HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Investigations of the structural organization of the Disrupted-in-Schizophrenia 1 (DISC1) protein, a major risk factor for mental illness

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences

at the Heinrich Heine University Düsseldorf

presented by

Antony Sravan Kumar Yerabham

from Kurnool

Düsseldorf, June 2017

from the institute for Institute for Neuropathology

at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at

Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Carsten Korth

Co-supervisor: Prof. Dr. Georg Groth

Date of the oral examination: 31st July 2017

Contents

Figures

Abstract

Major mental illnesses such as schizophrenia, major depression, bipolar disorder are devastating conditions affecting adaptive behavior and personality of an individual. Both genetic pre-disposition and environmental factors have been shown to be causal for mental illnesses. Nevertheless, the exact molecular etiology of these complex neuropsychiatric conditions remains unclear.

Disrupted-in-schizophrenia 1 (*DISC1*) is an important mental illness genetic risk factor. It was discovered at a genetic loci disrupted by a balanced chromosomal translocation (1; 11) (q42.1; q14.3) and showing strong linkage with mental illnesses in a large Scottish family. The gene was also reported to have a frame-shift mutation, thereby shortened, in an American family with mental illness. The DISC1 protein interacts with a large number of binding partners; many of them being themselves crucial players in various neuronal signalling pathways or brain diseases. Given the high neurological relevance, understanding of DISC1's functional mechanism is still only vague and one of the reasons for this is the lack of sufficient structural information. In this cumulative thesis, comprising three publications, I attempt to provide new insights of the DISC1 protein's structure and thereby its potential molecular actions by a functional-module based approach.

In the first manuscript, the hypothesis is put forward that DISC1 could be a scaffold protein due to its large protein-interaction network, therefore its features are compared to that of other well-known scaffolding proteins (such as the Ste5 and the AKAP proteins). From the available data, it was here found that DISC1 has many, but not all, the features of a scaffold protein. It is also observed that DISC1 is a 'Diverse Inhibitor of Signalling Cascades' as it regulates a wide range of enzymes of interrelated signalling cascades, mostly through inhibition. Finally, DISC1 acts in a similar manner to other scaffold proteins by orchestrating different signalling pathways in a modular fashion. Thus, it also potentially has multiples of these modules or localisation regions in various arrangements for efficient functioning.

In the second manuscript, studies on the DISC1 protein sub-domain structure were performed. This has previously proven difficult due to the high insolubility and aggregation propensity of the full length recombinant protein when expressed in bacteria. Here we employed a high-throughput Expression of Soluble Protein by Random Incremental Truncation (ESPRIT) technique to identify the structured regions within the human DISC1 protein, by assessing the solubility of tens of thousands of randomly truncated fragments of recombinant DISC1. Based on the assumption that structured regions in a protein are responsible for its functions, we identified and characterized four structured regions within the DISC1 protein, which we proposed to form the basic domain structure. These were named as D, I, S and C and are found at amino acids 257-383, 539-655, 635-738 and 691-836 respectively. The newly proposed domain structure gives a coherent explanation for the effects caused in the truncation caused by the Scottish chromosomal translocation. Based on this domain information, a mimic of the frame-shift mutant of DISC1 C region was characterized and found to form aberrantly multimeric and aggregated complexes with an unstable secondary structure. Thus, the newly proposed soluble regions of DISC1 protein provides a powerful platform for further investigations of the structure and function of the protein.

The third manuscript concerns the generation and characterization of a camelid single-chain antibody against the C-terminus of the DISC1 protein. This was further used in combination with the

recombinant DISC1 C region (described in manuscript 2), to obtain a structural envelope by small-angle X-ray (SAXS) analysis and partially validate the *ab-initio* modeling of the C region. From our model, several insights into structure of the DISC1 C region could be hypothesized relating to its functioning. This camelid antibody has a high potential to serve as a useful tool for a wide range of applications and further investigations of the DISC1 protein.

In conclusion, this thesis provides a functional-module based approach to investigate the structure of the DISC1 protein. The results obtained here gave novel insights into the structural organisation of the protein and hint towards its probable mode of functioning. The domain architecture obtained and the camelid antibody generated can facilitate further module based investigations and ultimately pave the way towards obtaining an experimental structure of DISC1 and understanding of its functional mechanism.

Zusammenfassung

Mentale Erkrankungen wie Schizophrenie, Depressionen und bipolare Störungen sind Krankheitsbilder mit verheerenden Auswirkungen auf das Verhalten und die Persönlichkeit eines Menschen. Obwohl die genaue Ätiologie dieser komplexen neuropsychiatrischen Erkrankungen bis heute unklar ist, sind bekanntermaßen sowohl eine genetische Prädisposition als auch Umweltfaktoren bei deren Entstehung von Bedeutung.

DISC1 ist einer der am besten untersuchten genetischen Risikofaktoren mentaler Erkrankungen. Es wurde erstmals an einem von einer balancierten chromosomalen (1; 11) (q42.1; q14.3) Translokation betroffenen genetischen Locus in einer schottischen Familie entdeckt und zeigt eine starke Assoziation mit verschiedenen mentalen Erkrankungen. Weiterhin wurde eine *Nonsense*-Mutation in einer amerikanischen Familie beschrieben, welche zu einem C-terminal verkürzten Protein führt. Bislang gibt es keine Evidenz für eine enzymatische Aktivität des DISC1-Proteins, allerdings interagiert es mit einer Vielzahl für neurologische Prozesse wichtiger Proteine. Trotz dieser neurologischen Relevanz ist die Funktion von DISC1 noch immer ungeklärt, und einer der wichtigsten Gründe dafür ist der Mangel an ausreichenden strukturellen Informationen. In dieser kumulativen Doktorarbeit, die auf drei wissenschaftlichen Publikationen beruht, möchte ich neue Einsichten bezüglich der Proteinstruktur von funktionalen Domänen des DISC1-Proteins (und damit deren Bedeutung für Funktion und Dysfunktion) darlegen.

Im ersten Manuskript werden die Eigenschaften des Gerüstproteins DISC1, das über ein weitreichendes Interaktionsnetzwerk verfügt, mit denen anderer bekannter Gerüstproteine, wie Ste5 und AKAP, verglichen. Anhand der verfügbaren publizierten Daten konnte im Übersichtsartikel gezeigt werden, dass DISC1 viele, aber nicht alle, Eigenschaften eines Gerüstproteins aufweist. Es wurde gezeigt, dass DISC1 als ein '*Diverse Inhibitor of Signalling Cascades*', also ein 'vielseitiger Inhibitor von Signalkaskaden' fungiert, welcher eine Vielzahl von Enzymen diverser zusammenhängender Signalkaskaden reguliert, meist durch deren Inhibition. DISC1 agiert vergleichbar mit anderen Gerüstproteinen, indem es verschiedene Signalkaskaden auf eine modulare Weise beeinflusst. Demnach hat DISC1 auch diverse Domänen, die in verschiedenen Gruppierungen für effiziente Funktionalität sorgen.

Im zweiten Manuskript habe ich nach technischen Lösungen gesucht, die eine effiziente Aufreinigung und Charakterisierung des in hohem Maße unlöslichen und aggregierenden DISC1 in Bakterien ermöglichen. Es wurde die Hochdurchsatzmethode *'Expression of Soluble Protein by Random Incremental Truncation* (ESPRIT)' Technik angewandt mit dem Ziel, strukturierte Regionen innerhalb des humanen DISC1 Proteins zu identifizieren. Dafür wurde zunächst die Löslichkeit von Zehntausenden zufällig trunkierten DISC1-Proteinfragmenten analysiert. Basierend auf der generellen Annahme, dass strukturierte Regionen innerhalb eines Proteins für dessen Funktion verantwortlich sind, wurden vier strukturierte Domänen innerhalb von DISC1 identifiziert und

charakterisiert. Aus den Daten geht hervor, dass diese eine grundlegende Domänenformation bilden. Die vier strukturierten Regionen wurden D, I, S und C genannt und umfassen die Aminosäuren 257-383, 539-655, 635-738 und 691-836. Diese vorgeschlagene Domänenstruktur liefert eine logische Erklärung für die durch die chromosomale Translokation in der schottischen Familie versursachten Effekte. Basierend auf den Domäneninformationen wurde auch eine Nachbildung der Leserahmen-Mutation innerhalb der DISC1-C-Region charakterisiert, welche abberante multimere und aggregierte Komplexe mit einer instabilen Sekundärstruktur bilden. Die von uns vorgeschlagenen löslichen Domänen von DISC1 liefern demnach eine wichtige Plattform für weitergehende Untersuchungen der Struktur und Funktion des Proteins.

Das dritte Manuskript beschäftigt sich mit der Generierung und Charakterisierung eines rekombinanten kameliden Antikörperfragmentes, das gegen den C-Terminus von DISC1 gerichtet ist. Dieser Antikörper wurde in Kombination mit der DISC1-C-Region benutzt (Beschreibung der C-region befindet sich im zweiten Manuskript), um strukturelle Hinweise via der *'small-angle X-ray scattering'* (SAXS)-Analyse zu erhalten und dadurch teilweise das *ab-initio* Modell der C-Region zu validieren.

Ausgehend von unserem Modell kann nun über verschiedene potentielle Einsichten bezüglich der Struktur und Funktion des DISC1-Proteins spekuliert werden. Der kamelide Antikörper birgt ein hohes Potential, um als nützliches Werkzeug für eine Vielzahl an Applikationen und weiterführender Untersuchungen des DISC1-Proteins zu dienen.

Zusammenfassend liefert die hier vorgestellte kumulative Doktorarbeit einen Ansatz zur Untersuchung des DISC1-Proteins, der auf dessen funktionalen Domänen basiert. Die hier vorgestellten Ergebnisse führen zu neuen Einsichten bezüglich der strukturellen Organisation des Proteins und deuten auf mögliche Funktionsweisen des DISC1-Proteins hin. Die erlangte Domänenarchitektur und der generierte kamelide Antikörper können weitergehende Modul-basierte Untersuchungen vereinfachen und auf diese Weise den Weg zu einer experimentellen Struktur des DISC1-Proteins und dem Verständnis seines funktionellen Mechanismus ebnen.

1. Introduction

1.1. What are mental illnesses

A mental illness can be defined as a behavioral condition that clearly causes dysfunction, suffering or deterioration of the regular life of an individual. Such a condition may be persistent or episodic, that is that it may either relapsing or single time. The signs and symptoms widely vary with the specific kind of mental illnesses, as different aspects of human behavior or personality become altered.

The behavior types affected can include anxiety, emotion or mood, perception abilities, eating behavior, sleep cycle, substance abuse, to name a few. Abnormalities in each of these behaviors are further categorized into several disorders based on the intensity and complexity. For example, anxiety related abnormal behaviors can be categorized as anxiety disorders, which may include specific phobias, generalized excessive anxiety, panic disorders etc. Mood related disorders generally involve an intense or sustained emotion, such as sadness or despair, and include clinical major depression, bipolar disorder, dysthymia and mania. Perception related disorders involve hallucinations, delusions or thought disorders these include schizophrenia and other schizoaffective disorders. Thereby the other behavior alterations also involve specific dysfunctions associated with them and could be categorized as individual disorders or in combination with several other disorders.

1.2. Implication of genetics in mental illnesses

Major mental illnesses such as schizophrenia, bipolar disorder and major depression are complex and devastating neuropsychiatric conditions with variable clinical symptoms.The etiology of these complex conditions remain largely unknown, nevertheless a large number of family, twin and adoptions studies in small isolated populations and very large populations suggest a strong role of genetics in these illnesses (*1-3*).

Many of the genes, such as Brain-derived neurotropic factor (*BDNF*), Glutamate metabotropic receptor 3 (*GRM3)* or Glutamate ionotropic receptor NMDA type subunit 2A *(GRIN2A),* identified during large scale genomic studies were already known or later found to play important roles in neurodevelopment and brain function (*4, 5*). Some of the genes obtained from genomic surveys were coding for voltage-gated calcium channel subunits for example *CACNA1C, CACNB2, CACNA1I* which are crucial in signalling pathways, such as glutamate signalling, and are implicated in diseases such as schizophrenia (*5*). Even some genes whose products are targets for validated antipsychotic drugs, for example Dopamine receptor D2 (*DRD2)*, have also been identified in these genomic studies (*6*). Considering all this information, the genetic studies are generally reliable for the predisposition of

psychiatric conditions at a broad functional level. In addition to the role of different genetic loci, the effect of different genetic polymorphisms, SNPs and mutations of these genetic loci have also been highlighted through genome wide studies for psychiatric illnesses (*7*).

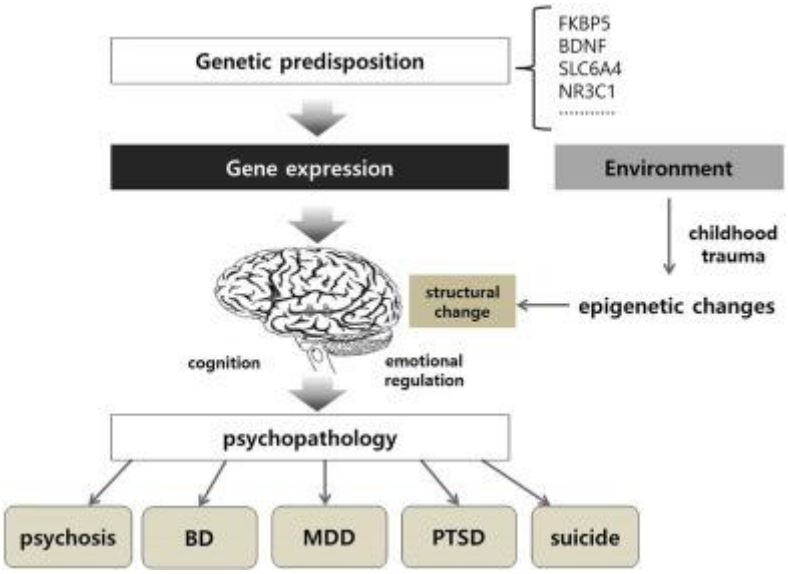

Figure 1: Effects of genetic predisposition and the environmental triggers on brain and behaviour. (Taken from Kim JS and Lee SH, 2016 (*8*))

Expression of genes which influence brain development and function are themselves influenced by environmental factors. These gene-environment interactions trigger structural changes in brain that also cause changes in cognition and emotional regulation.

1.3. Relevance of DISC1 to mental illnesses

Among the various genes implicated in psychiatric illness, Disrupted-in-schizophrenia 1 (*DISC1*) is one of the few highly investigated genes in the field. Thus, it is a promising candidate to understand the biological mechanisms underlying these conditions.

Identification of the *DISC1* gene occurred through strong genetic linkage of a balanced chromosomal translocation (1;11) (q42.1;14.3), which disrupted the *DISC1* genetic locus, with major mental illness in a large Scottish family (*9, 10*). Further research on this family, after the discovery of the translocation, strongly suggested co-inheritance of psychological afflictions with the mentioned translocation (*11, 12*). Later, a smaller American family were discovered to have a frame-shift mutation resulting in the deletion at the C-terminal of the DISC1 protein following amino acid 807, which therefore shares similarity with the Scottish pedigree and was linked to mental illness in that family (*13*). It was observed that the frame-shift-derived mutant form of DISC1 is involved with synaptic vesicle deficits; by depletion or

dysregulation of several synapse-related genes expression in human forebrain neurons (*14*). Linkage analysis of the *DISC1* locus to schizoaffective disorders was also observed in families from Finland, Ireland and the United Kingdom and resulted in providing supporting evidence for association of various *DISC1* single-nucleotide polymorphisms (SNPs) with schizophrenia, depression, autism and Asperger syndrome (*15, 16*). Also, some variants showed association with brain anatomical related and other phenotypes, including hippocampal grey matter volume, poor concentration, memory related issues, visuospatial ability, psychomotor processing and general cognitive ability (*17, 18*). The volume of the hippocampus has been reported in several studies to be reduced in patients with schizophrenia, implicating this region in the aetiology of this disease (*19*).

Transgenic mouse models with genetic alterations in *DISC1* exhibit several measurable phenotypical changes compatible with disturbed behavioural adaptation (*20*). Changes in dopamine levels, sensitivity and neuroanatomy were also observed in these DISC1 mutant mouse models (*21*). Emerging evidence has also been obtained through all these genetics and behavioural analyses, supporting the notion of role of DISC1 functioning and influence in phenotypes related to several mental illnesses.

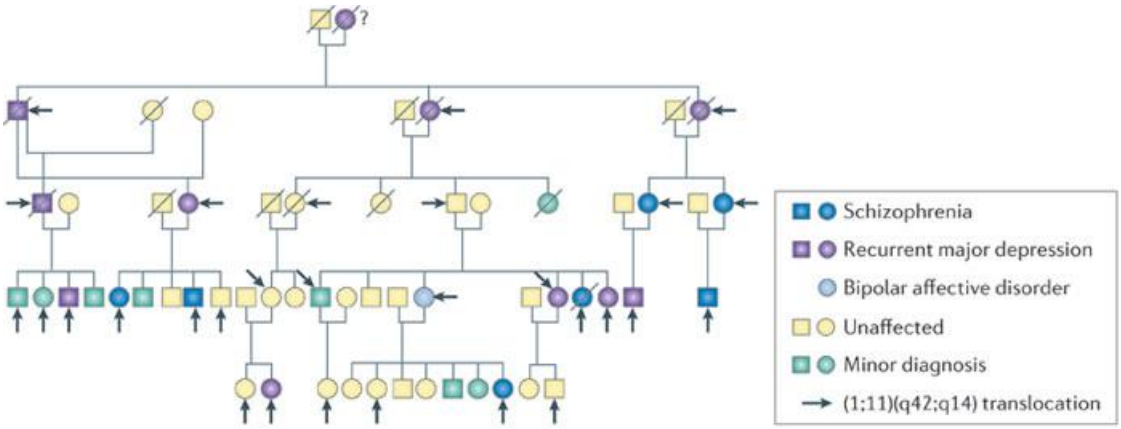

Figure 2: The discovery of the disrupted in schizophrenia 1 (DISC1) pedigree in a Scottish family. (Taken from Brandon and Sawa, 2011 (*22*))

1.4. Expression, interactions and functions of the DISC1 protein

The *DISC1* gene is ubiquitously expressed, as determined through northern-blot analysis, nevertheless many studies also showed a selective pattern of expression dependent on time and space (*10, 23, 24*). The protein is sub-cellularly located to the nucleus, centrosome, cytoplasm, axons and synapse, while mitochondrion is shown as the predominant site for endogenous DISC1 expression (*25-27*). At least two isoforms of DISC1 occupy internal mitochondria locations (*28*). The existence of many splice variants of DISC1 also leads to the

hypothesis that presence or absence of certain domains within the protein would influence the function and also the localisation of DISC1 (Table 1).

Cellular location	Associated functions	References
Centrosome	Involved in recruitment of kendrin, dynein, and dynactin subunits, LIS1, NDEL1, PCM1, ninein, and CAMDI to the centrosome. Also interacts here with PDE4B, PDE4D, NDE1, and BBS4. These interactions include those important in microtubule aster formation, neurite outgrowth, and neuronal migration.	(*29*), (*30*)
Cytoskeleton	Seen along both actin filaments and microtubules. Expression of truncated DISC1 leads to a disorganized microtubule network.	(*27*), (*26, 31*)
Growth cones	Found here in the hippocampus and is involved in recruiting proteins including LIS1, NDEL1, 14-3-3ε, and girdin via kinesin-based transport along the cytoskeleton. Also important for axonal elongation.	(*27, 32*)
Membranes	Found in membrane fractions where it interacts with APP, which is important in cortical precursor formation.	(*24*), (*33*), (*34*)
Mitochondria	Found at mitochondria present on microtubules and regulates their trafficking. Can cause mitochondria to form "ring" structures. Interacts with mitofilin within the mitochondria and is required for correct electron transport chain, monoamine oxidase, and Ca^{2+} activity.	(*25*), (*35*), (*36*)
Nucleus	Found at the chromatin, promyelocytic leukemia bodies, and outer layers of nuclear membrane. Represses ATF4 transcriptional activity and transcription of N-cadherin and alters sleep homeostasis in *Drosophila*.	(*25, 37, 38*)
Synapse	Seen at the postsynaptic density along with PDE4 and NDE1. Interacts with and inhibits TNIK here, leading to degradation of key synaptic proteins. Also affects dendritic spine formation through modulation of PSD-95/Kal-7/Rac1 complexes here.	(*39*), (*40*), (*41*)

Table 1 : DISC1 protein subcellular localization **(taken from Soares et al, 2011)**

Till date, DISC1 is not shown to possess any enzymatic activity on its own; however it possesses a large network of protein interactions and is also considered to act as a molecular hub or scaffolding protein guiding the protein-protein interactions (*42*). The extensive and highly connected DISC1 protein-protein network, sometimes referred to as the DISC1 "interactome", consist of more than 100 interaction partners harbouring proteins involved processes like neuronal migration, neurite outgrowth, signal transduction, cyclic adenosine monophosphate (cAMP) signalling, cytoskeletal modulation and translational regulation (*43, 44*). The connection between DISC1 and neurodevelopment through the processes such as neuronal proliferation and migration is of special interest, as conditions such schizophrenia are believed to result from neurodevelopmental disturbances (*45, 46*).

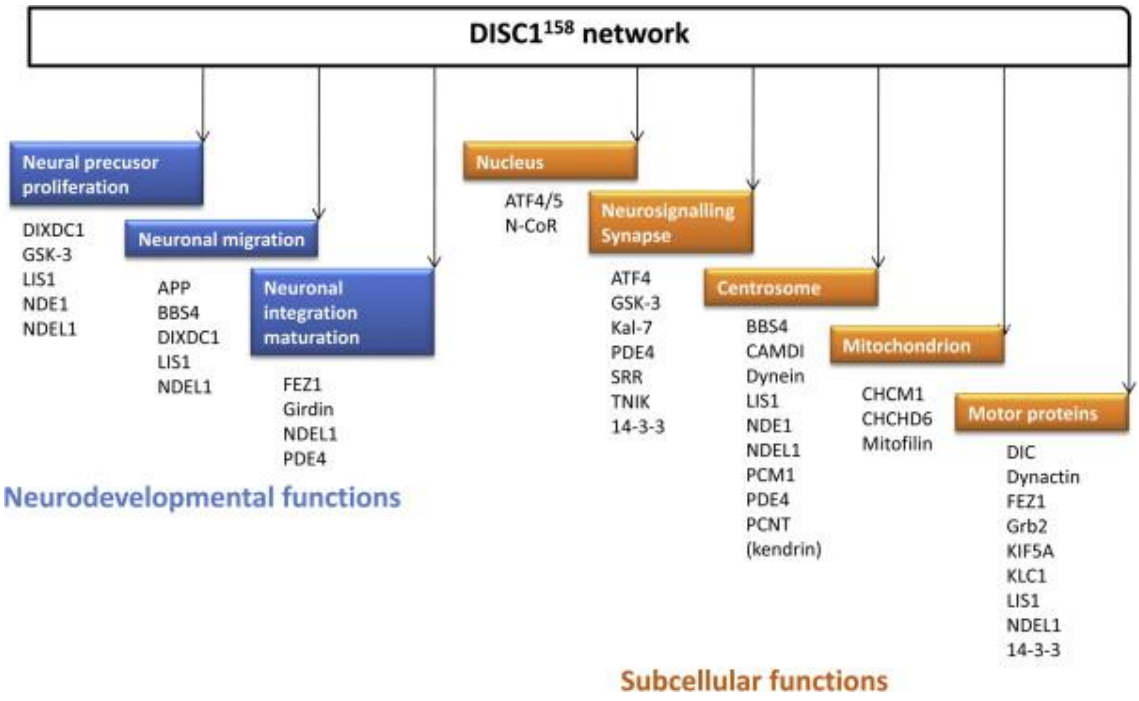

Figure 3: Schematic depicting DISC1 network and interactions by its functions. **(Taken from Lipina and Roder, 2014 (*20*))**

A lot of information regarding DISC1 protein functioning and expression has been obtained by using mice, rat, drosophila and zebrafish as animal models. In zebrafish, DISC1 is involved in forebrain development and regulates the tempo of neuronal integration into the brain, while guiding the positioning of new neurons (*47*). Adult mouse brain shows a restricted pattern of DISC1 expression, identified through in situ hybridization. Highest expression levels were seen in dentate gyrus of the hippocampus and lower expression in CA1-CA3 of the hippocampus, cerebellum, cerebral cortex and olfactory bulbs (*48*). High expression of DISC1 is observed during critical periods of brain development, particularly in the embryonic ventricular and sub-ventricular zones of the cortex, where neural progenitor cells are found (*49*). High expression of DISC1 is also seen in the hippocampus during development and the levels remain high in adult dentate gyrus and olfactory bulb, where the adult neurogenesis occurs. This suggests DISC1 is a critical regulator of both embryonic and adult neurogenesis (*50, 51*).

Mechanistic viewpoints to explain the role of DISC1 in neurogenesis have linked it to the Wnt signalling pathway (*52*). The changes resulting are similar in forebrain gene expression after loss of either DISC1 or Wnt8b function. It was also observed that the loss of gene expression due to non-functioning DISC1 was restored by employing a Gsk3β inhibitor; furthermore an inducible ß-catenin protein restores the DISC1 loss of function phenotype. In frame with these findings, a direct interaction between DISC1 and Gsk3β was observed in the late gestation mouse embryos (*53*). Research with the DISC1 variants cause a loss of function in Wnt/Gsk3ß signalling and disrupt brain development. The direct binding of the N-terminus of DISC1 protein to Gsk3ß stabilized ß-catenin levels, while down-regulation of DISC1 expression decreased ß-catenin levels and opposite effects were seen while overexpressed. Thus it can be interpreted that the cellular and behavioural results due to DISC1 knockdown

can be counteracted by hyperactivation of the canonical Wnt signalling pathway. So DISC1 acts as a crucial regulator of neural progenitor proliferation acting through positively modulating the Wnt signalling pathway by inhibition of Gsk3ß.

DISC1 is also implicated in other neurological processes including neuronal migration (*54*), neuronal outgrowth, axon formation, neuroplasticity and differentiation. It can be easily theorized that the DISC1 protein's involvement in the above mentioned processes would occur through its interaction with the microtubule/centrosome cascade components. There are number of centrosome and cytoskeletal proteins identified to bind DISC1, including PCM1, LIS1, MIPT3, MAP1A, dynactin, NDEL1 and NDE1 (*24*). Interaction of DISC1 was also observed with transport proteins such as α-actin2, β4-spectrin and FEZ1 (*26*). Furthermore, a lot of data suggest the presence of DISC1 at synapses principally at the post synaptic side and regulates synaptic plasticity (*41*).

Behaviour studies with animal models having the genetic alterations of DISC1 were consistent in displaying overlapping behavioural abnormalities. In addition to behavioural deficits, alteration in neurotransmission, neurochemical and signalling cascades are also observed also associations with the dopamine homeostasis (*21*). Given all this evidence, DISC1 clearly has a significant role in healthy functioning of the brain.

1.5. The DISC1 protein sequence, structural predictions and aggregation

In humans, the full length DISC1 protein is 854 amino acids long, with orthologues of it having also been reported in various vertebrate species (*28*). Nevertheless, the homology of the protein to other proteins of known function is relatively low, thereby limiting our ability to identify the functional domains by sequence comparison. It also appears that the DISC1 protein has evolved rapidly, having remarkably low sequence conservation among its orthologues. It possesses a very high percentage of amino acid composition for specific residues, namely serine, alanine and glycine in localised stretches within the protein, further obscuring its origins (*28, 55*).

Structural information of a protein is an important prerequisite to decipher its functional mechanism. In the absence of experimental data on the protein structure, bioinformatic structural predictions were made from the protein sequence of DISC1. From these bioinformatic predictions, it has been suggested that the N-terminal 325 amino acids of DISC1 would be disordered, while the rest of the protein consisting of α-helical or coiled-coil domains (*56*). Proteins having long α-helices can wrongly be predicted to form coiled coils, such as the C-terminal part of the UvrB protein which gives high prediction scores to exist as coiled coils but was shown to contain two α-helices with an antiparallel hairpin fold with left-hand twist (*57*). Such a structure is known as a UVR domain (*58*). Within the DISC1 protein, after extensive sequence analyses, two of these UVR domains (at residues 343-394 and 574-625) have been predicted (*55*).

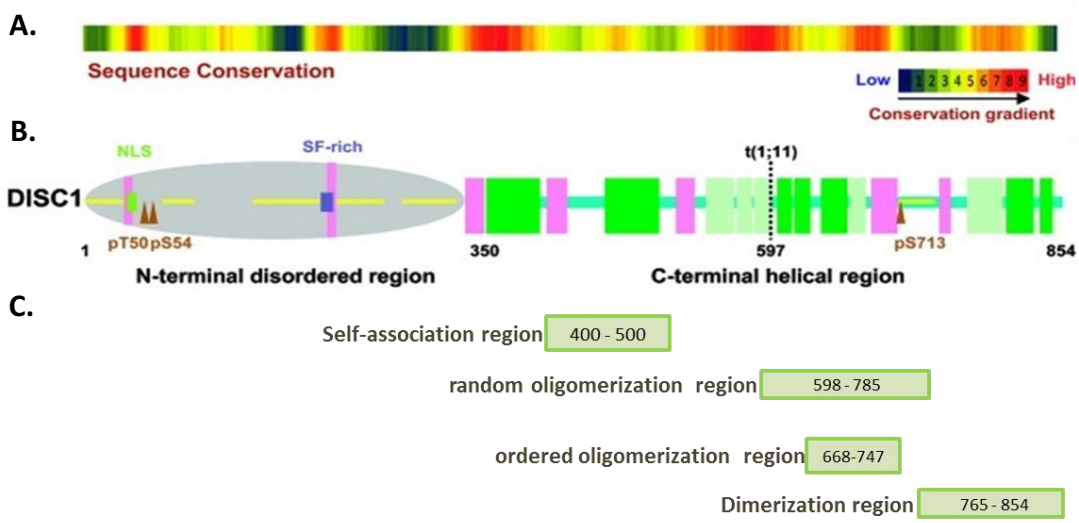

Figure 4: DISC1 protein schematic showing the disorder and helical content according to bioinformatics predictions. (adapted from Soares *et al*, 2011 (*59*))

A. The bar corresponds to the sequence conservation across the length of the protein.
B. The schematic of DISC1 protein structure prediction shows disorder by the oval shape and helical structure by bars.
C. The regions on the protein shown to possess self-association.

So far, the experimental delineation of the DISC1 protein structure has also been slow. Limited information could still be obtained, such as a self-association domain around the residues 400-500, which was deduced through mutational analysis (*31*). Certain characterisation studies also revealed regions in the C-terminal part of protein to form high-order oligomers, such as protein fragments comprising residues 598-785 which were shown to readily appear as oligomers and also control cell-invasiveness of the protein (*60, 61*). Similarly, in another study, residues 668-747 were identified to be responsible for the orderly assembly of DISC1 oligomers (*62*). Studies on the full length recombinant DISC1 protein provided evidence that it exists as octamers and dimers. Nevertheless, these characterisations have been carried on refolded protein or on ones fused to solubility tags such as the maltose binding protein (MBP) tag. The use of such a tag was necessary owing to the low solubility and high aggregation propensity of the recombinant DISC1 protein when expressed in bacteria. Thus the limited solubility of DISC1 remains as a narrow bottleneck to research and seriously limits a thorough characterisation of the protein for both structure and function

Given the low solubility of DISC1 when expressed recombinantly in bacteria, its aberrant multimerization in-vivo leading to aggregation, which has been demonstrated to induce aberrant behavior (*21, 63*). Nevertheless, proper or ordered multimerization of the DISC1 protein is crucial for its functioning. The binding of NDEL1 protein takes place only to octamers of DISC1 but not to its dimers or other higher-order oligomers, indicative of the

need for a specific assembly or quaternary structure of DISC1 (*56*). Hypothetically, multiple factors could be able to influence in the oligomerization of DISC1 such as posttranslational modifications, sequence variations and proteolytic activities. When DISC1 was transiently overexpressed in human neuroblastoma cells, DISC1 formed aggregates and was insoluble in the presence of ionic detergents. Further work then showed that such ionic detergent resistant aggregates of DISC1 protein could be detected in the brain samples only in patients with mental illness and not in normal controls, fortifying the hypothesis of aberrant multimerization of the DISC1 protein in these conditions (*56*).

Due to the scaffolding nature of DISC1, the abberant multimerization and aggregation, might result in the pull down of its binding partners. This could also lead to an imbalance of the crucial cytosolic pool for normal functioning of the brain (*61*). Obtaining a clearer picture of the domain organization of the DISC1 protein, and its characteristics, would dramatically ease the process understanding the conditions affecting its multimer formation. Also would pave a way to obtain the finer structural details deciphering its functional mechanism.

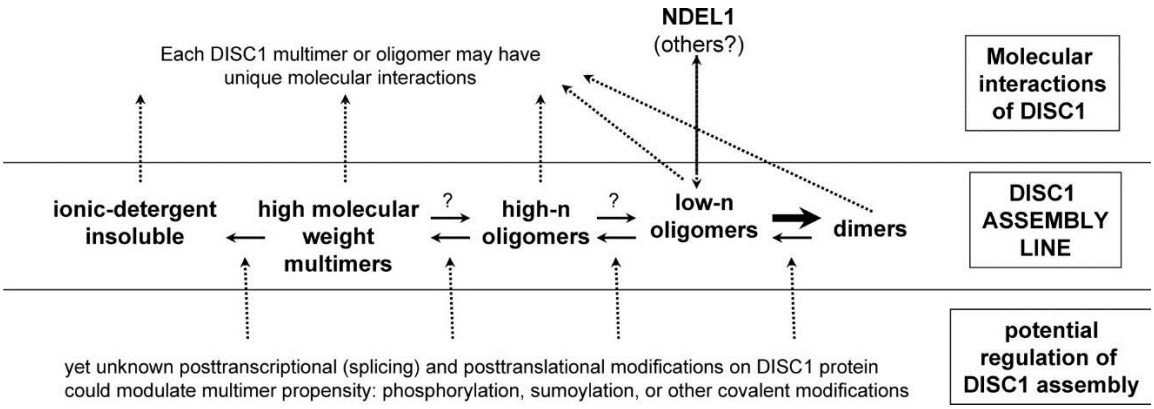

Figure 5: Depiction of possible scenarios of assembly and misaasembly of the DISC1 protein affecting its function. (Taken from Brandon *et al*, 2009(*52*))

1.6. Cloning of camelid single chain antibodies

The sera of camelids (camels, llamas, dromedaries) contain functional heavy chain only antibodies which are devoid of light chains, unlike the conventional ones which require domain pairing of the heavy chain and light chain N –terminal variable domains (VH and VL) for antigen binding (*64*). The single unpaired N-terminal variable domain, referred to as V_HH domains or nanobodies, of these heavy chain only antibodies are by themselves fully capable of specific antigen recognition and binding (*65*). These V_HH domains show numerous benefits compared to the conventional antibodies, which include higher yields of expression when produced in recombinant fashion in bacterial cells. They also possess relatively higher levels of stability and solubility and show a greater ease of delivery for therapeutic applications, while retaining high affinity and specificity towards the targeted antigen (*66*). V_HH antibodies have a large range of applications in diverse fields, from therapeutic interventions and diagnostics to basic research, notably in the investigation of protein structure and function

(*67, 68*). The antigen specific V_HH isolation from immune, non-immune or semisynthetic libraries is usually performed using well-established phage, yeast or ribosome display methods (*69-71*).

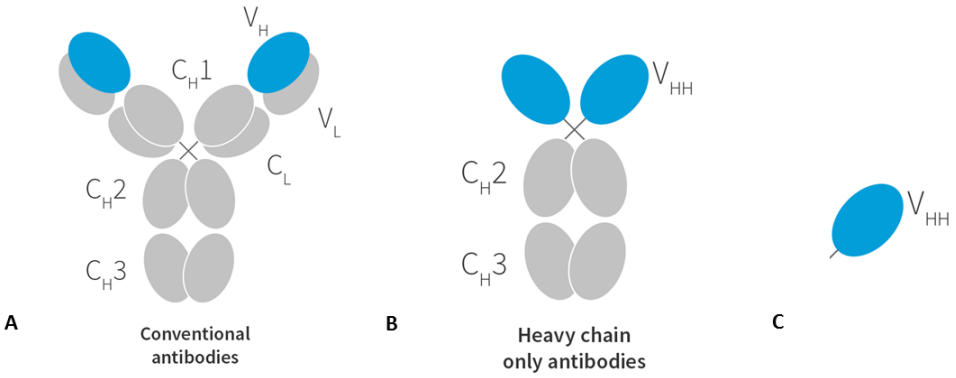

Figure 6 : Schematic depiction of the frame work of conventional, heavy chain only and VHH domain antibodies. **(Taken from http://www.ablynx.com/technology-innovation/understanding-nanobodies)**

A. The conventional antibody consists of heavy chains and light chains paired to each other by non-covalent interactions that are indicated by the antibody domains overlaying each other. The variable domains of the heavy chain and light chain are labelled as V_H and V_L respectively, whereas the constant domains are labelled as CH and CL respectively.
B. The heavy-chain only antibody consists of a variable domain V_HH, which does not require any sort of pairing. The constant domains are labelled as CH and these pair as depicted by the domain overlay.
C. V_HH domain antibody or nanobody is the variable domain alone from the camelid heavy chain only antibodies.

Phage display

Phage display technique is used to study protein-protein, protein-peptide and protein-DNA interactions through use of bacteriophages (*72*). This technique involves expression of a gene encoding the protein of interest inserted into a phage coat protein gene, resulting in the display of this protein of interest in the phage coat. These protein displaying phages could then be screened against the other proteins, peptides or DNA sequences, in order to detect any interaction between the two molecules. This technique is used to screen large libraries of proteins and thereby amplify the strong binders by a process called in-vitro selection or bio-panning.

The generic protocol of phage display screening to identify polypeptides that bind with high affinity to a desired target protein or DNA sequence involves:

1. Immobilization of the target molecule, for example to the wells of a microtiter plate.
2. A phage library, where each phage has the genetic sequence corresponding only a single coat protein of interest.
3. The phage library is added onto the wells coated with the target and incubated to facilitate binding.
4. After incubation, the wells are washed leaving only the phages binding to the target molecule to remain in the wells.
5. The attached phages are now eluted and used to infect the bacterial host to amplify the selected phages.
6. Steps 2 to 5 are repeated around 3 times, to enrich only the strong binders.
7. Finally the DNA of the strongest binders are sequenced to identify the interacting protein and cloned into desired plasmid vector for further use.

1.7. Techniques for the structural analysis of protein

1.7.1. Expression of soluble protein by random incremental truncation (ESPRIT)

When a full-length protein fails to express in soluble form or does not crystallise, it is common to instead study shorter fragments of the protein. Obtaining well behaving fragments by random trial and error is a time consuming and laborious process, especially when working on proteins with very little or no similarity to other known proteins. A shortcut for this lengthy process from DNA to protein is to create a large library of constructs with fragments of the corresponding gene and then screen them for those fragments which show soluble expression. Expression of soluble proteins by random incremental truncation (ESPRIT) technology is one such shortcut to screen for soluble regions within a protein from its genetic sequence (*73*).

The methodology of ESPRIT involves:

1. Random gene fragment construct library:

   The desired gene is cloned into a plasmid vector, termed pESPRIT002, with an N-terminal His tag and a C-terminal biotin acceptor peptide tag.
   Then a two-step process of random truncation is performed using Exonuclease III degradation. In the first step, the plasmid with the gene is linearized at one end of the gene and randomly truncated by Exonuclease III digestion thereafter a plasmid library is recovered. A second truncation is now performed on the pool of these plasmids from the other end of the gene and then re-ligated, thus generating gene fragment plasmid library, encoding versions of the gene randomly truncated from both ends.

   The sampling efficiency can be increased and number of clones to screen can be reduced by size fractionation on a agarose gel. For example, if a domain of 300 amino acids were expected to be screened, then constructs in the range of 200 to 400 amino acids can be isolated for testing.

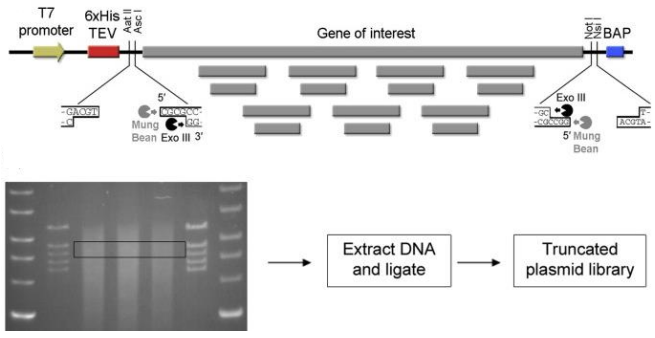

Figure 7: Generation of truncated gene plasmid library. (Taken from Yumerefendi *et al*, 2009 (*73*)).

2. Colony blot analysis:

Nearly 28,000 clones from the above generated library are then transformed into bacteria, which are induced for protein expression in a colony-array, usually in duplicates, with each colony expressing a unique truncated target gene insert.

The colony blots from the arrays are prepared by *in-situ* lysis on nitrocellulose membranes and the blots are probed with streptavidin in order to detect biotinylation of the C-terminal biotin acceptor peptide, which would indicate solubility. Additionally, the N-terminal histidine tag is also probed with monoclonal antibody against it; thus clones with both N and C –terminal tags by colony blot analysis were considered to be intact, non-degraded proteins, which were soluble. The intensity of detection is used as a scale to rank the constructs in the library according to their solubility.

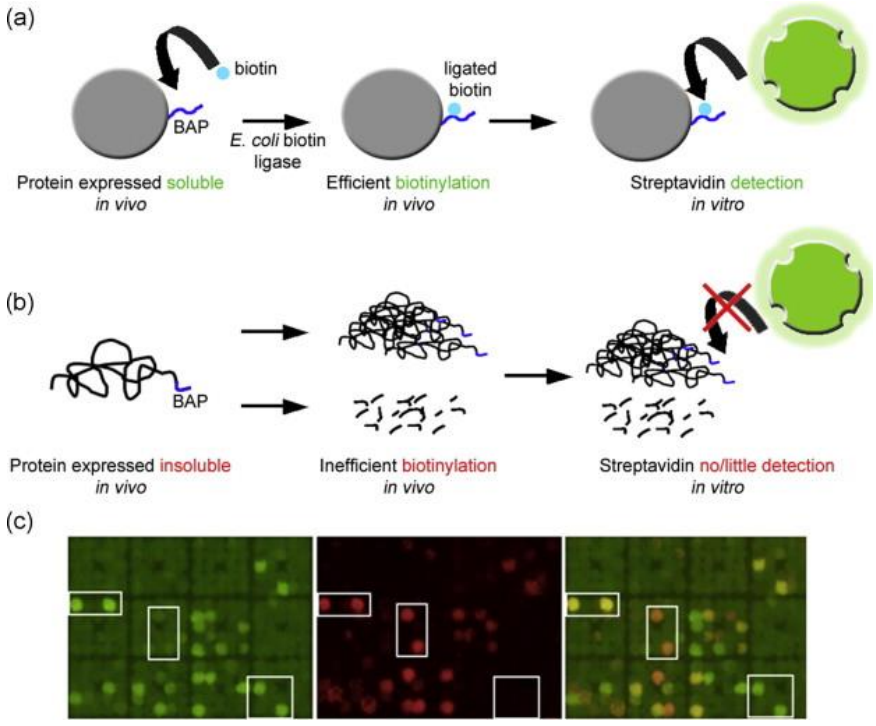

Figure 8: Colony array screen for soluble protein expressing gene constructs. (Taken from Yumerefendi_ *et al*, 2009 (*73*)).

3. Confirmation of the solubility by small scale protein expression and purification from liquid cultures:

In the second level of screening, the most efficiently biotinylated constructs are isolated and enriched through expression in a small scale format liquid culture format, which are then lysed and purified by Ni^{2+} NTA affinity resin. The recombinant

protein fragments yielded are visualised by SDS-PAGE, and then the constructs which encoded them are sequenced to determine the exact fragment which they contain. This second level distinguishes between the fully or partially soluble and insoluble constructs, thus selecting only for genetic constructs encoding highly soluble proteins.

The ESPRIT process could in principle be performed manually, but the whole process right from colony arraying, blotting, tag detection and small scale protein purification are performed by pre-programmed automatic robotic support. This makes this valuable high-throughput screen very feasible and in a short time duration.

1.7.2. Circular dichroism spectroscopy

Circular dichroism or CD is a valuable spectroscopic method to study all types of chiral molecules, but it is widely used in the study of large biological molecules. CD is the differential absorption of left-handed circularly polarized light (LCPL) and the right-handed circularly polarized light (RCPL) in the presence of molecule having one or more chiral groups.

Circular dichroism (CD) = $\Delta A(\lambda) = A(\lambda)_{LCPL} - A(\lambda)_{RCPL}$

Where A is absorption and λ is the wavelength

This technique involves the measurement of CD of molecules over a range of wavelengths. Such measurement of CD is performed to analyse the secondary structure or conformation of macromolecules, in particular proteins. The secondary structure of proteins is sensitive to changes in environment, such as interaction with other molecules, pH, temperature etc. Therefore structural, kinetic and thermodynamic information of proteins could be derived by CD spectroscopy.

The main unit of CD spectroscopy is termed as 'ellipticity', which is because of the conversion of the linearly polarized light to elliptically polarized form when passed through a circular dichroic sample. The elliptical polarized light is not fully circular but instead elliptical in shape because of the differential absorption of the right and left circularly polarized lights. This ellipticity when given as a mean of all the residues in the protein, is termed as mean residue ellipticity (degree.cm^2.dmol^{-1}).

Protein secondary structure is determined by CD in the 'far-UV' spectral range of wavelengths, which is 190-250nm. At these wavelengths, the peptide bond acts as the chiral group and the environment surrounding each one is recorded to give a spectrum. The α-helix, β-sheet and random coils are each giving rise to a characteristic shape and magnitude to the CD spectrum. Certain aspects of tertiary structure of a protein could also be obtained by CD in a 'near-UV' spectral region, which is 250-320nm. The aromatic amino acids and the disulphide bonds are the chiral group targeted at this spectral range.

CD spectroscopy gives relatively less structural information than the techniques such as X-ray crystallography and NMR spectroscopy. However, it is a quick and flexible method which is reliable, easily processed, require small amount of sample and can be used in variable conditions including the native physiological ones.

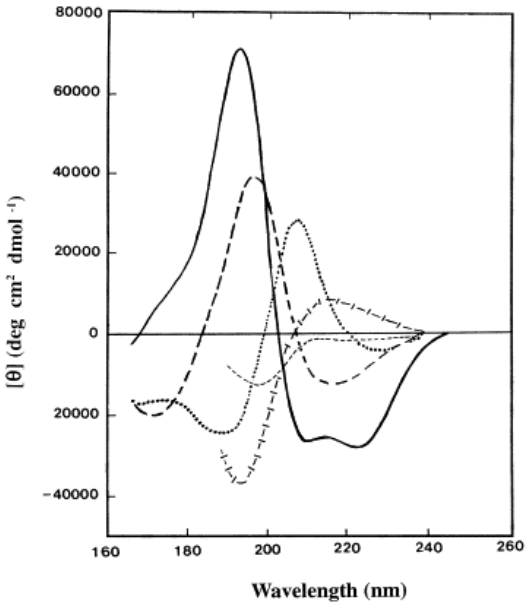

Figure 9: CD spectra for various secondary structures. (Taken from Kelly *et al, 2005.(74)*)

Solid line, α-helix; long dashed line, anti-parallel β-sheet ; short dashed line, irregular structure or random coil.

1.7.3. Small-angle X-ray scattering (SAXS)

Small-angle X-ray scattering is a scattering technique where the elastic scattering of X-rays by a homogenous protein sample are recorded at very low angles, typically below 1 degree. This angular range would contain information regarding the shape and size of molecules such as biological macromolecules. The structural information obtained is normally in the resolution range between 1nm and 25nm.

Biological applications of this technique are in the determination of the structure of a particle in terms of average particle size and shape. One could also obtain information regarding the surface-to-volume ratio of the samples. This method is non-destructive and typically involves the dispersion of the biological molecules in liquid.

The concept of the technique can be simply explained as follows: a homogenous protein sample is exposed to X-rays and the resulting scattering is registered in a detector. The scattering intensity is recorded as a function of momentum transfer. The intensity of scattering of the solvent alone is subtracted from the scattering of the solution. The random positions and orientation of particles in a homogenous sample result in a isotropic intensity distribution and is proportional to scattering of individual molecules.

The scattering curve derived here is directly sufficient for the determination of overall parameters of the protein in solution such as the radius of gyration, volume, molecular mass and folding state. Using mathematical calculation, even the surface of the molecule could be subtracted from the solution, thus the outer envelope or outer structure of the molecule in the solution can also be derived. This technique, when combined with the other structural methods such as X-ray crystallography, NMR spectroscopy and Cryo-electron microscopy is helpful in obtaining a holistic view to study a given molecule.

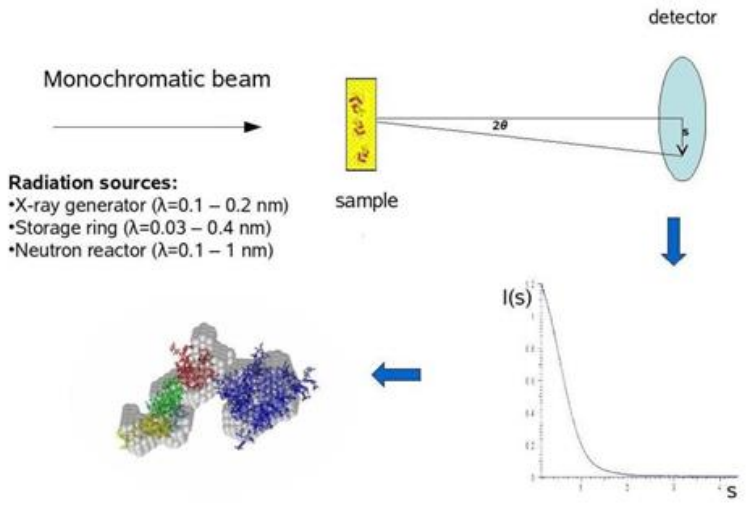

Figure 10: Flow chart depicting the scheme of SAXS experiment. **(taken from Wikipedia/biological small-angle scattering/Saxs scheme1.jpg).**

A homogenous sample when exposed to a X ray beam, the scattering in recorded as a scattering curve. A structural envelope is obtained from the scattering curve.

2. Aim of the thesis

The Disrupted-in-Schizophrenia 1 (DISC1) protein is understood to play crucial roles in neurological functions, however very little is understood about its functional mechanism. A singifcant reason for this is the lack of sufficient structural information, resulting from technical difficulties associated with experimentally characterizing this highly insoluble and aggregation-prone protein.

In this thesis, I aimed to obtain the domain architecture of the DISC1 protein by employing a functional-module based approach to address the technical problems associated with experimental characterization and search for the most soluble regions within the protein. I then applied the obtained knowledge in order to generate a picture of a potential functional mechanism and associated abnormalities.

3. Synopses of publications I, II, III

Synopsis I

Revisiting Disrupted-in-Schizophrenia 1 as a scaffold protein.

Antony S. K. Yerabham, Oliver H. Weiergräber, Nicholas J. Bradshaw, and Carsten Korth

Biological Chemistry 2013; 394(11): 1425–1437

DOI 10.1515/hsz-2013-0178

Author's contribution (60%):

- Design of the concept and content
- Co-preparation of the figures and tables
- Co-writing of the manuscript

DISC1 protein is known to bind to a large number of interaction partners and is therefore sometimes referred to as a scaffolding protein. In this review, this role of DISC1 as a scaffolding protein is looked through in detail and compared with well-known scaffolding proteins.

Here I, along with the senior author, established the characteristic features of the functioning of DISC1 and some of the established scaffold proteins. Additionally, I coordinated the insights generated into the structural features of DISC1 taking into account its interacting ability. We showed that DISC1 shares many but not all the properties of conventional scaffold proteins and here we also proposed an alternative name for DISC1 as 'Diverse Inhibitor of Signaling Cascades' due to the fact that it's major method of regulating enzymes is by inhibiting them, thus affecting several critical signaling pathways.

This review was the first to critically examine the scaffolding nature of DISC1 and compare it with the established criteria by which they are defined.

Synopsis II

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.

Antony S. K. Yerabham, Philippe J. Mas, Christina Decker, Dinesh C. Soares,
Oliver H. Weiergräber, Luitgard Nagel-Steger, Dieter Willbold. Darren J. Hart,
Nicholas J. Bradshaw, and Carsten Korth

Journal of Biological Chemistry 2017 Apr 21; 292(16):6468-6477

DOI: 10.1074/jbc.M116.773903

Author's Contribution (75%):

- Design of experimental setup
- Cloning and involvement in generation, plus the screening of the library
- Purification of all of the recombinant proteins
- Characterization of proteins by size exclusion chromatography and CD spectroscopy
- Data analysis and planning
- Co-writing the manuscript

In this study, I aimed for structural characterization of the DISC1 protein, which has until now been restricted largely to bioinformatics predictions.

Here we employed a high-throughput technique termed ESPRIT (Expression of Soluble Protein by Random Incremental Truncation), where a large library of 27,652 clones with randomly truncated *DISC1* gene fragments was generated and thereafter screened for the most soluble protein products, as protein solubility is an indication of the extent to which it is structured. I designed and cloned the DISC1 vector for the screen, which was then performed by our collaborators, and was responsible for testing and screening of the post promising "hits". I further characterized and identified the most soluble DISC1 protein fragments by size exclusion chromatography, circular dichroism spectroscopy

and analytical ultracentrifugation. Analytical ultracentrifugation was coordinated largely through the support of collaborators.

Based on different characteristics of the most structured regions within the DISC1 protein, I hypothesized the existence of four folded structural regions, named D, I, S and C, which form the structural basis of this protein. From the oligomericity and relative stability of individual regions, we speculated the functioning of the full length protein in vivo. Using the information obtained, we also investigated a pathological frame-shift mutant form of the DISC1 protein, by mimicking the condition in the C region and found it to possess a higher tendency to unfold and aggregate relative to the wild type C region.

Synopsis III

Biophysical insights from a single chain camelid antibody - Disrupted In Schizophrenia 1 protein (691-836) complex.

Antony S. K. Yerabham, Andreas Müller-Schiffmann, Tamar Ziehm, Andreas Stadler, Sabrina Köber, Xela Indhurkya, Rita Marreiros, Svenja V. Trossbach, Nicholas J. Bradshaw, Ingrid Prikulis, Oliver H. Weiergräber, and Carsten Korth

PLOS ONE - submitted

Author's Contribution (70%):

- Design of experimental setup
- Cloning and screening of the library
- Purification of all the proteins
- Characterization of proteins by size exclusion chromatography and western blot
- Data analysis and planning
- Co-writing the manuscript

The final 300 amino acids at the C-terminal end of the DISC1 protein includes many functionally important sites, including the disease-associated polymorphism S704C, the phosphorylation site at S713 and binding sites for multiple protein interaction partners. Notably, we had recently shown (in paper 2, described above) in the proposed structural organisation of the DISC1 protein that these 300 amino acids include three of the four regions described in the finding, namely I, S and C regions.

Here we report the generation of a camelid single chain antibody, also known as camelid V_HH, against the C-terminal of DISC1 and characterize its binding with the C region, in the process obtaining structural insights into this complex. I have screened, isolated and generated the camelid V_HH. Thereafter, I co-characterized its binding to the C region and was heavily involved in structural analysis by small-angle X-ray scattering (SAXS). This study also re-confirmed the findings of the C region, as described in Yerabham, *et al*, 2017 (paper 2, above).

4. Conclusions

Disrupted-in-Schizophrenia 1 (*DISC1*) is one of the most investigated genes in regard to mental illness, but there is a relative lack of mechanistic understanding regarding the protein it encodes. A major reason for this slow pace in revealing its functional mechanism is the dearth of sufficient structural information on the DISC1 protein. In the three papers which make up this cumulative thesis, I attempt to address this lack of structural understanding of DISC1.

Recombinant expression of human DISC1 in bacteria results in very low soluble and highly aggregating products, which stand as a major hurdle for the biophysical and structural characterization of the DISC1 protein. On the basis of this, to deal with this difficult protein I employed a high-throughput technique 'Expression of Soluble Protein by Random Incremental Truncation' (ESPRIT) to screen for the most soluble regions in the DISC1 protein between 150 – 300 amino acids in size (*73*). Through this high-throughput technique, structured regions of uncharacterized proteins are identified, by generation of a large library of random truncated protein fragments and screening for the most soluble ones. This was based on the general assumption that the structured or folded regions in a protein are more soluble than the loosely structured ones. These structured regions are often also responsible for the protein's functions.

Here I identified four folded regions within the DISC1 protein and propose them to form the basic domain architecture of the protein. The four regions are namely D region (from residues 249 to 399), I region (from residues 539 to 655), S region (from residues 635 to 738), and C region (from residues 691 to 836). Characterization of these regions individually showed differing oligomeric states and stabilities. The data presented here suggests the need for re-interpretation of the previous assumptions made regarding the predictions of the N-terminus of the protein to be predominantly disordered, because of the presence of the structured D region. Also, the prominent oligomeric form of the C region was deduced as an elongated monomer, rather than the previous prediction of this region existing as a dimer. As the dimeric I region and tetrameric S region overlap each other, while the protein fragments encompassing both the regions show mixed features, we hypothesize that at least two different structural conformations of the full-length DISC1 protein could exist. We also extrapolate that the conformer formation is probably a consequence of binding partners and their dependent regions, which stabilize the conformation thereby determine the functional fate of the formed complex.

The behavior of these structural elements also provide a valid explanation of the effect of the bisection of DISC1 by the Scottish translocation, which completely removes the more stable S and C regions while the I region is disrupted, thereby most likely triggering an incorrect folding of the remainder of the protein. Similarly, I have provided a partial explanation of the consequences of the frame-shift mutation of the *DISC1* gene, found in an American family, which disrupts the C region after the residue 807 followed by nine novel amino acid residues (*13*). When the frame-shift mutation was simulated using recombinant expressed C region protein, it led to the aberrant and higher oligomerization of the region relative to a wild type version. This was followed by unfolding and aggregation of the protein, which may parallel the ubiquitin mediated degradation as observed in the neurons generated using the pluripotent stem cells from the *DISC1* gene frameshift mutation affected individuals (*14*).

Using the soluble fragments of DISC1, various structural studies would be very feasible; nevertheless it is still a difficult target to obtain an experimental structure, due to its high flexibility. This probably exists in order to facilitate binding to multiple proteins. Usage of binding partners, either natural or artificial, would be an apt way to stabilize the flexible parts of the protein. Therefore I generated a camelid antibody against the DISC1 protein. The variable domain, or V_HH domain, of camelid heavy chain-only antibodies possess appreciable solubility and stability, and is individually capable of recognition and specific binding to an antigen. Currently these V_HH domains are widely used for stabilization of difficult proteins in structural investigations. Therefore, I generated a V_HH antibody (termed B5) against the C-terminus of the DISC1 protein using phage display methodology. We also characterized it by determining its binding site to be in the residue range 691 to 715 of the full length DISC1 protein. We thereafter obtained the first structural envelope of the DISC1 protein C region in complex with the B5 antibody by small-angle X-ray scattering (SAXS). This was further used to fit in the model of the DISC1 protein C region, thereby allowing interpretation of the structural folding of this region. This B5 antibody can be anticipated to serve as a useful tool for a wide range of potential applications to probe in-vitro and in-vivo investigations of the DISC1 protein and specifically target the DISC1 protein S and C regions. The camelid V_HH B5 generated here against the C-terminus of DISC1 is a potentially very useful tool for both structural and functional studies. The *ab-initio* model of the DISC1 C region is partially validated by fitting it into the structural envelope of the DISC1 C region–B5 complex, obtained through SAXS analysis. From our model several insights could be predicted: the DISC1 C region consists of three long α-helices, while the proline-rich motif (^{730}PPIPP^{734}) proposed to act as docking site for the Growth factor receptor-bound protein 2 (Grb2) SH3 domain exists in the loop region (*32*). Similarly amino acid S713, shown to be a phosphorylation site involved with switching the roles in DISC1 functioning from neuronal proliferation to neuronal migration (*54*), is located in the loop adjacent to the first helix and appears to be readily accessible to kinases.

Given the extensive binding ability of DISC1, it is predicted to act as a scaffold protein. Therefore, I reviewed this attribute of DISC1 protein alongside other well-known scaffold proteins, such as the Ste5 protein and the AKAP proteins. The mode of functioning, structural and organizational features of the well-known scaffold proteins were compared with that of DISC1 protein to obtain new insights on the functioning of the DISC1 protein. The characteristics of a typical scaffold protein are:

- Affecting a signaling pathway by binding to multiple components of that particular pathway
- Target or restrict specific proteins in localized environment or sub cellular compartment
- Enhancing or reducing the enzymatic activity of bound proteins
- Shielding the bound proteins from degradation, phosphorylation or any such modifications
- Being prominently structured as coiled coils.

From the available data, DISC1 does exhibit certain aspects of a typical scaffold protein such as it being able to bind multiple proteins of a particular signaling pathway, but also binds components of several inter-related pathways. Thus, it is functionally much broader in this aspect to that of a typical scaffold protein and strikingly DISC1, to the best of our knowledge, principally regulates the enzymatic protein by inhibition, with only few partial or indirect examples of enhancement. Furthermore, the structure of DISC1 is rich in α-helical content, like most scaffold proteins. Also DISC1's functioning is affected by its oligomeric state, which is typical for scaffold proteins as they orchestrate in a modular fashion, consisting of multiple domains in various arrangements for efficient

functioning. In summary, DISC1 therefore possess the features of a scaffold protein, but in an atypical way. Investigations into the various DISC1-related pathways individually, taking its scaffolding nature into consideration, would contribute to the better understanding of its mechanism of action.

Finally, the modular-based approach provided in this thesis has proven to be a very effective way to obtain a better understanding of the functional mechanism of the DISC1 protein. The domains proposed here and their characteristics hold a high potential for further investigations on the function, dysfunction and also experimental structure determination of this scaffold protein by individually targeting specific binding partners and pathways involved.

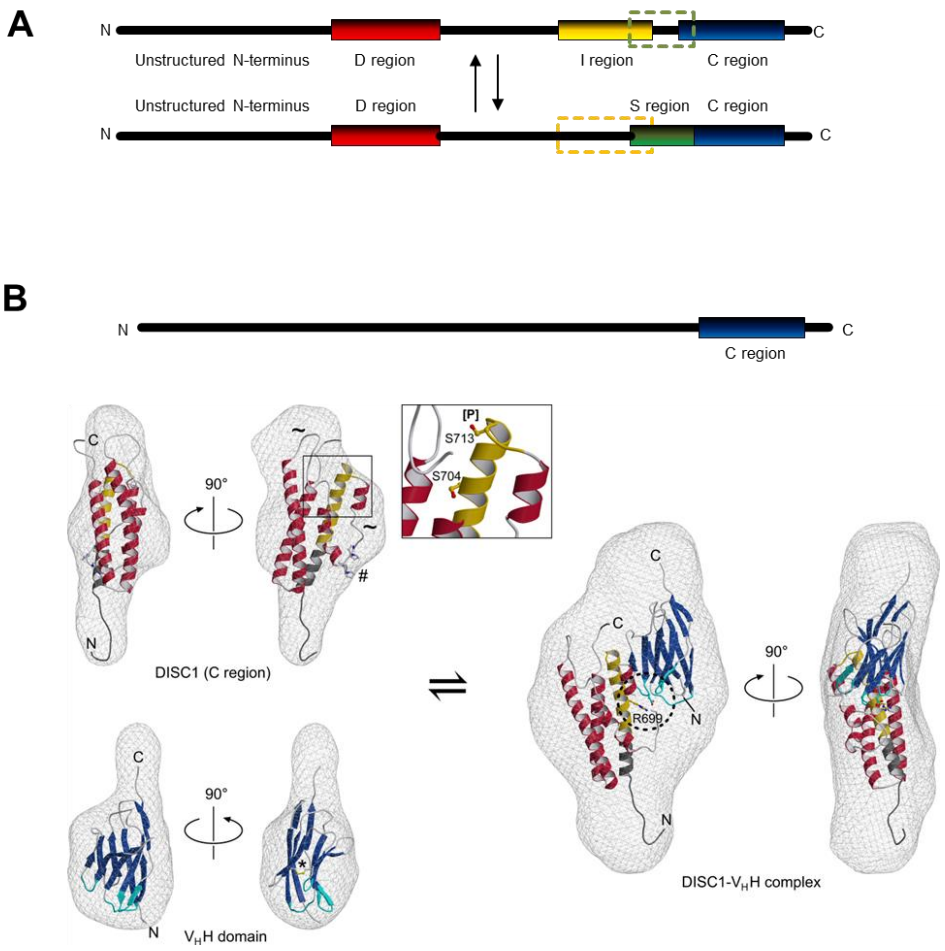

Figure 11: The structural organisation of the DISC1 protein and *ab-intio* model of the C region. **(Images adapted from manuscript 2 and 3 of this cumulative thesis)**

A. The proposed domain architecture of the DISC1 protein and the alternating conformers.

B. The structural envelope of the DISC1 protein C region, V_HH and their complex with *ab-intio* models fitted in predicting sites of prominence.

5. List of Publications

Peer-Reviewed Publications:

- Bradshaw NJ, **Yerabham ASK**, Marreiros R, Zhang T, Nagel-Steger L, Korth C., "An Unpredicted Aggregation-Critical Region of the Actin-Polymerizing Protein TRIOBP-1/Tara, Determined by Elucidation of its Domain Structure", ***The Journal of Biological Chemistry,*** 2017 doi: 10.1074/jbc.M116.767939

- **Yerabham ASK**, Mas PJ, Decker C, Soares DC, Weiergräber OH, Nagel-Steger L, Hart DJ, Bradshaw NJ, Korth C., "A Structural Organization for Disrupted in Schizophrenia 1, Identified by High Throughput Screening, Reveals Distinctly Folded Regions Which Are Bisected by Mental Illness-Related Mutations", ***The Journal of Biological Chemistry,*** 2017 292(16):6468-6477

- Ahmad S, Muthukumar S, Kuncha SK, Routh SB, **Yerabham ASK**, Hussain T, Kamarthapu V, Kruparani SP, Sankaranarayanan R., "Specificity and catalysis hardwired at the RNA-protein interface in a translational proofreading enzyme", ***Nature Communications***, 2015 26; 6:7552

- **Yerabham ASK**, Weiergräber OH, Bradshaw NJ, Korth C., "Revisiting disrupted-in-schizophrenia 1 as a scaffold protein", ***Biological Chemistry***, 2013 394 (11):1425-37

- **Sravankumar AS*[1]**, Ahmad S[1], Kruparani SP, Sankaranarayanan R., "Cloning, expression, purification, crystallization and preliminary X-ray crystallographic analyses of threonyl-tRNA synthetase editing domain from Aeropyrum pernix.", ***Acta Crystallographica Section F***, 2012 1; 68(Pt 11):1390-3

* Published as "Antony Sravankumar" rather than "Antony Sravan Kumar Yerabham", 1: Equal contribution

Manuscripts in Process:

- **Yerabham ASK** , Müller-Schiffmann A, Ziehm T, Stadler A, Köber S, Indhurkya X, Marreiros R, Trossbach SV, Bradshaw NJ, Prikulis I, Weiergräber OH, Korth C "Biophysical insights from a single chain camelid antibody - Disrupted In Schizophrenia 1 protein (691-836) complex." – *submitted.*

6. Acknowledgements

I owe my sincere gratitude to all those people who have made this thesis possible.

First of all I thank Prof. Dr. Carsten Korth, as he gave me this excellent opportunity to perform research at his laboratory. He has always given me a high degree of freedom to explore, collaborate and was available for valuable discussions, suggestions and critique. I really enjoyed his excellent mentoring throughout the course of my doctoral studies.

I thank Prof. Dr. Georg Groth for kindly agreeing to mentor my thesis as a member of the 'Mathematisch-Naturwissenschaftliche Fakultät'.

I especially express my sincere gratitude to Asst. Prof. Dr. Nicholas James Bradshaw and Dr. Andreas Müller-Schiffmann for their constant effort and tremendous patience in training and mentoring me very closely, both professionally and personally. Literally this thesis would not have been possible without their parent like care and support right from the beginning.

I would like to also especially thank PD. Dr. Oliver Weiergräber for his mentoring, compassionate attitude and immense support with all the work conducted at the Forschungszentrum Jülich.

I am very grateful to the former and current members of the Korth Lab and the other groups of the Institute of Neuropathology for their support during my time at the institute.

The warm atmosphere at the Korth Lab always made me feel it to be an extended family. I attempt to name everyone, in a random order: heartful thanks to Dr. Svenja Troßbach, Ms. Rita Marrieros, Dr. Sandra Vomund, Ms. Ingrid Prikulis, Dr. Verain Bader, Dr. Sarah Tschirner, Dr. Phillip Ottis, Ms. Sarah Swzed and the other members of Korth lab for their constant support, guidance and immense help with various chores in the lab. Apart from on a professional level, I also feel bonded at a personal level with the members of the group, which I will cherish all my life.

I thank Dr. Luitgard Nagel-Stager, Dr. Christina Decker and Mr. Tao Zhang for the productive collaboration and friendly support at the Institute for Physical Biology, HHU, Düsseldorf.

I thank Prof. Roland Riek for giving me an opportunity to work as a guest researcher in his laboaratory, Labor für Physikalische Chemie, ETH Zürich, Switzerland.

I want to thank 'Biostruct' and 'iBrain' graduate schools and STiBET for the financial support during my doctoral thesis.

I again specially thank Dr. Nicholas Bradshaw and Dr. Svenja Troßbach, for the immense help they provided me to compile this thesis.

The whole experience of living and studying in Germany was made very special by the numerous friends I gained during the period of my stay here; I thank each one of them from the bottom of my heart, whose love and support gave me a lot of strength during difficult times and also to learn the culture.

On this occasion I sincerely thank all my teachers, mentors right from my childhood, whose nurturing gave me the skills I attain right now.

Last, but most importantly, none of this would have been possible without the love, care and support from my dear family, my parents Prabhakara Reddy Yerabham, Vijaya Lakshmi Yerabham; my beloved sister's family Amala Sirisha Yerabham, Sunil Kommareddy and Ananya Kommareddy. I deeply appreciate their belief in me.

7. Bibliography

(1) Althoff, R.R., Faraone, S.V., Rettew, D.C., Morley, C.P., and Hudziak, J.J. (2005) Family, twin, adoption, and molecular genetic studies of juvenile bipolar disorder. *Bipolar disorders*. **7**, 598-609

(2) Cowan, W.M., Kopnisky, K.L., and Hyman, S.E. (2002) The human genome project and its impact on psychiatry. *Annual review of neuroscience*. **25**, 1-50

(3) Gottesman, II, and Shields, J. (1971) Schizophrenia: geneticism and environmentalism. *Human heredity*. **21**, 517-522

(4) Neves-Pereira, M., Mundo, E., Muglia, P., King, N., Macciardi, F., and Kennedy, J.L. (2002) The Brain-Derived Neurotrophic Factor Gene Confers Susceptibility to Bipolar Disorder: Evidence from a Family-Based Association Study. *The American Journal of Human Genetics*. **71**, 651-655

(5) Schizophrenia Working Group of the Psychiatric Genomics, C. (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. **511**, 421-427

(6) Serretti, A., and Smeraldi, E. (1999) Dopamine D2 receptor gene not associated with symptomatology of mood disorders. *American Journal of Medical Genetics*. **88**, 294-297

(7) Chang, M., Sun, L., Liu, X., Sun, W., Ji, M., Wang, Z., Wang, Y., and You, X. (2015) Evaluation of relationship between GRM3 polymorphisms and cognitive function in schizophrenia of Han Chinese. *Psychiatry Research*. **229**, 1043-1046

(8) Kim, J.S., and Lee, S.-H. (2016) Influence of interactions between genes and childhood trauma on refractoriness in psychiatric disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. **70**, 162-169

(9) St Clair, D., Blackwood, D., Muir, W., Walker, M., St Clair, D., Muir, W., Carothers, A., Spowart, G., Gosden, C., and Evans, H.J. (1990) Association within a family of a balanced autosomal translocation with major mental illness. *The Lancet*. **336**, 13-16

(10) Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A.M., Devon, R.S., Clair, D.M.S., Muir, W.J., Blackwood, D.H.R., and Porteous, D.J. (2000) Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Human Molecular Genetics*. **9**, 1415-1423

(11) Blackwood, D.H.R., Fordyce, A., Walker, M.T., St. Clair, D.M., Porteous, D.J., and 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**, 428-433

(12) Thomson, P.A., Duff, B., Blackwood, D.H., Romaniuk, L., Watson, A., Whalley, H.C., Li, X., Dauvermann, M.R., Moorhead, T.W., Bois, C., Ryan, N.M., Redpath, H., Hall, L., Morris, S.W., van Beek, E.J., Roberts, N., Porteous, D.J., St Clair, D., Whitcher, B., Dunlop, J., Brandon, N.J., Hughes, Z.A., Hall, J., McIntosh, A., and Lawrie, S.M. (2016) Balanced translocation linked to psychiatric disorder, glutamate, and cortical structure/function. *NPJ schizophrenia*. **2**, 16024

(13) Sachs, N.A., Sawa, A., Holmes, S.E., Ross, C.A., DeLisi, L.E., and Margolis, R.L. (2005) A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. *Molecular psychiatry*. **10**, 758-764

(14) Wen, Z., Nguyen, H.N., Guo, Z., Lalli, M.A., Wang, X., Su, Y., Kim, N.S., Yoon, K.J., Shin, J., Zhang, C., Makri, G., Nauen, D., Yu, H., Guzman, E., Chiang, C.H., Yoritomo, N., Kaibuchi, K., Zou, J., Christian, K.M., Cheng, L., Ross, C.A., Margolis, R.L., Chen, G., Kosik, K.S., Song, H., and Ming, G.L. (2014) Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature*. **515**, 414-418

(15) Ekelund, J., Hennah, W., Hiekkalinna, T., Parker, A., Meyer, J., Lonnqvist, J., and Peltonen, L. (2004) Replication of 1q42 linkage in Finnish schizophrenia pedigrees. *Molecular psychiatry*. **9**, 1037-1041

(16) Hamshere, M.L., Bennett, P., Williams, N., and et al. (2005) Genomewide linkage scan in schizoaffective disorder: Significant evidence for linkage at 1q42 close to disc1, and suggestive evidence at 22q11 and 19p13. *Archives of General Psychiatry*. **62**, 1081-1088
(17) Palo, O.M., Antila, M., Silander, K., Hennah, W., Kilpinen, H., Soronen, P., Tuulio-Henriksson, A., Kieseppä, T., Partonen, T., Lönnqvist, J., Peltonen, L., and Paunio, T. (2007) Association of distinct allelic haplotypes of DISC1 with psychotic and bipolar spectrum disorders and with underlying cognitive impairments. *Human Molecular Genetics*. **16**, 2517-2528
(18) Thomson, P.A., Harris, S.E., Starr, J.M., Whalley, L.J., Porteous, D.J., and Deary, I.J. (2005) Association between genotype at an exonic SNP in DISC1 and normal cognitive aging. *Neuroscience Letters*. **389**, 41-45
(19) Wright, I.C., Rabe-Hesketh, S., Woodruff, P.W., David, A.S., Murray, R.M., and Bullmore, E.T. (2000) Meta-analysis of regional brain volumes in schizophrenia. *The American journal of psychiatry*. **157**, 16-25
(20) Lipina, T.V., and Roder, J.C. (2014) Disrupted-In-Schizophrenia-1 (DISC1) interactome and mental disorders: Impact of mouse models. *Neuroscience & Biobehavioral Reviews*. **45**, 271-294
(21) Trossbach, S.V., Bader, V., Hecher, L., Pum, M.E., Masoud, S.T., Prikulis, I., Schable, S., de Souza Silva, M.A., Su, P., Boulat, B., Chwiesko, C., Poschmann, G., Stuhler, K., Lohr, K.M., Stout, K.A., Oskamp, A., Godsave, S.F., Muller-Schiffmann, A., Bilzer, T., Steiner, H., Peters, P.J., Bauer, A., Sauvage, M., Ramsey, A.J., Miller, G.W., Liu, F., Seeman, P., Brandon, N.J., Huston, J.P., and Korth, C. (2016) Misassembly of full-length Disrupted-in-Schizophrenia 1 protein is linked to altered dopamine homeostasis and behavioral deficits. *Molecular psychiatry*. **21**, 1561-1572
(22) Brandon, N.J., and Sawa, A. (2011) Linking neurodevelopmental and synaptic theories of mental illness through DISC1. *Nat Rev Neurosci*. **12**, 707-722
(23) Seshadri, S., Kamiya, A., Yokota, Y., Prikulis, I., Kano, S., Hayashi-Takagi, A., Stanco, A., Eom, T.Y., Rao, S., Ishizuka, K., Wong, P., Korth, C., Anton, E.S., and Sawa, A. (2010) Disrupted-in-Schizophrenia-1 expression is regulated by beta-site amyloid precursor protein cleaving enzyme-1-neuregulin cascade. *Proceedings of the National Academy of Sciences of the United States of America*. **107**, 5622-5627
(24) Ozeki, Y., Tomoda, T., Kleiderlein, J., Kamiya, A., Bord, L., Fujii, K., Okawa, M., Yamada, N., Hatten, M.E., Snyder, S.H., Ross, C.A., and Sawa, A. (2003) Disrupted-in-Schizophrenia-1 (DISC-1): mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 289-294
(25) James, R., Adams, R.R., Christie, S., Buchanan, S.R., Porteous, D.J., and Millar, J.K. (2004) Disrupted in Schizophrenia 1 (DISC1) is a multicompartmentalized protein that predominantly localizes to mitochondria. *Molecular and cellular neurosciences*. **26**, 112-122
(26) Morris, J.A., Kandpal, G., Ma, L., and 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**, 1591-1608
(27) Miyoshi, K., Honda, A., Baba, K., Taniguchi, M., Oono, K., Fujita, T., Kuroda, S., Katayama, T., and Tohyama, M. (2003) Disrupted-In-Schizophrenia 1, a candidate gene for schizophrenia, participates in neurite outgrowth. *Molecular psychiatry*. **8**, 685-694
(28) Taylor, M.S., Devon, R.S., Millar, J.K., and Porteous, D.J. (2003) Evolutionary constraints on the Disrupted in Schizophrenia locus. *Genomics*. **81**, 67-77
(29) Bradshaw, N.J., Ogawa, F., Antolin-Fontes, B., Chubb, J.E., Carlyle, B.C., Christie, S., Claessens, A., Porteous, D.J., and Millar, J.K. (2008) DISC1, PDE4B, and NDE1 at the centrosome and synapse. *Biochemical and biophysical research communications*. **377**, 1091-1096
(30) Bradshaw, N.J., Christie, S., Soares, D.C., Carlyle, B.C., Porteous, D.J., and Millar, J.K. (2009) NDE1 and NDEL1: multimerisation, alternate splicing and DISC1 interaction. *Neurosci Lett*. **449**, 228-233

(31) Kamiya, A., Kubo, K.-i., Tomoda, T., Takaki, M., Youn, R., Ozeki, Y., Sawamura, N., Park, U., Kudo, C., Okawa, M., Ross, C.A., Hatten, M.E., Nakajima, K., and Sawa, A. (2005) A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nat Cell Biol*. **7**, 1167-1178

(32) Shinoda, T., Taya, S., Tsuboi, D., Hikita, T., Matsuzawa, R., Kuroda, S., Iwamatsu, A., and Kaibuchi, K. (2007) DISC1 regulates neurotrophin-induced axon elongation via interaction with Grb2. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. **27**, 4-14

(33) Young-Pearse, T.L., Suth, S., Luth, E.S., Sawa, A., and Selkoe, D.J. (2010) Biochemical and functional interaction of disrupted-in-schizophrenia 1 and amyloid precursor protein regulates neuronal migration during mammalian cortical development. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. **30**, 10431-10440

(34) Brandon, N.J., Schurov, I., Camargo, L.M., Handford, E.J., Duran-Jimeniz, B., Hunt, P., Millar, J.K., Porteous, D.J., Shearman, M.S., and Whiting, P.J. (2005) Subcellular targeting of DISC1 is dependent on a domain independent from the Nudel binding site. *Molecular and cellular neurosciences*. **28**, 613-624

(35) Park, Y.U., Jeong, J., Lee, H., Mun, J.Y., Kim, J.H., Lee, J.S., Nguyen, M.D., Han, S.S., Suh, P.G., and Park, S.K. (2010) Disrupted-in-schizophrenia 1 (DISC1) plays essential roles in mitochondria in collaboration with Mitofilin. *Proceedings of the National Academy of Sciences of the United States of America*. **107**, 17785-17790

(36) Atkin, T.A., MacAskill, A.F., Brandon, N.J., and Kittler, J.T. (2011) Disrupted in Schizophrenia-1 regulates intracellular trafficking of mitochondria in neurons. *Molecular psychiatry*. **16**, 122-124, 121

(37) Hattori, T., Shimizu, S., Koyama, Y., Yamada, K., Kuwahara, R., Kumamoto, N., Matsuzaki, S., Ito, A., Katayama, T., and Tohyama, M. (2010) DISC1 regulates cell-cell adhesion, cell-matrix adhesion and neurite outgrowth. *Molecular psychiatry*. **15**, 778, 798-809

(38) Sawamura, N., Ando, T., Maruyama, Y., Fujimuro, M., Mochizuki, H., Honjo, K., Shimoda, M., Toda, H., Sawamura-Yamamoto, T., Makuch, L.A., Hayashi, A., Ishizuka, K., Cascella, N.G., Kamiya, A., Ishida, N., Tomoda, T., Hai, T., Furukubo-Tokunaga, K., and Sawa, A. (2008) Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. *Molecular psychiatry*. **13**, 1138-1148, 1069

(39) Clapcote, S.J., Lipina, T.V., Millar, J.K., Mackie, S., Christie, S., Ogawa, F., Lerch, J.P., Trimble, K., Uchiyama, M., Sakuraba, Y., Kaneda, H., Shiroishi, T., Houslay, M.D., Henkelman, R.M., Sled, J.G., Gondo, Y., Porteous, D.J., and Roder, J.C. (2007) Behavioral phenotypes of Disc1 missense mutations in mice. *Neuron*. **54**, 387-402

(40) Hayashi-Takagi, A., Takaki, M., Graziane, N., Seshadri, S., Murdoch, H., Dunlop, A.J., Makino, Y., Seshadri, A.J., Ishizuka, K., Srivastava, D.P., Xie, Z., Baraban, J.M., Houslay, M.D., Tomoda, T., Brandon, N.J., Kamiya, A., Yan, Z., Penzes, P., and Sawa, A. (2010) Disrupted-in-Schizophrenia 1 (DISC1) regulates spines of the glutamate synapse via Rac1. *Nature neuroscience*. **13**, 327-332

(41) Wang, Q., Charych, E.I., Pulito, V.L., Lee, J.B., Graziane, N.M., Crozier, R.A., Revilla-Sanchez, R., Kelly, M.P., Dunlop, A.J., Murdoch, H., Taylor, N., Xie, Y., Pausch, M., Hayashi-Takagi, A., Ishizuka, K., Seshadri, S., Bates, B., Kariya, K., Sawa, A., Weinberg, R.J., Moss, S.J., Houslay, M.D., Yan, Z., and Brandon, N.J. (2011) The psychiatric disease risk factors DISC1 and TNIK interact to regulate synapse composition and function. *Molecular psychiatry*. **16**, 1006-1023

(42) Yerabham, A.S., Weiergraber, O.H., Bradshaw, N.J., and Korth, C. (2013) Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. *Biological chemistry*. **394**, 1425-1437

(43) Camargo, L.M., Collura, V., Rain, J.C., Mizuguchi, K., Hermjakob, H., Kerrien, S., Bonnert, T.P., Whiting, P.J., and Brandon, N.J. (2007) Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. *Molecular psychiatry*. **12**, 74-86

(44) Porteous, D.J., Thomson, P., Brandon, N.J., and Millar, J.K. (2006) The genetics and biology of DISC1--an emerging role in psychosis and cognition. *Biological psychiatry*. **60**, 123-131
(45) Lewis, D.A., and Levitt, P. (2002) Schizophrenia as a disorder of neurodevelopment. *Annual review of neuroscience*. **25**, 409-432
(46) Jaaro-Peled, H., Hayashi-Takagi, A., Seshadri, S., Kamiya, A., Brandon, N.J., and Sawa, A. (2009) Neurodevelopmental mechanisms of schizophrenia: understanding disturbed postnatal brain maturation through neuregulin-1-ErbB4 and DISC1. *Trends in neurosciences*. **32**, 485-495
(47) Drerup, C.M., Wiora, H.M., Topczewski, J., and Morris, J.A. (2009) Disc1 regulates foxd3 and sox10 expression, affecting neural crest migration and differentiation. *Development*. **136**, 2623-2632
(48) Ma, L., Liu, Y., Ky, B., Shughrue, P.J., Austin, C.P., and Morris, J.A. (2002) Cloning and characterization of Disc1, the mouse ortholog of DISC1 (Disrupted-in-Schizophrenia 1). *Genomics*. **80**, 662-672
(49) Singh, K.K., Ge, X., Mao, Y., Drane, L., Meletis, K., Samuels, B.A., and Tsai, L.H. (2010) Dixdc1 is a critical regulator of DISC1 and embryonic cortical development. *Neuron*. **67**, 33-48
(50) Schurov, I.L., Handford, E.J., Brandon, N.J., and Whiting, P.J. (2004) Expression of disrupted in schizophrenia 1 (DISC1) protein in the adult and developing mouse brain indicates its role in neurodevelopment. *Molecular psychiatry*. **9**, 1100-1110
(51) Duan, X., Chang, J.H., Ge, S., Faulkner, R.L., Kim, J.Y., Kitabatake, Y., Liu, X.-b., Yang, C.-H., Jordan, J.D., Ma, D.K., Liu, C.Y., Ganesan, S., Cheng, H.-J., Ming, G.-l., Lu, B., and Song, H. (2007) Disrupted-In-Schizophrenia 1 Regulates Integration of Newly Generated Neurons in the Adult Brain. *Cell*. **130**, 1146-1158
(52) Brandon, N.J., Millar, J.K., Korth, C., Sive, H., Singh, K.K., and Sawa, A. (2009) Understanding the Role of DISC1 in Psychiatric Disease and during Normal Development. *The Journal of Neuroscience*. **29**, 12768-12775
(53) Mao, Y., Ge, X., Frank, C.L., Madison, J.M., Koehler, A.N., Doud, M.K., Tassa, C., Berry, E.M., Soda, T., Singh, K.K., Biechele, T., Petryshen, T.L., Moon, R.T., Haggarty, S.J., and Tsai, L.H. (2009) Disrupted in schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3beta/beta-catenin signaling. *Cell*. **136**, 1017-1031
(54) 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., and Sawa, A. (2011) DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature*. **473**, 92-96
(55) Sanchez-Pulido, L., and Ponting, C.P. (2011) Structure and evolutionary history of DISC1. *Human Molecular Genetics*. **20**, R175-R181
(56) Leliveld, S.R., Bader, V., Hendriks, P., Prikulis, I., Sajnani, G., Requena, J.R., and 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. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. **28**, 3839-3845
(57) Sohi, M., Alexandrovich, A., Moolenaar, G., Visse, R., Goosen, N., Vernede, X., Fontecilla-Camps, J.C., Champness, J., and Sanderson, M.R. (2000) Crystal structure of Escherichia coli UvrB C-terminal domain, and a model for UvrB-uvrC interaction. *FEBS letters*. **465**, 161-164
(58) Alexandrovich, A., Czisch, M., Frenkiel, T.A., Kelly, G.P., Goosen, N., Moolenaar, G.F., Chowdhry, B.Z., Sanderson, M.R., and Lane, A.N. (2001) Solution structure, hydrodynamics and thermodynamics of the UvrB C-terminal domain. *Journal of biomolecular structure & dynamics*. **19**, 219-236
(59) Soares, D.C., Carlyle, B.C., Bradshaw, N.J., and Porteous, D.J. (2011) DISC1: Structure, Function, and Therapeutic Potential for Major Mental Illness. *ACS chemical neuroscience*. **2**, 609-632

(60) Bader, V., Ottis, P., Pum, M., Huston, J.P., and Korth, C. (2012) Generation, purification, and characterization of cell-invasive DISC1 protein species. *Journal of visualized experiments : JoVE*. e4132
(61) Ottis, P., Bader, V., Trossbach, S.V., Kretzschmar, H., Michel, M., Leliveld, S.R., and Korth, C. (2011) Convergence of two independent mental disease genes on the protein level: recruitment of dysbindin to cell-invasive disrupted-in-schizophrenia 1 aggresomes. *Biological psychiatry*. **70**, 604-610
(62) Leliveld, S.R., Hendriks, P., Michel, M., Sajnani, G., Bader, V., Trossbach, S., Prikulis, I., Hartmann, R., Jonas, E., Willbold, D., Requena, J.R., and 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
(63) Barodia, S.K., Park, S.K., Ishizuka, K., Sawa, A., and Kamiya, A. (2015) Half-life of DISC1 protein and its pathological significance under hypoxia stress. *Neuroscience Research*. **97**, 1-6
(64) Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., and Hamers, R. (1993) Naturally occurring antibodies devoid of light chains. *Nature*. **363**, 446-448
(65) Muyldermans, S. (2013) Nanobodies: natural single-domain antibodies. *Annual review of biochemistry*. **82**, 775-797
(66) Li, T., Vandesquille, M., Koukouli, F., Dudeffant, C., Youssef, I., Lenormand, P., Ganneau, C., Maskos, U., Czech, C., Grueninger, F., Duyckaerts, C., Dhenain, M., Bay, S., Delatour, B., and Lafaye, P. (2016) Camelid single-domain antibodies: A versatile tool for in vivo imaging of extracellular and intracellular brain targets. *Journal of Controlled Release*. **243**, 1-10
(67) Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G.F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W.G.J., Kobilka, B.K., and Steyaert, J. (2014) A general protocol for the generation of Nanobodies for structural biology. *Nat. Protocols*. **9**, 674-693
(68) Thys, B., Saerens, D., Schotte, L., De Bleeser, G., Muyldermans, S., Hassanzadeh-Ghassabeh, G., and Rombaut, B. (2011) A simple quantitative affinity capturing assay of poliovirus antigens and subviral particles by single-domain antibodies using magnetic beads. *Journal of Virological Methods*. **173**, 300-305
(69) Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R., and Muyldermans, S. (1997) Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS letters*. **414**, 521-526
(70) Tanha, J., Dubuc, G., Hirama, T., Narang, S.A., and MacKenzie, C.R. (2002) Selection by phage display of llama conventional V(H) fragments with heavy chain antibody V(H)H properties. *Journal of immunological methods*. **263**, 97-109
(71) Goldman, E.R., Anderson, G.P., Liu, J.L., Delehanty, J.B., Sherwood, L.J., Osborn, L.E., Cummins, L.B., and Hayhurst, A. (2006) Facile generation of heat-stable antiviral and antitoxin single domain antibodies from a semisynthetic llama library. *Analytical chemistry*. **78**, 8245-8255
(72) Smith, G.P., and Petrenko, V.A. (1997) Phage Display. *Chemical Reviews*. **97**, 391-410
(73) Yumerefendi, H., Tarendeau, F., Mas, P.J., and Hart, D.J. (2010) ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. *Journal of structural biology*. **172**, 66-74
(74) Kelly, S.M., Jess, T.J., and Price, N.C. (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1751**, 119-139

8. Appendices

Publication I

Revisiting Disrupted-in-Schizophrenia 1 as a scaffold protein.

Antony S. K. Yerabham, Oliver H. Weiergräber, Nicholas J. Bradshaw, and Carsten Korth

Biological Chemistry 2013; 394(11): 1425–1437

DOI 10.1515/hsz-2013-0178

DOI 10.1515/hsz-2013-0178 — Biol. Chem. 2013; 394(11): 1425–1437

Review

Antony S.K. Yerabham, Oliver H. Weiergräber, Nicholas J. Bradshaw* and Carsten Korth*

Revisiting Disrupted-in-Schizophrenia 1 as a scaffold protein

Abstract: *Disrupted-in-Schizophrenia 1* (*DISC1*) is a widely-accepted genetic risk factor for schizophrenia and many other major mental illnesses. Traditionally DISC1 has been referred to as a 'scaffold protein' because of its ability to bind to a wide array of other proteins, including those of importance for neurodevelopment. Here, we review the characteristic properties shared between established scaffold proteins and DISC1. We find DISC1 to have many, but not all, of the characteristics of a scaffold protein, as it affects a considerable number of different, but related, signaling pathways, in most cases through inhibition of key enzymes. Using threading algorithms, the C-terminal portion of DISC1 could be mapped to extended helical structures, yet it may not closely resemble any of the known tertiary folds. While not completely fitting the classification of a classical scaffold protein, DISC1 does appear to be a tightly regulated and multi-faceted inhibitor of a wide range of enzymes from interrelated signaling cascades (Diverse Inhibitor of Signaling Cascades), which together contribute to neurodevelopment and synaptic homeostasis. Consequently, disruption of this complex regulation would be expected to lead to the range of major mental illnesses in which the *DISC1* gene has been implicated.

Keywords: DISC1; mental illness; protein-protein interactions; signaling pathways; structure; threading models.

***Corresponding authors: Nicholas J. Bradshaw and Carsten Korth,** Department of Neuropathology, Heinrich Heine University Düsseldorf, Moorenstrasse 5, D-40225 Düsseldorf, Germany, e-mail: nicholas.bradshaw@uni-duesseldorf.de; ckorth@uni-duesseldorf.de

Antony S.K. Yerabham: Department of Neuropathology, Heinrich Heine University Düsseldorf, Moorenstrasse 5, D-40225 Düsseldorf, Germany

Oliver H. Weiergräber: Institute of Complex Systems (ICS-6: Structural Biochemistry), Forschungszentrum Jülich, D-52425 Jülich, Germany

DISC1 – a major gene involved in behavioral control and chronic mental illness

Disrupted-in-Schizophrenia 1 (*DISC1*) is one of the highest profile risk factors for schizophrenia and other chronic mental illnesses (Chubb et al., 2008) having originally been identified as a gene that was disrupted by a chromosomal translocation (1;11) strongly linked to psychiatric illness in a large Scottish family (Millar et al., 2000; Blackwood et al., 2001). Since then, a substantial body of genetic evidence has accumulated, supporting the notion that *DISC1* is more generally associated with multiple chronic mental illnesses including schizophrenia, depression and autism (reviewed in: Chubb et al., 2008; Bradshaw and Porteous, 2012), as well as with cognitive dysfunction in healthy individuals (Callicott et al., 2005; Thomson et al., 2005). The concept of *DISC1* being a key player in behavioral control in supported independently by several rodent models expressing mutant or deleted variants of the *DISC1* gene (Shen et al., 2008; Brandon and Sawa, 2011; Kuroda et al., 2011).

The concept of DISC1 as a scaffold protein

Initial analysis of the amino acid sequence of the DISC1 protein yielded no clear indications as to its function beyond a similarity to certain structural proteins (Millar et al., 2000) and to date no evidence of enzymatic activity has been reported. However, yeast two-hybrid studies implicated the DISC1 protein in binding to numerous protein interaction partners (Millar et al., 2003; Miyoshi et al., 2003; Morris et al., 2003; Ozeki et al., 2003) using multiple distinct binding sites (reviewed in: Soares et al., 2011). Together this led to the suggestion that DISC1 is 'a multifunctional protein, which interacts via distinct domains with different components of the intracellular

machinery' (Morris et al., 2003) and which 'is likely to be critical to development and functioning of the brain by acting as a scaffold to bring together certain proteins, such as signal transduction molecules' (Millar et al., 2003). This was reinforced by further systematic yeast two-hybrid screening, contributing greatly to a list that now numbers more than 200 potential DISC1-interacting proteins, the 'DISC1 interactome', and the suggestion that DISC1 could act as a key hub protein for processes of potential importance to mental health (Camargo et al., 2007).

A diverse number of identified functions collectively explain many aspects of the behavioral control attributed to DISC1. Most prominent are functions in neuronal migration (Kamiya et al., 2005), the integration of adult-born neurons (Duan et al., 2007), Wnt signaling (Mao et al., 2009), the development of dopaminergic neurotransmission (Niwa et al., 2010), NMDA neurotransmission (Hayashi-Takagi et al., 2010) and GABA neurotransmission (Kim et al., 2012). These functions are not mutually exclusive, but so far no unifying theory has been presented to explain which function is most salient, and in which context.

In view of this large number of interaction partners and functions, it has become customary to refer to DISC1 as a scaffold or scaffolding protein, for example, in recent reviews by ourselves and others (Brandon and Sawa, 2011; Bradshaw and Porteous, 2012; Korth, 2012; Thomson et al., 2013). Looking back on a decade of structure-function research on DISC1, we set out to investigate to what extent its classification as a conventional scaffold protein is indeed backed by experimental evidence, in order to help develop an overview of the roles of DISC1 in the cell and in the pathology of major mental illness.

The characteristics of a typical scaffold protein

In order to evaluate whether DISC1 qualifies as a *bona fide* scaffold protein, it is first important to review a few well-established examples of this group of proteins. Scaffold proteins are normally conceived as modulators of cellular pathways, conventionally lacking enzymatic activity and instead exerting their function by simultaneously interacting with multiple macromolecules, offering 'a simple, flexible strategy for regulating selectivity in pathways, shaping output behaviours, and achieving new responses from pre-existing signalling components' (Good et al., 2011). They usually act as specificity elements that selectively channel signaling between its bound partners (Zeke et al., 2009).

The concept of scaffolding proteins has been established for almost two decades (Choi et al., 1994; Zeke et al., 2009). Among the first scaffold proteins to be thoroughly investigated was Ste5 (Sterile 5), which was known to affect the activity of three enzymes from the *Saccharomyces cerevisiae* mitogen-activated protein kinase (MAPK) pathway (Ste11, Ste7, Fus3), and so was initially assumed to exist as an upstream signaling element (Hasson et al., 1994). Further research instead demonstrated Ste5 to be responsible for tethering these proteins and forming a multi-kinase complex (Choi et al., 1994). Thus, Ste5 effectively brings enzyme-substrate pairs of the cascade into physical proximity, resulting in more efficient signal transduction. Mathematical modeling studies revealed that the active phosphorylated intermediate species of Ste11, Ste7 and Fus3, when bound by the Ste5 scaffold protein, have a minimal probability of being consumed by competing signaling pathways or by non-specific phosphatases within the cell (Locasale et al., 2007). These studies also demonstrated that the signaling proteins bound to a scaffold protein will display restricted enzymatic activity due to their spatial arrangement. Finally, the ability of Ste5 to regulate the pathway was dependent on whether a positive (Ste50) or negative (Msg5) regulatory molecule was bound to it (Bashor et al., 2008).

A-kinase anchoring proteins (AKAPs) are another group of scaffold proteins, classified by their capacity to associate with protein kinase A (PKA) (Diviani and Scott, 2001). These proteins act to recruit signaling components to specific subcellular locations, for example AKAP350/450 targets PKA and protein kinase C to sites within the centrosome (Shanks et al., 2002), while in the heart a shorter splice variant of it, Yotiao, brings together PKA, PP1, PDE4D3 and the potassium ion channel and thereby locally regulates the β-adrenergic stimulation (Li et al., 2012).

Additional common properties of scaffolding proteins can be seen from other examples in the literature. For example, the mammalian Ste5 homologue Kinase Suppressor of Ras (KSR), depending on its phosphorylation state, mediates the translocation of multiple MAPK signaling components from the cytoplasm to the plasma membrane (Muller et al., 2001). Inactivation No Afterpotential D () is another scaffold protein expressed in the photoreceptor cells of *Drosophila* which localize and promote phosphorylation of transient receptor potential Ca^{2+} channels, thereby enhancing the quick adaptability of the visual system in fruit flies (Popescu et al., 2006). Homer 2 and Homer 3 are cytoplasmic scaffold proteins

that negatively regulate T-cell activation in mammals by binding to the Nuclear Factor of Activated T cells protein and shielding it from dephosphorylation. Homer-deficient mice developed autoimmune-like pathology, implying the importance of proper regulation of T-cell activation by the Homer family of scaffold proteins (Huang et al., 2008).

While the overall three-dimensional folds of scaffold proteins are quite diverse, our analysis indicates that the presence of coiled coil regions is a widespread feature. This super-secondary structure offers several advantages to the orchestration of multi-protein complexes (Rose and Meier, 2004). First and foremost, these include the ability to homo- or hetero-oligomerize, resulting in formation of supercoils involving different numbers of helices. Since coiled coil proteins often contain distinct functional domains on the same polypeptide chain, this oligomerization propensity provides a first level of clustering diverse (catalytic or non-catalytic) activities. In a few cases, formation of coiled coil dimers has been shown to be subject to regulation, such as by temperature or phosphorylation state (Hurme et al., 1996; Szilak et al., 1997). An additional level of complexity is usually added by different types of protein-protein interactions, typically with signaling proteins and enzymes, which may be mediated by the coiled coil domains themselves or by associated modules.

Second, scaffolding functions are typically assigned to elongated or rod-like molecules, and long helical supercoils provide an efficient means of accomplishing this geometry. In addition to providing an extended surface for protein-protein interactions, such elongated structures may serve as molecular rulers or spacers, precisely defining both the distances and relative orientation of associated molecules. An interesting variation to this theme is found in proteins of the spectrin superfamily. In this case, rod-like structures are formed from a series of so-called spectrin repeats, which are relatively short coiled coils sometimes connected by continuous helices (Djinovic-Carugo et al., 2002). Despite obvious structural constraints, spectrin repeats have evolved to accomplish a surprising variety of interactions also involving non-coiled coil proteins.

Finally, coiled coil domains are also commonly found in polypeptides determining the shapes and mechanical properties of cells and organelles, such as centrosomal proteins, golgins and motor proteins (Rose and Meier, 2004). Besides their structural roles (akin to scaffolds in human construction works), many of these proteins are also involved in signal transduction networks, and may thus qualify as scaffolding proteins in the sense used in this review.

Therefore, there are a number of features common to many scaffold proteins, namely that they:

- Bind to multiple protein components of the same signaling pathway, increasing both the efficiency and specificity of enzymatic cascades.
- Target the bound proteins to specific subcellular compartments, restricting a particular response to the appropriate location in the cell.
- Regulate the signaling pathway by enhancing or reducing the enzymatic activity of bound proteins.
- Shield the activated signaling intermediates preventing from any non-specific activity and consumption such as dephosphorylation (Locasale et al., 2007; Shaw and Filbert, 2009).
- Often have a high coiled coil-forming propensity, facilitating diverse and complex protein interactions.

With these criteria in mind, we will address each of these points in turn, and investigate how well DISC1 fits with them, based on published literature of its cellular functions.

Does DISC1 fit with these criteria for a classical scaffold protein?

Binding multiple proteins of the same signaling pathway

Among the many known protein-binding partners of DISC1, there are numerous examples of pairs of proteins that also exist in complexes with each other, particularly at the centrosome (reviewed in: Wang and Brandon, 2011; Bradshaw and Porteous, 2012), including the motor protein dynein and its associated proteins dynactin, Lissencephaly 1 (LIS1), Nuclear Distribution Element 1 (NDE1) and NDE-Like 1 (NDEL1) (Millar et al., 2003; Morris et al., 2003; Ozeki et al., 2003; Brandon et al., 2004) as well as the proteins Pericentriolar Material 1 (PCM1) and Bardet-Biedl Syndrome 4 (BBS4), which are together important for neuronal migration (Kamiya et al., 2008). DISC1 and NDEL1 have each been seen to interact with the mitochondrial protein Mitofilin (Park et al., 2010). Given the multiple distinct protein binding regions of DISC1 (reviewed in: Soares et al., 2011), it is highly likely that DISC1 is capable of binding to multiple different interaction partners simultaneously.

Evidence for such direct interactions can come from recombinant protein binding studies, once the individual proteins have been established to complex in a pairwise

fashion in the cell. For example, the kinesin motor protein KIF5A and adaptor protein Growth factor Receptor-Bound protein 2 (GrB2) interact *in vitro* in the presence, but not absence, of recombinant DISC1, strongly suggesting that the three proteins complex, through simultaneous binding to DISC1 (Shinoda et al., 2007). KIF5A-NDEL1-DISC1 ternary complexes have also been demonstrated in this way (Taya et al., 2007). In an analogous approach, use of DISC1 knockdown by shRNA reduced the co-immunoprecipitation of neurodevelopmental proteins Fasciculation and Elongation protein Zeta-1 (FEZ1) and NDEL1 from neuronal progenitor lysates (Kang et al., 2011), elegantly providing evidence for FEZ1-DISC1-NDEL1 complexes in an *in vivo* system. Conversely, over-expression of DISC1 enhances the interactions of both the microtubule proteins LIS1 and NDEL1 (Brandon et al., 2004) and of the synaptic proteins Kalirin-7 (Kal-7) and Postsynaptic Density protein 95 (PSD95) (Hayashi-Takagi et al., 2010), implying that these proteins co-complex at least in part through simultaneous interaction with DISC1.

At a minimum, DISC1 therefore appears to act as a scaffold for protein complexes involved in neurite outgrowth and function. While to date no published, confirmed examples of complexes involving both DISC1 and at least two enzymes from the same signaling pathway have been reported, candidate pathways do exist. Specifically, DISC1 binds Glycogen Synthase Kinase 3β (GSK3β, Mao et al., 2009), a key kinase of the Wnt signaling cascade, as well as two other modulators of this pathway, DIX Domain Containing 1 (Dixdc1, Singh et al., 2010) and Girdin (also known as KIAA1212, Enomoto et al., 2009; Kim et al., 2009). Similarly, DISC1 interacts with a range of phosphodiesterase 4 (PDE4) isoforms (Millar et al., 2005; Murdoch et al., 2007), which degrade cAMP as part of a negative feedback involving PKA. While DISC1 is not known to directly regulate PKA, it does interact with both Activating Transcription Factor 4 (ATF4) and NDE1 (Millar et al., 2003; Morris et al., 2003; Burdick et al., 2008; Sawamura et al., 2008; Bradshaw et al., 2009), two PKA-substrates each found in complex with PDE4B (Elefteriou et al., 2005; Bradshaw et al., 2008). Thus it is plausible that DISC1-GSK3β-regulator or DISC1-PDE4-substrate of PKA complexes could exist *in vivo*, fulfilling a key criterion for DISC1 having a scaffold protein function in GSK3β/PDE4B signaling. Interconnection between the PDE4 and GSK3 pathways has also been described (Carlyle et al., 2011; Lipina et al., 2012).

The ability of DISC1 to bind multiple proteins simultaneously is likely to be facilitated by its propensity to form oligomers, and thus potentially to present more protein binding domains at the same time. Of particular interest in this respect, a C-terminal DISC1 (640–854) fragment expressed in *Escherichia coli* was shown to interact with NDEL1 when the fragment was in an octameric form in a cell free assay (Leliveld et al., 2008), a finding that was later corroborated with full-length DISC1 (Narayanan et al., 2011). These findings suggest that the multimerization state of DISC1 may critically influence its protein interactions (Brandon et al., 2009).

Recruitment of proteins to specific cellular locations

As a putative scaffold protein, DISC1 would be expected to recruit a subset of its protein interaction partners to specific subcellular locations in order to facilitate their functions in a regulated manner. This appears to be the case with BBS4, which is recruited to the centrosome by DISC1 in a phosphorylation-dependent manner (Ishizuka et al., 2011). This in turn leads to the recruitment of further proteins, PCM1 and ninein, to the centrosome, from where DISC1, BBS4 and PCM1 are together implicated in regulating radial neuron migration (Kamiya et al., 2008). In a similar manner, DISC1 over-expression leads to increased expression of ATF4 within the nucleus (Pletnikov et al., 2007), where DISC1 and ATF4 have been shown to cumulatively repress cAMP Response Element-dependent gene transcription (Sawamura et al., 2008).

A better understood example of DISC1-driven recruitment is its role in bringing proteins to the axonal tips (Kamiya et al., 2006; Shinoda et al., 2007; Taya et al., 2007; Enomoto et al., 2009), which it appears to do by simultaneously interacting with kinesin motor complexes and the cargo proteins, facilitating their transport in an anterograde direction from the centrosome along the microtubule network (Shinoda et al., 2007; Taya et al., 2007).

Additionally, over-expression of DISC1 has been reported to recruit several of its interaction partners to punctate structures within the cytoplasm (Morris et al., 2003; Burdick et al., 2008; Wang et al., 2011), although whether this represents physiological recruitment or the known ability of DISC1 to bring other proteins into insoluble aggregates when highly concentrated (Ottis et al., 2011; Bader et al., 2012) remains unclear. We do not see the inherent aggregation propensity of DISC1 to be in contradiction to a possible function in scaffolding. As noted earlier, the multimerization-dependent interactions of DISC1 with NDEL1 (Leliveld et al., 2008) and the aggregation-dependent interactions with dysbindin or Collapsin Response Mediator Protein 1 are likely to be in a continuum of DISC1 multimer-dependent interactions as outlined in Brandon et al. (2009).

DISC1 therefore appears to play roles in protein recruitment, notably to the centrosome and via kinesin-related transport along axons, although its direct relevance to enzymes or signaling pathways is unclear, as summarized in Table 1.

Ability to modulate the enzymatic activity of binding partners

DISC1 interacts with multiple enzymes for which it is not a known substrate, implicating it as a regulator of their activity, as has been demonstrated in several cases. DISC1 binds to various members of the PDE4 family, which are responsible for the breakdown of cAMP, a second messenger for various stimuli and crucial in learning, memory and mood regulation (Millar et al., 2005). DISC1 holds PDE4 in an inactive state and can then release specific isoforms in response to cAMP-activated PKA (Murdoch et al., 2007). It was also shown that DISC1, via PDE4, regulates NDE1 phosphorylation by PKA, thereby modulating the organization of the NDE1-NDEL1-LIS1 complex (Bradshaw et al., 2011).

DISC1 has also been shown to repress the transcription-regulating activity of ATF4 (Sawamura et al., 2008). It appears to do this while ATF4 is present at its DNA target sites by simultaneously binding to ATF4 and recruiting the transcriptional repression factor Nuclear receptor Co-Repressor (N-CoR) (Sawamura et al., 2008), in a process that is regulated by both dopamine and PKA signaling (Soda et al., 2013). It is of interest that one of the genes directly affected by DISC1 and ATF4 is *PDE4D*, demonstrating the interconnectedness of pathways involving DISC1 (Soda et al., 2013).

DISC1 also inhibits GSK3β activity by direct physical interaction, preventing it from phosphorylating β-catenin, with important consequences for neural progenitor cell proliferation during cortical development (Mao et al., 2009). DISC1 co-modulates this pathway with its interaction partner Dixdc1. NDEL1 also exists in this complex, and during neuronal migration, phosphorylation of Dixdc1 by cyclin-dependent kinase 5 facilitates the Dixdc1-NDEL1 interaction (Singh et al., 2010).

GSK3β is also inhibited by Akt signaling, which in turn is regulated by DISC1 through the protein Girdin (Enomoto et al., 2009). DISC1 binds to Girdin, preventing it from activating Akt, a key kinase that regulates actin polymerization and cell motility (Enomoto et al., 2009). This function of DISC1 and the Akt-mTOR pathway has been shown to be important in regulating neuronal development in the dentate gyrus of the hippocampus (Kim et al., 2009).

Other examples of DISC1 affecting the enzymatic activity of proteins include the oligopeptidase activity of NDEL1, which is inhibited by DISC1 interaction (Hayashi et al., 2005). Interestingly, this NDEL1 enzyme activity is reduced in patients with schizophrenia, indicating the importance of regulation of NDEL1's activity for normal functioning of the brain (Gadelha et al., 2013). DISC1 also plays a crucial role in neuronal connectivity, via modulation of Ras-related C3 botulinum substrate 1 (Rac1) signaling via triple functional domain protein (TRIO) and Kal-7. DISC1 binds to TRIO, facilitating its Rac1 guanine nucleotide exchange factor ability through inhibition of its alternative Rho activating function (Chen et al., 2011) thereby promoting axon guidance. Conversely, DISC1 inhibits the alternative Rac1 guanine nucleotide exchange factor Kal-7 (Hayashi-Takagi et al., 2010). DISC1 also interacts with and inhibits Traf2 and Nck-interacting kinase (TNIK), leading to increased degradation of several key postsynaptic density proteins (Wang et al., 2011).

For a subset of its interaction partners, DISC1 therefore appears to be crucial for modulating the timing of their activity, normally through inhibition of catalytic functions. Thus DISC1 appears to be not merely a stationary sequester but an active regulator of the enzymatic behavior of its binding partners.

Shielding of activated signaling intermediates

There are a couple of indications showing that DISC1 may be able to shield specific signaling cascade intermediates

Table 1 A list of protein binding partners of DISC1 known to be recruited by DISC1 to distinct subcellular locations.

Subcellular region	Proteins
Centrosome	Bardet-Biedl syndrome 4, CAMDI, dynein IC, Lissencephaly 1, Ninein, p150^{glued} and Pericentriolar Material 1
Distal parts of the axon	14-3-3ε, Girdin, Growth factor Receptor-Bound protein 2, Lissencephaly 1 and NDE-Like 1
Nucleus	Activating Transcription Factor 4
Aggregates/aggresome	Collapsin Response Mediator Protein 1 and dysbindin
Cytoplasmic punctate structures	ATF5, Nuclear Distribution Element 1, NDE-Like 1 and Traf2 and Nck-interacting kinase

See main text for citations.

from undesired consumption. One example is serine racemase (SR), the enzyme responsible for production of *D*-serine, a co-agonist of the NMDA receptor. DISC1 was seen to complex with SR, with over-expression of a DISC1 mutant leading to an increased rate of ubiquitination and degradation of SR, but no down-regulation at the mRNA level (Ma et al., 2013). Thus, DISC1 binding is essential for the stability of SR, protecting it from degradation and facilitating the production of *D*-serine.

Similarly, the mitochondrial inner membrane protein Mitofilin was shown to bind to DISC1, with shRNA-targeted depletion of DISC1 or over-expression of a mutant leading to mitochondrial dysfunction (Park et al., 2010) resembling the effect of Mitofilin knockdown. Interestingly, over-expression of Mitofilin in DISC1-depleted cells restored these mitochondrial functions, strongly implying that DISC1 stabilizes it, a theory reinforced by experiments showing that the absence of DISC1 triggers ubiquitination of Mitofilin (Park et al., 2010).

These two examples indicate that DISC1 possesses the ability to shield some of its binding partners, with SR providing an example directly relevant to signaling pathways, as predicted for a scaffold protein.

Structural basis for DISC1 as a scaffold protein

The precise three-dimensional structure of DISC1 is still unknown and the protein shows no obvious sequence similarity to any entry in the Protein Data Bank, precluding conventional homology modeling. Previous analyses have indicated that the N-terminal region of DISC1 (330–350 residues) is likely to be largely disordered, with the remainder of the protein consisting of helical or coiled-coil domains (Soares et al., 2011). Additionally, two putative UVR domains, each containing a pair of antiparallel helices connected by a hairpin, have been predicted (Sanchez-Pulido and Ponting, 2011). A strategy to yield tentative three-dimensional fold information when obvious homologues with known structure are unavailable is given by threading approaches, which refers to a group of algorithms designed to detect faint similarities between a protein of interest and sequences with a known fold. In contrast to classical alignment strategies, sequences are not compared directly but represented as profiles, which can be expressed as position-specific scoring matrices or, in more sophisticated implementations, as hidden Markov models (Dunbrack, 2006). The method is based on the assumption that polypeptides exhibiting similar patterns of side chains tend to fold in a similar way. Such fold recognition by threading does not, however, imply any statement on the evolutionary relationship between the proteins involved.

We have subjected the DISC1 sequence to several current structure prediction pipelines, including Phyre[2] (Kelley and Sternberg, 2009), I-TASSER (Roy et al., 2010) and SAM-T08 (Karplus et al., 1998), which make use of different algorithms for the detection of potential templates. Out of the large number of profile-based alignments obtained, several long-range matches are particularly noteworthy (Figure 1A). In some of these, the majority of the C-terminal portion of DISC1 is mapped to extended helical structures, such as the spectrin repeat region of α-actinin, or the channel-forming colicin Ia. A second group of alignments indicate that large parts of the molecule may resemble the importin-β structure, a solenoid fold made up by lateral association of numerous HEAT repeats (containing two α-helices each), resulting in a spring-like superstructure. The predictive value of the resulting models, however, is questionable as the sequence profile of the DISC1 C-terminal domain appears to be equally compatible with very different arrangements of α-helices, making any prediction highly ambiguous. More importantly than this, critical assessment of model quality yields relatively low scores for all candidates. The situation can be slightly improved by manual adjustment of templates and alignments as well as additional energy minimization cycles, resulting in the model shown in Figure 1B, but the overall quality is still inferior to that expected of good homology models. These low-quality scores are mainly related to problems in side-chain packing and solvent accessibility: despite the apparent high confidence of profile matches, the distribution of hydrophobic and polar/charged side chains deviates from expectations. In general, extended hydrophobic patches observed on a protein surface may also represent sites of high-affinity protein-protein interactions, consistent with a scaffolding role. In the case of DISC1, however, it appears more likely that its C-terminal domain contains helical elements in an arrangement that is either not yet represented in the PDB or cannot be matched by the algorithms currently available.

Discussion

In this review we have analyzed the published literature on the DISC1 protein in comparison with the characteristics typical of scaffold proteins in order to determine how well DISC1 fits with that label (summarized in Table 2).

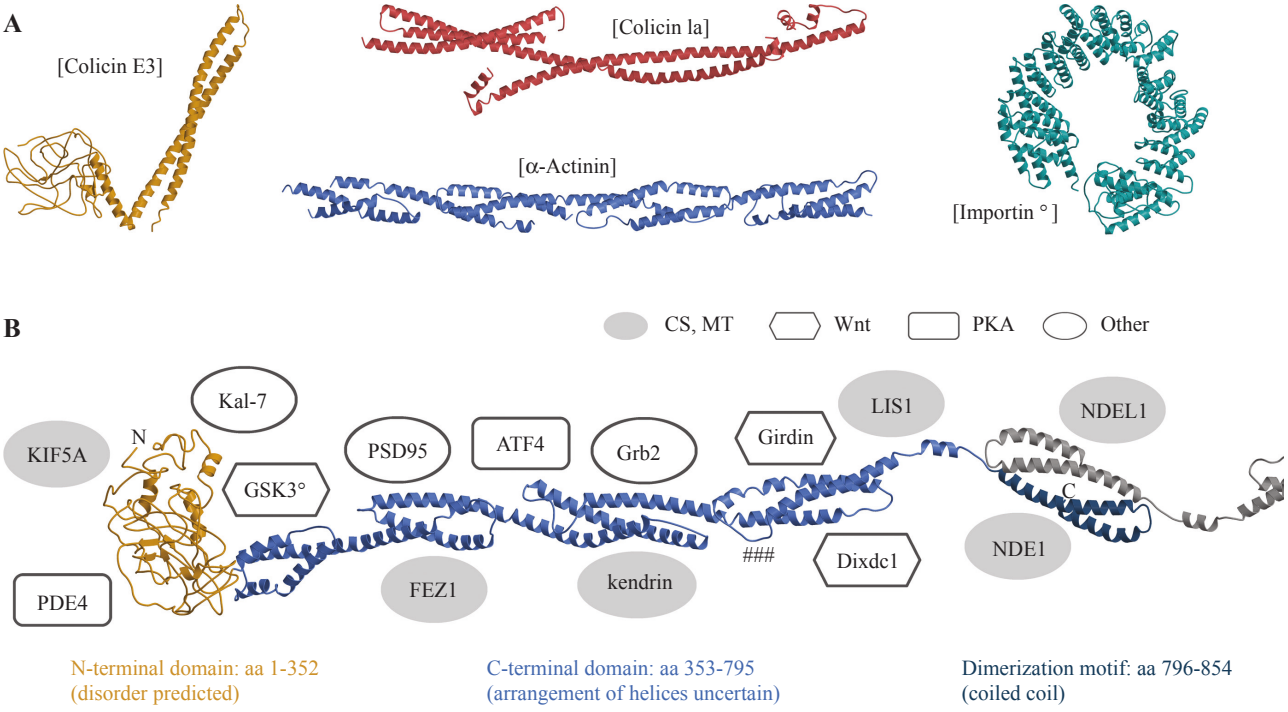

Figure 1 Speculative representations of possible structures of DISC1.
(A) Models derived from single templates (as indicated in square brackets), using profile-based alignments provided by Phyre² (for colicins and α-actinin) and I-TASSER (for importin β), respectively. Note the different arrangement of helices. (B) A speculative model of full-length DISC1 composed using MODELLER (Šali and Blundell, 1993) and modified versions of the Phyre² alignments. Specifically, amino acids 142–346, 351–773 and 782–829 were mapped to the N-terminal domain of colicin E3 (PDB code 1JCH, Soelaiman et al., 2001), the spectrin repeats of α-actinin (PDB code 1SJJ, Liu et al., 2004), and the coiled coil-domain of the transcription factor FOXP3 (PDB code 4I1L, Song et al., 2012), respectively. Regions without detectable profile matches (1–141 and 830–854) were built *ab initio* by POING (Jefferys et al., 2010), as implemented in Phyre². The resulting model was subjected to an additional energy minimization in the YASARA force field (Krieger et al., 2009). Coloring indicates three major segments of the molecule, comprising residues 1–352 (yellow), 353–795 (blue), and 796–854 (dark blue), respectively. The C-terminus of the partner molecule in a putative DISC1 dimer is shown in gray. Please note that the dimerization motif is being defined here by the terminal coiled coil region predicted to commence around residue 800. In order to visualize the scaffolding function of DISC1, several established interaction partners are indicated (CS, centrosomal protein; MT, microtubular transport; Wnt, Wnt/Glycogen Synthase Kinase 3β pathway; PKA, PKA/PDE4 signaling). Hashes denote a region implicated in DISC1 oligomerization.

Table 2 A summary of typical functions of scaffold proteins and their relevance to DISC1.

Scaffolding feature	Examples of DISC1 exhibiting the feature
Binding multiple proteins of the same signaling pathway	– Forms ternary complexes with KIF5-Growth factor Receptor-Bound protein 2, KIF5A-Nuclear Distribution Element 1, Fasciculation and Elongation protein Zeta-1-NDE-Like 1 and Kal7-Postsynaptic Density protein 95 – Potential signaling-related complexes with Glycogen Synthase Kinase 3β, DIX Domain Containing 1 and Girdin and with PDE4B and substrates of protein kinase A
Recruitment of proteins to specific cellular locations	– Upon phosphorylation recruits Bardet-Biedl syndrome 4 proteins to the centrosome – Recruits Activating Transcription Factor 4 to the nucleus – Role in protein transport via Kinesin to the axon
Modulating the enzymatic activity of binding partners	– Binds inactive PDE4B, in a protein kinase A-dependent manner – Repress Activating Transcription Factor 4 activity – Inhibits the activity of Glycogen Synthase Kinase 3β, Nuclear Distribution Element-Like 1 and Traf2 and Nck Interacting Kinase. Blocks AKT-mTOR signaling pathway via Girdin
Shielding activated signaling pathway intermediates from inappropriate consumption	– Prevents ubiquitination of serine racemase – Prevents ubiquitination of Mitofilin

From these data, it is clear that DISC1 exhibits certain aspects of a scaffold protein but contrasts with classical scaffold proteins, such as Ste5, which typically channel a single signal cascade by binding to multiple components of it. Instead DISC1 appears to be involved in the fine-tuning of a large number of inter-related signaling pathways with links to neuronal development and function (Figure 2). Strikingly there are no established examples, to the best of our knowledge, of DISC1 directly enhancing the activity of an enzyme through recruitment of substrates and co-factors or through anchoring the enzyme itself to the location of action, as would be predicted for scaffold proteins. Instead, the principal regulatory activity of DISC1 appears to be the inhibition of enzymes through direct physical contact, as is the case for Kal-7, GSK3β, PDE4, TNIK and the oligopeptidase activity of NDEL1 (Hayashi et al., 2005; Millar et al., 2005; Murdoch et al., 2007; Mao et al., 2009; Hayashi-Takagi et al., 2010; Wang et al., 2011) or through the recruitment of repressive cofactors, such as N-CoR to ATF4 (Sawamura et al., 2008). The partial exceptions are, firstly, SR which is stabilized through interaction with DISC1 (Ma et al., 2013), leading to a net increase in its activity and, secondly, that DISC1 can enhance the ability of TRIO to act as an exchange factor for Rac1, through the binding of DISC1 to a second active enzymatic site on TRIO, inhibiting this latter activity and favouring Rac1 as a substrate (Chen et al., 2011).

If DISC1 is to regulate such a diverse array of processes, it is also likely to be itself tightly regulated. In support of this, multiple PKA phosphorylation sites on DISC1 have been described which act as switches between Wnt-signaling, centrosome-related functions (Ishizuka et al., 2011) and regulation of ATF4-based transcriptional activity, respectively (Soda et al., 2013). Complex regulation of DISC1 expression is also suggested by the existence of over 40 splice variants of *DISC1* in the brain (Nakata

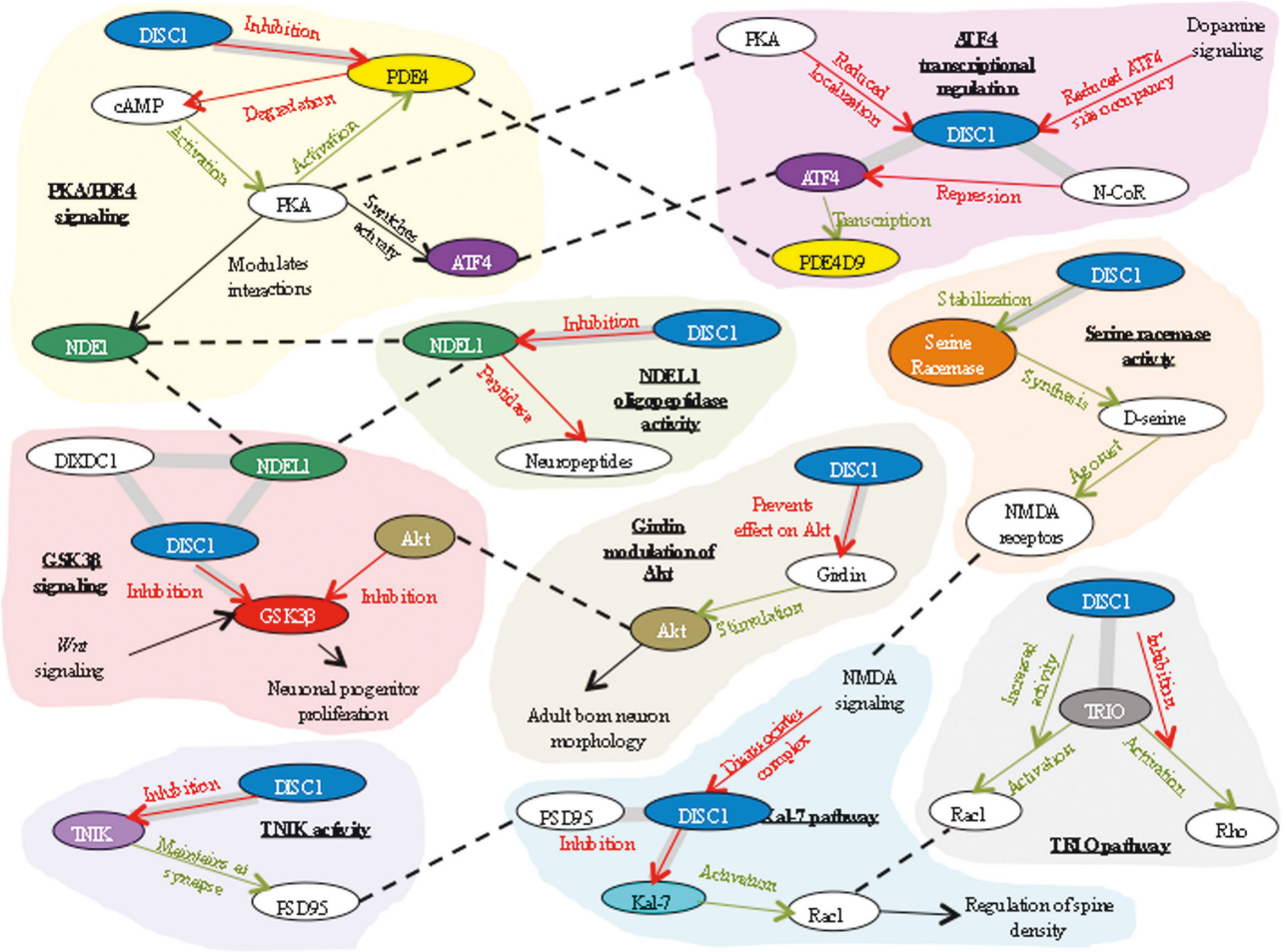

Figure 2 A schematic summary of nine interlinked signaling pathways in which DISC1 has been implicated.
Thick gray lines indicate direct protein interactions, arrows indicate interaction in the form of activation/stimulation (green), inhibition (red) or less easily defined effects (black). Identical or similar proteins found in multiple pathways are linked by hashed lines.

et al., 2009), some of which have been demonstrated to show altered protein-protein interactions (Newburn et al., 2011). This would allow for the existence of DISC1 species capable of only a subset of the functions of the full-length protein.

Another important factor likely to affect the scaffold-like functions of DISC1 is its multimerization. NDEL1 has been shown to interact with octameric, and depending on the assay used potentially dimeric, forms of DISC1, but not with higher-order oligomers (Leliveld et al., 2008; Narayanan et al., 2011), indicating that the interactions exhibit conformational specificity along with domain specificity. This also raises the obvious possibility of oligomer-specific interactions making DISC1 a multi-potent scaffold protein.

Parallels can be drawn between DISC1 and the known functions of the AKAP scaffold proteins. For example, like DISC1, AKAP9 has a complex pattern of alternate splicing (Welch et al., 2010) and like DISC1, its encoded proteins are involved in recruiting proteins to the centrosome (Shanks et al., 2002) among other locations, and can also inhibit a subset of their interaction partners, such as adenylate cyclases (Piggott et al., 2008). More generally, AKAPs form multivalent signaling complexes involving a large array of binding partners (Welch et al., 2010) and have been described as acting as the nucleus of a 'transduceosome' for a set of related signaling processes (Feliciello et al., 2001). DISC1 appears to be acting in a manner at least partially analogous to this; however, while AKAPs focus on the anchoring of the PKA protein and its substrates and associated molecules, DISC1 appears to instead form multiple seemingly distinct signaling complexes involving a wide array of signaling enzymes.

In conclusion, DISC1 possesses certain properties of a scaffold protein, but in a unique way. As it stands as a complex regulator for various interrelated pathways in neurodevelopment, unlike the classical single pathway scaffolds, it is very crucial for normal brain functioning. Further investigation of these DISC1-related pathways individually, considering its scaffolding nature, would likely contribute to a better understanding of DISC1's mechanism. Thus, in addition to being 'Disrupted in Schizophrenia', DISC1 also appears to be a 'Diverse Inhibitor of Signaling Complexes', and it is likely that modulation of this activity, for example by specific blocking of its interaction sites, would form the most promising lead for any future DISC1-inspired therapeutic intervention.

Acknowledgments: We gratefully acknowledge support (and training) from the International NRW Research School BioStruct, granted by the Ministry of Innovation, Science and Research of the State North Rhine-Westphalia, the Heinrich Heine University of Düsseldorf and the Entrepreneur Foundation at the Heinrich Heine University of Düsseldorf to C.K., as well as funding from NEURON-ERANET DISCover (BMBF 01EW1003) to C.K. and the Alexander von Humboldt Foundation to N.J.B.

Received May 14, 2013; accepted July 3, 2013; previously published online July 5, 2013

References

Bader, V., Tomppo, L., Trossbach, S.V., Bradshaw, N.J., Prikulis, I., Leliveld, S.R., Lin, C.-Y., Ishizuka, K., Sawa, A., Ramos, A., et al. (2012). Proteomic, genomic and translational approaches identify CRMP1 for a role in schizophrenia and its underlying traits. Hum. Mol. Genet. *21*, 4406–4418.

Bashor, C.J., Helman, N.C., Yan, S., and Lim, W.A. (2008). Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science *319*, 1539–1543.

Blackwood, D.H.R., Fordyce, A., Walker, M.T., St. Clair, D.M., Porteous, D.J., and 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., Christie, S., Soares, D.C., Carlyle, B.C., Porteous, D.J., and Millar, J.K. (2009). NDE1 and NDEL1: Multimerisation, alternate splicing and DISC1 interaction. Neurosci. Lett. *449*, 228–233.

Bradshaw, N.J., Ogawa, F., Antolin-Fontes, B., Chubb, J.E., Carlyle, B.C., Christie, S., Claessens, A., Porteous, D.J., and Millar, J.K. (2008). DISC1, PDE4B, and NDE1 at the centrosome and synapse. Biochem. Biophys. Res. Commun. *377*, 1091–1096.

Bradshaw, N.J. and Porteous, D.J. (2012). DISC1-binding proteins in neural development, signalling and schizophrenia. Neuropharmacology *62*, 1230–1241.

Bradshaw, N.J., Soares, D.C., Carlyle, B.C., Ogawa, F., Davidson-Smith, H., Christie, S., Mackie, S., Thomson, P.A., Porteous, D.J., and Millar, J.K. (2011). PKA phosphorylation of NDE1 is DISC1/PDE4-dependant and modulates its interaction with LIS1 and NDEL1. J. Neurosci. *31*, 9043–9054.

Brandon, N.J., Handford, E.J., Schurov, I., Rain, J.-C., Pelling, M., Duran-Jimeriz, B., Camargo, L.M., Oliver, K.R., Beher, D., Shearman, M.S., et al. (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*, 42–55.

Brandon, N.J., Millar, J.K., Korth, C., Sive, H., Singh, K.K., and Sawa, A. (2009). Understanding the role of DISC1 in psychiatric disease and during normal development. J. Neurosci. *29*, 12768–12775.

Brandon, N.J. and Sawa, A. (2011). Linking neurodevelopmental and synaptic theories of mental illness through DISC1. Nat. Rev. Neurosci. *12*, 707–722.

Burdick, K.E., Kamiya, A., Hodgkinson, C.A., Lencz, T., DeRosse, P., Ishizuka, K., Elashvili, S., Arai, H., Goldman, D., Sawa, A., et al. (2008). Elucidating the relationship between *DISC1*, *NDEL1*, and *NDE1* and the risk for schizophrenia: evidence of epistasis and competitive binding. Hum. Mol. Genet. *17*, 2462–2473.

Callicott, J.H., Straub, R.E., Pezawas, L., Egan, M.F., Mattay, V.S., Hariri, A.R., Verchinski, B.A., Meyer-Lindenberg, A., Balkissoon, R., Kolachana, B., et al. (2005). Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. Proc. Natl. Acad. Sci. USA *102*, 8627–8632.

Camargo, L.M., Collura, V., Rain, J.-C., Mizuguchi, K., Hermjakob, H., Kerrien, S., Bonnert, T.P., Whiting, P.J., and Brandon, N.J. (2007). Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. Mol. Psychiatry *12*, 74–86.

Carlyle, B.C., Mackie, S., Christie, S., Millar, J.K., and Porteous, D.J. (2011). Co-ordinated action of DISC1, PDE4B and GSK3b in modulation of cAMP signalling. Mol. Psychiatry *16*, 693–694.

Chen, S.Y., Huang, P.H., and Cheng, H.J. (2011). Disrupted-in-Schizophrenia 1-mediated axon guidance involves TRIO-RAC-PAK small GTPase pathway signaling. Proc. Natl. Acad. Sci. USA *108*, 5861–5866.

Choi, K.Y., Satterberg, B., Lyons, D.M., and Elion, E.A. (1994). Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in S. cerevisiae. Cell *78*, 499–512.

Chubb, J.E., Bradshaw, N.J., Soares, D.C., Porteous, D.J., and Millar, J.K. (2008). The *DISC* locus in psychiatric illness. Mol. Psychiatry *13*, 36–64.

Diviani, D., and Scott, J.D. (2001). AKAP signaling complexes at the cytoskeleton. J. Cell Sci. *114*, 1431–1437.

Djinovic-Carugo, K., Gautel, M., Ylanne, J., and Young, P. (2002). The spectrin repeat: a structural platform for cytoskeletal protein assemblies. FEBS Lett. *513*, 119–123.

Duan, X., Chang, J.H., Ge, S., Faulkner, R.L., Kim, J.Y., Kitabatake, Y., Liu, X.-b., Yang, C.-H., Jordan, J.D., Ma, D.K., et al. (2007). Disrupted-In-Schizophrenia 1 regulates integration of newly generated neurons in the adult brain. Cell *130*, 1146–1158.

Dunbrack, R.L., Jr. (2006). Sequence comparison and protein structure prediction. Curr. Opin. Struct. Biol. *16*, 374–384.

Elefteriou, F., Ahn, J.D., Takeda, S., Starbuck, M., Yang, X., Liu, X., Kondo, H., Richards, W.G., Bannon, T.W., Noda, M., et al. (2005). Leptin regulation of bone resorption by the sympathetic nervous system and CART. Nature *434*, 514–520.

Enomoto, A., Asai, N., Namba, T., Wang, Y., Kato, T., Tanaka, M., Tatsumi, H., Taya, S., Tsuboi, D., Kuroda, K., et al. (2009). Roles of disrupted-in-schizophrenia 1-interacting protein girdin in postnatal development of the dentate gyrus. Neuron *63*, 774–787.

Feliciello, A., Gottesman, M.E., and Avvedimento, E.V. (2001). The biological functions of A-kinase anchor proteins. J. Mol. Biol. *308*, 99–114.

Gadelha, A., Machado, M.F.M., Yonamine, C.M., Sato, J.R., Juliano, M.A., Oliveira, V., Bressan, R.A., and Hayashi, M.A.F. (2013). Plasma Ndel1 enzyme activity is reduced in patients with schizophrenia – A potential biomarker? J. Psychiatr. Res. *47*, 657–663.

Good, M.C., Zalatan, J.G., and Lim, W.A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. Science *332*, 680–686.

Hasson, M.S., Blinder, D., Thorner, J., and Jenness, D.D. (1994). Mutational activation of the STE5 gene product bypasses the requirement for G protein β and γ subunits in the yeast pheromone response pathway. Mol. Cell Biol. *14*, 1054–1065.

Hayashi-Takagi, A., Takaki, M., Graziane, N., Seshadri, S., Murdoch, H., Dunlop, A.J., Makino, Y., Seshadri, A.J., Ishizuka, K., Srivastava, D.P., et al. (2010). Disrupted-in-Schizophrenia 1 (DISC1) regulates spines of the glutamate synapse via Rac1. Nat. Neurosci. *13*, 327–332.

Hayashi, M.A.F., Portaro, F.C.V., Bastos, M.F., Guerreiro, J.R., Oliveira, V., Gorrao, S.S., Tambourgi, D.V., Sant'Anna, O.A., Whiting, P.J., Camargo, L.M., et al. (2005). Inhibition of NUDEL (nuclear distribution element-like)-oligopeptidase activity by disrupted-in-schizophrenia 1. Proc. Natl. Acad. Sci. USA *102*, 3828–3833.

Huang, G.N., Huso, D.L., Bouyain, S., Tu, J., McCorkell, K.A., May, M.J., Zhu, Y., Lutz, M., Collins, S., Dehoff, M., et al. (2008). NFAT binding and regulation of T cell activation by the cytoplasmic scaffolding Homer proteins. Science *319*, 476–481.

Hurme, R., Berndt, K.D., Namork, E., and Rhen, M. (1996). Additions and Corrections to DNA binding exerted by a bacterial gene regulator with an extensive coiled-coil domain. J. Biol. Chem. *271*, 17547.

Ishizuka, K., Kamiya, A., Oh, E.C., Kanki, H., Seshadri, S., Robinson, J.F., Murdoch, H., Dunlop, A.J., Kubo, K.-i., Furukori, K., et al. (2011). DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. Nature *473*, 92–96.

Jefferys, B.R., Kelley, L.A., and Sternberg, M.J.E. (2010). Protein folding requires crowd control in a simulated cell. J. Mol. Biol. *397*, 1329–1338.

Kamiya, A., Kubo, K.-i., Tomoda, T., Takaki, M., Youn, R., Ozeki, Y., Sawamura, N., Park, U., Kudo, C., Okawa, M., et al. (2005). A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. Nat. Cell Biol. *7*, 1167–1178.

Kamiya, A., Tan, P.L., Kubo, K., Engelhard, C., Ishizuka, K., Kubo, A., Tsukita, S., Pulver, A.E., Nakajima, K., Cascella, N.G., et al. (2008). Recruitment of PCM1 to the centrosome by the cooperative action of DISC1 and BBS4: a candidate for psychiatric illnesses. Arch. Gen. Psychiatry *65*, 996–1006.

Kamiya, A., Tomoda, T., Chang, J., Takaki, M., Zhan, C., Morita, M., Cascio, M.B., Elashvili, S., Koizumi, H., Takanezawa, Y., et al. (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.

Kang, E., Burdick, Katherine E., Kim, Ju Y., Duan, X., Guo, Junjie U., Sailor, Kurt A., Jung, D.-E., Ganesan, S., Choi, S., Pradhan, D., et al. (2011). Interaction between FEZ1 and DISC1 in regulation of neuronal development and risk for schizophrenia. Neuron *72*, 559–571.

Karplus, K., Barrett, C., and Hughey, R. (1998). Hidden Markov models for detecting remote protein homologies. Bioinformatics *14*, 846–856.

Kelley, L.A. and Sternberg, M.J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

Kim, J.Y., Duan, X., Liu, C.Y., Jang, M.H., Guo, J.U., Pow-anpongkul, N., Kang, E., Song, H., and Ming, G.L. (2009). DISC1 regulates new neuron development in the adult brain via modulation of AKT-mTOR signaling through KIAA1212. Neuron *63*, 761–773.

Kim, J.Y., Liu, C.Y., Zhang, F., Duan, X., Wen, Z., Song, J., Feighery, E., Lu, B., Rujescu, D., Clair, D.S., et al. (2012). Interplay between DISC1 and GABA signaling regulates neurogenesis in mice and risk for schizophrenia. Cell *148*, 1051–1064.

Korth, C. (2012). Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. Prion *6*, 1–8.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins *77*, 114–122.

Kuroda, K., Yamada, S., Tanaka, M., Iizuka, M., Yano, H., Mori, D., Tsuboi, D., Nishioka, T., Namba, T., Iizuka, Y., et al. (2011). Behavioral alterations associated with targeted disruption of exons 2 and 3 of the Disc1 gene in the mouse. Hum. Mol. Genet. *20*, 4666–4683.

Leliveld, S.R., Bader, V., Hendriks, P., Prikulis, I., Sajnani, G., Requena, J.R., and 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.

Li, Y., Chen, L., Kass, R.S., and Dessauer, C.W. (2012). The A-kinase anchoring protein yotiao facilitates complex formation between adenylyl cyclase type 9 and the IKs potassium channel in heart. J. Biol. Chem. *287*, 29815–29824.

Lipina, T.V., Wang, M., Liu, F., and Roder, J.C. (2012). Synergistic interactions between PDE4B and GSK-3: DISC1 mutant mice. Neuropharmacology *62*, 1252–1262.

Liu, J., Taylor, D.W., and Taylor, K.A. (2004). A 3-D reconstruction of smooth muscle α-actinin by CryoEM reveals two different conformations at the actin-binding region. J. Mol. Biol. *338*, 115–125.

Locasale, J.W., Shaw, A.S., and Chakraborty, A.K. (2007). Scaffold proteins confer diverse regulatory properties to protein kinase cascades. Proc. Natl. Acad. Sci. USA *104*, 13307–13312.

Ma, T.M., Abazyan, S., Abazyan, B., Nomura, J., Yang, C., Seshadri, S., Sawa, A., Snyder, S.H., and Pletnikov, M.V. (2013). Pathogenic disruption of DISC1-serine racemase binding elicits schizophrenia-like behavior via D-serine depletion. Mol. Psychiatry *18*, 557–567.

Mao, Y., Ge, X., Frank, C.L., Madison, J.M., Koehler, A.N., Doud, M.K., Tassa, C., Berry, E.M., Soda, T., Singh, K.K., et al. (2009). Disrupted in Schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3b/b-Catenin signaling. Cell *136*, 1017–1031.

Millar, J.K., Christie, S., and Porteous, D.J. (2003). Yeast two-hybrid screens implicate DISC1 in brain development and function. Biochem. Biophys. Res. Commun. *311*, 1019–1025.

Millar, J.K., Pickard, B.S., Mackie, S., James, R., Christie, S., Buchanan, S.R., Malloy, M.P., Chubb, J.E., Huston, E., Baille, G.S., et al. (2005). DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signalling. Science *310*, 1187–1191.

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., et al. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. Hum. Mol. Genet. *9*, 1415–1425.

Miyoshi, K., Honda, A., Baba, K., Taniguchi, M., Oono, K., Fujita, T., Kuroda, S., Katayama, T., and Tohyama, M. (2003). *Disrupted-in-Schizophrenia 1*, a candidate gene for schizophrenia, participates in neurite outgrowth. Mol. Psychiatry *8*, 685–694.

Morris, J.A., Kandpal, G., Ma, L., and 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*, 1591–1608.

Muller, J., Ory, S., Copeland, T., Piwnica-Worms, H., and Morrison, D.K. (2001). C-TAK1 regulates Ras signaling by phosphorylating the MAPK scaffold, KSR1. Mol. Cell *8*, 983–993.

Murdoch, H., Mackie, S., Collins, D.M., Hill, E.V., Bolger, G.B., Klussmann, E., Porteous, D.J., Millar, J.K., and Houslay, M.D. (2007). Isoform-selective susceptibility of DISC1/Phosphodiesterase-4 complexes to dissociation by elevated intracellular cAMP levels. J. Neurosci. *27*, 9513–9524.

Nakata, K., Lipska, B.K., Hyde, T.M., Ye, T., Newburn, E.N., Morita, Y., Vakkalanka, R., Barenboim, M., Sei, Y., Weinberger, D.R., et al. (2009). *DISC1* splice variants are upregulated in schizophrenia and associated with risk polymorphisms. Proc. Natl. Acad. Sci. USA *106*, 15873–15878.

Narayanan, S., Arthanari, H., Wolfe, M.S., and 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.

Newburn, E.N., Hyde, T.M., Ye, T., Morita, Y., Weinberger, D.R., Kleinman, J.E., and Lipska, B.K. (2011). Interactions of human truncated DISC1 proteins: implications for schizophrenia. Transl. Psychiatry *1*, e30.

Niwa, M., Kamiya, A., Murai, R., Kubo, K.-i., Gruber, A.J., Tomita, K., Lu, L., Tomisato, S., Jaaro-Peled, H., Seshadri, S., et al. (2010). Knockdown of DISC1 by in utero gene transfer disturbs postnatal dopaminergic maturation in the frontal cortex and leads to adult behavioral deficits. Neuron *65*, 480–489.

Ottis, P., Bader, V., Trossbach, S.V., Kretzschmar, H., Michel, M., Leliveld, S.R., and Korth, C. (2011). Convergence of two independent mental disease genes on the protein level: Recruitment of dysbindin to cell-invasive DISC1 aggresomes. Biol. Psychiatry *70*, 604–610.

Ozeki, Y., Tomoda, T., Kleiderlein, J., Kamiya, A., Bord, L., Fujii, K., Okawa, M., Yamada, N., Hatten, M.E., Snyder, S.H., et al. (2003). Disrupted-in-Schizophrenia-1 (DISC-1): mutant truncation prevents binding to NudE-like (NUDEL) and inhibits neurite outgrowth. Proc. Natl. Acad. Sci. USA *100*, 289–294.

Park, Y.-U., Jeong, J., Lee, H., Mun, J.Y., Kim, J.-H., Lee, J.S., Nguyen, M.D., Han, S.S., Suh, P.-G., and Park, S.K. (2010). *Disrupted-in-schizophrenia* 1 (*DISC1*) plays essential roles in mitochondria in collaboration with Mitofilin. Proc. Natl. Acad. Sci. USA *107*, 17785–17790.

Piggott, L.A., Bauman, A.L., Scott, J.D., and Dessauer, C.W. (2008). The A-kinase anchoring protein Yotiao binds and regulates adenylyl cyclase in brain. Proc. Natl. Acad. Sci. USA *105*, 13835–13840.

Pletnikov, M.V., Xu, Y., Ovanesov, M.V., Kamiya, A., Sawa, A., and Ross, C.A. (2007). PC12 cell model of inducible expression of mutant DISC1: new evidence for a dominant-negative mechanism of abnormal neuronal differentiation. Neurosci. Res. *58*, 234–244.

Popescu, D.C., Ham, A.J., and Shieh, B.H. (2006). Scaffolding protein INAD regulates deactivation of vision by promoting phosphorylation of transient receptor potential by eye protein kinase C in *Drosophila*. J. Neurosci. *26*, 8570–8577.

Rose, A. and Meier, I. (2004). Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. Cell Mol. Life Sci. *61*, 1996–2009.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. *5*, 725–738.

Šali, A. and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol *234*, 779–815.

Sanchez-Pulido, L. and Ponting, C.P. (2011). Structure and evolutionary history of DISC1. Hum. Mol. Genet. *20*, R175–R181.

Sawamura, N., Ando, T., Maruyama, Y., Fujimuro, M., Mochizuki, H., Honjo, K., Shimoda, M., Toda, H., Sawamura-Yamamoto, T., Makuch, L.A., et al. (2008). Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. Mol. Psychiatry *13*, 1138–1148.

Shanks, R.A., Steadman, B.T., Schmidt, P.H., and Goldenring, J.R. (2002). AKAP350 at the Golgi apparatus. I. Identification of a distinct Golgi apparatus targeting motif in AKAP350. J. Biol. Chem. *277*, 40967–40972.

Shaw, A.S. and Filbert, E.L. (2009). Scaffold proteins and immune-cell signalling. Nat. Rev. Immunol. *9*, 47–56.

Shen, S., Lang, B., Nakamoto, C., Zhang, F., Pu, J., Kuan, S.-L., Chatzi, C., He, S., Mackie, I., Brandon, N.J., et al. (2008). Schizophrenia-related neural and behavioral phenotypes in transgenic mice expressing truncated Disc1. J. Neurosci. *28*, 10893–10904.

Shinoda, T., Taya, S., Tsuboi, D., Hikita, T., Matsuzawa, R., Kuroda, S., Iwamatsu, A., and Kaibuchi, K. (2007). DISC1 regulates Neurotrophin-induced axon elongation via interaction with Grb2. J. Neurosci. *27*, 4–14.

Singh, K.K., Ge, X., Mao, Y., Drane, L., Meletis, K., Samuels, B.A., and Tsai, L.-H. (2010). Dixdc1 is a critical regulator of DISC1 and embryonic cortical development. Neuron *67*, 33–48.

Soares, D.C., Carlyle, B.C., Bradshaw, N.J., and Porteous, D.J. (2011). DISC1: structure, function, and therapeutic potential for major mental illness. ACS Chem. Neurosci. *2* 609–632.

Soda, T., Frank, C., Ishizuka, K., Baccarella, A., Park, Y.-U., Flood, Z., Park, S.K., Sawa, A., and Tsai, L.-H. (2013). DISC1-ATF4 transcriptional repression complex: dual regulation of the cAMP-PDE4 cascade by DISC1. Mol Psychiatry doi:10.1038/mp.2013.38.

Soelaiman, S., Jakes, K., Wu, N., Li, C., and Shoham, M. (2001). Crystal structure of colicin E3: implications for cell entry and ribosome inactivation. Mol. Cell *8*, 1053–1062.

Song, X., Li, B., Xiao, Y., Chen, C., Wang, Q., Liu, Y., Berezov, A., Xu, C., Gao, Y., Li, Z., et al. (2012). Structural and biological features of FOXP3 dimerization relevant to regulatory T cell function. Cell Rep. *1*, 665–675.

Szilak, L., Moitra, J., Krylov, D., and Vinson, C. (1997). Phosphorylation destabilizes α-helices. Nat. Struct. Biol. *4*, 112–114.

Taya, S., Shinoda, T., Tsuboi, D., Asaki, J., Nagai, K., Hikita, T., Kuroda, S., Kuroda, K., Shimizu, M., Hirotsune, S., et al. (2007). DISC1 regulates the transport of the NUDEL/LIS1/14-3-3epsilon complex through kinesin-1. J. Neurosci. *27*, 15–26.

Thomson, P.A., Harris, S.E., Starr, J.M., Whalley, L.J., Porteous, D.J., and Deary, I.J. (2005). Association between genotype at an exonic SNP in DISC1 and normal cognitive aging. Neurosci. Lett. *389*, 41–45.

Thomson, P.A., Malavasi, E.L.V., Grünewald, E., Soares, D.C., Borkowska, M., and Millar, J.K. (2013). DISC1 genetics, biology and psychiatric illness. Front. Biol. *8*, 1–31.

Wang, Q. and Brandon, N.J. (2011). Regulation of the cytoskeleton by Disrupted-in-Schizophrenia 1 (DISC1). Mol. Cell Neurosci. *48*, 359–364.

Wang, Q., Charych, E.I., Pulito, V.L., Lee, J.B., Graziane, N.M., Crozier, R.A., Revilla-Sanchez, R., Kelly, M.P., Dunlop, A.J., Murdoch, H., et al. (2011). The psychiatric disease risk factors DISC1 and TNIK interact to regulate synapse composition and function. Mol. Psychiatry *16*, 1006–1023.

Welch, E.J., Jones, B.W., and Scott, J.D. (2010). Networking with AKAPs: context-dependent regulation of anchored enzymes. Mol. Interv. *10*, 86–97.

Zeke, A., Lukacs, M., Lim, W.A., and Remenyi, A. (2009). Scaffolds: interaction platforms for cellular signalling circuits. Trends. Cell Biol. *19*, 364–374.

Antony Sravan Kumar Yerabham completed his Master's degree at the University of Hyderabad, India in the year 2010. Following this, he received training in Macromolecular Crystallography as a Project Junior Research Fellow in Dr. Rajan Sankaranarayanan's group at the Centre for Cellular and Molecular Biology, Hyderabad. Currently he is pursuing his Doctoral studies on the structure and function of the DISC1 protein and its interaction partners, in the work group of Prof. Dr. Carsten Korth, at the Department of Neuropathology, Heinrich Heine University, Düsseldorf, Germany.

Nicholas Bradshaw received his BSc in Natural Sciences (Biology with Physics) from Durham University in 2005 and undertook his doctorate at the University of Edinburgh under the supervision of Dr Kirsty Millar and Prof David Porteous on the interactions of the schizophrenia-related proteins DISC1 and NDE1. After receiving his PhD in 2009, Nick continued his research into the biochemical and biophysical properties of these and related proteins pertinent to major mental health in Edinburgh. In 2011 he received a postdoctoral fellowship from the Alexander von Humboldt Foundation to further develop his research in the laboratory of Prof Dr Carsten Korth at the Heinrich Heine University, Düsseldorf.

Carsten Korth did his MD at the LMU Munich and PhD at the VU Amsterdam. After a residency in psychiatry at the Max Planck Institute for Psychiatry, he moved to fundamental research in protein conformational disease, first to the University of Zurich where he also co-founded the company Prionics, then to the Institute for Neurodegenerative Diseases, University of California San Francisco (Prof Stanley Prusiner). Since 2002 he is independent investigator at the Heinrich Heine University of Düsseldorf focussing on protein pathology in brain diseases, particularly chronic mental illnesses like schizophrenia or recurrent affective disorders.

Publication II

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.

Antony S. K. Yerabham, Philippe J. Mas, Christina Decker, Dinesh C. Soares,
Oliver H. Weiergräber, Luitgard Nagel-Steger, Dieter Willbold. Darren J. Hart,
Nicholas J. Bradshaw, and Carsten Korth

Journal of Biological Chemistry 2017 Apr 21; 292(16):6468-6477

DOI: 10.1074/jbc.M116.773903

JBC Papers in Press. Published on March 1, 2017 as Manuscript M116.773903
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M116.773903

A Structural Organization for Disrupted in Schizophrenia 1, Identified by High-Throughput Screening, Reveals Distinctly Folded Regions Which Are Bisected by Mental Illness-Related Mutations

Antony S. K. Yerabham[‡], Philippe J. Mas[§], Christina Decker[¶], Dinesh C. Soares[|], Oliver H. Weiergräber[‡‡], Luitgard Nagel-Steger[¶,‡‡], Dieter Willbold[¶,‡‡]. Darren J. Hart[§§], Nicholas J. Bradshaw[‡,1], and Carsten Korth[‡,1]

From the [‡]Department of Neuropathology and the [¶]Institute of Physical Biology, Heinrich Heine University, 40225 Düsseldorf, Germany, [§]Integrated Structural Biology Grenoble (ISBG) CNRS, CEA, Université Grenoble Alpes, EMBL, 38042 Grenoble, France, the [|]MRC Human Genetics Unit/Centre for Genomic and Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, U.K., the [‡‡]Institute of Complex Systems (ICS-6: Structural Biochemistry), Forschungszentrum Jülich, 52425 Jülich, Germany and the [§§]Institut de Biologie Structurale (IBS), CEA, CNRS, Université Grenoble Alpes, 38044 Grenoble, France.

Running title: *DISC1 domain architecture*

[1]To whom correspondence should be addressed: Prof. Dr. Carsten Korth & Dr. Nicholas J. Bradshaw, Institut für Neuropathologie, Universitätskliniken Düsseldorf, Moorenstraße 5, 40225 Düsseldorf, Germany, Telephone: +49 211 8118653 (NJB), +49 211 8116153 (CK); Fax: +49 211 8117804; E-mail: nicholas.bradshaw@hhu.de (NJB), ckorth@hhu.de (CK)

Keywords: Analytical ultracentrifugation, Circular dichroism, DISC1, Domain structure, ESPRIT, Mental illness, Protein solubility, Protein stability, Recombinant protein expression, Schizophrenia

ABSTRACT

Disrupted in Schizophrenia 1 (DISC1) is a scaffolding protein of significant importance for neurodevelopment and a prominent candidate protein in the pathology of major mental illness. DISC1 modulates a number of critical neuronal signaling pathways through protein-protein interactions; however, the mechanism by which this occurs and how DISC1 causes mental illness is unclear, partly because knowledge of the structure of DISC1 is lacking. A lack of homology with known proteins has hindered attempts to define its domain composition. Here, we employed the high-throughput Expression of Soluble Proteins by Random Incremental Truncation (ESPRIT) technique to identify discretely folded regions of human DISC1 *via* solubility assessment of tens of thousands of fragments of recombinant DISC1. We identified four novel structured regions, named D, I, S, and C, at amino acids 257-383, 539-655, 635-738, and 691-836, respectively. One region (D) is located in a DISC1 section previously predicted to be unstructured. All regions encompass coiled-coil or α- helical structures and three are involved in DISC1 oligomerization. Crucially, three of these domains would be lost or disrupted in a chromosomal translocation event after amino acid 597, which has been strongly linked to major mental illness. Furthermore, we observed that a known illness-related frameshift mutation after amino acid 807 causes the C region to form aberrantly multimeric and aggregated complexes with an unstable secondary structure. This newly revealed domain architecture of DISC1, therefore, provides a powerful framework for understanding the critical role of this protein in a variety of devastating mental illnesses.

Disrupted in Schizophrenia 1 (*DISC1*) was initially identified as being of importance for mental illness due to its disruption by a chromosomal translocation in a Scottish family with schizophrenia, depression and related psychiatric illnesses (1,2). At the protein level, its translocation would cause loss of all of the amino acids C-terminal of residue 597 in the 854-amino acid long human DISC1 protein. A

second family was subsequently reported where a C-terminal frameshift mutation following amino acid 807 was linked to mental illness (3). Since then it has been established that the DISC1 protein acts as a scaffold, interacting with a considerable number of protein binding partners in order to modulate a wide variety of signaling pathways, many of which are of significant importance for neurodevelopment (4-6). Progress in understanding the mechanisms by which this modulation occurs has been limited, however, in part because of the dearth of information concerning the basic structure of DISC1.

To date, domain delineation has been limited to bioinformatics predictions that suggest the first 325 amino acids of DISC1 are disordered, while the remainder of the protein (from 326 or 350 onwards) is rich in α-helical and/or coiled-coil structure (1,7,8). In addition to a deduced self-association domain around residues 400-500, identified via a mutation analysis conducted in cell culture (6) several biophysical studies of on DISC1 protein have revealed distinct dimerization and oligomerization domains towards the C-terminus of DISC1 (9) and provided evidence that the full-length protein forms octamers and dimers (10). The major barriers impeding the characterization of the DISC1 protein are the low solubility and high aggregation propensity of recombinant DISC1, except for when fused to maltose binding protein (10) or when insoluble fragments are solubilized in urea and refolded *in vitro* (9). Moreover, the *DISC1* gene is also evolving very rapidly, with the encoded protein having a low degree of sequence conservation across species and lacking significant homology with any other known proteins (7,8,11,12). This hinders attempts to identify functional domains, which would facilitate the expression of discretely folded DISC1 regions.

To address this latter issue, we utilized the Expression of Soluble Protein by Random Incremental Truncation (ESPRIT) technology (13,14). This technique, developed to identify folded structural domains of uncharacterized proteins, involves the random truncation of a gene of interest at one or both ends. The ensuing library of tens of thousands of truncated variants is then expressed in bacteria and screened for solubility. Clones encoding the most soluble protein fragments are sequenced to identify construct boundaries before being further characterized, based on the hypothesis that such soluble, well-expressed protein fragments often correspond to distinctly folded domains within the full-length protein. This approach has previously led to the identification of unpredicted domains within several proteins, and in some instances to solved three-dimensional structures of such domains (13,15-17).

By employing this technique, we identified and characterized four novel folded regions of the human DISC1 protein that we propose as the basic domain architecture of the protein. Significantly, two mutations associated with psychiatric illness in families lie within these regions. These results, therefore, provide a powerful basis for further structural and functional studies of the DISC1 protein and its relevance to mental health.

Results

Identification of four novel structural regions within DISC1 by ESPRIT screening - A version of the wild-type human *DISC1* gene, codon-optimized for efficient expression in *E. coli*, was cloned into a plasmid vector encoding hexa-histidine (His_6) and biotin acceptor peptide (BAP) tags at the N- and C-termini respectively. The gene was then subjected to random incremental truncation from both the 5′ and 3′ ends, generating a library of truncated DISC1 gene fragments. This library was then constrained to those DNA fragments encoding approximately 150-400 amino acid (AA) protein fragments in order to focus on constructs of typical domain size. Twenty-one randomly picked clones were sequenced to ensure an even distribution of construct lengths within this range, as well as coverage of the entire length of the DISC1 protein

A total of 27,652 clones (72 × 384 well plates) were arrayed robotically onto nitrocellulose membranes to generate colony filters in which encoded DISC1 fragments were expressed. Identification of putatively soluble expression constructs was by colony blot hybridization to detect His_6 and BAP signals for each clone. The 94 highest double-signaling constructs from the array were isolated, expressed in liquid culture format and assayed directly for solubility by nickel-NTA purification, SDS-PAGE and western blotting. Thirty-one constructs yielded detectable DISC1 protein fragments which could be purified, and these were sequenced to identify construct boundaries. Most of these constructs are located

in the C-terminal half of the protein (AA 427-854) (Fig. 1A), with only two constructs being found in the N-terminal half (AA 1-426).

In order to determine which of these constructs were most likely to express folded, stable, soluble proteins representative of the domains of the DISC1 protein *in vivo*, the BAP tags were deleted from the constructs, and then each protein was purified and analyzed by size exclusion chromatography (SEC). Most of these recombinant proteins displayed a high aggregation propensity, eluting largely in the void volume and with a highly variable oligomeric state. However, six constructs, representing four different regions of the DISC1 protein, were soluble with consistently stable oligomeric states by SEC, indicative of folded, stable regions, as will be described below. We, therefore, propose these four newly defined regions to represent at least part of the basic domain architecture of the full-length DISC1 protein and will be referred to here (from N-terminus to C-terminus) as the D, I, S and C regions (Fig. 1A,B). Notably, however, two of the regions (S and C) lie C-terminal to the Scottish family translocation breakpoint, while the I region lies across it, implying that any truncated or fusion protein produced from the translocation chromosome would lack most of its major folded elements.

All four regions, when expressed as recombinant proteins at 0.5 mg/ml were stable, showing no signs of visible precipitation after 48 hours at room temperature (25 °C). Furthermore, the S region was stable for at least a week and the C region over even longer time periods, even at concentrations of 10 mg/ml. The biophysical characterization of these four regions is described below and is summarized in Table 1.

The D region is an unpredicted dimeric structure in the N-terminal half of DISC1 - We define the D region based on the two highly similar truncation constructs, corresponding to AA 249-383 and 257-400 of full-length DISC1. These constructs were the only soluble constructs detected in the N-terminal half of DISC1 during the solubility screen. SEC analysis of the purified D region revealed multiple species (Fig. 2A). As these fractions all contained the D region fragment, as assayed by SDS-PAGE, and the elution points were roughly consistent with the different species having molecular weights which were multiples of each other, these were interpreted as representing different oligomeric states of the D region. When eluted peaks were re-analyzed by SEC after 24 hours at 4 °C, the formerly most abundant species remained the most prominent, although with some shift in equilibrium towards a lower oligomeric state (Fig. 2A). Analytical ultracentrifugation (AUC) sedimentation equilibrium (SE) experiments on the most prominent oligomeric state revealed it to have a molecular weight (MW) of 38.8 ± 0.2 kDa, strongly indicative of a dimer (Fig. 2B, the dimer is predicted to be 35.4 kDa). Despite predictions that this region of DISC1 would incorporate a large swathe of protein disorder (8), circular dichroism (CD) of the D region yields a spectrum consistent with a protein of high α-helical content (Fig. 2C).

The I region forms a distinct helical oligomeric species with a high aggregation propensity - We define the I region based on a single ESPRIT construct encoding DISC1 AA 539-655. Like the other three proposed regions, the I region can be expressed in *E. coli* as a soluble protein with a stable, low-order oligomeric state (Fig. 2D); however unlike the other three regions there is an apparent upper limit to this, with the purified proteins exhibiting visible precipitation when concentrated above approximately 30 μM. Multiple other ESPRIT-derived constructs overlap the AA 539-655 segment and similarly precipitate, although these do so immediately upon purification and are found abundantly in the void volume after SEC (Fig. 2E). Of the truncated DISC1 species detected in this region by the solubility screen, the AA 539-655 fragment is, therefore, unique in being expressed as a soluble, stable protein under these conditions, and is therefore defined here as a structured region. SEC shows the I region is dimeric, although AUC SE analysis was not possible due to the upper concentration limit, while CD spectrometry showed it to be predominantly helical (Fig. 2F).

The S region is a highly elongated tetramer, which overlaps with neighboring regions - We define the S region based on a construct encoding DISC1 AA 635-738. It overlaps with both the I and C regions, and is a stable tetramer. Specifically, SEC demonstrates the region to express as multiple different oligomeric states, of which one was consistently seen to be the most abundant. After a 24 hour period, SEC of this principal oligomeric state

showed it retained its prominence (Fig. 3A). AUC SE analysis shows the preparations of this species have a mean MW of 68.8 ± 2.3 kDa, which is between a tetramer and a pentamer (predicted MW: 55.6 kDa, Fig. 3B). Of these, the tetramer is the more likely conformation, with the higher apparent MW being the result of larger oligomeric species forming during the concentration of tetrameric species at the bottom of the cell by AUC.

Notably, while recombinant I or S regions are soluble and stable when expressed alone, albeit with differing oligomeric states, constructs which span both of these regions are unstable with regard to their oligomerization and show low solubility (Fig. 2E). We, therefore, consider them to be distinct regions, potentially representing alternative conformations of full-length DISC1. AUC sedimentation velocity (SV) analysis resulted in a c(s) distribution with a weighted average frictional ratio of 1.8, suggesting that the tetramer is highly elongated, while CD confirms it to be mostly helical (Fig. 3C).

The C region is an elongated α-helical monomer, containing some disordered content - We defined the C region of DISC1 based on two constructs encoding AA 691-836 and 684-836 of full-length DISC1. Upon purification, the recombinant C region exists as both a prominent and a minor species according to SEC, however only the more prominent, lower molecular weight species was stable (Fig. 3D). AUC SE analysis showed this species to have a MW of 20.8 ± 0.2 kDa, strongly indicative of a monomer (Fig. 3E, predicted MW: 20.1 kDa), while AUC SV analysis of the species indicated it to possess an elongated shape (weighted average frictional ratio: 1.45). CD showed this region to be helical, with some disordered content (approximately 25% random coil, based on deconvolution using the CDSSTR method in DichroWeb with reference dataset 4 (18-21), Fig. 3F). An additional construct lies within the N-terminal half of the C region, representing AA 718-771 of DISC1. This fragment is also predicted as monomeric by SEC (Fig. 3G), while CD revealed a lack of regular secondary structure (Fig. 3H). This section of the C region thus appears to be responsible for the random coil content of the region, and is consistent with the previous prediction of a short unstructured segment in this area (8).

A frameshift mutation implicated in schizophrenia and schizoaffective disorder leads to an aggregation-prone C region - As a proof-of-principle to show how knowledge of the domain architecture of DISC1 can aid in the understanding of mental illness, we considered a frameshift mutation at AA 807 in DISC1 which was previously observed in the case of an American family with schizophrenia and schizoaffective disorders (3). This mutation leads to the addition of 9 amino acids of non-specific read-through followed by a premature stop codon (3). Notably, this mutation lies within our newly defined C region, and would, therefore, be predicted to disrupt the structure of this folded, soluble region.

Two recombinant C region proteins were therefore expressed in parallel in *E. coli*, one representing the wild-type sequence and one mimicking the truncation resulting from the frameshift mutation, with the addition of the resulting 9 amino acids and stop codon deriving from it. While this mutant protein still expressed at 50% the level of the wild-type protein (Fig 4A), its solubility decreased by 90% (Fig 4B), due to insoluble, and potentially unfolded, proteins being deposited in inclusion bodies. To investigate this further, the inclusion bodies of bacteria expressing wild-type or frameshift-carrying DISC1 C region were denatured with urea buffer and the recombinant C regions were then purified and refolded *in vitro*. The ensuing mutant C region elutes earlier in SEC than the wild-type version, in spite of its smaller number of amino acids (Fig. 4C,D). When analyzed further in AUC SE experiments, the mutant form was confirmed to exist as a larger protein complex, consistent with aberrant multimerization (average MWs determined through AUC: wild-type 20.5 kDa, mutant 25 kDa), but also to have a higher aggregation propensity (Fig. 4E,F). Finally, the mutant C region showed a lower level of α-helix forming propensity than the wild-type (Fig. 4G). The secondary structure signal detected by CD is lost over time when stored at room temperature, potentially due to precipitation, unlike the wild-type protein which remained a structurally stable entity under the same conditions.

Together, these experiments show that the effect of the frameshift mutation on the C region structure is sufficient to induce aberrant oligomer formation, ultimately leading to the formation of aggregating protein complexes,

despite the absence of regions which cause the region to form oligomers in its wild-type state.

Discussion

Understanding the structure of a protein is an important step towards understanding its function, however, progress on DISC1 has been hampered by insolubility of the recombinant protein and a lack of understanding of its domain architecture. Here, we have addressed these technical difficulties by using the ESPRIT library screening technique to scan for readily soluble and stable regions within the DISC1 protein. In previous experiments employing the ESPRIT technique, generally employed to investigate enzymes, distinct globular domains were determined: regions which were identified by multiple different ESPRIT constructs that varied only by a small number of amino acids encoded at the 5′ and/or 3′ ends of the cDNA (13,15-17). These constructs represented compactly folded domains with hydrophobic cores, with similar variants having only small unstructured terminal extensions. In contrast, and reflecting the nature of DISC1 as a primarily helical scaffold protein (5,8), this screen, when coupled to purification and biophysical analysis, instead identified four helical regions within DISC1 as being both soluble and structured. We have defined these as the D, I, S and C regions and propose them to represent distinct, structured regions of the human DISC1 protein, and which coincide with established binding sites for several of the key protein interaction partners of DISC1 (Fig. 5A).

Additionally, in light of this proposed domain architecture, some previous work into DISC1 must be re-evaluated, where fragments of the protein defined arbitrarily or based on bioinformatic predictions were expressed, to determine how they relate to this architecture. Protein fragments containing only part of these structured regions could, for example, appear to show loss or gain of function as a result of aberrant or a lack of folding.

A caveat to this work, however, is the extent to which protein expressed in large quantities in *E. coli* would be representative of the low levels of endogenous full-length human DISC1 expressed *in vivo*. That the four described regions are soluble even at very high concentrations, combined with their stable oligomeric state and secondary structure argues in favor of them representing distinct stable structural elements in DISC1, however, it is possible that others also exist not detectable by this approach. For example, the lack of other structured regions in the N-terminal half supports the idea of this section of the protein being unstructured, although the presence of the D region requires the redefining of the extent of this unstructured region as only stretching from the N-terminus to approximately AA 250.

The possibility must also be considered, however, that other structural elements exist which are not easily transcribed, translated or folded in *E. coli*, but which may be seen in mammals. Furthermore, it is likely given the role of DISC1 as a scaffold protein with many interaction partners (5) that DISC1 would not naturally exist *in vivo* outside of complexes with protein binding partners. Indeed, the regions of DISC1 which are unstable *in vitro*, instead most likely gain stability *in vivo* immediately after translation in the cell through interaction with binding partners. It is possible that such early interactions may direct individual DISC1 molecules towards specific cellular functions, through establishing its oligomeric state and setting a tertiary structure, limiting the number of potential other protein interaction sites which are exposed on the surface of the protein.

It is notable that, when expressed in isolation, the individual structured regions of DISC1 display differing oligomeric states. This is particularly striking in the case of the overlapping dimeric I region and tetrameric S regions, while constructs overlapping the two show mixed characteristics of both. We therefore put forward the hypothesis based on our data that at least two different structural conformations of full-length DISC1 could exist, based on which of the dimeric I or tetrameric S region was dominant as the intramolecular interaction domain (Fig. 5A, Table 1). The first of these would have at its core a dimeric I-region, from which monomeric C region and N-terminal sections extend, the simplest possible model of which would be a dimer (Fig. 5B), consistent with the observed dimer of full-length DISC1 by Narayanan *et al.* (10). The second would instead have a tetrameric core, based on that seen with the S region construct, and therefore represent a higher oligomeric state of DISC1 (Fig. 5C). This would likely represent the basis of the reported octameric species of DISC1 (10), with the additional oligomericity potentially arising through association of the D regions, through interactions between different

regions, or possibly with additional structural elements of DISC1 not detectable in this screen.

Of the four regions described here, the I region is the least stable, existing in solution at concentrations up to approximately 30 μM, but precipitating to insoluble aggregates above that. Longer ESPRIT-derived constructs that include the I region all show a strong tendency to form higher molecular weight aggregate complexes as well. Such insolubility, if also present *in vivo*, may contribute to the formation of the insoluble DISC1 aggregate species found in brains of a subset of patients with mental illness (22,23) or a transgenic rat overexpressing full-length DISC1 (24).

A consequence of our proposed domain structure is that the sequence encoding the I region would be bisected by the Scottish translocation (1), meaning that any protein translated from this locus would likely have an incorrectly folded I region and lack completely the highly stable S and C regions, the former of which would likely be critical for its higher-order oligomerization. A translocation-derived protein would thus be highly unstable, in addition to effects due to loss of protein-protein interaction sites.

Finally, we investigated the structural consequences of a frameshift mutation in the *DISC1* gene which has been linked to major mental illness in a family (3) and which causes synaptic vesicle deficits via depletion or dysregulation of the expression of several synapse-related genes in human forebrain neurons (25). In such neurons generated using induced pluripotent stem cells from this family, the total level of DISC1 protein was reduced, suggesting ubiquitin-mediated degradation of the mutant protein, despite the frameshift protein containing almost 95% of the DISC1 reading frame. The domain structure described here provides a partial explanation for this effect, as the frameshift directly disrupts the C region of DISC1. Furthermore, when the frameshift mutation was simulated in a recombinant C region protein, it led to the aberrant and higher oligomerization of the region (Fig. 4E), relative to a wild-type version, and to have a tendency to unfold and aggregate which, during early protein genesis could lead to ubiquitin-mediated protein degradation as described previously (25).

In this work, we have defined for the first time structural domains of human DISC1 based on experimental evidence, thus laying the foundations for understanding the functional architecture of the protein. Specifically, we revealed that the disordered N-terminal region of DISC1 is followed by a dimeric D region. Following a central stretch which is unstable at least when expressed in *E. coli*, presumably requiring either protein binding partners or mammalian-specific factors to facilitate its folding, lie the overlapping I and S regions which drive the oligomeric state of DISC1. Finally, at the extreme C-terminus harbors the monomeric, helical C domain which is seemingly involved in protein-protein interactions. While other structural domains, not amenable to expression in *E. coli*, may be added with time, this framework nevertheless has already provided insight into the mechanisms by which DISC1 is disrupted in major mental illness and will be a powerful resource for future studies into its structure and function.

Experimental procedures

Construct generation - A cDNA encoding for full-length human DISC1 (RefSeq accession number NP_061132.2) was codon optimized for expression in *E. coli* (GeneArt, Thermo Fisher Scientific, Regensburg, Germany) and then cloned between the *Asc*I and *Not*I sites of the pET-derived pESPRIT002 vector (26) for use in the ESPRIT screen. Ensuing constructs had their BAP tags removed by *Bsp*EI enzyme digestion prior to biophysical characterization. Additional subregions of DISC1 were subcloned from this vector and inserted into the pESPRIT vector at its *Aat*II and *Bsp*EI sites. The identity of all constructs was confirmed by sequencing.

ESPRIT - High throughput screening of DISC1 for soluble domains by incremental truncation was performed as described in detail previously (26). Briefly, a construct encoding DISC1 with N-terminal His_6 tags and C-terminal BAP tags was linearized at the 3′ end of the gene by restriction digest and truncated in a 3′ to 5′ direction with Exonuclease III and mung bean nuclease (both from New England Biolabs, Évry, France) and the resulting blunt ends polished with *Pfu* polymerase (Agilent Technologies, Les Ulis, France). The linear plasmid library of truncated constructs was religated back to a circular plasmid, then recovered by transformation of *E. coli* MACH1 (Thermo Fisher Scientific, Villebon-sur-Yvette, France), plating and subsequent plasmid

preparation from pooled colonies. This plasmid mix was then similarly truncated from the 3′ end. The linear plasmid library was electrophoresed on an agarose gel and plasmids containing constructs encoding 150-400 amino acid DISC1 fragments excised, ligated and recovered into *E. coli* MACH1. DNA sequencing of randomly selected plasmid inserts revealed inserts distributed along the entire length of DISC1. *E. coli* BL21 AI (Thermo Fisher Scientific) was transformed with the library and 27,642 clones picked robotically into 72 × 384 well plates, grown, then arrayed in duplicate onto nitrocellulose membranes (GE Healthcare, Vélizy-Villacoublay, France) over LB agar with antibiotics. Clones were grown, induced by transfer to fresh agar trays containing 50 µM biotin and 0.2 % (w/v) arabinose for 4 hours. Colonies on membranes were then lysed on NaOH-soaked filter paper, neutralized in buffer (27), blocked with Superblock (Thermo Fisher Scientific) and probed using anti-His (GE Healthcare) and rabbit anti-mouse Alexa Fluor 532 conjugate (Thermo Fisher Scientific) to detect the N-terminus and Streptavidin Alexa Fluor 488 (Thermo Fisher Scientific) to detect the C-terminus. The 3545 clones with highest anti-His signal were then ranked for *in vivo* biotinylation efficiency based upon the streptavidin signal and the highest signaling 94 clones were selected for expression and nickel-NTA purification analyses in larger scale cultures.

Recombinant soluble protein expression and purification - Plasmid vectors were transfected into BL21 AI cells and grown in Terrific Broth. Protein expression was induced by the addition of 0.2 % L-arabinose and 1mM IPTG for 16 hours at 25 °C. Bacterial pellets were stored at -80 °C and lysed by incubation in 25 mM Tris pH 7.4 / 150 mM NaCl / 5 mM imidazole / 1 mM DTT / 0.5% Triton-X 100 / 1 mM $MgCl_2$ containing lysozyme, DNaseI and protease inhibitors at room temperature. The insoluble pellet was spun down by centrifugation at 12,000 g for 45 minutes. The soluble fraction was then incubated with nickel-NTA (Qiagen, Hilden, Germany) for 45 minutes at room temperature and washed with 25 mM Tris pH 7.4 / 150 mM NaCl / 20 mM imidazole / 1mM DTT. Protein was eluted with the same buffer containing 500 mM imidazole and was then further purified by SEC. Where necessary, proteins were concentrated using Amicon Ultra centrifugation devices (Merck Millipore, Darmstadt, Germany). All proteins used for experiments were at least 95% pure.

Denaturing protein purification from the inclusion bodies and protein refolding - Plasmid vectors were transfected into *E. coli* BL21 AI cells and grown in Terrific Broth. Protein expression was induced by the addition of 0.2 % L-arabinose and 1 mM IPTG for 6 hours at 37^0C. Bacterial pellets were stored at -80 °C and lysed by incubation in 25 mM Tris pH 7.4 / 150 mM NaCl / 5 mM imidazole / 1 mM DTT / 0.5% Trition-X100/1 mM $MgCl_2$ containing lysozyme, DNaseI and protease inhibitors at room temperature. The insoluble pellet was spun down by centrifugation at 12,000 g for 45 minutes. The insoluble pellet obtained after the centrifugation of the cell lysate was dissolved in 25 mM Tris pH 7.4 / 150 mM NaCl / 5m M imidazole / 1 mM DTT / 8 M urea and then incubated with nickel-NTA (Qiagen) for 45 minutes at room temperature and washed with 25 mM Tris pH 7.4 / 150 mM NaCl / 20 mM imidazole / 1 mM DTT / 8 M Urea. Protein was eluted with the same buffer containing 500 mM imidazole. Refolding of the eluted protein is done by dialyzing in three steps for a period of 16-18 hours. The refolded protein was then further purified by SEC.

Gels and western blots - Protein samples were denatured in Laemmli buffer and run on SDS-polyacrylamide gels and then either directly stained with SYPRO Ruby Protein Gel Stain (Thermo Fisher Scientific) or InstantBlue Protein Stain (Expedeon, Swavesey, UK), or else transferred to nitrocellulose membranes. These were then stained accordingly with primary antibodies against the His_6 tag and DISC1 (14F2, raised against a peptide found in the C region and which was described previously (28)) as well as fluorescently-labeled streptavidin to detect the BAP tag. Protein signal was then detected using IRDye secondary antibodies (LI-COR Biosciences, Bad Homburg, Germany) on an Odyssey Clx infrared imaging system (LI-COR Biosciences).

Size exclusion chromatography - SEC was performed using and ÄKTA Pure system (GE Healthcare, Freiburg, Germany), cooled to 4-8 °C, with 25 mM Tris pH 7.4/150 mM NaCl/1mM TCEP, and using a HiLoad 16/600 Superdex 200pg column (GE Healthcare). In

some instances, to check the stability of the oligomeric state, specific eluted protein fractions were stored at 4 °C for 24 hours and then subjected to SEC again under the same conditions.

Circular dichroism - For circular dichroism experiments, protein was desalted into 25 mM sodium phosphate pH 7.0 / 150 mM NaF / 1 mM TCEP. Circular dichroism spectroscopic measurements were carried out on a JASCO J-815 spectrometer (JASCO, Groß-Umstadt, Germany). A 1 mm optical path length cuvette was used. The temperature was controlled at 20 °C. Spectra were recorded from □ = 260 nm to 185 nm at 1 nm resolution, 50 nm/min scan speed, and an integration time of 0.5 s. For signal improvement, ten accumulations were averaged. The obtained spectra were transformed to mean residue ellipticity after subtraction of the buffer spectra. Deconvolution of the data was performed on DichroWeb (18,19) using the CDSSTR method and reference dataset 4 (20,21).

Analytical ultracentrifugation - Sedimentation velocity centrifugation experiments at 50,000 rpm and 20 °C were carried out in a Beckman Optima XL-A (Beckman-Coulter, Brea, CA, USA), equipped with absorption optics and a four-hole rotor. Samples (volume 400 µL) were filled into standard aluminium double sector cells with quartz glass windows. Measurements were performed in absorbance mode at detection wavelengths 230 nm. Radial scans were recorded with 30 µm radial resolution at ~1.5 min intervals. The software package SEDFIT v.14.1 (www.analyticalultracentrifugation.com) was used for data evaluation. After editing time-invariant noise was calculated and subtracted. Continuous sedimentation coefficient distributions *c(s)* were determined with 0.05 S resolution and F-ratio = 0.95. Suitable *s*-value ranges between 0 and 20 S and f/f_0 between 1 and 4 were chosen. Buffer density and viscosity had been calculated with SEDNTERP v. 20111201 beta (29) (bitcwiki.sr.unh.edu/index.php). The partial specific volumes of the DISC1 protein fragments were calculated according to the method of Cohn and Edsall (30,31) as implemented in SEDNTERP.

Sedimentation equilibrium experiments were performed in standard aluminium double sector cells with quartz glass windows. Equilibria were established at multiple speeds. After equilibrium was reached concentration profiles were recorded with 10 µm radial resolution and averaging of seven single registrations per radial value. Data evaluation was performed using SEDPHAT v.10.55b (32).

Author contributions - ASKY, OHW, LN-S, DJH, NJB, and CK designed the research. ASKY, PJM, CD, and NJB performed experiments. ASKY, PJM, DCS, OHW, LN-S, DW, DJH, NJB, and CK analyzed the data. ASKY, NJB, and CK wrote the paper, with input from the other authors.

Acknowledgments - The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under BioStruct-X (grant agreement N°283570), as well as the "Biostruct" and "iBrain" graduate schools at the Heinrich Heine University, Düsseldorf. Additional funding came from the Seventh Framework Program (MC-IN "IN-SENS" #60761 to CK), the Brain Behavior Research Foundation (NARSAD Independent Investigator Award #20350 to CK), the Alexander von Humboldt Stiftung (1142747 to NJB), the Forschungskommission der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf (9772547 to NJB), the Fritz Thyssen Stiftung (10.14.2.140 to NJB) and the Impuls und Vernetzungs-Fonds der Helmholtzgemeinschaft (Portfolio programs "Technology and Medicine" and "Drug Research" to DW). The ESPRIT platform of the Grenoble Instruct Centre (ISBG: UMS 3518 CNRS-CEA-UJF-EMBL) receives current financial support from the French Infrastructure for Integrated Structural Biology (ANR-10-INSB-05-02) and the Grenoble Alliance for Integrated Structural & Cell Biology (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

Conflict of interest - The authors declare that they have no conflicts of interest with the contents of this article.

References

1. 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., and Porteous, D. J. (2000) Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* **9**, 1415-1425
2. Thomson, P. A., Duff, B., Blackwood, D. H. R., Romaniuk, L., Watson, A., Whalley, H. C., Li, X., Dauvermann, M. R., Moorhead, T. W. J., Bois, C., Ryan, N. M., Redpath, H., Hall, L., Morris, S. W., van Beek, E. J. R., Roberts, N., Porteous, D. J., St. Clair, D., Whitcher, B., Dunlop, J., Brandon, N. J., Hughes, Z. A., Hall, J., McIntosh, A., and Lawrie, S. M. (2016) Balanced translocation linked to psychiatric disorder, glutamate, and cortical structure/function. *NPJ Schizophr.* **2**, 16024
3. Sachs, N. A., Sawa, A., Holmes, S. E., Ross, C. A., DeLisi, L. E., and 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
4. Bradshaw, N. J., and Porteous, D. J. (2012) DISC1-binding proteins in neural development, signalling and schizophrenia. *Neuropharmacol.* **62**, 1230-1241
5. Yerabham, A. S. K., Weiergräber, O. H., Bradshaw, N. J., and Korth, C. (2013) Revisiting Disrupted in Schizophrenia 1 as a scaffold protein. *Biol. Chem.* **394**, 1425-1437
6. Kamiya, A., Kubo, K.-i., Tomoda, T., Takaki, M., Youn, R., Ozeki, Y., Sawamura, N., Park, U., Kudo, C., Okawa, M., Ross, C. A., Hatten, M. E., Nakajima, K., and Sawa, A. (2005) A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nat. Cell Biol.* **7**, 1167-1178
7. Taylor, M. S., Devon, R. S., Millar, J. K., and Porteous, D. J. (2003) Evolutionary constraints on the Disrupted in Schizophrenia locus. *Genomics* **81**, 67-77
8. Soares, D. C., Carlyle, B. C., Bradshaw, N. J., and Porteous, D. J. (2011) DISC1: Structure, function, and therapeutic potential for major mental illness. *ACS Chem. Neurosci.* **2** 609-632
9. Leliveld, S. R., Hendriks, P., Michel, M., Sajnani, G., Bader, V., Trossbach, S., Prikulis, I., Hartmann, R., Jonas, E., Willbold, D., Requena, J. R., and 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
10. Narayanan, S., Arthanari, H., Wolfe, M. S., and 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
11. Bord, L., Wheeler, J., Paek, M., Saleh, M., Lyons-Warren, A., Ross, C. A., Sawamura, N., and Sawa, A. (2006) Primate disrupted-in-schizophrenia-1 (DISC1): High divergence of a gene for major mental illnesses in recent evolutionary history. *Neurosci. Res.* **56**, 286-293
12. Sanchez-Pulido, L., and Ponting, C. P. (2011) Structure and evolutionary history of DISC1. *Hum. Mol. Genet.* **20**, R175-R181
13. Tarendeau, F., Boudet, J., Guilligay, D., Mas, P. J., Bougault, C. M., Boulo, S., Baudin, F., Ruigrok, R. W. H., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J.-P., and Hart, D. J. (2007) Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat. Struct. Mol. Biol.* **14**, 229-233
14. Yumerefendi, H., Tarendeau, F., Mas, P. J., and 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
15. Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W. H., Ortin, J., Hart, D. J., and Cusack, S. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* **15**, 500-506
16. Nadal, M., Mas, P. J., Blanco, A. G., Arnan, C., Solà, M., Hart, D. J., and Coll, M. (2010) Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16078-16083
17. Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., Marcellin, R., Melzer, R., Brown, E., Hart, D., Theißen, G., Silva, C. S., Parcy, F., Dumas, R., Nanao, M., and Zubieta,

C. (2014) Structural basis for the oligomerization of the MADS domain transcription factor SEPALLATA3 in *Arabidopsis*. *Plant Cell* **26**, 3603-3615
18. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucl. Acids Res.* **32**, W668-W673
19. Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **89**, 392-400
20. Manavalan, P., and Johnson Jr, W. C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76-85
21. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252-260
22. Leliveld, S. R., Bader, V., Hendriks, P., Prikulis, I., Sajnani, G., Requena, J. R., and 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
23. Korth, C. (2012) Aggregated proteins in schizophrenia and other chronic mental diseases: DISC1opathies. *Prion* **6**, 134-141
24. Trossbach, S. V., Bader, V., Hecher, L., Pum, M. E., Masoud, S. T., Prikulis, I., Schäble, S., de Souza Silva, M. A., Su, P., Boulat, B., Chwiesko, C., Poschmann, G., Stühler, K., Lohr, K. M., Stout, K. A., Oskamp, A., Godsave, S. F., Müller-Schiffmann, A., Bilzer, T., Steiner, H., Peters, P. J., Bauer, A., Sauvage, M., Ramsey, A. J., Miller, G. W., Liu, F., Seeman, P., Brandon, N. J., Huston, J. P., and Korth, C. (2016) Misassembly of full-length Disrupted-in-Schizophrenia 1 protein is linked to altered dopamine homeostasis and behavioral deficits. *Mol. Psychiatry* **21**, 1561-1572
25. Wen, Z., Nguyen, H. N., Guo, Z., Lalli, M. A., Wang, X., Su, Y., Kim, N.-S., Yoon, K.-J., Shin, J., Zhang, C., Makri, G., Nauen, D., Yu, H., Guzman, E., Chiang, C.-H., Yoritomo, N., Kaibuchi, K., Zou, J., Christian, K. M., Cheng, L., Ross, C. A., Margolis, R. L., Chen, G., Kosik, K. S., Song, H., and Ming, G.-l. (2014) Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature* **515**, 414-418
26. Sukackaite, R., Jensen, M. R., Mas, P. J., Blackledge, M., Buonomo, S. B., and Hart, D. J. (2014) Structural and biophysical characterization of murine Rif1 C terminus reveals high specificity for DNA cruciform structures. *J. Biol. Chem.* **289**, 13903-13911
27. Büssow, K., Cahill, D., Nietfeld, W., Bancroft, D., Scherzinger, E., Lehrach, H., and Walter, G. (1998) A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucl. Acids Res.* **26**, 5007-5008
28. Ottis, P., Bader, V., Trossbach, S. V., Kretzschmar, H., Michel, M., Leliveld, S. R., and Korth, C. (2011) Convergence of two independent mental disease genes on the protein level: Recruitment of dysbindin to cell-invasive Disrupted-in-Schizophrenia 1 aggresomes. *Biol. Psychiatry* **70**, 604-610
29. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S., Rowe, A., and Horton, J. C. eds.), Royal Society of Chemistry Cambridge. pp 90-125
30. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides*, Reinhold, New York
31. Durchschlag, H. (1986) *Thermodynamic Data for Biochemisty and Biotechnology*, Springer-Verlag, New York
32. Ghirlando, R. (2011) The analysis of macromolecular interactions by sedimentation equilibrium. *Methods* **54**, 145-156
33. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* (Walker, J. M. ed.), Humana Press Inc., Totowa, NJ. pp 571-607

34. 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., and 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
35. Murdoch, H., Mackie, S., Collins, D. M., Hill, E. V., Bolger, G. B., Klussmann, E., Porteous, D. J., Millar, J. K., and Houslay, M. D. (2007) Isoform-selective susceptibility of DISC1/Phosphodiesterase-4 complexes to dissociation by elevated intracellular cAMP levels. *J. Neurosci.* **27**, 9513-9524
36. Shinoda, T., Taya, S., Tsuboi, D., Hikita, T., Matsuzawa, R., Kuroda, S., Iwamatsu, A., and Kaibuchi, K. (2007) DISC1 regulates Neurotrophin-induced axon elongation via interaction with Grb2. *J. Neurosci.* **27**, 4-14
37. Burdick, K. E., Kamiya, A., Hodgkinson, C. A., Lencz, T., DeRosse, P., Ishizuka, K., Elashvili, S., Arai, H., Goldman, D., Sawa, A., and Malhotra, A. K. (2008) Elucidating the relationship between *DISC1*, *NDEL1*, and *NDE1* and the risk for schizophrenia: Evidence of epistasis and competitive binding. *Hum. Mol. Genet.* **17**, 2462-2473
38. Sawamura, N., Ando, T., Maruyama, Y., Fujimuro, M., Mochizuki, H., Honjo, K., Shimoda, M., Toda, H., Sawamura-Yamamoto, T., Makuch, L. A., Hayashi, A., Ishizuka, K., Cascella, N. G., Kamiya, A., Ishida, N., Tomoda, T., Hai, T., Furukubo-Tokunaga, K., and Sawa, A. (2008) Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. *Mol. Psychiatry* **13**, 1138-1148
39. Shimizu, S., Matsuzaki, S., Hattori, T., Kumamoto, N., Miyoshi, K., Katayama, T., and Tohyama, M. (2008) DISC1-kendrin interaction is involved in centrosomal microtubule network formation. *Biochem. Biophys. Res. Comm.* **377**, 1051-1056
40. Wang, Q., Charych, E. I., Pulito, V. L., Lee, J. B., Graziane, N. M., Crozier, R. A., Revilla-Sanchez, R., Kelly, M. P., Dunlop, A. J., Murdoch, H., Taylor, N., Xie, Y., Pausch, M., Hayashi-Takagi, A., Ishizuka, K., Seshadri, S., Bates, B., Kariya, K., Sawa, A., Weinberg, R. J., Moss, S. J., Houslay, M. D., Yan, Z., and Brandon, N. J. (2011) The psychiatric disease risk factors DISC1 and TNIK interact to regulate synapse composition and function. *Mol. Psychiatry* **16**, 1006-1023

Abbreviations

AA: Amino Acid, AUC: Analytical Ultracentrifugation, BAP: Biotin Acceptor Peptide, CD: Circular Dichroism, DISC1: Disrupted in Schizophrenia 1, ESPRIT: Expression of Soluble Proteins by Random Incremental Truncation, His_6: Hexa-Histidine, MW: Molecular Weight, SE: Sedimentation Equilibrium, SEC: Size Exclusion Chromatography, SV: Sedimentation Velocity.

Region	Amino acids in ESPRIT construct(s)	Calculated molecular weight, kDa	Predominant oligomeric state
D	249-383 257-399	14.9 (17.7) 16.2 (19.0)	Dimer
I	539-655	13.4 (16.2)	Dimer
S	635-738	11.9 (13.9)	Tetramer
C	684-836 691-836	17.3 (20.1) 16.6 (19.4)	Monomer

TABLE 1. **The four proposed structural regions of DISC1.** Amino acid positions of their prototypical ESPRIT constructs are shown, along with calculated molecular weight (determined using the ExPASy ProtParam tool (33); web.expasy.org/protparam) for the region alone, and for the His-tagged constructs used in this study (in brackets)

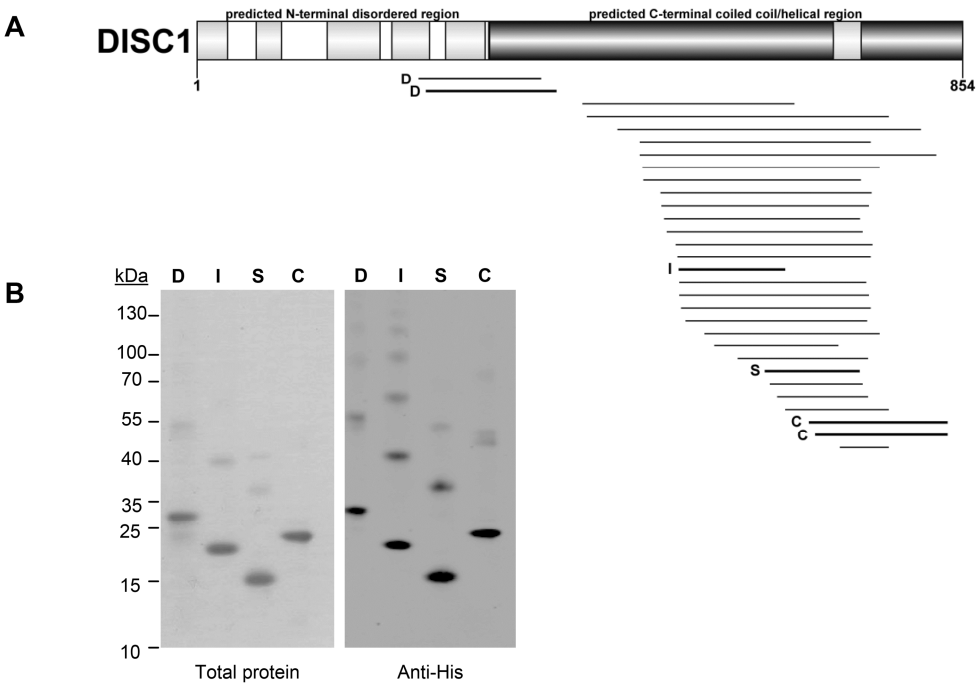

FIGURE 1. **Identification of structural regions of DISC1 through ESPRIT.** (A) The positions and lengths of soluble truncated DISC1 proteins following high throughput screening, shown relative to the predicted secondary structure of the full-length protein for comparison. Locations of a translocation breakpoint (at amino acid 597) and a frameshift mutation (at 807) both linked to major mental illness are also displayed. Of the soluble proteins, six were subsequently determined to yield consistent, stable proteins. These were interpreted to be representative of four structural regions of the DISC1 protein which were named D (represented by two clones), I, S (one clone each) and C (two clones). In the full-length DISC1 schematic, dark grey represents sections with a predicted coiled-coil or helix forming propensity, light grey represents regions predicted to be disordered. White regions were not predicted to be either helical or disordered. (B) SDS-PAGE Coomassie-stained gel (left panel) and western blot (right panel) of DISC1 protein fragments encoded by four of these clones, representing the soluble D, I, S and C regions. SDS-resistant oligomerization of some of the species is visible on the blot.

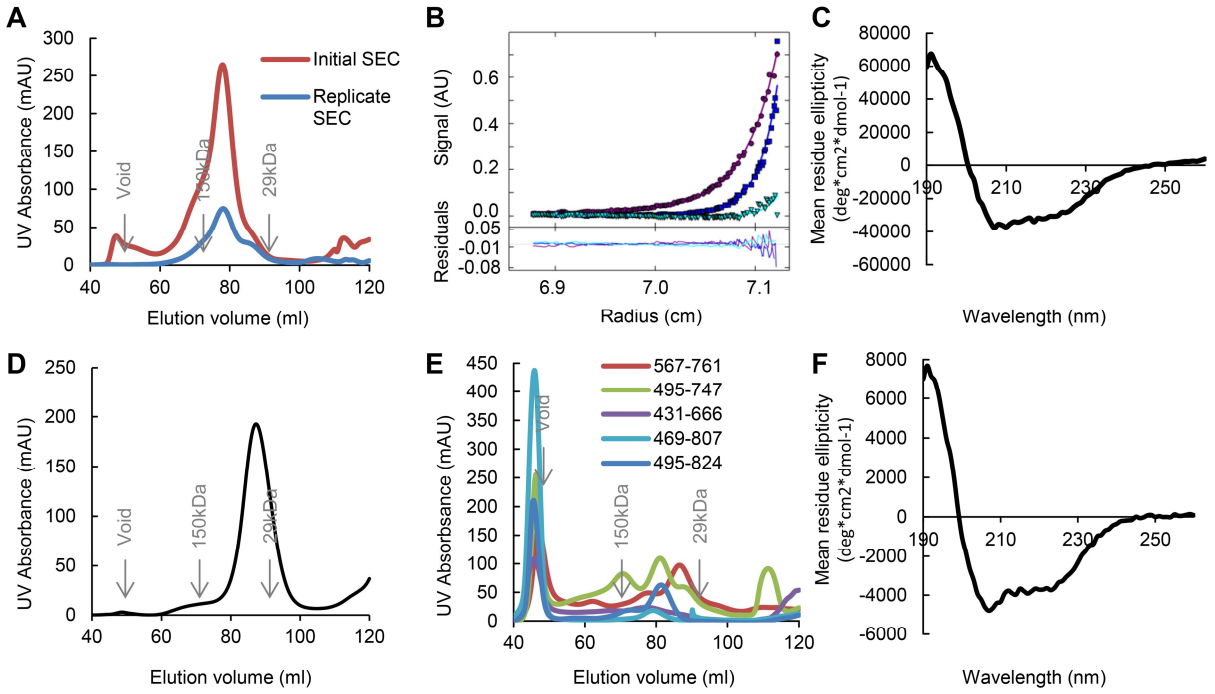

FIGURE 2. **Biophysical characterization of the D and I regions of DISC1**. (A) SEC of the DISC1 D region fragment covering amino acids 257-400. The affinity purified protein was subjected to SEC ("Initial SEC", 170 µM protein as input), with the major peak being eluted, concentrated and re-run on the SEC column 24 hours later ("Replicate SEC") to confirm protein stability. Two major species, one seemingly twice the size of the other were detected (B) AUC SE results of the smaller of these peaks. Concentration gradients at equilibrium were obtained at 29,300 (67,185 × g, cyan), 38,800 (117,816 × g, blue), and 46,300 rpm (167,765 × g, purple). The data had been fitted globally with a model for a single, ideal component. Residuals of the fit are shown in the graph attached below the main graph. Protein sample concentration: 32 µM. (C) Circular dichroism profile of the D region construct. Protein sample concentration: 17 µM (D) SEC of the DISC1 I region fragment covering amino acids 539-655 with a 100 µM protein input. (E) Representative SECs of other DISC1 fragments obtained from the ESPRIT screen which overlaps with the I domain. (F) CD spectrum of the I region construct. Protein sample concentration: 9 µM.

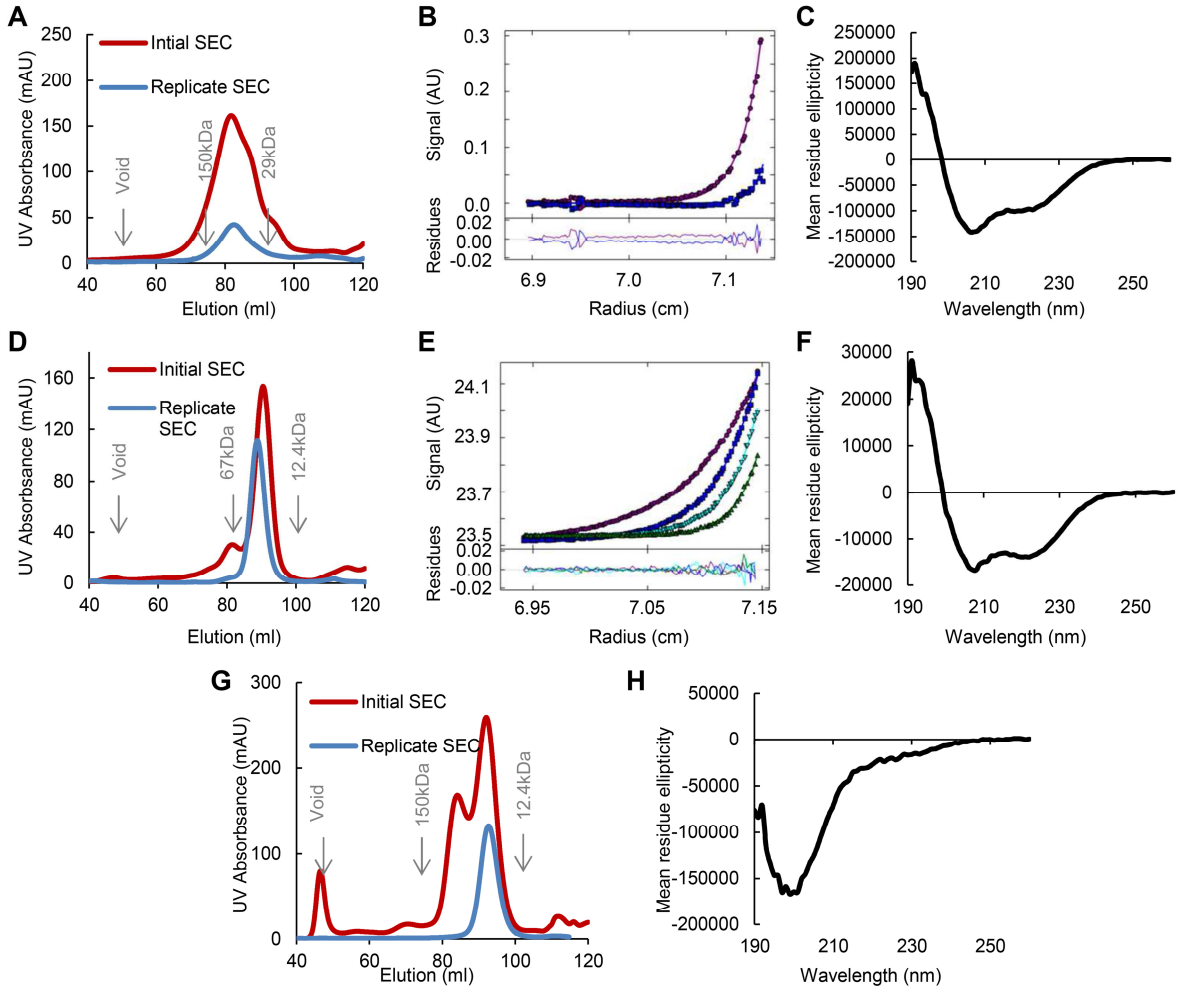

FIGURE 3. **Biophysical characterization of the S and C regions of DISC1**. **(A)** SEC of the DISC1 S region fragment covering amino acids 635-738. The affinity purified protein was subjected to SEC ("Initial SEC", 122 µM protein as input), with the principal eluted protein species (peak at approximately 75 ml) being collected, concentrated and re-run on the SEC column 24 hours later ("Replicate SEC") to confirm protein stability. (B) AUC SE results for the major eluted peak. Concentration gradient at equilibrium had been obtained at 29,300 rpm (67,185 × g, purple) and 38,800 rpm (117,816 × g, blue). The data have been fitted with a model for a single component. Residuals of the fit are shown in the graph attached below the main graph. Protein sample concentration: 44 µM. (C) CD profile of the S region construct. Protein sample concentration: 19 µM (D) SEC of the DISC1 C region fragment covering amino acids 691-836, with 57 µM protein as input. (E) AUC SE results for this sample. Concentration gradients at equilibrium had been obtained at 30,700 (70,759 × g, green), 39,700 (123,345 × g, cyan), 47,000 (172,876 × g, blue), and 53,200 rpm (221,495 × g, purple). The data had been fitted globally with a model for a single component. Residuals of the fit are shown in the graph attached below the main graph. Protein sample concentration: 19,4 µM. (F) CD spectrum of the C region construct. Protein sample concentration: 8.2 µM (G) SEC for a DISC1 fragment within the C region, containing amino acids 718-771. (H) CD spectrum of the 718-771 fragment. Protein sample concentration: 20 µM.

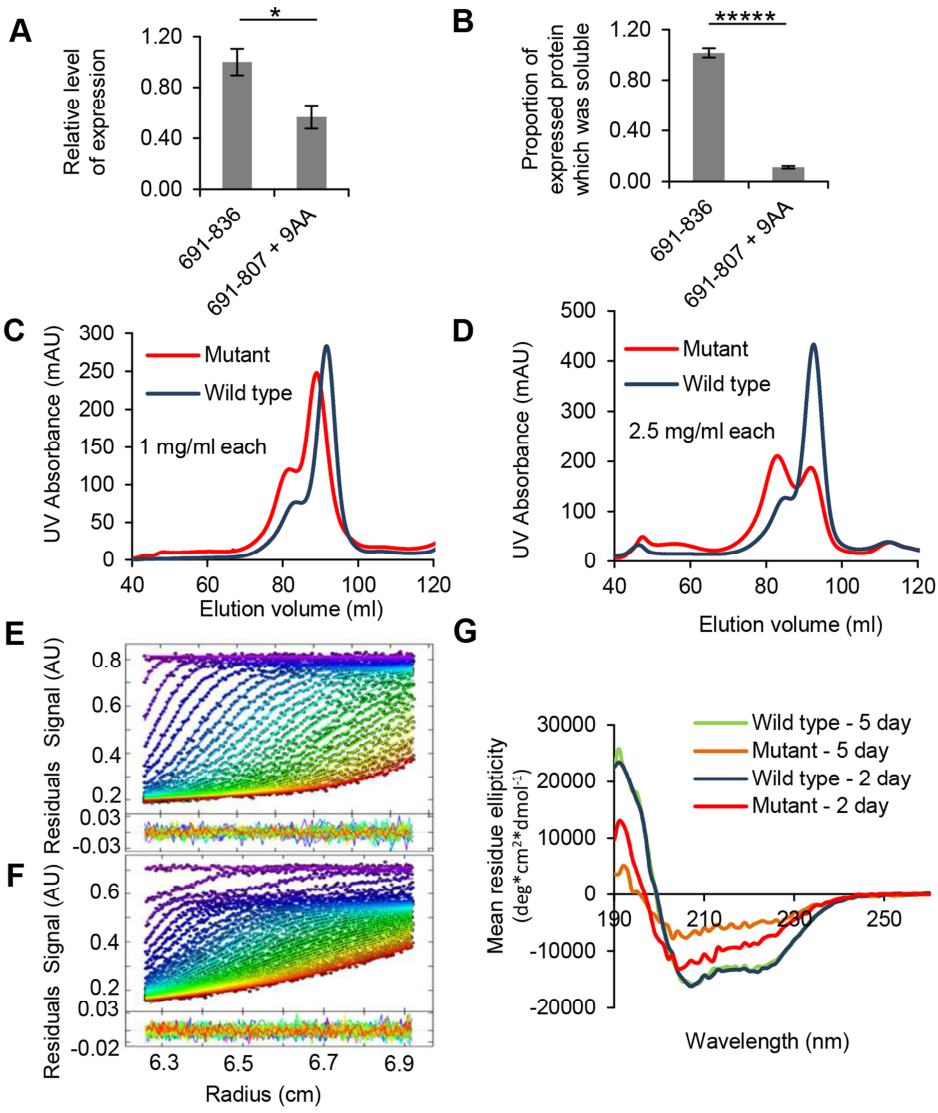

FIGURE 4. **A pathological frameshift mutation disrupts the stability, oligomeric state and secondary structure of the C region** (A) Quantified, relative level of DISC1 expression in bacterial cell lysates expressing either the wild-type DISC1 C region (AA 691-836) or a mutant version representing a frameshift mutation associated with mental illness (leading to truncation of the protein after AA 807, followed by 9 AA of non-specific read-through), n = 6, *: $p < 0.05$. (B) Following centrifugation of these lysates, the wild-type protein remains almost entirely in the supernatant, indicating solubility, while much of the mutant protein is instead lost in the pellet, indicating insolubility. n = 6, *****: $p < 10^{-6}$. (C) Comparison of the two proteins at two different concentrations (indicated in the figure), following refolding from a urea-solubilized pellet, by SEC, showing the mutant form to preferentially form higher molecular weight complexes. (D) This effect is enhanced with increasing protein concentration. (E) AUC SV results for the wild-type C regions obtained at 50,000 rpm (181,675 × g), 20 °C. Sedimentation profiles (dots) recorded over time from purple (initial scan) to red (final, scans taken at 3 minute intervals) are shown together with the *c*(*s*) fit results (lines). Below the graphs, the corresponding residuals are shown. Both AUC experiments performed using 33 µM protein. (F) Equivalent results for the mutant C regions, showing an increased propensity to form insoluble aggregates visible in the upper part of the sedimentation boundary as quickly sedimenting material. (G) Comparison of the CD spectra of the wild-type and mutant C regions, at a concentration of 5 µM each, showing the mutant to have reduced α-helical content, which decreases further over time stored at room temperature, while these is no obvious similar loss of structure for the wild-type.

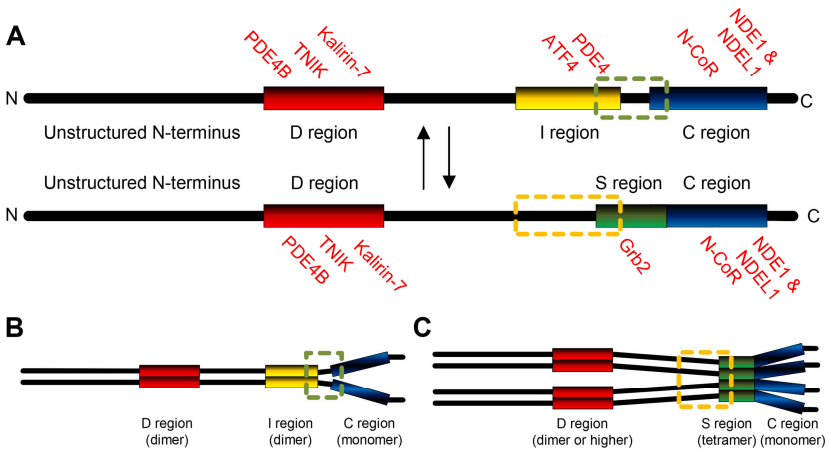

FIGURE 5. **Summaries of the domain structure and oligomerization of DISC1.** (A) Schematic of the domain structure for DISC1 proposed in this paper, with the I and S regions representing alternative configurations. Approximate locations at which known protein interaction partners bind to DISC1 are indicated in red, where these have been mapped at high precision to within one of the four regions (34-40), although this does not necessarily imply that they would interact with only one structural configuration of DISC1. (B) Schematic of a simple domain configuration for a DISC1 dimer, with the I and D regions both forming dimers. (C) Equivalent schematic for a DISC1 tetramer based on the S region, in which the D regions could form dimers or a tetramer. Note: that these figures are one-dimensional and do not take into account the folding of the structural regions or the potential for interaction between these regions.

A Structural Organization for Disrupted in Schizophrenia 1, Identified by High-Throughput Screening, Reveals Distinctly Folded Regions Which Are Bisected by Mental Illness-Related Mutations

Antony S. K. Yerabham, Philippe J. Mas, Christina Decker, Dinesh C. Soares, Oliver H. Weiergräber, Luitgard Nagel-Steger, Dieter Willbold, Darren J. Hart, Nicholas J. Bradshaw and Carsten Korth

J. Biol. Chem. published online March 1, 2017

Publication III

Biophysical insights from a single chain camelid antibody - Disrupted In Schizophrenia 1 protein (691-836) complex.

Antony S. K. Yerabham, Andreas Müller-Schiffmann, Tamar Ziehm, Andreas Stadler, Sabrina Köber, Xela Indhurkya, Rita Marreiros, Svenja V. Trossbach, Nicholas J. Bradshaw, Ingrid Prikulis, Oliver H. Weiergräber, and Carsten Korth

PLOS ONE - submitted

Biophysical insights from a single chain camelid antibody Disrupted-in-Schizophrenia 1 protein (691-836) complex

Antony S. K. Yerabham[1], Andreas Müller-Schiffmann[1], Tamar Ziehm[2], Andreas Stadler[3], Sabrina Köber[1], Xela Indhurkya[1], Rita Marreiros[1], Svenja V. Trossbach[1], Nicholas J. Bradshaw[1], Ingrid Prikulis[1], Oliver H. Weiergräber[2], Carsten Korth[1*]

[1] Department of Neuropathology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2] Institute of Complex Systems (ICS-6: Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany

[3] Jülich Centre for Neutron Science JCNS and Institute for Complex Systems ICS, Forschungszentrum Jülich, Jülich, Germany

KEYWORDS: Phage display, Small-angle X-ray scattering, Mental illness, DISC1, V_HH antibody fragment

*corresponding author:

E-mail: ckorth@hhu.de (CK)

Abstract

Accumulating evidence suggests an important role for the Disrupted-in-Schizophrenia 1 (DISC1) protein in neurodevelopment and chronic mental illnesses. In particular, the C-terminal 300 amino acids of DISC1 have been found to mediate important protein-protein interactions and to harbor functionally important phosphorylation sites and disease-associated polymorphisms. However, long disordered regions and oligomer-forming subdomains have so far impeded structural analysis. V_HH domains derived from camelid heavy chain only antibodies are minimal antigen binding modules with appreciable solubility and stability, which makes them best suited for the stabilization of proteins in structural investigations. Here, we generated and characterized a V_HH domain derived from an immunized *Llama glama* displaying high affinity for the human DISC1 C region (aa 691-836) by surface plasmon resonance, size exclusion chromatography and immunological techniques. The V_HH-DISC1 (C region) complex was also used for structural investigation using small-angle X-ray scattering analysis. In combination with molecular modeling, these data support predictions as to the three-dimensional fold of this DISC1 segment as well as the binding mode of our V_HH antibody.

Introduction

Disrupted in schizophrenia 1 (*DISC1*) was discovered as the gene disrupted by a balanced chromosomal translocation (1; 11) (q42.1; q14.3), which is strongly linked with schizophrenia and several other mental disorders in a large Scottish family [1]. In an American family with mental illness, a frame shift mutation following amino acid 807 in the DISC1 protein was reported [2]. DISC1 is known to interact with more than 150 other proteins [3], but has no known enzymatic activity, and is therefore considered to act as a scaffolding protein regulating the functions of its binding partners [4]. DISC1 is involved in cerebral cortex development and affects critical processes such as neuronal migration and proliferation [5,6]. Studies on animal models have associated pathological DISC1 protein variants with behavioral abnormalities [7], which were linked to disturbances in various neurotransmitter systems including dopamine [8,9].

The C-terminal region comprising amino acids 598-854 of human DISC1 protein (UniProtKB Q9NRI5) contains interaction sites for several binding partners such as NDE1, NDEL1, and LIS1, which are important in brain development and neuronal migration [10,11]. It also harbors functionally important sites such as the mental illness associated S704C polymorphism [12-14] and a phosphorylation site at S713, which was found to coordinate the switch of neuronal progenitor cell proliferation to migration during corticogenesis [15]. Moreover, a DISC1 protein fragment, consisting of amino acids 598-785, was shown to control unique characteristics such as cell invasiveness [16,17] and to contribute prominently to the structural order of the DISC1 protein, while amino acids 668-747 are essential for an orderly assembly of oligomers [18]. Recent findings on the structural organization of the DISC1 protein showed that the C-terminal region of the DISC1 protein covers three out of the four structural regions described i.e., the I region (aa 539-655), the S region (aa 635-738) and the C region (aa 691-836)[19]. The I region was shown to be the most aggregation prone zone

among all, whereas the S region is very stable and highly oligomeric and the C region previously thought to be dimeric was proved to exist as a monomer but to lose its high solubility due to the disease related frameshift mutation at amino acid 807 with nine novel amino acids. Therefore, structural information on the DISC1 protein C-terminal region is desirable for understanding the role of this region in the protein's normal function, solubility, and interaction properties, as well as the effects of truncations found in disease variants.

Single-domain camelid antibodies, also called V_HH domains or nanobodies, are engineered polypeptides used in a large range of applications [20]. They consist entirely of the variable domain (V_H) of heavy chain-only immunoglobulins, which occur in the sera of animals belonging to the Camelidae family. This variable domain alone is completely functional in recognizing a specific antigen, without the presence of a paired light chain variable (V_L) domain [21-23].

The numerous benefits of V_HH fragments compared to conventional antibodies include higher yields of expression when recombinantly produced in bacterial cells, higher levels of stability and solubility, and greater ease of delivery for therapeutic applications, while retaining high affinity and specificity towards the target antigen [24]. V_HH antibodies have proven applicable in diverse fields, from therapeutic interventions and diagnostics to basic research, including the investigation of protein structure and function [25-28].

Here we report generation of the first V_HH antibody directed against human DISC1, termed B5. Using biochemical and biophysical methodology, we quantitatively characterize the interaction of the antibody with DISC1 and determine its epitope. Moreover, the structures of the DISC1 C region [19] and its complex with V_HH B5 are investigated by small-angle X-ray scattering (SAXS) analysis and molecular modelling. These results provide insight into the binding mode of our V_HH antibody and suggest an explanation for its specificity for human vs. mouse DISC1.

Experimental procedures

Llama immunization

A female llama was immunized on days 0, 17, 32, 65 with 240 μg of purified recombinant human DISC1 fragment comprising residues 598-785 (DISC1 598-785) [18] in 0.5 mL PBS mixed with an equal volume of complete (day 0) or incomplete (all other days) Freund's adjuvant. At day 70, 550 mL of blood were taken and used for generating a phage library. The immunization procedure and blood withdrawal of the llama was contracted by the company Preclinics GmbH (Potsdam, Germany).

Construction of the phage library

500 mL of blood were collected 7 days after the final injection and diluted 1:1 with PBS. In total 6.9×10^8 lymphocytes were isolated via Ficoll Paque gradient centrifugation, pelleted and snap frozen in liquid nitrogen before storage at -80° C.

RNA was extracted using the RNeasy Mini Kit (Qiagen, Germany) and first strand cDNA was generated with Oligo(dT)18 primer and the Revert Aid M-MuLV Reverse Transkriptase Kit (Thermo Scientific). cDNA encoding V_HH was specifically amplified by PCR using sense and anti-sense hinge-specific primers "MJ 1.2.3 Back" and "CH2 + CH2b3" described in [29]. PCR bands between 620 and 670 bp were gel eluted and reamplified by a nested PCR with the primers MJ7 [29] and J_H [30]. These primers included respective *Sfi*I and *Not*I cleavage sites for subsequent cloning into the phagemid vector pIT2 (obtained from the human single fold scFv Tomlinson I+J libraries, Cambridge, UK) [31]. Ligated DNA was electro-transfected into electro-competent *E.coli* TG1 cells. The colonies of plated cells were counted, collected and frozen in growth medium containing 15% glycerol at -80° C. The size of the library was determined to be 1.2×10^7CFU (colony-forming units).

Phage display for selection

The phage display process was performed based on the Tomlinson I + J protocol [32]. The phage library generated was used for three rounds of panning to screen for the strongest DISC1 binding clone. The panning procedure was performed by coating the Nunc Maxisorp ELISA plates (Thermo Scientific, Denmark) with 0.1 mg/mL of the DISC1 protein fragment aa 598-785 in PBS buffer by overnight incubation at 4° C. The plates were blocked with 1 % skimmed milk powder and 1 % Casein (Hammerstein grade) in PBS buffer, for alternate rounds of panning, after coating with the DISC1 protein by similar incubation conditions as for coating. The phage library was panned onto these coated and blocked ELISA plates in the respective blocking solution, which were washed multiple times (8-10 times) with PBS before eluting the final binder phages. The elution was performed using 10 mg/mL trypsin solution dissolved in PBS. The trypsin eluted phage species in each round of panning were propagated in TG1 *E.coli*. The efficacy of each panning round was titrated by colony count at different dilutions of the propagated TG1 *E.coli* species infected by the trypsin eluted phages. Typically the titre of the TG1 *E.coli* cells infected with trypsin eluted phages increased in each consecutive round corresponding to a successful selection process. Thereafter the V_HH sequence of the clone selected through 3 rounds of panning from phage library was cloned into the plasmid vector pET 22b (by *Nco1* and *Not1*) for ease of overexpression.

Cloning and production of anti-DISC1 human Fc-V_HH antibody

The antiDISC1-V_HH antibody sequence was PCR amplified from the pET22b vector using forward primer (5'- agacggctgtgtcttcaggt -3') and reverse primer (5'- acctgaagacacagccgtct -3'). *Age1* and *Not1* sites were inserted in the PCR product using the forward 5'-gaaaaccggtatggcccaggtaaag-3' and reverse 5'-cttggcggccgcctactatgcggccgctg-3'primers and

then ligated into the expression vector pLHCX-F_C-fs downstream of the IgG signal sequence and followed by the human Fc coding sequence.

The expression vector pLHCX-F_C-fs is generated from a modified pLHCX vector (Clontech), where the multiple cloning site was modified using HindIII and ClaI, and a synthetic sequence

(ACCGGTCTCGAGGCGGCCGCGGCCAAAAAGGCCGGATCCGTTAACACCAAAAA ATGGCACGTGGCCGGCACGCGTGGGCCCGTCGAC) was inserted.

To obtain the coding sequence of the human Fc, mammalian cells were transiently transfected with the pLHCX-PrP-Fc vector containing the genomic human Fc sequence. RNA and subsequent cDNA was used to clone the coding sequence of the human Fc into the modified pLHCX vector in-frame at the restriction site *SalI*. The IgG signal sequence in expression vector pLHCX-F_C-fs, was cloned at HindIII and BamHI restriction sites.

The pLHCX_F_C-fs vector encoding the antibody was used to generate retroviruses according to the manufacturer's protocol (Clontech), which were used to infect HEK293 cells, causing stable expression of the V_HH antibody. HEK293 cells were then cultured in DMEM with 5% fetal calf serum, penicillin, streptomycin and hygromycin. Antibody secreted into the conditioned media (cell supernatant) was used further.

Purification of DISC1 protein fragments and V_HH B5 antibody

A DISC1 fragment comprising residues 598-785 ($DISC1^{598-785}$) was expressed and purified as described in [16]. The plasmid vectors pESPRIT002 DISC1 691-836, pESPRIT002 DISC1 539-655, pESPRIT002 DISC1 480-725, DISC1 pESPRIT002 598-715[19] were transfected into BL21 AI cells and grown in Terrific Broth (12 g tryptone, 24 g yeast extract, 4 ml glycerol autoclaved in 900 mL H_2O followed by addition of 100 mL filter sterilized 0.17 M monopotassium phosphate and 0.72 M dipotassium phosphate solution). Protein expression

was induced at OD_{600} = 0.8 by the addition of 0.2 % L-arabinose and 1 mM IPTG for 16 hours at 25^{0}C. The pET22b V_HH-B5 plasmid vector was transformed into the BL21 Rosetta strain (Novagen, USA) and protein expression was induced at OD_{600} = 0.8 with 1 mM IPTG at 18°C for 16 hours. Bacterial pellets were stored at -80 ^{0}C and lysed by incubation in 25 mM Tris pH 7.4/150 mM NaCl/5 mM imidazole/1 mM DTT/0.5% Trition-C100/20 mM $MgCl_2$ containing 0.25 mg/mL Lysozyme, DNaseI 40 U/mL and 2 mM PMSF at room temperature. The insoluble fraction was spun down by centrifugation at 6000x g for 45 minutes. The soluble fraction was then incubated with Ni^{2+}-NTA-Agarose (QIAgen) for 45 minutes at room temperature and washed with 25 mM Tris pH 7.4/150 mM NaCl/5 mM imidazole/1 mM DTT. Protein was eluted with the same buffer containing 500 mM imidazole and was then further purified by size exclusion chromatography (SEC) on Hiload 16/60 Superdex 200 pg column (GE Healthcare Bio-Sciences AB, Sweden) with a flow rate of 1 mL per minute at 4^0 C.

Monoclonal antibodies and cell lines

The α-human DISC1 human monoclonal antibody 14F2 has been described [17]. The mouse-Disc1 C-term monoclonal antibody was produced in rabbit as described previously [8,33].The 6x-His Epitope Tag antibody was commercially purchased from Thermo Scientific.

The NLF human neuroblastoma cell line was used for transfection with human and mouse DISC1.

Surface plasmon resonance (SPR)

SPR measurements were performed using a Biacore T200 instrument (GE Healthcare, Sweden) at 25 °C with PBS + 0.05 % Tween-20, pH 7.4 as running buffer. For preparation of

the flow cells, a CM5 sensor chip (GE Healthcare, Sweden) was activated with EDC / NHS (0.2 M / 0.05 M) and DISC1 fragment comprising the residues 691-836 ($DISC1^{691-836}$) (30 µg/mL) diluted in 10 mM sodium acetate, pH 4.0 was immobilized to a final level of 250 RU and the flow cell was deactivated with 1 M ethanolamine-HCl. A reference flow cell was activated and deactivated only. Afterwards, V_HH B5 at concentrations ranging from 19 nM to 1.5 µM was injected in a single cycle without regeneration. The sensorgrams were double referenced using the reference flow cell and a buffer cycle and evaluation was performed by plotting the respective response levels against the applied V_HH B5 concentrations. The curves were fitted using Langmuir´s 1:1 binding model (Hill function with n = 1, OriginPro 8.5G, OriginLab, Northampton, USA).

Small angle X-ray scattering (SAXS) data acquisition and processing

SAXS data were recorded using protein samples of >95 % purity, as judged by SDS-PAGE with Coomassie staining. In addition to the $DISC1^{691-836}$ protein and the V_HH B5 antibody, we investigated their complex after co-purification via SEC. SAXS was measured on beam line P12 at the PETRA III storage ring (DESY, Hamburg, Germany [34]) and on BM29 at the ESRF (Grenoble, France [35]). The X-ray wavelengths used on P12 and BM29 were 1.24 Å and 1 Å, respectively. An automatic robot was used for sample storage and loading of the solutions in the quartz capillary for X-ray exposure. Storage and measurement temperature was 20 ° C throughout all experiments. On P12 the exposure time was 100 ms and 20 frames were recorded, while on BM29 the exposure time was 2 s and 10 frames were taken. Buffer was measured before and after each protein sample. The individual frames were checked for the absence of radiation damage and the corresponding frames were merged. The scattering contribution of the buffer was subtracted from the measured intensities of the protein solutions. The buffer subtracted SAXS data were scaled by the protein concentrations

determined via absorbance at 280 nm. Final datasets were derived from measurements at protein concentrations of 2.8 mg/mL and 2.3 mg/mL for the DISC1$^{691-836}$ protein fragment and the complex with V_HH B5, whereas in case of the V_HH domain B5 recordings for 2.8 and 5.6 mg/mL were merged.

Data were analyzed using the *ATSAS* software package [36]. The radius of gyration (R_g) was derived from the Guinier approximation, as implemented in AUTORG [37], while CRYSOL [38] was used for calculation of theoretical values from atomic coordinates. The distance distribution function P(r) was determined using the program *DATGNOM* [37]. For each dataset, 20 *ab initio* reconstructions were generated using *DAMMIF* [39] or *GASBOR* [40], followed by averaging and filtering in *DAMAVER* [41]. The resulting consensus models were used in further analyses. For the purposes of fitting and visualization, they were converted into volumetric maps with *pdb2vol*, which is part of the *Situs* package [42].

Preparation of Structural Models

A starting model of the DISC1$^{691-836}$ protein including an N-terminal His tag was generated by *ab initio* prediction using *QUARK* [43]. This model provided a reasonable fit to the SAXS reconstruction, with a normalized spatial discrepancy (NSD) value of 1.83 as determined by *SUPCOMB [44]*. While it yielded a very good 3D-1D profile, with 92 % of all residues scoring higher than the confidence threshold of 0.2 as given by *Verify3D* [45], validation with *QMEAN* [46] indicated problems with torsion angles, corresponding to an overall *QMEAN* score of 0.54 and a Z-score of -2.29. Since visual examination of the model revealed poor geometry in several loop regions, it was subjected to iterative rebuilding using *MODELLER* [47], making extensive use of loop refinement algorithms. This procedure resulted in further improvement of the 3D-1D profile (94 % of residues exceeding 0.2) as well as an increase of the *QMEAN* score to 0.67 (Z-score -0.98). The final model comprises all 170 residues of the protein (including residues 691-836 of the human DISC1 sequence); according to validation

in *Coot* [48], all residues are located in the allowed regions of the Ramachandran diagram and do not display rotamer outliers. Compared to the initial *QUARK* model, the refined version features a better fit to the SAXS envelope, with an NSD of 1.36.

A homology model for the V_HH B5 antibody was built in *MODELLER*, using as template the crystal structure of a Llama nanobody (PDB ID 3EZJ, chain B [49]). This model contains residues 1-120, again without Ramachandran or rotamer outliers, and performs well in consistency tests using *Verify3d* (100 % scoring > 0.2) and *QMEAN* (score 0.77, Z-score 0.30).

The models of the $DISC1^{691\text{-}836}$ proteinand the V_HH B5 antibody were subjected to *in silico* docking simulations using *CLUSPRO* [50] in antibody mode [51]. The resulting models were examined for consistency with the SAXS reconstruction of the complex as well as experimental data narrowing down the binding site, and the best candidate selected for further analysis.

Fitting to SAXS envelopes

The *ab initio* SAXS reconstructions obtained from *DAMAVER* were initially fit to the corresponding structural models using *SUPCOMB*, which takes care of the handedness ambiguity of the SAXS envelopes. After conversion of the latter to *Situs*-format volumetric maps, the relative positioning of the model structures was optimized by correlation-based refinement, as implemented in *Sculptor* [52].

Data visualization

SAXS data were plotted using *Grace* (http://plasma-gate.weizmann.ac.il/Grace). Ribbon representations were generated with *POVScript+* [53] and *Raster3D* [54], applying secondary structure assignments provided by *DSSP* [55].

Results

Generation of an anti-DISC1 camelid single chain V_HH antibody

An anti-DISC1 camelid V_HH phage library was generated using mRNA obtained from a llama which had been immunized with the insoluble fraction of recombinant human $DISC1^{598-785}$ protein. This fragment forms multimers and may be largely responsible in forming the pathogenic misassembled DISC1 protein [18]. The insoluble $DISC1^{598-785}$ protein had been purified and refolded from inclusion bodies after overexpression in *E.coli*. The phage library generated was subjected to three rounds of panning by phage display in order to select the strongest binding clone against the $DISC1^{598-785}$ protein. (Fig 1A) There was a gradual increase in the phage titer in subsequent rounds of panning (3×10^3CFU/ml, 5×10^3CFU/ml, 1×10^7CFU/ml respectively, CFU=Colony Forming Units) indicating a successful selection process.

The strongest binder obtained, clone B5, was isolated and subcloned into the bacterial expression vector pET 22b. The camelid V_HH antibody was then recombinantly overexpressed in the periplasm of *E.coli* providing an oxidizing environment for proper folding and was then purified by Ni^{2+}-NTA affinity chromatography, followed by SEC. In SEC the V_HH B5 antibody protein eluted prominently as a monomer and very minimally in an oligomeric state (Fig 1 B, C).

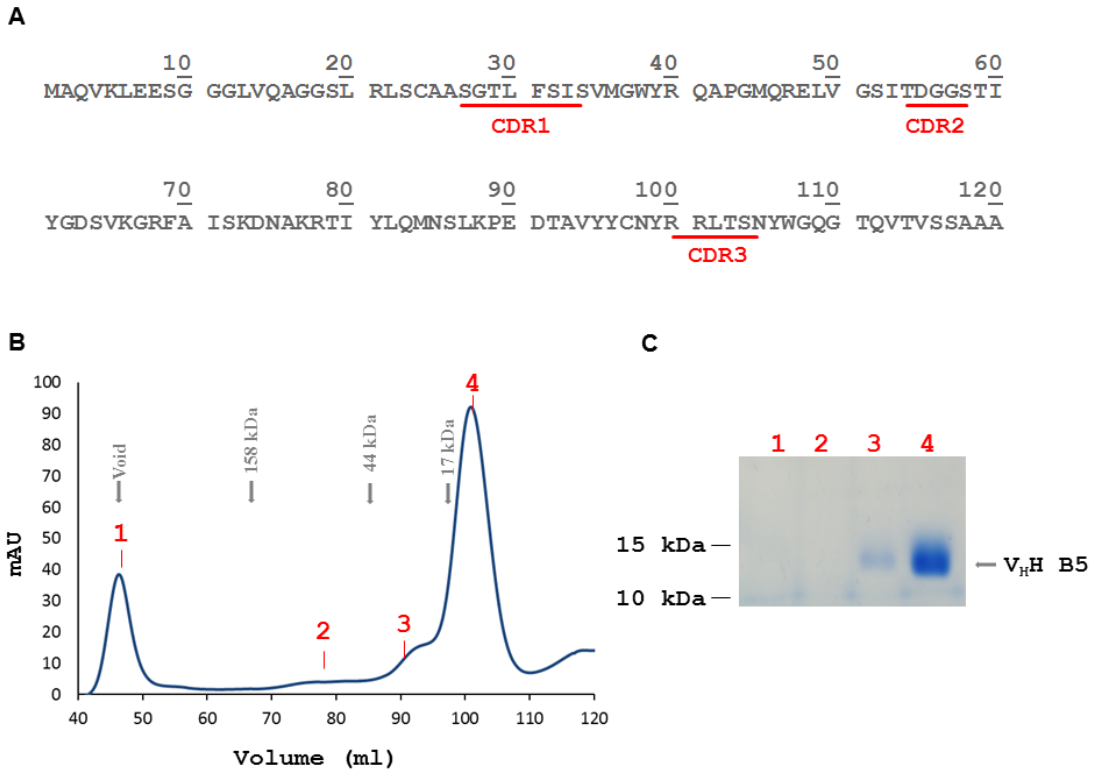

Figure 1. Sequence and purification of the generated anti-DISC1 V_HH B5 antibody.

(A) Sequence of the anti-DISC1 V_HH B5 protein showing the three complimentarity determining regions (CDRs). (B) SEC profile of the purified anti-DISC1 V_HH B5 antibody eluting with an apparent molecular mass of 13 kDa. (C) Coomassie stained SDS gel loaded with different SEC elution fractions (samples numbered in accordance with the chromatogram). Fraction 1 corresponds to the void volume, fractions 3 and, 4 contain the V_HH B5 antibody.

Interaction of V_HH B5 with DISC1 $^{691-836}$ protein

In order to confirm and further narrow down the binding of V_HH B5 to DISC1, we tested for its binding to the C region of DISC1 protein by SEC. DISC1$^{691-836}$ was referred to as 'C region' in the recent finding of the structural organisation of DISC1 protein [19]. This C region is a stable, distinct structural region within the DISC1 protein and is known to exist as a monomer, unlike the immunogen (DISC1$^{598-785}$) used, and also holds several physiologically relevant sites such as the disease variant S704C and the crucial phosphorylation site at S713 [15].

The SEC of an equimolar mixture of the DISC1$^{691-836}$ protein and the V_HH B5 antibody resulted in a shift of the resultant elution peak towards higher molecular weight species, compared to the individual proteins alone, and the protein mixture co-eluted as a complex in SEC (Fig 2 A), indicating binding of the V_HH antibody to the DISC1 protein fragment. The binding observed by SEC was further confirmed through SPR, with the binding affinity of the B5 antibody to the DISC1$^{691-836}$ protein was being characterized by a dissociation constant (K_D) of 139 nM. (Fig 2 B)

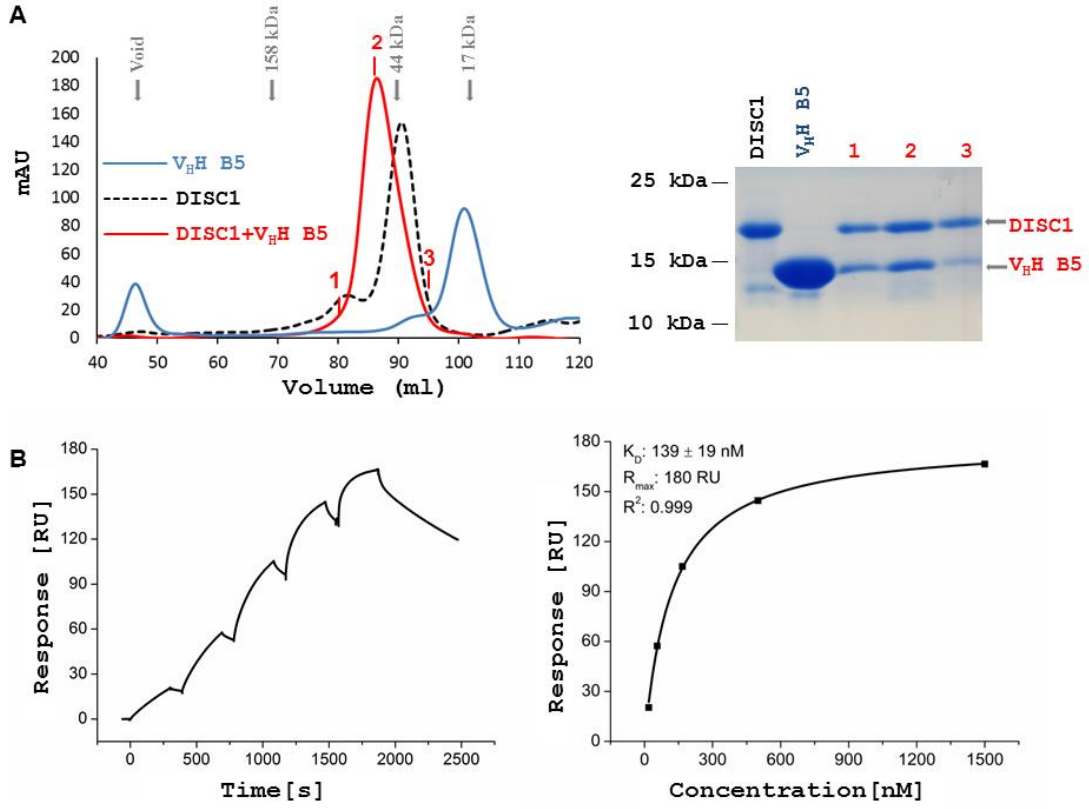

Figure 2. Interaction of the V_HH B5 antibody with DISC1$^{691-836}$ protein.

(A) The SEC profile demonstrates a shift in the elution peak of the DISC1 fragment and V_HH B5 antibody mixture, compared to the individual proteins. Fractions containing the DISC1-V_HH B5 co-elution peak (labelled 1, 2 and 3) were investigated by SDS-PAGE. (B) Representative SPR sensorgram showing the binding of a concentration series of V_HH B5 antibody to immobilized DISC1$^{691-836}$ protein and the corresponding fitting curve. The K_D value is presented as mean ± SD of three independent experiments.

Mapping of the V_HH antibody epitope on the DISC1 protein

In order to increase the avidity, as well as to expand the potential applications for functional analysis, the V_HH B5 was cloned into the plasmid vector pLHCX F_C-fs. This allows the secretion of V_HH B5, from transfected cells, as a human Fc (dimeric humanized heavy chain only antibody) with the V_HH B5 as the variable region. We first used this V_HH–Fc to further narrow down the binding epitope within DISC1, using DISC1 truncation constructs. The DISC1 truncation constructs obtained through ESPRIT (is a high-throughput screening technique to identify soluble region within any protein) [19] DISC1$^{691-836}$, DISC1$^{539-655}$, DISC1$^{480-725}$, DISC1$^{598-715}$ were probed by western blot using an anti-human secondary antibody. Here DISC1$^{691-836}$ and DISC1$^{539-655}$ are the positive and negative controls respectively, whereas DISC1$^{480-725}$ and DISC1$^{598-715}$ are the test samples. The constructs DISC1$^{691-836}$, DISC1$^{480-725}$ and DISC1$^{598-715}$ were clearly detected by V_HH–Fc antibody indicating the binding site to be within the aa residues 691-715. (Fig 3A) Similarly, NLF neuroblastoma cells transfected with full length human DISC1 or mouse Disc1 were used to test the specificity of the V_HH-Fc antibody. Intriguingly, only human DISC1 was detected and not mouse Disc1 (Fig 3B). An amino acid sequence alignment of the two proteins (UniProtKB

Q9NRI5 and Q811T9) in the 691–715 region (Figure 3C) reveals high overall similarity, with non-conservative exchanges (human > mouse) at positions 692 (E > K), 699 (R > Q), and 711 (R > G). The interaction of our V_HH B5 with human DISC1 is thus likely to involve at least one of these charged residues.

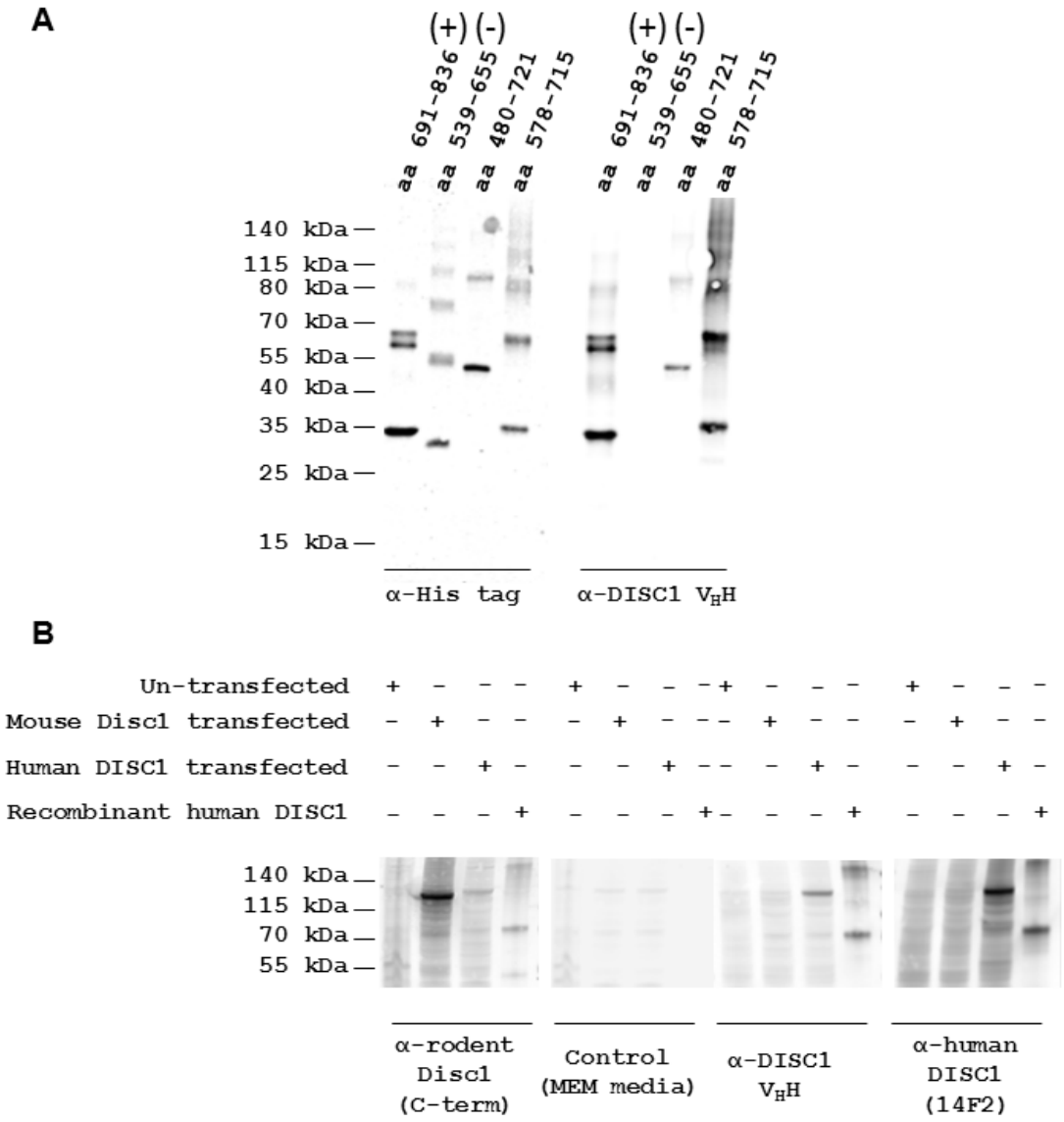

Figure 3. Mapping the binding site of the anti-DISC1 V_HH B5 antibody on the DISC1 protein.

(A) The binding epitope of the anti-DISC1 V_HH B5 antibody was mapped by western blot using truncated DISC1 protein constructs. (+) and (-) indicate positive and negative control, respectively. All constructs were detected by the anti-His tag antibody whereas only the constructs covering the epitope aa 691-715 were detected by anti-DISC1 V_HH B5 antibody (B) Detection of human and mouse DISC1 protein in the lysates oftransfected NLF cells in western blot by either anti-rodent Disc1C-term antibody, or control MEM media, or anti-DISC1 V_HH B5antibody or anti-human DISC114F2 antibody. Four samples tested are untransfected NLF cell lysate, mouse Disc1 transfected NLF cell lysate, human DISC1 transfected NLF cell lysate, recombinant human DISC1 (50ng). (C) Alignment of the human DISC1 epitope i.e., aa 691-715, with the corresponding region of mouse Disc1.

SAXS analysis of the DISC1 $^{691-836}$ protein

In order to further characterize complex formation between the DISC1$^{691-836}$ protein and the V_HH B5, SAXS experiments were performed. As shown in Figure 4A, the shapes of the scattering curves differ significantly between the V_HH B5 on the one hand and the DISC1$^{691-836}$ protein (as well as the complex) on the other hand. Indeed, the two components are expected to attain very different structures; while immunoglobulin domains are roughly ellipsoidal and mostly contain beta structure, the DISC1$^{691-836}$ protein is predicted to form several alpha helices with different lengths, resulting in a more elongated and possibly irregular shape[19]. In all three cases, the Guinier plot features a linear segment at low momentum transfer (inset), as expected for (essentially) monodisperse samples; the radius of gyration (R_g) can be derived from the slope of the respective linear fit, yielding values of 2.64

nm, 1.81 nm, and 3.07 nm for the DISC1$^{691-836}$ protein, the V_HH antibody and their complex, respectively. In addition, approximate volumes (and corresponding molecular masses) of the hydrated species can be determined by applying the Porod equation; in accordance with the R_g values, the complex appears to be significantly larger (76.8 nm^3/45.2 kDa) than the DISC1$^{691-836}$ protein alone (40.9 nm^3/24.1 kDa) or the V_HH antibody (21.3 nm^3/12.5 kDa). Hence, our SAXS analysis supports binding of the V_HH domain B5 to the DISC1$^{691-836}$ protein in solution. We note, however, that the size of the complex estimated by this method significantly exceeds the sum of the values for the two individual proteins. Given that both SEC and SPR results support a 1:1 stoichiometry, in accordance with the expected binding mode of a monoclonal antibody, this discrepancy is most likely due to a limited degree of oligomerization which was not detectable in the Guinier plot.

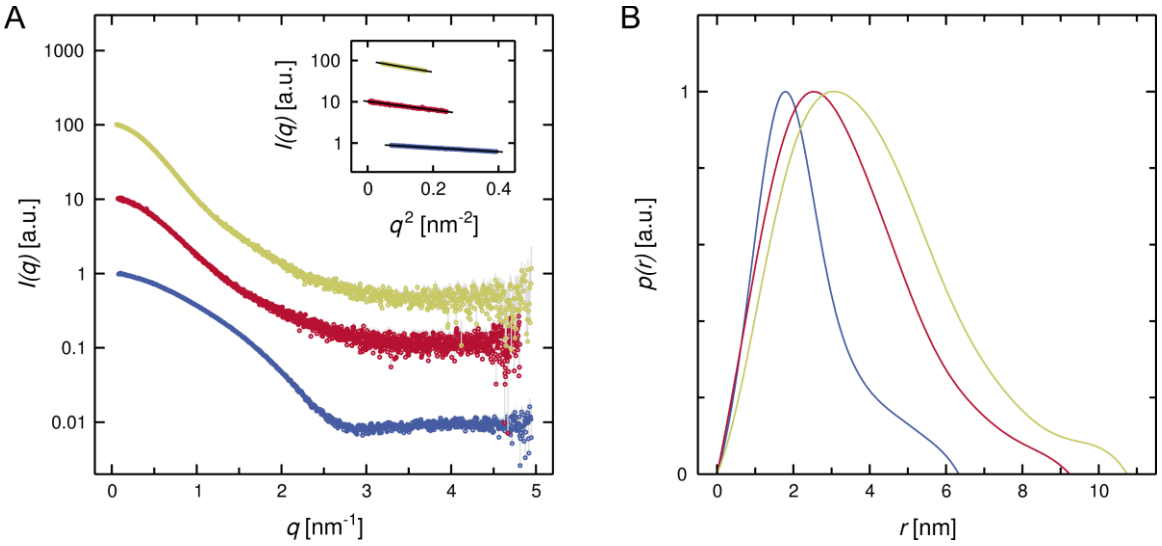

Figure 4. Investigation of DISC1-V_HH complex formation by SAXS.

(A) Primary scattering data (Intensity *I* as a function of momentum transfer *q*, the latter being defined as $4\pi \sin\theta / \lambda$) for the V_HH domain (blue), the DISC1$^{691-836}$ protein (red), and their complex (yellow). The inset shows the linear regions of the respective Guinier plots at low q values, as suggested by *AUTORG*; the slope of the linear fit is related to the radius of gyration

427 of the respective sample. (B) Distance distribution functions derived from the data shown in
428 (A), using the indirect Fourier transformation implemented in *DATGNOM*. For each species,
429 the intersection with the positive abscissa corresponds to the maximum particle diameter.
430 A.u., arbitrary units.

431

Discussion

In this study we report the first single domain antibody against the human DISC1 protein. V_HH B5 was raised against the DISC1$^{598–785}$ protein. Following recombinant expression and purification, binding of the V_HH B5 to the DISC1 protein C region (aa 691-836, previously identified to constitute a stable and soluble domain [19]) was demonstrated by three independent methods: size exclusion chromatography, surface plasmon resonance spectroscopy, and SAXS. The binding epitope of our anti-DISC1 V_HH B5 antibody was determined to reside between residues 691 and 715 of human DISC1. Moreover it was shown to be specific for human DISC1 and does not detect the mouse Disc1 protein expressed in transfected NLF cells. We also performed SAXS analysis of the DISC1$^{691-836}$ protein and its complex with the generated V_HH B5 antibody. The data on DISC1$^{691-836}$ confirmed our previous findings concerning the monomeric form and its extended shape.

In addition to quantitative parameters related to the size of the scattering particles, such as R_g and Porod volume, the reciprocal space diffraction profiles or, alternatively, the real-space distance distribution functions (Figure 4) can be used to calculate *ab-initio* reconstructions, which essentially indicate the outer shape of the scattering particles. These envelopes are shown in Figure 5 for the DISC1$^{691-836}$ protein, V_HH B5, and their complex. Moreover, we have developed molecular models to illustrate the structures of these polypeptides and their potential mode of interaction. It is important to note that the Protein Data Bank does not contain any structures with obvious sequence similarity to DISC1, precluding conventional homology modelling. We therefore resorted to the *ab-initio* approach implemented in *QUARK* [43] (for details refer to *Experimental Procedures*). While this type of structure assembly is inherently less reliable than template-based approaches, the DISC1$^{691-836}$ protein should represent a relatively favourable case due to its moderate size and predicted abundance of helical secondary structure [56]. Indeed, the resulting model is dominated by a bundle of three long α-helices, the N-terminal one being connected to the other two by a stretch containing

three shorter helical segments (Figure 5, upper left). In addition to the termini, two extended loops (residues 723-737 and 752-771) are suggested to display significant conformational freedom. Among those the 723-737 segment is particularly noteworthy since it contains a proline-rich motif (^{730}PPIPP^{734}). This proline-rich motif has been proposed to function as docking site for the Grb2 SH3 domain by mutational analyses [57].

Additional sites with known functional relevance in human DISC1 include S704 and S713, both of which are shown in a close-up view in Figure 5. Position 704 is affected by a single-nucleotide polymorphism leading to an exchange for cysteine, which has been associated with mental disorders [12]. According to our model, it is located in the first long α-helix and would not seem to be available for protein-protein interactions. Note that studies investigating the consequences of the S704C mutation revealed only moderate effects on the oligomerization propensity of C-terminal DISC1 fragments. [18,58]; this observation supports the notion that S704 may not be solvent-exposed. S713, on the other hand, has recently been demonstrated to be targeted by kinases in vivo, and its phosphorylation is thought to mediate the switch in DISC1 function from a proliferation-promoting to a migration-promoting state [15]. This residue is located in the loop adjacent to the first helix and is predicted to be readily accessible. In addition to a change in size and charge of the serine side chain itself, phosphorylation is likely to impart some degree of local rearrangement (possibly translating into long-range effects) due to the presence of basic side chains in its vicinity. It is reasonable to assume that such effects underlie the switch in the DISC1 interactome associated with S713 phosphorylation.

In contrast to the DISC1$^{691-836}$ protein, the structure of the V_HH B5 domain could be readily modelled based on a template with high sequence similarity (PDB ID 3EZJ). The nanobody displays the canonical V-type immunoglobulin fold, a sandwich composed of a five-stranded and a four-stranded β-sheet which are connected by a conserved disulfide bridge. As shown in Figure 5, both models fit the SAXS-derived envelopes reasonably well, with NSD values of

1.36 and 1.40 for the DISC1$^{691-836}$ protein and the V_HH B5 antibody, respectively. We also developed a model of the DISC1-V_HH B5 complex via *in-silico* docking simulations. One of the candidate arrangements places the arginine side chain at position 699 of the DISC1 protein in close proximity to the complementarity-determining region (CDR) loops of the nanobody, where it is able to form a salt bridge with an aspartic acid at position H53 (Clothia numbering) in CDR2. This model is consistent with our experimental evidence locating core interaction determinants in the 691-715 region of DISC1. Indeed, R699 is one out of three charged residues in this segment that are exchanged in mouse Disc1, possibly underlying the specificity of our V_HH antibody for the human orthologue.

We anticipate V_HH B5 to perform favorably in a wide range of potential applications. It should constitute an efficient DISC1 probe for use in in-vitro and in-vivo investigations and specifically target the S and C regions, which we have recently suggested to represent structural units of the DISC1 protein [19]. A relevant prospect would be to utilize V_HH B5 to gain mechanistic insight into the aggregation of the DISC1 disease variant carrying a frame shift mutation at residue 807[2], both with the full-length DISC1 protein and the isolated C region. Given the numerous reports of using single-domain antibodies for structural investigation of challenging proteins through techniques such as X-ray crystallography and cryo-electron microscopy, V_HH domains like the one described in this study may ultimately pave the way for experimental determination of the DISC1 structure, which has defied all efforts to date. [59,60].

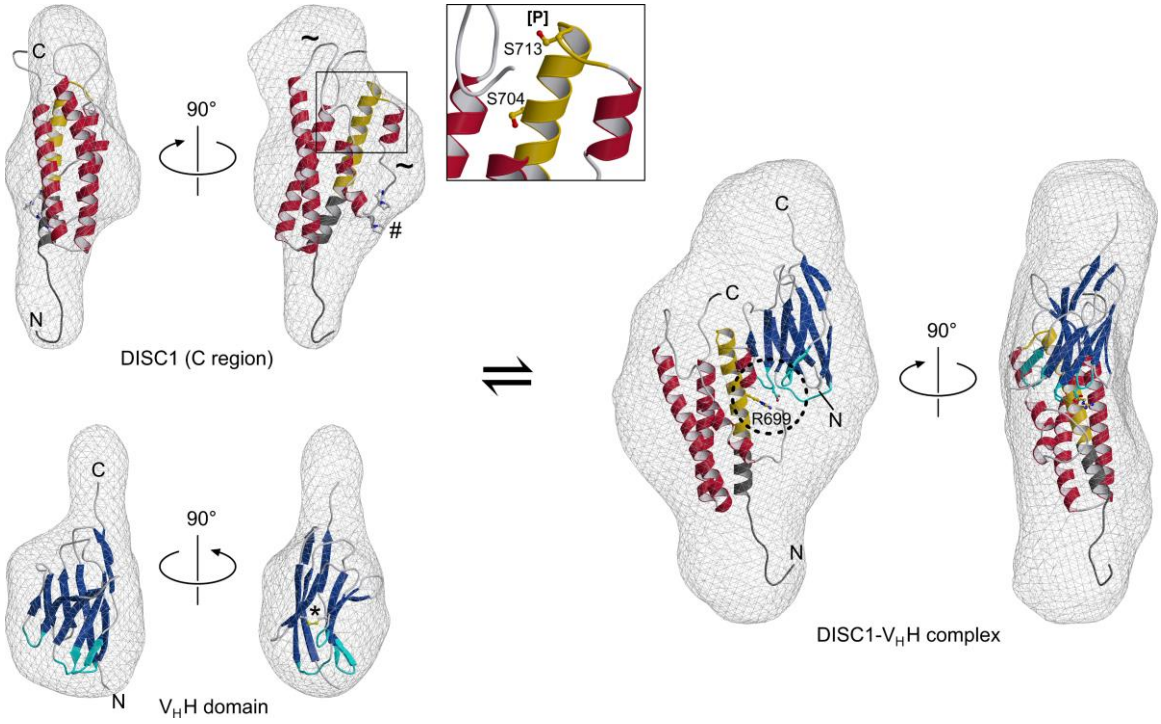

Figure 5. Ab-initio SAXS reconstructions and molecular models for the DISC1$^{691-836}$ protein, the V_HH domain, and their complex.

The hypothetical DISC1 model (upper left) features a single domain containing three longer and three shorter α-helices. Similar to the terminal segments, the two extended loops (tilde symbols) are suggested to be poorly ordered. The positions of functionally important residues S704 and S713 are indicated in a close-up view. Also note the proline-rich motif (ball-and-stick representation, marked by a hash), which is predicted to be readily accessible for protein-protein interactions. The 691–715 region suggested to contain the paratope for the nanobody is highlighted in gold, cloning artifacts at the termini are colored dark gray. The model of the V_HH antibody (lower left) displays a canonical immunoglobulin fold, including a conserved disulfide bridge (ball-and-stick model, marked by an asterisk). The CDR segments, as defined by the Clothia criteria [61], are highlighted in cyan. Finally, the DISC1 fragment aa 691-836 and the nanobody are predicted to interact in a complex (right) that is stabilized by a critical salt bridge between R699 of DISC1 and an aspartic acid in CDR2 of the V_HH molecule (dotted circle).

AUTHOR INFORMATION

Corresponding Author

Prof. Dr. Carsten Korth Institut für Neuropathologie, Universitätskliniken Düsseldorf, Moorenstraße 5, 40225 Düsseldorf, Germany, Telephone: +49 211 8116153 (CK); Fax: +49 211 8117804; ckorth@hhu.de (CK)

Present Addresses

NJB: Department of Biotechnology, University of Rijeka, Croatia

SK: Neurochirurgische Klinik, Universitätskliniken Düsseldorf, Düsseldorf, Germany

Author Contributions

ASKY, AMS, OHW, and CK designed the research. ASKY, TZ, SK, XI, RM and IP performed experiments. ASKY, AMS, AS, NJB, OHW and CK analyzed the data. ASKY, AMS, OHW, and CK wrote the paper, with input from the other authors.

Funding Sources

Funding was provided by NRW Graduate Schools BioStruct and HHU Graduate School iBRAIN for CK, as well as from EU Seventh Framework Program (MC-IN "IN-SENS" #60761 to CK) and the Brain Behavior Research Foundation (NARSAD Independent Investigator Award #20350 to CK), the Fritz Thyssen Stiftung (10.14.2.140 to NJB).

Acknowledgement

The authors are grateful to Prof. Dieter Willbold for his support in utilizing the facilities at the Institute of Complex Systems (Forschungszentrum Jülich, Germany), and to Vladimir Arinkin for recording SAXS data. We acknowledge the European Synchrotron Radiation Facility and the European Molecular Biology Laboratory Hamburg for provision of synchrotron radiation facilities and would like to thank Dr. Martha Brennich and Dr. Clement Blanchet for assistance in using beamlines BM29 and P12, respectively.

Abbreviations

aa, amino acid; CFU; colony-forming unit; CTR, carboxy-terminal region; CDR, complementarity determining region; PBS, phosphate buffer saline; PCR, polymerase chain reaction, SAXS; small angle X-ray scattering; SEC: size exclusion chromatography.

References

1. Millar JK, Wilson-Annan JC, Anderson S, Christie S, Taylor MS, et al. (2000) Disruption of two novel genes by a translocation co-segregating with schizophrenia. Human Molecular Genetics 9: 1415-1423.
2. Sachs NA, Sawa A, Holmes SE, Ross CA, DeLisi LE, et al. (2005) A frameshift mutation in Disrupted in Schizophrenia 1 in an American family with schizophrenia and schizoaffective disorder. Mol Psychiatry 10: 758-764.
3. Camargo LM, Collura V, Rain JC, Mizuguchi K, Hermjakob H, et al. (2007) Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. Mol Psychiatry 12: 74-86.
4. Yerabham AS, Weiergraber OH, Bradshaw NJ, Korth C (2013) Revisiting disrupted-in-schizophrenia 1 as a scaffold protein. Biol Chem 394: 1425-1437.
5. Kamiya A, Kubo K-i, Tomoda T, Takaki M, Youn R, et al. (2005) A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. Nat Cell Biol 7: 1167-1178.

6. Mao Y, Ge X, Frank CL, Madison JM, Koehler AN, et al. (2009) Disrupted in Schizophrenia 1 Regulates Neuronal Progenitor Proliferation via Modulation of GSK3β/β-Catenin Signaling. Cell 136: 1017-1031.
7. Lipina TV, Roder JC (2014) Disrupted-In-Schizophrenia-1 (DISC1) interactome and mental disorders: Impact of mouse models. Neuroscience & Biobehavioral Reviews 45: 271-294.
8. Trossbach SV, Bader V, Hecher L, Pum ME, Masoud ST, et al. (2016) Misassembly of full-length Disrupted-in-Schizophrenia 1 protein is linked to altered dopamine homeostasis and behavioral deficits. Mol Psychiatry 21: 1561-1572.
9. Dahoun T, Trossbach SV, Brandon NJ, Korth C, Howes OD (2017) The impact of Disrupted-in-Schizophrenia 1 (DISC1) on the dopaminergic system: a systematic review. Transl Psychiatry 7: e1015.
10. Brandon NJ, Handford EJ, Schurov I, Rain JC, Pelling M, et al. (2004) Disrupted in Schizophrenia 1 and Nudel form a neurodevelopmentally regulated protein complex: implications for schizophrenia and other major neurological disorders. Molecular and Cellular Neuroscience 25: 42-55.
11. Niethammer M, Smith DS, Ayala R, Peng J, Ko J, et al. (2000) NUDEL Is a Novel Cdk5 Substrate that Associates with LIS1 and Cytoplasmic Dynein. Neuron 28: 697-711.
12. Callicott JH, Straub RE, Pezawas L, Egan MF, Mattay VS, et al. (2005) Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. Proc Natl Acad Sci U S A 102: 8627-8632.
13. Hashimoto R, Numakawa T, Ohnishi T, Kumamaru E, Yagasaki Y, et al. (2006) Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. Hum Mol Genet 15: 3024-3033.
14. Chubb JE, Bradshaw NJ, Soares DC, Porteous DJ, Millar JK (2008) The DISC locus in psychiatric illness. Mol Psychiatry 13: 36-64.
15. Ishizuka K, Kamiya A, Oh EC, Kanki H, Seshadri S, et al. (2011) DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. Nature 473: 92-96.
16. Bader V, Ottis P, Pum M, Huston JP, Korth C (2012) Generation, purification, and characterization of cell-invasive DISC1 protein species. J Vis Exp: e4132.
17. Ottis P, Bader V, Trossbach SV, Kretzschmar H, Michel M, et al. (2011) Convergence of two independent mental disease genes on the protein level: recruitment of dysbindin to cell-invasive disrupted-in-schizophrenia 1 aggresomes. Biol Psychiatry 70: 604-610.
18. Leliveld SR, Hendriks P, Michel M, Sajnani G, Bader V, et al. (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.
19. Yerabham ASK, Mas PJ, Decker C, Soares DC, Weiergraber OH, et al. (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.
20. Hassanzadeh-Ghassabeh G, Devoogdt N, De Pauw P, Vincke C, Muyldermans S (2013) Nanobodies and their potential applications. Nanomedicine (Lond) 8: 1013-1026.
21. Dumoulin M, Conrath K, Van Meirhaeghe A, Meersman F, Heremans K, et al. (2002) Single-domain antibody fragments with high conformational stability. Protein Sci 11: 500-515.
22. Muyldermans S (2013) Nanobodies: natural single-domain antibodies. Annu Rev Biochem 82: 775-797.
23. Rudolph MJ, Vance DJ, Cassidy MS, Rong Y, Mantis NJ (2017) Structural Analysis of Single Domain Antibodies Bound to a Second Neutralizing Hot Spot on Ricin Toxin's Enzymatic Subunit. J Biol Chem 292: 872-883.
24. Li T, Vandesquille M, Koukouli F, Dudeffant C, Youssef I, et al. (2016) Camelid single-domain antibodies: A versatile tool for in vivo imaging of extracellular and intracellular brain targets. Journal of Controlled Release 243: 1-10.
25. Thys B, Saerens D, Schotte L, De Bleeser G, Muyldermans S, et al. (2011) A simple quantitative affinity capturing assay of poliovirus antigens and subviral particles by single-domain antibodies using magnetic beads. Journal of Virological Methods 173: 300-305.

26. Pardon E, Laeremans T, Triest S, Rasmussen SGF, Wohlkönig A, et al. (2014) A general protocol for the generation of Nanobodies for structural biology. Nat Protocols 9: 674-693.
27. Strauss M, Schotte L, Karunatilaka KS, Filman DJ, Hogle JM (2017) Cryo-electron Microscopy Structures of Expanded Poliovirus with VHHs Sample the Conformational Repertoire of the Expanded State. J Virol 91.
28. Bever CS, Dong J-X, Vasylieva N, Barnych B, Cui Y, et al. (2016) VHH antibodies: emerging reagents for the analysis of environmental chemicals. Analytical and Bioanalytical Chemistry 408: 5985-6002.
29. Doyle PJ, Arbabi-Ghahroudi M, Gaudette N, Furzer G, Savard ME, et al. (2008) Cloning, expression, and characterization of a single-domain antibody fragment with affinity for 15-acetyl-deoxynivalenol. Mol Immunol 45: 3703-3713.
30. Alvarez-Rueda N, Behar G, Ferre V, Pugniere M, Roquet F, et al. (2007) Generation of llama single-domain antibodies against methotrexate, a prototypical hapten. Mol Immunol 44: 1680-1690.
31. Griffiths AD, Williams SC, Hartley O, Tomlinson IM, Waterhouse P, et al. (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. The EMBO Journal 13: 3245-3260.
32. Wang Y, Zhang X, Zhang C, Liu Y, Liu X (2012) Isolation of single chain variable fragment (scFv) specific for Cry1C toxin from human single fold scFv libraries. Toxicon 60: 1290-1297.
33. Wang Q, Charych EI, Pulito VL, Lee JB, Graziane NM, et al. (2011) The psychiatric disease risk factors DISC1 and TNIK interact to regulate synapse composition and function. Mol Psychiatry 16: 1006-1023.
34. Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, et al. (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). J Appl Crystallogr 48: 431-443.
35. Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, et al. (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. J Synchrotron Radiat 20: 660-664.
36. Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. J Appl Crystallogr 45: 342-350.
37. Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI (2007) ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. Journal of Applied Crystallography 40: s223-s228.
38. Svergun D, Barberato C, Koch MHJ (1995) CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. Journal of Applied Crystallography 28: 768-773.
39. Franke D, Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Crystallogr 42: 342-346.
40. Svergun DI, Petoukhov MV, Koch MH (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys J 80: 2946-2953.
41. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography 36: 860-864.
42. Wriggers W (2012) Conventions and workflows for using Situs. Acta Crystallogr D Biol Crystallogr 68: 344-351.
43. Xu D, Zhang Y (2012) Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. Proteins 80: 1715-1735.
44. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. Journal of Applied Crystallography 34: 33-41.
45. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356: 83-85.
46. Benkert P, Tosatto SC, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. Proteins 71: 261-277.
47. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234: 779-815.

683 48. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta
684 Crystallogr D Biol Crystallogr 66: 486-501.
685 49. Korotkov KV, Pardon E, Steyaert J, Hol WG (2009) Crystal structure of the N-terminal domain of
686 the secretin GspD from ETEC determined with the assistance of a nanobody. Structure 17:
687 255-265.
688 50. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, et al. (2017) The ClusPro web server for protein-
689 protein docking. Nat Protoc 12: 255-278.
690 51. Brenke R, Hall DR, Chuang GY, Comeau SR, Bohnuud T, et al. (2012) Application of asymmetric
691 statistical potentials to antibody-protein docking. Bioinformatics 28: 2608-2614.
692 52. Wahle M, Wriggers W (2015) Multi-scale Visualization of Molecular Architecture Using Real-Time
693 Ambient Occlusion in Sculptor. PLoS Comput Biol 11: e1004516.
694 53. Fenn TD, Stamper GF, Morollo AA, Ringe D (2003) A side reaction of alanine racemase:
695 transamination of cycloserine. Biochemistry 42: 5775-5783.
696 54. Merritt EA, Bacon DJ (1997) Raster3D: photorealistic molecular graphics. Methods Enzymol 277:
697 505-524.
698 55. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of
699 hydrogen-bonded and geometrical features. Biopolymers 22: 2577-2637.
700 56. Zhang Y (2014) Interplay of I-TASSER and QUARK for template-based and ab initio protein
701 structure prediction in CASP10. Proteins 82 Suppl 2: 175-187.
702 57. Shinoda T, Taya S, Tsuboi D, Hikita T, Matsuzawa R, et al. (2007) DISC1 regulates neurotrophin-
703 induced axon elongation via interaction with Grb2. J Neurosci 27: 4-14.
704 58. Narayanan S, Arthanari H, Wolfe MS, Wagner G (2011) Molecular characterization of disrupted in
705 schizophrenia-1 risk variant S704C reveals the formation of altered oligomeric assembly. J
706 Biol Chem 286: 44266-44276.
707 59. Schotte L, Strauss M, Thys B, Halewyck H, Filman DJ, et al. (2014) Mechanism of action and
708 capsid-stabilizing properties of VHHs with an in vitro antipolioviral activity. J Virol 88: 4403-
709 4413.
710 60. Wu M, Park YJ, Pardon E, Turley S, Hayhurst A, et al. (2011) Structures of a key interaction protein
711 from the Trypanosoma brucei editosome in complex with single domain antibodies. J Struct
712 Biol 174: 124-136.
713 61. Al-Lazikani B, Lesk AM, Chothia C (1997) Standard conformations for the canonical structures of
714 immunoglobulins. J Mol Biol 273: 927-948.

715

Declaration

I hereby declare that the work presented here is reflection of my own independent efforts and had been conducted without any unauthorized assistance. Wherever contributions and consultation of other sources are involved, every effort is made to indicate this clearly with due reference to the literature and acknowledgement of collaborative research and discussions, if any. Moreover the dissertation has never been submitted in any form to any other institution.

Düsseldorf, June 2017

(Antony Sravan Kumar Yerabham)