldn't exceed 15 and AHI shouldn't exceed 10. We made one exception for a participant with an AHI of 16.3, as they exhibited no respiration-associated cortical arousal and normal daytime sleepiness. Ultimately, 15 out of 21 participants yielded complete datasets for metabolomics analysis, comprising 6 females with an average age of 26.2 ± 4.0 years, a BMI of 22.9 ± 1.8 , an AHI of 2.2 ± 4.1 , a PLMI of 0.8 ± 1.3 , and an Epworth sleepiness scale score of 4.6 ± 2.1 .

Study design. Our study design, previously detailed in (Hennecke et al., 2021) involved a rigorous sleep-wake protocol (Figure 1). Briefly, participants adhered to a consistent 9-hour sleep schedule for one week prior to the study, aligning with their natural sleep preferences (either 22.00 to 08.00 H or 23.00 to 09.00 H). This was verified using actigraphy and sleep diaries. During this pre-study phase, alcohol and caffeine intake were strictly prohibited.

Upon entering the sleep laboratory, participants underwent an adaptation night with 8 hours in bed (23.00 to 07.00 H or 00.00 to 08.00 H) to screen for any pre-existing sleep disorders. This was followed by a sequence of experimental nights: two baseline nights (23.00 H/00.00 H to 07.00 H/08.00 H), five experimental nights with reduced time in bed (5 hours, from 02.00 H/03.00 H to 07.00 H/08.00 H), one recovery night (23.00 H/00.00 H to 07.00 H/08.00 H), one night of total sleep deprivation, and a second recovery night (21.00 H/22.00 H to 07.00 H/08.00 H). Polysomnography was recorded during all sleep

opportunities. Two participants followed the protocol simultaneously, with constant staff supervision during wake periods, and only non-activating activities were permitted.

Study group 2:

All study procedures received approval by the same ethics committee and Federal Office as stated above, aligning with the Declaration of Helsinki. Participants provided written informed consent before enrollment, and the study was prospectively registered on January 19, 2022, with the German Clinical Trials Register (DRKS00027867).

We enrolled 40 healthy volunteers (14 females; mean age 27.5 ± 6.5 years, range 20–45). Exclusion criteria included chronic neurological or psychiatric disorders, head trauma, sleep disorders, alcohol or drug abuse, smoking, pregnancy or breastfeeding, current medication (excluding contraceptives), and regular caffeine consumption exceeding 500 mg/day. We verified these criteria through urinary tests for various substances, including Zolpidem, Propoxyphene, Cotinine, Amphetamine, Methamphetamine, Morphine, Methadone, Phencyclidine, Nortriptyline, Tetrahydrocannabinol, Secobarbital, Oxazepam, Benzoylecgonine (Cocaine), and human chorionic gonadotropin.

Study design. The study began with a seven-day ambulatory sleep satiation period at home, during which participants-maintained caffeine abstinence and adhered to their habitual sleep times. This was followed by an adaptation and baseline night in the laboratory, where sleep EEG was recorded. On the first day, participants underwent baseline psychomotor and sleepiness testing, alongside a PET measurement of synaptic density (not reported here). Two participants were studied concurrently at any given time.

Control participants were discharged after initial measurements but wore an actigraph. Upon their return the next day, actigraphy data confirmed the absence of daytime napping before a second set of measurements. Sleep-deprived participants remained in the laboratory between sessions, engaging in non-strenuous activities like conversations, video viewing, quiet games, reading, or internet browsing. To ensure wakefulness, participants were continuously monitored, and prohibited from closing their eyes for more than a typical blink, with prompt intervention by staff if necessary. Two samples were collected for each participant, the initial sample will be referred to as time point 1 and the latter sample as time point 2.

Sample Preparation

To analyze plasma metabolites, we first extracted them using a modified Bligh and Dyer (Bligh & Dyer, 1959) method. We combined 600 μL of each plasma sample with an equal volume of a 1:1 chloroform:methanol solution. This mixture was then vortex-mixed for one minute at room temperature to facilitate efficient extraction.

Following vortexing, the samples underwent centrifugation at 10,000 g to achieve distinct phase separation. The aqueous (upper) layer was then carefully collected and mixed with an equal volume of

chloroform. This mixture was again vortex-mixed for one minute at room temperature and centrifuged at 10,000 g for a second phase separation.

The purified aqueous layer was then meticulously collected and dried under vacuum using a SpeedVac concentrator set at 25 °C. Finally, the resulting dried extract was reconstituted in 660 µL of an aqueous buffer. This buffer contained 10 mM sodium phosphate (pH 7.4), 150 mM sodium chloride, 10% (v/v) deuterium oxide (D₂O), and 200 µM trimethylsilylpropanoic acid (TSP). TSP served as our internal reference for accurate chemical shift calibration and quantification during subsequent analyses.

NMR spectroscopy, data processing and metabolic profiling

We conducted NMR measurements using a Bruker Avance 900 MHz spectrometer equipped with a 5 mm probe. Spectra were acquired at a constant temperature of 293 K (20 °C) using a standard one-dimensional (1D) ¹H-NMR sequence with water pre-saturation (square-shaped pulse). Deuterium oxide (D₂O) provided the field-frequency lock signal, and chemical shifts (CS) were calibrated relative to the TSP peak at 0 ppm. All spectra were visualized and batch-processed using TopSpin software (version 4.0.6, Bruker BioSpin Ltd.).

The ¹H-NMR spectra acted as biochemical fingerprints for the metabolites in our samples. Before annotating chemical shifts, we performed extensive spectroscopic data pre-processing. This included Fourier transformation, phasing correction, baseline correction (using Chenomx spline), water region deletion, shim correction (line width of 1.2 Hz), apodisation (line broadening of 1Hz), and referencing control to TSP. Metabolite identification and quantification were then performed using the batch processing and legacy auto-fitting tool in Chenomx pro v10 Suite software (Chenomx Inc., Edmonton, Canada). Chemical shift annotations were made by comparing our spectra with the metabolite library containing 330 metabolites for untargeted metabolomics and 5 metabolites for targeted metabolomics that are available within the software.

Statistical Analysis

To identify metabolites relevant for sleep diagnostics, we performed statistical analyses in R, primarily utilizing pairwise t-tests and mixed linear modelling (from the lme4/lmerTest packages). Before analysis, we carefully curated the data for outliers. To minimize false-positive associations in the dataset, we only included metabolic traits with at least 300 non-missing values. Any identified outlier data were removed from subsequent analyses.

We performed log₂-transformed metabolite concentrations and then calculated their mean values. In our mixed linear models, time points of sample collection were treated as fixed effects. Participant age and sex were included as covariates, while individual participants were treated as random effects to account for within-subject variability. We set statistical significance at a p-value < 0.05 to identify potential biomarkers of interest.

Results

Metabolic fingerprinting of biomarkers associated with sleep disorders

We performed 1D ^1H -NMR metabolomics in an untargeted fashion to profile various metabolites in the processed blood samples of the volunteers in the study. Although, one of the major advantages of NMR spectroscopy lies in its ability in performing studies without altering the sample. We chose to perform organic phase separation using Bligh-Dyer (Bligh & Dyer, 1959) approach in removing lipophilic components from the sample. We deemed this as a necessary approach in our study because blood sample contains abundant number of large molecules in the form of proteins and lipid vesicles that leads to line-broadening effects, thereby resulting in lower resolution in spectrum. Following NMR experiments, using automated annotation of chemical shifts using Chenomx containing a library of 330 metabolites, we fingerprinted every sample by identifying and determining the concentrations of all the quantifiable metabolites from the sample. Thereafter, in order to comprehend the biochemical perspective of the acute sleep disorder from the annotated metabolites, we employed linear mixed regression analysis. Such an analysis is known to improve the quality of clinical metabolomics data described herein by biological variations of individuals. An individual parameter in the model is represented by fixed effect such as the metabolite concentrations challenged against longitudinal sampling along with covariates such as sex, BMI and age, random effect is represented by the participant themselves so as to separate between-subject variation from within-subject variation. Based on the P-values set to a significance level of $p < 0.05$, we identified five prospective metabolites (Figure 2, Figure S1 (concentration normalized plot) and Table 1), them being N-acetyl glutamine (NAG), Glycocholate, 3-hydroxyphenyl acetate (3-HPA), 4-Guanidobuturate, and Trimethylamine N-oxide (TMAO). These set of metabolites points to pathways associated to liver metabolism, gut microbiome interaction, and neurotransmitter and nitrogen cycling. Correction for multiple-testing yielded no significant findings therefore, we performed pairwise t-test as post-hoc analysis. Figure 3 (also see Figure 4, S2-S6 and Table 2) shows the trend of the metabolites over the course of the experiment in that, most of the metabolites show a progressively linear increase or decrease in the range of 10-30% (Table 1) of the metabolite concentration with Glycocholate displaying the largest concentration difference and TMAO the smallest between the baseline and sleep deprived conditions.

Post-hoc analysis

In the succeeding step, in order to retrospectively analysis the differences between the sampling points, we performed post-hoc analysis to gain further insights. It is because, measurements were performed on the samples from the same individuals (see Figure 1) at baseline (B), sleep restricted (E5), recovery (R), and sleep deprived (TSD) conditions, we employed pairwise t-test for participant going from time point B to E5, B to R, B to TSD etc revealed plenty of interesting hits that not only supports the findings of mixed-linear model but goes beyond in describing further significant metabolites.

As a first, the post-hoc analysis reveals that Glycocholate is a significant metabolite when test volunteer is progressing from E5 to R and TSD; TMAO is a significant metabolite moving from baseline to recovery; N-acetyl glutamine is a crucial metabolite when volunteer in baseline condition is sleep deprived; 3-Hydroxyphenyl acetate is a significant difference when comparing baseline and recovered phase; 4-Guanidiobutanoate is of significance when the “recovered” participant is sleep deprived.

In the next step, to understand the finer details of the longitudinal study, we compared baseline condition to the other time points (see Figures 3,4 and S2-S6). While B and TSD are well defined start- and end-points, E5 and R may represent myriad states of the participants based on the health and metabolic activities. To begin with, we constructed volcano plots (Figure 3) to identify the metabolites that are significantly affected in the longitudinal study. When comparing B vs E5 (Figure 4, Table 2), we observe a small number of significantly altered metabolites, for instance based on their concentrations, downregulation of TMAO (Figure 4), N6-Acetyllysine, to name a few (for full list of metabolites see Table 2) and a notably high upregulation in biochemicals like O-Phosphoethanolamine, Glutamine etc. When comparing baseline to “recovered” state (Figure S2, Table 2) we observe that metabolites like D-Threitol, and 2-Methylmalonate, are significantly upregulated, while others like 3-Hydroxyisobutyrate and 3-Hydroxyphenylacetate are downregulated. These findings beg to address the question, is recovered state, really recovered to baseline? Further analysis in the longitudinal direction reveals that the differences arising due to sleep deprivation (Figure S3, Table 2) include marked changes in the concentrations of biomolecules like L-Acetylglutamine, 2-Deoxyuridine, and Pyridoxine hinting towards metabolic differences in vitamin and nucleotide pathways. Continuing further in the grey zone of sleep restricted and deprived conditions (Figure S5, Table 2), we notice that energy currencies and anabolic metabolites like N-Acetylglutamine, 3-Hydroxyisobutyrate, and Histidine, among others are affected. Interestingly, when comparing sleep restricted and recovery (Figure S4, Table 2) metabolites influenced by gut microbiota like Glycocholate, and Formate, in addition to important co-regulatory metabolites like 4-Carboxyglutamate, and 2-Deoxyuridine. In the last set of comparative analysis between transition from recovered to sleep deprived conditions (Figure S7, Table 2), we notice severe disruption in catabolic pathways associated with sugars and polyols like D-Threitol and 1,3-Dihydroxyacetone.

Targetted metabolomics

We investigated whether our biomarker panel could be useful for monitoring therapeutic outcomes. To do this, we performed targeted metabolomics by deconvoluting the spectra obtained on samples from Study 2, quantifying the concentrations of only five specific metabolites mentioned above: N-acetylglucosamine (NAG), Glycocholate, 3-hydroxyisobutyrate (3-HPA), 4-Guanidinobutanoate, and Trimethylamine N-oxide (TMAO). Study 2 was particularly advantageous for this validation as it involved participants in their regular daily lives, allowing us to assess the real-world applicability of our initial findings.

Of the five metabolites, only 4-Guanidinobutanoate, 3-HPA, and TMAO were present in detectable amounts using NMR spectroscopy. Subsequent linear model testing, as described previously, revealed that only TMAO ($p = 0.044$) showed a statistically significant association. However, it's important to note that after correcting for multiple testing, no significant hits were observed.

While the single significant finding for TMAO in a real-life setting is promising, suggesting the potential to translate controlled sleep research to everyday situations, the limited number of significant biomarkers prompted a second set of post-hoc analyses. We then conducted untargeted metabolomics by deconvoluting the spectra of all participants in subject group 2 with all the 330 metabolites in the library, identifying and quantifying all metabolites from the samples, followed by both linear and pairwise t-tests to explore further associations.

Biomarker combinations and diagnostic utility (

Our initial comparative analysis between subjects and controls in Study 2 revealed significant variability among individual biomarkers. We identified fourteen metabolites (Figure 5 and S7): 1,3-Dimethylurate, 2-Aminoadipate, 2-Oxoisocaproate, 3-Hydroxyisobutyrate, 5-Aminolevulinate, Glucose, Isobutyrate, Lactose, Methylsuccinate, N-Nitrosodimethylamine, Propylene glycol, Pyruvate, Sarcosine, and Tryptophan. Many of these molecules are involved in mitochondrial metabolism (energy currencies), amino acid degradation, or are diet-related compounds. While this set of metabolites is broader than those found in Study 1, a consistent theme emerged: the involvement of nitrogen metabolism and the impacts of diet mediated by the gut microbiota.

Further refinement using pairwise t-tests (Figure S8) provided deeper insights into the distribution and concentrations of these samples. This analysis also uncovered additional significant metabolites, including: 2-Hydroxybutyrate, 3-Phenyllactate, 3-Chlorotyrosine, 4-Aminohippurate, Arabinitol, 5-Hydroxyindoleacetate (5-HIAA), Asparagine, Cystine, Guanidinoacetate, Galactitol, Homogentisate Phenylalanine, and Uracil. These metabolites largely stem from amino acid metabolism, Neurotransmitter pathways, Oxidative stress and inflammation, and metabolism in gut.

Discussion

This study provides insights into the metabolic alterations associated with sleep disorders, highlighting a targeted panel of circulating biomarkers in blood plasma using qNMR metabolomics. By focusing on specific metabolites, we demonstrate that variations in their plasma concentrations may serve as reliable indicators (Gehrman et al., 2018; Giskeodegard et al., 2015; Humer et al., 2020; Scholz et al., 2024; Sharma & Kavuru, 2010) of sleep-related pathophysiology.

Our findings reveal disruptions across multiple metabolic pathways. The altered metabolic landscape encompasses gut microbial metabolism, energy production pathways, amino acid metabolism (including neuro-regulators), vitamin synthesis, and nucleotide biosynthesis. Longitudinal comparison across participants leading a regular lifestyle further highlighted the influence of gut microbiota and vitamin

synthesis while also revealing shifts in carbohydrate metabolic pathways related to sugar alcohol production and altered energy metabolism under oxidative stress.

Longitudinal Metabolomic Findings from the Current Study

In our longitudinal analysis, significant alterations in metabolite levels were identified across six pairwise comparisons between baseline (B2), end of experiment (E5), recovery (R), and total sleep deprivation (TSD) conditions (Table 2). Many small-chained fatty acids (SCFA) particularly butyrate derivatives were repeatedly modulated, showing significant changes across different time points, highlighting its sensitivity to circadian rhythms. Additional metabolites frequently observed across comparisons included D-threitol, Glycerate, N-Carbamoylaspartate, and 2'-Deoxyuridine. Metabolites such as trimethylamine N-oxide (TMAO), methanol, and methionine were also differentially expressed under varied conditions, suggesting an underlying perturbation of energy and amino acid metabolic pathways in response to both circadian and sleep-driven factors.

Limitations and methodological challenges

Several methodological limitations warrant acknowledgment. While metabolic profiling offers considerable strength in constructing biochemical narratives through metabolite identification and pathway mapping, the chemical shift annotation process in NMR spectroscopy presents inherent challenges (Vignoli et al., 2019). The deconvolution of one-dimensional NMR spectra to identify and quantify metabolites can introduce errors due to overlapping chemical shifts between structurally similar molecules. Additionally, our relatively modest sample size precluded the application of multiple correction tests, which would have introduced excessive stringency in p-value adjustment potentially obscuring true associations. To address this limitation, we complemented our primary analysis with pairwise t-tests to provide additional statistical validation.

A significant limitation in our approach lies in the sample preparation methodology. The organic phase separation technique employed removes lipids and cholesterol contained within high-density and low-density lipoproteins (HDL and LDL). These compounds represent important risk factors in numerous neurological disorders, and their exclusion from our analysis represents a notable gap. Future studies would benefit from employing complementary analytical techniques that retain these lipophilic components while maintaining spectral resolution.

Comparison with previous findings

Our studies demonstrate that an individual's sleep state exerts significant influences on systemic metabolite levels as has also been previously reported (Scholz et al., 2024). The consistency of these observations across studies is particularly notable, especially for metabolites such as 3-Hydroxyisobutyrate and TMAO, which consistently emerge as perturbed. We observed significant alterations in the concentrations of Methionine, N-Acetylglutamine, Asparagine, and Methanol. These shifts collectively indicate a modulation of both amino acid and energy metabolism in response to acute

sleep deprivation and circadian misalignment, consistent with previous reports (Humer et al., 2020). Furthermore, the strongest congruence between our data with previous findings were noted to clinical studies with people suffering from insomnia (Gehrman et al., 2018; Humer et al., 2020), which revealed fluctuations in Fructose and Methanol that mirrored patterns described in relation to sleep deprivation and subsequent recovery periods. Additionally, changes in N6-Acetyllysine and Glycerate strongly suggest modulation of BCAA-related and glycolytic pathways. Further similarities between our studies were also observed when comparing these findings with the ones observed for diurnal urinary metabolite rhythms during a 24-hour sleep/wake cycle followed by a 24-hour period of total sleep deprivation (TSD, (Giskeodegard et al., 2015; Humer et al., 2020; Scholz et al., 2024)), which identified significant changes across a panel of metabolites such as 3-Hydroxyisobutyrate, Formate, TMAO, and Acetate. Such consistent similarities across independent investigations provides a powerful incentive to employ our studies for future clinical and research applications for precise diagnostics and targeted interventions in sleep disorders.

To test our scientific inkling, we compared our findings to the low molecular weight metabolites in the Nightingale Health Atlas ((Julkunen et al., 2023; Xu et al., 2023), Figure S9) under International Statistical Classification of Diseases and Related Health Problems (ICD) - G47 sleep disorder classification, we identified several novel biomarkers not currently included in standard panels. Biochemicals such as NAG, Glycocholate, 4-Guanidinobutanoate, D-Threitol, 2-Deoxyuridine, Pyridoxine, and TMAO represent unexplored metabolites in sleep disorder biomarker research. These compounds participate in critical pathways including nitrogen metabolism, bile acid signalling, microbial metabolism, and vitamin B6 dynamics, suggesting new biochemical mechanisms not captured by conventional NMR profiling approaches.

Among overlapping markers, Glutamine and Histidine are directly represented in the Nightingale panel. Both showed significantly decreased concentrations in association with sleep disorders, aligning with external research linking disrupted amino acid metabolism to neuroinflammation and circadian rhythm disturbances. Though not explicitly listed but generalized as SCFA in the Nightingale list, 3-Hydroxyisobutyrate functions within metabolic networks connected to glycolysis intermediates like Pyruvate and Lactate, which also showed reduced levels in our study. This pattern suggests impaired energy homeostasis as a central feature of sleep disruption.

Clinical and translational significance

The identification of specific metabolic signatures associated with sleep disruption offers several clinically relevant implications. First, these biomarkers may serve as objective indicators for diagnosing sleep disorders. Second, the metabolic profiles identified could help stratify patients into distinct subgroups based on underlying biochemical disruptions, enabling more targeted therapeutic approaches. Third, longitudinal monitoring of these metabolites might provide objective measures of treatment efficacy, allowing for personalized adjustment of interventions.

Our findings emphasize the relationship between sleep, metabolism, and mental health. The altered metabolic pathways, especially those involving amino acid metabolism, energy production, and gut microbial function offer deep insights into how sleep disruption may contribute to psychiatric vulnerability. Conversely, these same pathways may explain how primary psychiatric conditions can precipitate sleep disturbances, creating a self-reinforcing cycle of dysregulation.

The gut-brain axis emerges as a particularly important mediator in this relationship. Metabolites like TMAO, Glycocholate, 4-Guandinobutanoate, and 3-Hydroxyphenyl acetate directly reflect gut microbial activity and its influence on host metabolism. This observation aligns with emerging literature on the microbiome-gut-brain axis and its role in both sleep regulation and mental health (Dos Santos & Galie, 2024; Naufel et al., 2023; Roth et al., 2021).

Conclusion and perspective

This study demonstrates the utility of qNMR metabolomics (Cukkemane, 2023) for identifying circulating biomarkers associated with sleep disorders (Howard, 2005; Olson et al., 2009; Wickwire et al., 2017) in day-to-day life and work settings. The metabolic signatures identified highlight widespread disruptions across multiple biochemical pathways, particularly affecting gut microbial metabolism, energy production, amino acid metabolism, and vitamin synthesis. These findings not only enhance our understanding of the biochemical consequences of sleep disruption but also offer potential avenues for developing objective diagnostic tools and monitoring therapeutic interventions for neurological disorders that include neurodegenerative and psychiatric conditions.

While NMR-based platforms like Nightingale provide robust coverage of certain metabolic domains, our approach identified several clinically relevant metabolites previously unexplored in sleep disorder research. These novel biomarkers, particularly those involved in neurotransmission, the gut-liver axis, and micronutrient metabolism, represent promising targets for expanded diagnostic models. The qNMR-based approach offers considerable advantages for biomarker discovery and clinical translation, other avenues warrant exploration such as integration with mass spectrometry would provide more comprehensive coverage of the metabolome, particularly for low-abundance compounds. Development of sample preparation methods that retain lipophilic components while maintaining spectral quality would address a significant limitation of the current approach. Integrating metabolomic data (Julkunen et al., 2023; Xu et al., 2023) with other omics platforms such as genomics, transcriptomics, and proteomics can support a systems biology approach to better understand sleep disorders and their relationship with other CMIs. Utilizing qNMR metabolomics in a biomarker-driven diagnostic framework provides a reliable method for detecting and monitoring sleep disorders. Incorporating these insights into machine learning (ML, (Woelders et al., 2023)) and artificial intelligence (AI, (Beer et al., 2024)) models holds significant promise for developing scalable, automated diagnostic tools.

By advancing our understanding of the metabolic consequences of sleep disorders, we contribute to a growing body of research advocating for integrated diagnostic models (Depner et al., 2020; Mullington

et al., 2016) that recognize the physiological interconnectedness of sleep, metabolism, and mental health. This integrated approach may ultimately lead to more precise diagnostic tools, personalized treatment strategies, and improved outcomes for individuals living at the intersection of circadian, psychiatric, and metabolic dysfunction.

Author Contributions

E.-M.E., D.E., A.C., A.B. and D.A., designed research; A.C., K.B., E.-M.E. and D.E., performed research; A.C., K.B., D.L., E.-M.E. and D.E. contributed analytic tools; A.C., K.B., G.K., D.L., K.T., E.-M.E. and D.E. analyzed data; and A.C., E.-M.E., D.E., D.L wrote the paper, A.C., D.W., D.L., E.H., A.B., D.A., E.-M.E. and D.E revised the paper

Acknowledgments

The authors gratefully thank their colleagues from the DLR department of Sleep and Human Factors Research and from Research Center Jülich (INM-2) especially Judith Fronczek-Poncelet. The authors further gratefully acknowledge the financial support and endorsement from the DLR Management Board Young Research Group Leader Program and the Executive Board Member for Space Research and Technology, Helmholtz Information & Data Science Academy (HIDA) – trainee network to conduct the research activity.

Funding: This research was supported by internal institutional funds and the project SleepLess, which received funding from BMBF (grant # 01EW1808), FWO and FRQS under the frame of ERA-NET Neuron Cofund.

Access to the Jülich-Düsseldorf Biomolecular NMR Center jointly run by Forschungszentrum Jülich and HHU and support by the DFG (HE 3243/4-1 and INST 208/771-1 FUGG) are acknowledged.

Conflict of Interest: A.C., D.E., and E.-M.E. are co-inventors of patent (PCT/EP2024/073111) covering the application of biomarkers for sleep disorders. All the other authors declare no competing interests.

References.

- Beer, S., Elmenhorst, D., Bischof, G. N., Ramirez, A., Bauer, A., Drzezga, A., & Alzheimer's Disease Neuroimaging, I. (2024). Explainable artificial intelligence identifies an AQP4 polymorphism-based risk score associated with brain amyloid burden. *Neurobiol Aging*, 143, 19-29. <https://doi.org/10.1016/j.neurobiolaging.2024.08.002>
- Bligh, E. G., & Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can J Biochem Physiol*, 37(8), 911-917. <https://doi.org/10.1139/o59-099>
- Cukkemane, A., Elmenhorst, D., Elmenhorst, E-M (2023). *Biomarkers for the diagnosis of sleep disorders* (Germany Patent No. EPO).

- Cukkemane, A., Hurlemann, R., & Sanchez, V. G. (2025). Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. *Journal of Affective Disorders Reports*, 20, 100893. <https://doi.org/https://doi.org/10.1016/j.jadr.2025.100893>
- Depner, C. M., Cogswell, D. T., Bisesi, P. J., Markwald, R. R., Cruickshank-Quinn, C., Quinn, K., Melanson, E. L., Reisdorph, N., & Wright, K. P. (2020). Developing preliminary blood metabolomics-based biomarkers of insufficient sleep in humans. *Sleep*, 43(7). <https://doi.org/10.1093/sleep/zsz321>
- Dos Santos, A., & Galie, S. (2024). The Microbiota-Gut-Brain Axis in Metabolic Syndrome and Sleep Disorders: A Systematic Review. *Nutrients*, 16(3). <https://doi.org/10.3390/nu16030390>
- Edgar, N., & McClung, C. A. (2013). Major depressive disorder: a loss of circadian synchrony? *Bioessays*, 35(11), 940-944. <https://doi.org/10.1002/bies.201300086>
- Fink, A. M. (2024). Shift Work Sleep Disorder and Mental Health: An Integrative Review of Neurobiological, Sociological, and Psychological Perspectives With Public Policy Implications. *Policy Polit Nurs Pract*, 25(2), 94-102. <https://doi.org/10.1177/15271544241238752>
- Gehrman, P., Sengupta, A., Harders, E., Ubeydullah, E., Pack, A. I., & Weljie, A. (2018). Altered diurnal states in insomnia reflect peripheral hyperarousal and metabolic desynchrony: a preliminary study. *Sleep*, 41(5). <https://doi.org/10.1093/sleep/zsy043>
- Giskeodegard, G. F., Davies, S. K., Revell, V. L., Keun, H., & Skene, D. J. (2015). Diurnal rhythms in the human urine metabolome during sleep and total sleep deprivation. *Sci Rep*, 5, 14843. <https://doi.org/10.1038/srep14843>
- Hennecke, E., Lange, D., Steenbergen, F., Fronczek-Poncelet, J., Elmenhorst, D., Bauer, A., Aeschbach, D., & Elmenhorst, E. M. (2021). Adverse interaction effects of chronic and acute sleep deficits on spatial working memory but not on verbal working memory or declarative memory. *J Sleep Res*, 30(4), e13225. <https://doi.org/10.1111/jsr.13225>
- Howard, S. K. (2005). Sleep deprivation and physician performance: why should I care? *Proc (Bayl Univ Med Cent)*, 18(2), 108-112; discussion 112-103. <https://doi.org/10.1080/08998280.2005.11928045>
- Humer, E., Pieh, C., & Brandmayr, G. (2020). Metabolomics in Sleep, Insomnia and Sleep Apnea. *Int J Mol Sci*, 21(19). <https://doi.org/10.3390/ijms21197244>
- Julkunen, H., Cichonska, A., Tiainen, M., Koskela, H., Nybo, K., Makela, V., Nokso-Koivisto, J., Kristiansson, K., Perola, M., Salomaa, V., Jousilahti, P., Lundqvist, A., Kangas, A. J., Soininen, P., Barrett, J. C., & Wurtz, P. (2023). Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*, 14(1), 604. <https://doi.org/10.1038/s41467-023-36231-7>
- Li, C., Hu, Q., Zhang, D., Hoffstaedter, F., Bauer, A., & Elmenhorst, D. (2022). Neural correlates of affective control regions induced by common therapeutic strategies in major depressive

- disorders: An activation likelihood estimation meta-analysis study. *Neurosci Biobehav Rev*, 137, 104643. <https://doi.org/10.1016/j.neubiorev.2022.104643>
- McClung, C. A. (2013). How Might Circadian Rhythms Control Mood? Let Me Count the Ways. *Biological Psychiatry*, 74(4), 242-249. <https://doi.org/10.1016/j.biopsych.2013.02.019>
- Mullington, J. M., Abbott, S. M., Carroll, J. E., Davis, C. J., Dijk, D. J., Dinges, D. F., Gehrman, P. R., Ginsburg, G. S., Gozal, D., Haack, M., Lim, D. C., Macrea, M., Pack, A. I., Plante, D. T., Teske, J. A., & Zee, P. C. (2016). Developing Biomarker Arrays Predicting Sleep and Circadian-Coupled Risks to Health. *Sleep*, 39(4), 727-736. <https://doi.org/10.5665/sleep.5616>
- Naufel, M. F., Truzzi, G. M., Ferreira, C. M., & Coelho, F. M. S. (2023). The brain-gut-microbiota axis in the treatment of neurologic and psychiatric disorders. *Arq Neuropsiquiatr*, 81(7), 670-684. <https://doi.org/10.1055/s-0043-1767818> (O eixo microbiota-intestino-cerebro no tratamento de desordens neurológicas e psiquiátricas.)
- Olson, E. J., Drage, L. A., & Auger, R. R. (2009). Sleep deprivation, physician performance, and patient safety. *Chest*, 136(5), 1389-1396. <https://doi.org/10.1378/chest.08-1952>
- Prevention), C.-U. C. f. D. C. a. (2024, 14-05-2024). *Sleep Facts and Stats*. Centers for Disease Control and Prevention. <https://www.cdc.gov/sleep/data-research/facts-stats/index.html>
- Riemann, D., Benz, F., Dressle, R. J., Espie, C. A., Johann, A. F., Blanken, T. F., Leerssen, J., Wassing, R., Henry, A. L., Kyle, S. D., Spiegelhalder, K., & Van Someren, E. J. W. (2022). Insomnia disorder: State of the science and challenges for the future. *J Sleep Res*, 31(4), e13604. <https://doi.org/10.1111/jsr.13604>
- Riemann, D., Espie, C. A., Altena, E., Arnardottir, E. S., Baglioni, C., Bassetti, C. L. A., Bastien, C., Berzina, N., Bjorvatn, B., Dikeos, D., Dolenc Groselj, L., Ellis, J. G., Garcia-Borreguero, D., Geoffroy, P. A., Gjerstad, M., Goncalves, M., Hertenstein, E., Hoedlmoser, K., Hion, T., . . . Spiegelhalder, K. (2023). The European Insomnia Guideline: An update on the diagnosis and treatment of insomnia 2023. *J Sleep Res*, 32(6), e14035. <https://doi.org/10.1111/jsr.14035>
- Roth, W., Zadeh, K., Vekariya, R., Ge, Y., & Mohamadzadeh, M. (2021). Tryptophan Metabolism and Gut-Brain Homeostasis. *Int J Mol Sci*, 22(6). <https://doi.org/10.3390/ijms22062973>
- Scholz, M., Steuer, A. E., Dobay, A., Landolt, H. P., & Kraemer, T. (2024). Assessing the influence of sleep and sampling time on metabolites in oral fluid: implications for metabolomics studies. *Metabolomics*, 20(5), 97. <https://doi.org/10.1007/s11306-024-02158-3>
- Shah, A., Turkistani, A., Luenam, K., Yaqub, S., Ananias, P., Jose, A. M., Melo, J. P., & Mohammed, L. (2022). Is Shift Work Sleep Disorder a Risk Factor for Metabolic Syndrome and Its Components? A Systematic Review of Cross-Sectional Studies. *Metab Syndr Relat Disord*, 20(1), 1-10. <https://doi.org/10.1089/met.2021.0070>
- Sharma, S., & Kavuru, M. (2010). Sleep and metabolism: an overview. *Int J Endocrinol*, 2010. <https://doi.org/10.1155/2010/270832>

- Taillard, J., Sagaspe, P., Philip, P., & Bioulac, S. (2021). Sleep timing, chronotype and social jetlag: Impact on cognitive abilities and psychiatric disorders. *Biochem Pharmacol*, 191, 114438. <https://doi.org/10.1016/j.bcp.2021.114438>
- Trammer, R. A., Rooney, D., Benderoth, S., Wittkowski, M., Wenzel, J., & Elmenhorst, E. M. (2024). Effects of moderate alcohol consumption and hypobaric hypoxia: implications for passengers' sleep, oxygen saturation and heart rate on long-haul flights. *Thorax*, 79(10), 970-978. <https://doi.org/10.1136/thorax-2023-220998>
- van Os, J., & Kapur, S. (2009). Schizophrenia. *Lancet*, 374(9690), 635-645. [https://doi.org/10.1016/S0140-6736\(09\)60995-8](https://doi.org/10.1016/S0140-6736(09)60995-8)
- Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P., & Luchinat, C. (2019). High-Throughput Metabolomics by 1D NMR. *Angew Chem Int Ed Engl*, 58(4), 968-994. <https://doi.org/10.1002/anie.201804736>
- Wickwire, E. M., Geiger-Brown, J., Scharf, S. M., & Drake, C. L. (2017). Shift Work and Shift Work Sleep Disorder: Clinical and Organizational Perspectives. *Chest*, 151(5), 1156-1172. <https://doi.org/10.1016/j.chest.2016.12.007>
- Wishart, D. S. (2019). Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*, 99(4), 1819-1875. <https://doi.org/10.1152/physrev.00035.2018>
- Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., . . . Querengesser, L. (2007). HMDB: the human metabolome database. *Nucleic Acids Research*, 35, D521-D526. <https://doi.org/10.1093/nar/gkl923>
- Woelders, T., Revell, V. L., Middleton, B., Ackermann, K., Kayser, M., Raynaud, F. I., Skene, D. J., & Hut, R. A. (2023). Machine learning estimation of human body time using metabolomic profiling. *Proc Natl Acad Sci U S A*, 120(18), e2212685120. <https://doi.org/10.1073/pnas.2212685120>
- Xu, Y., Ritchie, S. C., Liang, Y., Timmers, P., Pietzner, M., Lannelongue, L., Lambert, S. A., Tahir, U. A., May-Wilson, S., Foguet, C., Johansson, A., Surendran, P., Nath, A. P., Persyn, E., Peters, J. E., Oliver-Williams, C., Deng, S., Prins, B., Luan, J., . . . Inouye, M. (2023). An atlas of genetic scores to predict multi-omic traits. *Nature*, 616(7955), 123-131. <https://doi.org/10.1038/s41586-023-05844-9>

Tables:

Table 1: The respective measured mean values of concentrations (in μM) for the five significantly altered metabolites that were identified using mixed linear model (see Figure 2) for the participant samples following the longitudinal experimentation:

	B2	E5	change vs	R1	change vs	TSD	change vs
	[μM]	[μM]	B2 [%]	[μM]	B2 [%]	[μM]	B2 [%]
3-hydroxyphenyl acetate	101.3	98.2	-3,06%	88.05	-13,08%	91.69	-9,49%
N-acetyl glutamine	25.69	29.15	13,47%	27.63	7,55%	29.16	13,51%
4-Guanidinobutanoate	11.48	11.13	-3,05%	10.742	-6,43%	15.12	31,71%
Glycocholate	16.19	15.64	-3,40%	13.44	-16,99%	13.5	-16,62%
Trimethylamine N-oxide	17.02	14.29	-16,04%	16.07	-5,58%	15.56	-8,58%

Table 2: The metabolites that are significantly up- (green) and down- (red) regulated in longitudinal study for the participants at different time points:

Metabolites (B2 vs E5)	Metabolites (B2 vs R)	Metabolites (B2 vs TSD)
Acetamide Asparagine N-Carbamoylaspartate N6-Acetyllysine O-phosphoethanolamine Trimethylamine N-oxide	3-Hydroxyisobutyrate 3-Hydroxyphenylacetate D-Threitol Ferulate Glycerate Methylmalonate O-Phosphocholine	2'-Deoxyuridine Fructose Isobutyrate Methanol Methionine N-acetylglutamine Pyridoxine
Metabolites (E5 vs R)	Metabolites (E5 vs TSD)	Metabolites (R vs TSD)
2-Hydroxy-3-methylvalerate 2-Phosphoglycerate 3,5-Dibromotyrosine 3-Hydroxyisobutyrate 3-Hydroxyphenylacetate 5-Hydroxylysine Aspartate Betaine Glutamate Glycerate Glycocholate Histidine N-acetylglutamine Proline	2'-Deoxyuridine 4-Carboxyglutamate Erythritol Formate Glycocholate N-Carbamoylaspartate	1,3-Dihydroxyacetone 4-Guanidobutanote D-Threitol Isovalerate Xylitol Xylose

Figures:

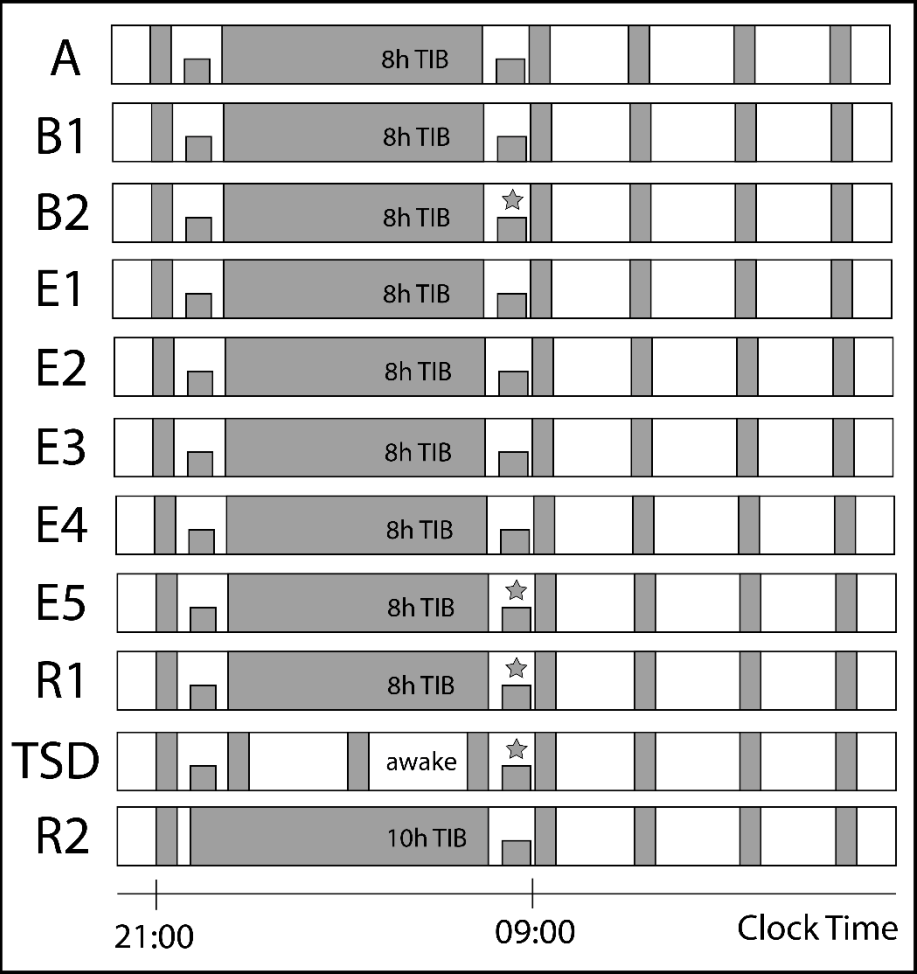


Figure 1 depicts the study design used in the present invention in order to illustrate when blood samples were taken. The volunteers were subjected to sleep-based experiments. Days are displayed as rows: A, adaptation; B2, baseline; E5, experimental day 5; R1 recovery; TSD is total sleep deprivation; TIB is time in bed. The time points at which the blood samples were collected for the NMR metabolomics experiments are shown as stars

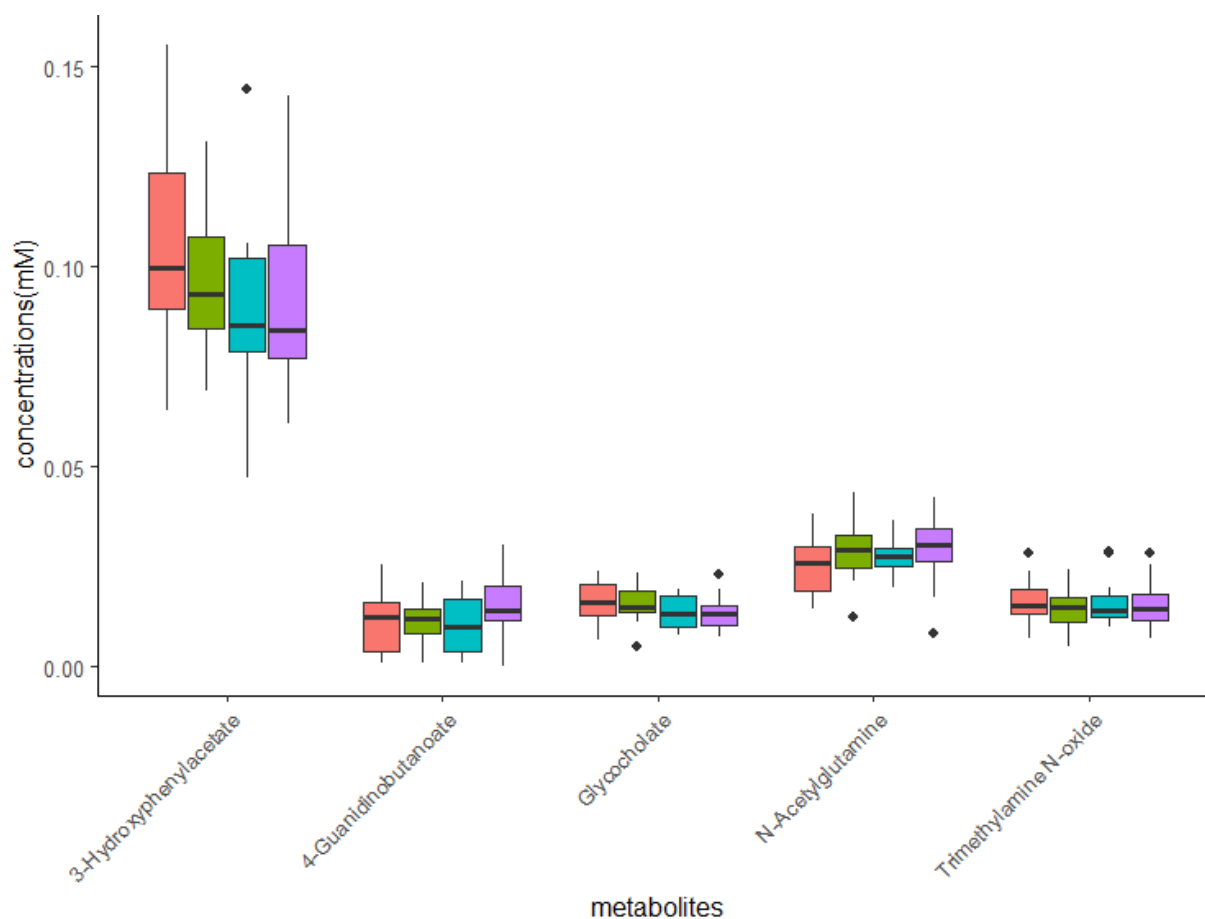


Figure 2 shows a graphical representation of a mixed linear model on the log 2 transformed data showing the most robust metabolites from plasma samples collected from various participants at different time points B2 (red), E5 (green), R1 (cyan) and TSD (purple, see e.g., figure1), wherein the respective compounds (3-HPA, 4-guanidinobutanoate, glycocholate, NAG, TMAO) are plotted against their respective concentrations given in mM in the aqueous phase from plasma of blood samples.

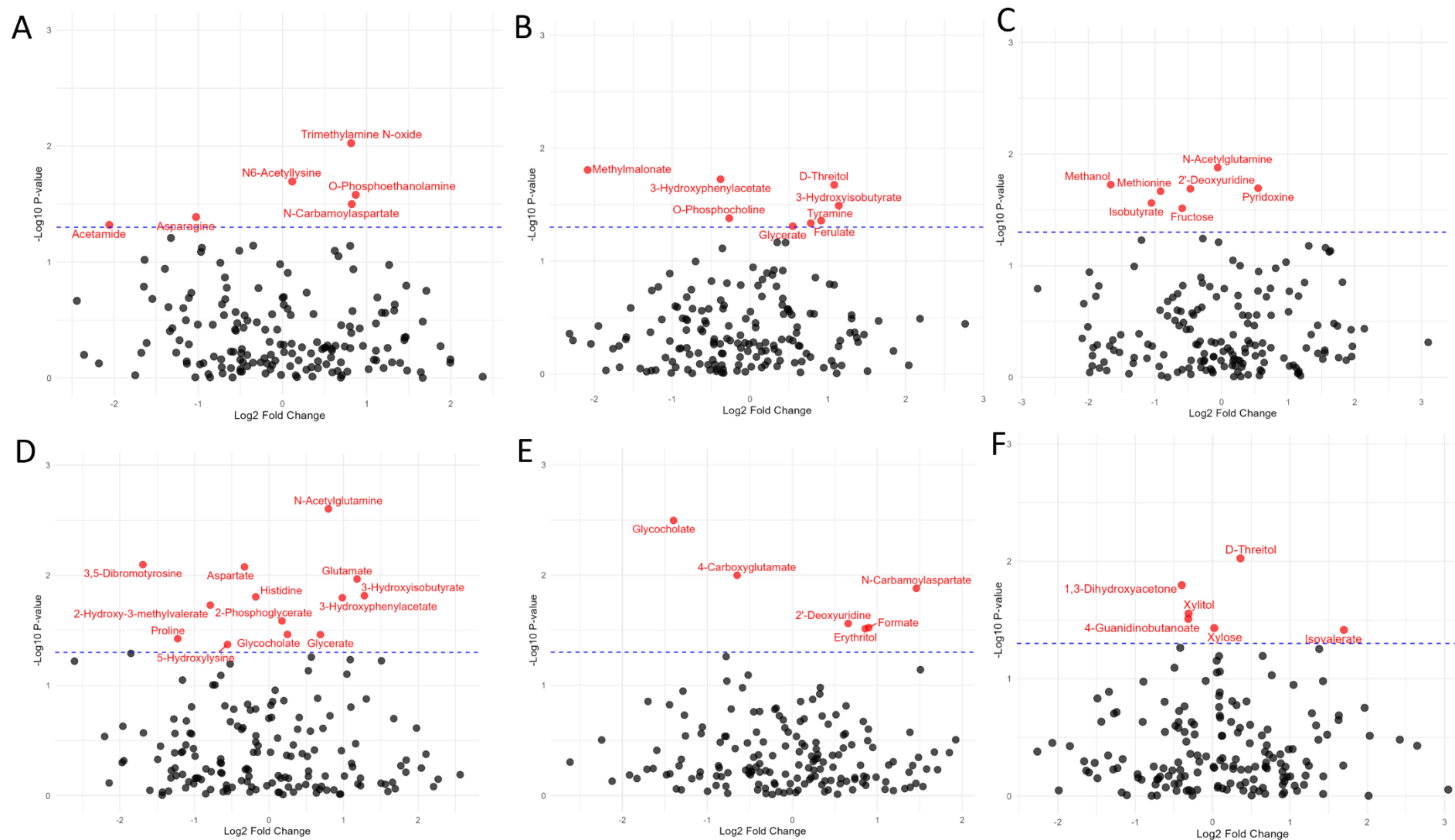


Figure 3 shows a volcano plot with the results of the paired t-tests comparing the metabolite intensities of the various test groups (A) B2 vs E5, (B) B2 vs R, (C) B2 vs TSD, (D) E5 vs R, (E) E5 vs TSD, and (F) R vs TSD.

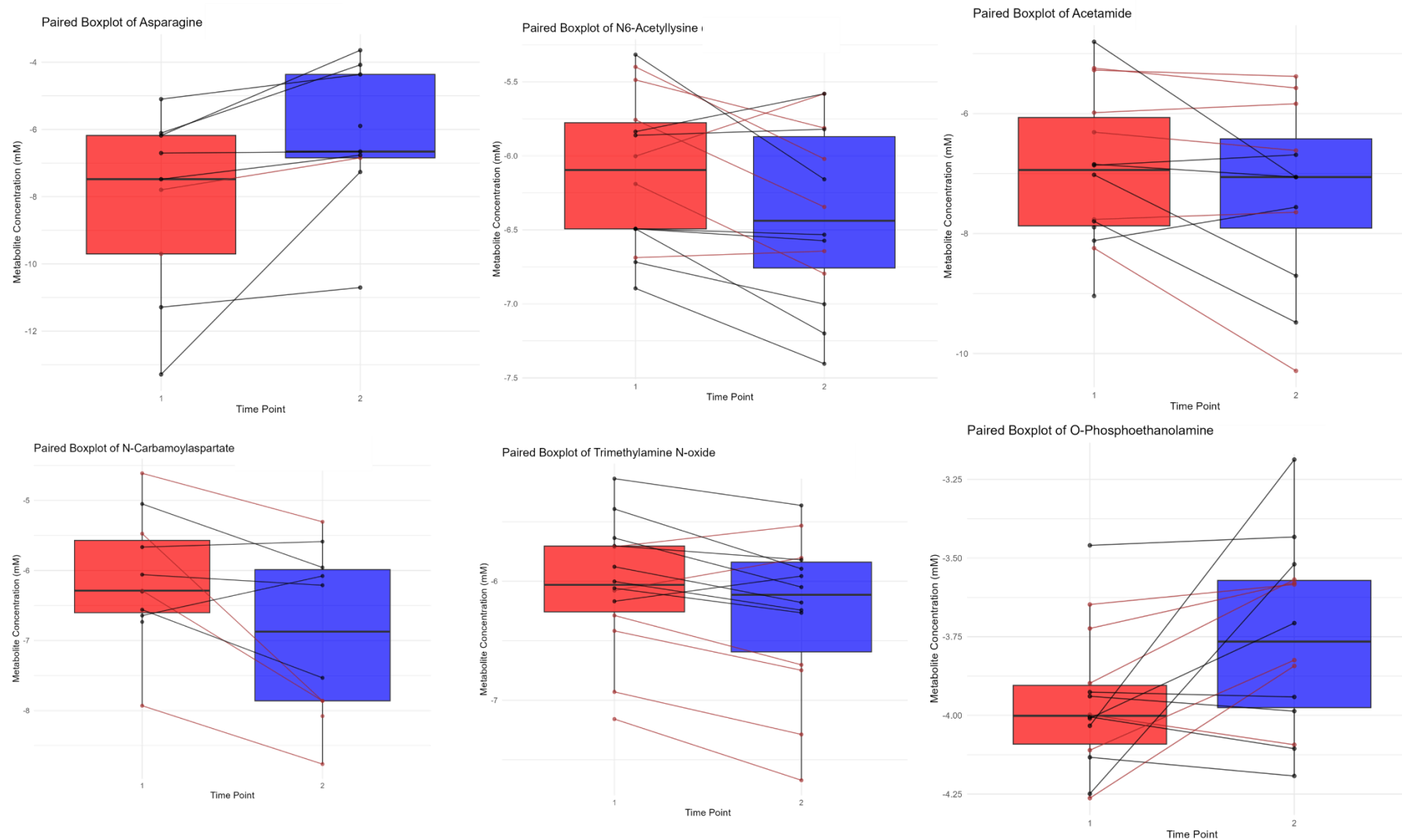


Figure 4 shows linear plots (females (brown) vs males (black)) of the participants overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups B2 and R.

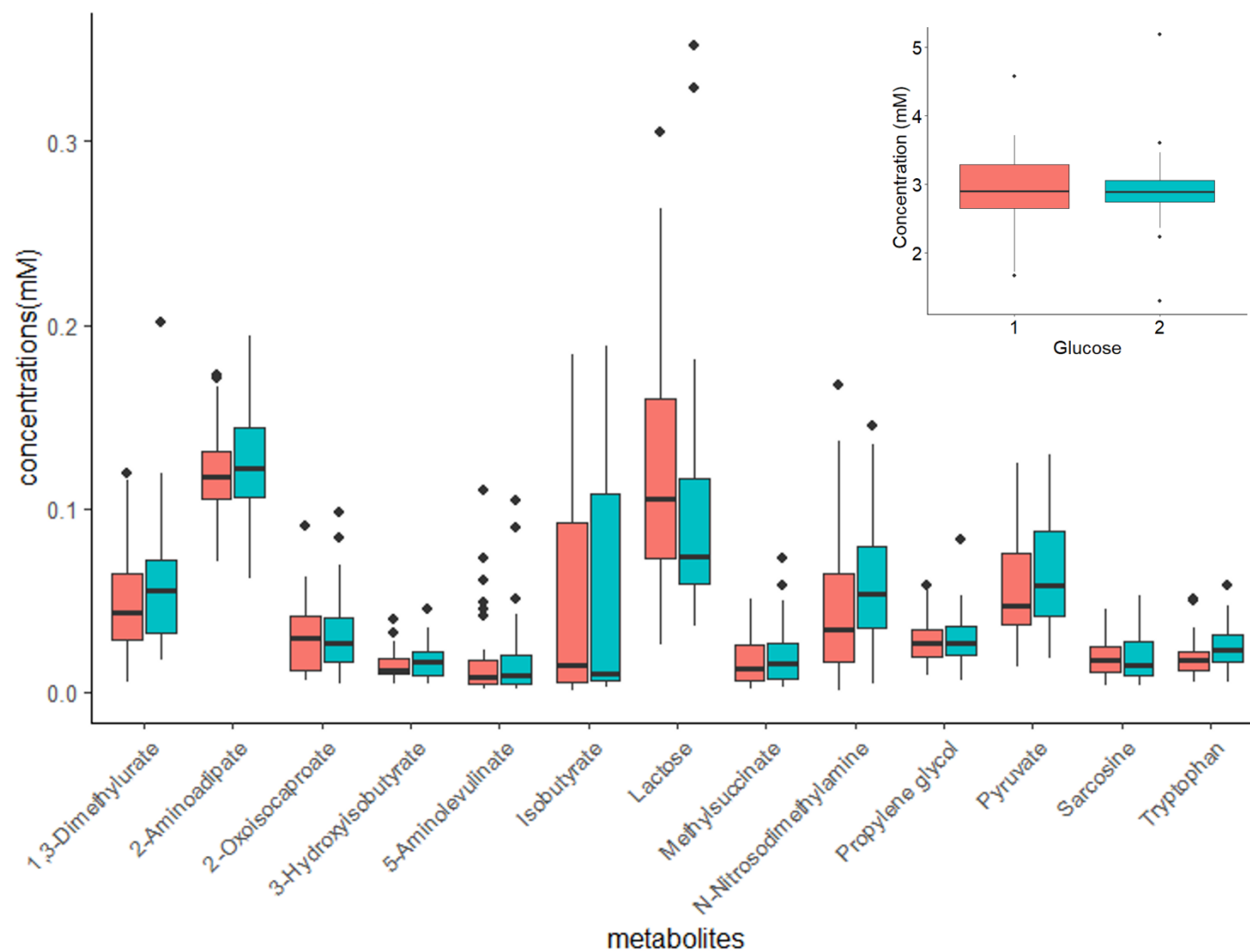


Figure 5 shows a graphical representation of a linear model on the log 2 transformed data showing the significant metabolites of participants from Study 2 from plasma collected plasma samples.

Toward precision diagnostics for circadian disruption: qNMR metabolomics and metabolic dysregulation

Abhishek Cukkemane^{i,k*}, Kevin Bochinsky^{j,l}, Denise Lange^{a,b}, Eva Hennecke^a, Anna L. Foerges, Simone Beer, Andreas Bauer^{f,g}, Daniel Aeschbach^{a,h,i}, Dieter Willbold^{j,k}, Gabi Kastenmüller^m, Alexander Drzezga, Jens Jordan, David Elmenhorst^{c,1}, Eva-Maria Elmenhorst^{a,b,1,*}

a. Department of Sleep and Human Factors Research, Institute of Aerospace Medicine, German Aerospace Center, 51170 Cologne, Germany

b. Institute of Occupational, Social, and Environmental Medicine, Medical Faculty, RWTH Aachen University, 52074 Aachen, Germany

c. Institute for Translational Psychiatry, University of Münster, 48149 Münster, Germany

d. Institute of Neuroscience and Medicine (INM-5), Forschungszentrum Jülich, 52425 Jülich, Germany

e. Institute of Radiochemistry and Experimental Molecular Imaging, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Straße 62, 50937 Cologne, Germany

f. Institute of Neuroscience and Medicine (INM-2), Forschungszentrum Jülich, 52425 Jülich, Germany

g. Department of Neurology, Medical Faculty, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

h. Institute of Experimental Epileptology and Cognition Research, University of Bonn, Medical Center, 53127 Bonn, Germany

i. Division of Sleep Medicine, Harvard Medical School, Sleep Division, Boston 02115, MA, USA

j. Institute of Biological Information Processing, Structural Biochemistry (IBI 7), Forschungszentrum Jülich, 52425 Jülich, Germany.

k. Heinrich Heine University Düsseldorf, Institut für Physikalische Biologie, Düsseldorf, Germany.

l. current address -Bruker Biospin, Ettlingen, Germany

m. Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, Neuherberg, Germany.

1. A.C, E.-M.E. and D.E. contributed equally to this work

* Corresponding Author: Eva-Maria Elmenhorst, MD
DLR-Institute of Aerospace Medicine,
51170 Cologne, Germany
Phone: +49 2203 601 4735
E-Mail: eva-maria.elmenhorst@dlr.de

Abhishek Cukkemane, PhD
Forschungszentrum Jülich,
52428 Jülich, Germany
Phone: +49 2461 61 9448
E-Mail: a.cukkemane@fz-juelich.de

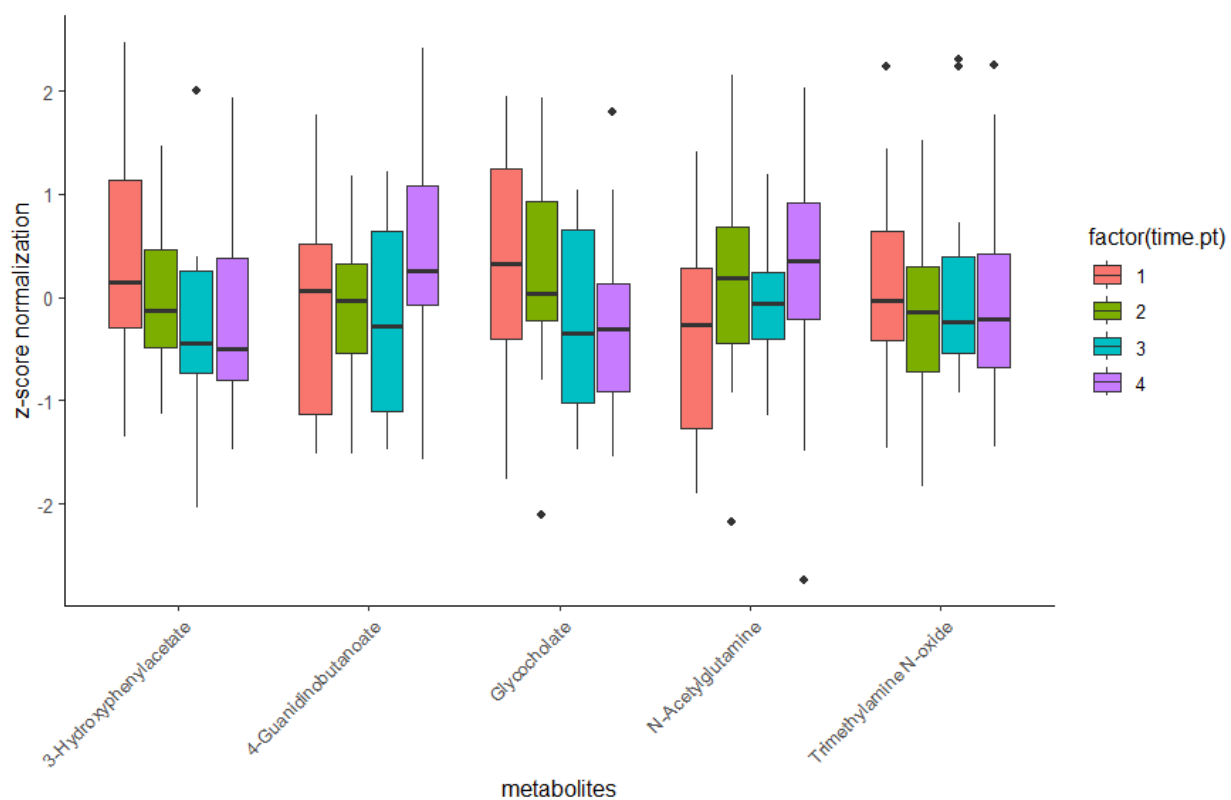


Figure S1 shows a normalized representation the metabolites (also see e.g., figure1), for the metabolites that show significant differences using the mixed-linear model for the different sample collection time points

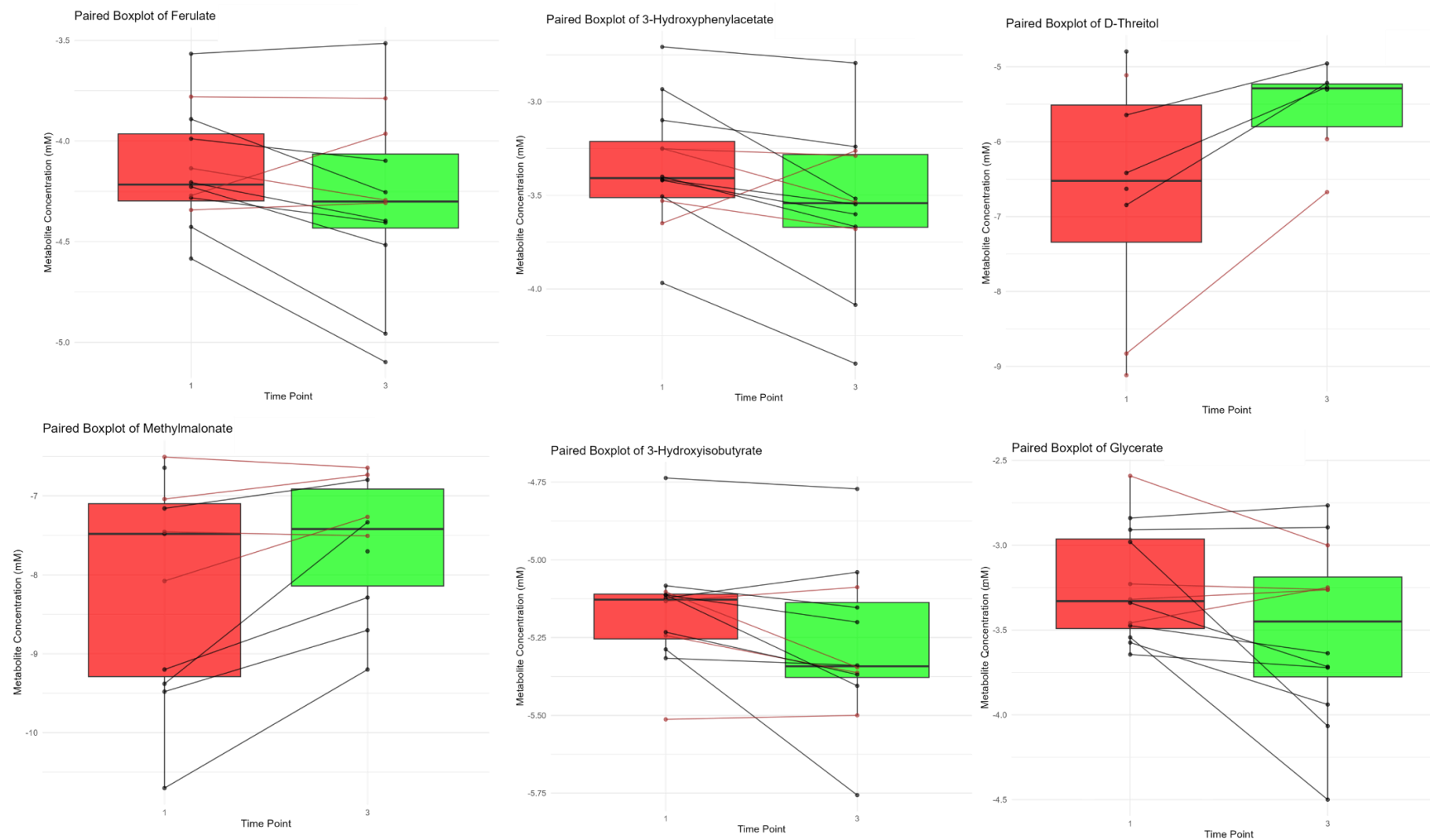


Figure S2 shows linear plots (females (brown) vs males (black)) of few significant metabolites from the participants in study group 1 overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups B and R.

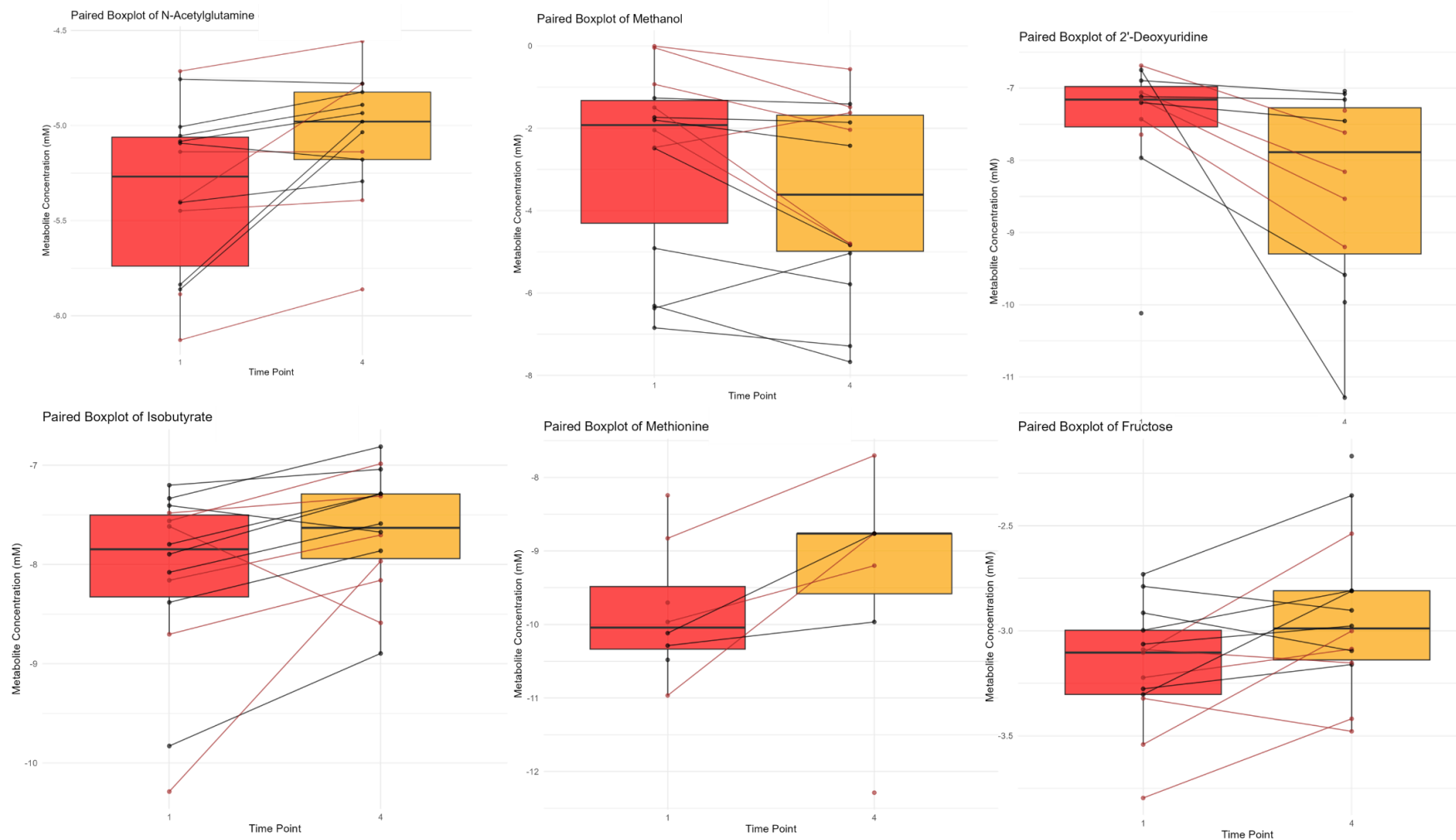


Figure S3 shows linear plots (females (brown) vs males (black)) of few significant metabolites from the participants in study group 1 overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups B and TSD.

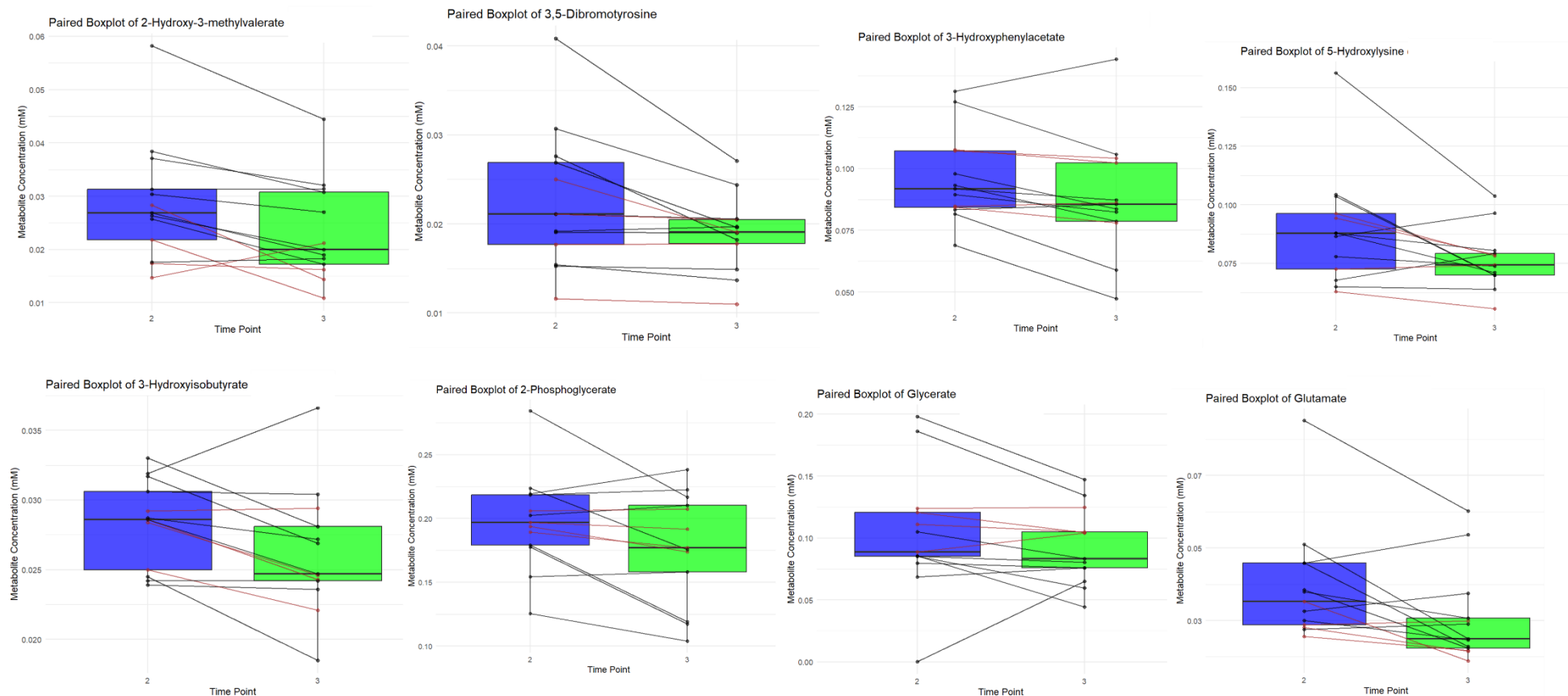
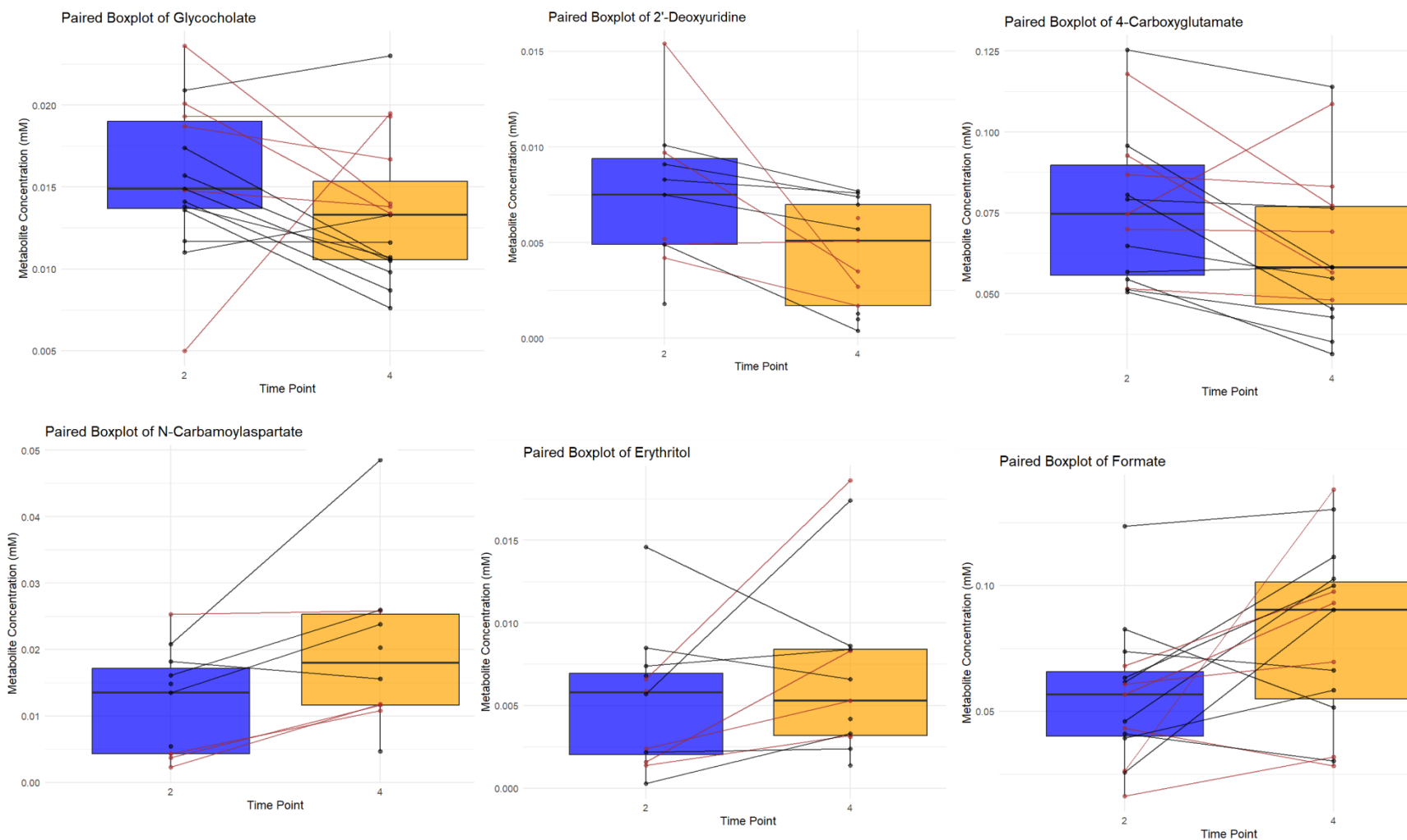


Figure S4 shows linear plots (females (brown) vs males (black)) of few significant metabolites from the participants in study group 1 overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups E5 and R.



9

Figure S5 shows linear plots (females (brown) vs males (black)) of significant metabolites from the participants in study group 1 overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups E5 and TSD.

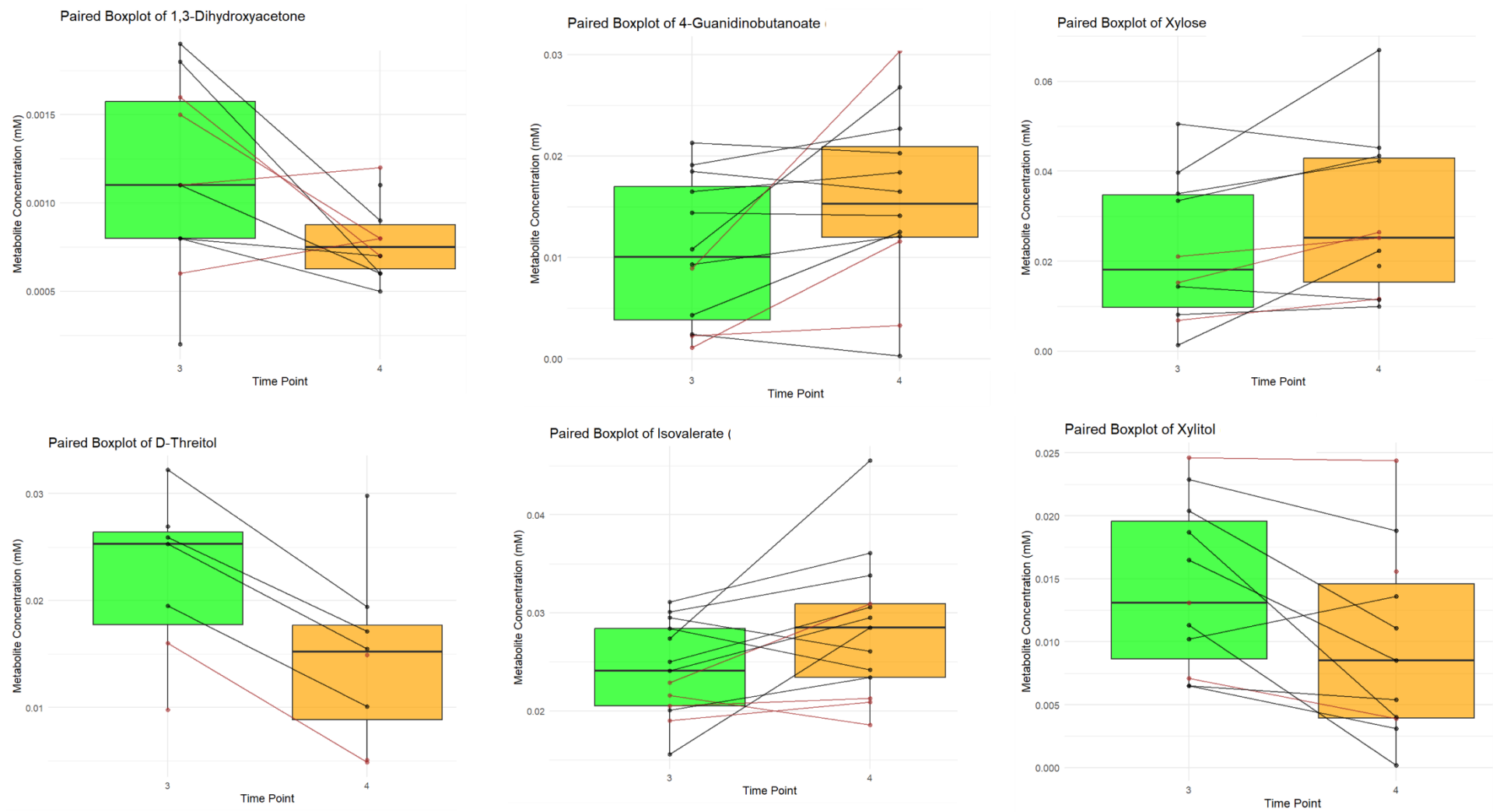


Figure S6 shows linear plots (females (brown) vs males (black)) of significant metabolites from the participants in study group 1 overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing the test groups R and TSD.

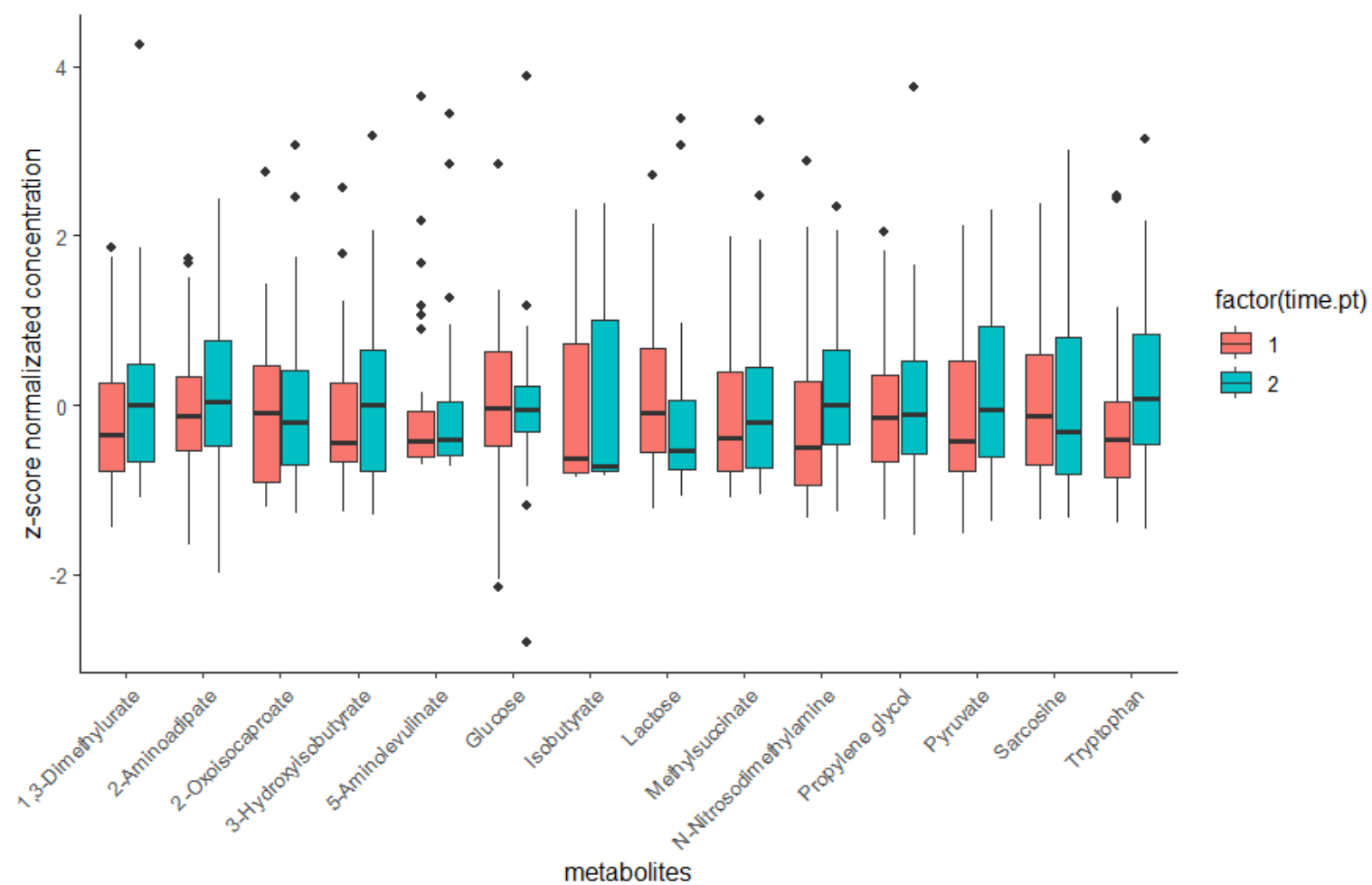


Figure S7 shows a normalized representation the metabolites (also see e.g., Figure 5), for the metabolites that show significant differences using the linear model for the different sample collection time points for subject 2

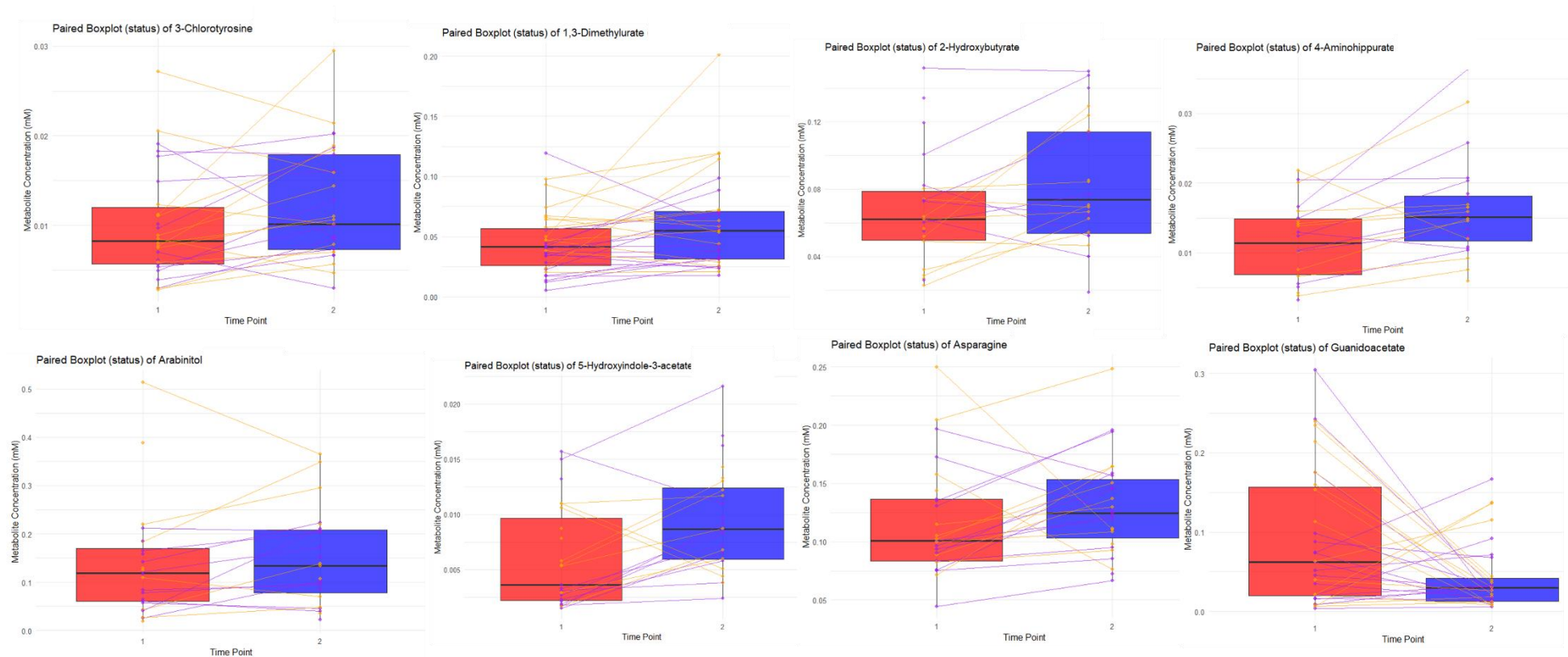


Figure S8 shows linear plots (subjects (orange) vs control (purple)) of the participants overlaid onto boxplots displaying the median values of the metabolite concentration dispersion for the identified significant for the significant biochemicals identified using the paired t-tests when comparing time points 1 and 2 for subject group 2.

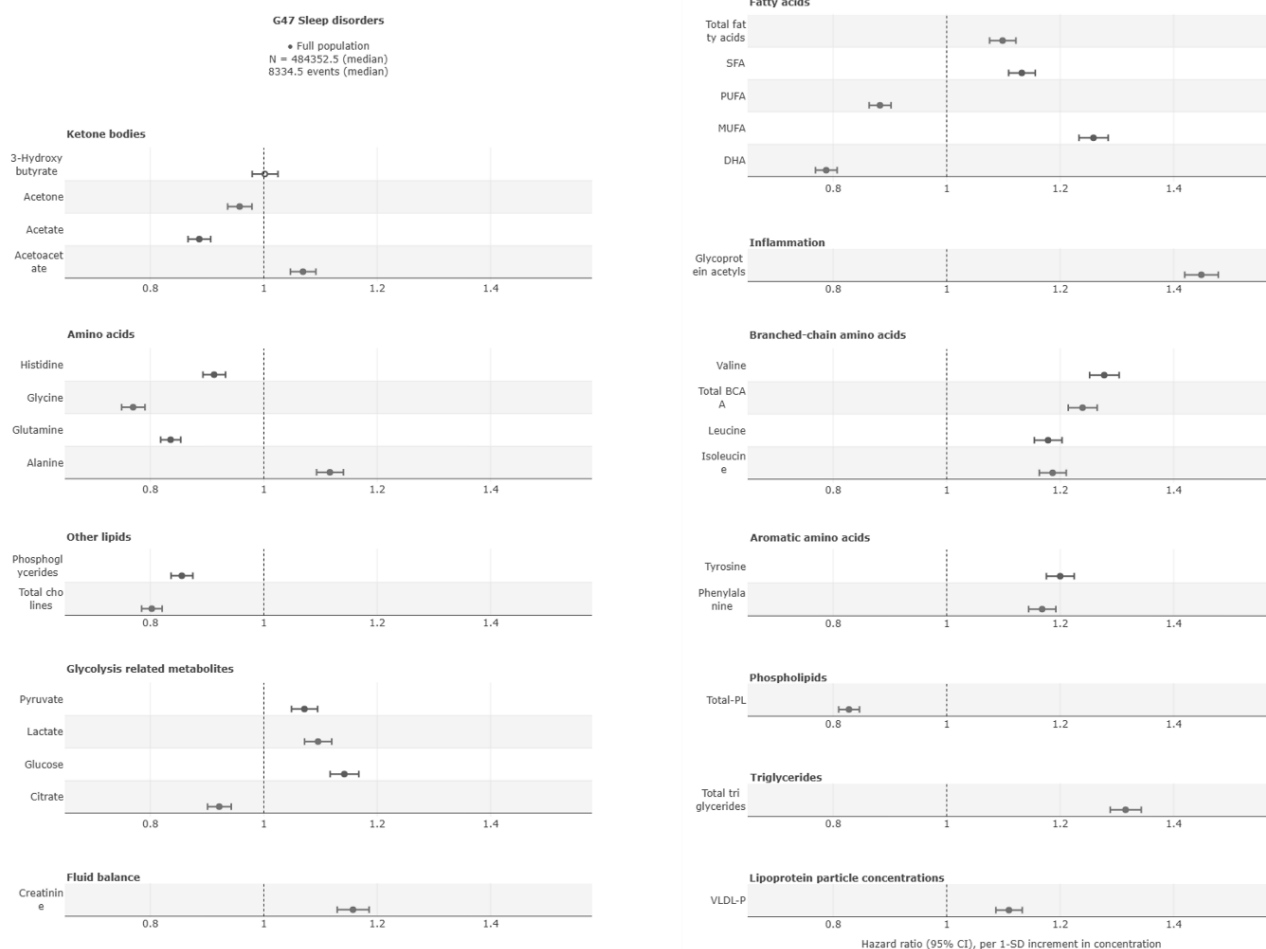


Figure S9: The image was generated based on the nightingale atlas (<https://research.nightingalehealth.com/atlas/forestplots>), where hazard ratios (with 95% confidence intervals) are presented for various metabolites grouped by low-molecular weight metabolite classes in relation to G47 sleep disorders. Each point and error bar on the forest plot represents the association strength of a metabolite with the disorder across a large population sample

5.4 Perspective: Developing diagnostic tools through insights into aberrant biochemical mechanisms

The diagnosis and treatment of chronic mental illnesses (CMIs) such as schizophrenia presents a formidable challenge in contemporary medicine. These disorders are highly heterogeneous with diverse symptomatic profiles. Adding to this complexity the current diagnostic paradigms heavily rely on subjective symptom-based assessments, that often lacks a strong biochemical foundation. Traditional psychiatric evaluations in hospitals and clinics do not capture the underlying molecular and/ or physiological disturbances that drives the underlying symptoms. In this context, with the recent advancements in metabolomics (Vignoli et al., 2019; Wishart, 2019), I was particularly fascinated with the potential of quantitative nuclear magnetic resonance (qNMR) spectroscopy for biochemical profiling.

Firstly, in collaboration with the Karl-Jaspers Klinik, Bad Zwischenahn, Germany, I performed qNMR metabolomics for plasma and urine metabolomic profiling in individuals. Using this technique, we were able to identify metabolites that were significantly different between patients and healthy controls. Furthermore, we were able to biochemical data with established clinical symptom scales used in clinics such as PANSS (Positive and Negative Syndrome Scale), BNSS (Brief Negative Symptom Scale), MADRS (Montgomery-Åsberg Depression Rating Scale), and GFS/GFR (Global Functioning: Social and Role). The results from such integration reveal those significant disruptions in amino acid and nitrogen metabolism, neurotransmitter biosynthesis, membrane dynamics, and energy pathways. Critically, these metabolic changes show symptom-specific associations. For instance, elevated creatinine levels were correlated with higher depression scores (MADRS), while altered glutamate concentrations were linked to the intensity of positive symptoms on the PANSS scale. These findings imply that schizophrenia consists of distinct biochemical subtypes, each corresponding to specific symptom dimensions valuable for diagnostic approach.

I am firmly of the opinion that such a hybrid strategy by combining clinical observation with objective metabolic markers offers a promising path toward precision diagnostics by not disrupting the traditional clinical methodology. Furthermore, it enables stratification of patients beyond mere symptoms, but by underlying biochemical profile. Such insights can further pave the way for personalized interventions, targeted drug development, and improved prognostic capabilities.

Further expanding on our strategy for psychiatric disorder diagnosis and management, I focused on sleep disorders. Disruption of sleep and circadian rhythms (Edgar & McClung, 2013; Sharma & Kavuru, 2010) is both a cause and consequence of many chronic mental illnesses (CMIs), including schizophrenia and other mood disorders. In this study, blood plasma (Cukkemane, 2023) from patients exhibiting sleep-wake disturbances was analysed to isolate for several small-molecule biomarkers. These included vital components such as energy metabolites (e.g., lactate, glucose), essential vitamins and cofactors (e.g., vitamin B6, NAD⁺), and molecules influenced by gut microbiota (e.g., short-chain

fatty acids). Such markers reflect disrupted metabolic pathways offering a biochemical snapshot of circadian misalignment and also provide insights into imbalance in CMIs.

The versatility of qNMR lies in its non-invasive, cost-efficient, ease-of-use, reproducible quantitative data from minimal biological input, and ease-in-interpretation. These features allow for longitudinal monitoring of circadian health, making it suitable for both acute diagnosis and chronic management. Moreover, when paired with machine learning algorithms, this biomarker panel can be used to train predictive models capable of diagnosing sleep disorders with high sensitivity and specificity. The implications of such an approach are far-reaching as it provides a novel method for assessment of circadian disruption among shift-work and people in complex professions. Also, it offers the possibility of early detection of the clinical onset of psychiatric symptoms providing a window for preventive intervention, and it opens avenues to personalized medicine, where treatment regimens can be tailored not only to psychiatric symptoms but also to individual circadian profiles.

Altogether, from both the studies, I am confident that the integration of qNMR metabolomics for mapping biochemical pathways, diagnostic framework for CMIs and for sleep disorders highlights the potential of the technique for routine exploratory research, and also scalable solution for precision diagnostics.

References

- Cukkemane, A., Elmenhorst, D., Elmenhorst, E-M (2023). *Biomarkers for the diagnosis of sleep disorders* (Germany Patent No. EPO).
- Edgar, N., & McClung, C. A. (2013). Major depressive disorder: a loss of circadian synchrony? *Bioessays*, 35(11), 940-944. <https://doi.org/10.1002/bies.201300086>
- Sharma, S., & Kavuru, M. (2010). Sleep and metabolism: an overview. *Int J Endocrinol*, 2010. <https://doi.org/10.1155/2010/270832>
- Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P., & Luchinat, C. (2019). High-Throughput Metabolomics by 1D NMR. *Angew Chem Int Ed Engl*, 58(4), 968-994. <https://doi.org/10.1002/anie.201804736>
- Wishart, D. S. (2019). Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol Rev*, 99(4), 1819-1875. <https://doi.org/10.1152/physrev.00035.2018>

6. Summary – Holistic approach in tackling schizophrenia

Schizophrenia is a profoundly complex and multifactorial disorder, which is best understood as a societal illness, shaped by the convergence of biological vulnerability, environmental and psychosocial elements, and systemic limitations within healthcare infrastructures. Therefore, I am of the opinion that addressing schizophrenia demands a holistic, multidisciplinary strategy that integrates insights from clinical psychiatry, molecular biology, and social work.

In this thesis, in pursuit of my habilitation study, I employed for a dual, translational approach to tackling schizophrenia. The first dimension focused on the molecular (dys)function of the Disrupted-in-Schizophrenia 1 (DISC1) protein, which is a key scaffold protein involved in neurodevelopment and cell proliferation and has been implicated in the disorder's pathophysiology. Through a combination of biophysical and structural biology studies, I demonstrated how conformational heterogeneity, and structural instability may be the underlying reason for the C-region of DISC1 contributing to chronic mental illnesses (CMIs). Alongside, we have been successful in developing DISC1-binding peptides, which has the potential for both therapeutic targeting and detection.

Parallely, I sought the application of nuclear magnetic resonance (NMR) spectroscopy-based metabolomics to uncover biochemical dysfunctions associated with schizophrenia. By profiling blood biomarkers and circadian disruptions, research studies herein provide insights into disrupted biochemical pathways, identify novel metabolic signatures for schizophrenia and sleep-related comorbidities, and diagnostic applications. Such molecular and metabolic insights provide a foundation for developing scalable, patient-specific treatment models aligned with the principles of precision psychiatry.

Beyond laboratory research, I actively engaged with stakeholders including clinicians, patients' families, diagnostic innovators, and pharmaceutical enterprises to understand the broader socio-medical landscape of the disorder. Qualitative analytics and natural language processing (NLP) of stakeholder perspectives underscore the emotional and operational challenges of managing CMIs, from social worker burnout to innovation bottlenecks. Based on which, I am of the opinion that a biopsychosocial model, advocating for patient-centred care strategies informed by diverse stakeholder insights and digital health integration.

By combining clinical, molecular, and metabolomic science with societal and health sciences, this thesis provides a template for future evidence-based, scalable, and personalized diagnostic and therapeutic solutions. The research reflects a commitment to my engagement in higher education and translational science, offering fresh perspectives for the future of mental health care.

7 Publications and Achievements

One can access my profile in the following pages

https://www.fz-juelich.de/profile/cukkemane_a

<https://www.linkedin.com/in/abhishek-cukkemane-31983057>

<http://orcid.org/0000-0002-1532-6869>

https://www.researchgate.net/profile/Abhishek_Cukkemane

Awards and achievements

- 2024 Norwegian Artificial Intelligence Research Consortium - Helmholtz Information & Data Science Academy Trainee Network
- 2023 Hansewissenschaftskolleg Delmenhorst Fellowship
- 2022 Forschungszentrum Jülich INNOVATION Start program
- 2022 Alexander von Humboldt – Research group linkage
- 2022 Helmholtz Information & Data Science Academy Trainee Network
- 2020 Helmholtz Enterprise Field Study
- 2019 Forschungszentrum Jülich fellowship programme for excellent guest scientists
- 2016 Start-up Grant, Ministry of Small & Medium Enterprises (MSME), Government of India
- 2016 Start-up Award, DST-AIT, Government of India - Embassy of Switzerland
- 2007 Poster Prize, European Biophysics Congress, London, UK
- 2004 State Eligibility test for teaching in Maharashtra & Goa
- 2004 All India GATE - Ranked 112
- 2003 NCCS-Summer Research Fellowship, India
- 2001 JNCASR-Summer Research Fellowship, India

Publications (*corresponding author):

1. **Cukkemane A***, Sanchez, VG (IN REVISION) Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. Journal of Affective Disorders Reports
2. **Cukkemane A***, Becker N, Kupreichyk T, Heise H, Willbold D, Weiergräber OH (2025) Snapshots of the aggregation pathway of the scaffold protein DISC1 and the implications for chronic mental illnesses: structure mirrors function. Journal of Structural Biology: X. <https://doi.org/10.1016/j.yjsbx.2025.100128>
3. **Cukkemane A***, Dingley AD, Mohrlüder J, Santiago-Schübel B, Weiergräber OH, Willbold D (2025) DISC1 mutants are functionally active but structurally unstable to choreograph their tasks. European Journal of Pharmaceutical Sciences. <https://doi.org/10.1016/j.ejps.2025.107148>

4. **Cukkemane A***, Hureleemann R, Sanchez, VG (2025) Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. Journal of Affective Disorders Reports.
<https://doi.org/10.1016/j.jadr.2025.100893>
5. **Cukkemane, A**, Elmenhorst D and Elmenhorst E-M (2023) Biomarkers for the diagnosis of sleep disorders. Reference number – PCT/EP2024/073111
https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2025040590&_cid=P11-MBI9SA-80557-1
6. **Cukkemane, A** and Willbold, D (2022) DISC1-binding peptides and the use thereof for the treatment and diagnosis of schizophrenia and other chronic mental disorders (CMDs). Reference number - PCT/EP2023/066796
https://patentscope.wipo.int/search/en/detail.jsf?docId=EP418639163&_fid=WO2023247625
7. **Cukkemane A***, Becker N, Zielinski M, Frieg B, Lakomek NA, Heise H, Schröder GF, Willbold D, Weiergräber OH. (2021) Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. Transl Psychiatry 11, 639.
<https://doi.org/10.1038/s41398-021-01765-1>
8. **Cukkemane, A.***, Kumar, P., and Sathyamoorthy, B. (2020) Understanding the role of synbiotics in functional foods and dietary therapeutics for health, communicable and non-communicable diseases using metabolomics footprint studies. Food Research International.
<https://doi.org/10.1016/j.foodres.2019.108679>
9. **Cukkemane, A** and Sathyamoorthy, B (2019) Highly nutritious synbiotic dairy drinks based on grain amylopectin material with efficient antimicrobial properties and a process preparation thereof. Reference number – 201921015221
https://patentscope.wipo.int/search/en/detail.jsf?docId=IN242391082&_cid=P11-MBQQMP-08168-1
10. Rane, J., Singh, A.J., George, P., Govindasamy, V., **Cukkemane, A.**, Raina, S.K., Chavan, P.M., Aher, L., V. S. Sunoj, J., Singh, N.P. (2019) Effect of Cow Urine-Based Bioformulations on Growth and Physiological Responses in Mungbean Under Soil Moisture Stress Conditions. PNAS, India Section B: Biological Sciences pp1-11
<https://link.springer.com/article/10.1007/s40011-019-01088-8>
11. Pawar, S., Kalyankar, V., Dhamangaonkar, B., Dagade, S., Waghmode, S., and **Cukkemane, A.*** (2017) Biochemical profiling of antifungal activity of betel leaf (*Piper betle* L.) extract and its significance in traditional medicine. Journal of Advanced Research in Biotechnology. 1(1):4-6
<https://sybiosisonlinepublishing.com/biotechnology/biotechnology16.php>

12. **Cukkemane, A** and Waghmode, S (2017) One pot synthesis of antimicrobial photo-active herbal core shell TiO₂ quantum dots. Reference number – 201721016132
https://patentscope.wipo.int/search/en/detail.jsf?docId=IN211800196&_cid=P11-MBQQMP-08168-1
13. **Cukkemane, A** and Waghmode, S (2017) One pot synthesis of photo-active carbon-based rod shaped TiO₂ quantum dots. Reference number – 201721016133
https://patentscope.wipo.int/search/en/detail.jsf?docId=IN211799183&_cid=P11-MBQQMP-08168-1
14. **Cukkemane, A** and Waghmode, S (2017) One pot synthesis of antimicrobial photo-active herbal core shell silver quantum dots. Reference number – 201721016135
https://patentscope.wipo.int/search/en/detail.jsf?docId=IN211799185&_cid=P11-MBQQMP-08168-1
15. Tank, J., Dhamangaonkar, B., Ukale, U.D., Cukkemane, N., **Cukkemane, A.A*** (2017) Biochemical and Microbiological analysis of different organic manures; their effect on germination of Coriandrum sativum (Cilantro) and Solanum melongena (Eggplant). Journal of Bioprocessing and Biotechniques. 7(1) <https://doi:10.4172/2155-9821.1000295>
16. Tank, J., Ukale, U.D., Mapare, V., Dhamangaonkar, B., Cukkemane, N, **Cukkemane, A.A*** (2017) Metabolic profiling of wastes from raw and ripe mango seeds to comprehend its beneficial medicinal and cosmetic properties for commercialization. Journal of Agricultural Studies. 5(1) <https://doi.org/10.5296/jas.v5i1.10234>
17. Cukkemane, N. and **Cukkemane A.A.*** (2017) Comprehending the biochemical basis of liquid organic manure to promote mass production by combining livestock and organic farming practices for rural development. The role of Vrikshayurveda and Traditional practices in Organic agriculture. Thakur, R.P. and Choudhary, S.L. (eds) 84-91
18. **Cukkemane, A.*** (2016) Socio-Commercial Agri-Biotech Model for Rural Development in India by Combining Livestock and Organic Farming Practices. Journal of Advanced Research in Biotechnology. 1(1):4-6
<https://symbiosisonlinepublishing.com/biotechnology/biotechnology12.php>
19. **Cukkemane, A.A.**, and Bhatevara, S. (2016) Back to our roots; Ayurvedic organic farming by scientific principles. LAP Lambert Academic Publishing. ISBN 978-3-330-00504-4
20. Ukale, U.D., Bhagwat R.V, Upadhyay S.K, Cukkemane, N, **Cukkemane, A.A*** (2016) Metabolic analysis of liquid formulations of organic manures and its influence on growth and yield of Solanum lycopersicum L. (tomato) crop in field. Biocatalysis and Agricultural Biotechnology 8:50-54 <https://doi.org/10.1016/j.bcab.2016.08.007>
21. Kaplan, M., **Cukkemane, A.**, van Zundert, G., C., Narasimhan, S., Daniëls, M., Mance, D., Waksman, G., Bonvin, A., M., Fronzes, R., Folkers, G., E., Baldus, M. (2015) Probing a cell-

- embedded megadalton protein complex by DNP-supported solid-state NMR. *Nature Methods* 12(7):649-652 <https://doi.org/10.1038/nmeth.3406>
22. Langlois, C., Ramboarina, S., **Cukkemane, A.**, Auzat, I., Chagot, B., Gilquin, B., Ignatiou, A., Petitpas, I., Kasotakis, E., Paternostre, M., White, H.E., Orlova, E.V., Baldus M., Tavares, P., and Zinn-Justin S. (2014) Bacteriophage SPP1 Tail Tube Protein self-assembles into β -structure rich tubes. *J. Biol. Chem.* [https://www.jbc.org/article/S0021-9258\(20\)49226-9/fulltext](https://www.jbc.org/article/S0021-9258(20)49226-9/fulltext)
 23. Moiset, G., Lopez, C., Rijpkema, E., Syga, L., Bartelds, R., **Cukkemane, A.**, Baldus, M., Poolman, B., and Marrink, S.J. (2013) "Disaccharides Impact the Lateral Organization of Lipid Membranes." *J. Am. Chem. Soc.* 134(46):16167-16175 <https://pubs.acs.org/doi/10.1021/ja505476c>
 24. **Cukkemane, A.** and Baldus, M. (2013) "Structural Characterization of a cyclic-nucleotide-activated K⁺-channel and its lipid environment using solid-state NMR spectroscopy." *Chem. Bio. Chem.* 14(14):1789-1798 <https://doi.org/10.1002/cbic.201300182>
 25. Cruijssen, v,d,E. Nand, D., Weingarth, M., Prokofyev, A., **Cukkemane, A.**, Horning, S., Becker, S., Bonvin, A., M., J., J., Hulse, R.,E., Perozo, E., Pongs, O., and Baldus, M. (2013). "The importance of the lipid-pore loop interface for potassium channel structure and function" *PNAS* 110 (32): 13008-13013 <https://www.pnas.org/doi/10.1073/pnas.1305563110>
 26. Peuker, S., **Cukkemane, A.**, Held, M., Noe, F., Kaupp, U., B. and Seifert, R. (2013) "Kinetics of ligand-receptor interaction reveals an induced-fit mode of binding in cyclic nucleotide-activated protein." *Biophys. J.* : 104 (1) :63-74 <https://doi.org/10.1016/j.bpj.2012.11.3816>
 27. Gradmann, S., Ader, C., Heinrich, I., Nand, D., Dittmann, M., **Cukkemane, A.**, van Dijk, M., Bonvin, A., M., J., J., Engelhard, M., and Baldus, M. (2012) "Rapid prediction of multi-dimensional NMR data sets." *J. Biomol. NMR* 54 (4):377-387 <https://doi.org/10.1007/s10858-012-9681-y>
 28. **Cukkemane, A.**, Nand, D., Gradmann, S., Weingarth, M., Kaupp, B., and Baldus, M. (2012). "Solid-state NMR [¹³C,¹⁵N] resonance assignments of the Nucleotide-Binding Domain of a bacterial Cyclic Nucleotide-Gated channel." *Biomol. NMR assign.* 6 (2):225-229. <https://doi.org/10.1007/s12104-012-9363-4>
 29. **Cukkemane, A.A.**, Renault, M., and Baldus, M. (2012). *Solid-State NMR on Membrane Proteins: Methods and Applications*. In: *NMR of Biomolecules: Towards Mechanistic Systems Biology*, Wiley Blackwell, ISBN 978-0-470-74608-0, pp. 407-418.
 30. Nand, D., **Cukkemane, A.**, Becker, S., and Baldus, M. (2012). "Fractional deuteration applied to biomolecular solid-state NMR spectroscopy." *J. Biomol. NMR* 52 (2):91-101. <https://doi.org/10.1007/s10858-011-9585-2>
 31. Ader, C., Nand, D., Renault, M., **Cukkemane, A.A.**, and Baldus, M. (2011). "Solid-State NMR on Membrane proteins: Methods and Applications." *Advances in Biomedical Spectroscopy: Biomolecular NMR spectroscopy*, IOS Press, ISBN 978-1-60750-694-2, 3: 115-136

32. **Cukkemane, A.**, Seifert, R., and Kaupp, U.B. (2011)". Cooperative and uncooperative cyclic-nucleotide-gated ion channels". Trends Biochem. Sci. 36(1): 55-64.
<https://doi.org/10.1016/j.tibs.2010.07.004>
33. Renault, M., **Cukkemane A.**, and Baldus, M. (2010). "Solid-state NMR spectroscopy on complex biomolecules / Festkörper-NMR-Spektroskopie an komplexen Biomolekülen" Angew. Chem. Int. Ed. Engl. 49(45): 8346-57.
<https://onlinelibrary.wiley.com/doi/10.1002/ange.201002823>
34. Zhang, Z., J. Buitenhuis, **Cukkemane, A.**, Brocker, M., Bott, M., and Dhont, J.K. (2010). "Charge reversal of the rodlike colloidal fd virus through surface chemical modification". Langmuir 26(13): 10593-9. <https://pubs.acs.org/doi/10.1021/la100740e>
35. **Cukkemane, A.**, Gruter, B., Novak, K., Gensch, T., Boenigk, W., Gerharz, T., Kaupp, U.B., and Seifert, R. (2007). "Subunits act independently in a cyclic nucleotide-activated K⁺-channel". EMBO Reports 8: 749-755. <https://doi.org/10.1038/sj.embor.7401025>

Acknowledgements

This thesis represents an ambitious undertaking and I would like to sincerely thank Prof Dr. Dieter Willbold for allowing me to study and pursuit my research objectives under his guidance.

On an equal note, my umpteen thanks to Prof. Dr. Oliver Weiergräber, who permitted me to boundary-less challenge many aspects of schizophrenia proteinopathy in DISC1. Likewise, I am very grateful to Dr. Renu Batra-Safferling and Dr. Thomas Gensch, both of whom also happen to be my PhD supervisors years ago for their continued motivation & unstinted support. I am highly obliged to Fr. Dorothea Stobbe and Ms. Astrid Wies for attending several tiny bits of issues in the admins.

Many a friend of mine helped reshape my career in the last five and half years, Nadine Jordan, Nachiket Kashikar, Nitin Joshi, David Elemenhorst, Eva-Maria Elemenhorst, Reinhard Seifert, Dagmar Wachten, Kevin Bochinsky, Henrike Heise, Nils Lakomek, Joachim Granzin, Tuyen Bujnicki, Janine Kutzsche, Anne(lisse) Cousin, Markus Tusche, Karin Hänel, Andrew Dingley, Taras Balandin, Sebastian Peuker, Shatanik Mukherjee, Pavlik Lettinga, and Smruti Phadnis for their motivation & unstinted support. All their help & cooperation in the hours of need was unforgettable.

I am grateful to my friends from IPB-HHU and IBI7-FZJ them being, Victoria, Dominik, Nicole, Nici, Nina M, Monika, Raphael, Lea, Bernd, Christian, James, Sabine P, Jolanta, Tugba, Imke, Sabine. We, Matthias, Rudi, Jenny, Birgit, Sabine Wi, Fritz, Volker, Phillip, Lothar, Ian, Esther, Sarah, Liang, Leonardo, Baran, Nina B, Luis, Jorg, Lena, Karoline, Suman, Benedikt, Hebah, Sahin, Mohammed, Robin, Wolfgang, Manuel, Hessam, Mohan, for their constant help, affection & cooperation.

I thankful to my mentor at the HMGU Munich, Prof. Dr. Gabi Kastenmüller and her Sysmet team of Matthias, Werner, Orhan, Bharadwaj, Daniela, Tong, Patrick, Leo, and Maria.

I would like to thank my mentor at the University of South Eastern, Norway, Dr. Veralia Gabriela Sanchez for educating me in social sciences and qualitative analytics, and an amazingly successful collaboration.

Umpteen thanks to the NEMO group of Prof. Dr. Dr. Rene Hurlemann at the Karl-Jaspers-Klinik, Wehnen, who have had a profound impact in my life into the world of schizophrenia. Good old friends from the group Dirk (now in Bochum), Danilo, Simon, and Max; Thanks guys. And my thanks also to Pia, Christina, Sofia, Diana, Marina, Marianne, Vanessa, Marc, Ulrike, and Christine.

A large part of my cross- and multi-disciplinary work + education was promoted and inspired at the Hanse Wissenschaftskolleg (HWK), Delmenhorst. I thank Prof. Kerstin Schill and Dorothe for providing me with the opportunity. I am highly obliged to the staff in Thuriid, Doris, Wolfgang, Michael Ka, Christine F, Christine I, Christine G, Kestin, Sabine, Bijan, Michael Ki, and Klaus. My HWK family educated me tremendously with their expertise, Anja, Mayoko, Andreiy, Larissa, Isaac, Clayton, Michaela, Benoit, Reza, Emmanuella, Alexander, Flora, Ida, Sudheshna, Sanchari, Christian Mo, Christian Fu, Janine, Ignacio, Laurel, Kufre, Jens, and Lucy.

I have been able to navigate this path with the support of my mother Lata, and my father Arun (who recently left us for his heavenly abode) and my younger brother Abhilekh and his wife Prajakta. My cousin Ms. Shwetha Joshi-Kelkar provided me with tremendous amount of support and guidelines. My family from Bhopal in Bharat, Srividya and little Perna. And lastly, I have been very lucky in the last two years with my partner Ms. Isabel Hartmann for supporting my research activity and above all being with me all the while.

And last but not the least I thank and apologies all my other teachers and friends whose name I missed out.

SORRY.



ABHISHEK CUKKEMANE

M.Sc., Ph.D.

CONTACT

CORRESPONDANCE ADDRESS

Düsseldorfer straÙe 17,
D-52428, Juelich, Germany
PHONE: +49-1522-3608739
EMAIL:

Abhishek.cukkemane@gmail.com

a.cukkemane@fz-juelich.de

[abhishek.cukkemane@uni-](mailto:abhishek.cukkemane@uni-duesseldorf.de)

duesseldorf.de

E-PROFILE

WEBSITE:

[https://www.linkedin.com/in/abhishek-](https://www.linkedin.com/in/abhishek-cukkemane-31983057)
[cukkemane-31983057](https://www.linkedin.com/in/abhishek-cukkemane-31983057)

<http://orcid.org/0000-0002-1532-6869>

[https://www.researchgate.net/profile/](https://www.researchgate.net/profile/Abhishek_Cukkemane)
[Abhishek_Cukkemane](https://www.researchgate.net/profile/Abhishek_Cukkemane)

[https://www.fz-](https://www.fz-juelich.de/profile/cukkemane_a)
[juelich.de/profile/cukkemane_a](https://www.fz-juelich.de/profile/cukkemane_a)

EXPERTISE

scientific communication
intellectual property management
ISO Management systems
e-marketing
social networking
data mining & OMICS
analytics
life sciences
pharmaceuticals
food technology
protein biochemistry
structural biology
agronomy & agriculture
research & teaching

WORK EXPERIENCE (ACADEMIA AND START-UPS)

Forschungszentrum Jülich & Universität Düsseldorf Germany; [Research Scientist [Group leader & Business development manager]

Nov 2019–Current

[Biophysical & Biochemical characterization of risk factors in Schizophrenia, bipolar disorders & Chronic mental illness]

Indian Institute of Science Education and Research, Bhopal India

[Post-Doctorate]

Oct 2018–Sep 2019

[Solution NMR spectroscopy & Metabolomics - Dairy, Probiotics & Nutraceuticals]

Cuukky Natural Sciences Pvt Ltd, Pune, India

[CEO & co-founder]

Jul 2018–May 2022

[Activities are related to agri- & food- biotech, future perspective in Dietary therapeutics & pharmaceuticals]

Microbiology Dept, Tuljaram Chaturchand College, Baramati, India

[Assistant Professor]

Dec 2014–Oct 2018

[Visiting Faculty & Research Scientist, Teaching UG, PG & Faculty]

Bijasu Agri Research Laboratory LLP, Pune, India; [CEO & CSO]

Dec 2014–Mar 2017

[With a CAPEX investment of 21 Lakh Rs (30,000 €), I set up an agri-biotech platform, trained team members; supporting internal and external projects; IP Management; networking & collaborations]

Cactus communication; [Scientific Editor, Remote]

Apr 2014–Oct 2014

[Contract based assessment and expertise sharing in subjects of chemical biology, natural sciences, and medicinal sciences for editing, proofreading, client requirements and communication.]

E-PROFILE

WEBSITE:

<https://www.linkedin.com/in/abhishek-cukkemane-31983057>

<http://orcid.org/0000-0002-1532-6869>

https://www.researchgate.net/profile/Abhishek_Cukkemane

https://www.fz-juelich.de/profile/cukkemane_a

EMAIL:

Abhishek.cukkemane@gmail.com
a.cukkemane@fz-juelich.de
abhishek.cukkemane@uni-duesseldorf.de

HOBBIES

Painting
Photography
Swimming
Racket sports

LANGUAGES

Kannada (Native spoken)
English (Fluent-spoken & writing)
Hindi (Fluent-spoken and basic-writing),
Marathi (Basic-spoken)
German (Basic-reading)
French (Basic-reading)

START-UP AWARDS

Ministry of Small Micro & Medium Enterprise: Among top 6 start-ups (Jan 2016)

Dept of Science and Technology-Swissnex; Academia-Industry training: Among top 10 start-ups (Oct 2016)

IP INDIA (REF ID)

[201721016132](#)
[201721016133](#)
[201721016135](#)
[201921015221](#)

EUROPEAN PATENT (REF ID)

EP22180453 (PCT/EP2023/066796)
EP102023122636 (PCT/EP2024/073111)

Utrecht University, Utrecht, The Netherlands; [Post-Doctorate]

Apr 2009–Mar 2014

[Biophysics, solid-state NMR spectroscopy, co-supervising Ph.D candidates, Teaching UG & PG students in chemistry]

Forschungszentrum Jülich & Forschungszentrum Caesar, Bonn, Germany; [Post-Doctorate]

Oct 2007–Mar 2009

Biochemistry, crystallography, spectroscopy, co-supervising Ph.D students]

EDUCATION

University of Cologne & Forschungszentrum Jülich, Germany [Ph.D. Biochemistry; Summa cum Laude]

2004–2007

[Structural & Functional studies of a Prokaryotic CNG Channel]

University of Pune, Pune, India; [M.Sc. Microbiology; O grade GPA 7.3/10]

2002–2004

[Production of bio-emulsifiers from *Halomonas* spp.]

University of Pune, Pune, India

[B.Sc. Microbiology & Honors in Industrial Microbiology; 1st class – 67%]

1999–2002

PROFESSIONAL TRAINING

HIFIS - Helmholtz Federated IT Services; [2021]

[First Steps in R-Programming, Basic statistics, Advanced statistics, multi-variate statistics, R-markdown]

Leadership –Schiel Consulting GmbH, Cologne, Germany; [2020]

[Leadership without disciplinary authority, Fit for Leadership]

Advanced Courses, World Intellectual Property Organization (WIPO), DL-101,201,203,301,302,317,318,320,450,501,502,503,506,511;

2018– 2020; 2017

[Patent (drafting & search); Trademark & GI; Copyright (incl. Software) & CMO; IP & Collective Management, TK & TCE; Mediation & Arbitration]

ISO Management Systems; Lead Auditor- Global Exemplar;

2017– 2018

[ISO 9001:2015 QMS; ISO 14001:2015 EMS; ISO 22000:2005 FSMS]

RELATED PROFESSIONAL ACTIVITIES

[ISO 9001:2015 QMS; ISO 14001:2015 EMS; ISO 22000:2005 FSMS]; [2018]

[Lead Auditor for Zenith Quality Assessors Pvt Ltd., Pune, India]

[ISO 9001:2015 QMS; ISO 14001:2015 EMS] [2018]

[6 months consultancy project at Sadguru Sri Sri Sugar Factory, Sangli, India for incorporating ISO standards, streamlining activity, and minimizing wasteful expenditure]

[ISO 22000:2005 FSMS] [2014]

[3 months consultancy project at IDEA Discovery Lab, Venture Center, Pune, India for Lactic acid bacteria-based technology]

[Scientific Editor] [2014]

[Served role of scientific editor at Cactus global, Mumbai for 3 months for editing manuscripts from non-native English-speaking countries]

AWARDS AND ACHIEVEMENTS

[Biologo Diagnostics-4-future conference- Startup PITCH competition 2024]

[When a vintage classic meets a smart tech: Schizophrenia diagnostics]

[Norwegian Artificial Intelligence Research Consortium - Helmholtz Information & Data Science Academy Trainee Network 2024]

[Using AI-driven models to comprehend the variables associated with schizophrenia management from social and healthcare perspectives]

[Hanse-Wissenschaftskolleg Fellowship] [2023]

[Solving the biochemical jigsaw puzzle of schizophrenia by combining traditional clinical diagnostic practices along with modern day metabolic profiling approaches]

[Biologo Diagnostics-4-future conference- Startup PITCH competition 2023]

[Diagnostics for sleep disorders]

[Forschungszentrum Jülich INNOVATION Start program] [2022]

[Combining NMR metabolomics, big-data, and AI for reliable schizophrenia diagnosis]

[Alexander von Humboldt – Research group linkage] [2022-2024]; Applied jointly with Doc.dr.sc. Dr. Nicholas J. Bradshaw, University of Rijeka, Croatia

[DISC1: Its Structure, Causes of Aggregation and Relevance to Disease]

[Helmholtz Information & Data Science Academy Trainee Network] [2022]

[Combining clinical qNMR metabolomics, data analytics and machine learning approaches as a fast and reliable schizophrenia diagnostic tool]

[Helmholtz Enterprise Field Study] [2020]

[NMR metabolomics]

[Forschungszentrum Jülich fellowship programme for excellent guest scientists] [2019]

[Start-up Awards] [2016]

[MSME (Govt of India); DST-AIT Govt of India-Swiss Embassy]

[Poster Prize, European Biophysics Congress] [2007]

[Ion channel Physiology at EBSA- Imperial College London, UK]

[Indian Academic Scholarships]; [2002-2004]

NCCS-SRFP India; GATE- All India Rank 112; SET-Maharashtra & Goa, India; JNCASR-SRFP India - Bose Res. Inst., Kolkatta, India]

COMPUTER SKILLS

- **Operating systems:** DOS, Windows OS and Debian Linux systems
- **Pro-Software:** Sigma plot, Origin, Pymol, Rasmol, Bruker Topspin (Data Analysis software), MS Office, LaTeX, Structure Calculation (X-plor NIH), and Molecular Dynamic simulation using Gromacs
- **Programming languages:** R and php

TEACHING EXPERIENCE AND QUALIFICATIONS

- **2019-Current**, Heinrich-Heine-Universität, Düsseldorf; V423; Molecular Biophysics: X-ray Structure Analysis
- <https://www.biologiestudium.hhu.de/bachelorstudiengaenge/bachelor-of-science-biologie/modulhandbuecher>
- **2019-Current**, Heinrich-Heine-Universität, Düsseldorf; M-4456 Biomolecular Crystallography
- <https://www.biologiestudium.hhu.de/master-biologie>
- **2014-2018**, visiting faculty and Research scientist at Tuljaram Chaturchand College, Baramati
- **2014-2017**, visiting Lecturer at Savitribai Phule Pune University, Pune, and affiliated colleges viz. Modern College & Abasaheb Garware College
- **2014**, Ph.D. examiner, University of Milan (UDSM), Italy
- **2013**, Bijvoet Symposium, Utrecht, the Netherlands. Poster session – Jury Member 2013: Conducted lectures and practicals for the “Molecular and Cellular life sciences course”
- **Subject:** Basics of solid-state NMR spectroscopy and its application to biomolecular research
- **2009-2013:** Lecturer at the Bijvoet Summer school “Exploring Nature’s Molecular Machines”,

- **Subject:** Membrane proteins studied by Solid-state NMR spectroscopy
- **2004 SET qualified** – Eligible for lectureship of Bachelors and Masters Students in science in the state of Maharashtra (India).

CONFERENCES, WORKSHOPS AND PRESENTATIONS

Oral presentations:

- **Hanse-Wissenschaftskolleg 2024**, Delmenhorst Germany, The two sides of the schizophrenic coin: heads => science, tails => society
- **Helmholtz Innovation Lab 2024**, Potsdam Germany, SuFIDA Innovation lab, From then..... until NOW
- **Helmholtz-Israel Bootcamp 2024**, Berlin Germany, Diagnostic tool-kits for epileptic seizure
- **Karl-Jaspers-Klinik 2024**, Bad Zwischenahn Germany, The biophysics and Biochemistry of schizophrenia
- **Helmholtz Information & Data Science Academy- Norwegian Artificial Intelligence Research Consortium 2024**, Vesftold Norway, Schizophrenia Diagnostics – Market Survey and biomarker analysis
- **Young Entrepreneurs in Science – Boehringer Ingelheim Venture Fund 2024**, Jülich (online), Drug development for schizophrenia
- **Startup transfer.NRW 2024**, Düsseldorf Germany, A start-up for diagnosing sleep disorders for shift work jobs that is deep-learning tunable: Deep sleep
- **Biologo Diagnostics-4-future conference 2024**, Constance Germany, When a vintage classic meets a smart tech: Schizophrenia diagnostics
- **Biologo Diagnostics-4-future conference 2023**, Constance Germany, Diagnostics for sleep disorders
- **University of Rijeka 2022**, Rijeka Croatia, DISC(1)erning the structural & functional attributes of Disrupted-in-Schizophrenia (isoform 1) protein in chronic mental illnesses
- **Biologo Diagnostics-4-future conference 2022**, Constance Germany, The SuFIDA technology: Evolution of ELISA - counting single molecules
- **bio-N³MR 2020**, DISC(1)erning the structural & functional attributes of disrupted-in-schizophrenia (isoform 1) protein in chronic mental illnesses
- **Microbiovision (Savitribai Phule Pune University):** A national level workshop for students. **2017**, Sharadabai Pawar Mahavidyalaya, Barmati, India
- Academia Industry Training, **Department of Science and Technology (DST)**, India and Swissnex, Switzerland, **2016**
- **Binding-Gating relationship:** An ssNMR case study of a cyclic Nucleotide-activated channel." **2013:** Future of Biophysics. Erice, Sicily, Italy
- "Deciphering binding-gating relationship: An ssNMR case study of a cyclic-nucleotide-activated channel." **2012:** 3rd **International Workshop on Expression, Structure and Function of Membrane Proteins.** Florence, Italy- September 23-27.
- **"Protein solid-state NMR and beyond."** **2011:** 3rd EU-CCPN NMR Conference. Oeiras, Portugal- September 28-30.

Poster Presentations:

- **2016**, 3rd International Conference on "Herbal and Synthetic Drug Studies (HSDS-2016). Abeda Inamdar, Pune, India, Participation
- **2013**, Future of Biophysics, Erice, Italy. "Ligand-binding domain displays conformational plasticity that is essential for gating of a cyclic nucleotide-activated K⁺-channel in lipid membranes"
- **2012**, ICMBRS, Lyon, France. "Cyclic AMP regulation of a tetrameric-160 kDa cyclic-nucleotide-regulated K⁺-channel in lipid membranes as seen by multidimensional solid-state NMR
- **2012**, Bijvoet Symposium, Utrecht, the Netherlands. "Probing multiple conformations of cyclic-nucleotide-binding domain by solid-state NMR."
- **2011**, CHAINS (Chemistry as Innovating Science), Utrecht, the Netherlands. "Probing multiple conformations of cyclic-nucleotide-binding domain by solid-state NMR."
- **2011**, 7th Alpine Conference on solid-state NMR, Chamonix Mount-Blanc, France. "Probing multiple conformations of cyclic-nucleotide-binding domain by solid-state NMR."
- **2011**, New Frontiers in Ion channel Physiology, Hamburg-Blankenese, Germany. "Structural Polymorphism correlates with functional landscape of a cyclic-nucleotide binding-domain."
- **2011**, Bijvoet Symposium, Utrecht, the Netherlands. "Structural Polymorphism correlates with functional landscape of a cyclic-nucleotide binding-domain."
- **2010**, EUROMAR/ISMAR, Florence, Italy. "Structural characterization of a cyclic-nucleotide regulated K⁺-channel using solid-state NMR Spectroscopy."
- **2010**, Bijvoet Symposium, Utrecht, the Netherlands. "Structural characterization of a cyclic-nucleotide regulated K⁺-channel using solid-state NMR Spectroscopy."
- **2009** Membrane and Modules, Berlin, Germany. "Structural characterization of a cyclic-nucleotide regulated K⁺-channel using solid-state NMR Spectroscopy."
- **2009** NWO-CW Annual meeting, Veldhoven, the Netherlands. "Structural characterization of a cyclic-nucleotide regulated K⁺-channel using solid-state NMR Spectroscopy."
- **2009** Bijvoet Symposium, Utrecht, the Netherlands. Participation
- **2008** EMBO electron microscopy and Crystallography, Gif-sur-Yvette, Paris, France. Participation
- **2007** European Biophysics Congress, London, UK. Poster Prize. "Functional characterization of a prokaryotic cyclic nucleotide-gated channel."

PUBLICATION LIST

Scientific articles in preparation (*corresponding author)

- **Cukkemane A***, Bochinsky, K, Lange D, Hennecke E, Foerges AL, Beer S, Bauer A, Aeschbach D, Willbold D, Kastenmüller G, Drzezga A, Jordan J, Elmenhorst D, Elmenhorst EM (IN PREPARATION) The Human Metabolome Quantified During Sleep Loss
- Bujnicki T, Cousin A, Schwarten S, Hänel K, Rehn F, König B, Victoria Kraemer-Schulien V, Pils M, Blömeke L, Bannach O, Cukkemane A, Willbold D (IN PREPARATION) Digital direct detection (ThreeD) of single biomolecules revealed femtomolar sensitivity
- **Cukkemane A***, Postin D, Bochinsky K, Scheele D, Willbold D, Hurlmann R (IN PREPARATION) Metabolic analysis of blood biomarkers reveal lacunae in clinical diagnosis of schizophrenia

Scientific articles and communications (*corresponding author)

- **Cukkemane A***, Sanchez, VG (ACCEPTED) Mapping emotional perceptions from stakeholders' survey using natural language processing in the management of chronic mental illnesses – perspectives from qualitative analytics. Journal of Affective Disorders Reports
- **Cukkemane A***, Becker N, Kupreichyk T, Heise H, Willbold D, Weiergräber OH (2025) Snapshots of the aggregation pathway of the scaffold protein DISC1 and the implications for chronic mental illnesses: structure mirrors function. Journal of Structural Biology: X. <https://doi.org/10.1016/j.yjsbx.2025.100128>
- **Cukkemane A***, Dingley AD, Mohrlüder J, Santiago-Schübel B, Weiergräber OH, Willbold D (2025) DISC1 mutants are functionally active but structurally unstable to choreograph their tasks. European Journal of Pharmaceutical Sciences. <https://doi.org/10.1016/j.ejps.2025.107148>
- **Cukkemane A***, Hurelemann R, Sanchez, VG (2025) Qualitative analytics from stakeholder perspective describes the biopsychosocial model for patient-centred approach to managing psychiatric ailments. Journal of Affective Disorders Reports. <https://doi.org/10.1016/j.jadr.2025.100893>
- **Cukkemane A***, Becker N, Zielinski M, Frieg B, Lakomek NA, Heise H, Schröder GF, Willbold D, Weiergräber OH. (2021) Conformational heterogeneity coupled with beta-fibril formation of a scaffold protein involved in chronic mental illnesses. Transl Psychiatry 11, 639. <https://doi.org/10.1038/s41398-021-01765-1>
- **Cukkemane, A.***, Kumar, P., and Sathyamoorthy, B. (2020) Understanding the role of synbiotics in functional foods and dietary therapeutics for health, communicable and non-communicable diseases using metabolomics footprint studies. Food Research International. <https://doi.org/10.1016/j.foodres.2019.108679>
- Rane, J., Singh, A.J., George, P., Govindasamy, V., **Cukkemane, A.**, Raina, S.K., Chavan, P.M., Aher, L., V. S. Sunoj, J., Singh, N.P. (2019) Effect of Cow Urine-Based Bioformulations on Growth and Physiological Responses in Mungbean Under Soil Moisture Stress Conditions. PNAS, India Section B: Biological Sciences pp1-11 <https://link.springer.com/article/10.1007/s40011-019-01088-8>
- Pawar, S., Kalyankar, V., Dhamangaonkar, B., Dagade, S., Waghmode, S., and **Cukkemane, A.*** (2017) Biochemical profiling of antifungal activity of betel leaf (Piper betle L.) extract and its significance in traditional medicine. Journal of Advanced Research in Biotechnology. 1(1):4-6 <https://symbiosisonlinepublishing.com/biotechnology/biotechnology16.php>
- Tank, J., Dhamangaonkar, B., Ukale, U.D., Cukkemane, N., **Cukkemane, A.A*** (2017) Biochemical and Microbiological analysis of different organic manures; their effect on germination of Coriandrum sativum (Cilantro) and Solanum melongena (Eggplant). Journal of Bioprocessing and Biotechniques. 7(1) <https://doi.org/10.4172/2155-9821.1000295>
- Tank, J., Ukale, U.D., Mapare, V., Dhamangaonkar, B., Cukkemane, N, **Cukkemane, A.A*** (2017) Metabolic profiling of wastes from raw and ripe mango seeds to comprehend its beneficial medicinal and cosmetic properties for commercialization. Journal of Agricultural Studies. 5(1) <https://doi.org/10.5296/jas.v5i1.10234>
- Ukale, U.D., Bhagwat R.V, Upadhyay S.K, Cukkemane, N, **Cukkemane, A.A*** (2016) Metabolic analysis of liquid formulations of organic manures and its influence on growth and yield of Solanum lycopersicum L. (tomato) crop in field. Biocatalysis and Agricultural Biotechnology 8:50-54 <https://doi.org/10.1016/j.bcab.2016.08.007>
- Kaplan, M., **Cukkemane, A.**, van Zundert, G., C., Narasimhan, S., Daniëls, M., Mance, D., Waksman, G., Bonvin, A., M., Fronzes, R., Folkers, G., E., Baldus, M. (2015) Probing a cell-embedded megadalton protein complex by DNP-supported solid-state NMR. Nature Methods 12(7):649-652 <https://doi.org/10.1038/nmeth.3406>
- Langlois, C., Ramboarina, S., **Cukkemane, A.**, Auzat, I., Chagot, B., Gilquin, B., Ignatiou, A., Petitpas, I., Kasotakis, E., Paternostre, M., White, H.E., Orlova, E.V., Baldus M., Tavares, P., and Zinn-Justin S. (2014) Bacteriophage SPP1 Tail Tube Protein self-assembles into β -structure rich tubes. J. Biol. Chem. [https://www.jbc.org/article/S0021-9258\(20\)49226-9/fulltext](https://www.jbc.org/article/S0021-9258(20)49226-9/fulltext)
- Moiset, G., Lopez, C., Rijpkema, E., Syga, L., Bartelds, R., **Cukkemane, A.**, Baldus, M., Poolman, B., and Marrink, S.J. (2013) "Disaccharides Impact the Lateral Organization of Lipid Membranes." J. Am. Chem. Soc.: 134(46):16167-16175 <https://pubs.acs.org/doi/10.1021/ja505476c>
- **Cukkemane, A.** and Baldus, M. (2013) "Structural Characterization of a cyclic-nucleotide-activated K⁺-channel and its lipid environment using solid-state NMR spectroscopy." Chem. Bio. Chem.: 14(14):1789-1798 <https://doi.org/10.1002/cbic.201300182>
- Crujisen, v.d.E. Nand, D., Weingarh, M., Prokofyev, A., **Cukkemane, A.**, Horning, S., Becker, S., Bonvin, A., M., J., J., Hulse, R., E., Perozo, E., Pongs, O., and Baldus, M. (2013). "The importance of the lipid-pore loop interface for potassium channel structure and function" PNAS 110 (32): 13008-13013 <https://www.pnas.org/doi/10.1073/pnas.1305563110>

- Peuker, S., **Cukkemane, A.**, Held, M., Noe, F., Kaupp, U., B. and Seifert, R. (2013) "Kinetics of ligand-receptor interaction reveals an induced-fit mode of binding in cyclic nucleotide-activated protein." *Biophys. J.* : 104 (1) :63-74
<https://doi.org/10.1016/j.bpj.2012.11.3816>
- Gradmann, S., Ader, C., Heinrich, I., Nand, D., Dittmann, M., **Cukkemane, A.**, van Dijk, M., Bonvin, A., M., J., J., Engelhard, M., and Baldus, M. (2012) "Rapid prediction of multi-dimensional NMR data sets." *J. Biomol. NMR* 54 (4):377-387 <https://doi.org/10.1007/s10858-012-9681-y>
- **Cukkemane, A.**, Nand, D., Gradmann, S., Weingarh, M., Kaupp, B., and Baldus, M. (2012). "Solid-state NMR [¹³C,¹⁵N] resonance assignments of the Nucleotide-Binding Domain of a bacterial Cyclic Nucleotide-Gated channel." *Biomol. NMR assign.* 6 (2):225-229. <https://doi.org/10.1007/s12104-012-9363-4>
- Nand, D., **Cukkemane, A.**, Becker, S., and Baldus, M. (2012). "Fractional deuteration applied to biomolecular solid-state NMR spectroscopy." *J. Biomol. NMR* 52 (2):91-101. <https://doi.org/10.1007/s10858-011-9585-2>
- Zhang, Z., J. Buitenhuis, **Cukkemane, A.**, Brocker, M., Bott, M., and Dhont, J.K. (2010). "Charge reversal of the rodlike colloidal fd virus through surface chemical modification." *Langmuir* 26(13): 10593-9.
<https://pubs.acs.org/doi/10.1021/la100740e>
- **Cukkemane, A.**, Grueter, B., Novak, K., Gensch, T., Boenigk, W., Gerharz, T., Kaupp, U.B., and Seifert, R. (2007). "Subunits act independently in a cyclic nucleotide-activated K⁺-channel." *EMBO Reports* 8: 749-755.
<https://doi.org/10.1038/sj.embor.7401025>

Reviews and Book Chapters (*corresponding author)

- Cukkemane, N. and **Cukkemane A.A.*** (2017) Comprehending the biochemical basis of liquid organic manure to promote mass production by combining livestock and organic farming practices for rural development. The role of Vrikshayurveda and Traditional practices in Organic agriculture. Thakur, R.P. and Choudhary, S.L. (eds) 84-91
- **Cukkemane, A.*** (2016) Socio-Commercial Agri-Biotech Model for Rural Development in India by Combining Livestock and Organic Farming Practices. *Journal of Advanced Research in Biotechnology.* 1(1):4-6
<https://symbiosisonlinepublishing.com/biotechnology/biotechnology12.php>
- **Cukkemane, A.A.**, and Bhatevara, S. (2016) Back to our roots; Ayurvedic organic farming by scientific principles. LAP Lambert Academic Publishing. ISBN 978-3-330-00504-4
- **Cukkemane, A.A.**, Renault, M., and Baldus, M. (2012). Solid-State NMR on Membrane Proteins: Methods and Applications. In: *NMR of Biomolecules: Towards Mechanistic Systems Biology*, Wiley Blackwell, ISBN 978-0-470-74608-0, pp. 407-418.
- Ader, C., Nand, D., Renault, M., **Cukkemane, A.A.**, and Baldus, M. (2011). "Solid-State NMR on Membrane proteins: Methods and Applications." *Advances in Biomedical Spectroscopy: Biomolecular NMR spectroscopy*, IOS Press, ISBN 978-1-60750-694-2, 3: 115-136
- **Cukkemane, A.**, Seifert, R., and Kaupp, U.B. (2011)". Cooperative and uncooperative cyclic-nucleotide-gated ion channels." *Trends Biochem. Sci.* 36(1): 55-64. <https://doi.org/10.1016/j.tibs.2010.07.004>
- Renault, M., **Cukkemane A.**, and Baldus, M. (2010). "Solid-state NMR spectroscopy on complex biomolecules / Festkörper-NMR-Spektroskopie an komplexen Biomolekülen" *Angew. Chem. Int. Ed. Engl.* 49(45): 8346-57.
<https://onlinelibrary.wiley.com/doi/10.1002/ange.201002823>