Structural characterization of intrinsically disordered prostate apoptosis response factor 4 (Par-4)

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

Aziz Tumeh
from Qafeen/Tulkarm/Palestine

Jülich, June 2024

from the institute IBI/7- structural biochemistry
at the Forschungszentrum Jülich

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Jörg Labahn
Co-supervisor: Prof. Oliver H. Weiergräber

Date of the oral examination: 01. 10. 2024

Dedicated to

my Parents

my wife

my son

Table of Contents

Abbreviations

AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride
AGE	Agraose Gel Electrophoresis
Akt1	Protein kinase B
ATP	Adenosine triphosphate
Apaf-1	Apoptotic Protease Activating Factor 1
aPKC	atypical PKC
BeStSel	Beta Structure Selection
cAMP	Cyclic adenosine monophosphate
CC	Coiled Coil
CD	Circular Dichroism
cl-Par-4	Cleaved Par-4
CRM-1	chromosome region maintenance protein 1
CV	Column Volume
DAPK3	death associated protein kinase 3
DESY	Deutsches Elektronen-Synchrotron
DISC	death-inducing signaling complex
DHAP	dihydroxy acetophenone
DLS	Dynamic Light Scattering
DMSO	dimethyl sulfoxide
EDTA	Ethylenediamine tetraacetic acid
ER	endoplasmic reticulum
ESRF	European Synchrotron Radiation Facility
FADD	Fas-associated death domain
GRP78	Glucose regulated protein 78
HRP	horseradish peroxidase
IPTG	Isopropyl β- d-1-thiogalactopyranoside
kDa	Kilodalton
LB	Lysogeny Broth
LC/MS	Liquid Chromatography/Mass Spectrometry

LZ	Leucine Zipper
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time-of-Flight
MBP	Maltose-Binding protein
MD	Megadaltone
MIR	multiple isomorphous replacement
MRE	Mean Residual Ellipticity
MS	Mass spectrometry
MW	Molecular Weight
MWCO	Molecular Weight CutOff
NCS	non-crystallographic symmetry
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NLS1	Nuclear Localization Signal 1
NLS2	Nuclear Localization Signal 2
NMR	Nuclear Magnetic Resonance
PAF	Par-4 amino-terminal fragment
Par-4	Prostate Apoptosis Response factor 4
Par-4 CC EXS	Extended region from CC domain to SAC domain
PAWR	pro-apoptotic WT1 regulator
PCR	Polymerase Chain Reaction
PD	polydispersity
PEG	polyethylene glycol
PEI	Polyethyleneimine
PFO	Perfluoropolyether
PKA	Protein kinase A
PKCζ	Protein kinase C, zeta
PML-NBs	promyelocytic leukemia nuclear bodies
PMSF	phenylmethylsulfonyl fluoride
PNGase F	Peptide -N-Glycosidase F
Psi	Pound per square inch
PTMs	post-translation modifications
PVDF	polyvinylidene fluoride
RMS	Root Mean Square
RT	Room Temperature

SAC	Selective for Apoptosis induction in cancer cells
SDS-PAGE	Sodium Dodecyl Sulphate–PolyAcrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid-D
TGE	Transient Gene Expression
TLS	Translation/Libration/Screw
Tm	Melting temperature
TNF-α	Tumour Necrosis Factor alpha
UV-VIS	Ultraviolet-visible

List of figures

List of tables

1 Introduction

1.1 Apoptosis

Apoptosis, or programmed cell death, is a tightly regulated process that plays a crucial role in maintaining cellular homeostasis, development, and tissue integrity (Elmore, 2007). It serves as a fundamental mechanism for eliminating unwanted or damaged cells, allowing for the controlled removal of cells without triggering an inflammatory response. Apoptosis is essential for various physiological processes, including embryonic development, immune response, tissue remodeling, and maintenance of proper cell populations in adult organisms. Dysregulation of apoptosis is implicated in the development and progression of numerous diseases, including cancer, neurodegenerative disorders, autoimmune diseases, and cardiovascular diseases (Saikumar et al., 1999). Insufficient apoptosis can lead to the accumulation of abnormal or damaged cells, contributing to tumor formation and autoimmune disorders. Conversely, excessive apoptosis can result in tissue degeneration and neurodegenerative diseases, such as Alzheimer's and Parkinson's (Jin & El-Deiry, 2005).

Apoptosis is regulated by a complex network of signaling pathways, involving numerous key players. The two major pathways involved in apoptosis are the intrinsic (mitochondrial) pathway and the extrinsic (death receptor) pathway. These pathways converge to activate a cascade of proteolytic enzymes called caspases, which ultimately execute the apoptotic process (Galluzzi et al., 2015; Green & Llambi, 2015).

The intrinsic pathway is triggered by various intracellular signals, including DNA damage, oxidative stress, and loss of cell survival factors. Key players in this pathway include Bcl-2 family proteins, such as Bax and Bak, which regulate mitochondrial outer membrane permeabilization, releasing cytochrome c and activating caspases. Cytochrome c, upon release from mitochondria, binds to Apaf-1 (apoptotic protease-activating factor 1) and forms the apoptosome complex, leading to caspase activation (Hongmei, 2012; Ouyang et al., 2012).

The extrinsic pathway, on the other hand, is initiated by the binding of death ligands, such as TNF-α and Fas ligand, to specific cell surface death receptors like TNFR1 and Fas receptor. Death receptors recruit and activate adapter proteins like FADD (Fas-associated death domain), initiating the formation of the death-inducing signaling complex (DISC). Caspase-8, activated at the DISC, can directly activate effector caspases such as caspase-3 or trigger the intrinsic pathway through cleavage of Bcl-2 family proteins (Hongmei, 2012; Schultz & Harringto Jr, 2003).

1.2 Prostate Apoptosis Response Factor 4 (Par-4)

1.2.1 Par-4 discovery

The Prostate Apoptosis Response Factor 4 (Par-4), or PAWR (PRKC, apoptosis, WT1, regulator), is a versatile protein initially identified through differential display screening of androgen-dependent and -independent rat prostate cancer cells. Where Its expression was found to be upregulated in cells undergoing apoptosis (Sells et al., 1994). What sets Par-4 apart is its exceptional ability to induce apoptosis specifically in certain cancer cells when ectopically expressed or activated, while leaving normal or immortalized cells unharmed (Hebbar et al., 2012).

Par-4's significance spans across various species with its widespread expression in different tissues. However, its depletion, mutation, or inactivation has been observed in numerous cancer types (Boehrer et al., 2002; Cook et al., 1999; Moreno-Bueno et al., 2007; Nagai et al., 2010). The human Par-4 gene resides on chromosome 12q21.31, a region frequently deleted in pancreatic and gastric cancers, as well as tumors of male germ cells (Johnstone et al., 1998). This downregulation of Par-4 is linked to tumor recurrence and reduced patient survival (Mabe et al., 2018).

The cellular localization of Par-4 can vary depending on the cell type and its activation status. Par-4 is predominantly found in the cytoplasm in normal cells but translocates to the nucleus upon specific stimuli or interactions with partner proteins in most cancer cells (El-Guendy, 2003) . This dynamic localization allows Par-4 to exert its pro-apoptotic function and potentially influence gene expression in the nucleus. The

expression patterns of Par-4 can differ between normal cells and cancer cells. In normal cells, Par-4 expression is generally low, while in certain cancer cells, including prostate cancer, it is upregulated (El-Guendy & Rangnekar, 2003). This differential expression suggests the potential involvement of Par-4 in tumorigenesis and the regulation of cellular processes associated with cancer development.

Par-4's functional repertoire extends beyond apoptosis induction, as it is involved in various cellular processes and implicated in several diseases. It acts as a pro-apoptotic protein, initiating apoptosis in response to diverse stimuli, employing various mechanisms such as inhibiting anti-apoptotic proteins, modulating mitochondrial function, and activating caspases (El-Guendy & Rangnekar, 2003). In addition, it serves as a tumor suppressor in various cancers where its loss is observed (Shrestha-Bhattarai & Rangnekar, 2010). Aligned with its significant involvement in apoptosis, the ectopic expression of Par-4 has been demonstrated to sensitize a wide range of cultured cells, such as NIH3T3 cells, neurons, melanoma, prostate, and colon cancer cells, to apoptotic stimuli and chemotherapeutic agents(Diaz-Meco et al., 1999; Díaz-Meco et al., 1996; Guo et al., 2001; Sells et al., 1997). Moreover, elevated Par-4 levels have been identified in neurons affected by neurodegenerative diseases and HIV encephalitis (Duan et al., 2000; Guo et al., 1998; Kruman et al., 1999; Pedersen et al., 2000). In addition to apoptosis, PAR-4 plays roles in regulating autophagy, cell migration, and inflammation, as it interacts with various proteins and signaling pathways, suggesting its involvement in diverse physiological and pathological contexts (Cheratta et al., 2021).

1.2.2 Functional Domains of Par-4

The Par-4 protein in humans is comprised of 340 amino acids, while the mouse Par-4 has 333 amino acids and the rat Par-4 has 332 amino acids. Across species, Par-4 exhibits several functionally conserved domains, highlighting their evolutionary significance. Comparative analysis of Par-4 sequences among vertebrates reveals a notable degree of conservation, underscoring the fundamental role of Par-4 in the regulation of apoptosis across different species (Ranganathan & Rangnekar, 2005).

When comparing the rat Par-4 sequence with those of mouse and human Par-4 through UniProt, it was found that the sequence identity was 90% and 72%, respectively. Although specific residues may differ, the overall structural and functional domains of Par-4 remain well-preserved (Figure 1.2.2.1). This conservation supports the notion that the function of Par-4 is intricately governed through interactions with other proteins, intercellular trafficking processes, and post-translational modifications, ensuring precise regulation of its activity (Cheratta et al., 2021).

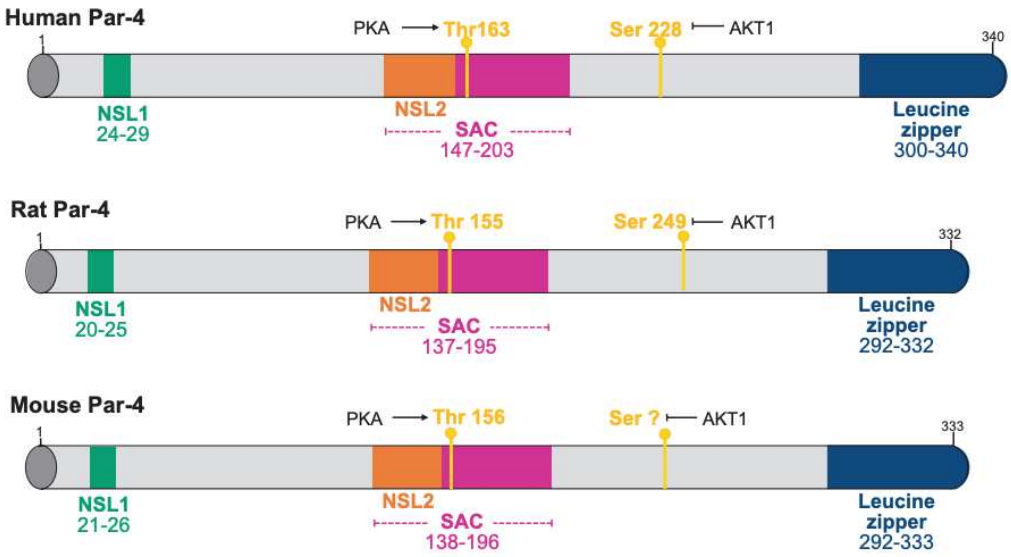


Figure 1.2.2.1 Functional domains of Par-4 (figure adapted from (Ganguly et al., 2022))

The functional domains within Par-4 exhibit a high degree of conservation among human, rat, and mouse species. The numerical values correspond to their respective amino acid positions. Specifically, NLS1 denotes Nuclear Localization Signal 1, NLS2 represents Nuclear Localization Signal 2, SAC signifies Selectivity for Apoptosis in Cancer Cells, and LZ stands for the Leucine Zipper domain. Notably, Ser249 in rat Par-4 serves a function similar to that of Ser228 in human Par-4.

Par-4 is a multidomain protein. Its N-terminus contains two putative nuclear localization sequences (NLS1 and NLS2), VASA domain (in human but not in rat), and the SAC (Selective for Apoptosis in Cancer cells) domain, which is the minimal amino acid sequence to induce apoptosis (El-Guendy, 2003; Goswami et al., 2005; Zhao et al., 2007). The NLS2 sequence is located in the SAC domain. It is essential for the translocation of Par-4 from the cytoplasm to the nucleus and is necessary to cause

apoptosis (El-Guendy, 2003). Notably, deletion nutation of NLS2 but not NLS1 hampers the nuclear translocation and apoptosis induction by Par-4. Although NLS1 is conserved through evolution, no clear function has been assigned to it (Hebbar et al., 2012). Human Par-4 contains a VASA domain which is responsible for preventing Fbox45-mediated degradation of intracellular Par-4, allowing the release of intracellular Par-4 from Fbxo45 for apoptosis of cancer cells (Hebbar et al., 2017). SAC domain was initially identified through serial deletion mutations from both amino and hydroxy terminus, consists of 59 amino acids and 100% conserved in rodent and human Par-4 (El-Guendy et al., 2003). The apoptotic activity of SAC is dependent on both, NLS2 which allows the nuclear entry of SAC , and protein kinase A (PKA) which phosphorylates Thr155 within the SAC (El-Guendy, 2003; Gurumurthy et al., 2005). The ability of SAC or Par-4 to selectively trigger apoptosis in cancer cells is attributed to the elevated endogenous activity of PKA in cancer cells compared to normal cells. Notably, upregulation of PKA level in normal cells by stimulants such as cAMP led to sensitization of the cells to SAC-mediated apoptosis (Gurumurthy et al., 2005). The SAC domain has the ability to trigger apoptosis in cancer cells regardless of their sensitivity or resistance to full-length Par-4-mediated apoptosis, and yet does not induce apoptosis in healthy cell (Zhao & Rangnekar, 2008). SAC lacks Akt1 phosphorylation site, which is located between the SAC domain and C-terminus at Ser249 residue. When Ser249 is phosphorylated by Akt1, Par-4 will remain in the cytoplasm (Goswami et al., 2005). This could potentially provide an explanation for how SAC can induce apoptosis not only in Par-4-sensitive cancer cells but also in Par-4-resistant cancer cells, including those with heightened Akt1 activity (Goswami et al., 2005). SAC localizes in the nucleus of both healthy and cancer cells. However, its apoptotic function is specifically observed in cancer cells. This is because, like Par-4, SAC activation requires Thr phosphorylation by PKA, which is elevated in cancer cells compared to normal cells. Thus, SAC's apoptotic activity is uniquely triggered in the cancer cell environment due to this specific biochemical pathway (Gurumurthy et al., 2005) (Ranganathan & Rangnekar, 2005) .

While SAC domain has been mainly assigned for apoptosis induction by inhibiting NF-κB activity, it was shown to interact with other partner proteins and activate Fas pathway and induce apoptosis (El-Guendy et al., 2003). Notably, SAC domain can

inhibit Akt1 activity, which is known to suppress apoptotic pathways, through binding to its activator Glucose regulated protein 78 (GRP78)(Burikhanov et al., 2009; Sato et al., 2010). Moreover, as a result of paracrine apoptosis induction, binding of extracellular Par-4 through its SAC domain to GRP78 caused ER stress and cell death by activation of FADD/caspase-8/caspase-3 pathway (Burikhanov et al., 2009).

The carboxyl-terminal of Par-4 protein contains a leucine zipper domain (LZ) residue (292-332), and a predicated nuclear export signal (NES). The LZ domain is important for homodimeric as well as heterodimeric protein-protein interactions, and regulation of par-4 function. The significance of LZ domain is exemplified by Par-4 deletion mutants that lack this domain and consequently cannot sensitize melanoma cells to apoptotic stimuli. (Sells et al., 1997). Likewise, PC12 cells become insensitive to trophic factor withdrawal or exposure to Aβ-42 when the Par-4 LZ domain is deleted (Guo et al., 1998).

Several notable protein partners have been identified that interact with the LZ domain of Par-4, including Wilm's tumor WT1 protein, PKCζ, DAPK3, THAP1, TOP1, AATF, UACA, and Akt1 (de Thonel et al., 2014; Díaz-Meco et al., 1996; El-Guendy & Rangnekar, 2003; Shrestha-Bhattarai & Rangnekar, 2010; Tiruttani Subhramanyam et al., 2017). One of the key findings is that Par-4 binds to and inhibits the activity of PKCζ, leading to the inhibition of NF-κB activity. However, this inhibition can be countered by the formation of a ternary complex involving p62 (Chang et al., 2002). Additionally, Par-4 plays a crucial role in regulating the processing of the β-amyloid precursor protein, with its interaction with AATF blocking this process(Guo & Xie, 2004). Moreover, Par-4 interacts with DAPK3 and Amida, leading to their relocation from the nucleus to the cytoskeleton when co-expressed, which triggers cytoskeletal reorganization and induction of apoptosis (Boosen et al., 2005; Page et al., 1999). Another significant interaction involves Par-4 and WT1, which transcriptionally downregulates BCL-2, resulting in apoptosis induction(Ranganathan & Rangnekar, 2005). Furthermore, Par-4 interacts with THAP1, a protein associated with promyelocytic leukemia nuclear bodies (PML-NBs), and contains a DNA-binding THAP domain (Roussigne et al., 2003).

The NES is predicated to reside at residues 299–310 (MIGKLKEEIDLL) in human Par-4 (rat Par-4, 295–301residues) (Tiruttani Subhramanyam et al., 2017). This sequence is conserved across human, rat, and mouse Par-4 and is typically concealed upon dimerization. This unique characteristic potentially allows Par-4 to evade nuclear export and remain within the nucleus, enabling it to regulate the activities of topoisomerase 1 (TOP1) or NF-κB (Azmi et al., 2013; Goswami et al., 2008; Tiruttani Subhramanyam et al., 2017). Selective inhibitors that mimic nuclear export analogs have been developed to bind to the chromosome region maintenance protein 1 (CRM-1), inhibiting its interaction with Par-4 and subsequent its nuclear export (Azmi et al., 2013). Consequently, this inhibition leads to suppressed proliferation and induction of apoptosis, which can be hindered by Par-4 knockdown. These findings suggest that abnormal interactions involving the NES of Par-4 may facilitate nuclear expulsion and impede apoptosis (Azmi et al., 2013).

1.2.3 Caspase mediated cleavage of Par-4

Caspases, a distinctive family of cysteine proteases, play a central role in orchestrating apoptosis networks. Their cleavage of substrates can lead to diverse outcomes, such as gain-of-function, loss-of-function, or functional modifications (Julien & Wells, 2017). Intriguingly, the N-terminal region of human Par-4 is a target for caspase-3 cleavage during apoptosis, specifically at the site EEPD131↓G (rat EEPD123↓S), resulting in the formation of two distinct fragments: a 15 kDa (12 kDa in rat) amino-terminal fragment termed Par-4 amino-terminal fragment (PAF) which is suggested to serve as a decoy that protects full-length Par-4 from ubiquitin-based degradation (Hebbar et al., 2017), and a 25 kDa (24 kDa in rat) carboxy-terminal fragment referred to as cleaved Par-4 (cl-Par-4) (Chaudhry et al., 2012; Thayyullathil et al., 2013). Remarkably, cl-Par-4 encompasses both the SAC domain and the CC domain, whereas PAF harbors the NLS1 and VASA region. The cl-Par-4, containing NLS2, efficiently translocates into the nucleus, where it executes its apoptotic function, while PAF, through its VASA segment, counteracts the Fbxo45-mediated degradation of full-length Par-4 (Hebbar et al., 2017). Notably, when cells are transfected with the D131A mutant Par-4, rendering it resistant to caspase cleavage during cisplatin treatment, a significant decrease in apoptotic cells is observed, indicating that caspase-mediated

cleavage enhances Par-4's gain-of-function (Chaudhry et al., 2012). Additionally, caspase-8 also generates cl-Par-4 in response to TNFα- and UV-induced apoptosis (Treude et al., 2014). Interestingly, cl-Par-4 formation is induced by various anticancer agents, effectively eradicating cancer cells (Brasseur et al., 2016; Chaudhry et al., 2012; Guo et al., 2019; Rahman et al., 2019; Thayyullathil et al., 2013; Treude et al., 2014).

1.2.4 Role of Par-4 in Apoptosis

Apoptosis, orchestrated by Par-4, encompasses the concurrent inhibition of cell survival pathways and activation of the apoptotic pathway. It is the combined action of these pathways that culminates in Par-4-induced apoptosis, as neither pathway alone is sufficient. As illustrated in figure 1.2.4, upon Par-4 expression, Fas and FasL are trafficked to the cell membrane, facilitating the interaction between Fas and FADD. This interaction leads to the recruitment of FADD and procaspase-8, ultimately forming the death-inducing signaling complex (DISC). Through this complex, procaspase-8 is activated into caspase-8, which further triggers downstream effectors (Chakraborty et al., 2001). Cytoplasmic Par-4 exerts inhibitory effects on NF-κB activation by binding aPKC within the cytoplasm, leading to the initiation of apoptosis. The nuclear translocation of Par-4 is facilitated by PKA phosphorylation at T155 in the SAC domain. Under the influence of various apoptotic stimuli, Par-4 relocates to the nucleus, where it hinders NF-κB-mediated pro-survival signals and promotes apoptosis induction through its SAC domain. In human tumors, simply activating the apoptotic cascade is inadequate to induce cell death due to the concurrent activation of cell survival pathways. Notably, cancer cells often exhibit elevated NF-κB transcriptional activity, resulting in the expression of anti-apoptotic genes like XIAP, c-IAP1, and c-IAP2, which impede apoptosis progression. However, Par-4 has the capability to translocate to the nucleus and inhibit NF-κB transcriptional activity, thereby overcoming apoptosis resistance. Additionally, studies have shown that Par-4 in the cytosol suppresses NF-κB-dependent gene transcription by inhibiting the translocation of the NF-κB subunit RelA to the nucleus, achieved through atypical PKC inhibition (Chakraborty et al., 2001).

In addition to its nuclear role, intracellular Par-4 is also found in the endoplasmic reticulum (ER), where it exists in either an unbound state or forms complexes with GRP78. In response to ER stress, the Par-4/GRP78 complex translocates to the cell surface. Par-4 secretion can occur either upon dissociation from GRP78, before its translocation to the plasma membrane, or after reaching the cell surface. Furthermore, unbound Par-4 may be secreted through GRP78-independent pathways. The secreted Par-4 binds to GRP78 on the cell surface, enhancing the ER stress response and triggering FADD/caspase-8-dependent apoptosis via the extrinsic pathway. For the translocation of GRP78 from the ER to the membrane, the intracellular function of Par-4 is crucial in facilitating the activation of the extrinsic apoptosis pathway induced by secreted Par-4 (Burikhanov et al., 2009).

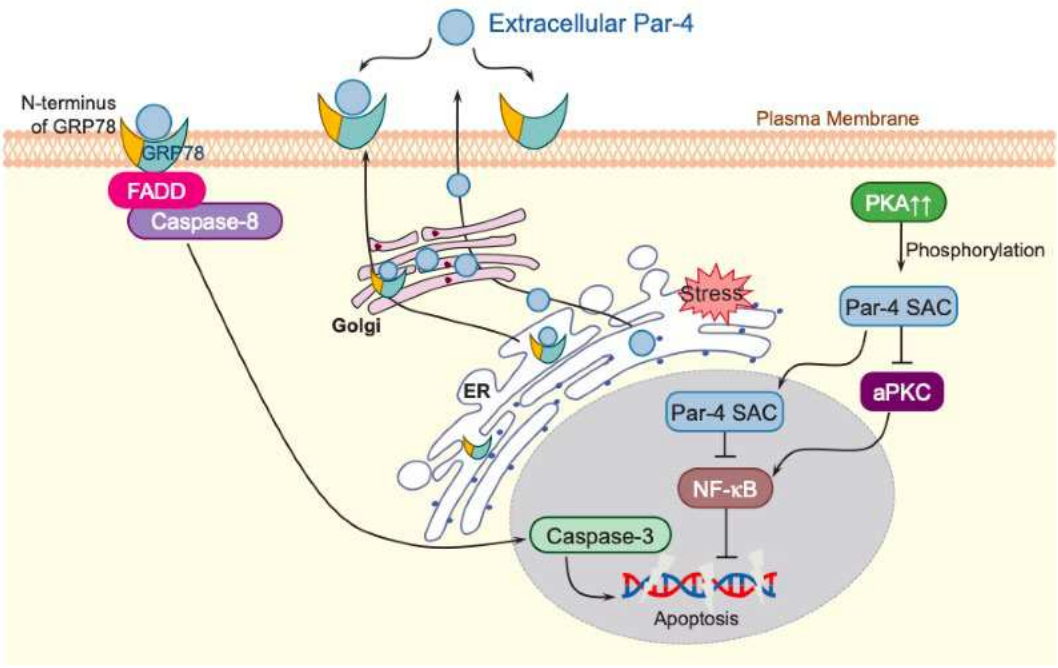


Figure 1.2.4.1 Main components of apoptosis mediated by intracellular and extracellular Par-4 (figure adapted from (Ganguly et al., 2022)

The trafficking of Fas/FasL to the cell membrane and NF-κB inhibition are dependent on the phosphorylation of Par-4 at T155 by PKA, whose enzymatic activity is elevated in cancer cells compared to normal cells. This selective phosphorylation by PKA grants Par-4 the ability to induce apoptosis exclusively in cancer cells (Gurumurthy et al., 2005). Notably, the PKA phosphorylation site in Par-4 (T155 in rat Par-4 or T163 in

human Par-4) is crucial, as an alanine mutation of this residue abolishes the apoptotic potential of ectopic Par-4. Interestingly, while endogenous Par-4 may not induce apoptosis due to an active intracellular mechanism involving inhibitory phosphorylation by cell survival kinases such as Akt, it can trigger apoptosis when cells are exposed to exogenous treatments that elevate PKA activity and T155 (or T163) phosphorylation (Gurumurthy et al., 2005). These diverse mechanisms highlight the multifaceted role of Par-4 in promoting apoptosis and its potential as a therapeutic target for cancer treatment.

1.2.5 Structural Analysis of Par-4

Bioinformatic analysis of Par-4's primary structure, using the disorder prediction algorithm DisEMBL, suggest a mostly intrinsically disordered protein. Notably, the C-terminal CC domain represents the primary region with an ordered structure. Additionally, three shorter stretches of predicted order are found near each nuclear localization signal (NLS1 near the N-terminus and NLS2 within the SAC domain) as well as in proximity to the C-terminus of the SAC domain (Libich et al., 2009).

Extensive efforts have been dedicated to investigating the structure of Par-4 since its discovery. However, due to the challenges in obtaining a well-defined structured state for this protein, only approximately one-third of its total structure (C-terminal) (residues 254-332) has been successfully determined.

Interestingly, under physiological conditions, the C-terminal region of Par-4 was predicted to form natively unfolded state (Libich et al., 2009). However, by modulating the pH and temperature, a self-associated coiled coil structure can be induced. The C-terminus exhibits a predominance of unfolded monomers under physiological conditions, but the scenario shifts as the pH and temperature decrease, favoring the formation of the coiled coil structure (Dutta et al., 2001). Notably, the LZ component demonstrates a remarkable ability to adapt its structural conformation in response to its environment, oscillating between a partially ordered monomer state and a predominantly coiled-coil dimer state (Schwalbe et al., 2010).

The C-terminal region of Par-4 (residues 254-332), encompassing the LZ domain (292-332), has been predicted to adopt a coiled coil structure (Libich et al., 2009). Excitingly, this prediction was experimentally validated using X-ray crystallography, which revealed a canonical coiled-coil arrangement within the CC domain (Tiruttani Subhramanyam et al., 2014). The CC domain consists of two right-handed helices running parallel to each other, forming a left-handed super helical twist. Remarkably, the N-terminal segment of the CC domain plays a crucial role in homodimer formation and protein folding, while the LZ domain in the C-terminal region favors heterodimerization with partner proteins. The formation of the coiled-coil structure is directed by eleven heptad repeats in their sequence denoted as (abcdefg)n, where hydrophobic amino acids predominantly occupy positions a and d (Tiruttani Subhramanyam et al., 2017).

Generally, coiled-coil structures achieve stability through the formation of interhelical salt bridges (Meier et al., 2002). In contrast, intrahelical salt bridges have been found to play a role in stabilizing monomeric α-helices (Burkhard et al., 2002; Kammerer et al., 2001). The structural analysis of Par-4 CC reveals the presence of both interhelical and intrahelical salt bridges, specifically of the g–e′, d–e′, and g–c varieties, as illustrated in Figure 1.2.5 a. Notably, interactions such as Arg263-Glu268′, Glu270-Arg275′, and Glu312-Lys317′ fall within the g–e′ category, while Glu274-Arg275′ and Lys288-Glu289′ are classified as d–e′ type interactions (1.2.5 b). Furthermore, two intrahelical salt bridges of the g–c type, involving Glu270-Arg273 and Glu277-Arg280, contribute to the structural stability. Additionally, the CC domain harbors a predicted nuclear export signal (NES) spanning residues 295-301, which regulates cellular localization (Tiruttani Subhramanyam et al., 2017).

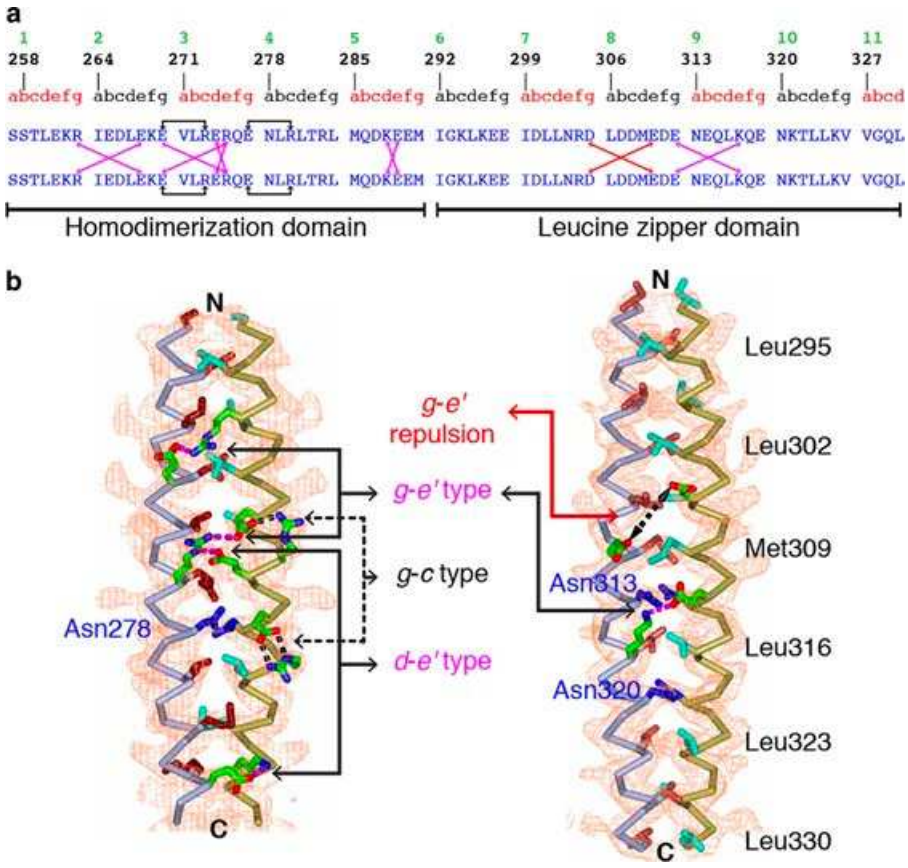


Figure 1.2.5.1 Homodimeric interactions of Par-4 CC (figure adapted from (Tiruttani Subhramanyam et al., 2017))

(a) The protein sequence of Par-4CC is aligned in a heptad register based on its structure, as determined by the SOCKET program. In the illustration, double-headed arrows are used to indicate salt bridges: black arrows for intrahelical salt bridges and magenta arrows for interhelical salt bridges. The red double-headed arrow illustrates charge repulsion between helices. Sequence numbering follows the UniProt ID: Q62627. The depicted region of Par-4CC corresponds to the domains responsible for homodimerization and leucine zippers. (b) The crystal structure of Par-4CC's homodimeric coiled-coil reveals two distinct domains: the homodimerization domain on the left and the leucine zipper domain on the right. Residues at the coiled-coil interface are highlighted with crimson for chain A (ice blue) and cyan for chain B (gold). The depiction includes interhelical and intrahelical salt bridges.

Regarding the SAC domain (residues 137-195), CD experiments have revealed the presence of residual structure, predominantly exhibiting a random coil profile. Notably, a SAC-T155E mutant, designed to mimic phosphorylation, displayed distinct properties compared to the native protein. CD and NMR spectroscopies demonstrated a shift towards the pre-molten globular state for this mutant, providing insights into the conformational dynamics of Par-4 (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, University of Duesseldorf, 2011).

1.3 Aim of the study

Understanding the structure of Par-4 is of paramount importance for several reasons. Firstly, it would shed light on the molecular mechanisms underlying its functional role in apoptosis regulation and its interaction with other partner proteins. By determining the precise arrangement of its domains, their spatial orientations, and potential inter-domain interactions, we can gain insights into how Par-4 participates in apoptotic signaling pathways. Furthermore, investigating the structure of Par-4 would enable the identification of key residues and regions that are critical for its interactions with other proteins. This knowledge could facilitate the design of targeted therapeutics that modulate these interactions and potentially provide novel approaches for cancer treatment.

Based on the aforementioned points, this project aimed to investigate and determine the structure of the region beyond the CC domain, based on the following hypotheses:

1. Phosphorylation of Ser249 and Thr248 could induce stability in the CC domain structure and help resolve the crystal structure, especially the missing parts of the short chains and the N-terminals of the chains.

2. Extending the N-terminal region of the CC domain to include the SAC domain could help stabilize the CC domain, SAC domain, and the linker region. Since this is the active form of Par-4 in inducing apoptosis, this structure is in itself functionally meaningful.

3. Binding the leucine zipper domain of Par-4 to its partner protein P62 PB1 domain could induce structural changes in the regulatory domain of Par-4, the CC domain, and affect regions outside the CC domain, either by structural induction or by binding to parts of the PB1 domain filaments apart from the CC domain.

The proteins will be subjected to biophysical characterization to assess their stability and folding, and, if possible, to proceed to a structural investigation that would allow understanding the function of Par-4 at a structural level.

2 Materials and methods

2.1 Instruments and materials

Table 2.1.1 Instruments used

Instruments	Company
Akta explorer 100	GE Healthcare Life Science
Autoflex maX	Bruker
AVIV model CD 425	AVIV
Benchtop centrifuge 5417R	Eppendorf
Benchtop centrifuge 1-14K	Sigma
Microfluidizer M-110P	Microfluidics International Corporation
Centrifuge 5804R	Eppendorf
ChemiDoc XRS gel system	Bio-Rad
H1001-M Incu-Shaker	Benchmark
Incubator shaker	INFORS Unitron
Maxisafe 2020	Thermo Scientific
Mini sub cell GT	Bio-Rad
Mini-PROTEAN tetra cell	Bio-Rad
Mosquito LCP robot	SPT Labtech
NanoDropTM 2000 Spectrophotometer	Thermo Scientific
T100 Thermal cycler	Bio-Rad
T80/T80+ UV-VIS spectrophotometer	PG Instruments Limited
Thermomixer R	Eppendorf
Trans-Blot Turbo Transfer System	Bio-Rad

Spectrafuge mini	Labnet
Spectrosize 300	Xtal concepts
Superspeed Centrifuge Sorvall LYNX 6000	Thermo Scientific

Table 2.1.2 Materials used

Materials	**Company**
Amicon Ultra centrifugal device	Millipore
Deoxyribonuclease-I from bovine pancreas	Applichem
ESF 921 Insect Cell Culture Medium, Protein Free	Expression systems
FastDigest restriction enzymes	Thermo Scientific, New England Biolabs
FastDigest DpnI	Thermo Scientific
GelRed Nucleic acid stain	Biotium Inc.
GeneJET Gel Extraction Kit	Thermo Scientific
GeneJET Plasmid Maxiprep Kit	Thermo Scientific
GeneJET Plasmid Miniprep Kit	Thermo Scientific
Gene Ruler DNA ladder	Thermo Scientific
lysozyme	Applichem
Nanosep MF and NAB Centrifugal Devices	Pall Laboratory
Phusion Hot Start II DNA Polymerase	Thermo Scientific
Polyethylenimine “Max” (PEI MAX) 40 kDa	Polysciences
PVDF membrane	GE Healthcare Life Science
Precision Plus Protein™ WesternC™ Blotting Standards	Bio-Rad

Precision Protein™ StrepTactin-HRP Conjugate	Bio-Rad
Pre-stained Protein Ladder	Jena Bioscience
Size-exclusion chromatography column	GE Healthcare Life Science
SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Scientific
Strep-Tactin®XT 4Flow® resin	IBA Lifesciences
T4 DNA ligase	Thermo Scientific
T4 Polynucleotide Kinase	Thermo Scientific

2.2 Biological materials

Table 2.2.1 Bacterial strains used

Strain	Genotype	Source	Purpose
Top10	F– mcrA Δ (mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ (ara leu) 7697 galU galK rpsL (StrR) endA1 nupG	Invitrogen	Plasmid amplification
BL21-CodonPlus (DE3)-RIL strain	E. coli B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [argU ileY leuW Camr]	Agilent Technologies	Protein expression

Table 2.2.2 Insect cell strain used

Strain	Species of origin	Source	Purpose
BTI-TN-5B1-4	Trichoplusia ni	Oxford expression technologies	Protein expression

Table 2.2.3 Expression vectors

Vector	Antibiotic resistance	Source	Expression system	DNA inserted
pET-11a	Ampicillin	Novagen	Bacteria	Par-4 CC EXS/CC domain
pOpiE-2	Ampicillin	Thermofisher scientific	Insect	Par-4 CC EXS
pOPTM	Ampicillin	ERC3/FZJ	Bacteria	P62 PB1 1-122 a a

Table 2.2.4 Media for bacterial culture

Medium	Compositions
Lysogeny Broth (LB)	Yeast extract (5 g/l) Tryptone (10 g/l) Sodium chloride (10 g/l)
DYT	Yeast extract (10 g/l) Tryptone (16 g/l) Sodium chloride (5 g/l)

All media were dissolved in Milli-Q water and autoclaved at 121°C for 20 minutes.

Table 2.2.5 Auto-induction medium

Component	Stock	Composition for 1 Liter
ZY (1 L)	10 g Tryptone 5 g Yeast Extract	928 ml
20 x NPS (1 L)	66 g (NH4)2SO4 136 g KH2PO4 142 g Na2HPO4	50 ml
50 x 5052 (1 L)	250 g Glycerol 25 g Glucose 100 g α-Lactose	20 ml
1 M MgSO4 (0.1 L)	12.04 g MgSO4	1 ml

All components stocks were dissolved in water and autoclaved separately.

2.3 Buffers

Table 2.3.1 Buffers for gel electrophoresis

Buffers	Compositions
Agarose gel running buffer (50X TAE)	2M Tris-base 1 M acetic acid 50 mM EDTA pH:8.0
SDS-Sample buffer (5X)	62 mM Tris, 2 % (w/v) SDS 5 % (v/v) β-Mercaptoethanol 20%(w/v) Glycerol 0.2 % (w/v) Bromophenol blue
SDS-PAGE running buffer	25 mM Tris 192 mM Glycine 0.1 % (w/v) SDS%
Blue silver staining solution	10% Phosphoric acid 10% Ammonium sulfate 1.2g/L Coomassie G-250 20% (v/v) Methanol
Transfer Buffer for Western blot	39 mM Glycine 48 mM Tris base 20% (v/v) Methanol
TBS buffer	20 mM Tris base pH 7.6 137 mM NaCl
TBS-T buffer	TBS buffer 0.1% (w/v) Tween 20
Blocking buffer	TBS-T buffer 5% (w/v) non-fat milk

Table 2.3.2 Sodium Dodecyl Sulphate–PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

Components	10 % resolving gel (20 ml)	12 % resolving gel (20 ml)	15 % resolving gel (20 ml)	5 % stacking gel (5 ml)
H_2O	7.9 ml	6.6 ml	4.6 ml	3.4 ml
30% Acrylamide	6.7 ml	8.0 ml	10.0 ml	0.83 ml
1.5 M Tris, pH 8.8	5.0 ml	5.0 ml	5.0 ml	-
1 M Tris, pH 6.8	-	-	-	0.63 ml
10 % SDS	0.2 ml	0.2 ml	0.2 ml	0.05 ml
10 % Ammonium Persulfate	0.2 ml	0.2 ml	0.2 ml	0.05 ml
TEMED	0.008 ml	0.008 ml	0.008 ml	0.005 ml

Table 2.3.3 Buffers for MBP-P62 PB1 protein purification

Buffers	Compositions
Lysis Buffer	50 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM TCEP, 1 complete protease inhibitor tablet per 50 mL lysis buffer.
Wash Buffer	50 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM TCEP.
Elution Buffer	50 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM TCEP, 10 mM maltose.
Filament formation buffer	50 mM HEPES pH 8.0, 50 mM NaCl, 0.5 mM TCEP**.**

Table 2.3.4 Buffers for Par-4 proteins purification

Buffers	Compositions
Lysis Buffer for Par-4 CC EXS	50 mM Tris pH 7.5, 300 mM NaCl, 10 mM $MgCl_2$, 0.5 mg/ml Lysozyme, 15 ug/ml DNase I, 1 uM E64, 1uM Pepstatin A, 1 uM Leupeptin, 1 mM AEBSF, 1 mM PMSF, 1 mM Benzamindine
Lysis Buffer for CC domain and its mutants	50 mM Tris pH 7.5, 200 mM $MgCl_2$, 0.5 mg/ml Lysozyme, 15 ug/ml DNase I, 1 uM E64, 1uM Pepstatin A, 1 uM Leupeptin, 1 mM AEBSF, 1 mM PMSF, 1 mM Benzamindine
Strep-Tactin XT resin Wash Buffer 1 (Equilibration Buffer) for CC EXS	50 mM Tris pH 7.5, 300 mM NaCl, 4 mM EDTA
Strep-Tactin XT resin Wash Buffer 2 (Equilibration Buffer) for CC domain	50 mM Tris pH 7.5, 200 mM $MgCl_2$
Buffer 1 for lowering pH to 4.5	50 mM KH_2PO pH 4.5, 1 M NaCl, 4 mM EDTA
Buffer 2 for lowering pH to 4.5	50 mM KH_2PO pH 4.5
Elution Buffer 1	50 mM KH_2PO pH 3.5
Elution Buffer 2	50 mM KH_2PO/ K_2HPO pH 7.5, 150 mM NaCl
SEC Buffer 1	50 mM KH_2PO pH 3.5
SEC Buffer 2	50 mM KH_2PO/ K_2HPO, pH 7.5, 150 mM NaCl

2.4 Cloning

Molecular cloning refers to the process of generating recombinant DNA molecules by combining specific DNA fragments or genes with suitable cloning vectors, typically plasmids. Cloning plays a crucial role in protein production, involving the isolation, modification, and insertion of a specific DNA sequence into a vector for expression (Old & Primrose, 1981). Polymerase Chain Reaction (PCR) is utilized with primers that incorporate tags and restriction sites. PCR amplifies the DNA fragment, which can be modified, digested with restriction enzymes, and ligated into a vector (Bej et al., 1991; Saiki, 1989). The ligated construct is then introduced into host cells through transformation. Following transformation, the presence and integrity of the cloned DNA sequence are confirmed through techniques such as PCR or DNA sequencing (Rong et al., 2002).

Site-directed mutagenesis is a widely employed technique in molecular biology for introducing precise alterations at specific positions within a target DNA sequence. It enables researchers to investigate the functional consequences of specific genetic changes by creating desired mutations, insertions, deletions, or substitutions in a gene of interest. The technique plays a pivotal role in studying structure-function relationships, protein engineering, drug development, and understanding disease-causing mutations (Kunkel, 1985; Zheng et al., 2004).

The Par-4 construct (made by Dr. Ulf Benno Eidhoff) as well as strep-tagged Par-4 CC construct (made by Dr. Udaya Kumar Tiruttani Subhramanyan) from our Lab stocks were used as templates in this project. The region of interest in the Par-4 construct (Appendix I: DNA and protein sequences) was amplified by directional cloning using gene specific primers that were designed to contain restriction sites of NdeI and BamHI in forward and reverse primers respectively for the protein to be expressed in E. coli, whereas, the primers used for the protein to be expressed in insect cells contain restriction sites of BamHI and AvrII in forward and reverse primers respectively (Appendix II Primers used). The forward primer was designed to introduce a Strep tag.

A Strep-tagged Par-4 CC construct (residues 240–332) was used as a template for creating phosphorylation-mimicking mutants as well as salt bridge mutants of the CC domain. The phosphorylation-mimicking mutants include serine at position 249 (numbering with respect to the Par-4 protein sequence), which was mutated to aspartic acid to create the S249D mutant and to glutamic acid to create the S249E mutant. Similarly, threonine at position 248 (T248) was mutated to create T248D and T248E mutants. Additionally, various charged amino acids responsible for stabilizing the CC domain by forming salt bridges were mutated to neutral amino acids. These mutations are E268Q, E274Q, K288M, K317M, and D305N.

PCR was performed in a thermal cycler with all the components needed and their respective concentrations. The primers were designed using the online primer design software from Thermo Fisher, and the primers of mutations were designed using the online software NEBaseChanger from New England BioLabs. The forward primer was designed to introduce a Strep tag. The primers were ordered from BioTeZ Berlin-Buch GmbH, Germany and integrated DNA Technologies (IDT), Leuven, Belgium and resuspended in ddH2O to get a final concentration of 50 µM.
The PCR reactions were performed as follow:

Table 2.4.1 PCR reaction

Components	20 µL reaction	Final concentration
5X Phusion HF Buffer	4 µL	1x
10 mM dNTPs	0.4 µL	200 µM
Forward primer	0.2 µL	0.5 µM
Reverse primer	0.2 µL	0.5 µM
Template DNA	1.5 µL	5 ng
DMSO	0.6 µL	3%
Phusion Hot Start II DNA Polymerase	0.2 µL	0.02 U/µL
H_2O	Up to 20 µL	

The PCR experiment was performed with the gradient annealing temperatures from 52 °C to 68 °C and the extension time of the amplicon was set to 15 seconds per 1 kb according to the manufacturer's protocol.

Table 2.4.2 PCR program

Cycle step	Temperature	Duration	No. of cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	30-35
Annealing	52-68 °C	30 s	
Extension	72 °C	15 s per kb	
Final Extension	72 °C	10 mins	1

The PCR product was verified using a 1.5% agarose gel electrophoresis (AGE) with a voltage of 120V for 45 mins. DNA bands were visualized by staining the gel with GelRed stain and were extracted using GeneJET Gel Extraction Kit from Thermo Fisher. The concentration of the purified PCR product was measured according to the absorption at 260 nm using NanoDrop spectrometer. The PCR product (insert) and the vector of interest (pET11a for bacterial expression, and pOPIE2 for insect cell expression) were digested by relative restriction enzymes at 37 °C for 30 mins. Proper controls were used to check the digestion by AGE. Subsequently the double digested products of the insert and the vector were extracted using Gel Extraction Kit. The purified insert and vector DNA were ligated using T4 DNA ligase from ThermoFisher. With ATP as a cofactor, the ligase catalyzes the formation of phosphodiester bonds between adjacent nucleotides, joining the fragments together (Tomkinson et al., 2006). A molar ratio of 1:3 was used for vector and insert using the rapid ligation kit under 22 °C for 30 mins.

In the site-directed mutagenesis experiment, the DpnI restriction enzyme that specifically digests methylated parental DNA, while leaving unmethylated newly synthesized DNA intact (Li et al., 1999), was incubated with the PCR product at 37 °C

for one hour to eliminate template DNA contamination. The target DNA was then analyzed using a 1% agarose gel and extracted using a Gel Extraction Kit. To facilitate subsequent ligation reactions, the linear DNA fragments underwent 5'-end phosphorylation using ThermoFisher T4 polynucleotide kinase, which transfers the gamma phosphate from ATP to the 5'-end of DNA or RNA (Cameron & Uhlenbeck, 1977). Following phosphorylation, ligation of the linear DNA fragments was performed using a T4 DNA ligase. This ligation reaction was carried out overnight at 16 °C, allowing the linear DNA fragments to circularize and form recombinant plasmids.

The above mentioned reactions are listed below:

Table 2.4.3 Restriction enzyme digestion reaction

Components	**50 µL reaction**
10X FastDigest Green Buffer	5 µl
DNA	1 µg
FastDigest enzyme I	1 µl
FastDigest enzyme II	1 µl
Nuclease-free water	Up to 50 µl

Table 2.4.4 DpnI digestion

Components	**40 µL reaction**
10X Reaction Buffer	4 µl
DNA	30 µl
Dpn I	1 µl
Nuclease-free water	Up to 40 µl

Table 2.4.5 DNA phosphorylation reaction

Components	50 µL reaction
10x reaction buffer A for T4 Polynucleotide Kinase	5 µl
DNA (from DpnI digestion)	39 µl
10mM ATP	4 µl
T4 Polynucleotide Kinase	2 µl

Table 2.4.6 Ligation reaction

Components	20 µL reaction
10X Ligase Reaction Buffer	2 µl
Linear vector DNA	100 ng
Insert DNA	1:3 molar ratio over vector
T4 DNA Ligase	0.5 µl
Nuclease-free water	Up to 20 µl

Ultra-competent cells (TOP10 strain) derived from Escherichia coli were incubated with the ligated product on ice for 15 mins followed by 90-second heat shock treatment at 42 °C. After the mixture was incubated on ice for 2 mins, 900 µl of LB medium was added to the mixture and incubated at 37 °C for 1 hour with 400 rpm shaking. After that, the mixture was centrifuged at 3000 xg for 2 mins and 900 µl supernatant was discarded and the cell pellet was resuspended in the remained 100 µl and plated on LB agar plate supplemented with 100 µg/ml carbenicillin (carb) and incubated at 37 °C overnight.

Several individual colonies were selected and inoculated into 5 ml of LB liquid medium. The cultures were then incubated overnight at 37 °C with shaking at 180

rpm. The bacterial cultures were harvested by centrifugation at room temperature under 6000 xg , and the plasmids were extracted using the GeneJET Plasmid Miniprep Kit. To confirm the presence of the desired plasmids, a double digestion with appropriate restriction enzymes was performed. Additionally, DNA sequencing was carried out by Microsynth company to verify the sequence integrity of the plasmids before utilizing them for protein expression.

2.5 In vivo protein overexpression

2.5.1 Bacterial expression of Par-4 CC domain and mutants

Recombinant protein production in Escherichia coli was achieved through the utilization of the pET expression system, which effectively harnesses the T7 RNA polymerase (Studier & Moffatt, 1986). This system exploits the unique recognition of the T7 phage promoter by the T7 RNA polymerase, excluding interaction with the E. coli RNA polymerase (Studier et al., 1990). Expression of the target genes is successfully initiated by introducing the T7 phage or by inducing strains harboring integrated copies of the polymerase gene, such as BL21DE3 (Studier & Moffatt, 1986). By employing this strategic approach, robust protein production is achieved, enabling comprehensive studies and characterization of the desired proteins. This system offers a powerful tool for efficient and controlled recombinant protein expression in E. coli (Studier & Moffatt, 1986; Studier et al., 1990).

The expression of CC domain mutants was carried out based on the published protocol (Tiruttani Subhramanyam et al., 2014). The transformant colonies of CC domain mutants of RIL strain of Escherichia coli BL21(DE3) were selected on an LB–agar plate with 100µg/ml carbenicillin (carb). Transformant colonies were subsequently inoculated into 50 ml LB medium supplemented with 100 µg/ml (carb) and 2% glucose. The pre-culture was incubated overnight at 37 °C with 180 rpm in a shaker incubator to allow for primary culture growth. A required volume of primary culture to get a percentage of 1-1.5% was inoculated into a DYT medium main culture, and was allowed to grow in a shaker incubator at 37 °C at 120 rpm. The expression was

induced with 1 mM Isopropyl-ß-D-1-thiogalactopyranoside (IPTG) when OD600 reached 0.6 -0.8 and incubated at 37°C with 120 rpm for 4 hours.

The bacterial culture was harvested by centrifugation at 4500 xg for 20 minutes at 4 °C, and the resulting cell pellet was subsequently stored at -80 °C to maintain its integrity until the purification process was initiated.

2.5.2 Bacterial expression of Par-4 CC EXS

The plasmid encoding the desired protein was introduced into ultra-competent cells of E. coli BL21(DE3) strains, specifically RIL strain, using a 90-second heat shock at 42 °C. The transformed cells were then selected by plating on lysogeny broth (LB) agar containing of 100 µg/ml carbenicillin (carb). Transformant colonies were subsequently inoculated into 50 ml LB medium supplemented with 100 µg/ml (carb) and 2% glucose to control the basal expression of pET system (Zhang et al., 2003). The pre-culture was incubated overnight at 37 °C with 180 rpm in a shaker incubator to allow for primary culture growth. A required volume of primary culture to get a percentage of 1-1.5% was inoculated into a DYT medium main culture, and was allowed to grow in a shaker incubator at 37 °C at 120 rpm. A 1:5 volume ratio of the culture volume to the flask volume was maintained to ensure better aeration (Rodríguez-Aparicio et al., 1988). The expression was induced with 0.6 mM IPTG when OD_{600} reached 0.8 -0.9 and incubated at 16 °C with 120 rpm for 18 hours.
The bacterial culture was harvested by centrifugation at 4500 xg for 20 minutes at 4 °C, and the resulting cell pellet was subsequently stored at -80 °C to maintain its integrity until the purification process was initiated.

2.5.3 Insect cell expression of Par-4 CC EXS

Protein expression in insect cells presents numerous advantages, making it an attractive choice for recombinant protein production. Insect cell lines offer a wide range of posttranslational modifications, including phosphorylation of serine (Héricourt et al., 2000), tyrosine sulfation (Thompson et al., 2017), Glycosylation and palmitoylation (Grünewald et al., 1996), which are crucial for the functional and

structural integrity of proteins. Moreover, insect cells are easier to handle compared to mammalian cells, simplifying experimental procedures. Another significant benefit is that insect cells can be cultivated at 27°C in serum-free media without the need for additional CO2 aeration, resulting in reduced costs and minimizing the risk of contamination with human pathogens (Stacey & Possee, 1996). These advantages make the insect cell expression system a versatile and cost-effective platform for efficient recombinant protein production.

The virus-free plasmid DNA-based transient gene expression system (TGE) was used for insect cell transfection. This system involves direct delivery of plasmid DNA into insect cells, enabling transient gene expression and protein production. Polyethyleneimine (PEI) is a commonly employed transfection agent known for its high efficiency in insect cell transfection (Boussif et al., 1995). PEI forms stable complexes with DNA, facilitating their uptake into host cells. The TGE approach offers simplicity, versatility, and scalability for various biotechnological applications (Carinhas et al., 2011; Kost et al., 2005).

For the transient transfection, the pOpiE2 plasmid was used. Notably, this vector encompasses not only the immediate early promoter OpiE2 but also includes the highly efficient IE1 terminator and enhancing sequences (Bleckmann et al., 2016). This comprehensive design ensures optimal gene expression control and robust protein production in High Five cells.

The BTI-Tn-5B1-4 cell line, derived from Trichoplusia ni (cabbage looper), commonly known as Hi5 cells, was used for the protein production. Frozen Hi5 cells were thawed and resuspended in 10 ml of ESF 921 serum-free media and divided into two T25 flasks and incubated at 28 °C for 1 hour after viability assessment. When living cells adhered (usually took 1 h), the media was changed with fresh one to remove DMSO toxicity and the cells were incubated until they have reached >80% confluency before sub-culturing. Cells were passaged 2-3 times until the viability reached above 90 % to start the suspension culture. Viability was determined by counting cells with a hemocytometer after combining equivalent volumes of trypan blue solution (Cadena-Herrera et al., 2015).

Suspension culture was prepared by adding the cells from the T25 flasks into fresh media in Erlenmeyer flask to get a final cell density of 0.8-1×10^6 cells/mL and incubated in a shaker at 28 °C and 120 rpm. A 1:4 volume ratio of the culture volume to the flask volume was maintained to ensure better aeration. When the cell density reached 4-6×10^6 cells/ml, a new passage was made to maintain the cultures in the exponential growth phase. Cells were passaged 2-3 times per week.

For insect cell transfection, cell cultures with a density of 5-7×10^6 cells/mL were centrifuged at 180 x g at room temperature (RT) for 10 mins, and the cells were resuspended in fresh media to have a density of 4×10^6 cells/ml. A 40 kDa PEI (polyethyleneimine) from Polysciences was prepared following the manufacturer's protocol, resulting in a stock solution concentration of 1 mg/ml. The DNA and PEI were added separately to the culture with the ratios 1µg DNA /1×10^6 cells and 1 µg DNA: 4 µg PEI.

Cells were incubated in a shaker at 28 °C and 120 rpm. Test samples of 1 ml were collected at different times from the start of expression and were analyzed by poly-acrylamide gel electrophoresis (PAGE) for protein expression. Cells were harvested when the viability percentage was about 50 % which usually took two days post-transfection. Centrifugation at 4000 x g for 15 min at 4 °C was used to harvest the cultures. Cell pellets were stored at -80 °C until purification process was initiated.

2.5.4 Bacterial expression of P62 PB1

The expression of P62 PB1 was done according to the published protocol (Tarafder et al., 2019). The plasmid encoding the P62 PB1 protein was introduced into ultra-competent E. coli BL21(DE3) RIL cells using a 90-second heat shock at 42 °C. The transformed cells were then plated on lysogeny broth (LB) agar containing 100 µg/ml carbenicillin (carb) for selection. Selected colonies were inoculated into 50-150 ml of LB medium with 100 µg/ml carb and incubated overnight at 37 °C with shaking at 200 rpm to allow for primary culture growth. For the main culture, 25 ml of the saturated overnight culture was added to each 1 L of ZY autoinduction medium containing 100 µg/ml carb in a 2 L baffled flask. The cells were incubated at 37 °C for 6 hours with shaking at 200 rpm. After this period, the cells were allowed to rest at room

temperature while the shaking incubator was cooled to 20 °C. Once the incubator reached 20 °C, the cells were returned and incubated overnight at 20 °C with shaking at 200 rpm. The bacterial culture was then harvested by centrifugation at 4500 xg for 20 minutes at 4 °C, and the resulting cell pellet was stored at -80 °C until the purification process commenced.

2.6 Protein purification

All purifications described in this thesis were performed in the cold room (5- 10°C) and protein samples were always handled on ice.

2.6.1 Affinity chromatography

2.6.1.1 Affinity chromatography of Par-4 proteins

Affinity chromatography based on the Strep tag purification system was performed as an initial major purification step for the protein. The strep-tag is a short peptide sequence consisting of eight amino acids (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) that exhibits high affinity for streptavidin or Strep-Tactin resin. This tag can be genetically fused to the target protein, allowing for specific and efficient purification (Schmidt & Skerra, 1994). The strep-tagged protein selectively binds to the resin via the strong interaction between the strep-tag and Strep-Tactin. After loading the lysate, the column is washed to remove non-specifically bound contaminants. Elution of the Strep-tagged protein can be accomplished using either a competitive elution strategy with biotin which has a higher affinity for Strep-Tactin than the strep-tag or by lowering the pH to disrupt the tetramer structure of Strep-Tactin and facilitate protein elution.

Cell pellets were re-suspended in lysis buffer with the ratio of 10 ml per gram cell pellet combined with the protease inhibitor cocktail, deoxyribonuclease-I (15 µg/ml) and lysozyme (0.5 mg/ml). The re-suspension was incubated in cold-room for 1 hour with 200 rpm speed stirring. The re-suspension was passed 4-5 times through microfluidizer with a pressure between 15,000- 20,000 psi. The lysate was clarified by centrifuging at 57000 xg for 45 min at 4 °C. The supernatant was added to Strep-

Tactin XT resin pre-equilibrated with at least 5 bed volumes of the wash buffer and incubated for 2 hours at 4 °C with end-over-end rotation. After loading the resin-protein mixture to an Econo-Pac chromatography column, the flow-through was collected with slow flow rate. The column was then washed with 4-5 bed volumes of wash buffer to remove unbound proteins. Protein was eluted either by lowering pH to 3.50 or by using 50 mM biotin at pH 7.5. A fractions of 2 ml elution were collected and protein concentration was measured. Samples were collected from each step of the purification for later purification assessment.

2.6.1.2 Affinity chromatography of P62 PB1 domain

The purification of PB1 was done based on the published protocol (Tarafder et al., 2019).

The cell pellets were thawed by incubating the tubes in water at RT, and then resuspended thoroughly with ice-cold MBP lysis buffer (10 mL/g of cell pellet). The re-suspension was incubated in cold-room for 1 hour with 200 rpm speed stirring. The re-suspension was passed 4-5 times through microfluidizer with a pressure between 15,000- 20,000 psi.

The supernatant from the centrifugation step at 57,000 g was incubated with pre-equilibrated amylose resin using 5 column volumes (CV) of wash buffer for 1 hour at 4°C with end-over-end rotation. The lysate/resin mixture was then transferred to an Econo-Pac chromatography column, allowing the resin to retain the target protein as the lysate flowed through. The column was washed with 40 CV of wash buffer to remove unbound proteins. Elution was performed using an elution buffer containing 10 mM maltose, with fractions collected at 1 CV intervals. Elution was monitored using a 96-well plate, where 60 µL of Bio-Rad protein assay dye reagent was added to each well, followed by 5 µL of each elution fraction. The formation of a blue color indicated the presence of protein. The appropriate fractions were pooled and dialyzed against 2 L of MBP wash buffer overnight at 4°C with gentle stirring, using dialysis membranes with a molecular weight cutoff (MWCO) below 30 kDa. After dialysis, the sample was recovered, and the protein concentration was measured using a NanoDrop. The

protein was aliquoted, snap-frozen in liquid nitrogen, and stored at −80°C until further processing.

To form the P62 filaments, the MBP tag should be cleaved and NaCl salt should be lowered. This process involves cleavage with TEV, followed by dialysis using a 100 kDa MWCO membrane. The high molecular weight cut-off of the membrane permits the elimination of cleaved MBP, protease, and monomeric p62, while retaining p62 filaments that have a significantly greater molecular weight than 100 kDa.

2.6.2 Size exclusion chromatography

Size exclusion chromatography (SEC) was performed as a final purification step for the protein in this study. SEC, a prevalent technique in molecular separation, is based on the principle of selectively filtering molecules according to their size (Silberring et al., 2004). This method relies on a gel matrix composed of spherical beads that possess specific pore size distributions. The intricate interplay of inclusion and exclusion of molecules within these pores drives the separation process. Smaller molecules effectively diffuse into the pores, thereby experiencing hindered flow through the chromatographic column based on their respective sizes. Conversely, larger molecules are incapable of permeating the pore structure and are eluted in the void volume of the column. Consequently, molecules are efficiently resolved based on their size as they traverse the chromatographic column, ultimately eluting in a descending order relative to their molecular weight (MW) (Brusotti et al., 2018). In this study, the Superdex 200 Increase 10/300 GL and HiLoad 16/600 Superdex 75 pg size exclusion chromatography (SEC) columns were utilized for precise molecular weight fractionation. Prior to loading onto the SEC, the protein sample underwent filtration through a 0.22 μm filter to ensure sample purity from aggregation.

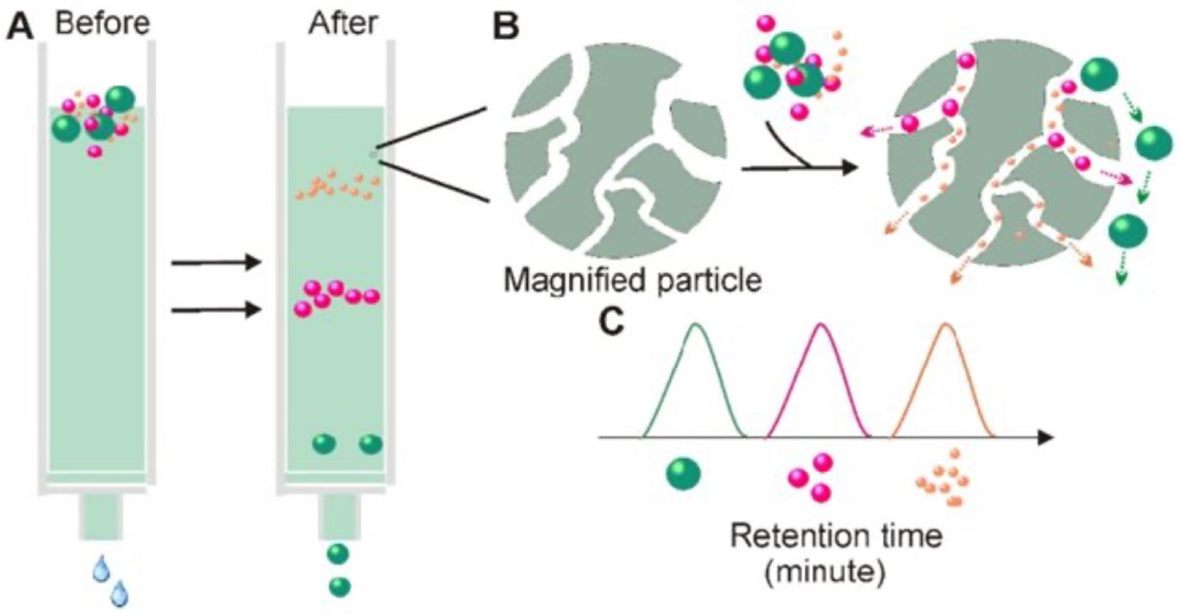


Figure 2.6.2.1 Principle of Size-exclusion chromatography (picture adapted from (Yang et al., 2020)).

When a solution passes through a stationary phase consisting of porous resin particles, molecules with hydrodynamic radii smaller than that of the pores of the stationary phase enter into the pores for longer traffic distance, whereas larger molecules move directly around the resin since they cannot enter the pores. This difference in travel distance results in different retention times for the molecules, which allows them to be separated based on size.

Affinity elution fractions were pooled and concentrated using Amicon® Ultra centrifugal device with a proper molecular weight cutoff membrane at 2700 xg speed. Every 10 min the sample in the concentrator was mixed thoroughly with a pipette. Immediately, concentrated sample was filtered through a 0.2 µm centrifugal filter at 20.000 xg and injected to a 500 µL sample loop onto the SEC pre-equilibrated with at least 3 bed volumes of the SEC buffer. The SEC was conducted under a flow rate of 0.4 mL/min, a pressure limit of 2.5 mPa and an elution volume of 1 column volume with 500 µL for each fraction. Samples from SEC elution fractions were collected for analysis. Protein samples were used for subsequent biophysical analysis and/or crystallization experiments.

2.7 Protein identification and Characterization

2.7.1 Protein estimation and concentration

Protein concentration was determined using UV-VIS or Nanodrop Spectrophotometers, measuring the absorbance within a wavelength range of 350 nm to 200 nm. The absorption at 280 nm, primarily attributed to Tryptophan and Tyrosine residues (Edelhoch, 1967), was utilized for quantification, employing a theoretical extinction coefficient based on the amino acid composition (Gill & Von Hippel, 1989). The theoretical molar extinction coefficient of the protein was calculated using the online Expasy ProtParam tool (https://web.expasy.org/protparam/) (Gasteiger et al., 2005).

Considering the issue of DNA contamination that was reported earlier (Dr. Jan Kubicek's PhD thesis, university of duesseldorf, 2005 and Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, university of duesseldorf, 2011), the A_{260}/A_{280} ratio was monitored to check the DNA contamination and verify the protein free of DNA (Warburg, 1942).

The protein concentration was calculated using the Beer-Lambert law (Gill & Von Hippel, 1989) : $A = \varepsilon lc$, where A represents the absorbance, ε is the molar extinction coefficient, l denotes the path length of light in centimeters, and c signifies the concentration of the solution.

2.7.2 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a widely employed technique for the separation of proteins based on their molecular weight (Laemmli, 1970; Towbin et al., 1979). To initiate the process, protein samples collected during the purification and SEC elution fractions were mixed with 4X SDS-loading buffer and incubated at 95 °C for 10 minutes. This step ensures denaturation of the proteins and allows sodium dodecyl sulfate (SDS) to bind to the positively charged and hydrophobic residues, facilitating uniform migration. The prepared samples were loaded onto a 15% polyacrylamide gel and subjected to electrophoresis

at 80V voltage for the stacking gel and 120V voltage for the resolving gel, until the front dye migrated out of the gel. To determine protein sizes, a prestained protein marker was included as a reference. Following electrophoresis, the gels were stained with a blue silver staining solution and incubated overnight at room temperature with gentle shaking. To remove background staining, the gels were destained through multiple washes with Milli Q water.

2.7.3 Western Blotting

Western blotting, a widely utilized analytical technique in protein analysis, enables the detection and characterization of specific proteins within complex biological samples (Moritz, 2017). Following the separation of proteins through SDS-PAGE, the transfer of proteins from the gel onto polyvinylidene fluoride (PVDF) membranes was accomplished using the Trans-Blot Turbo Transfer System (Bio-Rad) via electrophoresis. To minimize non-specific binding, the membranes were incubated in a blocking solution composed of 5% non-fat dry milk dissolved in TBS-T buffer at 4 °C overnight with gentle shaking. The membranes were subsequently washed two times with TBS-T buffer at room temperature for 10 minutes per round. Precision Protein StrepTactin-horseradish peroxidase (HRP) Conjugate antibody, diluted in TBS-T buffer, was then added to the membranes and incubated for 2 hours at room temperature. Following antibody incubation, the membranes were washed twice with TBS-T buffer to remove unbound antibodies, followed by an additional washing step using TBS buffer for at least 30 minutes to remove the residual TBS-T buffer. Notably, Strep-tag detection can be achieved without the need for secondary antibodies. To facilitate detection, the enzyme horseradish peroxidase catalyzed the oxidation of luminol in the presence of H_2O_2 (Kricka & Ji, 1995), and the resulting chemiluminescent signal was captured using the ChemiDoc MP Imaging System.

2.7.4 De-glycosylation Assay

One of the main advantages of expressing the protein in eukaryotic cells like insect cells and mammalian cells is that the high possibility of post-translation modifications (PTMs) such as glycosylation and phosphorylation that may induce folding and

stability for the protein. Purified mutated Par-4 CC EXS from Hi5 cells was analyzed for PTMs by enzymatic removal of the glycosylation moiety and mass spectroscopy.

Peptide-N-glycosidase F (PNGase F) is an enzyme that cleaves between the innermost GlcNAc and asparagine residues of high mannose, hybrid, and complex oligo-saccharides from N-linked glycoproteins, converting the asparagine residue to aspartic acid and liberating ammonia and the intact glycan. It is thus correctly termed an amidase, or amido-hydrolase, related in activity to the lysosomal glycosyl-asparaginases (Norris et al., 1994; Plummer Jr et al., 1984). Endoglycosidase H (endo H) cleaves N-linked glycans between the two N-acetylglucosamine (GlcNAc) residues in the core region of the glycan chain on high-mannose and hybrid, but not complex, glycans (Freeze & Kranz, 2008).

PNGase F and Endo H enzymes were purchased from New England Biolabs and Promega respectively. The reactions were done following the manufacturers protocols and samples were further analyzed by blue silver stained SDS-PAGE and anti-strep tag western blot.

2.7.5 Pull-down assay

A pull-down assay is an in vitro method used to detect or confirm protein-protein interactions. It is similar to co-immunoprecipitation but uses a purified, tagged protein (bait) to capture interacting proteins (prey) rather than immobilized antibodies. The bait is immobilized on an affinity ligand specific to its tag, and then incubated with prey proteins. The interacting complexes are then eluted using an eluting buffer depending on the affinity ligand. Proper controls, such as a positive control with only the bait protein and negative controls to detect nonspecific binding, are essential. The resulting protein fractions are analyzed by SDS-PAGE and visualized through gel staining or western blotting (Louche et al., 2017).

The pull-down assay was done to study the interaction between the Par-4 CC EXS and its partner protein P62 PB1 domain, and to get the proteins as a complex, which will be used for further biophysical characterization.

In this study, Par-4 protein was incubated with pre-equilibrated Strep-Tactin XT resin for one hour at 4 °C. Subsequently, MBP-P62 PB1 protein was added to the Par-4-Strep-Tactin resin complex and incubated overnight at 4 °C. After loading the resin-protein mixture onto an Econo-Pac chromatography column, the flow-through was collected, and the column was washed with 4-5 bed volumes of wash buffer to remove unbound proteins. The bound proteins were eluted using 50 mM biotin at pH 7.5, with 1 ml fractions collected. Samples were taken at each purification step. The pull-down assay results were analyzed by SDS-PAGE and visualized through gel staining and western blotting.

2.7.6 Mass Spectrometry

Mass spectrometry (MS) is a powerful analytical technique that measures the mass-to-charge ratio (m/z) of ions to identify and characterize various compounds in a sample. This technique has garnered considerable attention due to its sensitivity, accuracy, and ability to handle complex mixtures, making it an indispensable tool in modern research (Aebersold & Mann, 2003; Karas & Hillenkamp, 1988). One of the remarkable applications of mass spectrometry is Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) mass spectrometry. MALDI-TOF operates on the principle of ionizing analyte molecules by co-crystallizing them with a matrix compound, followed by laser irradiation to promote ionization (Tanaka et al., 1988). The ionized molecules are then accelerated in a vacuum tube based on their m/z ratios and arrive at a detector where the time-of-flight is measured. This time-of-flight information allows for the precise determination of mass and consequently identification of compounds (Hillenkamp & Karas, 1990). MALDI-TOF is utilized for its versatility and speed in analyzing a wide range of biomolecules, such as proteins, peptides, nucleic acids, lipids, carbohydrates, and synthetic polymers. This technique has revolutionized the field of proteomics, enabling rapid protein identification and profiling in biological samples (Karas et al., 2000).

In this study, a purified protein sample was analyzed at EMBL Hamburg's Protein characterization facility using a Bruker Autoflex™ Speed MALDI/TOF System. A filtered protein sample with 40uM concentration in addition to Bruker protein standard II (20 – 66 kDa) were prepared by combining them with 0.1% trifluoroacetic

acid-D (TFA) and supersaturated dihydroxy acetophenone (DHAP) matrix using the volume ratio 1: 1 :1, and 0.5 µL from each mixture was then spotted in duplicates onto MALDI target plate. After spotting, the samples were allowed to dry before being measured using the AutoflexÒ maX instrument (Bruker LabScape – Daltonics). Data acquisition was performed using the flexControl software (Bruker), and subsequent data analysis was carried out using the flexAnalysis software (Bruker).

Another notable application of mass spectrometry is Liquid Chromatography/Mass Spectrometry (LC/MS). LC/MS combines the separating power of liquid chromatography with the selective and sensitive detection capabilities of mass spectrometry. In LC/MS, analyte molecules are separated based on their interactions with a stationary phase in the chromatography column, and subsequently, eluted compounds are ionized in the mass spectrometer for mass analysis (John et al., 2004).

To determine the protein sequence and possible post-translation modifications, the protein bands were excised from the gel and in-gel trypsin digestion and in-gel acid hydrolysis were carried out to generate peptide fragments for identification. Subsequently, LC-MS/MS analysis was performed using either an Agilent 1290 Infinity HPLC system for LC, and Agilent 6550 iFunnel Q-TOF system for MS or LC-MS/MS on a Dionex Ultimate 3000 HPLC coupled to an Orbitrap Fusion Lumos System.

2.7.7 Dynamic Light Scattering

Dynamic Light Scattering (DLS) stands as an invaluable technique for probing the size distribution and diffusion characteristics of particles or molecules within solution. By observing the intricate fluctuations in scattered light intensity induced by Brownian motion, DLS provides crucial insights into particle sizes, aggregation states, and molecular interactions. This rapid and non-invasive method has found extensive application, particularly in the investigation of biomolecules like proteins and nanoparticles, across diverse scientific domains (Leipertz & Fröba, 2005).

The interpretation of DLS data depends on the Stokes-Einstein equation, a fundamental relationship that links the diffusion coefficient (D), the hydrodynamic radius (R), and the temperature (T):

$$D = \frac{k_B T}{6\pi\eta R_h}$$

In this equation, kB denotes the Boltzmann constant, T signifies the absolute temperature, η represents the solvent viscosity, and Rh stands for the hydrodynamic radius of the particle under study. The equation underscores that the diffusion coefficient is inversely proportional to the hydrodynamic radius. This signifies that particles with larger hydrodynamic radii exhibit slower diffusion rates compared to their smaller counterparts.

In the field of protein research using DLS, the correlation curve is obtained from the measured intensity autocorrelation function. This curve provides information about the rate of intensity fluctuations, which correlates directly with the protein's diffusion coefficient. The incorporation of the Stokes-Einstein equation, along with temperature and solvent viscosity, enables the determination of the protein's hydrodynamic radius. By leveraging this equation in the interpretation of DLS data, a precise understanding of protein sizes will be gained, thereby unraveling vital insights into their behavior and interactions within the solution environment.

DLS measurements were performed on the Spectrosize 300 (Xtal concepts). All the measurements were performed at 4 °C.

2.7.8 Circular Dichroism

Circular dichroism (CD) spectroscopy offers a rapid and effective means of assessing the secondary structure, folding dynamics, and binding characteristics of proteins (Kelly & Price, 2000). Essentially, CD spectroscopy quantifies the differential absorption of left-handed (L) and right-handed (R) circularly polarized light, providing valuable insights into the chiroptical properties of proteins (Fasman, 2013). When the absorption of left-handed (L) and right-handed (R) circularly polarized light

differs, the transmitted light will exhibit elliptical polarization that can be measured by CD spectrometer as a function of the wavelength of incident light, generating a CD spectrum (Kim et al., 2006). Typically, the far-UV CD (190-250 nm) is used to estimate the secondary structure content of a protein as different regular secondary structures found in proteins show distinct spectra in this region. Among these spectra, α-helical proteins spectra that have negative bands at 222 nm and 208 nm, along with a positive band at 193 nm, antiparallel β-pleated sheets (β-helices) spectra that display negative bands at 218 nm and positive bands at 195 nm, and disordered proteins spectra that give a minimal ellipticity above 210 nm and negative bands near 195 nm (Micsonai et al., 2015).

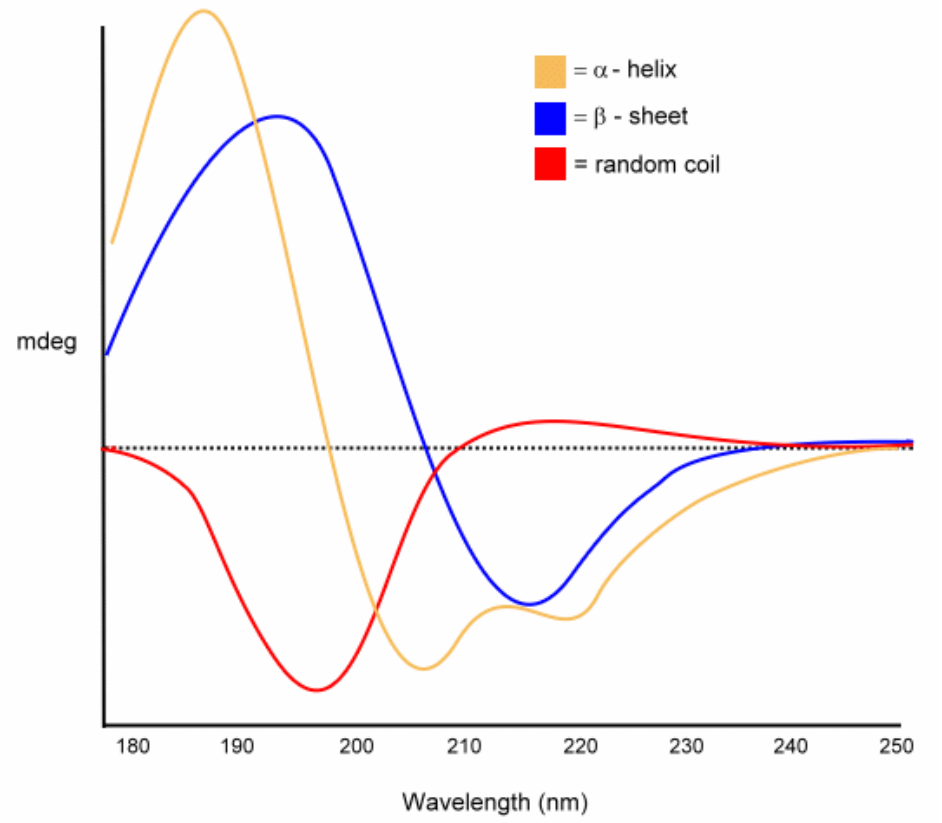


Figure 2.7.8.1 Different protein secondary structure by CD (adapted from (Dodero et al., 2011))

α-helical proteins typically exhibit characteristic negative peaks at 222 nm and 208 nm, along with positive peaks at 193 nm. Conversely, proteins featuring distinct anti-parallel β-pleated sheets display negative peak at 218 nm and positive peak at 195 nm. In contrast, disordered proteins tend to show low ellipticity readings beyond 210 nm and negative peak around 195 nm.

These different spectra can be analyzed by normalizing to the mean residual ellipticity (MRE) using the following formula

$$\mathrm{MRE} = \frac{\theta * \mathrm{MRW}}{10 * c * d}$$

where θ is the measured ellipticity (degrees), MRW is the mean residual weight of the protein (molecular weight / No. of amino acids - 1), d is the path length (cm) and c is concentration of the protein (g/ml).

Zeba Spin Desalting Column for buffer exchange was used upon need. CD spectra for the protein samples were measured using Aviv CD425 spectrometer at 4 °C, and temperature scans were performed from 4 °C to 98 °C at a heating rate of 2 °C per round with an averaging time of 9 seconds. As the accurate concentration of the protein is essential for a CD spectra deconvolution, BCA assay was used to determine the protein concentrations. All data were collected using a 0.1 cm path length quartz cuvette. The buffer baselines were subtracted from the sample spectra. The deconvolution of the CD spectra were calculated using BeStSel (Beta Structure Selection); an online tool used for the determination of secondary structure and recognition of protein folds from circular dichroism spectra (Micsonai et al., 2018).

The melting temperature of the proteins were determined through the temperature scanning from 4°C to 98° C with an increment of 2°C. The buffer CD signal was subtracted from each temperature scan. After that, the different temperature scans spectra were plotted against the wavelength range 190-260 nm. Depending on the dominant secondary structure exists in the protein, the effect of temperature gradient on the CD intensity can be assessed by plotting the intensity at 222 nm or 217 nm against the different temperature scans. The linear shifts in CD between folded and unfolded proteins was adjusted prior to plotting the data as a function of temperature. Based on the proposed equation for nonlinearly fit CD data by Santoro and Bolen (Santoro & Bolen, 1988), the calculations were made. Equation 1 was rewritten based on their formula as follows:

$$y = \frac{(a_N + b_N x) + (a_D + b_D x) e^{E\left(\frac{1}{x0} - \frac{1}{x}\right)}}{1 + e^{E\left(\frac{1}{x0} - \frac{1}{x}\right)}}$$

In the context of this theoretical framework, the linear function characterizing the original, intact state before any unfolding is donated as (aN + bNx), while the state after denaturation is symbolized as (aD + bDx). The significant concept highlighted by Santoro and Bolen is that the variable E embodies the unfolding process's free energy and corresponds to the extrapolated interception point of the unfolding phenomenon. When the second derivative of the fitted graph is set equal to zero, the resulting solution provides the precise x-value at which the slope of the original curve reaches its maximum, effectively indicating the protein's melting temperature.

In addition, Investigating the fluctuations in circular dichroism (CD) concerning temperature, especially at specific wavelengths, introduces a distinctive method for unveiling the thermodynamics associated with a change in state, specifically unfolding. This encompasses deducing the van't Hoff enthalpy (ΔH) and entropy (ΔS) related to unfolding, pinpointing the midpoint of the unfolding transition (Tm), and determining the free energy (ΔG) of unfolding. Furthermore, examining CD spectra collected with varying temperatures can be instrumental in discerning the presence of intermediate stages in the unfolding or folding process of a protein.

When a protein's CD undergoes temperature-dependent changes, this transformation can be utilized to analyze the thermodynamics associated with its unfolding or folding. In the most straightforward scenario, a molecule experiences an unfolding transition between two states: the folded state (N) and the unfolded state (D).

The value of y (which can be the ellipticity signal) at any temperature is the sum contributed by the native state and from the denatured states, which are both present in different fractional concentrations from 0 – 1, which can be illustrated as

$$y=(fNyN)+(fDyD).$$

where fN is the fraction native and yN is the contribution to the dependent variable y from the native state, and $fD$is the fraction denatured and yD is the

contribution to the dependent variable y from the denatured state. Conservation gives the following equation.

$$1=f_N+f_D \text{ or } f_N=1-f_D$$

Substituting (2) into (1) gives

$$y=(1-f_D)\, yN+(f_D y_D)=y_N-f_D y_N+(f_D y_D)$$

Rearranging this equation gives

$$\mathrm{f}D = \frac{\mathrm{y} - yN}{\mathrm{y}D - yN}$$

By substituting (2) and (3) into the expression for the equilibrium constant for the reaction N⟵⟶DN⟵⟶D, this gives:

$$\mathrm{Keq} = \frac{[\mathrm{D}]\mathrm{eq}}{[\mathrm{N}]\mathrm{eq}} = \frac{\mathrm{fD}}{\mathrm{fN}} = \frac{\mathrm{fD}}{1 - \mathrm{fD}}$$

From this, ΔG_0 can be calculated as below

$$\Delta G_0 = -RT\ln Keq = -RT\ln\left(\frac{\mathrm{fD}}{1-\mathrm{fD}}\right).$$

Where R is the Gas constant =8.314 J/mol, and T is the absolute temperature (Kelvin).

Knowing Keq, ΔH_0 and ΔS_0 can be calculated from Gibbs free energy equation as below since a semi-log plot of $\ln K_{eq}$ vs 1/T is a straight line with a slope of $-\Delta H_0 R$ and a y-intercept of $\Delta S_0/R$.

$$\Delta G_0 = \Delta H_0 - T\Delta S_0 = -RT\ln K_{eq}$$

$$\ln K_{eq} = -\frac{\Delta H_0 - T\Delta S_0}{RT}$$

$$\ln K_{eq} = -\frac{\Delta H_0}{RT} + \frac{\Delta S_0}{R}$$

At the melting temperature (Tm), both the folded fraction and denatured fraction are existed equally, Keq = 1 and the $\Delta G_0 = 0$. At this temperature,

$$\Delta H_0 = T\Delta S_0$$

2.9 X-ray crystallography

X-ray crystallography is a widely utilized technique in protein structural biology, enabling the determination of three-dimensional atomic structures of proteins (Mertens & Svergun, 2010). The process of X-ray crystallography begins with the growth of high-quality single crystals of the protein of interest through a crystallization process. These protein crystals are then exposed to a beam of X-rays. As X-rays interact with the electron clouds of atoms within the crystal lattice, they scatter in characteristic patterns. The resulting diffraction pattern, recorded on a detector, contains valuable information about the arrangement of atoms within the protein(Drenth, 2007).

Despite the challenge posed by the crystallographic phase problem, methods such as multiple isomorphous replacement and molecular replacement allow researchers to construct accurate atomic models(Guex et al., 2009). X-ray crystallography has significantly advanced our understanding of protein structures, providing crucial insights into active sites, binding pockets, and overall folding patterns, which is essential for deciphering protein functions and guiding drug discovery efforts(Shi, 2014).

The process of crystallization, encompassing various chemical species like proteins, involves two intimately connected yet distinct phases: nucleation and growth. Nucleation presents the most challenging aspect both theoretically and experimentally as it constitutes a first-order phase transition, orchestrating the transformation of molecules from a completely disordered state to an organized one.

This transition is thought to occur through the development of partially ordered or paracrystalline intermediates, exemplified by protein aggregates with short-range order. Eventually, this process leads to the emergence of small, highly ordered assemblies known as critical nuclei (McPherson & Gavira, 2014). These Critical nuclei must be evaluated considering molecular dimensions, supersaturation, and surface free energy of molecular addition(McPherson & Gavira, 2014). Although critical nuclear size has been elucidated in only a few systems, and in certain cases, it has been explored solely for two-dimensional nuclei developing on the surfaces of pre-existing crystals (Malkin et al., 1996). In concentrated protein solutions, the formation of a distinctive 'liquid protein phase' leads to critical nuclei exhibiting comprehensive order(Haas & Drenth, 1999; Piazza, 1999).

In contrast, the growth of macromolecular crystals is better understood and characterized compared to nucleation. The mechanisms of macromolecular crystal growth, such as dislocation growth, two-dimensional nucleation, normal growth, and three-dimensional nucleation, have been reasonably well-established (Malkin et al., 1995; McPherson et al., 2000). Both nucleation and growth are critically influenced by the supersaturation of the mother liquor giving rise to the crystals. Supersaturation serves as the driving force behind both processes, influencing their occurrence, extent, and the governing kinetics(McPherson & Gavira, 2014).

Through the integration of automated robotic systems, researchers achieve precise and efficient control over protein crystallization conditions. This innovative approach not only

expedites the discovery of protein structures but also propels advancements in drug development. By automating the laborious steps of protein crystallization, the technology allows for extensive experimentation and significantly improves the likelihood of obtaining high-quality crystals suitable for X-ray crystallography investigations(Jenkins et al., 2013).

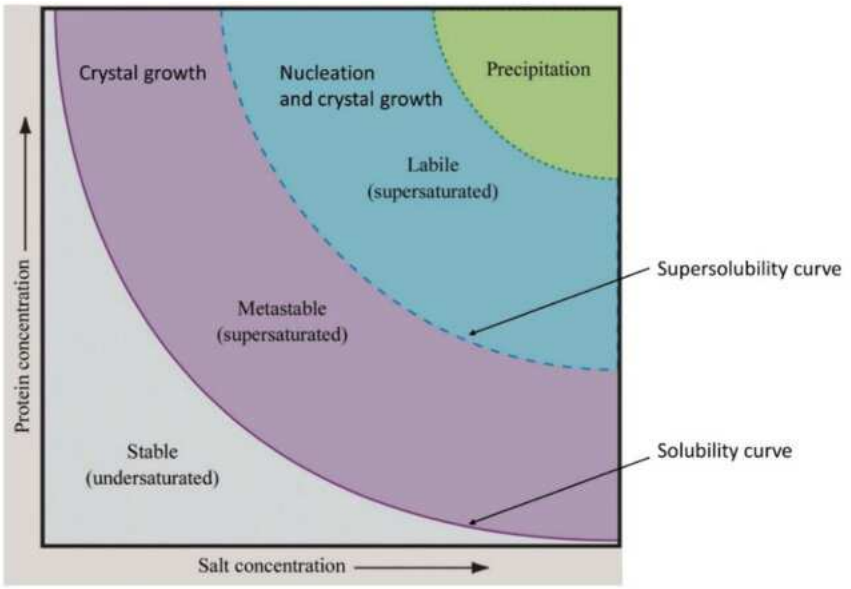


Figure 2.8.1.1 The phase diagram for the crystallization of macromolecules (Figure adapted from (McPherson & Gavira, 2014))

The solubility diagram is divided into undersaturation and supersaturation regions via a line indicating maximum solubility at specific precipitant concentrations, encompassing salt or polymer. This line epitomizes equilibrium between the solid and free-molecule phases. Supersaturation further subdivides into the enigmatic metastable and lability realms. In metastable regions, nuclei develop into crystals sans nucleation. In the lability region, both phenomena might occur. The high supersaturation segment, labeled the precipitation zone, holds the highest likelihood for the desired outcome. Crystals sprout solely from supersaturated solutions; achieving this with the protein of interest is the primary goal.

Utilizing the sitting drop and hanging drop vapor diffusion strategies, protein crystallization was pursued. This method encompassed an aqueous droplet containing the protein and crystallization agents, each at a lesser concentration than what's needed for crystal generation. The aqueous droplet gradually reached equilibrium with a reservoir buffer, leading to a gradual concentration increase of the components within the protein droplet. During this process of diffusion the crystallization trial experiments proceeds through a range of conditions thereby conducting a self-screening process (Chayen, 1998). Mosquito LCP robot was utilized for preparing the sitting drop in 96 well plates, in which 70 µl of crystallization buffer was used per well and different protein: reservoir buffer ratios were screened in total volume of 300 nl. Hanging drop was prepared manually, in which 750 µl of crystallization buffer was used per well and different protein: reservoir buffer ratios

were screened in total volume of 2 µl. The plates were incubated at different temperatures i.e., 16°C, 19°C, and 22°C.

2.9.1 Crystallization of CC domain and mutants

The crystallization of the CC domain and its mutants were performed according to the published paper (Tiruttani Subhramanyam et al., 2014). Briefly, a freshly purified protein sample at a concentration of 40-45 mg/ml was used for setting crystallization plates using the sitting drop vapor diffusion approach in which T-butanol with a concentration of 37 – 42 % was used as precipitant. A volume of 1 µl protein and 1 µl precipitant were mixed with 1 ml reservoir buffer volume. Crystallization plates were incubated at 22°C.

2.9.2 Crystallization of mutated Par-4 CC EXS protein

Always freshly prepared protein was used for setting the crystallization plates. Protein in SEC buffer (SEC buffer 1) was concentrated to 10 and 15 mg/ml, filtered and subjected for crystallization. Different screens were used for the initial crystallization experiments; screens from Nextal suites including, pH clear, PEGs, PEGII, (I, II, III, IV) of JCSG cores, MPD, and $AmSO_4$, screens from molecular dimension such as MemGold, MemGold 2, and proPlex, screens from Jena bioscience such as JBScreen wizard 1 and 2, JBScreen pentaerythritol 1 and 2. In addition, a crystallization plate was set up using different concentrations of tertiary butanol (t-butanol) as a precipitant (the precipitant used in crystallizing the CC domain (Tiruttani Subhramanyam et al., 2014)).

Based on the results of the first round of crystallization, a fine screen was made. The Nextal pH clear screen settings that led to needle crystals were composed of: 0.8 and 1.0 M $AmSO_4$ in HEPES pH 7.0, Tris-HCl pH 8.0 and Bicine pH 9.0.

Different approaches were utilized in order to improve the crystals quality such as in situ proteolysis and seeding. In situ proteolysis technique involves introducing a protease enzyme into the crystallization drop containing the protein sample, allowing enzymatic digestion to occur in the same well where the crystals are being grown. The

goal is to enhance crystal quality by selectively removing flexible or disordered regions of the protein, leading to improved diffraction quality for X-ray analysis. Different enzymes were tested for their cleavage activity on mutated Par-4 CC EXS including, trypsin, chymotrypsin, papain and subtilisin with different enzyme: protein weight ratios (1:100, 1:500, 1:1000, 1:2000, 1:5000, 1:10000 w/w). Immediately after mixing the Protein with the desired enzyme at different weight ratios, the crystallization plates were set up using the fine screen.

Seeding is a technique used in protein crystallography to improve the quality of crystals. It involves adding small crystals of the protein of interest to a solution of the protein. The small crystals act as a template for the growth of larger, more perfect crystals(Krauss et al., 2013). The crystals that grow from the seed are often larger and more perfect than those that grow without seeding. Mutated Par-4 CC EXS crystals were harvested and added to a microcentrifuge tube containing a glass bead and a stabilizing solution (mother liquor). Then, the crystals were vortexed and crushed. Different dilutions were prepared from the seed stock. The crystallization plates were set up after mixing the protein with different seed dilutions and incubated at different temperatures (16°C, 19°C, and 22°C).

2.9.3 Cryo-preservation and cryo-cooling of crystals

When exposed to a synchrotron X-ray beam at room temperature, many crystals experience a rapid decrease in diffraction power, typically within seconds. Continuing data collection under these conditions can exacerbate radiation damage due to prolonged exposure. Common indications of radiation damage include diminished diffraction power, loss of high-resolution data, expansion of the unit cell volume, and structural damage to covalent bonds. This damage can result in incomplete or non-isomorphous datasets, leading to inconclusive structural insights. To mitigate radiation damage, crystals are often cooled to 100 K using a cryo-stream of nitrogen during data collection at synchrotron sources, allowing for higher X-ray exposure doses without compromising data quality. This widespread adoption of cryogenic temperatures has become standard practice in structural biology (Garman, 2010).
However, cooling untreated protein crystals to cryogenic temperatures can introduce mechanical stress due to the expansion of water volume and the phase transition to

hexagonal ice. To prevent ice formation-induced damage, cryo-protectants are utilized to enhance the vitrification of water within the crystal and surrounding solvent. Careful selection of cryo-protectants is crucial to avoid disrupting the crystalline order through adverse interactions with the protein molecules. This strategy helps maintain crystal integrity during the cooling process, ensuring the preservation of high-quality diffraction data essential for structural analysis (Chang & Zhao, 2021).

Various cryoprotectants were assessed for their efficacy in preserving mutated Par-4 CC EXS crystals, including different concentrations of glycerol, polyethylene glycol (PEG), and dimethyl sulfoxide (DMSO). After fishing out the crystals using cryoloops, they were soaked in a cryoprotectant solution containing the precipitation buffer, followed by immersion in a cryo-stream of nitrogen at 100 K.

For CC domain and the its mutants, Perfluoropolyether (PFO) or LV cryo oil (LVCO-1) was overlaid on the sitting drop immediately after opening the crystallization experiment. Crystals were fished with cryoloops and were soaked in PFO before plunging into cryo-stream of nitrogen at 100 K.

2.9.4 Data collection and data processing

Studying crystal structures using X-ray diffraction is like using a super-powerful microscope to see how molecules are shaped. Typically, the wavelength of the light employed is usually in the vicinity of 1.5 Å, approximately the length of a carbon-carbon single bond, allowing, theoretically, the resolution of individual atoms. In microscopy, a specimen is illuminated by a light beam, and the resulting scattered light forms a focused image via a lens. Similarly, in X-ray diffraction, there is no X-ray lens to converge the scattered X-rays from the crystal's electrons. Consequently, direct interpretation through image acquisition is unattainable. Instead, crystallographers measure the intensities of the scattered X-rays from the crystal and utilize these, along with supplementary data, to generate an image computationally.

The reason to use crystal is that detecting the signal from just one molecule would be too faint to see clearly. The crystal behaves like a three-dimensional diffraction

grating, scattering the incident X-ray beam only in specific directions. As each molecule within the crystal contributes to this scattering collectively, and considering the immense number of molecules present (often exceeding 10^{15} in most crystals), it becomes feasible to capture the scattered signal before any radiation-induced harm occurs to the crystal.

The phenomenon of diffraction can be illustrated as the combined effect of each atom within the unit cell scattering X-rays. Each atom contributes to the diffraction with its own amplitude, determined by factors such as its electron count and size, and a phase, influenced by its position in the unit cell (Hickman & Davies, 1997). When these individual contributions are summed up, they form a vector characterized by an amplitude denoted as |F(h)| and a phase represented by α(h), where 'h' corresponds to (hkl), with h, k, and l serving as Miller indices defining the diffraction peaks. The amplitude |F(h)| and phase α(h) of the vector can be expressed by a structure factor termed F(h). This factor is solely dependent on the scatterer's structure and can be mathematically expressed in terms of the fj, which are individual scattering factors for each atom 'j' present in the unit cell, along with the atomic coordinates xj (Hickman & Davies, 1997).

$$F(\underline{h}) = \Sigma_j f_j \exp 2\pi i(\underline{h} \cdot \mathrm{x}_j) = |F(\underline{h})| \exp i\alpha(\underline{h})$$

By systematically varying the orientation of the crystal in relation to the X-ray beam, the complete diffraction pattern of the crystal can be captured and analyzed (Hickman & Davies, 1997). While crystallographers can measure the amplitudes of the diffraction intensities, represented by |F(h)|, directly from the recorded data, determining the phase angles, denoted as α(h), poses a challenge due to experimental limitations. To overcome this hurdle, alternative methods are employed to recover the missing phase information (Hickman & Davies, 1997). Once both the amplitudes and phases of the scattered X-rays are known, they are utilized to generate an electron-density map, providing insights into the arrangement of atoms within the crystal's repeating unit, known as the unit cell. The electron density (ρ) at any given position (xyz) within the unit cell can be mathematically described using a specific

$$\rho(x, y, z) = \Sigma_{hkl} |F(h, k, l)| \exp i\alpha(hkl) \exp -2\pi i(hx + ky + lz)$$

equation, incorporating complex numbers (with i denoting the square root of -1), enabling further structural analysis and interpretation (Hickman & Davies, 1997).

The crystals were tested for diffraction and the datasets for the diffracted ones were collected at different beamlines (Table 2.9.4.1) at the European Synchrotron Radiation Facility (ESRF, Grenoble, FRANCE). All measurements were performed in a cryo-stream of nitrogen at 100 K.

Once crystals of a protein diffract X-rays to a reasonable resolution, the next crucial step is to obtain an interpretable electron-density map, which requires accurate phase angles (α(hkl)).

Table 2.9.4.1 different beamlines used for data collection

Beamline	Type	Energy
ID23-1	Tunable	6 – 20 keV
ID23-2	Microfocus (fixed energy)	14.20 keV
ID30A-3	Microfocus (fixed energy)	12.9 keV
ID30B	Tunable	6 – 20 keV

Many methods have been employed for recovering the phases, all of which have their own limitations. Direct methods involve prediction of unknown phases mathematically based on the positivity and atomicity of electron density and are limited to small molecules or substructures of macromolecules (<100 atoms) (Johnson, 1985).

The multiple isomorphous replacement (MIR) technique involves attaching a heavy atom to the protein within its crystal. Typically, crystals are soaked in a heavy-atom solution, taking advantage of weak attachment sites on protein surfaces. Alternatively, the protein can be cocrystallized with a heavy-atom compound (Hickman & Davies, 1997; Ke, 1997; Taylor, 2010). The heavy atom's electron density leads to measurable differences in X-ray scattering intensities, aiding in locating its binding site. Multiple

crystals with various heavy atoms binding to distinct protein sites allow phase calculations for protein structure factors, facilitating electron-density map generation (Ke, 1997; Taylor, 2010). Despite challenges like accurately measuring small intensity differences and non-isomorphism between native and heavy-atom crystals, regions fitting the known sequence can usually be identified. This enables the construction of a model consistent with the protein's polypeptide backbone and side chains (Hickman & Davies, 1997; Taylor, 2010).

Anomalous-scattering methods also utilize heavy atoms, capitalizing on their behavior near absorption edges to scatter X-rays anomalously, altering both amplitude and phase (Hendrickson et al., 1985; Hickman & Davies, 1997). This technique, applied with a single heavy-atom derivative, initially identifies heavy atom binding sites and subsequently computes crystal phase angles (Cianci et al., 2005). While feasible at a single wavelength, the method yields more pronounced effects near or at the absorption edge. By measuring diffraction intensities at wavelengths both near and distant from the absorption edge, optimal intensity discrepancies are observed and utilized for structure solution (Hickman & Davies, 1997). The advent of tunable X-ray sources, like synchrotron radiation, has enhanced this approach, facilitating the development of multiwavelength anomalous dispersion (MAD) procedures. Wayne Hendrickson and colleagues pioneered this method, employing selenomethionine to substitute for methionine as a suitable anomalous scatterer (Hendrickson et al., 1985). A notable advantage of anomalous-scattering methods is their ability to determine structure using data from a single crystal harboring a suitable heavy atom (Hendrickson et al., 1985; Hickman & Davies, 1997).

The molecular-replacement method is utilized when there's an expectation of structural similarity between a target protein and a known protein structure. This similarity could arise from various scenarios, such as mutant proteins crystallized in different forms or proteins with high sequence homology to known structures (Evans & McCoy, 2008). The process involves two stages: rotation function and translation function. Initially, the rotation function determines the required rotation to align the known molecule's orientation with that of the unknown molecule within the unit cell. Subsequently, the translation function determines the precise position of the search molecule in the cell. This entails iteratively placing the search molecule in the unit cell,

avoiding unfavorable contacts, and seeking a position where observed and calculated structure factors align reasonably well. Molecular replacement is particularly effective when the two molecules share significant structural resemblance (Evans & McCoy, 2008) (Hickman & Davies, 1997). Several software packages, such as AMORE, MERLOT, and CCP-4, support the molecular-replacement method (Hickman & Davies, 1997; McCoy, 2007).

After that, the data were indexed and integrated with IMOSFLM (Battye et al., 2011), and validated and scaled using Aimless (Winn et al., 2011) . Molecular replacement using the crystal structure of colied coil (CC) domain of PAWR (PDB ID: 5FIY) was performed using MOLREP in CCP4 suit (Vagin & Teplyakov, 1997). Phenix and coot were utilized in order to carry out rebuilding and model refining (Afonine et al., 2012; Emsley & Cowtan, 2004; Emsley et al., 2010; Wrapp et al., 2020).

3 Results

3.1 Effect of phosphorylation-mimicking, salt bridge-removal, and charge repulsion removal mutations on CC domain structure

The region beyond the CC domain was predicted to be disordered. Par-4 contains a phosphorylation site for Akt1 at Ser249, which is close to the crystallographically solved structure of the CC domain (254-332). Since protein phosphorylation plays a crucial role in protein stability and folding, we proposed that phosphorylating Ser249 could induce structural changes in the CC domain and/or in the region beyond it, potentially enhancing folding and stability especailly in the region beyond the CC domain. Therefore, phosphorylation-mimicking mutations at Ser249 and Thr248 were created to investigate the biophysical and structural characterization of this region upon phosphorylation. These mutations are T248D, T248E, S249D, and S249E.

In addition, the structural analysis of the CC domain reveals the presence of various salt bridge interactions that stabilize the coiled-coil structure. The structure also shows single charge repulsion at residue Asp305 in the CC domain. The impact of some of these interactions on CC domain stability and folding was investigated by assessing the biophysical characteristics after salt bridge and charge repulsion removal mutations. These mutations are E268Q, E274Q, K288M, D305N, and K317M.

This results section shows the effects of these mutations on the CC domain structure.

3.1.1 Cloning of CC domain and mutants

The strep-tagged Par-4 CC construct (made by Dr. Udaya Kumar Tiruttani Subhramanyan) from our Lab stocks was used as a template for the creation of CC domain's mutants using Site-directed mutagenesis. The sequence integrity of all the mutants' plasmids was confirmed by Microsynth company using the T7 forward and T7 reverse primers.

3.1.2 Protein expression of CC domain and mutants

The expression of the CC domain and the mutants was achieved according to the published paper by Dr. Udaya Kumar Tiruttani Subhramanyan from our group (Tiruttani Subhramanyam et al., 2014). The expression was induced with 1 mM (IPTG) when OD600 reached 0.6 -0.8 and incubated at 37°C with 120 rpm for 4 hours. All the proteins were soluble at 37°C with decent amount.

3.1.3 Purification of CC domain and mutants

The purification of the CC domain and the mutants was achieved according to the published paper (Tiruttani Subhramanyam et al., 2014). Figure (3.1.3.1 A) shows the SDS-PAGE analysis of CC domain purification where a single band of protein can be seen. Figure (3.1.3.1 B) shows the SDS-PAGE of the SEC elution fractions obtained from gel filtration.

The theoretical molecular weights of the CC domain and the mutants are around 12.3 kDa. The SDS-PAGE for the SEC fractions shows a single band that ran at the molecular weight around 12 kDa. The SEC profile (figure 3.1.3.2) shows that the protein elutes with an apparent molecular weight of 95.30 kDa, which indicates an octamer (12.37490 kDa * 8 = 98.9 kDa).

The mutants as well, eluted with the same molecular weight of the CC domain, i.e., 95.30 kDa. This indicates that the mutations did not change the oligomerization state of the protein as all eluted at the same elution volume.

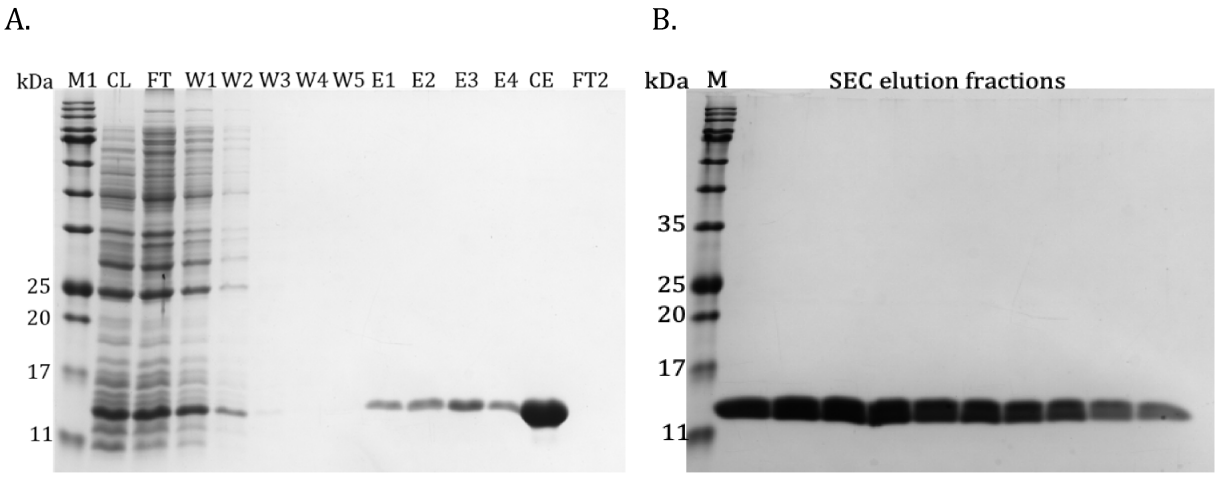


Figure 3.1.3.1 SDS-PAGE of CC domain purification

A. SDS-PAGE analysis of CC domain purification (blue silver stain), M- pre-stained marker; CL- cell lysate; FT flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator. B SDS-PAGE analysis of mutated Par-4 CC EXS different SEC elution fractions.

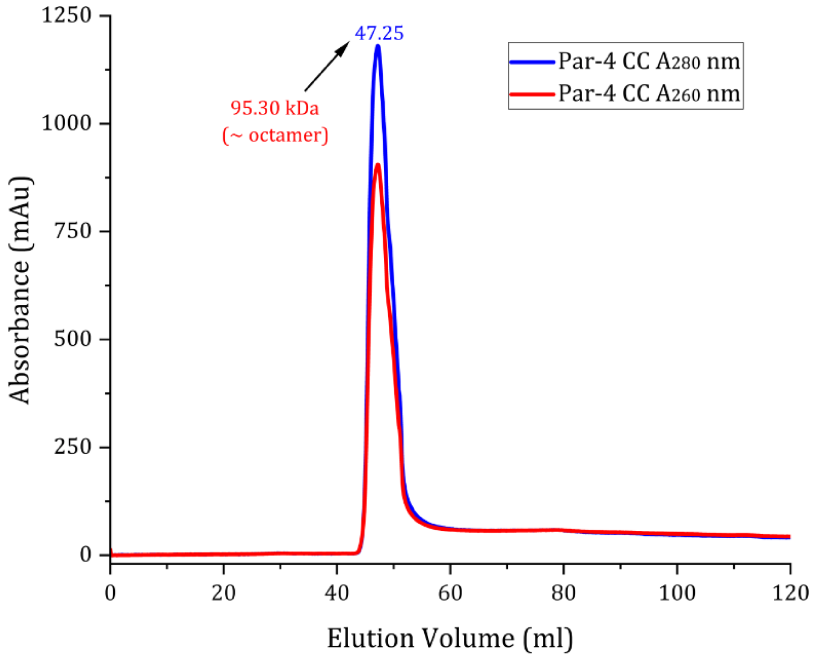


Figure 3.1.3.2 Superdex 75 SEC profile of CC domain

3.1.4 Identity and integrity of the proteins

The purified proteins were subjected to different techniques in order to verify its identify and integrity including western blot, mass spectrometry and LC/MS mass.

3.1.4.1 Western Blot results for CC domain and mutants

The anti-strep tag antibody was used in western blot detection since the proteins have a strep-tag. Figures (3.1.4.1. 1 A and B) show the western blot detection of the purification profile obtained from affinity chromatography and SEC elution fractions for CC domain and mutants.

The purified proteins are western positive which indicates the N- terminal end is intact in the proteins post purification.

A. B.

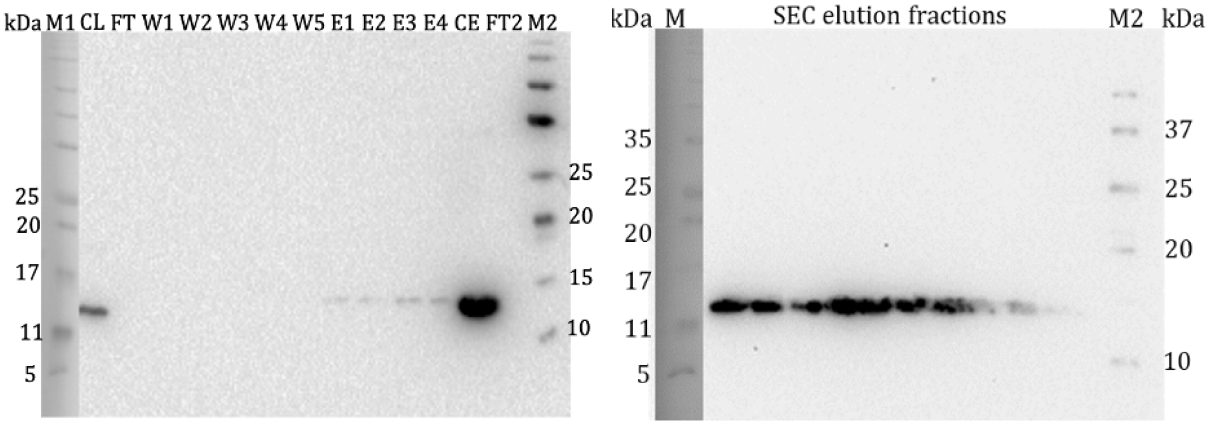


Figure 3.4.1.1.1 Anti-strep western blot of CC domain purification.

A. Anti-strep western blot analysis of CC domain purification, M1- pre-stained marker; CL- cell lysate; FT1 flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator; M1 - Precision Plus Protein™ WesternC™ Blotting Standards. B Western blot analysis of CC domain different SEC elution fractions.

3.1.4.2 LC/MS mass results for CC domain and mutants

The integrity of the proteins was evaluated using LC/MS mass spectrometry, with in-gel trypsin digestion performed to determine the N-terminal and C-terminal of the proteins. Trypsin cleaves specific peptide bonds on the C-terminal side of lysine and arginine residues. The resulting peptides were analyzed using the Agilent 6550 iFunnel Q-TOF system and Agilent liquid chromatography/mass spectrometry (LC/MS) databases and libraries. Detection of multiple peptides with identical N- or C-termini indicates the N- or C-terminus of a protein. The amino acid sequence coverage for the CC domain and its mutants was 96.2% (Figure 3.4.2.1). The N-terminal methionine residue was found to be cleaved, likely due to methionine aminopeptidase, which co-translationally removes the N-terminal methionine from nascent proteins. This cleavage often occurs when the second residue in the primary sequence is small and uncharged (e.g., Ala, Cys, Gly, Pro, Ser, Thr, or Val)(Varland et al., 2015). The CC domain sequence contains Met-Ala at the first two positions, explaining the methionine cleavage. Overall, the LC/MS results indicate that the purified protein is full-length and uncut during the purification process. Additionally, post-translational modification analysis showed that the protein did not have any modifications.

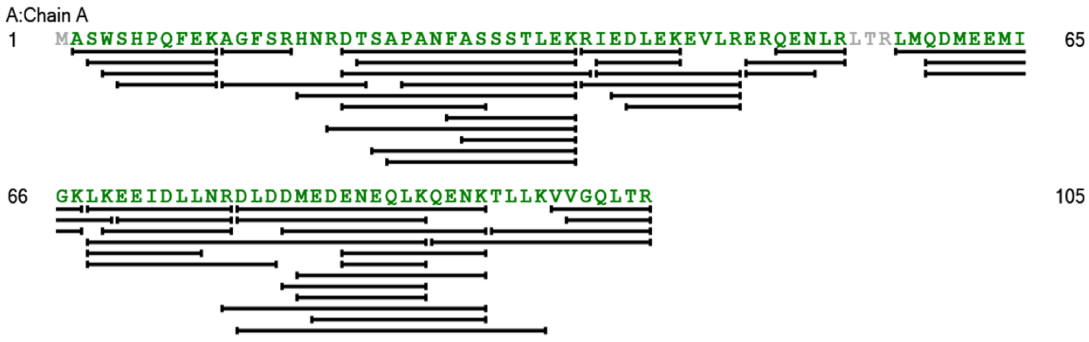


Figure 3.4.2.1 Peptides identification of mutated Par-4 CC EXS by LC-MS/MS analysis

LC-MS/MS gave N terminal end to C terminal end sequence coverage. A LC-MS/MS result CC domain

3.1.4.3 Mass spectrometry results for CC domain and mutants

Mass spectrometry was employed to determine the proteins' molecular weights more accurately. A protein standard with molecular weights ranging from 4 to 20 kDa was used to calibrate the mass spectrometer, ensuring reliable analysis by MALDI-TOF for molecular weights within that range. The MS result for the CC domain (Figure 3.1.4.3.1) shows two peaks: the strongest peak at 12.243 kDa, which matches the M^{1+} species of the CC domain, and a weaker peak at 6.121 kDa, corresponding to the M^{2+} species of the CC domain. The MS results did not exactly match the theoretical molecular weights of the CC domain. This discrepancy could be due to the cleavage of the N-terminal methionine, as previously observed in LC/MS mass results, since the difference is equal to the molecular weight of methionine (131 Da).

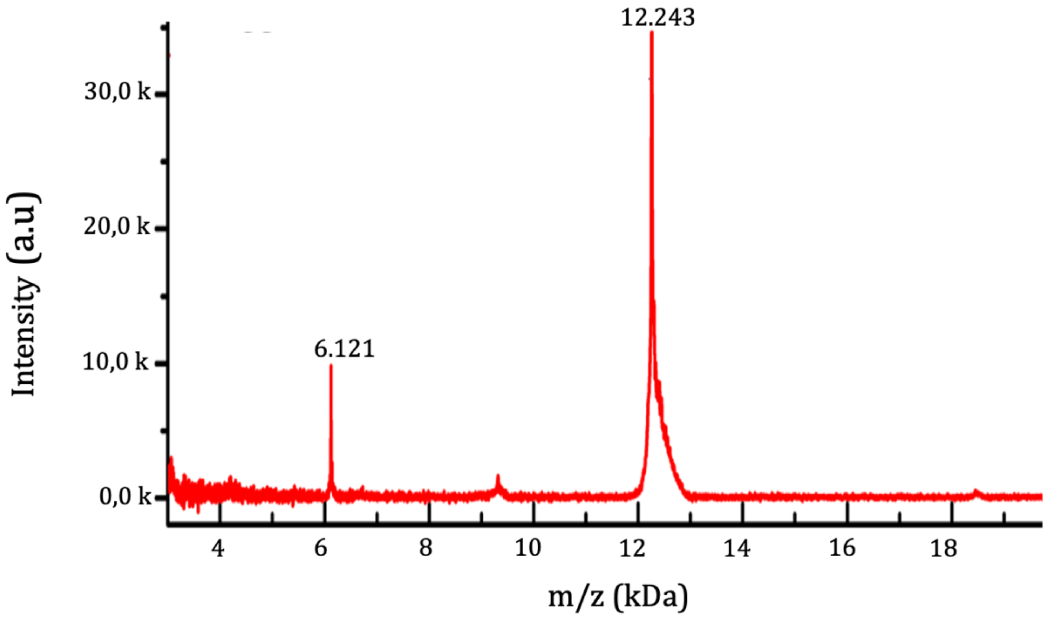


Figure 3.1.4.3.1 MALDI-TOF mass spectrum of CC domain

Since the evaluation of proteins' purity, integrity, and quantity have all been done, the subsequent step was to study the biophysical characterization of the protein.

3.1.5 Biophysical Characterization

Biophysical characterization techniques play a fundamental role in protein studies due to their ability to provide detailed insights into the secondary and tertiary structure, folding, stability, dynamic properties, as well as, the aggregation and higher oligomerization of the protein.

In order to assess the above-mentioned properties, the purified proteins were subjected to different biophysical techniques including Dynamic Light Scattering (DLS), Circular Dichroism (CD), and X-crystallography.

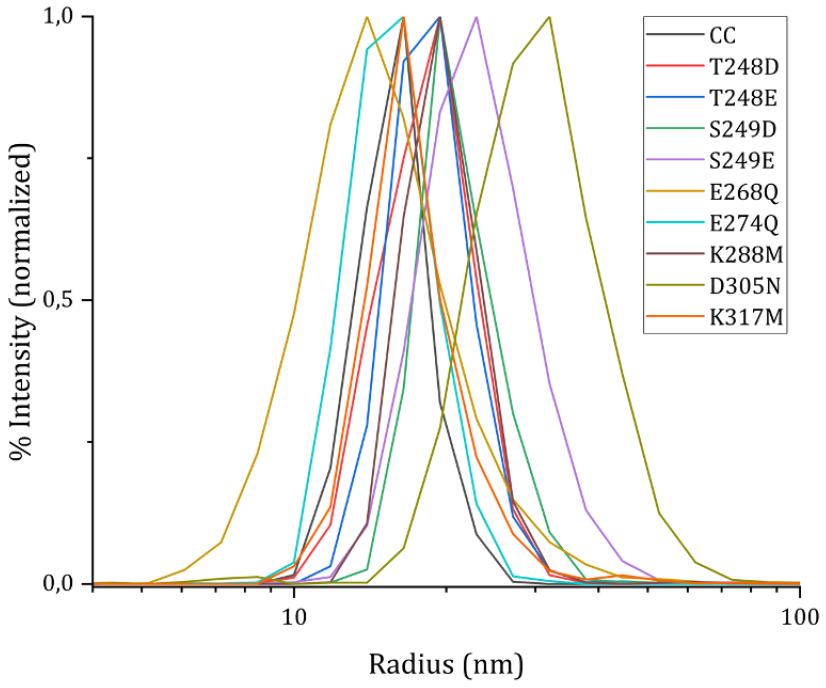
3.1.5.1 Dynamic Light Scattering results for CC domain and mutants

The DLS results of the CC domain and its mutants indicate that all the proteins formed highly aggregated states, with molecular weights exceeding 1.5 MD (Figures 3.1.5.1.1 A and B; Table 3.1.5.1.1). Additionally, all the proteins, except for the T248D, E268Q, E274Q, and D305N mutants, exhibited low polydispersity. Among all the mutants, D305N had the highest impact on particle size and PD%. Furthermore, the high oligomerization state of the CC domain and its mutants was assessed by calculating the number of molecules present in each protein. The results showed that the E268Q, E274Q, and K317M mutants had the lowest number of molecules compared to the CC domain, indicating that these mutations could reduce the high oligomerization of the protein. In contrast, the D305N mutant had the highest significant number of molecules, which could be due to the removal of the repulsive force between the helices. Moreover, the phosphorylation-mimicking mutants especially S249E exhibited higher oligomerization states compared to the CC domain.

Table 3.1.5.1.1 DLS results for CC domain and mutants

Protein	Radius (nm)	PD%	M.wt (MDa)	Intensity	No. of molecules
CC	16.57	14.7	2.084	1	171
T248D	18.93	19.53	2.852	1	234
T248E	18.43	16.8	2.663	1	218
S249D	18.97	8.3	2.844	1	233
S249E	22.65	11.1	4.28	1	351
E268Q	14.92	21.4	1.637	1	134
E274Q	16.23	18.0	1.988	1	163
K288M	20.62	13.6	3.447	1	283
D305N	32.28	22.5	9.663	1	792
K317M	15.74	7.9	1.855	1	152

A.



B.

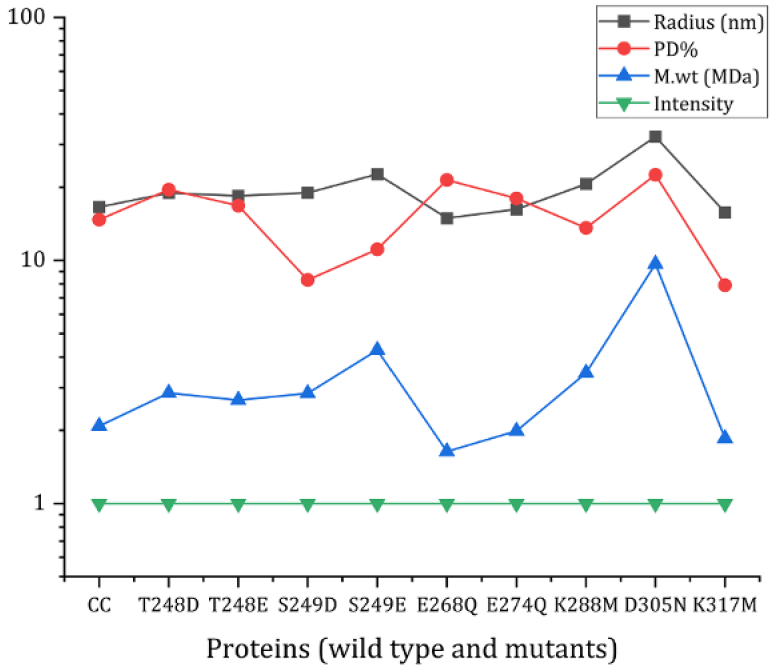


Figure 3.1.5.1.1 Particle size analysis of CC domain and mutants using DLS

A. Intensity percentage vs radius of different protein samples at acidic and neutral conditions. B. Log 10 of Radius, polydispersity %, molecular weight and intensity of different protein samples at acidic and neutral conditions.

3.1.5.2 Assessment of the secondary structure of the CC domain and its mutants by CD

The CD measurements were conducted on the CC domain and its mutants under acidic conditions in a 50 mM potassium phosphate buffer. Analysis of the far-UV spectra revealed an α-helical profile characterized by double minima at 208 nm and 222 nm, along with maxima around 193 nm (Figure 3.1.5.2.1). Notably, the MRE222/MRE208 ratio, indicative of the coiled coil conformation (Schwalbe et al., 2010) , was consistently greater than 1.0 for both the CC domain and its mutants (Table 3.1.5.2.2). Comparing these ratios to the CC domain (1.093), the mutants S249D and S249E exhibited the highest ratios, while the mutants E268Q and E274Q displayed the lowest ratios.

The CD spectra deconvolution analysis revealed a consistent 100% helix content across all proteins (table 3.1.5.2.1). However, variations were observed in the percentage of residues belonging to helix 1 (the regular portion of helices at the center) and helix 2 (the distorted portion at each end). In the CC domain, 86 residues were identified in helix 1, with 19 residues in helix 2, out of a total of 105 residues. Interestingly, calculations indicated an increase in the number of residues within helix 1 for all mutants, except for the K288M mutant, which showed a reduction of one residue compared to the CC domain in the same helix. Notably, the D305N mutant stands out as the only variant with all 105 residues located in helix 1 among the proteins analyzed.

Overall, all the proteins showed a 100 % helical structure in solution.

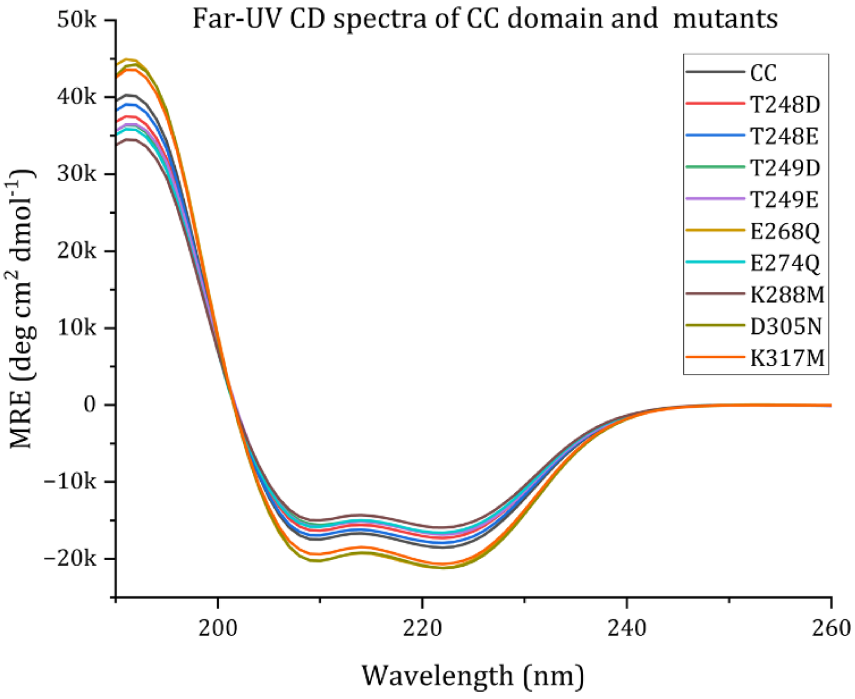


Figure 3.1.5.2.1 Far-UV CD spectra of CC domain and mutants

Table 3.1.5.2.1 Deconvolution of the CD spectra for CC domain and mutants

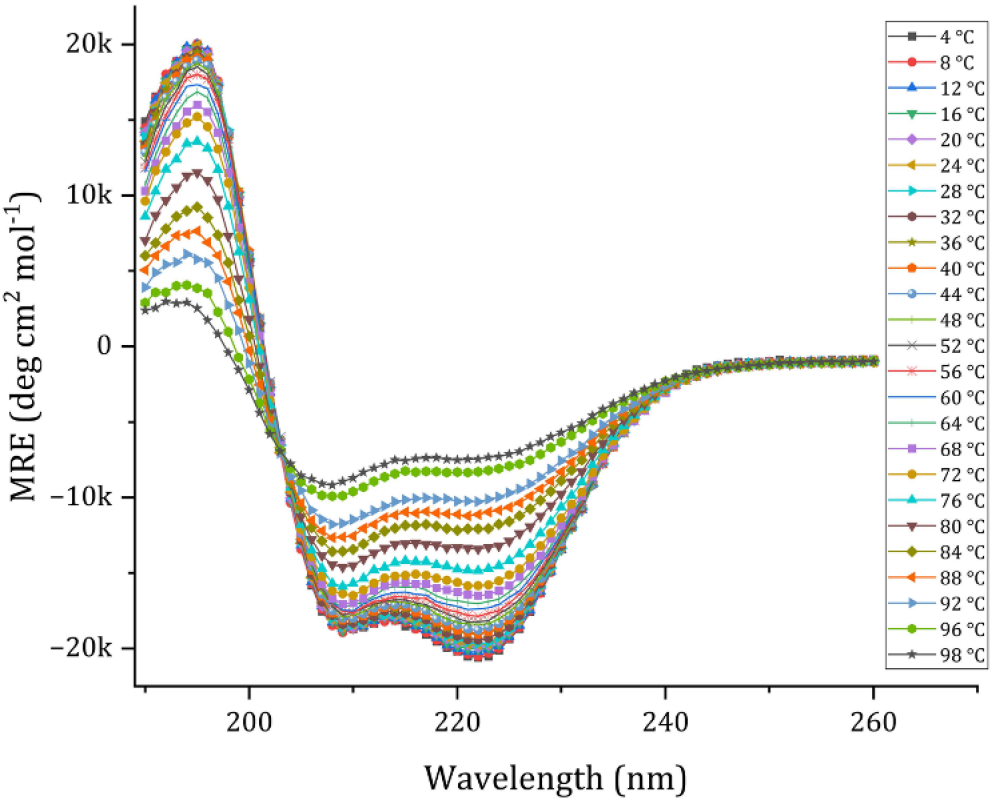
Protein	Helix %		Anti-parallel %	parallel %	Turns %	Other %	Total %
	Helix 1 %	Helix 2 %					
CC	81.5	18.5	0	0	0	0	100%
T248D	90.5	9.5	0	0	0	0	100%
T248E	91.5	8.5	0	0	0	0	100%
S249D	90.2	9.8	0	0	0	0	100%
S249E	91.5	8.5	0	0	0	0	100%
E268Q	91.1	8.9	0	0	0	0	100%
E274Q	85.7	14.3	0	0	0	0	100%
K288M	81.3	18.7	0	0	0	0	100%
D305N	100	0	0	0	0	0	100%
K317M	93.0	7.0	0	0	0	0	100%

Table 3.1.5.2.2 Number of residues in helix 1 and helix 2 for CC domain and mutants

Protein	MRE222/MRE208 ratio	No. of residues in helix 1 (regular)	No. of residues in helix 2 (distorted)
CC	1.093	86	19
T248D	1.0935	95	10
T248E	1.0878	95	10
S249D	1.1098	95	10
S249E	1.1014	96	9
E268Q	1.078	96	9
E274Q	1.079	90	15
K288M	1.097	85	20
D305N	1.080	105	0
K317M	1.099	98	7

The melting curves of the CC domain and its mutants were generated by plotting the CD intensity at 222 nm against the temperature scan, as changes in CD intensity at both 222 nm and 208 nm were observed during thermal denaturation (Figure 3.1.5.2.2 A). The results revealed a significant increase in melting temperature for the mutants compared to the CC domain (Figure 3.1.5.2.2 B). Notably, phosphorylation-mimicking mutants exhibited nearly identical melting temperatures, averaging around 77.4±0.2 °C, while mutations within the CC domain resulted in lower melting temperatures, averaging around 75.0±0.5 °C. Interestingly, the K317M mutant displayed the highest melting temperature among all mutants, reaching 80.0 °C. Surprisingly, all the mutations in the CC domain were created to remove the salt bridge interactions and thus decrease the CC domain stability, except for D305N, which was created to remove the charge repulsion. However, these findings suggest higher stability for all mutants compared to the native CC domain.

A.



B.

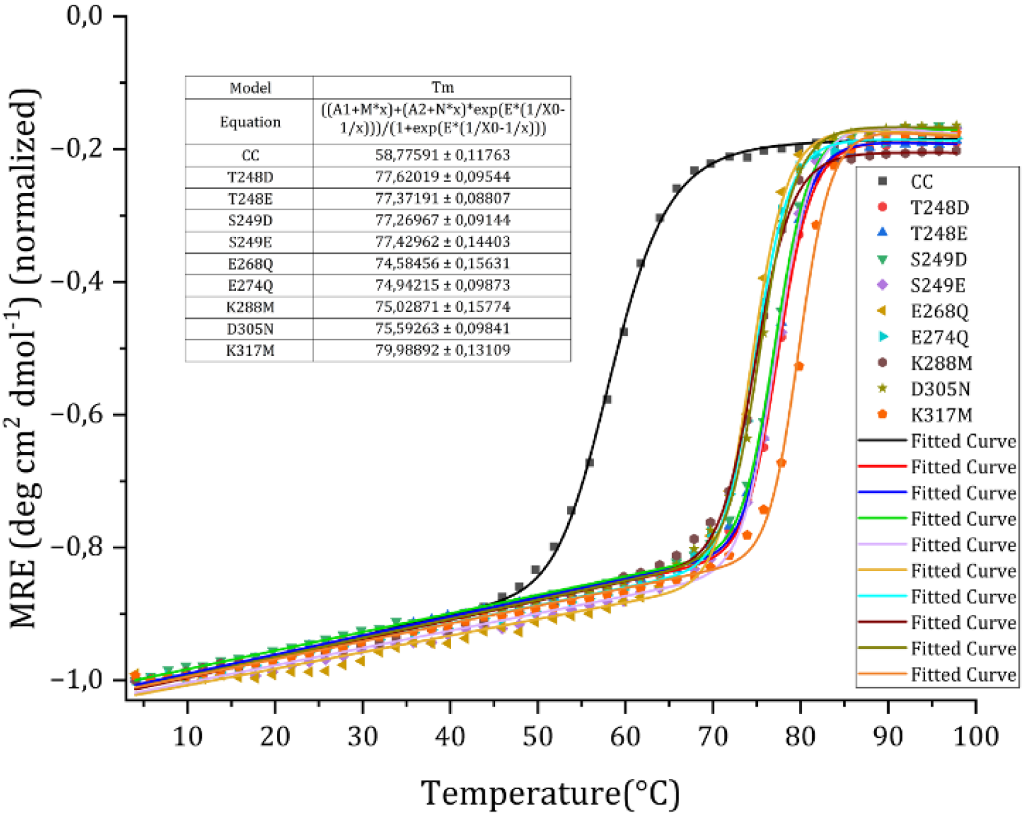


Figure 3.1.5.2.2 Thermal stability of CC domain and mutants

A. Temperature scanning of CC domain in 4 °C increments from 4 to 98 °C from by Circular Dichroism Spectroscopy. B. Thermal stability (Tm) of CC domain and its mutants at 222 nm.

The impact of mutations on protein stability was assessed by calculating thermodynamic parameters for each mutant and the CC domain. The Van 't Hoff plot (figure 3.1.5.2.3) illustrates the slope and intercept for each protein, detailed in the accompanying table 3.1.5.2.2. Thermodynamic parameters, including ΔH and ΔS at the melting temperature, are also summarized in the table 3.1.5.2.3. Additionally, stability curves for the proteins are depicted in the figure figure 3.1.5.2.6. The results demonstrate a significant increase in ΔG Unfolding for all mutants compared to the CC domain. Particularly, phosphorylation-mimicking mutants and the K317M mutant exhibit the highest ΔG Unfolding among all proteins, indicating their superior stability. Furthermore, the positive values of ΔΔG (ΔG mutant - ΔG CC domain) and ΔΔH (ΔH mutant - ΔH CC domain) suggest that mutants require higher free energy and more heat for the denaturing reaction compared to the CC domain, underscoring their enhanced stability (figures 3.1.5.2.4 and 3.1.5.2.5).

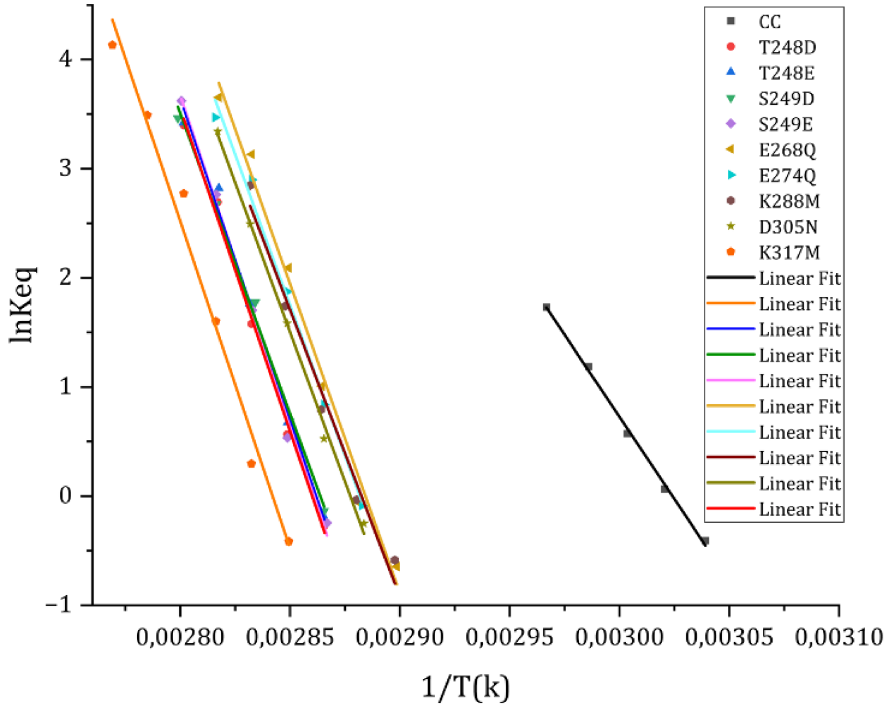


Figure 3.1.5.2.3 Van 't Hoff plots of CC domain and mutants

Table 3.1.5.2.2 Slopes and intercepts for CC domain and mutants obtained from Van 't Hoff plot

Protein	Slope	Slope error	Intercept	Intercept error	R-square
CC	-30101.2	905.1	91.0	2.7	0.9964
T248D	-59110.8	2437.8	169.1	6.9	0.9932
T248E	-58669.9	3038.2	167.9	8.6	0.9893
S249D	-55043.5	2369.8	157.6	6.7	0.9930
S249E	-60289.9	2630.1	172.5	7.5	0.9924
E268Q	-56676.4	2923.4	163.5	8.4	0.9870
E274Q	-55162.1	2913.6	159.0	8.3	0.9890
K288M	-52510.4	4132.0	151.4	11.8	0.9757
D305N	-54816.9	1799.2	157.7	5.1	0.9957
K317M	-59881.9	3837.1	170.2	10.7	0.9798

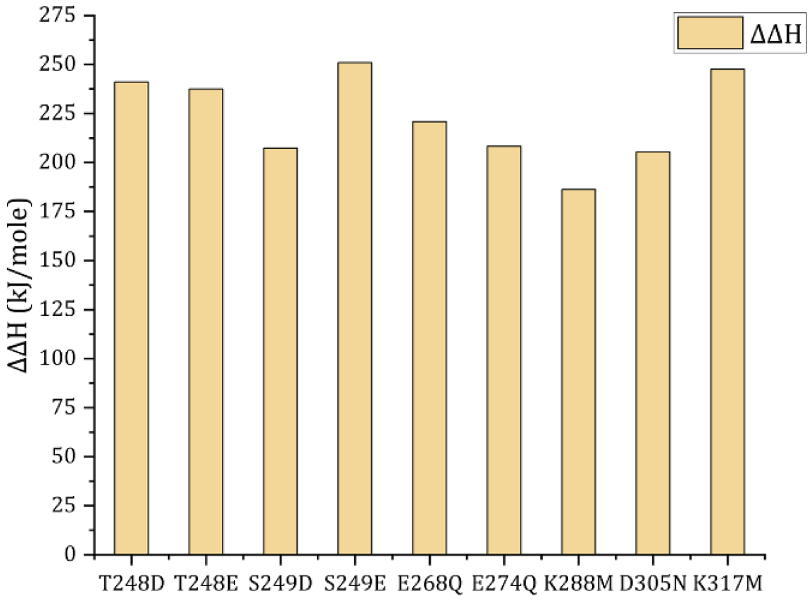


Figure 3.1.5.2.4 Enthalpy changes of mutants after applying ΔΔH: ΔH (mutant)- ΔH (CC)

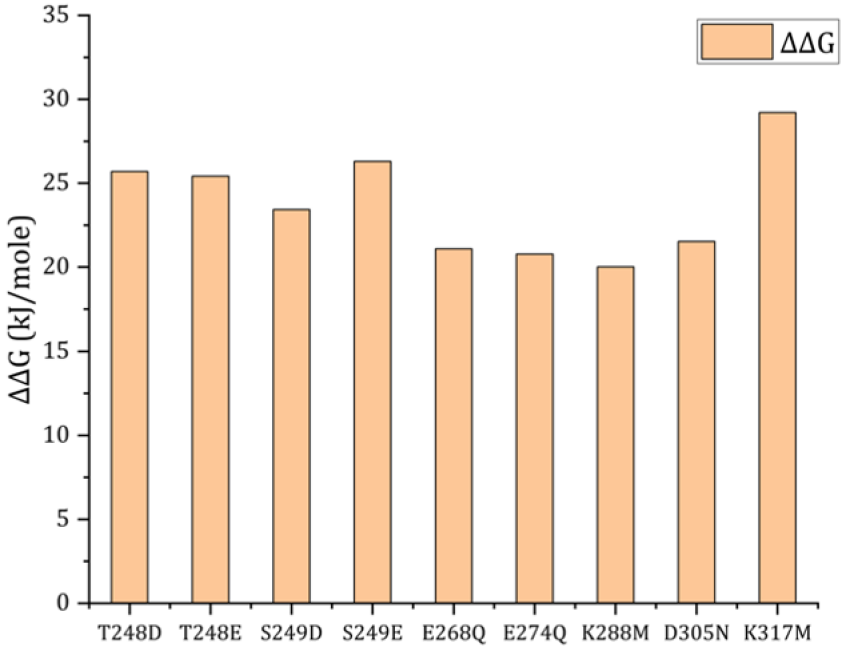


Figure 3.1.5.2.5 Free energy changes of mutants after applying ΔΔG: ΔG (mutant)- ΔG (CC)

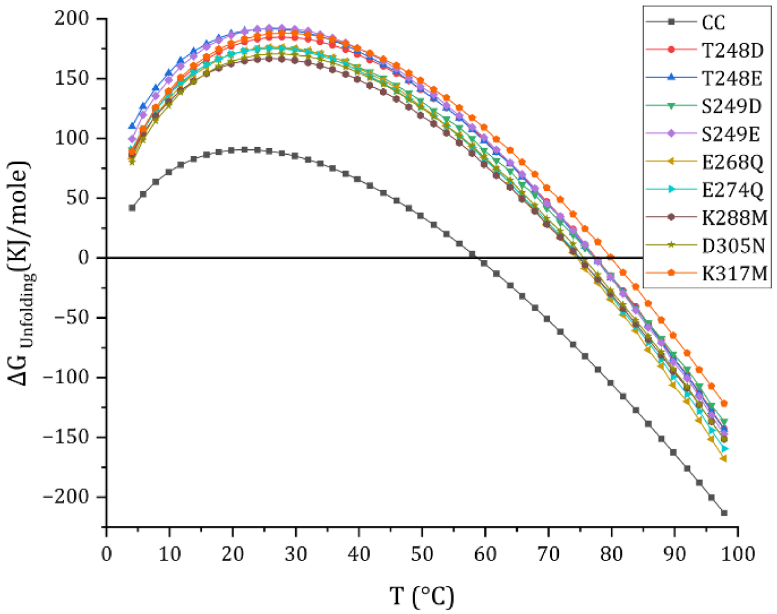


Figure 3.1.5.2.6 Stability curves of CC domain and its mutants

Table 3.1.5.2.3 Thermodynamic parameters for CC domain and mutants

Protein	Tm (°C)	ΔH_{Tm} (kJ/mol)	ΔS_{Tm} (kJ/k.mol)	ΔCp (kJ/k.mol)	ΔTm	ΔΔH (kJ/mol)	ΔΔG (kJ/mol)
CC	58.8	250.3	0.757	4.35	____	_____	_____
T248D	77.6	491.1	1.406	6.14	18.8	240.8	24.78
T248E	77.4	487.8	1.396	5.74	18.6	237.5	24.68
S249D	77.3	457.6	1.311	5.58	18.5	207.3	22.87
S249E	77.4	501.3	1.434	6.10	18.6	251.0	25.28
E268Q	74.6	471.2	1.359	6.17	15.8	220.9	21.00
E274Q	74.9	458.6	1.322	5.80	16.1	208.3	20.77
K288M	75.0	436.6	1.258	5.52	16.2	186.3	19.50
D305N	75.6	455.7	1.311	5.88	16.8	205.4	18.98
K317M	80.0	497.9	1.415	6.00	21.2	247.6	27.23

3.1.6 Crystallization of Proteins

3.1.6.1 Crystallization of CC domain and mutants

The crystallization of the CC domain and the mutants was achieved according to the published paper (Tiruttani Subhramanyam et al., 2014), which used 37-42% t-butanol as a precipitant in 100 mM sodium citrate/citric acid buffer at pH 5.9. The same t-butanol percentages and buffer with a pH range from 5.7 to 6.1 were used to crystallize the proteins.

Immediately after addition of crystallization buffer precipitation of the proteins sample was observed (Figure 3.1.6.1.1A), but the samples usually cleared within an hour (Figure 3.1.6.1.1 B). Phase transformation of precipitated proteins to a gel-like phase was observed (Figure 3.1.6.1.1 C) when 24 well plate with 1ml of reservoir buffer and 1-3 µl of sitting drop was used. The successful crystallization experiments of the CC domain and mutants had 37 and 38% t-butanol in 100 mM sodium citrate/citric acid pH 6.1.

The proteins crystals appeared randomly from the cleared crystallization droplets as well as from the gel-like phase within a time range from 10 days to a couple of weeks. All the mutations within the CC domain i.e., salt bridge removal mutations, did not form any crystals and the droplets remained clear even after 6 months of incubation, implying no nucleation initiated. The crystals appeared only for CC domain in addition to the phosphorylation mimicking mutants except for T248E. The shape of the crystals for CC domain and the phosphorylation mimicking mutants looked tetragonal (Figure 3.1.6.1.1 E), sometimes they formed clusters (Figure 3.1.6.1.1 D).

The crystals (Figures 3.1.6.1.2 A1, B1, C1) diffracted at 3.16 Å, 3.02 Å, and 2.98 Å for T248D, S249D, and S249E, respectively. The diffraction patterns for the crystals are shown in (Figures 3.1.6.1.2 A2, B2, C2).

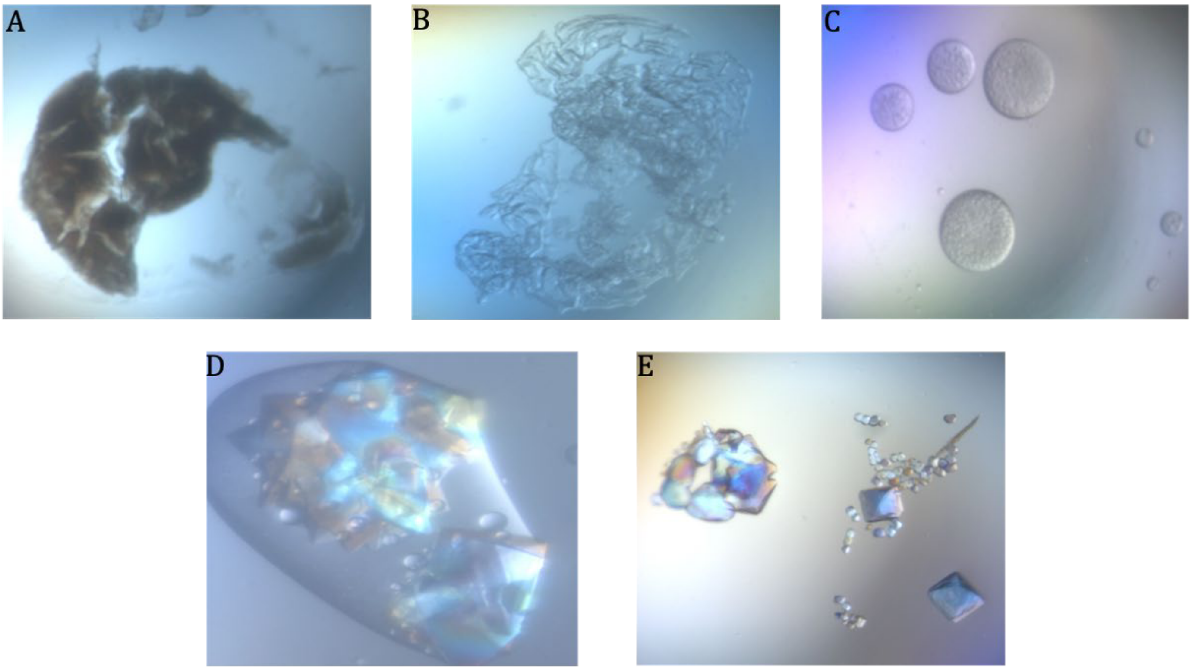
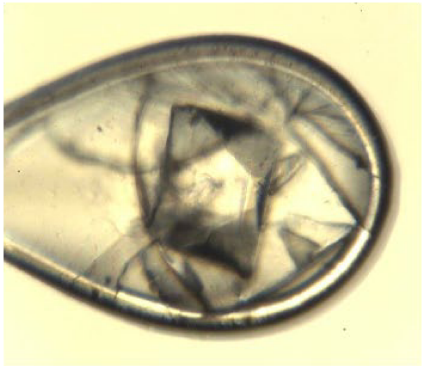


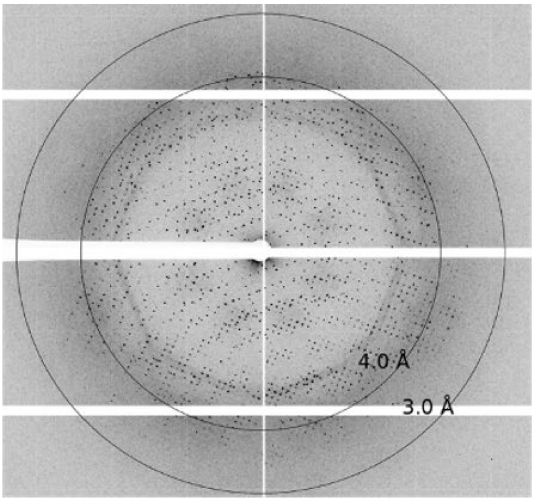
Figure 3.1.6.1.1 Crystallization profile for CC domain and mutants

A- Precipitated protein, B-Cleared droplet, C- Gel-like phase, D- Crystal cluster, E- Crystal.

A1



A2



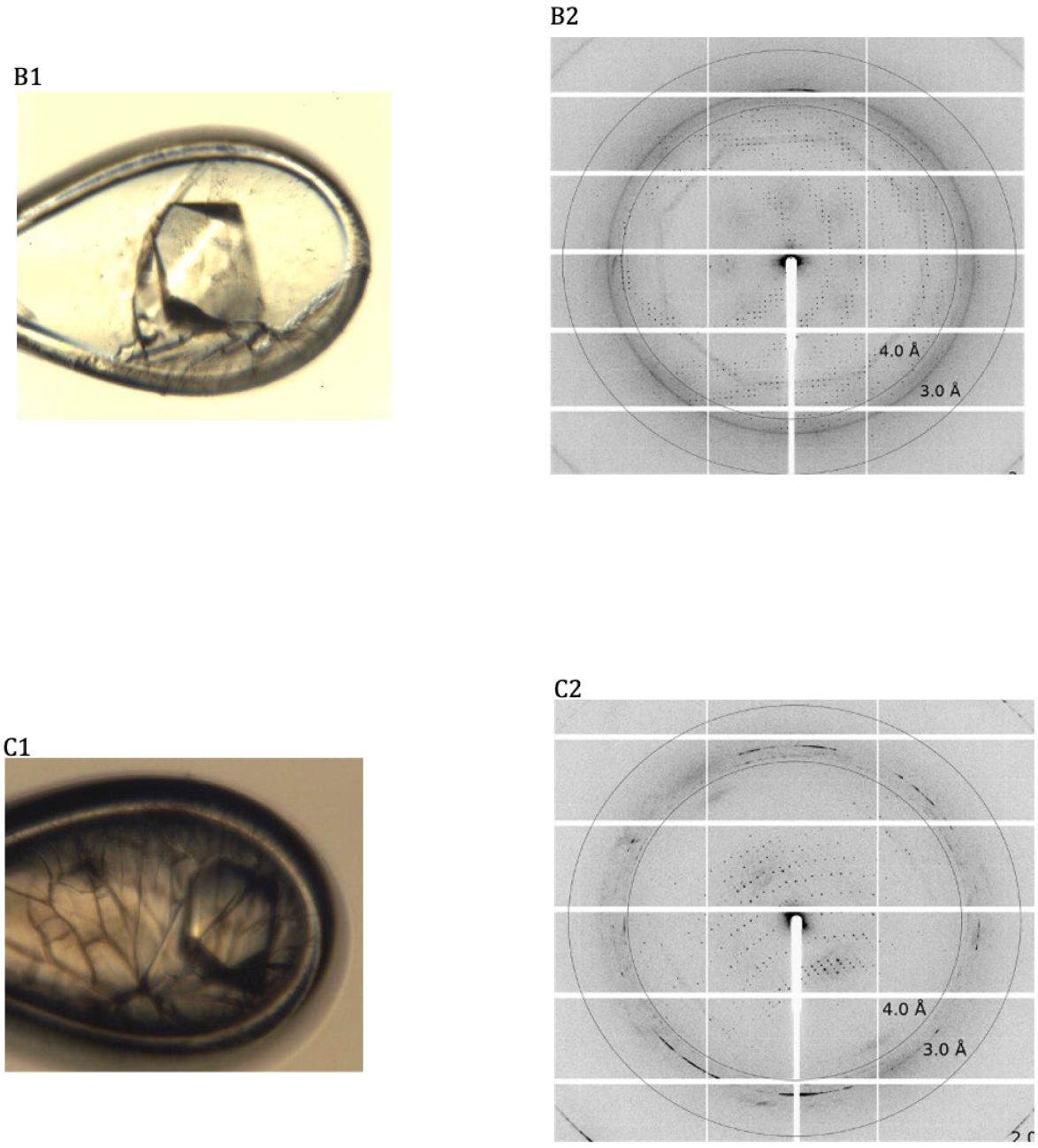


Figure 3.1.6.1.2 Crystals and diffraction images for different mutants

A1: S249D crystal, A2: Diffraction pattern for S249D mutant obtained at 3.02 Å. B1: S249E crystal, B2: Diffraction pattern for S249E mutant obtained at 2.98 Å. C1: T248D crystal, C2: Diffraction pattern for T248D mutant obtained at 3.16 Å.

3.1.6.2 Data collecting and processing

The crystals were diffracted at the European Synchrotron Radiation Facility (ESRF, Grenoble, FRANCE). Before data collection, the MOSFLM program's auto-indexing feature was utilized to determine the crystal's unit-cell parameters and orientation, using four diffraction images taken at a 90° angles apart. The diffraction data revealed a Laue class of 4/mmm, indicating a point group of 422, with potential primitive tetragonal space groups ranging from No. 89 to 96 (Aroyo, 2013). This result was expected, as the CC domain crystal had previously shown the same space group. However, the tetragonal space group was insufficient for completing the structure (Tiruttani Subhramanyam et al., 2014). Consequently, the data were integrated into three different space groups with different symmetry reductions ($P4_32_12$, $P2_12_12_1$, and P1) to analyze if, for parts of the molecule, the crystallographic symmetry may be reduced.

The most likely point group and possible space group were determined using Aimless program from CCP4 suite (Evans, 2011). The space group $P4_32_12$ was suggested for all the integrated data sets. The solvent content for each space group was calculated using the MATTHEWS_COEF program from the CCP4 suite, based on the unit cell volume and the molecular weight of the protein (Matthews, 1968). The solvent content was calculated to be 47.8%, corresponding to seven molecules per asymmetric unit in the space group $P4_32_12$. The data statistics are in table (3.1.6.2.1), table (3.1.6.2.2) and table (3.1.6.2.3).

The molecular replacement was carried using the CC domain model in MOLREP program (Vagin & Teplyakov, 1997). The model consists of seven chains in which five chains (ABCDE) are long and two chains (FG) are short. Each of them forms homodimeric coiled-coil structures. Chain E forms a crystallographic homodimer with a symmetry equivalent molecule. The model includes residues 254–331 for chains A and C, 255–232 for chains B and E, 256–332 for chain D, and residues 255–294 for chains F and G (figure 3.1.6.2.1).

The models of the mutants were found to be identical to the CC domain model, featuring five long helices and two short ones. However, the helices in the S249D and S249E models were positioned differently within the unit cell due to chains translation and rotation (3.1.6.2.2 A), while the T248D mutant appeared similar to the CC domain model (3.1.6.2.2 B). Despite these differences, the helices of all three mutants formed three non-crystallographic dimers (two long and one short), with the remaining helix forming a crystallographic dimer with a symmetry-equivalent molecule. The helices A, B, C, D, and E showed poor electron density at their N-terminals, whereas helices F and G exhibited poor electron density at both terminal halves, which did not allow model completion.

Several attempts were made to improve the density and complete the model. This included density modification using the DM program and BUCCANEER, a program that performs statistical chain tracing by identifying connected alpha-carbon positions using a likelihood-based density target (Cowtan, 2006) from CCP4 suite, as well as the Autobuild program from phenix suite. However, the built models were fragmented and did not include new structures compared to the MOLREP results.

Another attempt included examining a lower symmtery space groups, i.e. $P2_12_12_1$ and P1. The resulting models, however, were similar to the $P4_32_12$ models except that they had 10 long helices and 4 short helices in space group $P2_12_12_1$, and 40 long helices and 16 short helices in P1 space group, which is reasonable considering their non-isomorphic subgroup relation with $P4_32_12$ space group. Consequently, a crystallographic dimer was not observed in the models from space group $P2_12_12_1$ (figure 3.6.2.1.2 C), whereas 2 crystallographic dimers were obsereved in the models from space group P1.

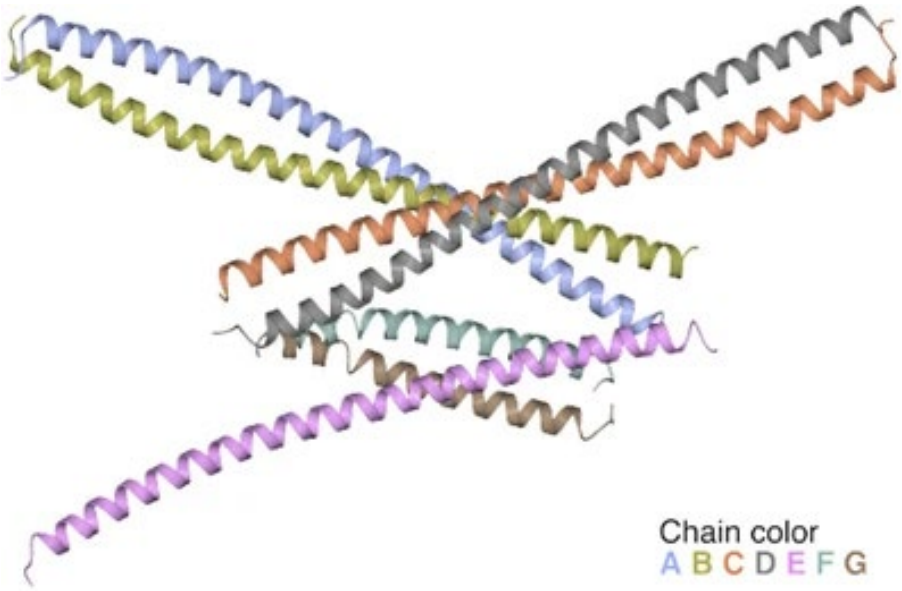


Figure 3.1.6.2.1 Asymmetric unit of the CC domain crystal (figure adopted from (Tiruttani Subhramanyam et al., 2017)

The asymmetric unit contains seven molecules, which are shown with different colors. Each of them forms homodimeric coiled-coil structures. Chain E forms a crystallographic homodimer with a symmetry equivalent molecule.

Table 3.1.6.2.1 Data processing statistics of mutant T248D for space groups P43212, P212121 and P1

Mutant	**T248D**		
Wavelength (Å)	0.87313		
Resolution range (Å)	84.11-3.40		
Unit cell a, b, c	115.12, 115.12, 123.20		
α, β, γ (°)	90, 90, 90		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Total reflections	150170	150775	152623
Unique reflections	11938	23149	85700
Multiplicity	12.6	6.5	1.8
Completeness (%)	100.0	99.9	98.6

Mean I/sigma(I)	10.9	13.2	4.9
R-merge	0.126	0.066	0.073
R-meas	0.132	0.071	0.104
R-pim	0.037	0.028	0.073
CC1/2	0.994	0.999	0.974

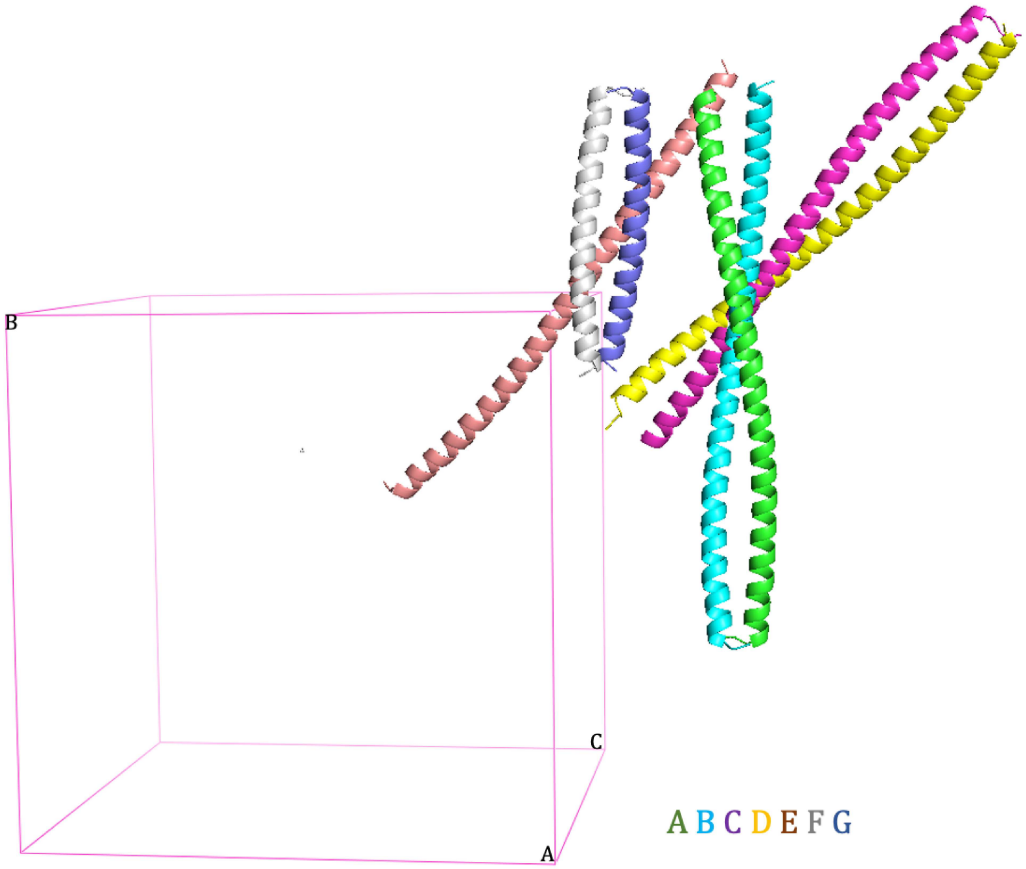
Table 3.1.6.2.2 Data processing statistics of mutant S249D for space groups P43212, P212121 and P1

Mutant	**S249D**		
Wavelength (Å)	0.96770		
Resolution range (Å)	81.44-3.40		
Unit cell a, b, c	115.18, 115.18, 122.96		
α, β, γ (°)	90, 90, 90		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Total reflections	153471	153533	153731
Unique reflections	11913	23104	79303
Multiplicity	12.9	6.6	1.9
Completeness (%)	99.9	99.9	91.3
Mean I/sigma(I)	18.4	13.0	4.4
R-merge	0.08	0.078	0.061
R-meas	0.084	0.085	0.086
R-pim	0.024	0.033	0.061
CC1/2	0.998	0.997	0.992

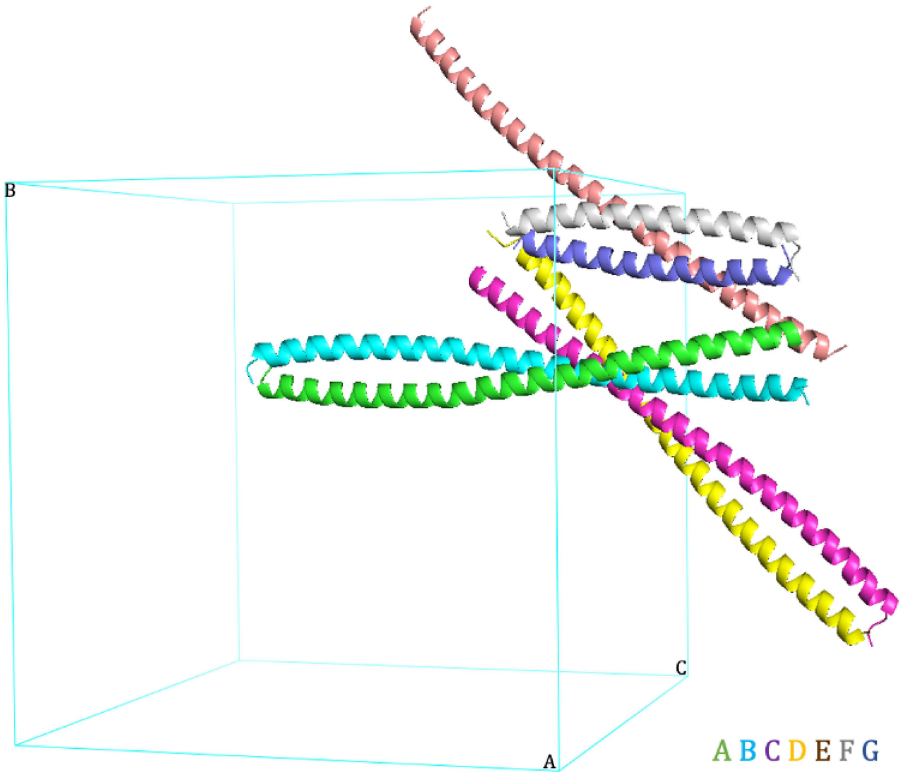
Table 3.1.6.2.3 Data processing statistics of mutant S249E for space groups P43212, P212121 and P1

Mutant	**S249E**		
Wavelength (Å)	0.87313		
Resolution range (Å)	84.77-3.70		
Unit cell a, b, c	114.73, 114.73, 125.81		
α, β, γ (°)	90, 90, 90		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Total reflections	116575	116612	118110
Unique reflections	9458	18276	67115
Multiplicity	12.3	6.4	1.8
Completeness (%)	100.0	99.9	98.1
Mean I/sigma(I)	7.1	9.0	2.6
R-merge	0.139	0.128	0.154
R-meas	0.145	0.139	0.217
R-pim	0.041	0.055	0.154
CC1/2	0.997	0.992	0.873

A.



B.



C.

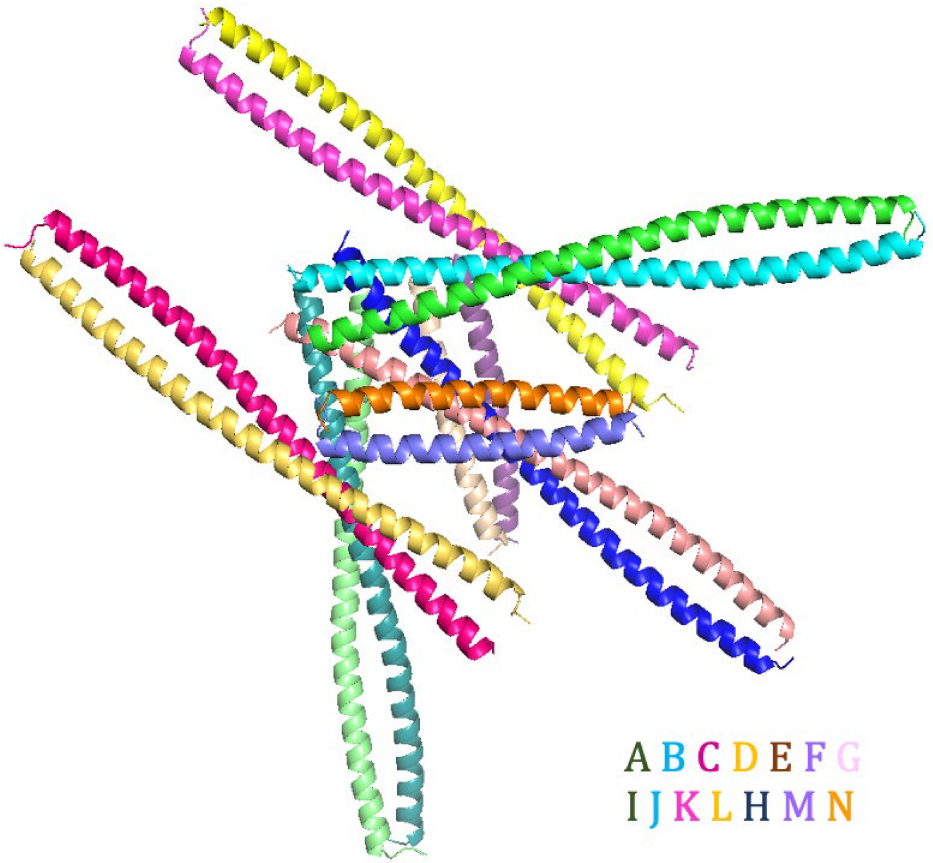


Figure 3.1.6.1.2 Models of mutants T248D, S249D and S249E in space groups P43212 and P212121.

A. Model of mutant T248D in space group $P4_32_12$. B. Model of mutant S248D and S249E in space group $P4_32_12$. C. Model of mutants T248D, S249D and S249E in space group $P2_12_12$.

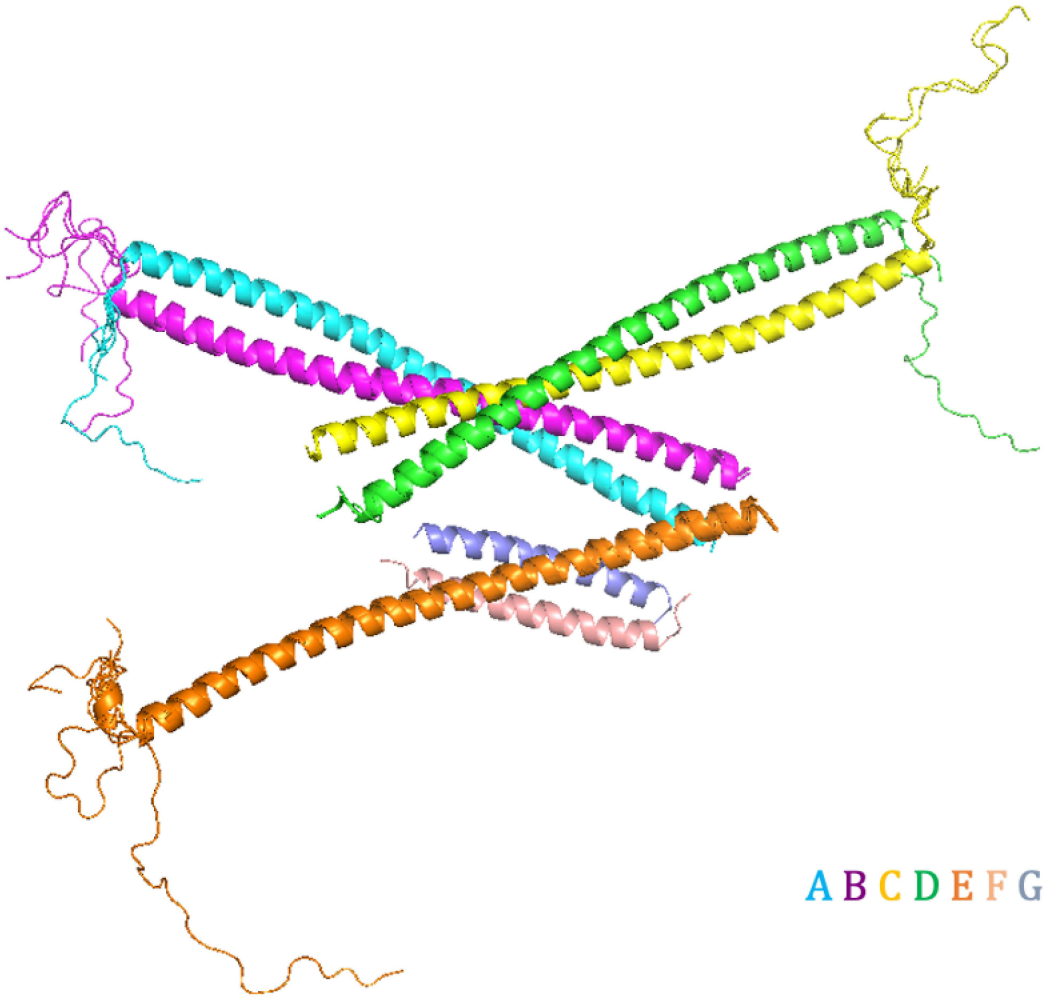
3.1.6.3 Model building and Refinement

Despite numerous attempts to enhance the density at the helices' N-terminal and the C-terminal of the FG dimer, only a few residues of some helices could be constructed as polyalanine at their N-terminals (figures 3.1.6.3.1 A and 3.1.6.3.1 B). These residues were manually built using COOT, a graphical program within the CCP4 suite for model building and refinement, using a contour level of 0.5-0.6 root-mean-square (rms) (figure 3.1.6.3.3). The weak density did not allow for the construction of the side chains of these residues, and among these residues, the phosphoryaltion mimicking residues,i.e., S249D, S249E, and T248D.

The models were refined using phenix software suite (Adams et al., 2010). Different refinement parameters inculding rigid body, atomic displacement parameter (B-factor), simulated annealing, and TLS (Translation/Libration/Screw) were selected prior to the refinement. Torsion-based NCS restraints, which allow local differences between NCS-related molecules, were used for all chains during the refnements.

These refinements are intended to reduce the R-value, which indicates the level of agreement between the observed structure factor amplitudes and those calculated from the model. To prevent overfitting, the model was evaluated using the free R-value, which is calculated from a (free) subset of amplitudes (5%) that have not been used for refinement (Brünger, 1993; Kleywegt & Jones, 1997). The refinement statistics are shown in table 3.1.6.3.1, table 3.1.6.3.2, table 3.1.6.3.3 and table 3.1.6.3.4.

A.



B.

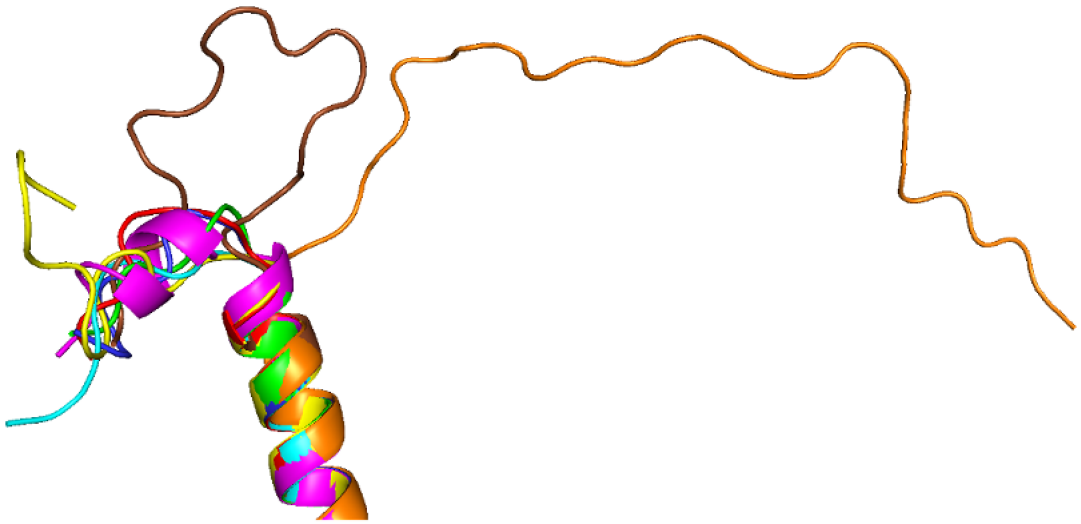


Figure 3.1.6.3.1 Model of mutants T248D, S249D and S249E in space group P1.

A. Model of mutants in P1, few residues were built at the N-terminal of some chains in space group P1 based on the density. B. Fitted N-terminals of E chain with its local symmetry chains.

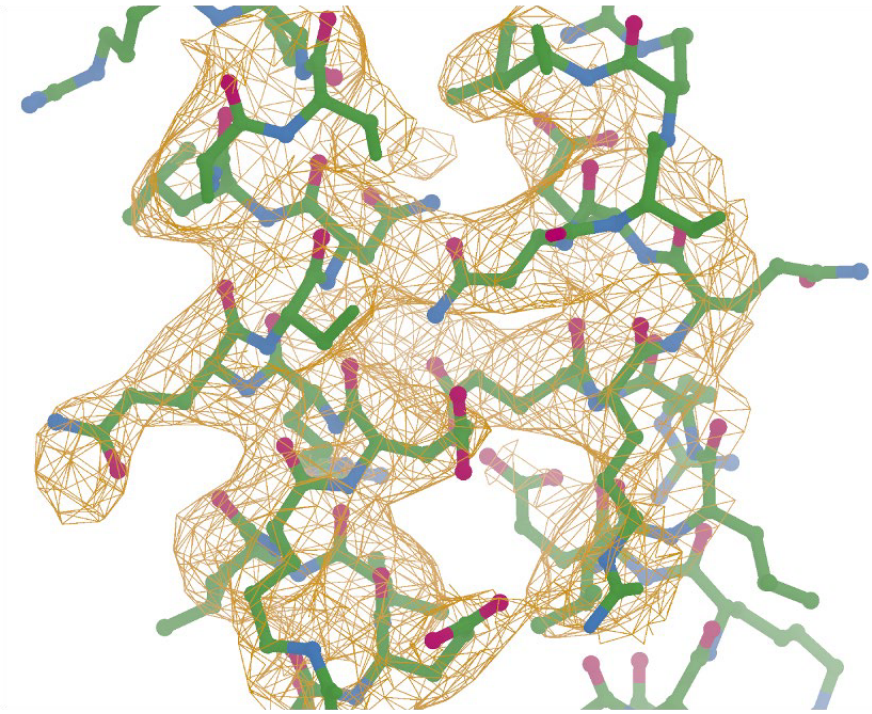


Figure 3.1.6.3.2 Residues of coiled coil structure in space group P1 covered by electron density map

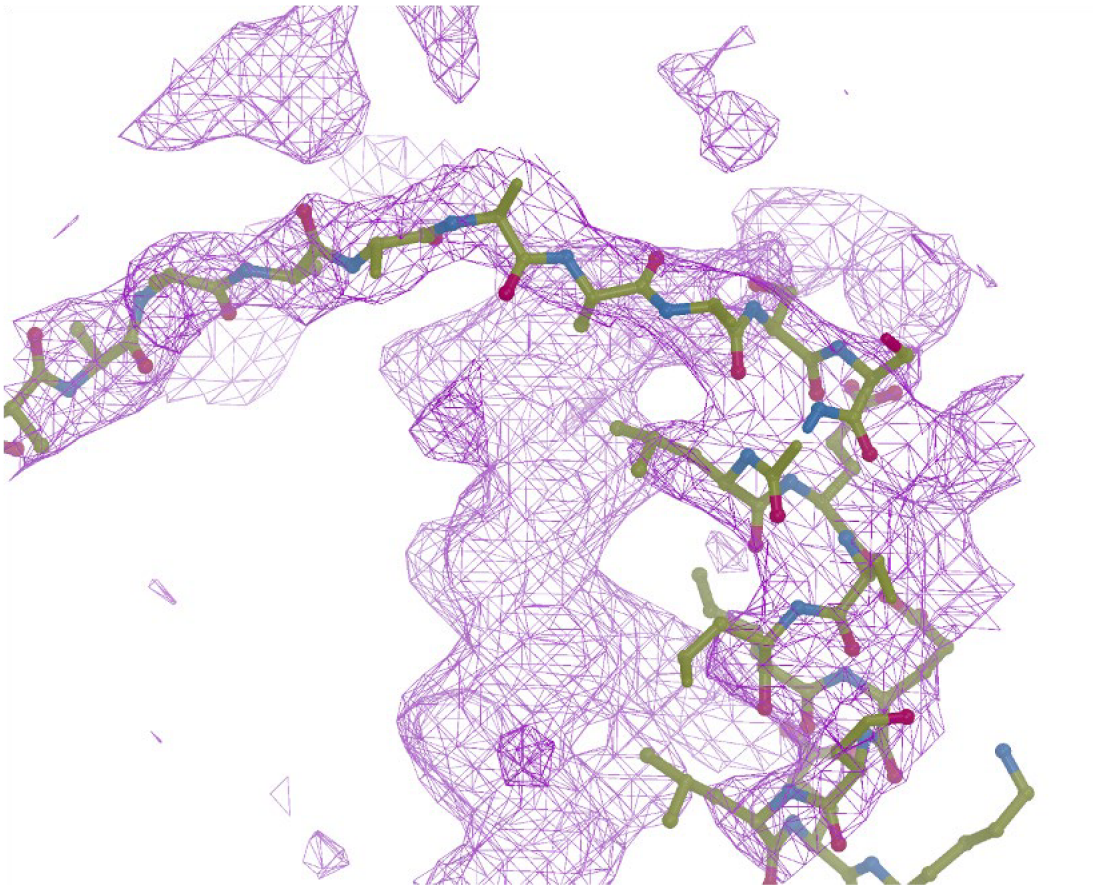


Figure 3.1.6.3.3 Polyalanine residues built at the N-terminus in space group P1 showing weak electron densities

Table 3.1.6.3.1 Model refinement statistics of mutant T248D in different space groups

Mutant	**T248D**		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Reflections used in refinement	11827	23021	85087
Reflections used for R-free	594	1167	4280
R-work	0.2536	0.2669	0.2575
R-free	0.3418	0.3703	0.3444
RMS (bonds)	0.0095	0.0099	0.0103

RMS (angles)	1.270	1.253	1.427
Ramachandran favored (%)	88.02	92.18	90.82
Ramachandran allowed (%)	9.32	6.88	6.62
Ramachandran outliers (%)	2.66	0.94	2.56
Rotamers outliers (%)	1.00	0.67	0.61
Clash score	13.51	16.74	16.89
Average B-factor	118.66	138.12	137.15

Table 3.1.6.3.2 Model refinement statistics of mutant S248D in different space groups

Mutant	**S249D**		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Reflections used in refinement	11843	22998	79060
Reflections used for R-free	593	1166	3890
R-work	0.2633	0.2672	0.2454
R-free	0.3055	0.3572	0.3056
RMS (bonds)	0.0119	0.0094	0.01
RMS (angles)	1.489	1.186	1.366
Ramachandran favored (%)	90.49	94.44	92.01
Ramachandran allowed (%)	6.66	4.24	5.92
Ramachandran outliers (%)	2.85	1.32	2.07
Rotamers outliers (%)	0.20	0.38	0.16
Clash score	14.50	13.27	13.28
Average B-factor	137.39	134.17	123.70

Table 3.1.6.3.3 Model refinement statistics of mutant S248E in different space groups

Mutant	S249E		
Space group	$P4_32_12$	$P2_12_12_1$	P1
Reflections used in refinement	9386	18163	65913
Reflections used for R-free	493	954	3421
R-work	0.2534	0.2985	0.2954
R-free	0.2996	0.3257	0.3081
RMS (bonds)	0.0069	0.003	0.0054
RMS (angles)	1.094	0.506	0.984
Ramachandran favored (%)	94.87	97.08	93.99
Ramachandran allowed (%)	3.42	2.92	4.56
Ramachandran outliers (%)	1.71	0.00	1.45
Rotamers outliers (%)	0.20	0.19	0.13
Clash score	7.03	5.87	8.6
Average B-factor	159.05	150.19	120.00

Table 3.1.6.3.4 R-work and R-free of T248D, S249D and S249E in space groups P43212, P212121 and P1

	T248D		S249D		S249E	
	R-work	**R-free**	**R-work**	**R-free**	**R-work**	**R-free**
$P4_32_12$	0.2536	0.3418	0.2633	0.3055	0.2534	0.2996
$P2_12_12_1$	0.2669	0.3703	0.2672	0.3572	0.2985	0.3257
P1	0.2575	0.3444	0.2454	0.3056	0.2954	0.3081

Different attempts were made to improve the density at the N-terminals of the chains and the C-terminals of the short helices, but none were successful. Integrating the data in space group P1 was chosen to better understand the structure at the N-terminus. The N-terminals of the chains were elongated differently based on the density, and only polyalanine residues were built due to missing density for the side chains.

When the local symmetries were mapped on their chains (Figure 3.1.6.3.1), the elongated N-terminals did not match, explaining why in higher symmetry models for the missing parts of the structure could not be obtained. This indicates that the N-terminal part of the CC domain is disordered.

Additionally, the CD spectroscopy results showed that all the proteins, including the native one, have a 100% helical structure in solution. However, in the crystal structure, the N-terminus could not be obtained in higher symmetry due to the disorder, highlighting that multiple conformations exist within the crystals for the N-terminus

In conclusion, the phosphorylation-mimicking mutations did not significantly enhance the folding and stability of the chains' N-terminals.

3.2 Structural Characterization of the Extended Region from the CC Domain to the SAC Domain (Par-4 CC EXS)

Since the phosphorylation-mimicking and salt bridge mutations did not reveal a distinctive structure beyond the CC domain or improve the structure of the short helices (i.e., FG chain), another hypothesis was proposed to gain a better understanding of the crystal structure of the CC domain and the region outside it.

Par-4 has been linked to apoptosis. However, the active form of Par-4 that induces apoptosis in cells is the region encompassing residues 124-332, resulting from the cleavage of Par-4 at the site EEPD123↓S by Caspase 3. This region includes both the SAC domain and the CC domain, as well as the linker between them. This suggests that this fragment should be able to induce apoptosis if it is properly structured, post-translationally modified, or bound to partner protein(s).

Based on this, another hypothesis was formulated by extending the region from the CC domain to include the SAC, assuming that this extension could stabilize and help in characterizing the structure of the CC domain and the region beyond it. The new construct is called Par-4 CC EXS.

To achieve this, the insect cell expression system was utilized alongside the bacterial expression system, as it can offer post-translational modifications that aid in folding and stabilizing the protein.

This section presents the biophysical and structural characterization results of Par-4 CC EXS.

3.2.1 Cloning of Par-4 CC EXS

The residues corresponding to the region 136-332 (named short extended region from coiled coil domain to SAC domain, CC EXS) (Figure 3.2.1.1) were cloned into pET11a and pOPIE2 vectors for protein expression in bacterial and insect cell systems, respectively. The Strep tag sequence was incorporated in the forward primer, while the reverse primer included a stop codon. Several random colonies were picked for plasmid amplification. Plasmids were checked by enzyme double digestion with respective enzymes and PCR. Once the bands corresponding to the insert size were visualized by agarose gel, the plasmids were sent for Sanger sequencing. Microsynth's Sanger sequencing with T7 forward and T7 reverse primers were used to verify the nucleotide sequence in the pET11a plasmid, while OPIE2 forward and IE1 reverse primers were used for nucleotide sequence verification in the pOPIE2 plasmid.

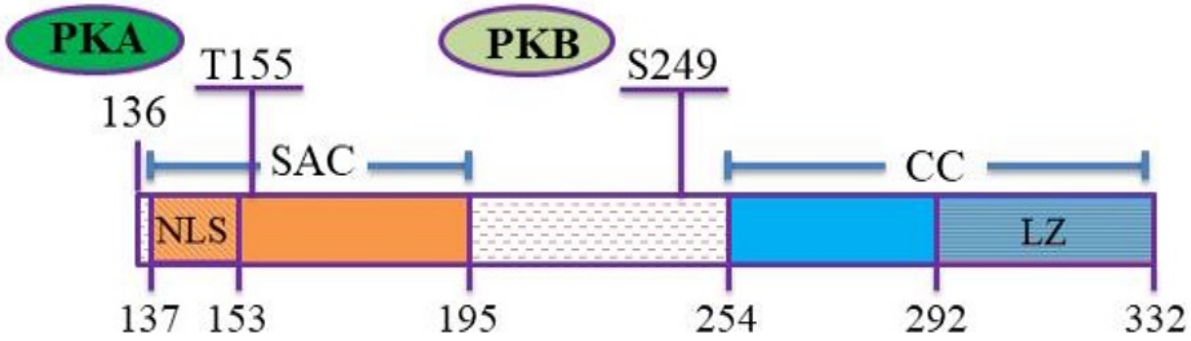


Figure 3.2.1.1 Schematic diagram for Par-4 CC EXS construct

The Par-4 CC EXS DNA insert was effectively cloned in its respective vectors, which were then utilized to overexpress the protein in E. coli and insect cell expression systems.

3.2.2 Protein expression

3.2.2.1 Expression of Par-4 CC EXS in E. coli

The expression of Par-4 CC EXS in E. coli was assessed for the amount and solubility of the protein using different temperatures (37 °C, 30 °C, 20°C, and 16 °C), strains (RIL and RP strains of E. coli BL21(DE3)), and IPTG concentrations (0.4 mM, 0.6 mM, 0.8 mM, and 1.0 mM). The solubility of the protein was evaluated by lysing small amount of the cell culture grown at each condition. The cells were lysed on ice with a sonicator set to 65% amplitude for 30 seconds on and 1 minute off for 10 cycles. After that, the samples were centrifuged at a speed 1000 x g for 10 minutes to pellet down the unopened cells, and the supernatants were additionally centrifuged at 10000 x g for a period of 30 minutes to pellet down the inclusion bodies. Following, the cell pellet (unopened cell), the inclusion bodies and the supernatant were analyzed using SDS-PAGE.

According to the results obtained, it was found that at temperatures of 37°C and 30°C, the protein was mostly insoluble, whereas at 20°C and 16°C, there was a shift in solubility, with the highest solubility observed at 16°C. Comparing the expression level and solubility in RIL and RP strains, the protein expressed in RIL was slightly higher than in RP. The cells were induced with different IPTG concentrations when the OD600 reached 0.8-0.9. IPTG results did not show a significant change in protein amount and solubility, with a minor increase when 0.6 mM was added.

In conclusion, in accordance with the findings presented above, Par-4 CC EXS was overexpressed in the RIL strain of E. coli BL21(DE3) cells by inducing expression using a final concentration of 0.6 mM IPTG and incubating at 16°C, 120 rpm for 18 hours. Several liters of cultures were grown, and the cells were harvested by centrifugation at 5000 x g for 15 minutes. The cell pellets were stored at -80°C for future use.

3.2.2.2 Expression of Par-4 CC EXS in Insect cell

The expression of Par-4 CC EXS in Hi5 insect cell was done according to the published protocol by Bleckmann et al (Bleckmann et al., 2019). Briefly, a suspension culture with a cell density of 4×10^6 cells/mL was prepared, and DNA and PEI were added separately to the culture at ratios of 1 µg DNA per 1×10^6 cells and 1 µg DNA to 4 µg PEI. After cell transfection, samples were taken from the cell culture at different time intervals (0 h, 24 h, 36 h, 48 h, and 60 h) to assess protein expression levels and cell viability. Figure 3.2.2.2.1 shows the protein expression levels at different time intervals post-transfection in Hi5 insect cells. After 60 h post-transfection, cell viability was found to be 28%, whereas after 48 h post-transfection, it was 57%. Therefore, harvesting cells after 60 h was excluded due to low cell viability. Regarding the amount of protein expressed, there was a slight difference in protein expression levels between the 36 h and 48 h intervals, except that viability decreased. Based on these results and considering cell viability, the cells were harvested 48 h post-transfection at 4000 x g for 15 minutes, and the cell pellet was stored at -80°C until the start of purification.

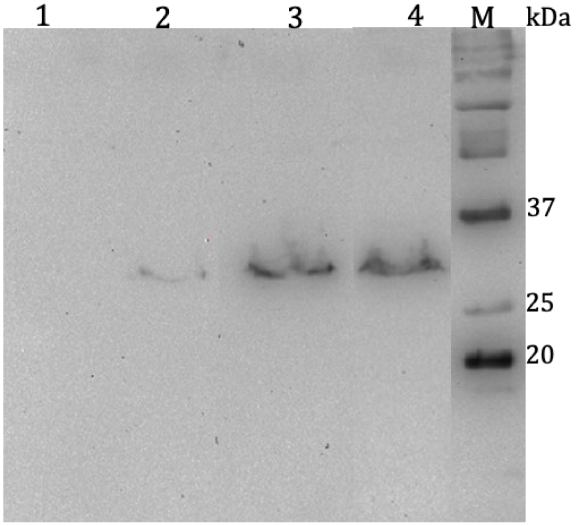


Figure 3.2.2.1 Anti-strep western blot for Time-dependent expression Profile for Par-4 CC EXS in Hi5 cells

1. Non-transfected Hi5 cells, 2. 24 hours post-transfection, 3. 36 hours post-transfection, 4. 48 hours post-transfection. M: Precision Plus Protein™ WesternC™ Blotting Standards.

3.2.3 Proteins purification

Regardless of the expression system used to express the protein, the same purification protocol was followed. After cell lysis, the obtained supernatant was used for subsequent protein purification. The purification process was based on affinity chromatography using Strep-Tactin XT resin, which binds specifically to the fused Strep-tag in the proteins. The affinity-purified protein was then subjected to size exclusion chromatography (SEC) to separate the target proteins from any contaminant proteins and to obtain highly purified protein.

Depending on the pH condition used to elute the protein, two elution buffers (elution buffer 1 and elution buffer 2) were used in the affinity chromatography, followed by two SEC buffers (SEC buffer 1 and SEC buffer 2) (Table 2.3.3) in the gel filtration chromatography. The CC domain and the mutants were purified under acidic conditions, while Par-4 CC EXS was purified under both acidic and neutral pH conditions.

3.2.3.1 Purification of Par-4 CC EXS under acidic condition

3.2.3.1.1 Purification of Par-4 CC EXS expressed in E. coli

Figure 3.2.3.1.1.1 A shows the SDS-PAGE of the SEC elution fractions obtained from gel filtration under acidic conditions for the protein expressed in E. coli. The figure shows four bands with different molecular sizes: band 1, band 2, band 3, and band 4, with estimated molecular sizes of 40 kDa, 34 kDa, 18 kDa, and 16 kDa, respectively. None of these molecular sizes match the theoretical molecular weight of Par-4 CC EXS, which is 24.0838 kDa.

The identity of these bands was checked by western blot using a Strep-tag antibody. The western blot (Figure 3.2.3.1.1.1 B) showed that band 1 and band 2 were western blot positive, whereas band 3 and band 4 were western blot negative.

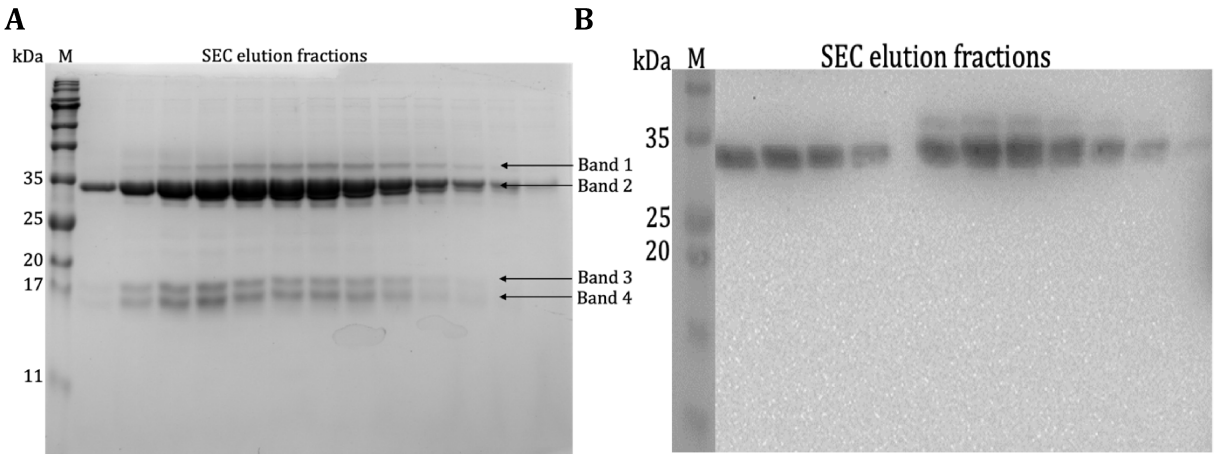


Figure 3.3.2.1.1.1 SDS-PAGE and Anti-strep western of Par-4 CC EXS from E. coli

A. Blue silver stained SDS-PAGE of SEC elution fractions for Par-4 CC EXS from E. coli. B. Anti-strep western blot for SEC elution fractions for Par-4 CC EXS from E. coli.

Obtaining a highly purified protein is essential for any further biophysical and structural investigations. Therefore, several approaches were attempted to eliminate these lower bands. However, none of these approaches successfully removed the bands.

3.2.3.1.2 Purification of Par-4 CC EXS expressed in Hi5 insect cells

Figure 3.2.3.1.2.1 A shows the SDS-PAGE of the SEC elution fractions obtained from gel filtration under acidic conditions for the protein expressed in Hi5 insect cells. Interestingly, the same bands with the same molecular sizes, except for band 1, were observed in the Par-4 CC EXS purification from Hi5 insect cells. The band with an estimated molecular size of 34 kDa was the only western blot positive band when an anti-Strep tag western blot was applied (Figure 3.2.3.1.2.1 B).

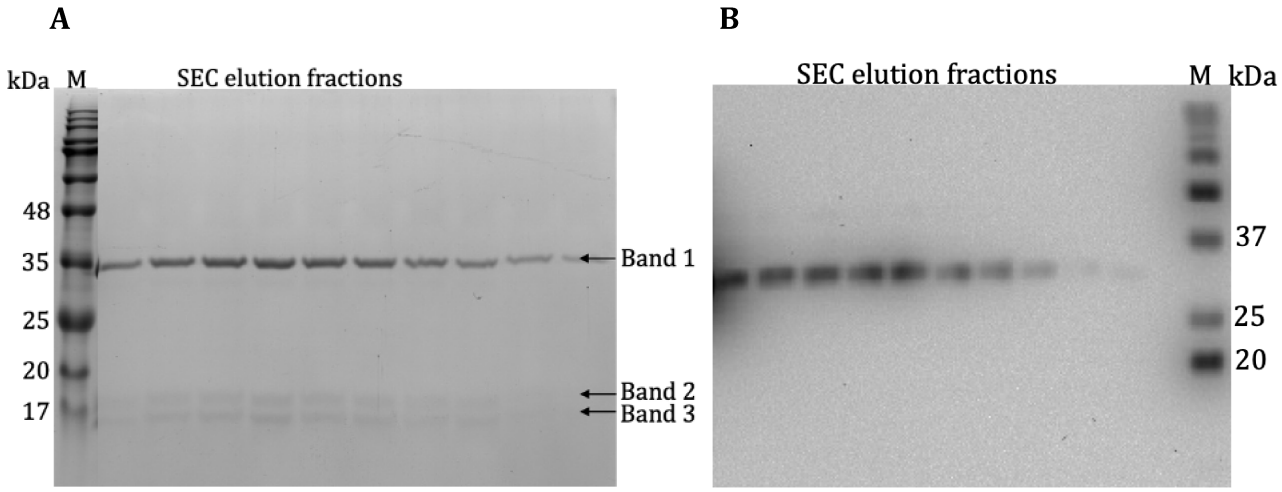


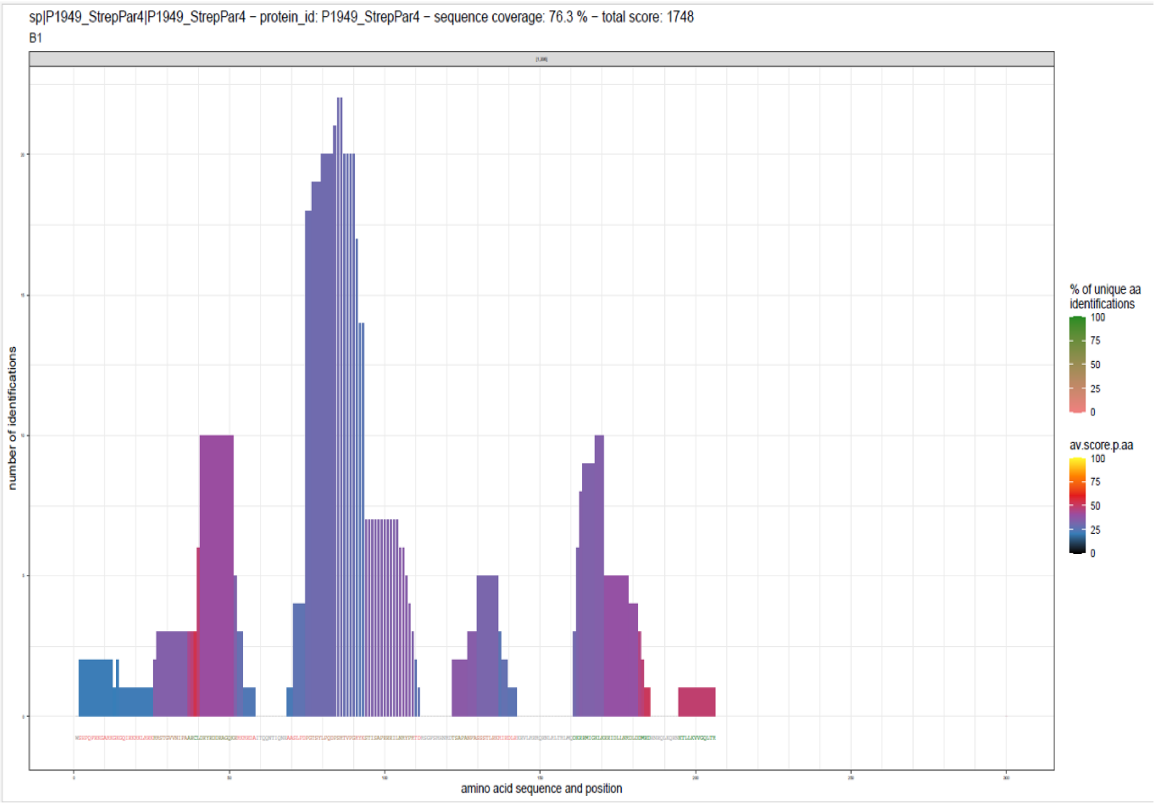
Figure 3.2.3.1.2.1 SDS-PAGE and Anti-strep western of Par-4 CC EXS from Hi5 insect cell

A. Blue silver stained SDS-PAGE of SEC elution fractions for Par-4 CC EXS from Hi5 insect cells B. Anti-strep western blot for SEC elution fractions for Par-4 CC EXS from Hi5 insect cells

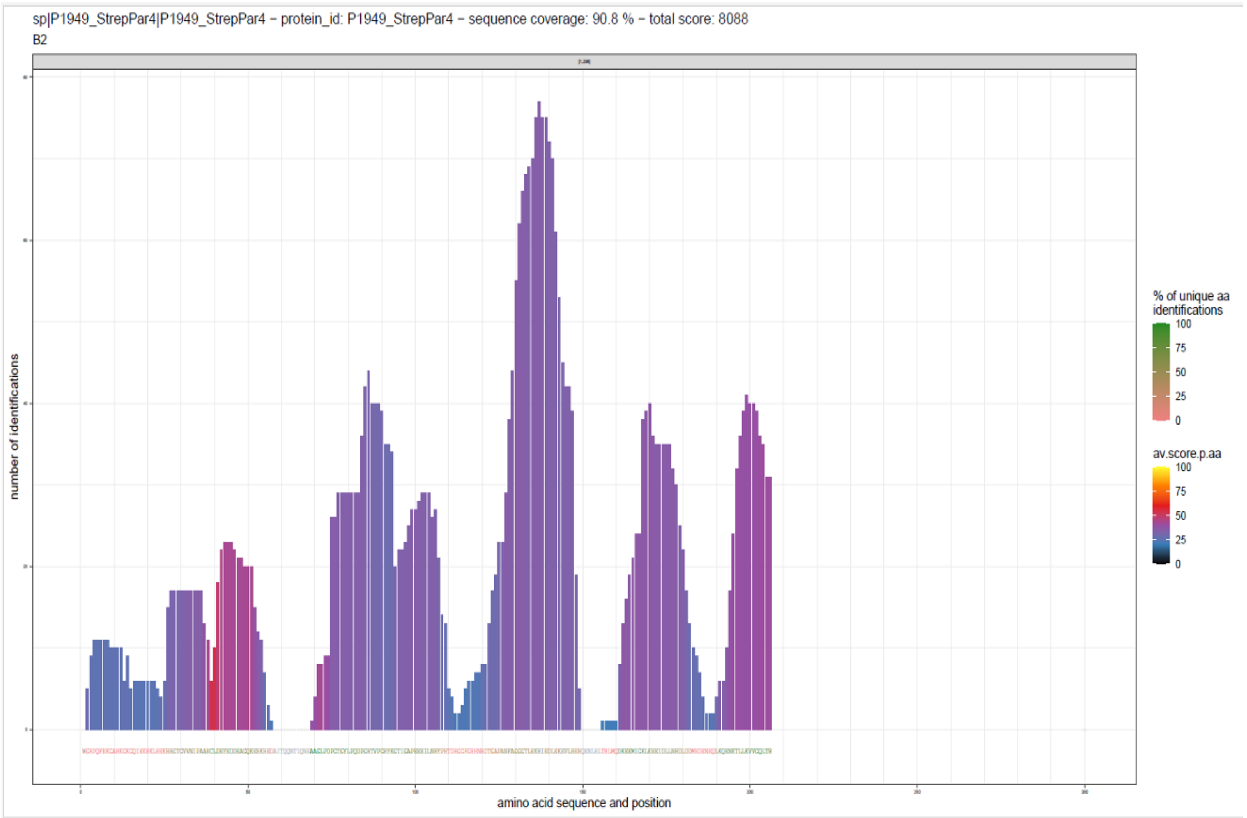
3.2.3.1.3 Protein bands identification

To identify the lower bands (18 kDa and 16 kDa) as well as the western-positive bands, the SDS-PAGE gels for both purifications were sent to the EMBL-Heidelberg Proteomics Core Facility for LC-Mass/Mass analysis. An in-gel acid hydrolysis workflow was used, as hydrochloric acid cleaves proteins arbitrarily into their constituent amino acids, aiding in the identification of the amino acid sequence. The generated amino acids were analyzed using the Lumos system. Figures 3.2.3.1.3.1 A-E show the results obtained from the LC-Mass/Mass analysis. The bands with estimated molecular sizes of 40 kDa and 34 kDa appear to be the full length of Par-4 CC EXS, whereas the bands with estimated molecular sizes of 18 kDa and 16 kDa are degraded Par-4 CC EXS protein. The results also showed that the degradation occurred at the same peptide, between the aspartic acid (D) at position 209 and proline (P) at position 210, for both proteins expressed in E. coli and Hi5 cells.

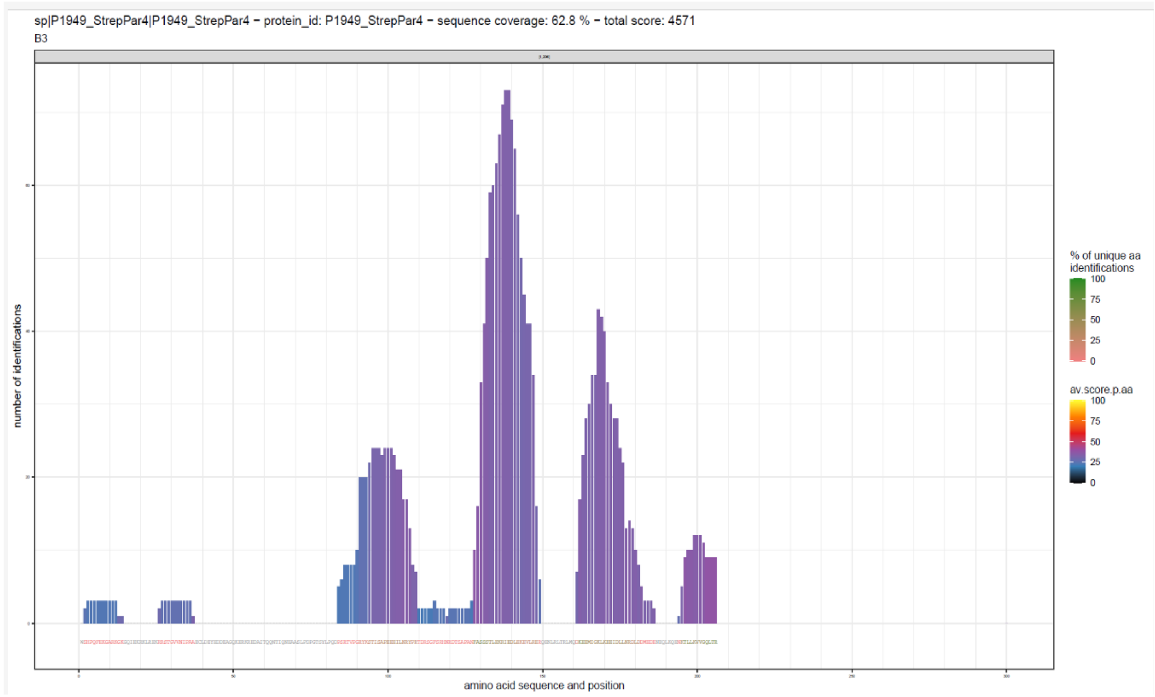
A


sp|P1949_StrepPar4|P1949_StrepPar4 – protein_id: P1949_StrepPar4 – sequence coverage: 76.3 % – total score: 1748
B1
number of identifications
amino acid sequence and position
% of unique aa identifications
100
75
50
25
0
av.score.p.aa
100
75
50
25
0

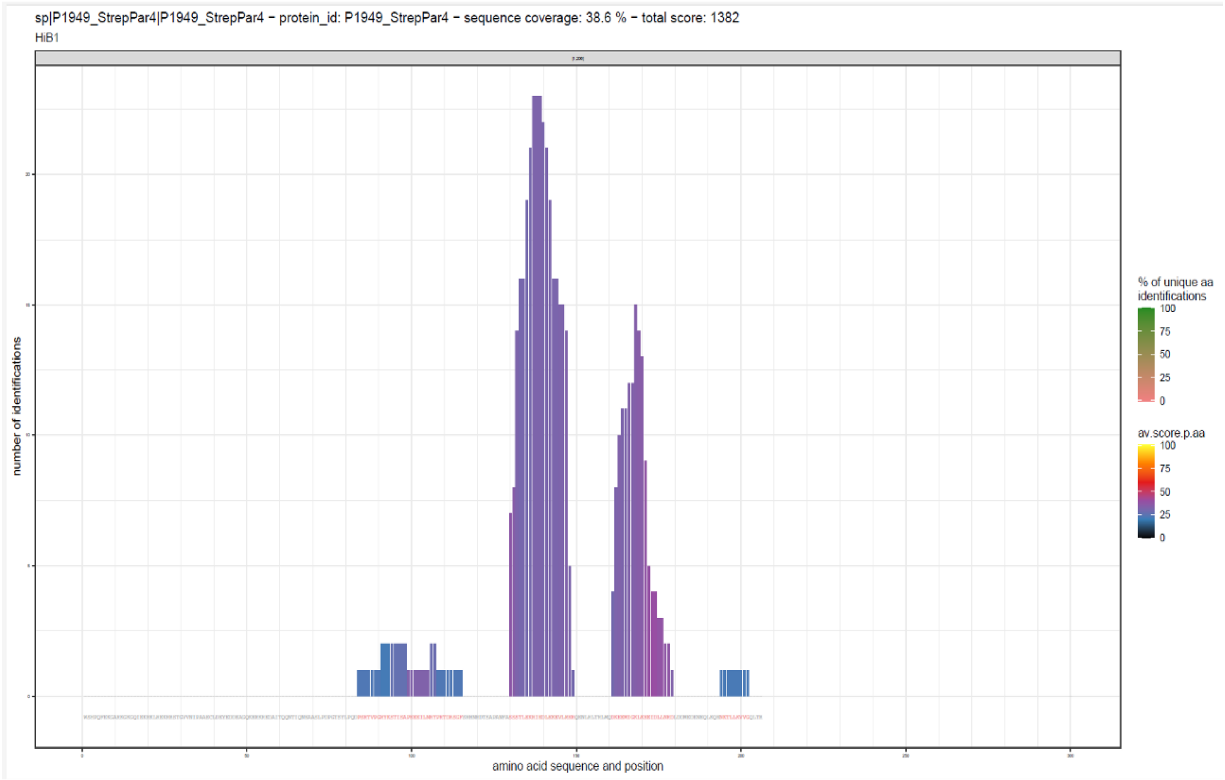
B.


sp|P1949_StrepPar4|P1949_StrepPar4 – protein_id: P1949_StrepPar4 – sequence coverage: 90.8 % – total score: 8088
B2
number of identifications
amino acid sequence and position
% of unique aa identifications
100
75
50
25
0
av.score.p.aa
100
75
50
25
0

C.


sp|P1949_StrepPar4|P1949_StrepPar4 – protein_id: P1949_StrepPar4 – sequence coverage: 62.8 % – total score: 4571
B3
number of identifications
amino acid sequence and position
% of unique aa identifications
100
75
50
25
0
av.score.p.aa
100
75
50
25
0

D.


sp|P1949_StrepPar4|P1949_StrepPar4 – protein_id: P1949_StrepPar4 – sequence coverage: 38.6 % – total score: 1382
HiB1
number of identifications
amino acid sequence and position
% of unique aa identifications
100
75
50
25
0
av.score.p.aa
100
75
50
25
0

E.

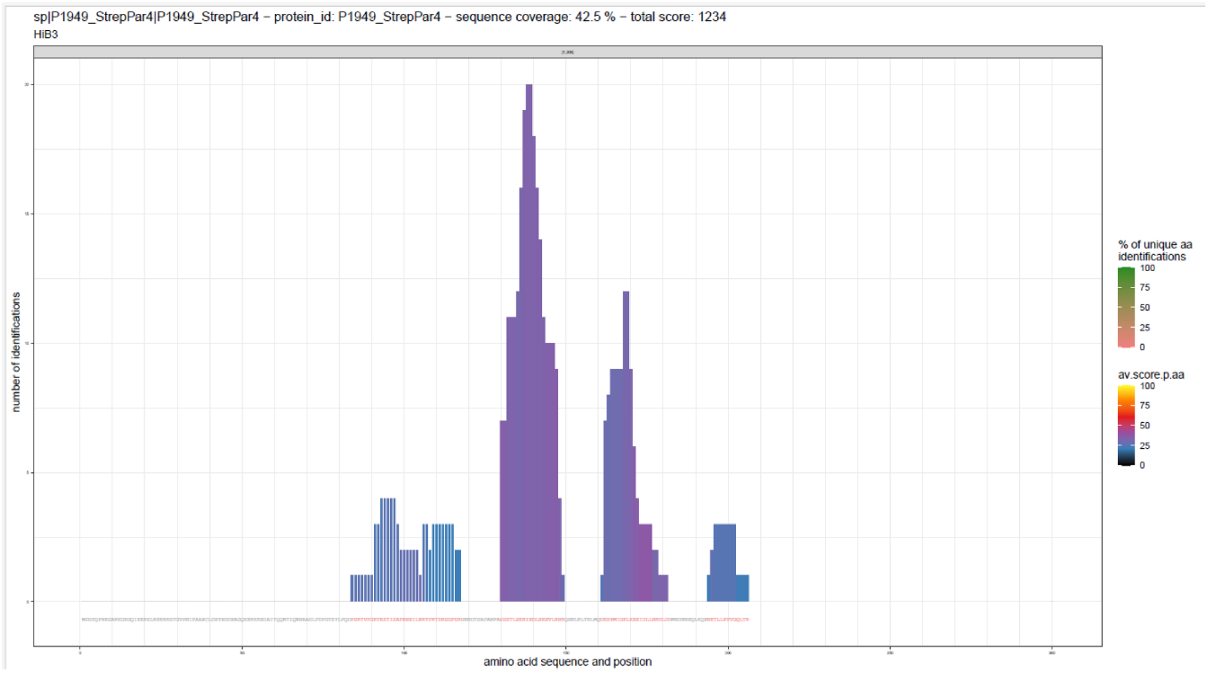


Figure 3.2.3.1.3.1 Par-4 CC EXS bands identification by LC-MS/MS analysis

LC-MS/MS gave N terminal end to C terminal end sequence coverage. A. Band 1 from E. coli B. Band 2 from E. coli. C. Band 3 from E. coli. D. Band 1 from Hi5. E. Band 3 from Hi5, where, Y-axis is number of identifications and X-axis is amino acid sequence and position.

To identify the reason for degradation, the amino acid sequence of Par-4 CC EXS was analyzed for cleavage using the online software PeptideCutter prediction tool from Expasy (Gasteiger et al., 2005).

The prediction indicated that the cleavage of the DP peptide bond is due to the presence of formic acid. But, formic acid was not used at any step during purification. However, it could suggest that the acidic pH itself, rather than formic acid, is the cause of cleavage since both the elution buffer and SEC buffer have a pH of 3.50.

To confirm that acidolysis is the reason for cleavage, Par-4 CC EXS was eluted from the affinity column under neutral pH of 7.5. Figure 3.2.3.1.3.2 shows the SDS-PAGE of the affinity elution fractions at neutral pH 7.5. Indeed, no degradation of the protein

was observed, which supports the hypothesis that acidolysis was the cause of the cleavage.

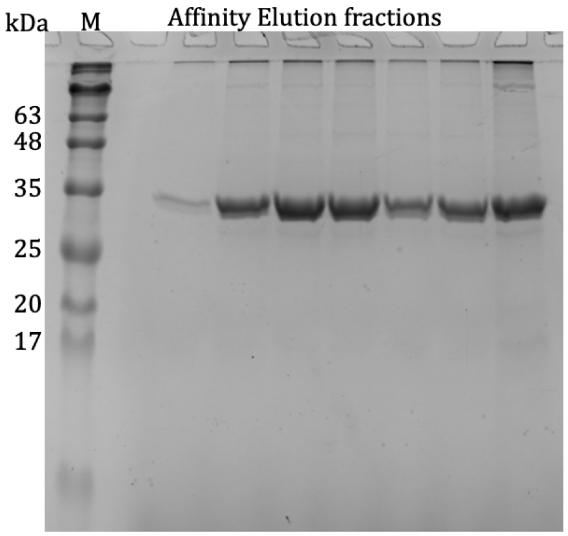


Figure 3.2.3.1.3.2 SDS-PAGE of Par-4 CC EXS affinity elution fractions at pH 7.5.

Moving forward with this project required preserving the protein's integrity, which could only be achieved by either purifying the protein at a neutral pH of 7.5 or utilizing the site-directed mutagenesis technique.

3.2.3.2 Purification of Par-4 CC EXS under neutral condition

The SEC profile (Figure 3.2.3.2.1) reveals that the protein elutes with two apparent molecular weights (32.9 kDa and 16.2 kDa). A molecular weight of 32.9 kDa suggests the protein exists as a dimer, while the 16.2 kDa weight suggests the protein exists as monomer. The deviation from the expected molecular weight could be explained by a non-globular shape.

The amount of purified protein obtained from SEC was only sufficient for biophysical characterization, such as CD and DLS, but not for setting crystallization plates. When the corresponding SEC elution fractions were concentrated to achieve a suitable concentration for CD measurement, some protein aggregated, and the maximum concentration reached was 0.25 mg/ml. This suggests that neutral conditions are

unfavorable for the protein, as both solubility and folding (to be discussed later) decrease, leading to protein aggregation.

Based on these results and observations, purifying the protein under acidic conditions was the only choice to obtain a sufficient amount with a reasonable concentration for setting crystallization plates and further structural studies.

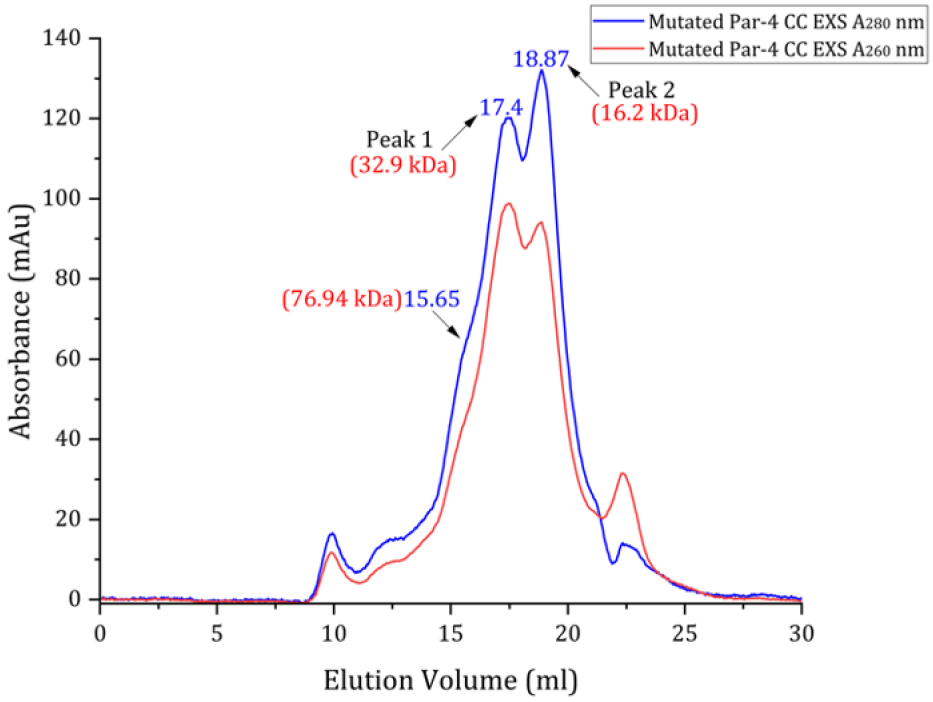


Figure 3.2.3.2.1 Superdex 200 increase SEC profile of Par-4 CC EXS from E. coli under neutral condition

The other approach in moving forward with this protein (Par-4 CC EXS) involved using site-directed mutagenesis to replace an aspartic acid (D) residue with a glutamic acid residue (E) in the peptide DP. In the Par-4 CC EXS protein, there are two aspartic acid residues followed by proline located at positions 200 and 209 (numbered according to rat full-length Par-4). While mass spectrometry results only showed cleavage at position 209, which was the N-terminal of the degraded protein, there was no direct evidence that position 200 is acidolysis resistant, despite being predicted as a cleavage site by acidolysis through the Expasy peptide cutter website. Therefore, three mutated Par-4 CC EXS proteins were created: one with the mutation at position

200 only, one at position 209 only, and one with mutations at both 200 and 209 by changing D to E.

Purification results for the mutated protein from both expression systems under acidic conditions showed that only the mutation at position 209 was sufficient to make the protein resistant to acidic cleavage. This indicates that this acidolysis site is likely exposed to the solvent. The subsequent results discussed are related to the mutated Par-4 CC EXS D209E.

3.2.3.3 Purification of mutated Par-4 CC EXS under acidic condition

3.2.3.3.1 Purification of mutated Par-4 CC EXS expressed in E. coli

Once the D209E mutation was confirmed by DNA sequencing, the protein was expressed and purified as described in section 3.2.3.1. SDS-PAGE was used to assess the purity of the protein after affinity purification and size exclusion chromatography (SEC). Figure 3.2.3.3.1.1 A shows the SDS-PAGE of the purification profile obtained from affinity chromatography under acidic conditions for the mutated protein expressed in E. coli. Figure 3.2.3.3.1.1 B shows the SDS-PAGE of the SEC elution fractions obtained from gel filtration under acidic conditions for the mutated protein expressed in E. coli. The theoretical molecular weight of the mutated protein is 24.09783 kDa. The SDS-PAGE for the SEC fractions shows a single band that ran at a molecular weight of around 34 kDa.

Figure 3.2.3.3.1.2 shows the SEC chromatogram under acidic conditions for the mutated protein expressed in E. coli. The SEC profile indicates that the mutated protein elutes with two apparent molecular weights (189.4 and 100.4 kDa), corresponding to an octamer (24.09783 kDa * 8 = 192.8 kDa) and a tetramer (24.09783 kDa * 4 = 96.4 kDa), with a higher presence of the octamer.

In summary, a pure mutated Par-4 CC EXS protein was produced with a yield of 7.5 mg from a 4.5 g cell pellet obtained from 1 L of culture. The purified protein was used for biophysical characterization and setting crystallization plates.

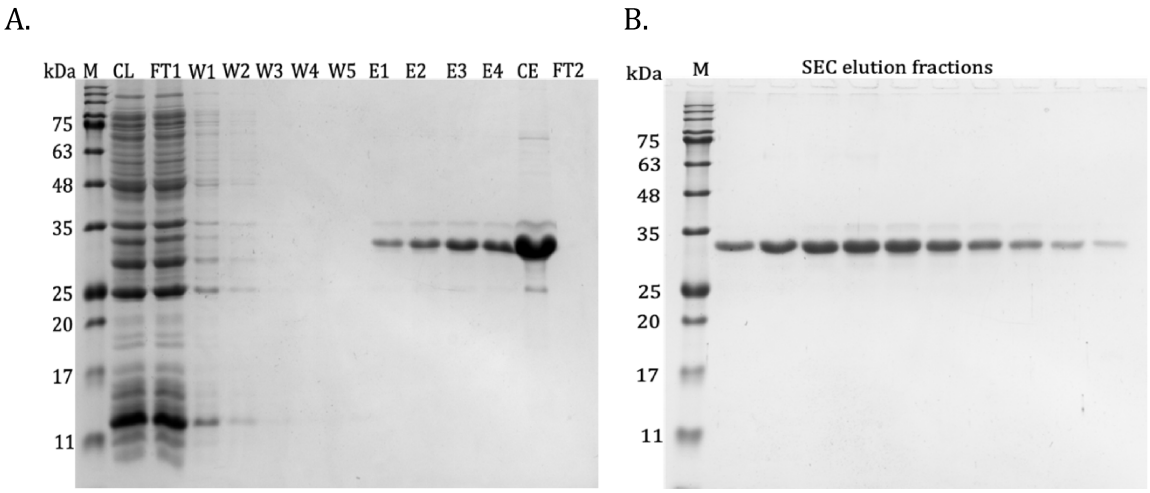


Figure 3.2.3.3.1.1 SDS-PAGE of mutated Par-4 CC EXS purification from E. coli

A. SDS-PAGE analysis of mutated Par-4 CC EXS purification (blue silver stain), M- pre-stained marker; CL- cell lysate; FT1 flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator. B SDS-PAGE analysis of mutated Par-4 CC EXS different SEC elution fractions.

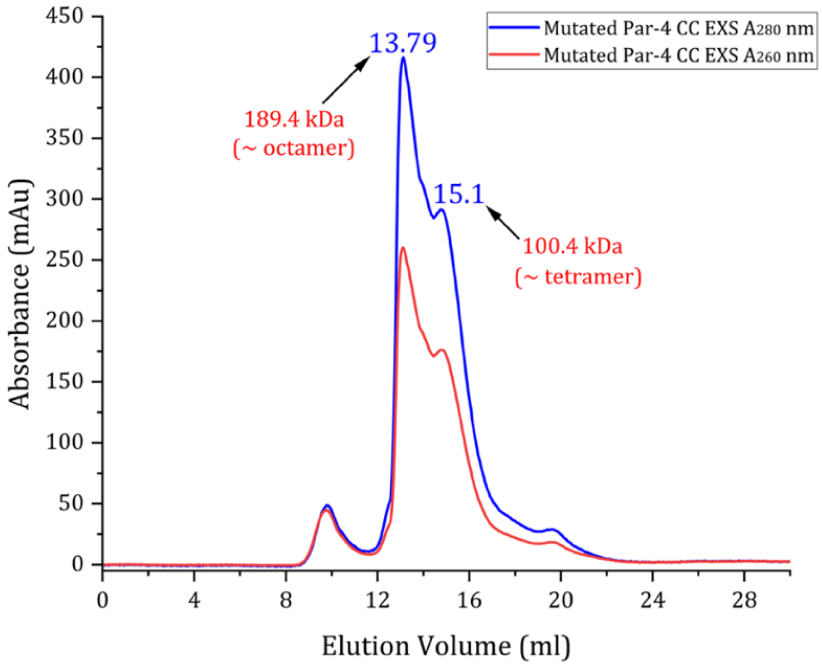


Figure 3.2.3.3.1.2 Superdex 200 increase SEC profile of mutated Par-4 CC EXS from E. coli under acidic condition

When different concentrations of sodium chloride (NaCl) (50 mM, 100 mM, 200 mM, 300 mM, and 400 mM) were included in the SEC buffer under acidic conditions, a minor shift to the right in the SEC peak was observed (Figure 3.2.3.3.1.3). However, another small peak appeared at the void volume of the SEC even with the lowest salt

concentration (50 mM). This peak increased as the salt concentration increased, indicating the formation of higher oligomerization states (aggregation). The SEC profile in the presence of 400 mM salt concentration was not included in the figure, as the protein aggregated and blocked the column.

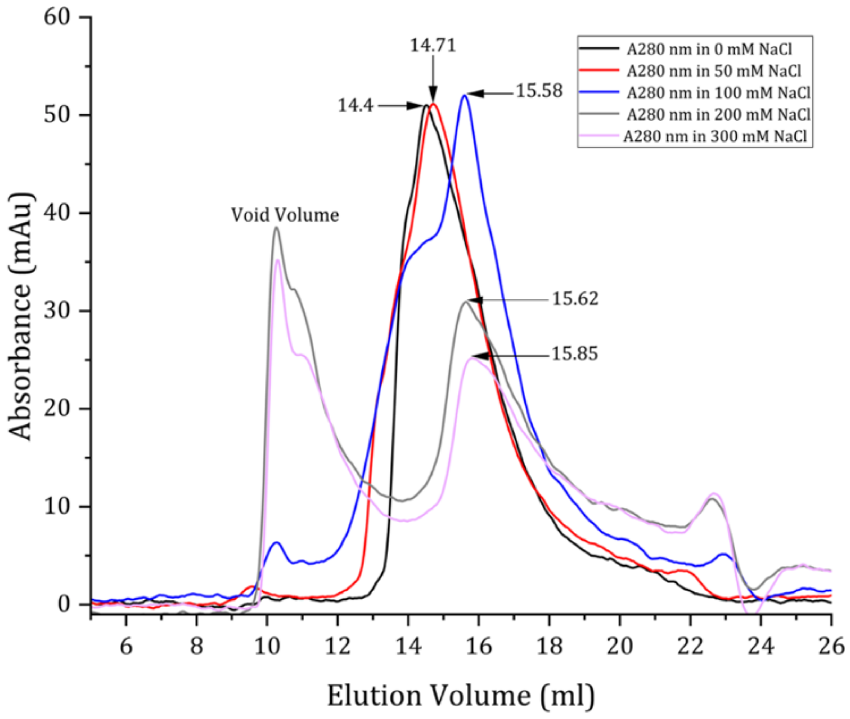


Figure 3.2.3.3.1.3 Superdex 200 increase SEC profile of mutated Par-4 CC EXS in the presence of different NaCl concentrations

3.2.3.3.2 Purification of mutated Par-4 CC EXS expressed in Hi5 cells

Figure 3.2.3.3.2.1 A shows the SDS-PAGE of the purification profile obtained from affinity chromatography under acidic conditions for the mutated protein expressed in Hi5 cells. Figure 3.2.3.3.2.1 B shows the SDS-PAGE of the SEC elution fractions

obtained from gel filtration under acidic conditions for the mutated protein expressed in Hi5 cells. The protein obtained from Hi5 cells exhibits the same mobility pattern in SDS-PAGE (higher molecular weight) and the same elution volume in size exclusion chromatography as the protein purified from E. coli (Figure 3.2.3.3.2.2). Approximately 100-150 µg of mutated Par-4 CC EXS could be purified from a 1 g Hi5 cell pellet obtained from 30 ml of culture. The purified protein was used for biophysical characterization.

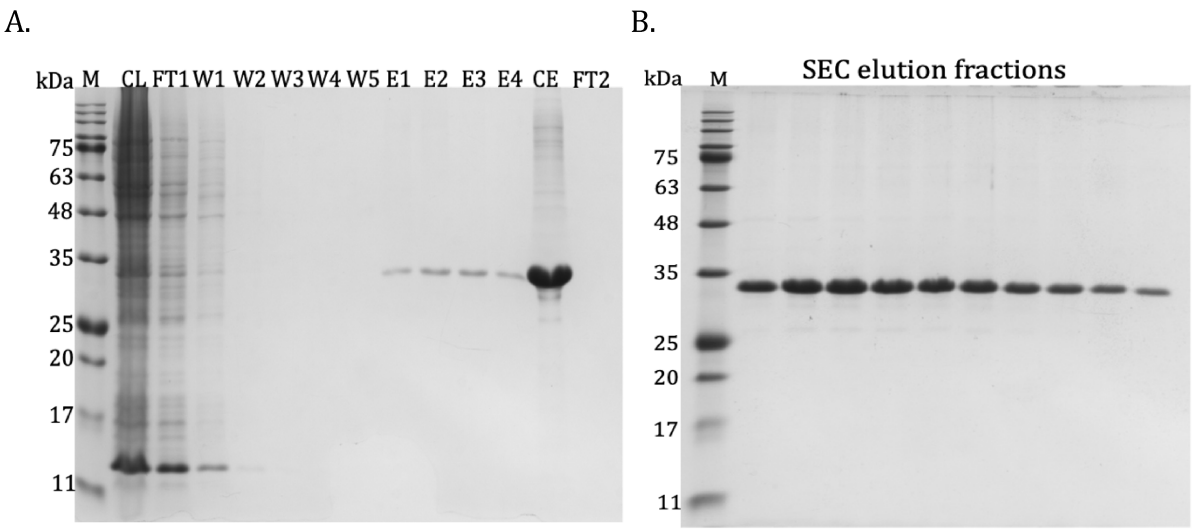


Figure 3.2.3.3.2.1 SDS-PAGE of mutated Par-4 CC EXS purification from Hi5 cells

A. SDS-PAGE analysis of mutated Par-4 CC EXS purification (blue silver stain), M- pre-stained marker; CL- cell lysate; FT1 flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator. B SDS-PAGE analysis of mutated Par-4 CC EXS different SEC elution fractions.

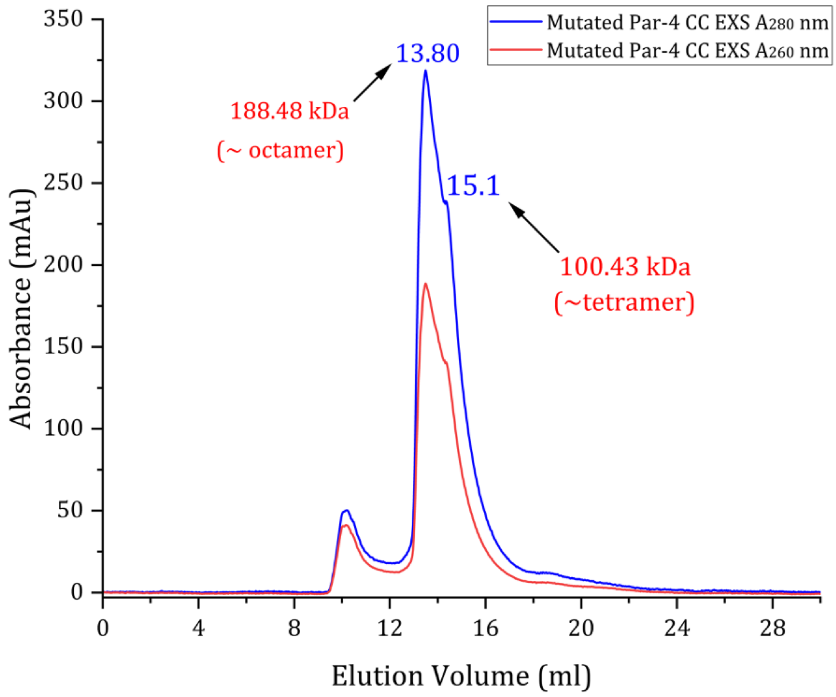


Figure 3.2.3.3.2.2 Superdex 200 increase SEC profile of mutated Par-4 CC EXS from Hi5 cells under acidic condition

According to the results mentioned above, no differences were observed in protein behavior, oligomerization, and homogeneity between the insect cell and E. coli expression systems.

3.2.4 Identity and integrity of the proteins

The purified proteins were subjected to different techniques in order to verify its identify and integrity including western blot, mass spectrometry and LC/MS mass.

3.2.4.1 Western Blot results for mutated Par-4 CC EXS

The anti-strep tag antibody was used in western blot detection since the proteins have a strep-tag. Figures 3.2.4.1.1 A and B and 3.2.4.1.2 A and B show the western blot detection of the purification profile obtained from affinity chromatography under acidic conditions and the SEC elution fractions for the mutated protein expressed in E. coli and Hi5 cells, respectively. The purified proteins are western-positive, indicating that the N-terminal end remains intact post-purification.

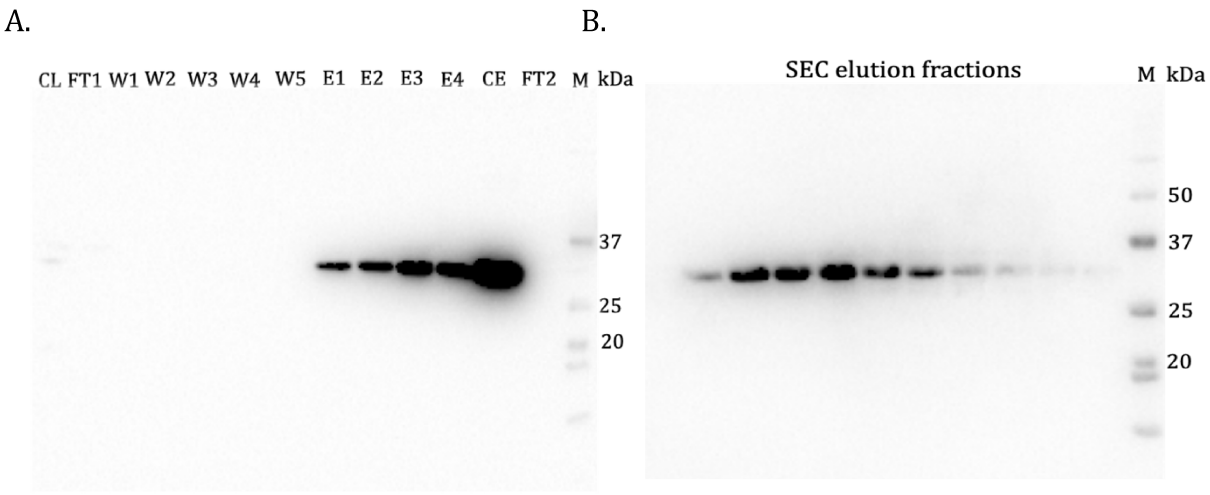


Figure 3.2.4.1.1 Anti-strep western blot of mutated Par-4 CC EXS purification from E. coli

A. Anti-strep western blot analysis of mutated Par-4 CC EXS purification, M- Precision Plus Protein™ WesternC™ Blotting Standards; CL- cell lysate; FT1 flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator. B Western blot analysis of mutated Par-4 CC EXS different SEC elution fractions.

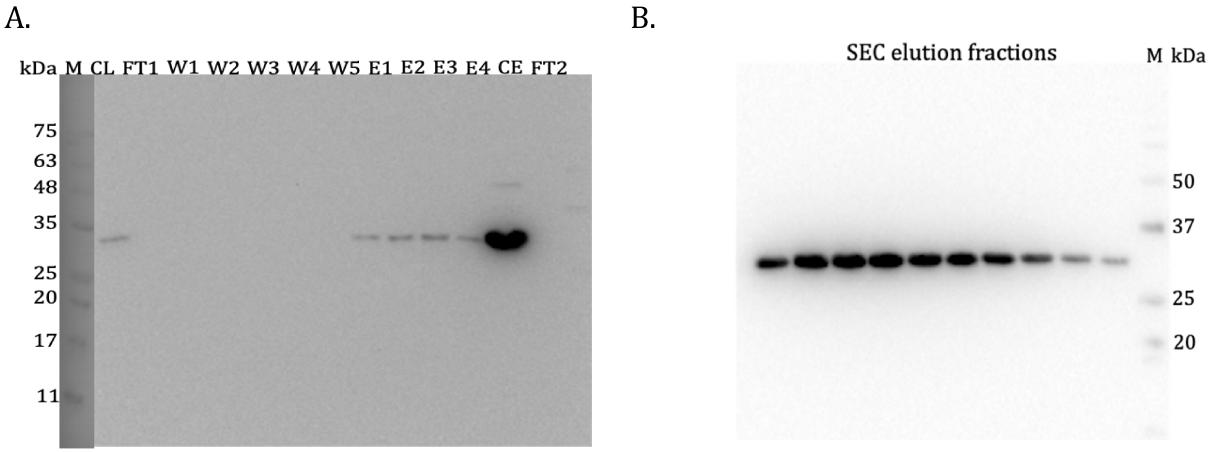


Figure 3.2.4.1.2 Anti-strep western blot of mutated Par-4 CC EXS purification from Hi5 cells

A. Anti-strep western blot analysis of mutated Par-4 CC EXS purification, M- M- pre-stained marker; CL- cell lysate; FT1 flow through; W1, W2, W3, W4, W5- different washing steps; E1, E2, E3, E4- different affinity elution fractions; CE concentrated affinity elution fractions; FT2- flow through from the concentrator. B Western blot analysis of mutated Par-4 CC EXS different SEC elution fractions; M - Precision Plus Protein™ WesternC™ Blotting Standards.

3.2.4.2 LC/MS mass results for mutated Par-4 CC EXS

The integrity of the proteins was evaluated using LC/MS mass spectrometry. In-gel trypsin digestion was performed to determine the N-terminal and C-terminal ends of the proteins. Trypsin cleaves at specific peptide bonds on the C-terminal side of lysine and arginine residues. The peptides were then analyzed on an Agilent 6550 iFunnel Q-TOF system using Agilent liquid chromatography/mass spectrometry (LC/MS) databases and libraries. The detection of multiple peptides with identical N- or C-termini indicates the respective terminus of a protein. The amino acid sequence coverage for the mutated Par-4 CC EXS from E. coli and Hi5 expression systems was 93.3% and 98%, respectively (Figure 3.2.4.2.1 A and B; Table 3.2.4.2.1).

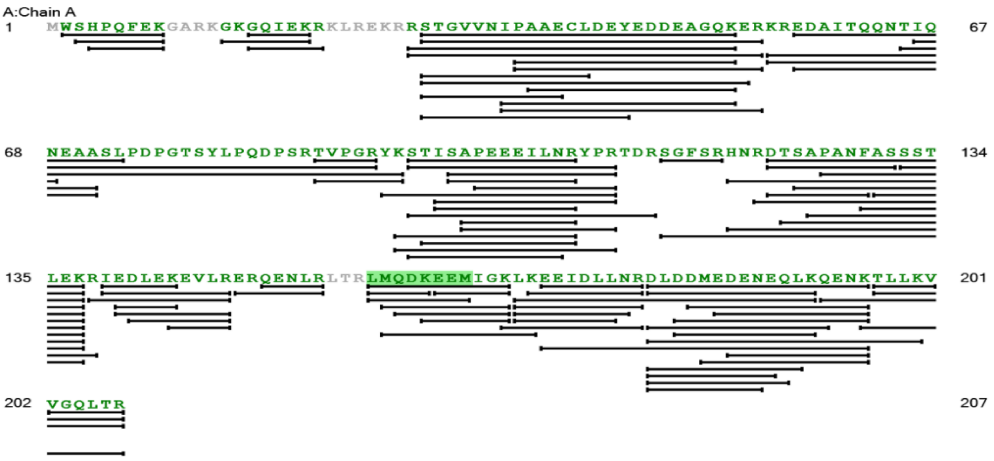
The N-terminal methionine residue in the sequence of the protein expressed in E. coli was found to be cleaved, whereas it was intact in the protein sequence expressed in Hi5 cells. This cleavage of methionine was also obsereved for LC/MS results of CC domain, which could be due to same reason i.e., methionine aminopeptidase, as both proteins; Par-4 CC EXS and CC domain that were expressed in E. coli have Met-Ala at the their N-terminals whereas the sequence expressed in Hi5 contains Met-Tyr at the same positions, which could explain the methionine cleavage.

Overall, the LC/MS results indicate that the purified proteins are full-length and uncut during the purification process. Furthermore, post-translational modification analysis showed that all the proteins did not have any modifications.

Table 3.2.4.2.1 N terminal to C terminal sequence coverage from LC-mass/mass

Protein	Expression system	Sequence coverage
Mutated Par-4 CC EXS	E. coli	93.3%
Mutated Par-4 CC EXS	Hi5 insect cells	98%

A.



B.

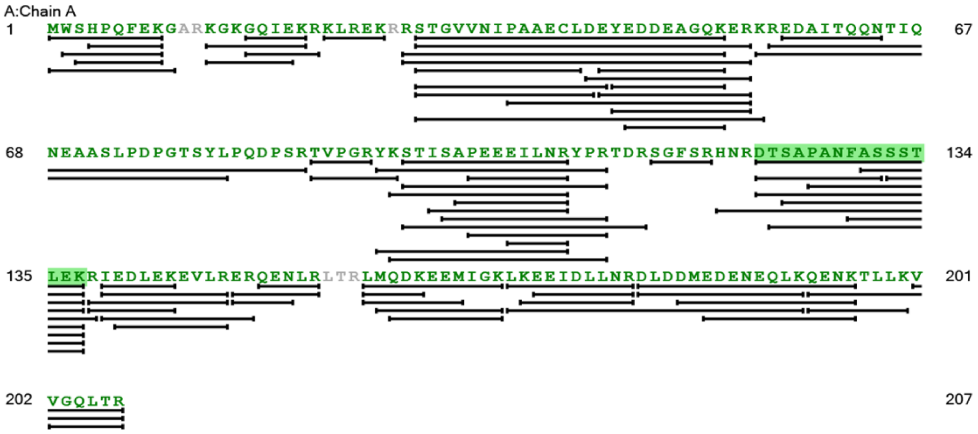


Figure 3.2.4.2.1: Peptides identification of mutated Par-4 CC EXS by LC-MS/MS analysis

LC-MS/MS gave N terminal end to C terminal end sequence coverage. A LC-MS/MS result CC domain and mutants. B. LC-MS/MS result for mutated Par-4 CC EXS from E. coli. C. LC-MS/MS result for mutated Par-4 CC EXS from Hi5 cell.

3.2.4.3 Mass spectrometry results for mutated Par-4 CC EXS

Mass spectrometry was used to determine the molecular weights of the proteins more accurately. A protein standard with molecular weights ranging from 10 to 66 kDa was used to calibrate the mass spectrometer to perform a reliable analysis by MALDI-TOF for molecular weights within that range. The MS result for the mutated Par-4 CC (Figure 3.2.4.3.1) shows three peaks. The strongest peak appears at 23.966 kDa/e, which matches the M^{1+} species of the mutated Par-4 CC EXS. The M^{2+} species peak

appears at 11.982 kDa/e, and a third weak peak at 47.927 kDa/e matches the single charge ion of the dimer ($2M^{1+}$ species).

The MS result for the protein did not match exactly with the theoretical molecular weight of the mutated CC EXS. This difference could be due to the cleavage of the N-terminal methionine, as previously observed in the LC/MS result, since the difference is equal to the molecular weight of methionine (131 Da)

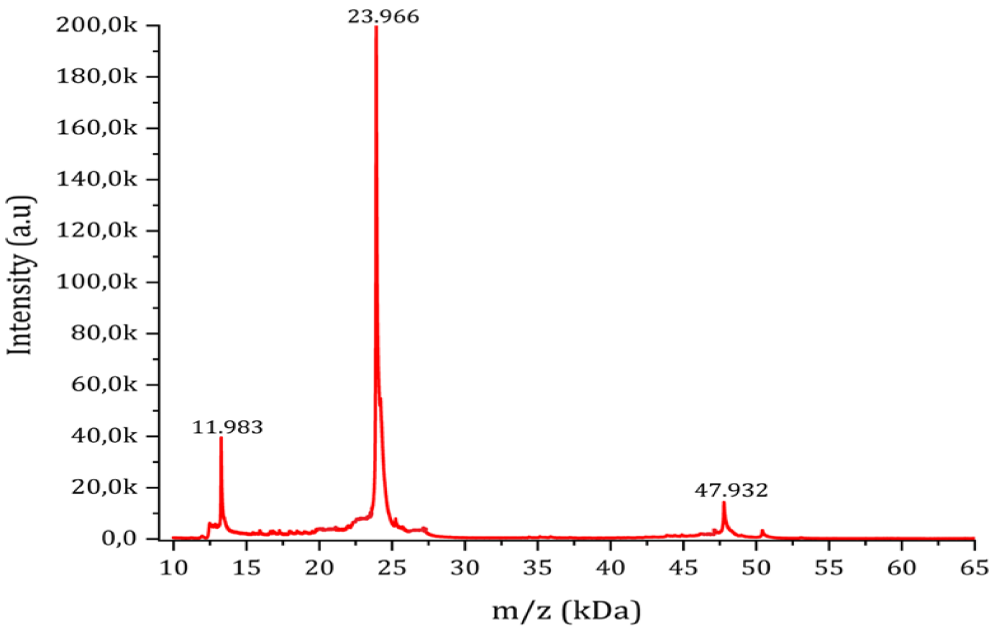


Figure 3.2.4.3.1 MALDI-TOF mass spectrum of mutated Par-4 CC EXS

3.2.4.4 De-glycosylation assay results for mutated Par-4 CC EXS

The purified mutated Par-4 CC EXS protein from Hi5 cells was treated with PNGase and Endo H enzymes to investigate the presence of N-glycosylation. Figures 3.2.4.4.1 A and B and 3.2.4.4.2 A and B show the blue silver stain SDS-PAGE and anti-strep tag western blot results for the PNGase and Endo H reactions, respectively. Since the protein was in SEC buffer at pH 3.5, each reaction was performed under both acidic and neutral conditions to ensure maximum enzyme activity. The treated samples with both enzymes did not show any difference in the band patterns on SDS-PAGE and the anti-strep western blot, implying that the sites are not accessible for glycosylation.

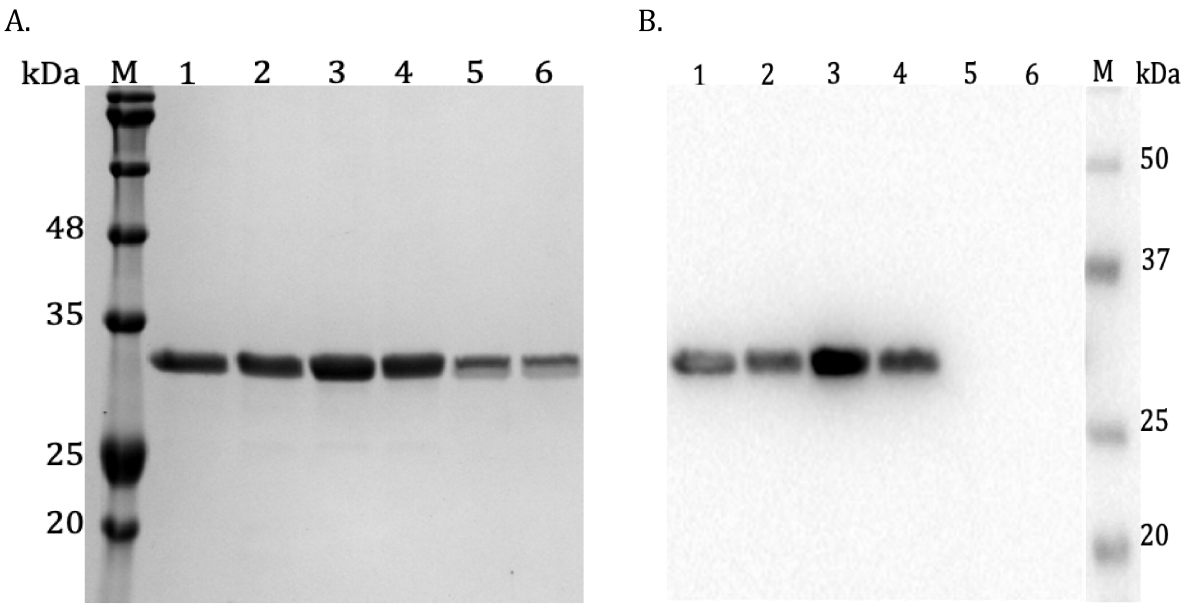


Figure 3.2.4.4.1 SDS-PAGE and anti-strep western blot analysis of PNGase reactions for mutated Par-4 CC EXS from Hi5 cells.

A. SDS-PAGE, 1. Protein and PNGase under acidic pH, 2. Protein alone under acidic pH, 3. Protein and PNGase under neutral pH, 4. Protein alone under neutral pH, 5. PNGase alone under acidic pH, 6. PNGase alone under neutral pH. B. Anti-strep western blot, 1. Protein and PNGase under acidic pH, 2. Protein alone under acidic pH, 3. Protein and PNGase under neutral pH, 4. Protein alone under neutral pH, 5. PNGase alone under acidic pH, 6. PNGase alone under neutral pH.

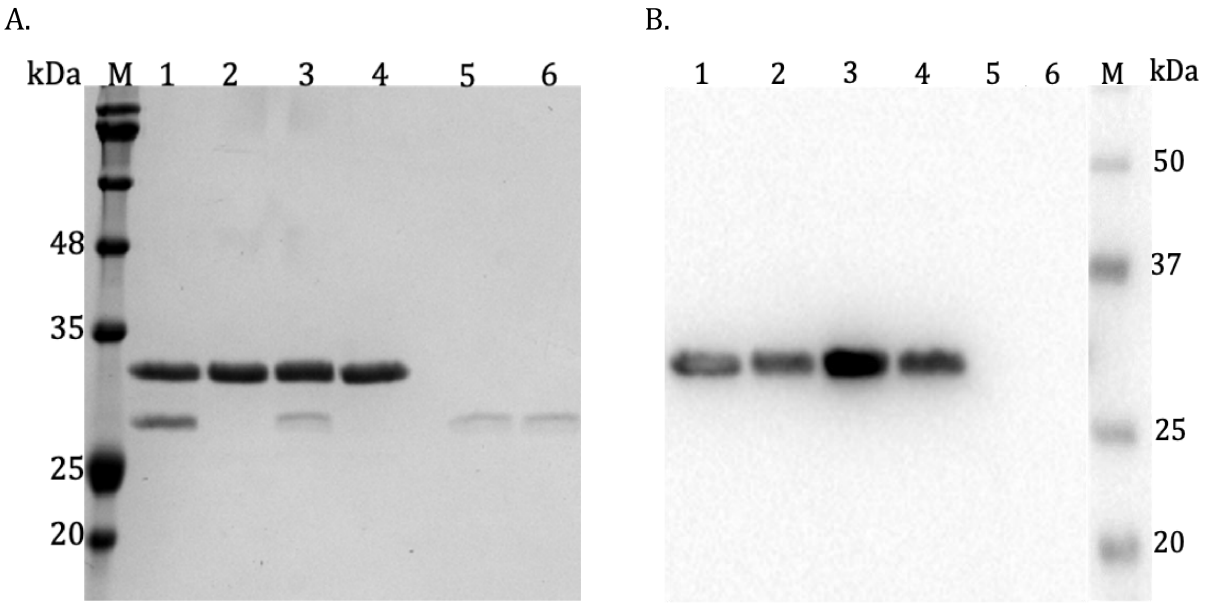


Figure 3.2.4.4.2: SDS-PAGE and anti-strep western blot analysis of Endo H reactions for mutated Par-4 CC EXS from Hi5 cells.

A. SDS-PAGE, 1. Protein and Endo H under acidic pH, 2. Protein alone under acidic pH, 3. Protein and Endo H under neutral pH, 4. Protein alone under neutral pH, 5. Endo H alone under acidic pH, 6. Endo H alone under neutral pH. B. Anti-strep western blot, 1. Protein and PNGase under acidic pH, 2. Protein alone under acidic pH, 3. Protein and PNGase under neutral pH, 4. Protein alone under neutral pH, 5. PNGase alone under acidic pH, 6. PNGase alone under neutral pH.

Since the evaluation of proteins' purity, integrity, and quantity have all been done, the subsequent step was to study the biophysical characterization of the protein.

3.2.5 Biophysical Characterization

Biophysical characterization techniques play a fundamental role in protein studies due to their ability to provide detailed insights into the secondary and tertiary structure, folding, stability, dynamic properties, as well as aggregation and higher oligomerization of the protein.
To assess these properties, the purified proteins from E. coli and Hi5 were subjected to various biophysical techniques, including Dynamic Light Scattering (DLS), Circular Dichroism (CD), and X-ray crystallography.

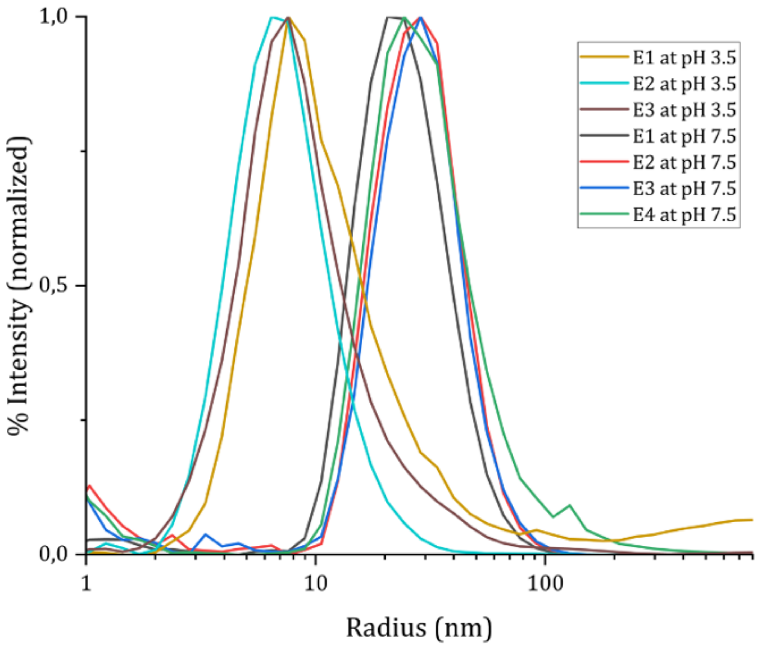
3.2.5.1 Dynamic Light Scattering results for mutated Par-4 CC EXS

Different protein samples from the SEC peaks, including octamer, octamer/tetramer, and tetramer fractions at pH 3.5, as well as various affinity elution fractions at pH 7.5, were subjected to Dynamic Light Scattering (DLS) to assess their purity, homogeneity, and oligomerization behavior. Figure 3.2.5.1.1 A shows the radius size versus the percentage intensity. Figure 3.2.5.1.1 B and Table 3.2.5.1.1 present the radius size, polydispersity percentage (PD %), molecular weight (MW), and intensity for each sample.

The results indicate that samples under acidic conditions have significantly smaller radius sizes and consequently smaller molecular weights compared to those under neutral conditions. This suggests that the protonation of Asp and Glu residues decreases protein aggregation. Furthermore, all samples under both conditions exhibit high polydispersity, indicating the presence of various species in solution for each sample.

Based on the aforementioned results, the neutral condition has a negative effect on protein folding and solubility, leading to protein aggregation, which was not observed under acidic conditions. Therefore, the protein was purified under acidic conditions for setting the crystallization plates.

A.



B.

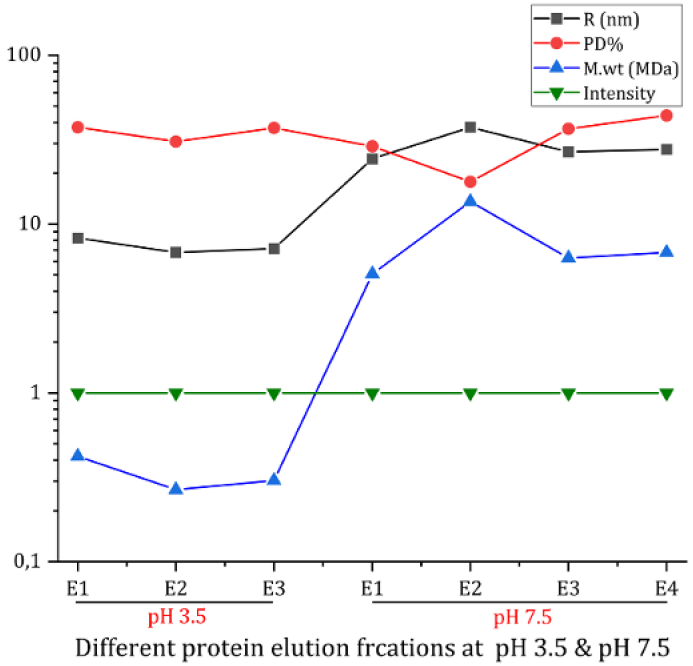


Figure 3.2.5.1.1 Particle size analysis of mutated Par-4 CC EXS under acidic and neutral conditions using DLS

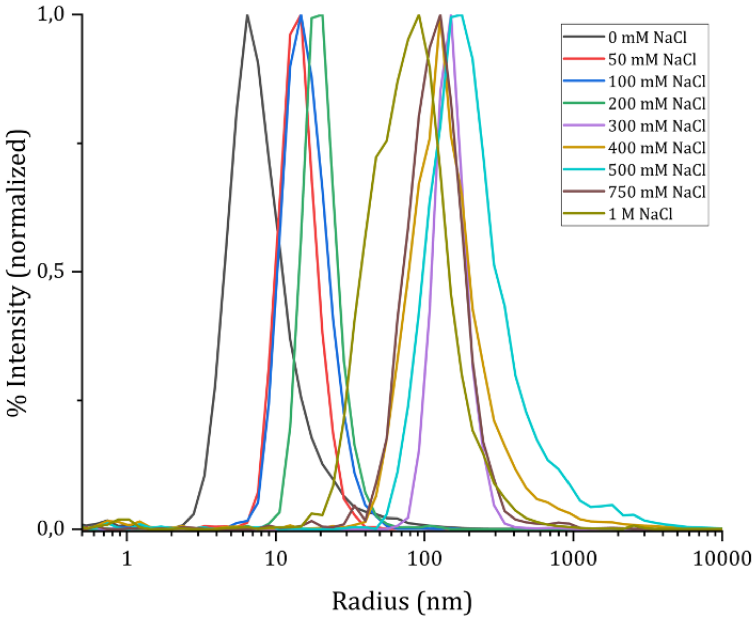
A. Intensity percentage vs radius of different protein samples under acidic and neutral conditions. B. Log 10 of Radius, polydispersity %, molecular weight and intensity of different protein samples under acidic and neutral conditions.

Table 3.2.5.1.1 DLS results for mutated Par-4 CC EXS under neutral and acidic conditions

Protein sample	pH	Radius (nm)	PD %	MW	Intensity
E1 (SEC octamer fraction)	3.5	8.27	37.5	421.9 kDa	1
E2 (SEC octamer/tetramer fraction)	3.5	6.78	30.8	267.3 kDa	1
E3 (SEC tetramer fraction)	3.5	7.16	37.1	303.3 kDa	1
E1	7.5	24.36	28.9	5.060 MDa	1
E2	7.5	37.44	17.8	13.59 MDa	1
E3	7.5	26.77	36.7	6.284 MDa	1
E4	7.5	27.68	44.0	6.786 MDa	1

Additionally, the effect of different sodium chloride (NaCl) salt concentrations on protein behavior under acidic conditions was assessed using DLS. The results revealed a relative increase in the radius size and molecular weight of the protein at concentrations between 50 mM and 200 mM NaCl compared to samples without salt. However, at concentrations between 300 mM and 1 M NaCl, both the radius size and molecular weight significantly increased, indicating a high level of oligomerization (Figures 3.2.5.1.2 A and B; Table 3.2.5.1.2). Furthermore, concentrations between 50 mM and 300 mM NaCl showed high polydispersity, whereas concentrations of 400 mM, 750 mM, and 1 M exhibited low polydispersity. Notably, the results highlighted that 500 mM NaCl had the most pronounced effect on the protein, resulting in the largest radius size, molecular weight, and highest polydispersity.

A.



B.

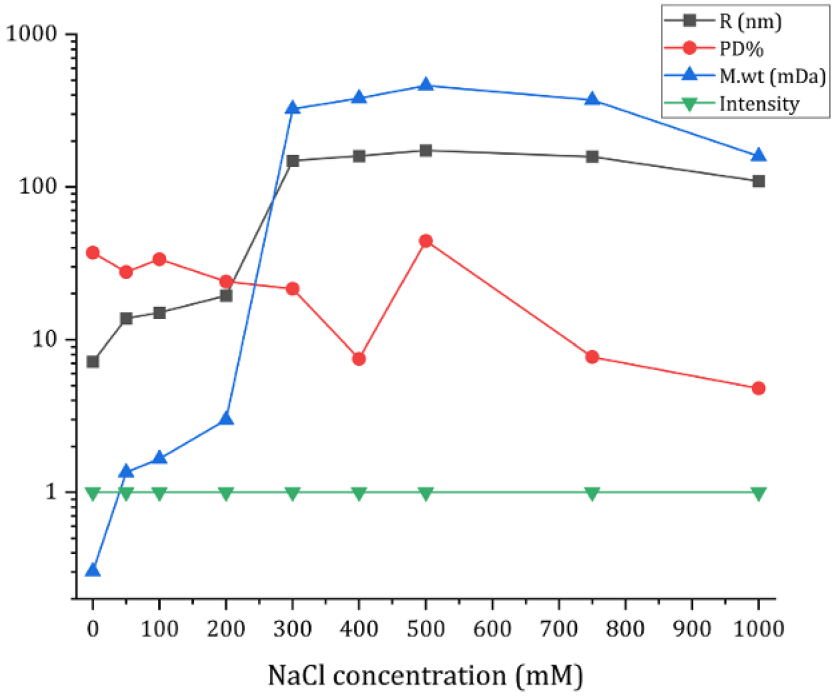


Figure 3.2.5.1.2 Particle size analysis of mutated Par-4 CC EXS as a function of salt concentration under acidic condition using DLS

A. Intensity percentage vs radius of different protein samples under acidic and neutral conditions in NaCl concentrations ranging from 0.0 to 1.0 M. B. Log 10 of Radius, polydispersity %, molecular weight and intensity of different protein samples under acidic and neutral conditions in NaCl concentrations ranging from 0.0 to 1.0 M.

Table 3.2.5.1.2 DLS results for mutated Par-4 CC EXS under acidic conditions as a function of NaCl salt concentration

NaCl salt concentration	Radius (nm)	PD %	MW (MDa)	Intensity
0 mM	7.16	37.1	0.3033	1
50 mM	13.73	27.7	1.352	1
100 mM	15.01	33.6	1.661	1
200 mM	19.37	24	2.986	1
300 mM	148.66	21.5	324.1	1
400 mM	159.59	7.5	381.5	1
500 mM	173.28	44.3	461.1	1
750 mM	157.75	7.7	371.5	1
1000 mM	109.05	4.8	158.9	1

3.2.5.2 Assessment of the secondary structure of Par-4 CC EXS by CD

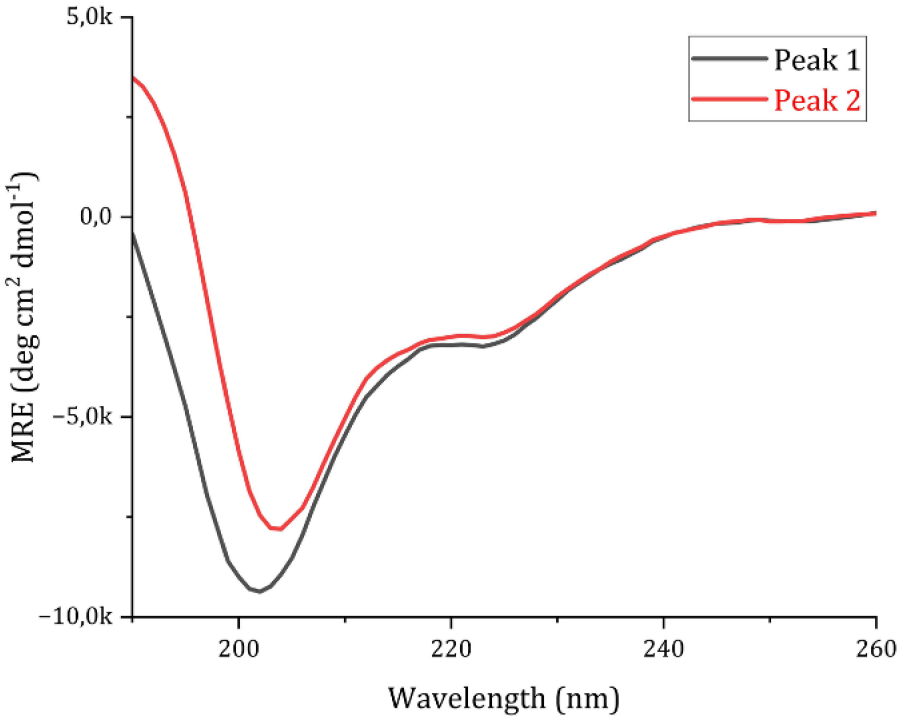
The circular dichroism (CD) measurements for mutated Par-4 CC EXS were conducted at both neutral and acidic pH, as well as in the absence and presence of various salt concentrations. At acidic pH, CD measurements for both proteins (from E. coli and Hi5) were performed in 50 mM potassium phosphate buffer (SEC buffer 1). Additionally, for the protein obtained from E. coli, CD measurements were taken in the absence and presence of different concentrations of potassium fluoride and sodium fluoride salts (ranging from 50 to 500 mM), as well as 50 mM magnesium sulfate salt, all in the same buffer. At neutral pH, CD measurements were conducted in 50 mM potassium phosphate buffer at pH 7.5, supplemented with 100 mM potassium fluoride

Under neutral pH, samples from the SEC peak 1 and peak 2 were used in the CD measurement. The CD spectra of both samples showed a deep negative peak near 200 nm indicating its predominant unfolded monomeric state with a small shoulder near 222 nm indicating the presence of some residual structure (figure 3.2.5.2.1 A).

At acidic pH, samples from the SEC octamer peak, tetramer peak and between the two peaks were used in the CD measurement. The Far UV- spectra of the three different SEC fractions showed an α-helical profile with double minima at 208 nm and 222 nm (Figure 3.2.5.2.1 B). In addition, a maximum near 193 nm was only observed in the octamer SEC fraction which indicates a more α-helix secondary structure exists in the octamer fraction compared to the other two fractions. The CD spectra deconvolution for the SEC fractions at both conditions are tabulated in table 3.2.5.2.1.

The CD results showed that the acidic condition is more favorable for inducing protein folding and stability than the neutral pH. In addition, no difference in the secondary structure content between the protein expressed in E. coli and the one expressed in Hi5 cells was observed. This indicates that utilizing the Hi5 insect cell system for protein expression did not add any benefits in inducing extra structure and folding in the protein, which was the main reason for using this expression system. From now on, the protein will be called mutated Par-4 CC EXS regardless to the expression system.

A.



B.

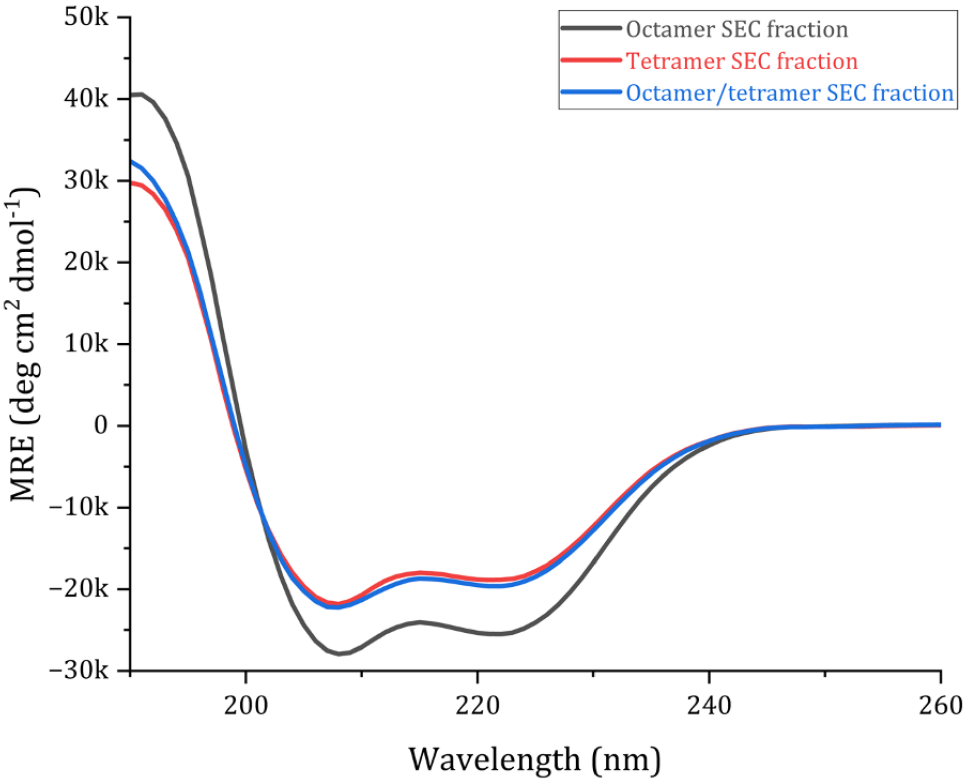


Figure 3.2.5.2.1 Far-UV CD spectra of mutated Par-4 CC EXS from E. coli and Hi5 cells under neutral and acidic condition

A. CD spectra for two different SEC peaks at neutral pH. B. CD spectra for different SEC elution fractions, Black CD spectrum corresponds to the octamer peak in the SEC. Red CD spectrum corresponds to the tetramer peak in the SEC. Blue CD spectrum corresponds to the overlapped fractions between octamer and tetramer SEC peaks.

Table 3.2.5.2.1 Deconvolution of the CD spectra for different SEC fractions of mutated Par-4 CC EXS under neutral and acidic conditions expressed in E. coli and Hi5 insect cells

Source	Buffer pH	SEC Fraction	Helix %	Antiparallel %	Parallel %	Turn %	Other %
E. coli	7.5	Peak 1	4.3	27.3	0.9	17.4	50.1
E. coli	7.5	Peak 2	5.0	29.2	0.0	16.6	49.2
E. coli	3.5	Octamer	63.5	8.6	0.0	1.7	26.2
E. coli	3.5	Octamer/ Tetramer	54.0	8.6	0.0	5.1	32.3
E. coli	3.5	Tetramer	49.4	8.6	0.0	9.7	32.3
Hi5	7.5	Peak 1	4.3	27.5	1.0	17.3	49.9
Hi5	7.5	Peak 2	4.9	29.2	0.1	16.7	49.1
Hi5	3.5	Octamer	63.6	8.6	0.0	1.6	26.2
Hi5	3.5	Octamer/ Tetramer	54.0	8.6	0.0	5.1	32.2
Hi5	3.5	Tetramer	49.4	8.6	0.0	9.7	32.3

The structure of the CC domain (residues 254–332) which is presented in 39.8% of the Par-4 CC EXS (residues 136–332) was shown to form a coiled coil structure at the same condition, i.e., acidic environment (Tiruttani Subhramanyam et al., 2017).

The table 3.2.5.2.2 shows the secondary structures exist out of the CC domain in each SEC fraction under the acidic condition.

Table 3.2.5.2.2 Secondary structures percentage presented in the region beyond the CC domain

Source	Buffer pH	SEC Fraction	Helix %	Antiparallel %	Parallel %	Turn %
E. coli	3.5	Octamer	23.7	8.6	0.0	1.7
E. coli	3.5	Octamer/ Tetramer	14.1	8.7	0.0	5.1
E. coli	3.5	Tetramer	9.6	8.6	0.0	9.7
Hi5	3.5	Octamer	23.8	8.6	0.0	1.6
Hi5	3.5	Octamer/ Tetramer	14.2	8.6	0.0	5.1
Hi5	3.5	Tetramer	9.6	8.6	0.0	5.7

Upon addition of varying concentrations of potassium fluoride (KF), sodium fluoride (NaF), and 50 mM magnesium sulfate (MgSO4) salts to the CD buffer at pH 3.5, changes in the CD spectra deconvolution were observed for the mutated Par-4 CC EXS. Generally, the deconvolution of spectra indicated a decrease in α-helix content and an increase in β-sheet and turn contents, which is found to be mostly salt concentration dependent tables 3.2.5.2.3, 3.2.5.2.4, and 3.2.5.2.5. The figures 3.2.5.2.2 A, B, and C depict the CD spectra for mutated Par-4 CC EXS in the presence of different concentrations of KF, NaF, and MgSO4, respectively. Remarkably, none of the tested salts, across their varied concentrations, induced additional structure, folding, or stability in the protein compared to the one without salts.

Table 3.2.5.2.3 Deconvolution of the CD spectra for mutated Par-4 CC EXS in the presence of varying concentrations of KF salt

Salt type	Salt concentration	Helix %	Antiparallel %	Parallel %	Turn %	Other %
KF	0	49.4	8.6	0.0	9.7	32.3
KF	50 mM	39.0	14.4	0.0	12.9	33.7
KF	100 mM	38.7	14.7	0.0	12.7	33.9
KF	150 mM	36.1	16.1	0.0	13.1	34.7
KF	200 mM	35.7	16.0	0.0	12.7	35.6
KF	300 mM	31.1	17.5	0.0	14.5	36.9
KF	400 mM	30.8	20.6	0.1	13.5	35.0
KF	500 mM	28.0	18.8	0.5	11.7	41.0

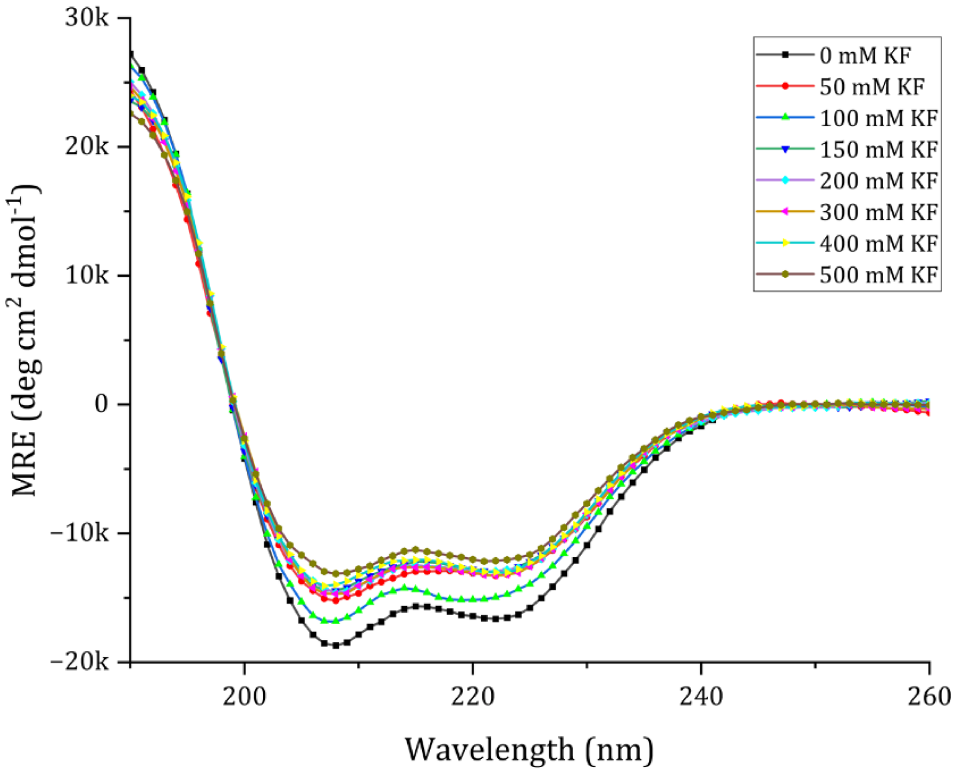
Table 3.2.5.2.4 Deconvolution of the CD spectra for mutated Par-4 CC EXS in the presence of varying concentrations of NaF salt

Salt type	Salt concentration	Helix %	Antiparallel %	Parallel %	Turn %	Other %
NaF	0	49.4	8.6	0.0	9.7	32.3
NaF	50 mM	38.6	14.1	0.0	13.0	34.3
NaF	100 mM	39.3	13.8	0.1	12.4	34.4
NaF	150 mM	36.0	16.4	0.0	13.0	34.6
NaF	200 mM	35.7	15.7	0.0	12.4	36.2
NaF	300 mM	31.5	17.7	0.0	15.2	35.6
NaF	400 mM	31.3	18.9	0.8	13.3	35.7
NaF	500 mM	28.7	19.2	0.1	12.8	39.2

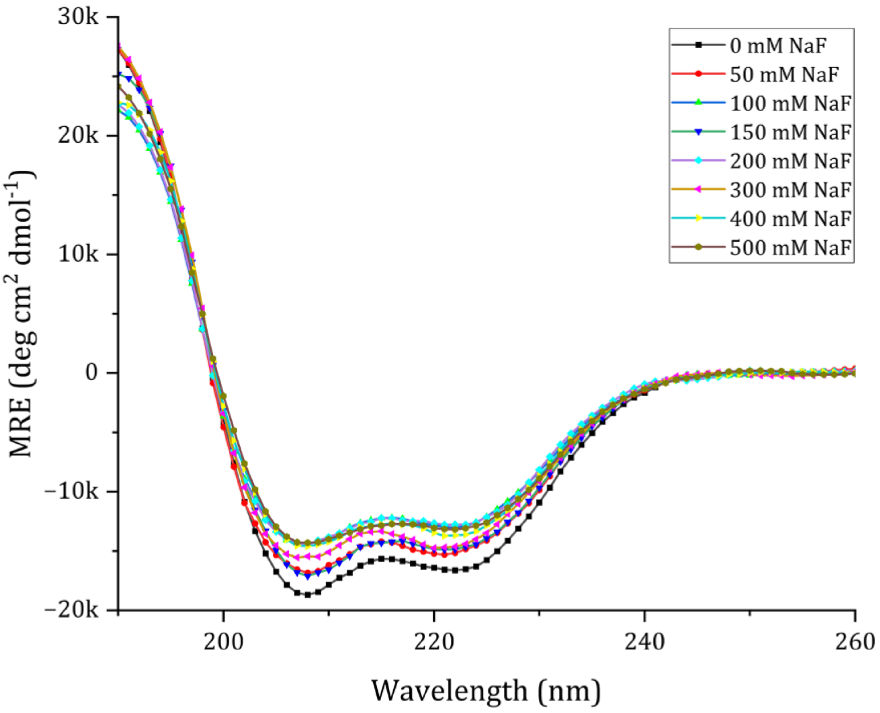
Table 3.2.5.2.5 Deconvolution of the CD spectra for mutated Par-4 CC EXS in the presence of 50 mM MgSO4 salt

Salt type	Salt concentration	Helix %	Antiparallel %	Parallel %	Turn %	Other %
$MgSO_4$	0	49.4	8.6	0.0	9.7	32.3
$MgSO_4$	50 mM	27.9	17.5	2.3	13.0	39.3

A.



B.



C.

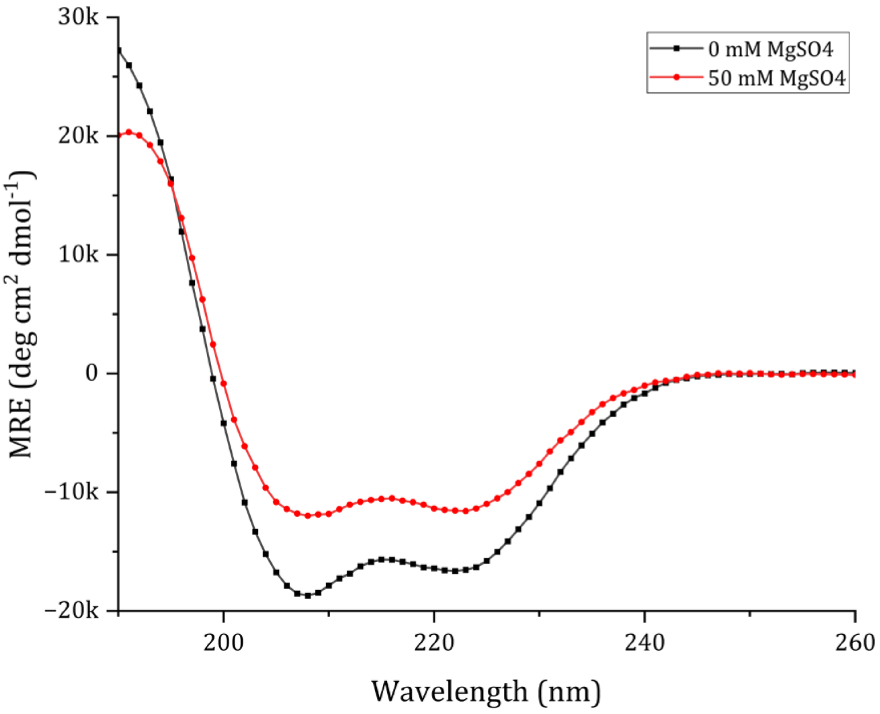
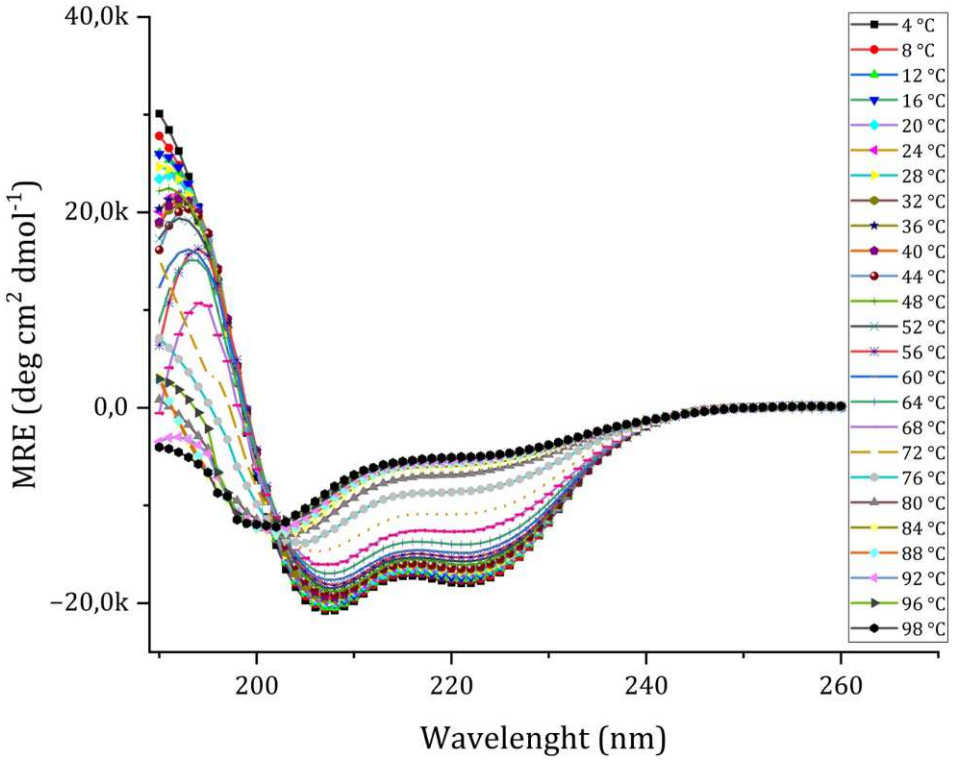


Figure 3.2.5.2.2 Far-UV CD spectra of mutated Par-4 CC EXS in different salts concentrations

A. Black CD spectrum corresponds to the octamer SEC fraction without added salt. The other colored CD spectra correspond to different KF salt concentrations starting from 50mM to 500 mM. B. Black CD spectrum corresponds to the octamer SEC fraction without added salt. The other colored CD spectra correspond to different NaF salt concentrations starting from 50mM to 500 mM. C. Black CD spectrum corresponds to the octamer SEC fraction without added salt. Red CD spectrum corresponds to the octamer SEC fraction with 50 mM MgSO4.

The Far-UV CD spectra reveal a gradual decrease in negative minima as temperature increases. The accompanying figure 3.2.5.2.3 A illustrates the CD temperature gradient scan for mutated Par-4 CC EXS from E. coli, depicting the evolving relationship between its folded and unfolded states. The denatured protein exhibits flattened local minima at 222 nm and 208 nm, whereas the intact protein displays two distinct minima at these wavelengths, alongside a maximum at 200-190 nm. To determine the protein's melting temperature (Tm) from the CD spectra, the emission intensity at 222 nm was compared across three different samples: octamer, octamer/tetramer, and tetramer SEC fractions, against the temperature gradient depicted in the 3.2.5.2.3 B. The results indicated varying Tm values for the samples, measured at 75.92 °C, 75.1 °C, and 73.95 °C for the octamer, octamer/tetramer, and tetramer SEC fractions, respectively. However, the thermal stability curves at 222 nm for the three fractions appear almost identical, suggesting that the oligomerization state of the protein is not influenced by changes in secondary structure.

A.



B

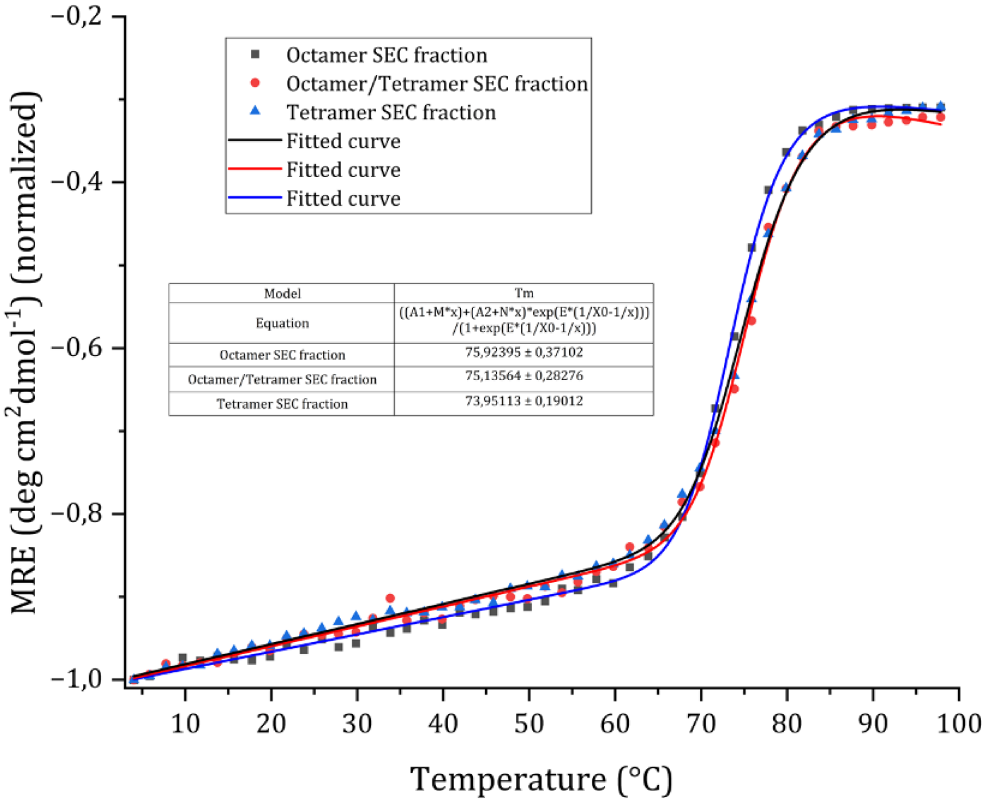


Figure 3.2.5.2.3 Thermal stability of mutated Par-4 CC EXS

A. Temperature scanning of mutated Par-4 CC EXS in 4 °C increments from 4 to 98 °C from by Circular Dichroism Spectroscopy. B. Thermal stability (Tm) of different SEC peaks of mutated Par-4 CC EXS at 222 nm.

The thermodynamic parameters for the three samples were calculated from the thermal denaturaing curves obtained from the CD data to assess their stability. Figure 3.2.5.2.4 shows the Van 't Hoff plot for the three SEC fractions and the obtained slopes and intercepts for each fraction are shown in table 3.2.5.2.6. The thermodynamic parameters for each fraction are tabulated in table 3.2.5.2.7.

Changes in the enthaply of the unfolded fractions of the samples at their melting temperatures are depicted in the figure 3.2.5.2.5. The results showed a relative difference in ΔGUnfolding at the melting temperatures among the samples, wherein the octamer sample had the highest value, and the lowest value was observed in the tetramer fraction (Figure 3.2.5.2.6). A positive increase in ΔGUnfolding indicates that more energy is required for the reaction to take place, rendering it non-spontaneous. Consequently, this suggests that the protein is more stable. Based on the results, the octamer fraction appears to be the most stable compared to the other two fractions.

Table 3.2.5.2.6 Slopes and intercepts for different SEC fractions of mutated Par-4 CC EXS obtained from Van 't Hoff plot

Protein	Slope	Slope error	Intercept	Intercept error	R-square
Octamer SEC fraction	-37147.6	1965.1	107.0	5.6	0.986
Octamer/Tetramer SEC fraction	-36380.1	1598.1	105.0	4.5	0.99
Tetramer SEC fraction	-33944.1	908.1	98.0	2.6	0.995

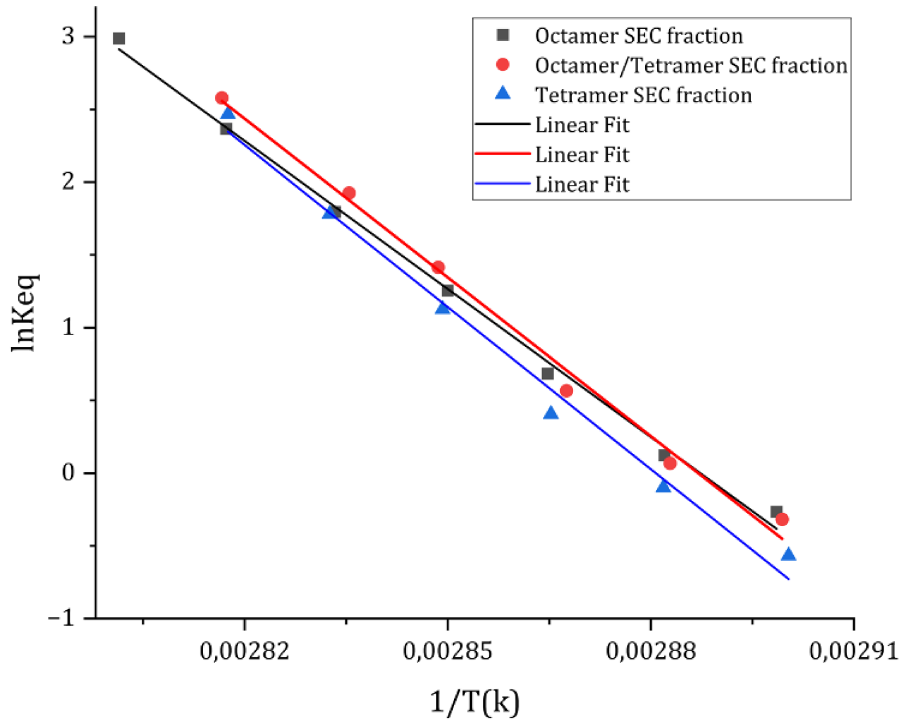


Figure 3.2.5.2.4 Van 't Hoff plots for different SEC fractions of mutated Par-4 CC EXS

Black: Linear fit of octamer SEC fraction. Red: Linear fit of octamer/tetramer SEC fraction. Blue: Linear fit of tetramer SEC fraction.

Table 3.2.5.2.7 Thermodynamic parameters for different SEC fractions of mutated Par-4 CC EXS

Protein	Tm (°C)	Tm (k)	ΔH_{Tm} (kJ/mol)	ΔS_{Tm} (kJ/k.mol)	ΔCp (kJ/k.mol)
Octamer SEC fraction	75.92	349.1	308.8	0.89	3.95
Octamer/Tetramer SEC fraction	75.13	348.3	302.5	0.873	3.78
Tetramer SEC fraction	73.95	347.1	282.2	0.815	3.54

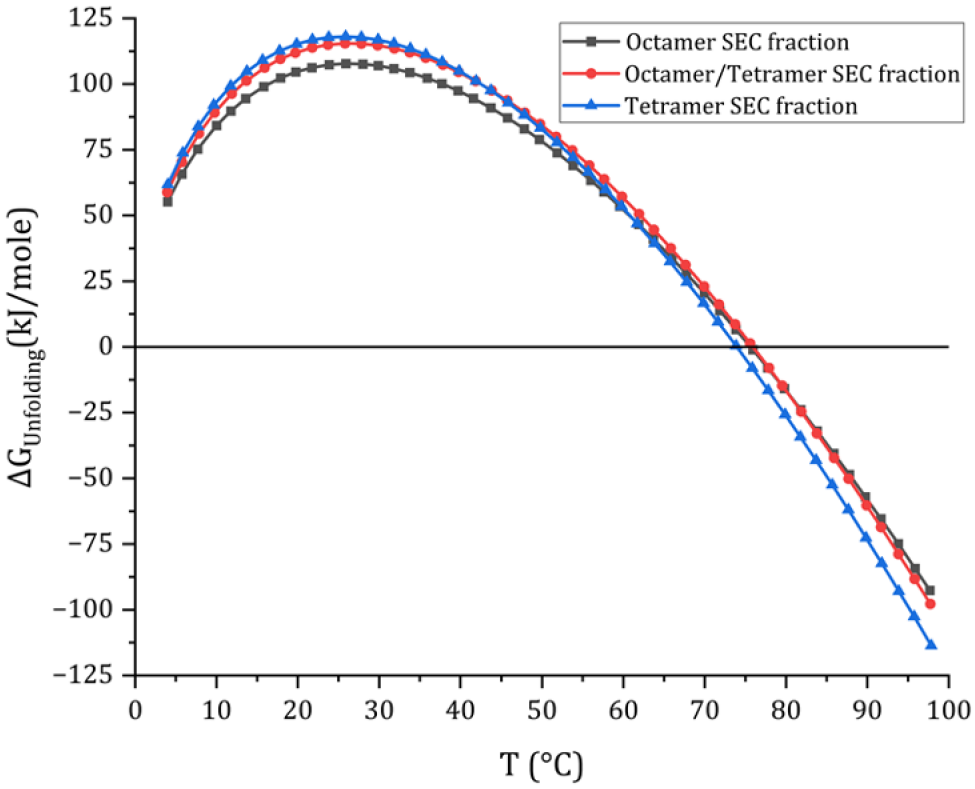


Figure 3.2.5.2.5 Stability curves for different SEC fractions of mutated Par-4 CC EXS

Black: Thermal stability curve of octamer SEC fraction. Red: Thermal stability curve of octamer/tetramer SEC fraction. Blue: Thermal stability curve of tetramer SEC fraction.

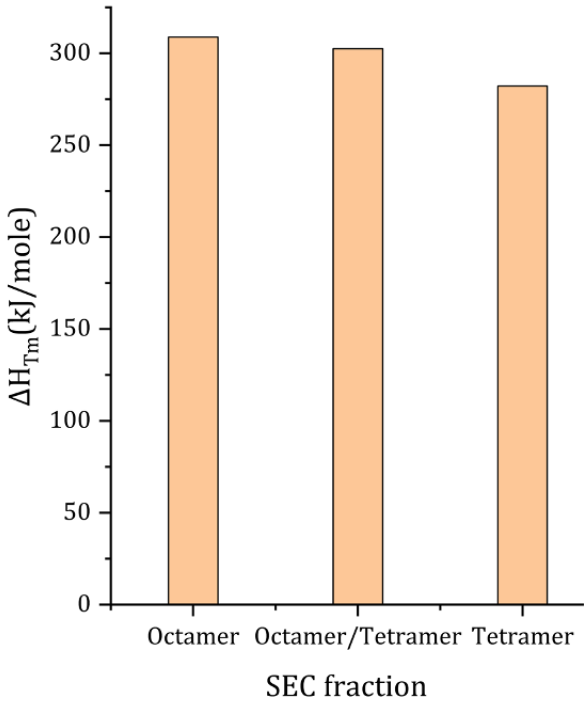
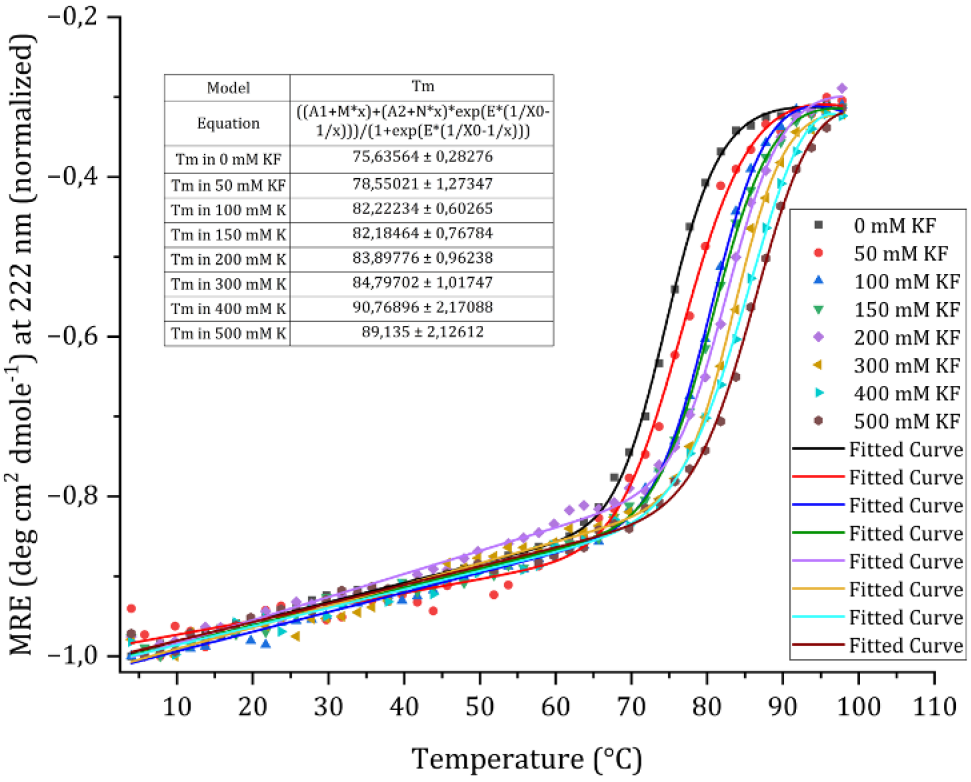


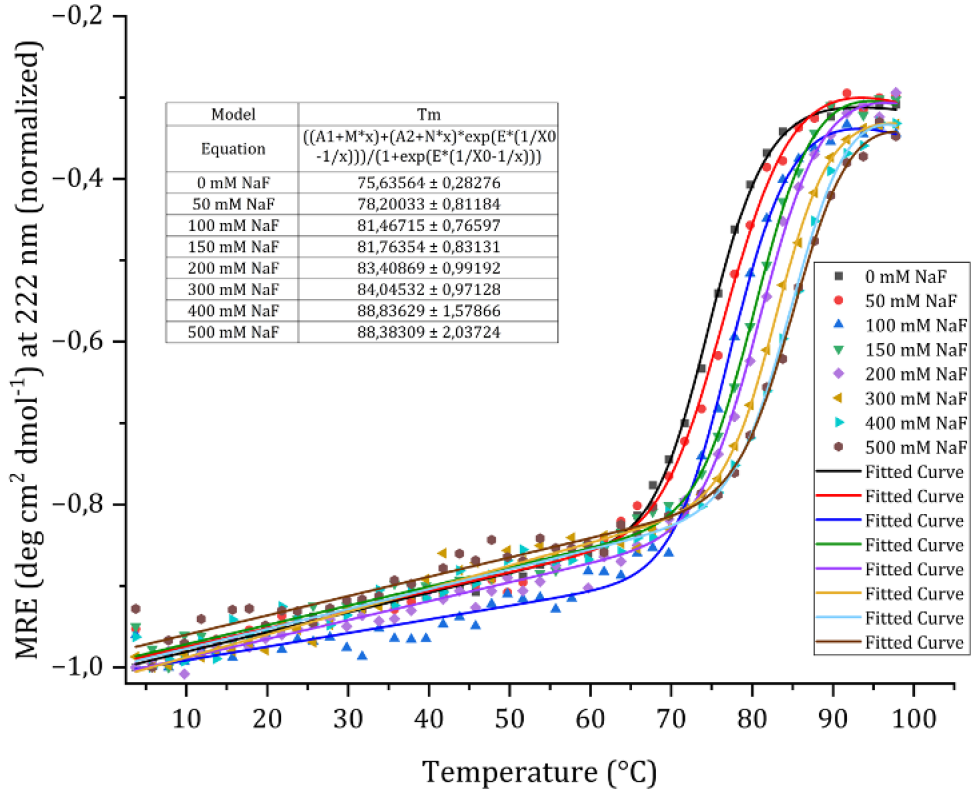
Figure 3.2.5.2.6 Enthalpy changes for different SEC fractions of mutated Par-4 CC EXS

When the temperature gradient was performed for the protein samples in different salt concentrations, i.e., KF and NaF salts, an increase in the Tm was observed and was found to be nearly proportional to the salt concentrations (Figures 3.2.5.2.7 A and B). The results indicated that there was no significant difference in the melting temperature of the protein when the same concentration of KF and NaF salts were included in the CD buffer. The highest melting temperature was observed when the protein was exposed to a concentration of 400 mM for both KF and NaF salts.

A.



B.



Figures 3.2.5.2.7 Thermal stability of mutated Par-4 CC EXS in different salts concentrations

A. Tm in different KF salt concentrations B. Tm in different NaF salt concentrations

3.2.6 Crystallization results of mutated Par-4 CC EXS

Various crystallization screens were employed to crystallize mutated Par-4 CC EXS (section 2.8). The plates were incubated at 16°C. After 4-5 months, needle-like crystals emerged in Nextal pH clear in three separate experiments, all sharing the same precipitant (0.8 M ammonium sulfate) but varying pH conditions: 100 mM HEPES pH 7.0, 100 mM Tris-HCl pH 8.0, and 100 mM Bicine. These crystals were then prepared for X-ray diffraction experiments.

Multiple cryoprotectants, including glycerol and polyethylene glycol (PEG) at different concentrations (10%-30%), were tested for crystal cryopreservation. The needle-shaped crystals were mounted using cryo-loops with a diameter ranging from 0.1 to 0.2 millimeters, flash-frozen in liquid nitrogen, and subsequently transferred to transport containers for storage in liquid nitrogen.

The crystals were subjected to diffraction testing at the PETRA III P11 beamline. One of the crystals exhibited weak diffraction, with only a few spots that were barely discernible. To improve resolution, efforts were made to enhance crystal quality.

The confirmation of protein crystal formation was achieved through mounting, washing, and running the crystals on SDS-PAGE (Figure 3.2.6.1). The results revealed that the crystallized protein was degraded, with a molecular weight around 13 kDa, whereas the mutated Par-4 CC EXS has a molecular weight of 23.980 kDa.

Subsequently, the identification of the crystallized protein was conducted using LC/MS mass spectrometry. In-gel trypsin digestion was performed to determine the N-terminal and C-terminal of the protein. The LC/MS results showed that only the C-terminal region of the mutated Par-4 CC EXS could be identified, with the amino acid sequence found to be 246-332 (numbering according to the full-length Par-4 protein) (Figure 3.2.6.2).

Comparison with the previously crystallized CC domain (254-332) revealed that only 8 additional amino acids were included in the needle crystal compared to the CC

domain crystal. Notably, among these 10 residues, Ser249, the phosphorylation site of Akt1, was included.

A fine crystallization screen was prepared with ammonium sulfate (AmSO4) concentrations ranging from 0.6 M to 1.2 M, utilizing various buffers including MES buffer pH 5.8-6.5, HEPES buffer pH 6.7-8.6, Tris-HCl buffer pH 7.0-9.2, and Bicine buffer pH 7.6-9.0. Crystallization plates were then set up and incubated at different temperatures, including 16°C, 19°C, and 22°C

In addition, in situ proteolysis using different weight ratios (w/w) (1:100, 1:250, 1:500, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:7000, and 1:10000) of trypsin: protein were tested for trypsin cleavage activity and cleaved fragments size determination at 19°C, and 22°C. Figures (3.2.6.3 A and B) and (3.2.6.4 A and B) show the SDS-PAGE and anti-strep western blot for trypsin cleavage activity with different trypsin: protein ratios. The results indicated that all tested trypsin ratios produced fragments of the same size, with differences in band intensities attributed to varying amounts of trypsin used.

Furthermore, plates without trypsin showed no crystal growth even after 5 months of incubation at the aforementioned temperatures. The addition of trypsin accelerated crystal growth, with the rate depending on the temperature. Needle crystals appeared after 3 months in plates incubated at 16°C and 19°C, while it took 1-2 weeks for crystals to form in plates incubated at 22°C (Figure 3.2.6.5).

Among all the conditions in the fine screen, needle crystals only appeared in the conditions of 0.8-1.0 M ammonium sulfate (AmSO4), pH 7-8 HEPES, and Tris-HCl, with larger crystals observed in 0.8 M AmSO4. Regarding the trypsin ratios, needle crystals appeared within 2 weeks at ratios of 1:1000 and 1:2000, while it took longer for crystals to form at a ratio of 1:5000, resulting in fewer and smaller crystals. No crystals appeared at a ratio of 1:500.

To increase crystal size, higher protein concentrations of 25 mg/ml and 35 mg/ml were used, but this only resulted in an increase in the number of crystals with smaller

sizes. Seeding, both homogeneous and heterogeneous, did not improve crystal size despite using various seed dilutions.

The needle crystals were subjected to diffraction testing at different beamlines at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France

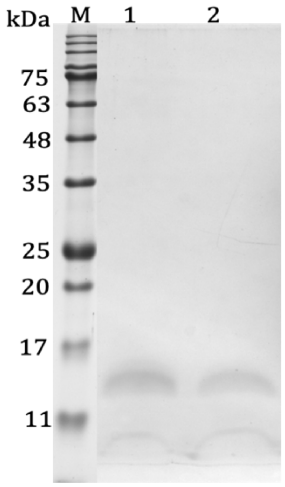


Figure 3.2.6.1 SDS-PAGE of needle crystal grown in Tris-HCl pH 8.0, and 0.8 M AmSO4

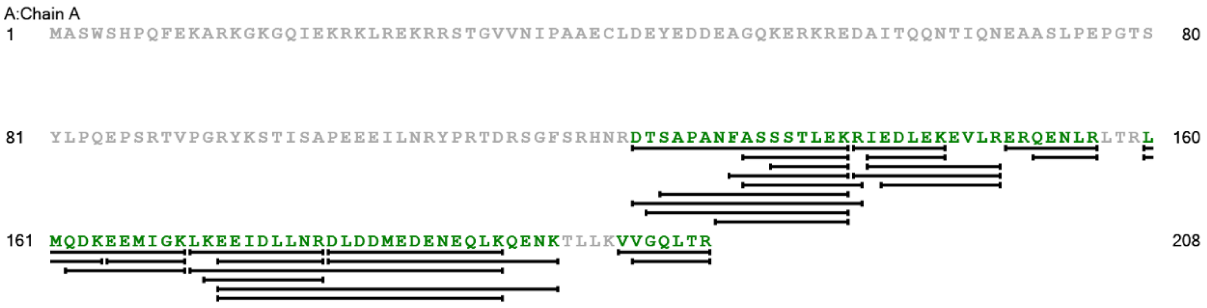


Figure 3.2.6.2 LC-MS/MS analysis of the needle crystal grown in Tris-HCl pH 8.0, and 0.8 M AmSO4

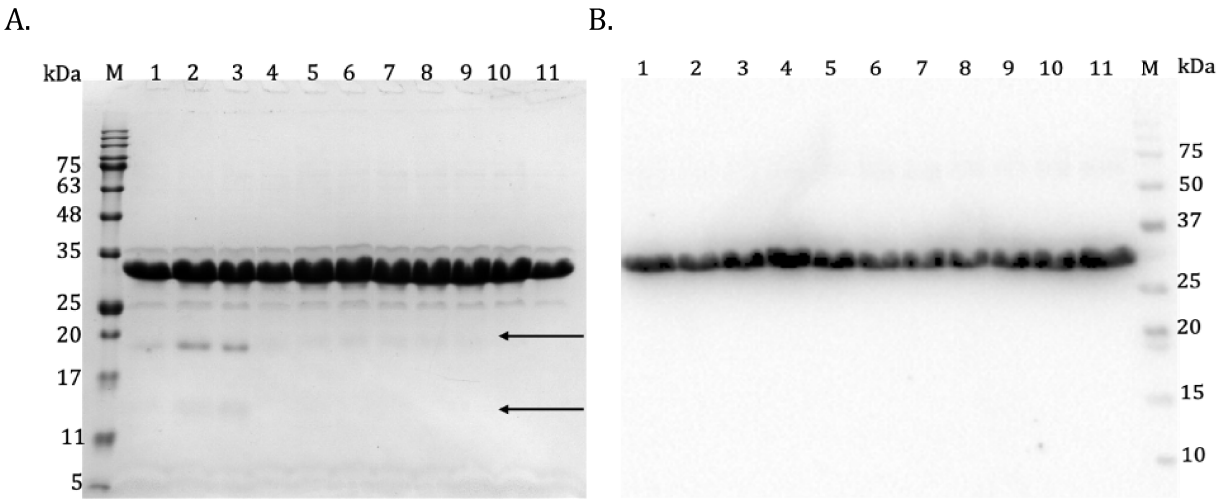


Figure 3.2.6.3 SDS-PAGE and Anti-strep western blot analysis for trypsin activity on mutated Par-4 CC EXS at 22 °C

A. SDS-PAGE, different ratios (w/w) trypsin: protein at 22 °C, 1. 1:100, 2. 1:250, 3. 1:500, 4. 1:1000, 5. 1:2000, 6. 1:3000, 7. 1:4000, 8. 1:5000, 9. 1:7000, 1:10000, 11. 1:0 (control), B. Anti-strep western blot for different ratios (w/w) trypsin: protein at 22 °C, same samples and order.

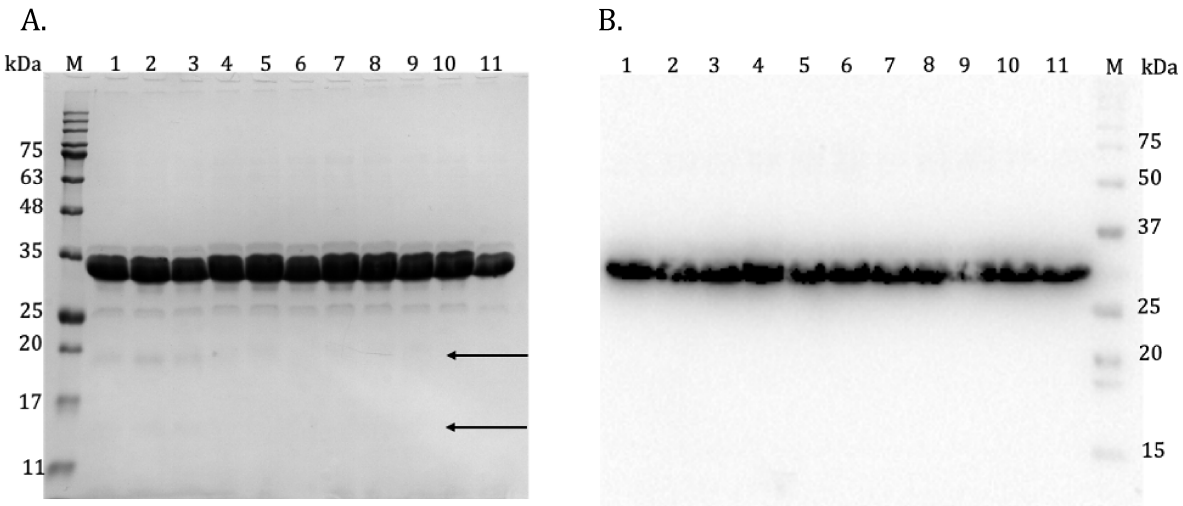


Figure 3.2.6.4 SDS-PAGE and Anti-strep western blot analysis for trypsin activity on mutated Par-4 CC EXS at 19 °C

A. SDS-PAGE, different ratios (w/w) trypsin: protein at 19 °C, 1. 1:100, 2. 1:250, 3. 1:500, 4. 1:1000, 5. 1:2000, 6. 1:3000, 7. 1:4000, 8. 1:5000, 9. 1:7000, 1:10000, 11. 1:0 (control), B. Anti-strep western blot for different ratios (w/w) trypsin: protein at 19 °C, same samples and order.

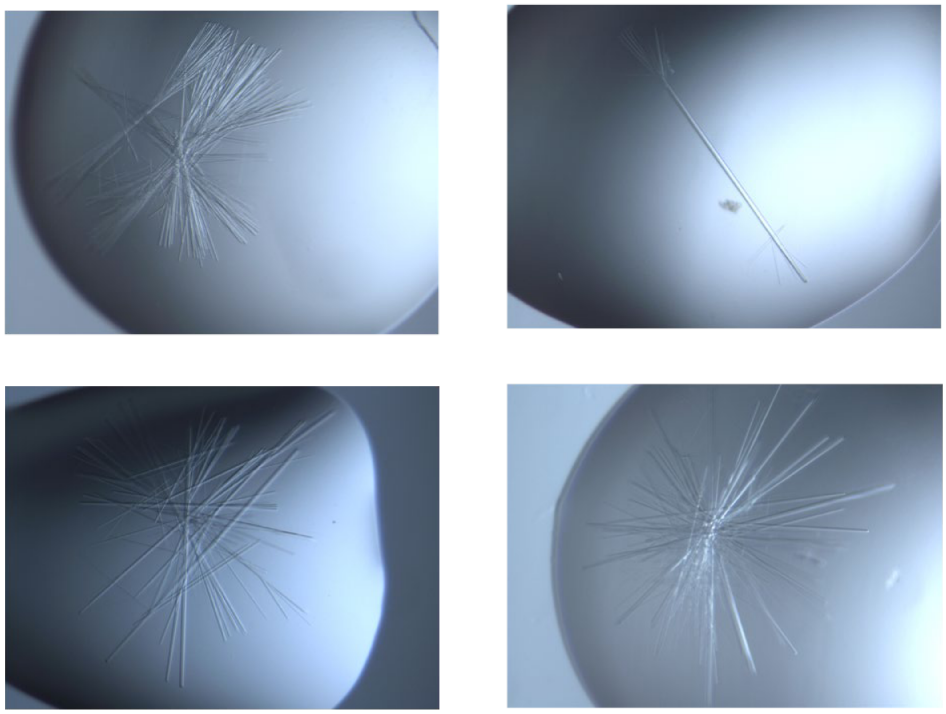


Figure 3.2.6.5 Mutated Par-4 CC EXS needle crystals

3.2.7 Alphafold model of Par-4 CC EXS

The Par-4 CC EXS protein contains three main parts: the CC domain, the SAC domain and a linker between them. Since the CC domain forms a dimer (Tiruttani Subhramanyam et al., 2017), the Par-4 CC EXS was modeled as a dimer using AlphaFold. Additionally, the SAC domain and the linker were also modeled as a dimer for the prediction (figures 3.2.7.1 A,B). However, when the SAC domain model was mapped to the Par-4 CC EXS model, they did not match even though both models share an alpha helix structure at the residues (Lys 14-Thr 31) (figure 3.2.7.2). Therefore, the Par-4 CC EXS, the SAC domain and the linker were modeled as a monomers for the predication (figures 3.2.7.3 A,B and C).

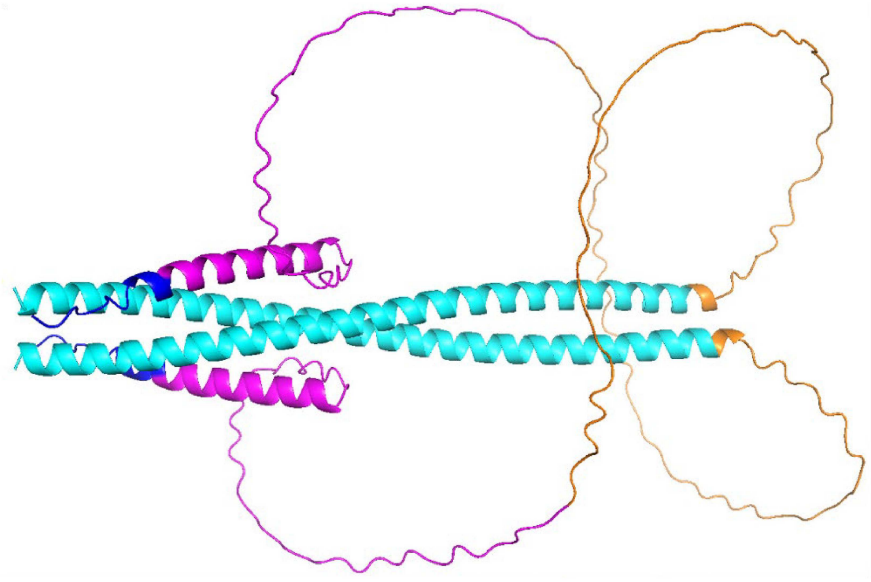
Intrestingly, both predicted models share the same alpha helix residues (Glu 49- Ala 71) which is different from the dimer models. Mapping the SAC domain to Par-4 CC EXS as monomers worked well. To have the fitted models as a dimers, the monomeric Par-4 CC EXS was mapped to the dimeric Par-4 CC EXS at their CC domain, as the

structure of this domain is consereved and crystallographically solved (figure 3.2.7.4). The monomric SAC domain model was then mapped and used to generate the dimer based on the generated dimeric Par-4 CC EXS. The generated dimer of Par-4 CC EXS model showed an additional helical structure at residues (Lys 16- Arg 28), that was not obsereved in the generated dimer of SAC domain model (figure 3.2.7.5).

Based on the AlphaFold model of Par-4 CC EXS, it can be concluded that the presence of the CC domain in this model helped, for unclear reasons, to induce folding in the SAC domain, as the linker in both models (as a monomer and generated dimer) did not show a clear secondary structure.

The model also shows different regions, especially in the linker, that have low tendency to form secondary structures, indicating instability in the folding of this protein. This could explain the different structures and oligomerization of this protein, as previously shown in the biophysical characterization results (i.e., SEC, DLS, and CD). Additionally, this instability also explains the long time needed for needle crystal formation, as these regions are floppy and unstable, requiring long incubation times to be properly degraded.

A.



B.

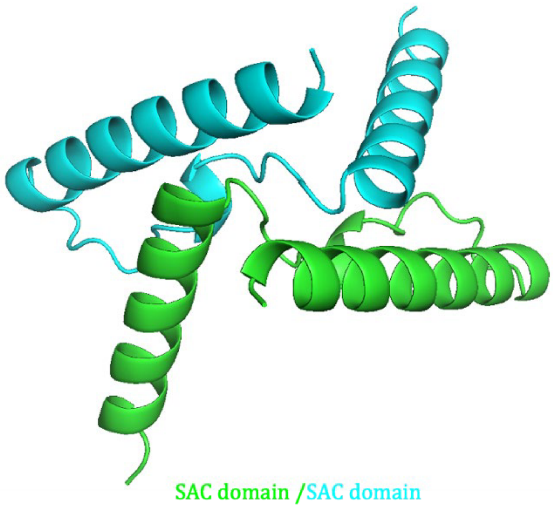


Figure 3.2.7.1 Alphafold model prediction of Par-4 CC EXS and SAC domain as dimers

A. Par-4 CC EXS dimer model, the dimeric domains have the same color. B. SAC domain dimer model, the dimeric domains have different color.

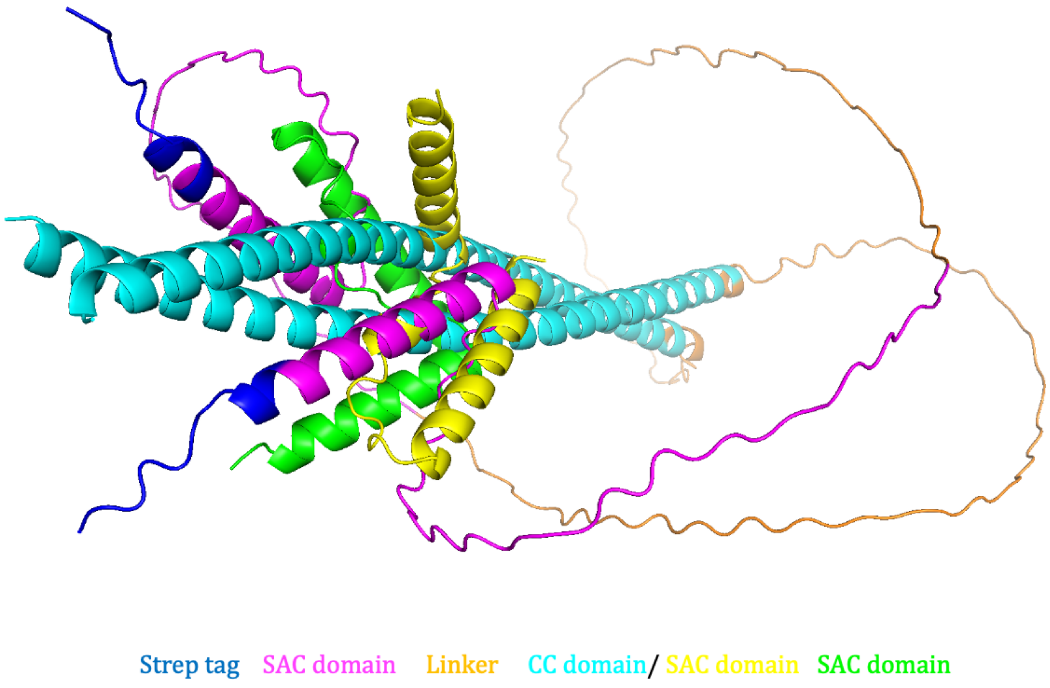
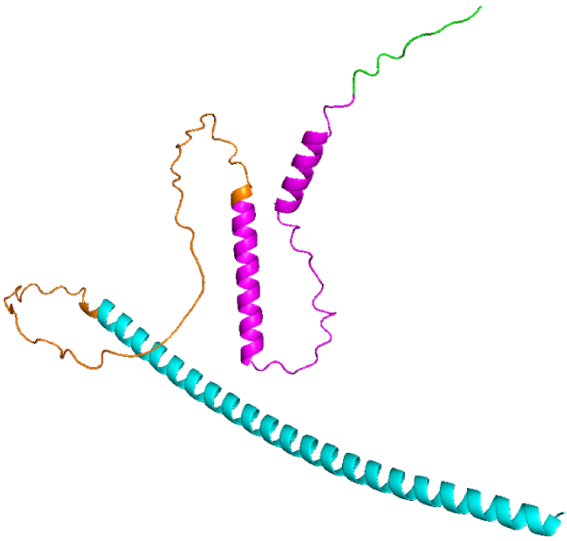
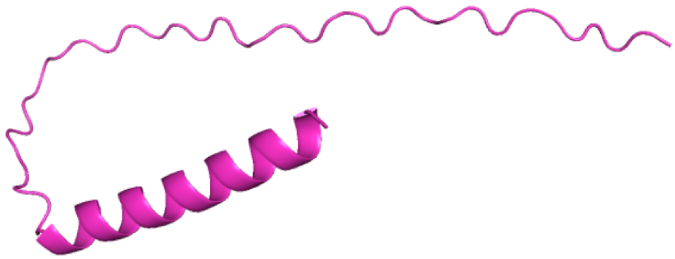


Figure 3.2.7.2 Fitted Alphafold dimeric SAC domain to Alphafold dimeric Par-4 CC EXS

A.



B.



C.

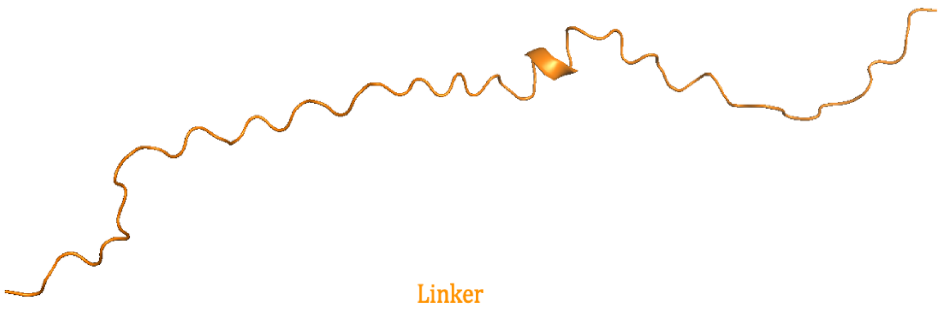
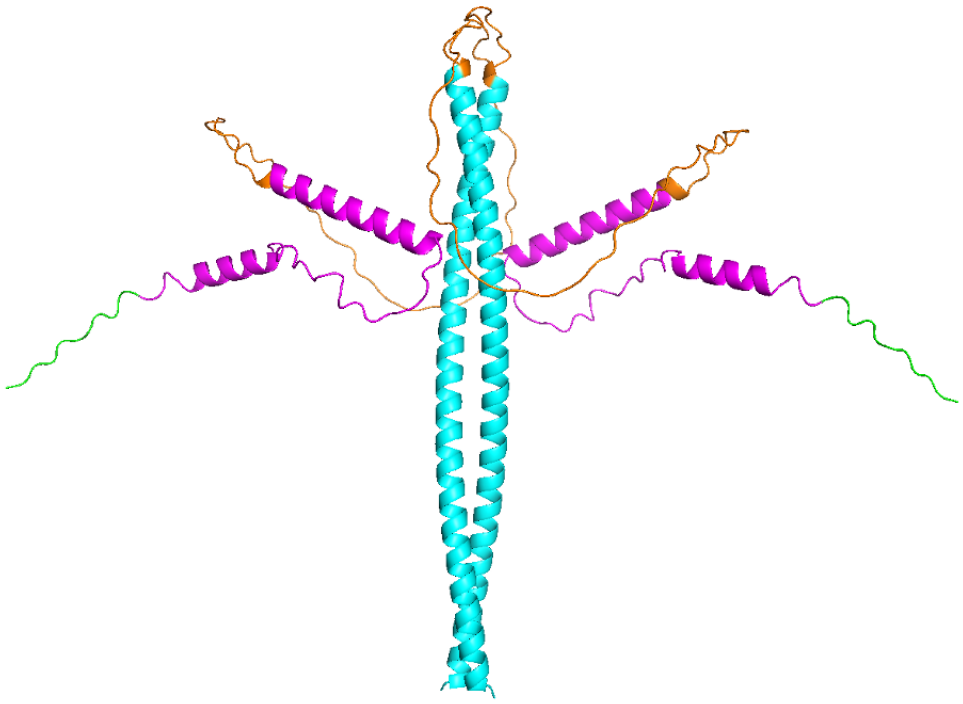


Figure 3.2.7.3 Alphafold model prediction of Par-4 CC EXS, SAC domain and linker as monomers

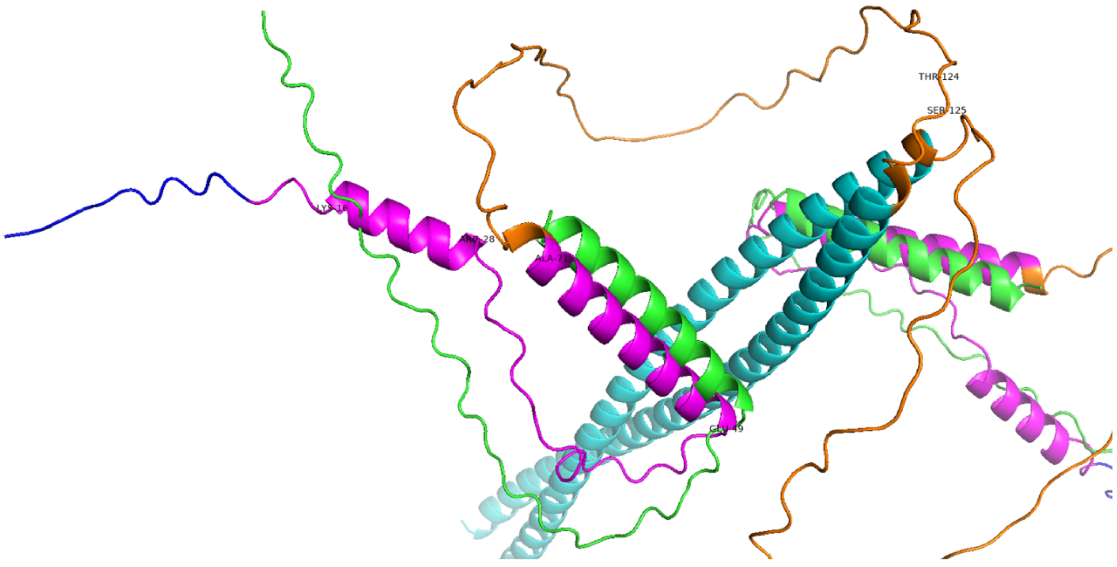
A. Alphafold model prediction of Par-4 CC EXS as a monomer, B. Alphafold model prediction of SAC domain as a monomer, C. Alphafold model prediction of linker as a monomer.

A.



Strep tag SAC domain Linker CC domain

Figure 3.2.7.4 Par-4 CC EXS dimer model generated from its monomeric Alphafold model



Strep tag SAC domain Linker CC domain/ mapped SAC domain

Figure 3.2.7.5 Fitted generated dimer of SAC domain to generated dimer of Par-4 CC EXS

Par-4 CC EXS had to be mutated in this study because the native protein was cleaved by acidolysis. The biophysical characterization results showed that the acidic condition is more favorable for protein folding and stability compared to the neutral condition.

The results also revealed that the protein exhibited different oligomerization states, conformations, and secondary structure content, indicating the instability of protein folding. Despite the mutation made to stabilize the protein, its crystal was found to be the result of a degraded protein, with 8 additional residues found compared to the CC domain crystal. This also suggests the instability of protein folding.

AlphaFold predicted some alpha-helical structure in the SAC domain, whereas no structure was predicted in the linker. This prediction was confirmed as the crystallized protein was cleaved in the linker region, which was also found to be unfolded in the model generated by AlphaFold.

In conclusion, extending the region from the CC domain to the SAC domain could induce folding in the overall structure in solution, especially in the SAC domain based on the AlphaFold model. However, the crystal still requires the removal of the N-terminal region as it is not stable. Therefore, another hypothesis was proposed to stabilize the protein by binding it with a partner protein.

3.3 Investigation of the Interaction Between Par-4 CC EXS and the P62 PB1 Domain

Although CD spectroscopy and Alphafold results showed some secondary structures in the extended region of the CC domain, the Par-4 CC EXS crystal was found to be degraded, indicating that the full protein is not properly folded, especially the linker region. Additionally, despite all attempts, the crystal did not diffract. Therefore, the next approach was to bind Par-4 with a partner protein to obtain the structure of the CC and the extended region in a more stabilized and folded state.

Previous studies have reported that the leucine zipper domain of Par-4 binds to the PB1 domain of P62 (Chang et al., 2002). This interaction was validated through a pull-down assay using cell lysates containing these domains. Recent research on the P62 PB1 domain has shown that it forms filaments, and the structure of these filaments has been determined using Cryo-EM techniques (Tarafder et al., 2019). Additionally, it has been demonstrated that Par-4 binding to actin filaments is essential for Par-4/Dlk-mediated apoptosis (Vetterkind et al., 2005) , suggesting that Par-4 could also bind to P62 PB1 filament.

Based on these findings, we hypothesized that P62 PB1 filaments could stabilize regions outside the CC domain, even though the P62 PB1 domain binds specifically to the leucine zipper domain. Subsequently, the complex protein-filament can be used in Cryo-EM to determine the structure of Par-4 CC EXS.

This section presents the results of the interaction study between Par-4 CC EXS and both the monomeric and filament forms of the P62 PB1 domain.

3.3.1 The interaction between Par-4 CC EXS and P61 PB1

The interaction between Par-4 CC EXS and P62 PB1 domain was investigated in two ways, based on the P62 PB1 protein conformation; as a monomer and as a filament. The interaction was investigated via pull down assay under the neutral pH 7.5.

In order to have the P62 filament, the MBP should be cleaved and NaCl concentration in the buffer should be lowered.

3.3.1.1 Par-4 CC EXS interaction with monomeric MBP-P61 PB1

In this part of the study, Par-4 CC EXS was incubated with Strep-Tactin XT resin for 2 hours at 4 °C. Following this, MBP-P62 PB1 was added to the Par-4 CC EXS-resin complex in a molar ratio of 1:1 (Par-4 CC EXS: MBP-P62 PB1) and incubated overnight at 4 °C. As a negative control, MBP alone was added to the Par-4 CC EXS-resin complex at the same molar ratio and incubated under the same conditions.

The figures 3.3.1.1.1 A and B and figures 3.3.1.1.1 C and D show the SDS-PAGE and western blot analysis results of the pull-down assay for Par-4 CC EXS with MBP-P62 PB1 and with MBP alone respectively. The results indicate that MBP-P62 PB1 was present in the elution fractions of Par-4 CC EXS, whereas no MBP protein was observed in the elution fractions. This suggests that Par-4 CC EXS interacts specifically with the P62 PB1 domain but not with the MBP protein. In addition, the bands intenisty look similar which could indicate that the binding is in 1:1 ratio.

To ensure that the presence of MBP-P62 PB1 in the elution fractions was not due to nonspecific binding to the Strep-Tactin resin, MBP-P62 PB1 and MBP proteins were incubated with the resin separately. The figures 3.3.1.1.2 A and B show that neither of these proteins were present in the elution fractions, further confirming the specific interaction between Par-4 CC EXS and MBP-P62 PB1.

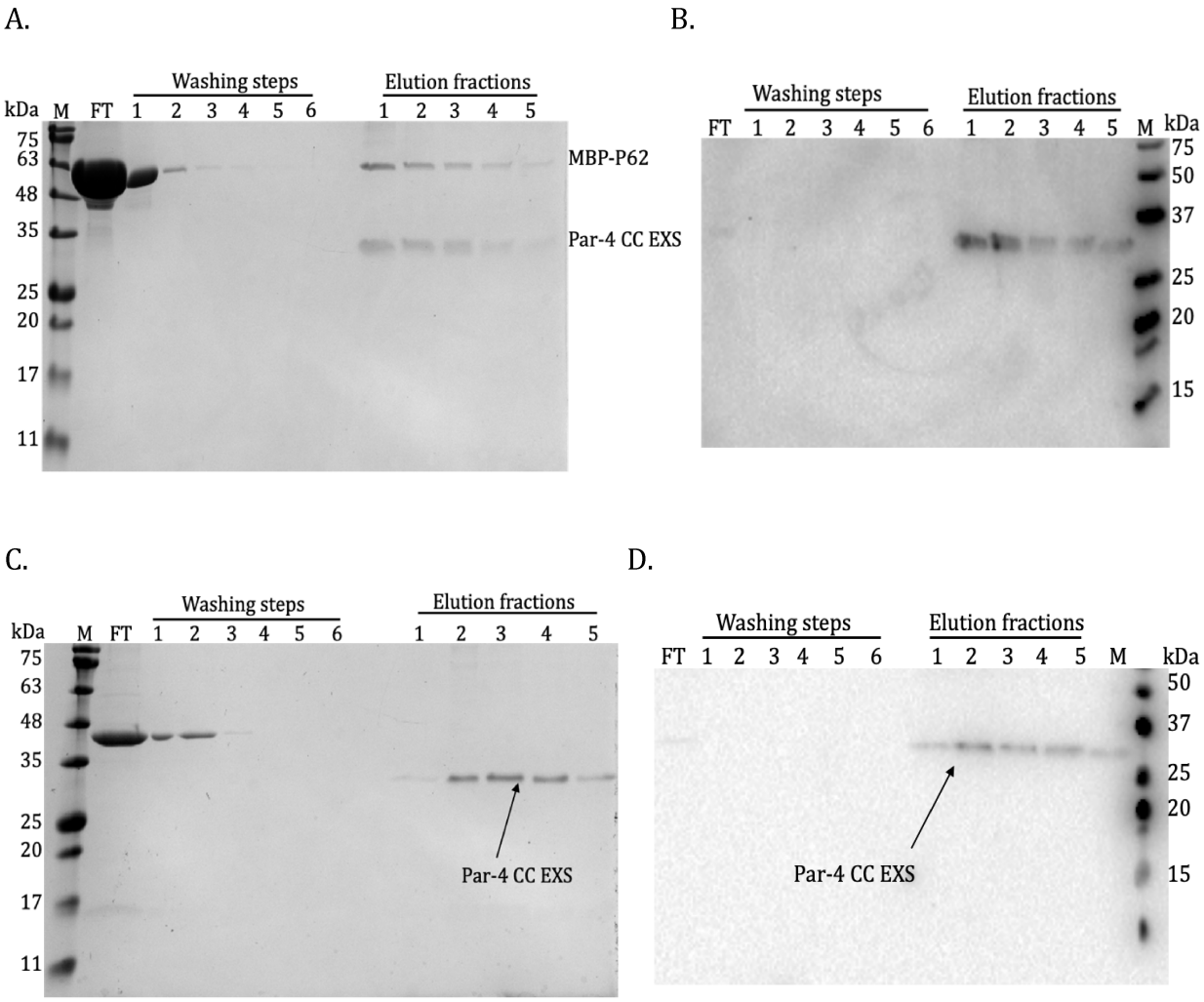


Figure 3.3.1.1.1 SDS-PAGE and Western blot analysis of Pull-down assay for Par-4 CC EXS with monomeric MBP-P62 PB1 and with MBP

A. SDS-PAGE analysis of the pull-down assay for Par-4 CC EXS and MBP-P62 PB1, FT: Flow through, 1-6: washing steps, 1-5: elution fractions. B. Western blot analysis of the pull-down assay for Par-4 CC EXS and MBP-P62 PB1, FT: Flow through, 1-6: washing steps, 1-5: elution fractions. C. SDS-PAGE analysis of the pull-down assay for Par-4 CC EXS and MBP, FT: Flow through, 1-6: washing steps, 1-5: elution fractions. D. Western blot analysis of the pull-down assay for Par-4 CC EXS and MBP, FT: Flow through, 1-6: washing steps, 1-5: elution fractions.

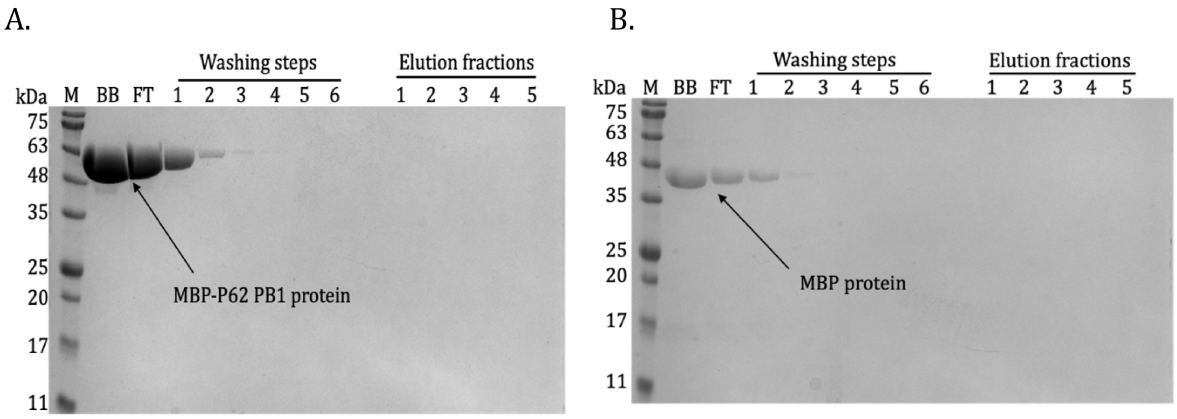


Figure 3.3.1.1.2 SDS-PAGE analysis of the possible nonspecific interaction of Strep-Tactin resin with monomeric MBP-P62 PB1 and MBP (negative controls)

A. SDS-PAGE analysis of the possible nonspecific interaction of Strep-Tactin resin with monomeric MBP-P62 PB1, BB: before binding, FT: Flow through, 1-6: washing steps, 1-5: elution fractions. B. SDS-PAGE analysis of the possible nonspecific interaction of Strep-Tactin resin with MBP, BB: before binding, FT: Flow through, 1-6: washing steps, 1-5: elution fractions.

3.3.1.2 Par-4 CC EXS interaction with P62 PB1 filament

In this part of the study, P62 filaments were obtained by removing the MBP tag and lowering the salt concentration. According to the published protocol for P62 filament formation, it was not possible to keep the filament soluble; once the MBP tag was removed, the P62 PB1 filament formed but aggregated, while the solubility of monomeric P62 was concentration-dependent. Therefore, to investigate whether Par-4 CC EXS interacts with the filament, Par-4 CC EXS was added to the cleavage reaction (MBP-P62 PB1 with TEV) and incubated overnight at 4 °C. The next day, the mixture was centrifuged to pellet the filament. The pellet was washed several times and analyzed by SDS-PAGE and western blot. The supernatant was incubated with pre-equilibrated Strep-Tactin resin for 2 hours at 4 °C, followed by a pull-down assay similar to the aforementioned section.

The figures 3.3.1.2.1 A and B show the SDS-PAGE and western blot analysis results of the pellet (filament) and the pull-down assay. The pellet analysis indicates that Par-4 CC EXS is present in the pellet, along with MBP and uncleaved MBP-P62 PB1. Comparing the band intensities, the Par-4 CC EXS band intensity is similar to the

uncleaved MBP-P62 PB1 but not the P62 PB1 band, suggesting that Par-4 CC EXS preferably binds to the monomeric P62 PB1. This similarity in band intensity was also observed in the monomeric pull-down assay (previous section).

The analysis of the pull-down results of the supernatant also revealed the presence of P62 PB1 in the elution fractions (Figure 3.3.1.2.2 A and B). However, the band intensities did not align with the previous observation of a 1:1 binding ratio. This discrepancy could be due to the presence of small filaments remaining in the supernatant after centrifugation. Additionally, it suggests that the actual binding ratio might be higher than previously observed, as the MBP tag, which is significantly larger (42 kDa) compared to the P62 PB1 domain (14 kDa), might have interfered with the binding accuracy in earlier experiments. Alternatively, it could indicate that the binding occurs between Par-4 CC EXS and the filamentous form of P62 PB1, which would explain the higher band intensity of P62 PB1 compared to Par-4 CC EXS, as multiple monomeric P62 PB1 units are required to form the filament.

In conclusion, the pull-down assay confirmed the interaction between Par-4 CC EXS and the P62 PB1 domain. However, it is difficult to determine which form of P62 PB1 is required for this interaction based on the pull-down assay alone.

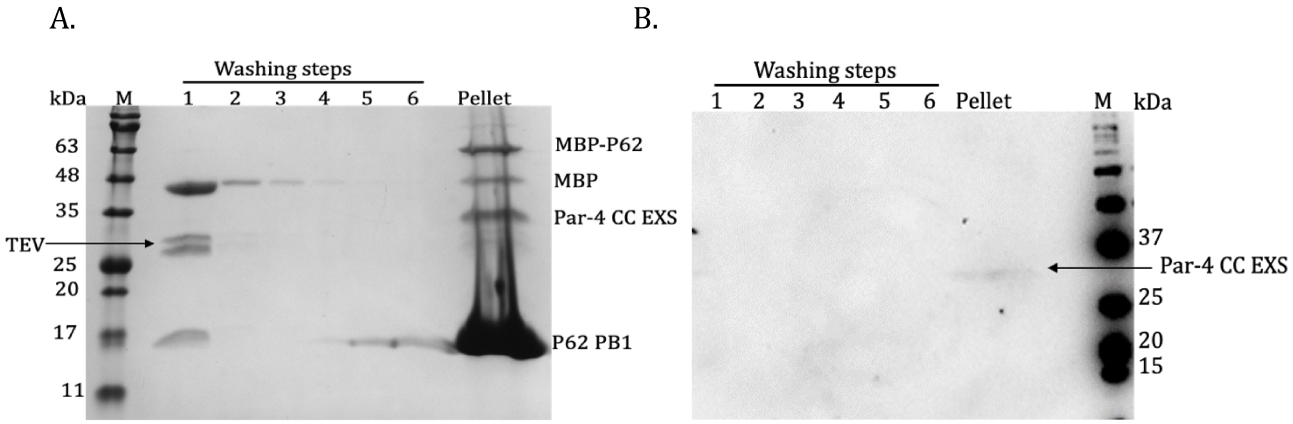


Figure 3.3.1.2.1 SDS-PAGE and Western blot analysis of the P62 pellet after cleavage centrifugation step

A. SDS-PAGE analysis of the P62 pellet, 1-6: washing steps, pellet. B. Western blot analysis of the P62 pellet, 1-6: washing steps, pellet. B. Western blot analysis

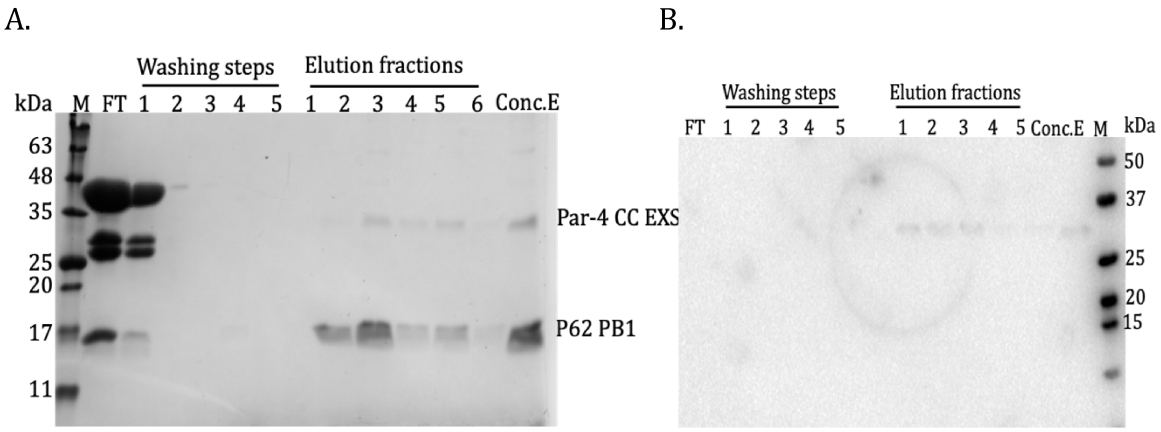


Figure 3.3.1.2.2 SDS-PAGE and Western blot analysis of Pull-down assay for the supernatant of cleavage centrifugation step

A. SDS-PAGE analysis of the pull-down assay for the supernatant of cleavage centrifugation step, FT: Flow through, 1-5: washing steps, 1-6: elution fractions, conc. E: concentrated elution fraction. B. Western blot analysis of the pull-down assay for the supernatant of cleavage centrifugation step, FT: Flow through, 1-5: washing steps, 1-5: elution fractions, conc. E: concentrated elution fraction.

Results from the pull-down assay confirmed the interaction between Par-4 CC EXS and P62 PB1. The findings showed that Par-4 CC EXS interacts with both monomeric and filament forms of the PB1 domain. However, binding of Par-4 CC EXS to P62 PB1 resulted in aggregation and precipitation under all conditions, making this complex unsuitable for crystallographic analysis.

4 Discussion

4.1 Overexpression, purification, and biophysical characterization of the proteins

Par-4 is a 332 amino acid proapoptotic protein with tumor suppressor activity, considered an intrinsically disordered protein. Its C-terminal (residues 254-332) mainly forms a coiled-coil structure and is involved in many interactions. And it had been suggested that it is an intrinsically unstructured protein. The structure of the region beyond the CC domain, including the SAC domain which is responsible for the selective induction of apoptosis in cancer cells, has not yet been determined and is predicted to be disordered. While the structure beyond the CC domain remains elusive, the crystallographic analysis of the CC domain has highlighted disorder in this region. Therefore, this study aimed to characterize both the CC domain and the extended region including the SAC domain. To achieve this, several approaches were employed. Phosphorylation-mimicking and salt bridge mutations were introduced into the CC domain, and an extended Par-4 variant (Par-4 CC EXS) was constructed to stabilize the conformation and analyze it. Additionally, the partner protein P62 PB1 was expressed and purified for binding studies.

In this study, a strep-tagged Par-4 CC construct (residues 240–332) was used as a template to create phosphorylation-mimicking mutants, salt bridge removal mutants, and a charge repulsion removal mutant. The phosphorylation-mimicking mutants included Ser249 and Thr248, while the other mutations were located within the CC domain. The phosphorylation-mimicking mutants were T248D, T248E, S249D, and S249E, whereas the salt bridge removal mutants were E268Q, E274Q, K288M, and K317M. The charge repulsion removal mutant was D305N. Additionally, the DNA sequence of the Par-4 gene corresponding to residues 136-332 was cloned to create a construct known as Par-4 CC EXS, which includes an extended region from the CC domain to the SAC domain which allowed to investigate the role of the different domains with regard to folding.

The E. coli system is always the first choice for recombinant proteins production due to its cost-effectiveness, rapid growth, and well-established molecular tools. E. coli yields high quantities of properly folded proteins, making it versatile for research with easy scale-up capabilities (Chen, 2012; Francis & Page, 2010). E. coli BL21(DE3) RIL cells strain was utilized for the proteins production. This strain is able to express proteins with high level of AGG/AGA (arginine), AUA (isoleucine), and CUA (leucine) codons since it co-expresses the tRNAs related to these codons (Nouri et al., 2016).

Utilizing the E. coli expression system to express the CC domain was sufficient to have the protein soluble and properly folded (Tiruttani Subhramanyam et al., 2014). Therefore, the expression of the mutants was achieved same as the CC domain expression, i.e., at 37°C. All the mutants were found to be soluble and folded.

This did not apply for Par-4 CC EXS expression as the protein was found to be in the inclusion bodies, and therefore, the expression had to be induced at lower temperature 16 °C. This was because, at lower temperature, the metabolic processes of E. coli become slow, and protein expression decreases, that will make the protein folding more efficient and therefore, the protein becomes soluble (Kaur et al., 2018).

The protein folding issues were reported in literature and in our own work which led to the suggestion that par-4 may be a natively unfolded protein especially in the region out of the CC domain (Dr. Jan Kubicek's PhD thesis, Heinrich-Heine University of Duesseldorf, 2005 and Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, Heinrich-Heine University of Duesseldorf, 2011). Therefore, the insect cell expression system was utilized beside the E. coli system for Par-4 CC EXS protein expression since it offers some post-translation modifications including phosphorylation of serine (Héricourt et al., 2000), tyrosine sulfation (Thompson et al., 2017), Glycosylation and palmitoylation (Grünewald et al., 1996), that may help in inducing the folding and the stability of protein.

The purification of the CC domain and the mutants was achieved according to the published paper (Tiruttani Subhramanyam et al., 2014). The SEC profile showed a

single peak for all the proteins implying that they exist as one species in solution. The apparent molecular weight given by SEC for all the proteins was 95.30 kDa, which corresponds to an octamer (12.3 kDa * 8 = 98.4 kDa). The molecular weights provided by SEC were inherently imprecise, as they also depend on the protein's shape. Having all the proteins at the same elution volume indicates that the mutations did not affect the oligomerization behavior of the proteins, while the agreement between apparent Mw and calculated suggested that the octamer adopts a globular shape.

Since the CC domain is included in the Par-4 CC EXS, the purification under the acidic condition was the first choice as this condition is favorable for the CC domain folding (Dutta et al., 2001). However, a degradation occurred in the purified protein from both expression systems under this condition. The degradation was due to the acidic pH, which cleaved the peptide bond between aspartic acid (D) at position 209 and the following proline (P). Previous studies have shown that this peptide bond is susceptible to cleavage under acidic conditions, even in week acidic environment (Lamed et al., 2001; Yang, 2015).

The mechanism underlying this process, as depicted in Figure (4.1.1), bears a resemblance to aspartimide formation. In this mechanism, the amide nitrogen atom originating from the proline backbone initiates an attack on the carboxyl side chain of the preceding aspartic acid residue, leading to the formation of an instable cationic imide intermediate 18 (Piszkiewicz et al., 1970). This intermediate readily undergoes hydrolysis, resulting in the formation of peptidyl fragments 19 and 20, where the C-terminus corresponds to the aspartic acid residue, and the N-terminus corresponds to the proline residue.

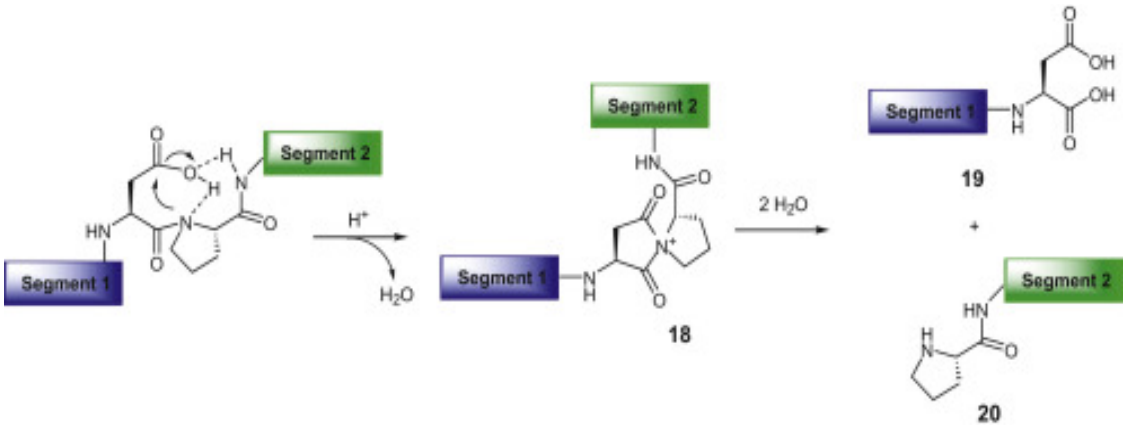


Figure 4.1.1 Proposed mechanism of -Asp-Pro- acidolysis (Figure adapted from (Yang, 2015)

Interestingly, while Par-4 CC EXS contains two peptide fragments of DP moieties located at the D200-P201, and D209-P210 (numbering according to rat full length Par-4), only the D209-P210 peptide suffers from the acidolysis. This was confirmed by mutating the D200 and D209 separately into E in addition to a third mutant that has both residues mutated. The mutation D200E did not make the protein resistant to hydrolysis, whereas D209E and the one with double mutations did. This indicates that the mutation D209E was sufficient to make the protein acidic cleavage resistant. It was reported that secondary structure surrounding the DP sequence is crucial in determining if the acidolysis side reaction is likely to occur or not (Shimon et al., 1997) (Lamed et al., 2001). This implies that the peptide at residues D209-P210 is likely to be exposed to the solvent.

The SEC profiles of the mutated CC EXS from both expression systems under acidic pH (Figures 3.2.3.3.1.2 and Figure 3.2.3.3.2.2) revealed that the protein eluted at two peaks with apparent molecular weights of 189.4 kDa and 100.4 kDa implying that utlizing the Hi5 insect cell expression system did not affect the protein behavior, oligomerization, and homogeneity. The theoretical molecular weight of the mutated Par-4 CC EXS is 24.09783 kDa. This indicates that the protein elutes as both an octamer (24.09783 kDa * 8 = 192.8 kDa) and a tetramer (24.09783 kDa * 4 = 96.4 kDa), with a higher presence of the octameric form. This observation indicates the presence of different oligomeric species in solution.

Since the SEC profile of the CC domain, which constitutes 39.8% of the Par-4 CC EXS, exhibited a single homogenized peak suggesting the stability of the CC domain under acidic conditions, and because no previous studies have reported any interactions between the CC domain and regions beyond it, this could imply that the regions outside of the CC domain lack proper folding, leading to structural variability, and different oligomerization states.

In addition, the oligomerization of Par-4 CC EXS was investigated by introducing varying sodium chloride salt concentrations (50 mM, 100 mM, 200 mM, 300 mM, and 400 mM) into the SEC buffer under acidic conditions. The results revealed a rightward shift in the SEC peak compared to the control without salt, indicating a reduction in protein oligomerization. However, an additional peak emerged in the void volume of the SEC even at the lowest salt concentration (50 mM), suggesting the formation of a higher oligomerization state (aggregation). Notably, as the salt concentration increased, the lower molecular weight in the SEC decreased while the void volume peak increased, indicating a salt concentration-dependent increase in higher oligomerization. These observations lead to the conclusion that sodium chloride salt negatively impacts protein homogeneity due to oligomerization, even at low concentrations.

Additionally, Par-4 CC EXS was purified under neutral conditions to further explore its biophysical characteristics. Upon elution from the affinity column, a visible turbidity was noted in the elution fractions, observed both by the naked eye and through light scattering in a UV-visible spectrophotometer. Although a centrifugation step at 20,000 xg clarified the fractions momentarily, the turbidity reappeared and intensified during protein concentration, resulting in significant protein loss due to aggregation. This highlights the concentration-dependent nature of the aggregation phenomenon.

Analysis of the pellet obtained from the centrifugation step confirmed the presence of Par-4 CC EXS protein. This aggregation behavior was not unexpected, considering previous findings that both the CC domain and SAC are inherently unfolded under physiological conditions (Libich et al., 2009) (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, Heinrich-Heine University of Duesseldorf, 2011)

In addition, the protein eluted from the SEC (3.2.3.2.1) at two peaks with apparent molecular weights of 32.9 kDa and 16.2 kDa. The molecular weight 32.9 kDa could indicate that the protein existed as a dimer, whereas the molecular 16.2 kDa could indicate a monomer. The deviation from the expected Mw could be explained by a non-globular shape. This indicates the protein is not properly folded or unfolded at all under the neutral pH. These results emphasize the previous findings about the CC domain and SAC domain regarding the folding issue and stability of protein at neutral pH ((Dutta et al., 2001; Tiruttani Subhramanyam et al., 2014) (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, Heinrich-Heine University of Duesseldorf, 2011).

The highest concentration that could be reached under neutral condition was below 0.25 mg/ml, whereas, under acidic condition, a concentration of 40 mg/ml was reached and it could be higher. This indicates that acidic pH induces folding and stability for Par-4 CC EXS compared to neutral pH involving Asp and Glu protonation. Based on these observations, purifying the protein at acidic condition was the best choice to study the structural characterization of the protein.

The purified proteins obtained from SEC were subjected to SDS-PAGE to assess their purity. A single band was consistently observed for all proteins, indicating high purity levels. However, it's noteworthy that the estimated molecular weight of Par-4 CC EXS, based on the SDS-PAGE marker proteins, was approximately 34 kDa, notably higher than its theoretical molecular weight of 24.09783 kDa. This discrepancy may be attributed to the unusually high content of basic amino acids in Par-4 CC EXS, causing the protein to migrate at a higher apparent molecular weight.

In contrast, both the CC domain and the mutants exhibited bands with estimated molecular weights around 12 kDa, which closely align with their theoretical molecular weight of 12.3 kDa. This consistency suggests accurate migration of these proteins on the SDS-PAGE gel.

After obtaining the purified proteins in good quantity, the identity and integrity of the proteins were checked. As all the proteins have an N-terminal strep-tag, the identity

was confirmed through the western blot technique. All the protein showed a positive signal indicating the intactness of the N-terminal end. However, the N to C terminal completeness was not confirmed yet. This was investigated using LC-MS/mass, and mass spectroscopy. The LC-MS/mass showed the intactness of the C-terminal end for all the proteins. However, the N-terminal methionine residue in all the proteins expressed in E. coli was found to be cleaved, whereas, it was intact in the protein expressed in Hi5 cells. This cleavage of methionine could be due to methionine aminopeptidase, which Co-translationally removes the N-terminal methionine from nascent proteins. The N-terminal methionine is often cleaved when the second residue in the primary sequence is small and uncharged (Met-Ala-, Cys, Gly, Pro, Ser, Thr, or Val)(Varland et al., 2015). All the proteins that were expressed in E. coli i.e., CC domain, mutants and mutated Par-4 CC EXS, contain Met-Ala at the first two positions in their sequences, whereas, the mutated Par-4 CC EXS sequence that was expressed in Hi5 cells contains Met-Trp at the same positions, which could explain the reason of detecting the methionine in the protein expressed in Hi5 cells but not in the ones expressed in E. coli. The LC-MS/mass did not show any post-translation modification for any of the proteins.

The mass spectroscopy for the mutated Par-4 CC EXS gave a molecular weight of 23.966 kDa which was 131 Da less than the theoretically calculated molecular weight (24.09783 kDa). The same difference in the molecular weight was also observed for the CC domain mass spectroscopy result. This difference in the molecular weight i.e., 131 Da matches the molecular weight of the methionine amino acid, which emphasizes the result obtained LC-MS/mass. Based on the discussion made for mass spectroscopy results, it can be concluded that the theoretical molecular weights match the ones obtained from mass spectroscopy. This indicates the absence of any post-translation modifications in these proteins.

The LC-MS/mass and mass spectroscopy confirmed the N to C terminal sequence coverage and the molecular weight, concluding that the full-length mutated Par-4 CC EXS, CC domain, and the mutants were expressed and purified successfully.

After confirming the identity and integrity of the mutated Par-4 CC EXS expressed in Hi5 cells, further validation of the mass spectrometry results was pursued through a de-glycosylation reaction using PNGase and Endo H enzymes. The results showed no discernible difference in mobility between the enzyme-treated samples and the controls on SDS-PAGE, implying that the sites are not accessible for glycosylation. This confirms the findings from the mass spectrometry analysis, suggesting that post-translational modification is not the natural pathway to stabilize or induce protein structure/folding.

The obtained proteins were then subjected to biophysical characterization techniques to evaluate the impact of the mutations on the CC domain and the effect of pH on mutated Par-4 CC EXS.

DLS analysis was conducted to examine the oligomerization state and homogeneity of the proteins. The results revealed a high oligomerization state for the CC domain and the mutants. Interestingly, with the exception of mutants T248D, E268Q, E274Q, and D305N, all proteins exhibited low polydispersity, suggesting homogeneous particles in solution and more defined state for these mutations.

Notably, mutant D305N exhibited the highest number of molecules in its oligomerization state, while mutants E268Q, E274Q, and K317M showed the lowest, indicating that polydispersity and oligomerization are not strictly correlated. Mutant D305N also had the most significant impact on particle size, number of molecules, and polydispersity, possibly due to the removal of repulsive forces between the helices, suggesting a dependence of oligomerization on charge interactions. Furthermore, mutants E268Q, E274Q, and K317M exhibited a lower number of molecules in their oligomers compared to the CC domain, suggesting that these mutations may reduce the protein's tendency for high oligomerization, implying that these charged amino acids participate in intermolecular interactions that promote oligomerization or aggregation. Additionally, phosphorylation-mimicking mutants, especially S249E, showed higher oligomerization states compared to the CC domain, indicating that charge interactions contribute to protein oligomerization and aggregation.

The DLS result for the mutated Par-4 CC EXS showed that the samples under the neutral condition have significant larger radius size, and hence higher molecular weight (aggregates) compared to the ones under the acidic condition. This implies that Asp and Glu protonation decreases aggregation. This difference was observed once the protein was eluted from the affinity column as the turbidity was observed in the elution fractions and increased during protein concentrating under the neutral pH, whereas at the acidic pH, the elution fractions were cleared even in the high concentrated sample. The results also showed that the protein in E2 sample under the neutral pH had the largest radius size and molecular weight among the samples at same pH, which is true since most of the protein eluted in E2 sample, which indicates that the aggregation is concentration dependent. It can be concluded from these results that the acidic condition is more favorable for the protein folding and stability compared to the neutral pH which led to protein aggregation with putatively randomized assembly. This explains the behavior of the CC domain (C-terminal of the Par-4 CC EXS) which is found to form natively unfolded state under physiological conditions (Libich et al., 2009), but a self-associated coiled coil structure can be induced under acidic condition (Dutta et al., 2001). This is also true for the SAC domain (N-terminal of the Par-4 CC EXS) as it exhibits a random coil structure at neutral conditions (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, University of Duesseldorf, 2011).

Although the acidic condition induced better folding and stability for the protein compared to the neutral condition, the samples exhibited high polydispersity under both conditions. This indicates the presence of different species (oligomers) in the solutions, with a greater number of conformations than in the CC domain, making it harder to obtain crystals. The reason of the high polydispersity under the neutral pH is mostly due to the unfolded state of the protein (all the domains), as previously mentioned, which led to different conformations of the protein, high aggregates and hence to high polydispersity, whereas under the acidic pH, the reason of the high polydispersity is mostly due to the region beyond the CC domain, which lacks a certain folding pattern, which for sure leads to high polydispersity but smaller particle size and hence molecular weight compared to the ones under neutral pH since the CC domain is properly folded. Overall, the neutral condition has a negative effect on the

protein folding and solubility as the protein aggregated, which could not be observed at acidic condition.

The effect of different sodium chloride concentrations on the mutated Par-4 CC EXS under the acidic condition was also assessed using the DLS. The results showed almost no difference in protein behavior and heterogeneity in absence and presence of up to 200 mM NaCl except a relative increase in radius size and molecular weight in samples with salt. However, at NaCl concentrations above 200 mM (300 mM-1 M), the results showed a significant increase in the particle size and molecular weight indicating the presence of an extreme oligomerization state (aggregates) at these concentrations. Interestingly, only the concentrations; 400 mM, 750 mM and 1 M showed a low polydispersity which indicates the protein is homogenized in these concentrations and in more defined state. The results also showed that the 500 mM NaCl has the extreme negative effect on the protein as the largest radius size, molecular weight and polydispersity were resulted in this concentration. Based on these results, it can be concluded that NaCl salt has a negative effect on protein's particle size and molecular weight even at low concentration. But it has a positive effect on protein's homogeneity at high concentration.

The secondary structure of the obtained proteins, including the CC domain and mutants under acidic condition, as well as ,the mutated Par-4 CC EXS under both conditions, was investigated using far-UV CD spectrometry. Additionally, the protein's stability and thermodynamic parameters were assessed through the determination of the melting temperature, calculated from the CD data at the wavelength 222 nm. Temperature scans were conducted from 4 °C to 98 °C at a rate of change of 2 °C.

The CD results of the CC domain and its mutants displayed an α-helical profile with double minima at 208 nm and 222 nm, along with a maximum around 193 nm. The spectra also showed a MRE222/MRE208 ratio greater than 1 for all samples, indicating a coiled coil conformation (Schwalbe et al., 2010). This implies that none of the mutations disrupted the coiled coil structure. Compared to the wild-type CC domain, all the phosphorylation-mimicking mutations showed an increase in the number of residues forming helical structures. This suggests that these mutations induce some helical structural induction, similar to the results obtained from CD

spectroscopy and NMR for the phosphorylation site Thr155 (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, Heinrich-Heine University of Duesseldorf, 2011). Additionally, the charge repulsion mutation D305N showed the highest number of residues forming helical structures among all the mutants, likely due to the removal of charge repulsion. Moreover, the salt bridge removal mutants E268Q, E274Q, and K317M also showed an increase in the number of residues forming helical structures, indicating that these mutations are essential for helix folding. In contrast, the mutant K288M showed one residue less in the helical structure compared to the native CC domain, implying its minimal effect on helical structure formation.

The secondary structure stability of the CC domain and its mutants results showed a significant increase in melting temperature for the mutants compared to the wildtype CC domain. Notably, phosphorylation-mimicking mutants exhibited nearly identical melting temperatures, averaging around 77.4±0.2°C, while mutations within the CC domain resulted in a lower melting temperature, averaging around 75.0±0.5°C. Interestingly, the K317M mutant displayed the highest melting temperature among all mutants, reaching 80.0°C. This indicates that the phosphorylation-mimicking mutants and the K317M mutant have less effect on stability, implying there is less interaction between residues in these areas.

A more comprehensive assessment of the stability of the CC domain and the mutants was achieved by calculating the changes in free energy and the heat required for each reaction to occur. All the mutants showed a significant increase in ΔG unfolding compared to the CC domain. Closer inspection reveals that all mutations exhibited an increase in ΔCp, suggesting that the native amino acid composition allows structural changes more easily than any of the mutants. This is possibly a functional feature of the CC domain.

The CD result of Par-4 CC EXS under neutral condition showed it is predominant unfolded monomeric state with the presence of some residual structure.This was not unexpected as the main domains in the protein, i.e., CC domain and SAC domain were found to be unfolded in monomeric state under the neutral pH (Dutta et al., 2001; Libich et al., 2009) (Dr. Udaya Kumar Tiruttani Subhramanyan's PhD thesis, Heinrich-

Heine University of Duesseldorf, 2011). The unfolded state could explain the aggregation behavior in the affinity elution fractions which was confirmed by the DLS results as all showed a high aggregation of the protein. Based on these results and observations, it can be concluded that the neutral pH is not favorable for Par-4 CC EXS.

On the contrary, CD results for the mutated Par-4 CC EXS under acidic conditions revealed improved protein folding compared to the neutral state. Samples from the SEC peaks, specifically the octamer peak, octamer/tetramer, and tetramer peak at the acidic condition from both expression systems, E. coli and Hi5 cells, were utilized for CD measurements. Importantly, CD spectra deconvolution showed no differences between the protein fractions purified from E. coli and those from Hi5 cells. This suggests that utilizing the Hi5 insect cell expression system did not induce additional structural folding in the protein. These findings align with the results from LC-MS/mass and de-glycosylation reactions, as both showed the absence of post-translational modifications and consistent with the SEC profile obtained from E. coli.

All protein samples from both expressions exhibited α-helical profile, with only the octamer spectra showing a maximum around 193 nm indicating the presence of more α-helix. Indeed, CD spectra deconvolution unveiled distinct secondary structure content percentages among them, with the highest α-helix content observed in the octamer, followed by the octamer/tetramer, and the least in the tetramer. This decrease in α-helix content accompanied with an increase in turn content among the samples from same expression system. The CC domain which is presented with 39.8% of the mutated Par-4 CC EXS protein has shown to form a coiled coil structure under acidic condition. This means that the difference in the secondary structure contents among the samples is caused by the region beyond the CC domain, indicating varying degree of folding and stability for the part out of the CC domain. This also explains the high polydispersity for these samples in DLS results.

The CD results also showed a salt dependent decrease in α-helix and increase in β sheet, turn as well as in random coiled contents when different concentrations of Sodium fluoride, Potassium fluoride, and Magnesium sulfate salts were included in CD buffer. Notably, the significant change in the secondary structure resulted at the

concentration above 200 mM of both KF and NaF, and in 50 mM $MgSO_4$. Regarding the Magnesium sulfate salt, only one concentration was used; 50 mM, since the higher concentrations led to Magnesium sulfate precipitation, and hence protein aggregation. Interestingly, the SEC profile result and DLS result (previously discussed) also showed a significant high oligomerization at salt concentrations above 200 mM. This suggests that a salt concentration exceeding 200 mM altered the secondary structure of the protein by increasing the β-sheet and turn contents while decreasing the α-helix contents. This structural shift likely contributed to the formation of high protein oligomerization (aggregates).

The secondary structure stability of the different SEC fractions under acidic conditions was assessed through the melting temperature computed from CD ttemperature scans, ranging from 4°C to 98°C, with a rate of change of 2 degrees. As a result, the measured melting temperatures for the three fractions differed slightly, with values of 75.92°C, 75.1°C, and 73.95°C for the octamer, octamer/tetramer, and tetramer SEC fractions, respectively, indicating that the octamer fraction exhibited the highest stability among the fractions. A more comprehensive assessment of the stability of the three SEC fractions was achieved by calculating the changes in free energy and the heat required for each reaction to occur. The calculated parameters showed that the octamer fraction required the highest amount of heat, and hence free energy, to reach the equilibrium between the folded and denatured protein fractions, occurring at a temperature of 75.92°C, confirming its highest stability. The reason for the different stability among the fractions is mostly due to the percentage of α-helix present in each fraction. A higher percentage of α-helix contributes to greater protein stability. Moreover, the occurrence of polar or charged side chains in the helix can facilitate additional interactions with other side chains in the helix or with other elements outside of the helical structure, imparting further stability (Pauling et al., 1951). However, the thermal stability curves at 222 nm for the three fractions appear almost identical, suggesting that the oligomerization state of the protein is not significantly influenced by changes in secondary structure.

In addition, the stability of mutated Par-4 CC EXS protein was also assessed in the presence of different salts concentrations; i.e., KF and NaF salts. Even though the CD

spectra deconvolution in the presence of these salts showed a salt concentration-dependent decrease in the α-helix and increase in the β sheet contents which affects inversely on Tm (Kumar et al., 2000), the measured melting temperatures increased in a concentration-dependent manner for both salts. The reason of this increase in the Tm could be due to the salt concentration dependent increase in the oligomerization state that was observed in DLS results. Moreover, there was no significant difference in the melting temperature of the protein when the same concentration of KF and NaF salts were included in the CD buffer. Though salt increases stability, the nature of the salt did not have any significant effect.

4.2 Structural analysis of CC domain mutants

The successful crystallization experiments of the CC domain and mutants used 37 and 38% t-butanol in 100 mM sodium citrate/citric acid at pH 6.1. The behavior of the mutants was similar to that of the native CC domain. Crystals appeared for the CC domain and the phosphorylation-mimicking mutants, except for T248E, and appeared tetragonal with packing identical to the CC domain crystals. No crystals formed from the other mutations in the CC domain, implying that charges have a profound effect on molecular interactions.

The structures of the T248D, S249D, and S249E mutants were determined using the CC domain model. During the indexing process for all mutants, the tetragonal metrics consistently showed the highest symmetry with minimal lattice distortion. Integration and scaling were performed in the space group $P4_32_12$. All mutant models in this space group displayed seven helices (five long and two short), similar to the CC domain model. It is noteworthy that the organization of the protein in the asymmetric unit, respectively unit cell appeared to be unrelated to the tetrameric and octameric states observed in solution.

Despite several refinement cycles with different parameters, the electron density for the helices' N-terminal or the short helices' C-terminal did not improve and hindered model completion in space group $P4_32_12$. Although the R-free values of the refined models in the space group $P4_32_12$ were 0.3418, 0.3055, and 0.2996 for the T248D,

S249D, and S249E mutants respectively, which is an indication of general correctness of the models, integrating the data into different space group with different symmetry reductions ($P2_12_12_1$, and P1) was necessary to analyze if, all parts of the molecule obey the tetragonal symmetry or crystallographic symmetry had to be reduced.

For the possible subgroup $P2_12_12_1$, one crystallographic symmetry element of $P4_32_12$ must be replaced by a non-crystallographic symmetry element, which requires the presence of a twinning operator (see Table 4.2.1).

When twinned refinement was performed, the R-free values increased to 0.3703, 0.3572, and 0.3257 for T248D, S249D, and S249E respectively, indicating that either the twining is not correct or the space group. However, this did not help in resolving the structure of the missing parts as the density did not improve. Therefore, it was necessary to reduce the symmetry further to space group P1.

Table 4.2.1 Twinning analysis of the mutants' data in the space group P212121 by different methods in phenix

	Britton alpha	H alpha	Maximum likelihood method (ML alpha)	Operator
T248D	0.435	0.458	0.478	K, h, -l
S249D	0.437	0.458	0.478	K, h, -l
S249E	0.409	0.419	0.478	K, h, -l

Reducing the symmetry from the space group $P4_32_12$ to P1 by applying the $P4_32_12$ symmetry operations resulted in 56 chains: 40 long helices and 16 short helices. The models in P1 space group showed some fragmented electron density in stark contrast to data processed in $P4_32_12$ which indicates that the N-terminal part of the CC domain is disordered, but the density patterns for different molecules show a different distribution of conformation for the N-terminus. The R free values for the models in space group P1 were 0.3444, 0.3056, and 0.3081 for the mutants T248D, S249D, and

S249E, respectively. Superposing each chain with its symmetries showed disorder in the orientation and length of the built residues at their N-terminals (see figure 3.1.6.3.1), explaining the absence of the helices' N-terminal in space group $P4_32_12$. For the short helices, the density did not allow elongation of the C-terminals at all implying that even the helices observed for the other molecules are disordered.

The model in space group P1 revealed that the phosphorylation mimicking did help in folding and stabilizing this region. Since the CD spectroscopy results showed a 100 % helical structure for all the protein, it could be concluded that the appearance of long and short dimers must be related to their different environment within the asymmetric unit, i.e. the stability of the structure depends on the packing of molecules and crystal formation.

4.3 Crystallization of mutated Par-4 CC EXS

Different crystallization screens were used for initial screening test including, pH clear, PEGs, PEGIIs cores, MPD, and $AmSO_4$. Two different protein concentrations were used for setting the crystallization plates; 10 mg/ml and 15 mg/ml.

Small needle crystals appeared only under certain conditions of the pH clear after 4-5 months and their size did not improve. The P11 Petra beamline at DESY in Hamburg was used to test these needle crystals for diffraction. One of the crystals weakly diffracted with few spots that were hardly observed. These needle crystals were analyzed on SDS-PAGE and subsequently LC-MS/mass, and they turned out to be a degraded protein. These results suggest the presence of flexible part in the protein that hindered the full protein to crystallize and thus got degraded because the plates were incubated for a long time.

Comparing to the CC domain that was crystallized before (254-332) (Tiruttani Subhramanyam et al., 2014), only 8 additional amino acids were included to the needle crystals. These findings also suggest that the region outside the CC domain is flexible and lacks stable folding, contributing to the protein's heterogeneity. This further emphasizes the findings from DLS, which revealed high polydispersity and

sample heterogeneity, and CD, which indicated variations in secondary structure conformations.

Resetting crystallization plates with the same conditions but a higher protein concentration, up to 35 mg/ml, did not accelerate the crystal growth rate. Consequently, various optimization strategies were employed to enhance crystal size and growth speed. These strategies included adjusting crystallization conditions, employing in situ proteolysis, and using both homogeneous and heterogeneous seeding approaches. Additionally, the plates were subjected to incubation at different temperatures. The in-situ proteolysis method, involving trypsin at a 1:1000 protein ratio and an increased incubation temperature of 22 °C, resulted in larger needle crystals within 1-2 weeks. However, the growth was observed in length rather than width, leading to no improvement in crystal diffraction. The conditions yielding crystals were 0.8/1.0 M AmSO4, 100 mM HEPES pH 7.0/100 mM Tris-HCl pH 8.0. For cryopreservation, a different percentages (10%-30%) of glycerol and polyethylene glycol (PEG) were used to prevent ice crystal formation. Despite employing these approaches, none proved successful in producing large and diffracting crystals.

The Alphafold model of Par-4 CC EXS predicted the linker region between the CC domain and the SAC domain to be unstructured, whereas about 60 % of the residues in the SAC domain were predicted to form helical structrue which is in a good agreement with results obtained from CD spectroscopy. This leaves the linker region to be the high susceptible for degradation. Indeed, the needle crystal was found to be degraded in the linker region close the CC domain.

Forming needle crystals from degraded protein emphasizes the disordered state of the protein in the region outside of the CC domain, which must be cleaved for crystal formation and packing. Nevertheless, the presence of additional structure in the region out CC domain in solution, as shown by CD spectroscopy, indicates protein folding in that region especially in SAC domain as predicted by Alphafold.

4.4 Stabilizing the region beyond the CC domain by binding it to the P62 PB1 domain

The previously discussed results have shown that the region beyond the CC domain, especially the linker region, lacks a well-defined folded conformation, hindering crystal formation for that region to the SAC domain. The PB1 domain of the P62 protein has been reported to bind to the leucine zipper domain of Par-4. Moreover, recent research on the P62 PB1 domain has shown that it forms filaments, with their structure determined using Cryo-EM techniques. These P62 PB1 filaments could stabilize regions outside the CC domain, even though the P62 PB1 domain binds specifically to the leucine zipper domain. Subsequently, the complex protein-filament could be used in Cryo-EM to determine the structure of Par-4 CC EXS.

Results obtained from the pull-down assay confirmed the interaction between Par-4 CC EXS and P62 PB1. The results showed that Par-4 CC EXS interacts with monomeric and filament forms of PB1 domain. However, upon binding of Par-4 CC EXS to the p62 PB1, aggregation and precipitation were observed under all conditions which renders this complex unsuitable for crystallographic analysis.

Conclusion

The comprehensive analysis conducted in this study offers several key insights into the behavior and structural properties of Par-4 constructs. Par-4 constructs exhibit maximum structure and minimum oligomerization under acidic conditions, unlike the conditions present in the nucleus or cytoplasm where the protein typically resides. This observation underscores the unique behavior of Par-4 in different pH environments and suggests that its structural properties are highly sensitive to pH changes.

The detailed analysis of the effect of single amino acid mutations showed that the protein is stabilized by any residue exchange, suggesting that structural instability is a biological feature. This points to a multitude of possible interactions with the protein, which has been identified to bind to Par-4. Therefore, it is suggested that intrinsically unstructured small proteins like Par-4 can achieve multiple binding modes through restructuring, rather than through different binding epitopes as commonly found in larger proteins.

In this case well defined structures require the study of the respective complexes of Par-4. Even that may prove difficult because the observed oligomerization, respectively aggregation poses a serious problem for the classical methods crystallography, NMR and cryo-EM. A major result of this work is the control of the aggregation state which may allow study of the octamer or bigger complexes with partner proteins by EM. Though it remains to be determined which of the multitude of partner proteins form complexes under acidic conditions.

CD spectroscopy findings indicate a structured region in solution, particularly within the SAC domain, aligning with computational predictions like those from Alphafold. This observation not only sheds light on the potential folding propensity of specific segments but also validates computational models predicting structural features of Par-4. Additionally, it suggests that the different oligomerization states and aggregation could be due to the instability of the tertiary structure.

Stabilizing the linker region through interactions with partner proteins presents a crucial avenue for future research. By leveraging the binding capabilities of interacting proteins, it may be possible to confer stability to the inherently disordered linker region, facilitating structural studies.

In conclusion, while the structural instability of Par-4 presents significant challenges, it also highlights the protein's unique biological adaptability. Future studies utilizing advanced techniques and exploring protein interactions hold the potential to unravel the elusive structure of Par-4, advancing our understanding of its functional mechanisms.

Summary

Par-4, a tumor suppressor protein, plays a crucial role in inducing apoptosis in cancer cells, primarily through its N-terminal SAC domain. Structurally, Par-4 is recognized as an intrinsically disordered protein with two significant domains: the SAC domain responsible for selective apoptosis and the C-terminal coiled-coil (CC) domain, essential for multiple interactions. While the structure beyond the CC domain remains elusive, the crystallographic analysis of the CC domain has highlighted disorder in this region. This study aims to characterize both the CC domain and the extended region including the SAC domain. To achieve this, several approaches were employed. Phosphorylation-mimicking, salt bridge removal, and charge repulsion removal mutations were introduced into the CC domain, and an extended Par-4 variant (Par-4 CC EXS) was constructed to stabilize the conformation and analyze it. Expressions in insect cells and E. coli were utilized for Par-4 CC EXS expression. Additionally, the partner protein P62 PB1 was expressed and purified for binding studies. The Par-4 proteins, tagged at the N-terminal with strep tags, were purified using affinity chromatography and size exclusion chromatography (SEC). However, during purification, it was observed that Par-4 CC EXS underwent degradation under acidic conditions, necessitating mutation to enhance its resistance to acidolysis.

SEC analysis revealed that all CC domain mutants exhibited similar elution volumes, suggesting minimal impact of the mutations on protein oligomerization. In contrast, SEC analysis of Par-4 CC EXS under acidic conditions unveiled diverse oligomeric forms. Additionally, the presence of Sodium chloride exacerbated aggregation behavior in Par-4 CC EXS, as evidenced by SEC results.

To validate the proteins' integrity and identity, western blot analysis was performed utilizing anti-strep antibodies, confirming the preservation of the N-terminal sequence. Furthermore, Mass spectrometry (MALDI TOF) and LC-Mass/mass analysis confirmed the presence of full-length proteins. Interestingly, Par-4 CC EXS expressed in insect cells exhibited no post-translational modifications, as evidenced by the results of LC-Mass/Mass and PNGase and Endo H digestions. Subsequent biophysical characterizations assessing the proteins' oligomerization state and homogeneity in solution using Dynamic Light Scattering (DLS) showed for the mutant proteins different effects on protein oligomerization and polydispersity, implying that

polydispersity and oligomerization are not strictly correlated, though both depend on charge interactions. The results from DLS revealed that Par-4 CC EXS exhibited a significantly smaller particle size under acidic conditions compared to neutral conditions indicating that Asp and Glu protonation decreases aggregation. Additionally, the protein displayed high polydispersity under acidic and neutral to alkaline conditions implying the presence of different conformations with different degrees of folding and stability. Furthermore, NaCl salt was observed to have a negative effect on Par-4 CC EXS particle size even at low concentration, though at high concentrations it had a positive effect on protein homogeneity.

Circular Dichroism (CD) spectroscopy was utilized to evaluate the secondary structure of the proteins: All the muations in the CC domain showed a MRE222/MRE208 ratio greater than 1 and an increase in the number of residues forming helical structure implying the conservation of coiled coil structure and induction of protein folding. CD results for the mutated Par-4 CC EXS under acidic conditions revealed improved protein folding compared to the neutral state suggesting that its structural properties are highly sensitive to pH changes. No secondary structure differences were observed between proteins purified from E. coli and Hi5 cells, indicating no additional folding in Hi5 cells. Different secondary structure contents were observed among different SEC fractions under acidic condition, suggesting the instability of secondary and tertiary structure. Varying concentrations of KF, NaF, and 50 mM $MgSO_4$ changed the secondary structure towards more beta sheet and turn contents. Melting temperature (Tm) computations were conducted by observing alterations in CD spectra at different temperatures, with calculation of thermodynamic parameters. Compared to native CC domain, all the mutants showed a significantly higher Tm implying stabilization of the CC domain upon these mutations. In addition, All the mutants showed a significant increase in ΔG unfolding and ΔCp compared to the CC domain suggesting the native amino acid composition allows structural changes more easily than any of the mutants. This is possibly a functional feature of the CC domain. The Tm for Par-4 CC EXS fractions were almost identical, suggesting that the oligomerization state of the protein is not significantly influenced by changes in secondary structure. Though the presence of salt increases the Tm and hence the stability of Par-4 CC EXS, the nature of the salt did not have any significant effect.

The structural elucidation of the mutants T248D, S249D, and S249E required integrating data in space group P1, as the N-terminal helices exhibited multiple conformations within the crystals, indicating disorder in the N-termini. The R-free values were 0.3444, 0.3056, and 0.3081 for the mutants T248D, S249D, and S249E, respectively. Conversely, attempts to resolve the structure of Par-4 CC EXS crystals were thwarted due to protein degradation, precluding diffraction despite exhaustive efforts. This degradation underscores the disorder within the region spanning the CC domain and SAC domain (linker), a finding consistent with the Alphafold model of Par-4 CC EXS, which highlighted the disorder in this region. In an effort to stabilize Par-4 CC EXS, its interaction with P62 PB1 was probed using a pull-down assay. While Par-4 CC EXS was found to bind to P62 PB1, it also exhibited complete aggregation and precipitation.

This study reveals that Par-4 constructs exhibit maximum structure and minimal oligomerization under acidic conditions, indicating pH sensitivity. All single amino acid mutations introduced stabilize Par-4, suggesting its structural instability may be a biological feature which enables diverse binding modes. Despite the challenges posed by this instability, it highlights Par-4's unique biological adaptability. Given the observed variability of the protein future research should focus on single particle analysis using EM to uncover the elusive details of the Par-4 structure to enhance our understanding of its functional mechanisms.

Zusammenfassung

Par-4, ein Tumorsuppressorprotein, spielt eine entscheidende Rolle bei der Induktion von Apoptose in Krebszellen, hauptsächlich durch seine N-terminale SAC-Domäne. Strukturell gilt Par-4 als intrinsisch ungeordnetes Protein mit zwei bedeutenden Domänen: der SAC-Domäne, die für selektive Apoptose verantwortlich ist, und der C-terminalen Coiled-Coil (CC)-Domäne, die für mehrere Interaktionen essentiell ist. Während die Struktur jenseits der CC-Domäne noch unklar ist, hat die kristallographische Analyse der CC-Domäne Unordnung in dieser Region hervorgehoben. Diese Studie zielt darauf ab, sowohl die CC-Domäne als auch die erweiterte Region einschließlich der SAC-Domäne zu charakterisieren. Dazu wurden mehrere Ansätze verwendet. Es wurden Phosphorylations-mimetische, Salzbrückenentfernungs- und Ladungsreduktions-mutationen in die CC-Domäne eingeführt, sowie ein längeres Par-4-Konstrukt (Par-4 CC EXS)erzeugt, um die Konformation zu stabilisieren und zu analysieren. Expressionen in Insektenzellen und E. coli wurden für die Expression von Par-4 CC EXS genutzt. Zusätzlich wurde das Partnerprotein P62 PB1 für Bindungsstudien exprimiert und gereinigt. Die Par-4-Proteine, die am N-Terminus mit Strep-Tags markiert waren, wurden mittels Affinitätschromatographie und Größenausschlusschromatographie (SEC) gereinigt. Während der Reinigung wurde beobachtet, dass Par-4 CC EXS unter sauren Bedingungen abgebaut wurde, was die Notwendigkeit einer Mutation zur Verbesserung der Säurebeständigkeit zeigte.

Die SEC-Analyse zeigte, dass alle CC-Domänen-Mutanten ähnliche Elutionsvolumina aufwiesen, was auf einen minimalen Einfluss der Mutationen auf die Protein-Oligomerisierung hinweist. Im Gegensatz dazu offenbarte die SEC-Analyse von Par-4 CC EXS unter sauren Bedingungen verschiedene oligomere Formen. Zusätzlich verschärfte das Vorhandensein von Natriumchlorid das Aggregationsverhalten von Par-4 CC EXS, wie durch die SEC-Ergebnisse belegt wurde.

Zur Validierung der Integrität und Identität der Proteine wurde eine Western-Blot-Analyse unter Verwendung von Anti-Strep-Antikörpern durchgeführt, die die Erhaltung der N-terminalen Sequenz bestätigte. Darüber hinaus bestätigten Massenspektrometrie (MALDI TOF) und LC-Mass/Mass-Analyse die Anwesenheit vollständiger Proteine. Interessanterweise zeigte Par-4 CC EXS, exprimiert in

Insektenzellen, keine posttranslationale Modifikationen, wie durch die Ergebnisse von LC-Mass/Mass- und PNGase- sowie Endo H-Verdauungen belegt wurde. Nachfolgende biophysikalische Charakterisierungen zur Beurteilung des Oligomerisierungsstatus und der Homogenität der Proteine in Lösung mittels Dynamischer Lichtstreuung (DLS) zeigten unterschiedliche Effekte auf die Oligomerisierung und Polydispersität der mutierten Proteine, was darauf hindeutet, dass Polydispersität und Oligomerisierung nicht streng korreliert sind, obwohl beide von Ladungswechselwirkungen abhängen. Die DLS Ergebnisse zeigten, dass Par-4 CC EXS unter sauren Bedingungen eine signifikant kleinere Partikelgröße aufwies als in neutralen Bedingungen, was darauf hinweist, dass die Protonierung von Asp und Glu die Aggregation verringert. Zusätzlich zeigte das Protein unter sauren und neutral-alkalischen Bedingungen eine hohe Polydispersität, was auf das Vorhandensein unterschiedlicher Konformationen mit verschiedenen Graden von Faltung und Stabilität hinweist. Des Weiteren wurde beobachtet, dass NaCl-Salz selbst in geringen Konzentrationen einen negativen Einfluss auf die Partikelgröße von Par-4 CC EXS hat, obwohl es bei hohen Konzentrationen eine positive Wirkung auf die Proteinhomogenität zeigt. Zirkular Dichroismus (CD)-Spektroskopie wurde verwendet, um die Sekundärstruktur der Proteine zu bewerten: Alle Mutationen in der CC-Domäne zeigten ein MRE222/MRE208-Verhältnis größer als 1 und eine Zunahme der Aminosäurereste, die eine helikale Struktur bilden, was auf die Erhaltung der Coiled-Coil-Struktur und die Induktion der Proteinfaltung hinweist. CD-Ergebnisse für das mutierte Par-4 CC EXS unter sauren Bedingungen zeigten eine verbesserte Proteinfaltung im Vergleich zum neutralen Zustand, was darauf hindeutet, dass seine strukturellen Eigenschaften stark pH-sensitiv sind. Es wurden keine Unterschiede in der Sekundärstruktur zwischen Proteinen festgestellt, die aus E. coli und Hi5-Zellen gereinigt wurden. Unterschiedliche Gehalte an Sekundärstruktur wurden unter sauren Bedingungen in verschiedenen SEC-Fraktionen beobachtet, was auf die Instabilität der Sekundär- und Tertiärstruktur zeigt. Verschiedene Konzentrationen von KF, NaF und 50 mM $MgSO_4$ veränderten die Sekundärstruktur hin zu mehr Beta-Faltblatt- und Turns. Die Schmelztemperatur (Tm) wurde durch die Beobachtung von Änderungen in den CD-Spektren bei verschiedenen Temperaturen bestimmt, und mit die thermodynamischern Parameter kalkuliert: Im Vergleich zur nativen CC-Domäne zeigten alle Mutanten eine signifikant höhere Tm, was auf eine

Stabilisierung der CC-Domäne durch diese Mutationen hindeutet. Darüber hinaus wiesen alle Mutanten eine signifikante Zunahme von ΔG-Unfolding und ΔCp im Vergleich zur CC-Domäne auf, was darauf hindeutet, dass die natürliche Aminosäurezusammensetzung strukturelle Veränderungen leichter ermöglicht als bei den Mutanten. Dies ist möglicherweise ein funktionales Merkmal der CC-Domäne. Die Tm für die Par-4 CC EXS-Fraktionen waren nahezu identisch, was darauf hindeutet, dass der Oligomerisierungsstatus des Proteins nicht signifikant von Veränderungen in der Sekundärstruktur beeinflusst wird. Obwohl das Vorhandensein von Salz die Tm und damit die Stabilität von Par-4 CC EXS erhöht, hatte die Art des Salzes keinen signifikanten Effekt.

Die strukturelle Aufklärung der Mutanten T248D, S249D und S249E erforderte die Integration von Daten in der Raumgruppe P1, da die N-terminalen Helices innerhalb der Kristalle mehrere Konformationen zeigten, was auf Unordnung in den N-Termini hinweist. Die R-frei-Werte betrugen 0,3444, 0,3056 bzw. 0,3081 für die Mutanten T248D, S249D und S249E. Versuche, die Struktur der Par-4 CC EXS-Kristalle aufzulösen, wurden hingegen durch Proteinabbau vereitelt, was trotz umfangreicher Bemühungen die Beugung verhinderte. Dieser Abbau unterstreicht die Unordnung in der Region, die die CC-Domäne und die SAC-Domäne (Linker) umfasst, was mit dem Alphafold-Modell von Par-4 CC EXS übereinstimmt. Um Par-4 CC EXS zu stabilisieren, wurde seine Interaktion mit P62 PB1 mittels eines Pull-Down-Assays untersucht. Obwohl Par-4 CC EXS an P62 PB1 bindet, zeigte es auch vollständige Aggregation und Ausfällung.

Diese Studie zeigt, dass Par-4-Konstrukte unter sauren Bedingungen maximale Struktur und minimale Oligomerisierung aufweisen, was auf eine pH-Sensitivität hinweist. Alle eingeführten einzelnen Aminosäuremutationen stabilisieren Par-4, was darauf hindeutet, dass seine strukturelle Instabilität möglicherweise ein biologisches Merkmal ist, das diverse Bindungsmodi ermöglicht. Trotz der Herausforderungen durch diese Instabilität hebt dies die besondere biologische Anpassungsfähigkeit von Par-4 hervor. Angesichts der beobachteten Variabilität des Proteins sollte zukünftige Forschung auf die Einzelpartikelanalyse mittels Elektronenmikroskopie abzielen, um die verborgenen Details der Par-4-Struktur aufzudecken und unser Verständnis ihrer Funktionsmechanismen zu verbessern.

Bibliography

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., & Grosse-Kunstleve, R. W. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 213-221.

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *nature*, *422*(6928), 198-207.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallographica Section D: Biological Crystallography*, *68*(4), 352-367.

Aroyo, M. I. (2013). *International tables for crystallography*. Wiley Online Library.

Azmi, A. S., Aboukameel, A., Bao, B., Sarkar, F. H., Philip, P. A., Kauffman, M., Shacham, S., & Mohammad, R. M. (2013). Selective inhibitors of nuclear export block pancreatic cancer cell proliferation and reduce tumor growth in mice. *Gastroenterology*, *144*(2), 447-456.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 271-281.

Bej, A. K., Mahbubani, M. H., & Atlas, R. M. (1991). Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their applications. *Critical reviews in biochemistry and molecular biology*, *26*(3-4), 301-334.

Bleckmann, M., Schürig, M., Chen, F.-F., Yen, Z.-Z., Lindemann, N., Meyer, S., Spehr, J., & van den Heuvel, J. (2016). Identification of essential genetic baculoviral elements for recombinant protein expression by transactivation in Sf21 insect cells. *PloS one*, *11*(3), e0149424.

Bleckmann, M., Schürig, M., Endres, M., Samuels, A., Gebauer, D., Konisch, N., & van den Heuvel, J. (2019). Identifying parameters to improve the reproducibility of transient gene expression in High Five cells. *PloS one*, *14*(6), e0217878.

Boehrer, S., Chow, K. U., Beske, F., Kukoc-Zivojnov, N., Puccetti, E., Ruthardt, M., Baum, C., Rangnekar, V. M., Hoelzer, D., & Mitrou, P. S. (2002). In lymphatic cells par-4 sensitizes to apoptosis by down-regulating bcl-2 and promoting disruption of mitochondrial membrane potential and caspase activation. *Cancer research*, *62*(6), 1768-1775.

Boosen, M., Vetterkind, S., Koplin, A., Illenberger, S., & Preuss, U. (2005). Par-4-mediated recruitment of Amida to the actin cytoskeleton leads to the induction of apoptosis. *Experimental cell research*, *311*(2), 177-191.

Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., & Behr, J.-P. (1995). A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proceedings of the National Academy of Sciences*, *92*(16), 7297-7301. https://www.pnas.org/doi/pdf/10.1073/pnas.92.16.7297

Brasseur, K., Fabi, F., Adam, P., Parent, S., Lessard, L., & Asselin, E. (2016). Post-translational regulation of the cleaved fragment of Par-4 in ovarian and endometrial cancer cells. *Oncotarget*, *7*(24), 36971.

Brünger, A. T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallographica Section D: Biological Crystallography*, *49*(1), 24-36.

Brusotti, G., Calleri, E., Colombo, R., Massolini, G., Rinaldi, F., & Temporini, C. (2018). Advances on size exclusion chromatography and applications on the analysis of protein biopharmaceuticals and protein aggregates: a mini review. *Chromatographia*, *81*(1), 3-23.

Burikhanov, R., Zhao, Y., Goswami, A., Qiu, S., Schwarze, S. R., & Rangnekar, V. M. (2009). The tumor suppressor Par-4 activates an extrinsic pathway for apoptosis. *Cell*, *138*(2), 377-388.

Burkhard, P., Ivaninskii, S., & Lustig, A. (2002). Improving coiled-coil stability by optimizing ionic interactions. *Journal of molecular biology*, *318*(3), 901-910.

Cadena-Herrera, D., Esparza-De Lara, J. E., Ramírez-Ibañez, N. D., López-Morales, C. A., Pérez, N. O., Flores-Ortiz, L. F., & Medina-Rivero, E. (2015). Validation of three viable-cell counting methods: Manual, semi-automated, and automated. *Biotechnology Reports*, *7*, 9-16.

Cameron, V., & Uhlenbeck, O. C. (1977). 3'-Phosphatase activity in T4 polynucleotide kinase. *Biochemistry*, *16*(23), 5120-5126.

Carinhas, N., Robitaille, A. M., Moes, S., Carrondo, M. J. T., Jenoe, P., Oliveira, R., & Alves, P. M. (2011). Quantitative proteomics of Spodoptera frugiperda cells during growth and baculovirus infection. *PloS one*, *6*(10), e26444.

Chakraborty, M., Qiu, S. G., Vasudevan, K. M., & Rangnekar, V. M. (2001). Par-4 drives trafficking and activation of Fas and Fasl to induce prostate cancer cell apoptosis and tumor regression. *Cancer research*, *61*(19), 7255-7263.

Chang, S., Kim, J. H., & Shin, J. (2002). p62 forms a ternary complex with PKCζ and PAR-4 and antagonizes PAR-4-induced PKCζ inhibition. *FEBS letters*, *510*(1-2), 57-61.

Chang, T., & Zhao, G. (2021). Ice inhibition for cryopreservation: materials, strategies, and challenges. *Advanced Science*, *8*(6), 2002425.

Chaudhry, P., Singh, M., Parent, S., & Asselin, E. (2012). Prostate apoptosis response 4 (Par-4), a novel substrate of caspase-3 during apoptosis activation. *Molecular and cellular biology*, *32*(4), 826-839.

Chayen, N. E. (1998). Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallographica Section D: Biological Crystallography*, *54*(1), 8-15.

Chen, R. (2012). Bacterial expression systems for recombinant protein production: E. coli and beyond. *Biotechnology advances*, *30*(5), 1102-1107.

Cheratta, A. R., Thayyullathil, F., Pallichankandy, S., Subburayan, K., Alakkal, A., & Galadari, S. (2021). Prostate apoptosis response-4 and tumor suppression: It's not just about apoptosis anymore. *Cell Death & Disease*, *12*(1), 47.

Cianci, M., Helliwell, J., Helliwell, M., Kaucic, V., Logar, N., Mali, G., & Tusar, N. (2005). Anomalous scattering in structural chemistry and biology. *Crystallography Reviews*, *11*(4), 245-335.

Cook, J., Krishnan, S., Ananth, S., Sells, S. F., Shi, Y., Walther, M. M., Linehan, W. M., Sukhatme, V. P., Weinstein, M. H., & Rangnekar, V. M. (1999). Decreased expression of the pro-apoptotic protein Par-4 in renal cell carcinoma. *Oncogene*, *18*(5), 1205-1208.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography*, *62*(9), 1002-1011.

de Thonel, A., Hazoume, A., Kochin, V., Isoniemi, K., Jego, G., Fourmaux, E., Hammann, A., Mjahed, H., Filhol, O., & Micheau, O. (2014). Regulation of the proapoptotic functions of prostate apoptosis response-4 (Par-4) by casein kinase 2 in prostate cancer cells. *Cell Death & Disease*, *5*(1), e1016-e1016.

Diaz-Meco, M. a. T., Lallena, M. a.-J., Monjas, A., Frutos, S., & Moscat, J. (1999). Inactivation of the inhibitory κB protein kinase/nuclear factor κB pathway by Par-4 expression potentiates tumor necrosis factor α-induced apoptosis. *Journal of Biological Chemistry*, *274*(28), 19606-19612.

Díaz-Meco, M. T., Municio, M. M., Frutos, S., Sanchez, P., Lozano, J., Sanz, L., & Moscat, J. (1996). The product of par-4, a gene induced during apoptosis, interacts selectively with the atypical isoforms of protein kinase C. *Cell*, *86*(5), 777-786.

Dodero, V. I., Quirolo, Z. B., & Sequeira, M. A. (2011). Biomolecular studies by circular dichroism.

Drenth, J. (2007). *Principles of protein X-ray crystallography*. Springer Science & Business Media.

Duan, W., Guo, Z., & Mattson, M. P. (2000). Participation of par-4 in the degeneration of striatal neurons induced by metabolic compromise with 3-nitropropionic acid. *Experimental neurology*, *165*(1), 1-11.

Dutta, K., Alexandrov, A., Huang, H., & Pascal, S. M. (2001). pH-induced folding of an apoptotic coiled coil. *Protein Science*, *10*(12), 2531-2540.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, *6*(7), 1948-1954.

El-Guendy, N. (2003). Identification of a Unique Core Domain of Par-4 Sufficient for Selective Apoptosis Induction in Cancer Cells. *Molecular and cellular biology*, *23:16*.

El-Guendy, N., & Rangnekar, V. M. (2003). Apoptosis by Par-4 in cancer and neurodegenerative diseases. *Experimental cell research*, *283*(1), 51-66.

El-Guendy, N., Zhao, Y., Gurumurthy, S., Burikhanov, R., & Rangnekar, V. M. (2003). Identification of a unique core domain of par-4 sufficient for selective apoptosis induction in cancer cells. *Mol Cell Biol*, *23*(16), 5516-5525. https://doi.org/10.1128/mcb.23.16.5516-5525.2003

Elmore, S. (2007). Apoptosis: a review of programmed cell death. *Toxicologic pathology*, *35*(4), 495-516.

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, *60*(12), 2126-2132.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 486-501.

Evans, P., & McCoy, A. (2008). An introduction to molecular replacement. *Acta Crystallographica Section D: Biological Crystallography*, *64*(1), 1-10.

Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 282-292.

Fasman, G. D. (2013). *Circular dichroism and the conformational analysis of biomolecules*. Springer Science & Business Media.

Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in E. coli. *Current protocols in protein science*, *61*(1), 5.24. 21-25.24. 29.

Freeze, H. H., & Kranz, C. (2008). Endoglycosidase and glycoamidase release of N-linked glycans. *Current protocols in immunology*, *83*(1), 8.15. 11-18.15. 26.

Galluzzi, L., Pedro, B.-S., Vitale, I., Aaronson, S., Abrams, J., Adam, D., Alnemri, E., Altucci, L., Andrews, D., & Annicchiarico-Petruzzelli, M. (2015). Essential versus accessory aspects of cell death: recommendations of the NCCD 2015. *Cell Death & Differentiation*, *22*(1), 58-73.

Ganguly, S., Burikhanov, R., Qiu, S., & Rangnekar, V. M. (2022). Discovery and overview of Par-4. *Tumor Suppressor Par-4: Structural Features, Molecular Mechanisms and Function*, 1-59.

Garman, E. F. (2010). Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 339-351.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server*. Springer.

Gill, S. C., & Von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry*, *182*(2), 319-326.

Goswami, A., Burikhanov, R., de Thonel, A., Fujita, N., Goswami, M., Zhao, Y., Eriksson, J. E., Tsuruo, T., & Rangnekar, V. M. (2005). Binding and phosphorylation of par-4 by akt is essential for cancer cell survival. *Molecular cell*, *20*(1), 33-44.

Goswami, A., Qiu, S., Dexheimer, T. S., Ranganathan, P., Burikhanov, R., Pommier, Y., & Rangnekar, V. M. (2008). Par-4 binds to topoisomerase 1 and attenuates its DNA relaxation activity. *Cancer research*, *68*(15), 6190-6198.

Green, D. R., & Llambi, F. (2015). Cell death signaling. *Cold Spring Harbor perspectives in biology*, *7*(12), a006080.

Grünewald, S., Haase, W., Reiländer, H., & Michel, H. (1996). Glycosylation, palmitoylation, and localization of the human D2S receptor in baculovirus-infected insect cells. *Biochemistry*, *35*(48), 15149-15161.

Guex, N., Peitsch, M. C., & Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, *30*(S1), S162-S173.

Guo, C., Yu, S., Davis, A. T., Wang, H., Green, J. E., & Ahmed, K. (2001). A potential role of nuclear matrix-associated protein kinase CK2 in protection against drug-induced apoptosis in cancer cells. *Journal of Biological Chemistry*, *276*(8), 5992-5999.

Guo, H., Treude, F., Krämer, O. H., Lüscher, B., & Hartkamp, J. (2019). PAR-4 overcomes chemo-resistance in breast cancer cells by antagonizing cIAP1. *Scientific Reports*, *9*(1), 8755.

Guo, Q., Fu, W., Xie, J., Luo, H., Sells, S. F., Geddes, J., Bondada, V., Rangnekar, V. M., & Mattson, M. P. (1998). Par-4 is a mediator of neuronal degeneration associated with the pathogenesis of Alzheimer disease. *Nature medicine*, *4*(8), 957-962.

Guo, Q., & Xie, J. (2004). AATF inhibits aberrant production of amyloid β peptide 1-42 by interacting directly with Par-4. *Journal of Biological Chemistry*, *279*(6), 4596-4603.

Gurumurthy, S., Goswami, A., Vasudevan, K. M., & Rangnekar, V. M. (2005). Phosphorylation of Par-4 by protein kinase A is critical for apoptosis. *Molecular and cellular biology*, *25*(3), 1146-1161.

Haas, C., & Drenth, J. (1999). Understanding protein crystallization on the basis of the phase diagram. *Journal of crystal growth*, *196*(2-4), 388-394.

Hebbar, N., Burikhanov, R., Shukla, N., Qiu, S., Zhao, Y., Elenitoba-Johnson, K. S., & Rangnekar, V. M. (2017). A naturally generated decoy of the prostate apoptosis response-4 protein overcomes therapy resistance in tumors. *Cancer research*, *77*(15), 4039-4050.

Hebbar, N., Wang, C., & Rangnekar, V. M. (2012). Mechanisms of apoptosis by the tumor suppressor Par-4. *Journal of cellular physiology*, *227*(12), 3715-3721.

Hendrickson, W. A., Smith, J. L., & Sheriff, S. (1985). Direct phase determination based on anomalous scattering. In *Methods in enzymology* (Vol. 115, pp. 41-55). Elsevier.

Héricourt, F., Blanc, S., Redeker, V., & Jupin, I. (2000). Evidence for phosphorylation and ubiquitinylation of the turnip yellow mosaic virus RNA-dependent RNA polymerase domain expressed in a baculovirus–insect cell system. *Biochemical Journal*, *349*(2), 417-425.

Hickman, A. B., & Davies, D. R. (1997). Principles of Macromolecular X-Ray Crystallography. *Current protocols in protein science*, *10*(1), 17.13. 11-17.13. 15.

Hillenkamp, F., & Karas, M. (1990). [12] Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. In *Methods in enzymology* (Vol. 193, pp. 280-295). Elsevier.

Hongmei, Z. (2012). Extrinsic and intrinsic apoptosis signal pathway review. In *Apoptosis and medicine*. InTechOpen.

Jenkins, J., Smith, D., & Gaisford, W. (2013). mosquito® Crystal and mosquito® LCP: fast, reliable automation of protein crystallisation drop set-up. *Acta Crystallogr A Found Crystallogr*, *69*, s656.

Jin, Z., & El-Deiry, W. S. (2005). Overview of cell death signaling pathways. *Cancer biology & therapy*, *4*(2), 147-171.

John, H., Walden, M., Schäfer, S., Genz, S., & Forssmann, W.-G. (2004). Analytical procedures for quantification of peptides in pharmaceutical research by liquid chromatography–mass spectrometry. *Analytical and bioanalytical chemistry*, *378*, 883-897.

Johnson, L. (1985). Protein crystallography. In *New Comprehensive Biochemistry* (Vol. 11, pp. 347-415). Elsevier.

Johnstone, R. W., Tommerup, N., Hansen, C., Vissing, H., & Shi, Y. (1998). Mapping of the human PAWR (par-4) gene to chromosome 12q21. *Genomics*, *53*(2), 241-243.

Julien, O., & Wells, J. A. (2017). Caspases and their substrates. *Cell Death & Differentiation*, *24*(8), 1380-1389.

Kammerer, R. A., Jaravine, V. A., Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Garcıa-Echeverrıa, C., Alexandrescu, A. T., Engel, J., & Steinmetz, M. O. (2001). An intrahelical salt bridge within the trigger site stabilizes the GCN4 leucine zipper. *Journal of Biological Chemistry*, *276*(17), 13685-13688.

Karas, M., Glückmann, M., & Schäfer, J. (2000). Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *Journal of Mass Spectrometry*, *35*(1), 1-12.

Karas, M., & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry*, *60*(20), 2299-2301.

Kaur, J., Kumar, A., & Kaur, J. (2018). Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. *International Journal of Biological Macromolecules*, *106*, 803-822.

Ke, H. (1997). [25] Overview of isomorphous replacement phasing. In *Methods in enzymology* (Vol. 276, pp. 448-461). Elsevier.

Kelly, S. M., & Price, N. C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current protein and peptide science*, *1*(4), 349-384.

Kim, M.-J., Yoo, S.-J., & Kim, D.-Y. (2006). A supramolecular chiroptical switch using an amorphous azobenzene polymer. *Advanced Functional Materials*, *16*(16), 2089-2094.

Kleywegt, G. J., & Jones, T. A. (1997). [11] Model building and refinement practice. In *Methods in enzymology* (Vol. 277, pp. 208-230). Elsevier.

Kost, T. A., Condreay, J. P., & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature biotechnology*, *23*(5), 567-575.

Krauss, I. R., Merlino, A., Vergara, A., & Sica, F. (2013). An overview of biological macromolecule crystallization. *International journal of molecular sciences*, *14*(6), 11643-11691.

Kricka, L. J., & Ji, X. (1995). 4-phenylylboronic acid: A new type of enhancer for the horseradish peroxidase catalysed chemiluminescent oxidation of luminol. *Journal of Bioluminescence and Chemiluminescence*, *10*(1), 49-54.

Kruman, I. I., Nath, A., Maragos, W. F., Chan, S. L., Jones, M., Rangnekar, V. M., Jakel, R. J., & Mattson, M. P. (1999). Evidence that Par-4 participates in the pathogenesis of HIV encephalitis. *The American journal of pathology*, *155*(1), 39-46.

Kumar, S., Tsai, C.-J., & Nussinov, R. (2000). Factors enhancing protein thermostability. *Protein engineering*, *13*(3), 179-191.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of the National Academy of Sciences*, *82*(2), 488-492.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature*, *227*(5259), 680-685.

Lamed, R., Kenig, R., Morag, E., Yaron, S., Shoham, Y., & Bayer, E. A. (2001). Nonproteolytic cleavage of aspartyl proline bonds in the cellulosomal scaffoldin subunit from Clostridium thermocellum. *Applied biochemistry and biotechnology*, *90*, 67-73.

Leipertz, A., & Fröba, A. P. (2005). Diffusion measurements in fluids by dynamic light scattering. In *Diffusion in Condensed Matter: Methods, Materials, Models* (pp. 579-618). Springer.

Li, F., Liu, S.-l., & Mullins, J. I. (1999). Site-directed mutagenesis using uracil-containing double-stranded DNA templates and DpnI digestion. *Biotechniques*, *27*(4), 734-738.

Libich, D. S., Schwalbe, M., Kate, S., Venugopal, H., Claridge, J. K., Edwards, P. J., Dutta, K., & Pascal, S. M. (2009). Intrinsic disorder and coiled-coil formation in prostate apoptosis response factor 4. *the FEBS Journal*, *276*(14), 3710-3728.

Louche, A., Salcedo, S. P., & Bigot, S. (2017). Protein–protein interactions: pull-down assays. *bacterial Protein secretion systems: methods and protocols*, 247-255.

Mabe, N. W., Fox, D. B., Lupo, R., Decker, A. E., Phelps, S. N., Thompson, J. W., & Alvarez, J. V. (2018). Epigenetic silencing of tumor suppressor Par-4 promotes chemoresistance in recurrent breast cancer. *The Journal of clinical investigation*, *128*(10), 4413-4428.

Malkin, A., Kuznetsov, Y. G., Glantz, W., & McPherson, A. (1996). Atomic force microscopy studies of surface morphology and growth kinetics in thaumatin crystallization. *The Journal of Physical Chemistry*, *100*(28), 11736-11743.

Malkin, A., Kuznetsov, Y. G., Land, T., DeYoreo, J., & McPherson, A. (1995). Mechanisms of growth for protein and virus crystals. *Nature structural biology*, *2*(11), 956-959.

Matthews, B. W. (1968). Solvent content of protein crystals. *Journal of molecular biology*, *33*(2), 491-497.

McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D: Biological Crystallography*, *63*(1), 32-41.

McPherson, A., & Gavira, J. A. (2014). Introduction to protein crystallization. *Acta Crystallographica Section F: Structural Biology Communications*, *70*(1), 2-20.

McPherson, A., Malkin, A., & Kuznetsov, Y. G. (2000). Atomic force microscopy in the study of macromolecular crystal growth. *Annual review of biophysics and biomolecular structure*, *29*(1), 361-410.

Meier, M., Lustig, A., Aebi, U., & Burkhard, P. (2002). Removing an interhelical salt bridge abolishes coiled-coil formation in a de novo designed peptide. *Journal of structural biology*, *137*(1-2), 65-72.

Mertens, H. D., & Svergun, D. I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of structural biology*, *172*(1), 128-141.

Micsonai, A., Wien, F., Bulyáki, É., Kun, J., Moussong, É., Lee, Y.-H., Goto, Y., Réfrégiers, M., & Kardos, J. (2018). BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic acids research*, *46*(W1), W315-W322.

Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., & Kardos, J. (2015). Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences*, *112*(24), E3095-E3103.

Moreno-Bueno, G., Fernandez-Marcos, P. J., Collado, M., Tendero, M. J., Rodriguez-Pinilla, S. M., Garcia-Cao, I., Hardisson, D., Diaz-Meco, M. T., Moscat, J., & Serrano, M. (2007). Inactivation of the candidate tumor suppressor par-4 in endometrial cancer. *Cancer research*, *67*(5), 1927-1934.

Moritz, C. P. (2017). Tubulin or not tubulin: heading toward total protein staining as loading control in western blots. *Proteomics*, *17*(20), 1600189.

Nagai, M. A., Gerhard, R., Salaorni, S., Fregnani, J. H. T. G., Nonogaki, S., Netto, M. M., & Soares, F. A. (2010). Down-regulation of the candidate tumor suppressor gene PAR-4 is associated with poor prognosis in breast cancer. *International journal of oncology*, *37*(1), 41-49.

Norris, G. E., Stillman, T. J., Anderson, B. F., & Baker, E. N. (1994). The three-dimensional structure of PNGase F, a glycosyl asparaginase from Flavobacterium meningosepticum. *Structure*, *2*(11), 1049-1059.

Nouri, H. R., Karkhah, A., Varasteh, A., & Sankian, M. (2016). Expression of a chimeric allergen with high rare codons content in codon bias-adjusted Escherichia coli: Escherichia coli BL21 (DE3)-codon plus RIL as an efficient host. *Current microbiology*, *73*, 91-98.

Old, R. W., & Primrose, S. B. (1981). *Principles of gene manipulation: an introduction to genetic engineering* (Vol. 2). Univ of California Press.

Ouyang, L., Shi, Z., Zhao, S., Wang, F. T., Zhou, T. T., Liu, B., & Bao, J. K. (2012). Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. *Cell proliferation*, *45*(6), 487-498.

Page, G., Kögel, D., Rangnekar, V., & Scheidtmann, K. H. (1999). Interaction partners of Dlk/ZIP kinase: co-expression of Dlk/ZIP kinase and Par-4 results in cytoplasmic retention and apoptosis. *Oncogene*, *18*(51), 7265-7273.

Pauling, L., Corey, R. B., & Branson, H. R. (1951). The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences*, *37*(4), 205-211.

Pedersen, W. A., Luo, H., Kruman, I., Kasarskis, E., & Mattson, M. P. (2000). The prostate apoptosis response-4 protein participates in motor neuron degeneration in amyotrophic lateral sclerosis. *The FASEB journal*, *14*(7), 913-924.

Piazza, R. (1999). Interactions in protein solutions near crystallisation: a colloid physics approach. *Journal of crystal growth*, *196*(2-4), 415-423.

Piszkiewicz, D., Landon, M., & Smith, E. L. (1970). Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations. *Biochemical and biophysical research communications*, *40*(5), 1173-1178.

Plummer Jr, T. H., Elder, J. H., Alexander, S., Phelan, A., & Tarentino, A. L. (1984). Demonstration of peptide: N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. *Journal of Biological Chemistry*, *259*(17), 10700-10704.

Rahman, A., Pallichankandy, S., Thayyullathil, F., & Galadari, S. (2019). Critical role of H2O2 in mediating sanguinarine-induced apoptosis in prostate cancer cells via facilitating ceramide generation, ERK1/2 phosphorylation, and Par-4 cleavage. *Free Radical Biology and Medicine*, *134*, 527-544.

Ranganathan, P., & Rangnekar, V. M. (2005). Regulation of cancer cell survival by Par-4. *Annals of the New York Academy of Sciences*, *1059*(1), 76-85.

Rodríguez-Aparicio, L., Reglero, A., Ortiz, A., & Luengo, J. (1988). Effect of physical and chemical conditions on the production of colominic acid by Escherichia coli in a defined medium. *Applied microbiology and biotechnology*, *27*, 474-483.

Rong, Y. S., Titen, S. W., Xie, H. B., Golic, M. M., Bastiani, M., Bandyopadhyay, P., Olivera, B. M., Brodsky, M., Rubin, G. M., & Golic, K. G. (2002). Targeted mutagenesis by homologous recombination in D. melanogaster. *Genes & development*, *16*(12), 1568-1581.

Roussigne, M., Cayrol, C., Clouaire, T., Amalric, F., & Girard, J.-P. (2003). THAP1 is a nuclear proapoptotic factor that links prostate-apoptosis-response-4 (Par-4) to PML nuclear bodies. *Oncogene*, *22*(16), 2432-2442.

Saiki, R. K. (1989). The design and optimization of the PCR. *PCR technology*, 7-16.

Saikumar, P., Dong, Z., Mikhailov, V., Denton, M., Weinberg, J. M., & Venkatachalam, M. A. (1999). Apoptosis: definition, mechanisms, and relevance to disease. *The American journal of medicine*, *107*(5), 489-506.

Santoro, M. M., & Bolen, D. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl. alpha.-chymotrypsin using different denaturants. *Biochemistry*, *27*(21), 8063-8068.

Sato, M., Yao, V. J., Arap, W., & Pasqualini, R. (2010). GRP78 signaling hub: A receptor for targeted tumor therapy. *Advances in genetics*, *69*, 97-114.

Schmidt, T. G., & Skerra, A. (1994). One-step affinity purification of bacterially produced proteins by means of the “Strep tag” and immobilized recombinant core streptavidin. *Journal of chromatography A*, *676*(2), 337-345.

Schultz, D. R., & Harringto Jr, W. J. (2003). Apoptosis: programmed cell death at a molecular level. Seminars in arthritis and rheumatism,

Schwalbe, M., Dutta, K., Libich, D. S., Venugopal, H., Claridge, J. K., Gell, D. A., Mackay, J. P., Edwards, P. J., & Pascal, S. M. (2010). Two-state conformational equilibrium in the Par-4 leucine zipper domain. *Proteins: Structure, Function, and Bioinformatics*, *78*(11), 2433-2449.

Sells, S. F., Han, S.-S., Muthukkumar, S., Maddiwar, N., Johnstone, R., Boghaert, E., Gillis, D., Liu, G., Nair, P., & Monnig, S. (1997). Expression and function of the leucine zipper protein Par-4 in apoptosis. *Molecular and cellular biology*.

Sells, S. F., Wood, D., Joshi-Barve, S. S., Muthukumar, S., Jacob, R., Crist, S., Humphreys, S., & Rangnekar, V. (1994). Commonality of the gene programs induced by effectors of apoptosis in androgen-dependent and-independent prostate cells. *Cell Growth and Differentiation*, *5*, 457-457.

Shi, Y. (2014). A glimpse of structural biology through X-ray crystallography. *Cell*, *159*(5), 995-1014.

Shimon, L. J., Bayer, E. A., Morag, E., Lamed, R., Yaron, S., Shoham, Y., & Frolow, F. (1997). A cohesin domain from Clostridium thermocellum: the crystal structure provides new insights into cellulosome assembly. *Structure*, *5*(3), 381-390.

Shrestha-Bhattarai, T., & Rangnekar, V. M. (2010). Cancer-selective apoptotic effects of extracellular and intracellular Par-4. *Oncogene*, *29*(27), 3873-3880.

Silberring, J., Kowalczuk, M., Bergquist, J., Kraj, A., Suder, P., Dylag, T., Smoluch, M., Chervet, J.-P., & Ekman, R. (2004). Size-exclusion chromatography. In *Journal of Chromatography Library* (Vol. 69, pp. 213-251). Elsevier.

Stacey, G., & Possee, R. (1996). Safety aspects of insect cell culture. *Cytotechnology*, *20*, 299-304.

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology*, *189*(1), 113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990). [6] Use of T7 RNA polymerase to direct expression of cloned genes. In *Methods in enzymology* (Vol. 185, pp. 60-89). Elsevier.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., & Matsuo, T. (1988). Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid communications in mass spectrometry*, *2*(8), 151-153.

Tarafder, A. K., Guesdon, A., Kuhm, T., & Sachse, C. (2019). Recombinant expression, purification, and assembly of p62 filaments. *Autophagy: Methods and Protocols*, 3-15.

Taylor, G. L. (2010). Introduction to phasing. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 325-338.

Thayyullathil, F., Pallichankandy, S., Rahman, A., Kizhakkayil, J., Chathoth, S., Patel, M., & Galadari, S. (2013). Caspase-3 mediated release of SAC domain containing fragment from Par-4 is necessary for the sphingosine-induced apoptosis in Jurkat cells. *Journal of molecular signaling*, *8*(1), 1-13.

Thompson, R. E., Liu, X., Ripoll-Rozada, J., Alonso-García, N., Parker, B. L., Pereira, P. J. B., & Payne, R. J. (2017). Tyrosine sulfation modulates activity of tick-derived thrombin inhibitors. *Nature chemistry*, *9*(9), 909-917.

Tiruttani Subhramanyam, U. K., Kubicek, J., Eidhoff, U. B., & Labahn, J. (2014). Cloning, expression, purification, crystallization and preliminary crystallographic analysis of the C-terminal domain of Par-4 (PAWR). *Acta Crystallographica Section F: Structural Biology Communications*, *70*(9), 1224-1227.

Tiruttani Subhramanyam, U. K., Kubicek, J., Eidhoff, U. B., & Labahn, J. (2017). Structural basis for the regulatory interactions of proapoptotic Par-4. *Cell Death & Differentiation*, *24*(9), 1540-1547.

Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., & Ellenberger, T. (2006). DNA ligases: structure, reaction mechanism, and function. *Chemical reviews*, *106*(2), 687-699.

Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, *76*(9), 4350-4354.

Treude, F., Kappes, F., Fahrenkamp, D., Müller-Newen, G., Dajas-Bailador, F., Krämer, O. H., Lüscher, B., & Hartkamp, J. (2014). Caspase-8-mediated PAR-4 cleavage is required for TNFα-induced apoptosis. *Oncotarget*, *5*(10), 2988.

Vagin, A., & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *Journal of applied crystallography*, *30*(6), 1022-1025.

Varland, S., Osberg, C., & Arnesen, T. (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics*, *15*(14), 2385-2401.

Vetterkind, S., Illenberger, S., Kubicek, J., Boosen, M., Appel, S., Naim, H. Y., Scheidtmann, K.-H., & Preuss, U. (2005). Binding of Par-4 to the actin cytoskeleton is essential for Par-4/Dlk-mediated apoptosis. *Experimental cell research*, *305*(2), 392-408.

Warburg, O. (1942). Isolation and crystallization of enolase. *Biochemische Zeitschrift*, *310*, 384-421.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., & McCoy, A. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 235-242.

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, *367*(6483), 1260-1263.

Yang, D., Zhang, W., Zhang, H., Zhang, F., Chen, L., Ma, L., Larcher, L. M., Chen, S., Liu, N., & Zhao, Q. (2020). Progress, opportunity, and perspective on exosome isolation-efforts for efficient exosome-based theranostics. *Theranostics*, *10*(8), 3684.

Yang, Y. (2015). *Side reactions in peptide synthesis*. Academic Press.

Zhang, Y., Taiming, L., & Liu, J. (2003). Low temperature and glucose enhanced T7 RNA polymerase-based plasmid stability for increasing expression of glucagon-like peptide-2 in Escherichia coli. *Protein expression and purification*, *29*(1), 132-139.

Zhao, Y., Burikhanov, R., Qiu, S., Lele, S. M., Jennings, C. D., Bondada, S., Spear, B., & Rangnekar, V. M. (2007). Cancer resistance in transgenic mice expressing the SAC module of Par-4. *Cancer research*, *67*(19), 9276-9285.

Zhao, Y., & Rangnekar, V. M. (2008). Apoptosis and tumor resistance conferred by Par-4. *Cancer biology & therapy*, *7*(12), 1867-1874.

Zheng, L., Baumann, U., & Reymond, J.-L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic acids research*, *32*(14), e115-e115.

Appendices

Appendix I: DNA and protein sequences

Par-4 (without tag) DNA sequence

```
  1 ATGGCGACCG GCGGCTATCG GAGCAGCGGC AGCACCACGG ACTTCCTGGA
 51 GGAGTGGAAA GCGAAGCGCG AGAAGATGCG CGCCAAGCAG AACCCCGTGG
101 GCCCGGGTTC GAGCGGCGGG GATCCAGCCG CCAAGTCCCC TGCGGGACCG
151 CTCGCCCAGA CTACGGCCGC GGGGACCTCG GAACTCAACC ACGGCCCCGC
201 CGGCGCGGCC GCACCTGCCG CCCCCGGGCC GGGCGCCCTG AACTGCGCTC
251 ACGGCTCGTC CGCGCTGCCC CGCGGGGCTC CCGGCTCCCG GCGGCCGGAG
301 GACGAGTGTC CTATTGCCGC TGGGGCCGCG GGAGCACCCG CGTCCCGGGG
351 AGACGAGGAG GAGCCGGATA GCGCCCCGGA GAAGGGCCGC AGCTCGGGGC
401 CCAGCGCCAG GAAAGGCAAA GGGCAGATCG AGAAGAGGAA GCTGCGGGAG
451 AAGCGCCGCT CCACCGGCGT GGTCAACATC CCCGCGGCGG AGTGCTTAGA
501 TGAGTACGAA GATGACGAAG CAGGACAGAA GGAACGGAAG CGAGAGGATG
551 CTATCACACA GCAGAACACC ATCCAGAATG AAGCTGCGAG CCTCCCAGAT
601 CCAGGAACCT CCTACCTGCC CCAGGACCCG TCGAGAACAG TCCCAGGCAG
651 ATACAAAAGC ACAATCAGTG CCCCAGAAGA AGAAATCTTA AATAGATATC
701 CCCGAACAGA TAGAAGTGGC TTCAGTAGAC ACAACAGAGA TACCAGTGCG
751 CCTGCTAACT TCGCTTCAAG TAGCACCTTG GAAAAGAGAA TTGAAGATCT
801 TGAGAAGGAA GTCTTGAGAG AAAGGCAAGA AAACCTTCGA CTTACGAGGC
851 TGATGCAAGA TAAAGAAGAA ATGATTGGAA AACTCAAGGA AGAGATTGAT
901 TTGTTAAATA GAGACCTCGA TGACATGGAA GACGAAAACG AGCAACTAAA
951 GCAGGAAAAT AAAACTCTTT TGAAAGTTGT TGGGCAGCTG ACAAGGTAG
```

Par-4 (without tag) protein sequence

```
  1 MATGGYRSSG STTDFLEEWK AKREKMRAKQ NPVGPGSSGG DPAAKSPAGP
 51 LAQTTAAGTS ELNHGPAGAA APAAPGPGAL NCAHGSSALP RGAPGSRRPE
101 DECPIAAGAA GAPASRGDEE EPDSAPEKGR SSGPSARKGK GQIEKRKLRE
151 KRRSTGVVNI PAAECLDEYE DDEAGQKERK REDAITQQNT IQNEAASLPD
201 PGTSYLPQDP SRTVPGRYKS TISAPEEEIL NRYPRTDRSG FSRHNRDTSA
251 PANFASSSTL EKRIEDLEKE VLRERQENLR LTRLMQDKEE MIGKLKEEID
301 LLNRDLDDME DENEQLKQEN KTLLKVVGQL TR-
```

Green - Coiled coil domain; Red - SAC domain; Orange-Linker

Par-4 CC EXS (with strep tag) DNA sequence

```
  1 ATGGCTAGCT GGAGCCACCC GCAGTTCGAG AAGGCGAGGA AAGGCAAAGG
 51 GCAGATCGAG AAGAGGAAGC TGCGGGAGAA GCGCCGCTCC ACCGGCGTGG
101 TCAACATCCC CGCGGCGGAG TGCTTAGATG AGTACGAAGA TGACGAAGCA
151 GGACAGAAGG AACGGAAGCG AGAGGATGCT ATCACACAGC AGAACACCAT
201 CCAGAATGAA GCTGCGAGCC TCCCAGATCC AGGAACCTCC TACCTGCCCC
251 AGGACCCGTC GAGAACAGTC CCAGGCAGAT ACAAAAGCAC AATCAGTGCC
301 CCAGAAGAAG AAATCTTAAA TAGATATCCC CGAACAGATA GAAGTGGCTT
351 CAGTAGACAC AACAGAGATA CCAGTGCGCC TGCTAACTTC GCTTCAAGTA
401 GCACCTTGGA AAAGAGAATT GAAGATCTTG AGAAGGAAGT CTTGAGAGAA
451 AGGCAAGAAA ACCTTCGACT TACGAGGCTG ATGCAAGATA AAGAAGAAAT
501 GATTGGAAAA CTCAAGGAAG AGATTGATTT GTTAAATAGA GACCTCGATG
551 ACATGGAAGA CGAAAACGAG CAACTAAAGC AGGAAAATAA AACTCTTTTG
601 AAAGTTGTTG GGCAGCTGAC AAGGTAG
```

Par-4 CC EXS (with strep tag) protein sequence

```
  1 MASWSHPQFE KARKGKGQIE KRKLREKRRS TGVVNIPAAE CLDEYEDDEA
 51 GQKERKREDA ITQQNTIQNE AASLPDPGTS YLPQDPSRTV PGRYKSTISA
101 PEEEILNRYP RTDRSGFSRH NRDTSAPANF ASSSTLEKRI EDLEKEVLRE
151 RQENLRLTRL MQDKEEMIGK LKEEIDLLNR DLDDMEDENE QLKQENKTLL
201 KVVGQLTR-
```

Blue- Strep tag RED - SAC domain , Green - Coiled coil domain, Orange -Linker

CC (with strep tag) DNA sequence

```
  1 ATGGCTAGCT GGAGCCACCC GCAGTTCGAG AAGGCGGGCT TCAGTAGACA
 51 CAACAGAGAT ACCAGTGCGC CTGCTAACTT CGCTTCAAGT AGCACCTTGG
101 AAAAGAGAAT TGAAGATCTT GAGAAGGAAG TCTTGAGAGA AAGGCAAGAA
151 AACCTTCGAC TTACGAGGCT GATGCAAGAT AAAGAAGAAA TGATTGGAAA
201 ACTCAAGGAA GAGATTGATT TGTTAAATAG AGACCTCGAT GACATGGAAG
251 ACGAAAACGA GCAACTAAAG CAGGAAAATA AAACTCTTTT GAAAGTTGTT
301 GGGCAGCTGA CAAGGTAG
```

CC (with strep tag) protein sequence

```
  1 MASWSHPQFE KAGFSRHNRD TSAPANFASS STLEKRIEDL EKEVLRERQE
 51 NLRLTRLMQD KEEMIGKLKE EIDLLNRDLD DMEDENEQLK QENKTLLKVV
101 GQLTR-
```

Appendix II: Primers used

Primers used for cloning (5' –> 3')

(The number in the primer name refers to the residue position in Par-4 protein sequence)

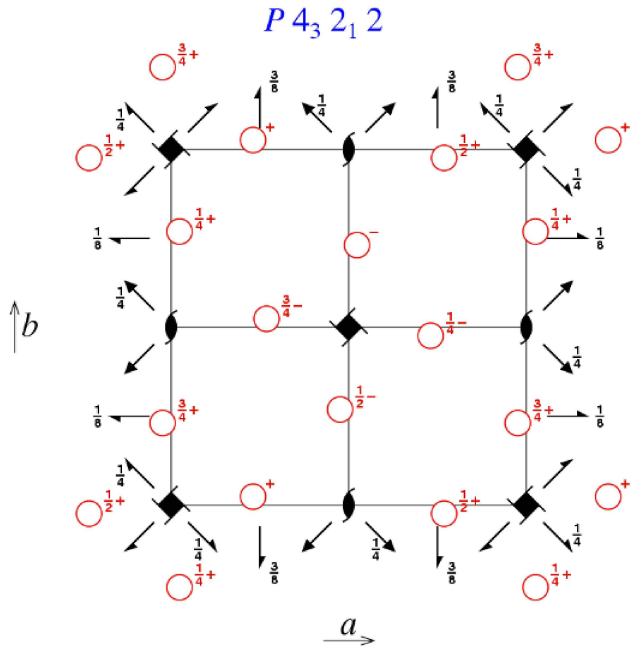
Primer name	sequence	No of base pair	CG %	Tm °C
Par-4 CC EXS FP *	GGCATATGGCTAGCTGGAGCCACCCGCAGTTCGAGAAGGCGAGGAAAGGCAAAGGGCAGATCGAGAAGAGGAAGCTGCGG	80	60	74.5
Par-4 CC EXS RP *	CCTGGATCCCTACCTTGTCAGCTGCCCAAC	30	60	67.3
Par-4 CC EXS FP **	GCAGGATCCATGTGGAGCCACCCGCAGTTCGAAAAAGGGGCCAGGAAAGGCAAAGGGCAG	60	60	74.7
Par-4 CC EXS RP**	CGGCCTAGGCTACCTTGTCAGCTGCCCAAC	30	63.3	69
Par-4 CC EXS D209E FP	CCTGCCCCAGGAGCCGTCGAGAA	23	69.6	68.1
Par-4 CC EXS D209E FP	TAGGAGGTTCCTGGATCTGGGAGG	24	58.3	62.1
Par-4 CC EXS D200E FP	GAGCCTCCCAGAGCCAGGAACCT	23	65.2	65.8
Par-4 CC EXS D200E RP	GCAGCTTCATTCTGGATGGTGTTC	24	50	59.2
CC S249E FP	CAGAGATACCGAGGCGCCTGCTAAC	25	60	62.8
CC S249E RP	TTGTGTCTACTGAAGCCC	18	50	51.6
CC S249D FP	CAGAGATACCGACGCGCCTGCTAAC	25	60	63
CC S249D RP	TTGTGTCTACTGAAGCCC	18	50	51.6
CC T248E FP	CAACAGAGATGAGAGTGCGCCTGCTAAC	28	53.6	62.5
CC T248E RP	TGTCTACTGAAGCCCGCC	18	61.1	57.2
CC T248D FP	CAACAGAGATGACAGTGCGCCTGC	24	58.3	62.4
CC T248D RP	TGTCTACTGAAGCCCGCC	18	61.1	57.2
CC E268Q FP	TGAAGATCTTCAGAAGGAAGTCTTG	25	40	54.4
CC E268Q RP	ATTCTCTTTTCCAAGGTG	18	38.9	46.8

CC E274Q FP	AGTCTTGAGACAAAGGCAAGAAAAC	25	40	55.8
CC E274Q RP	TCCTTCTCAAGATCTTCAATTC	22	36.4	50
CC K288M FP	GATGCAAGATATGGAAGAAATGATTGG AAAACTC	34	35.3	58.1
CC K288M RP	AGCCTCGTAAGTCGAAGG	18	55.6	53.9
CC K317M FP	CGAGCAACTAATGCAGGAAAATAAAAC	27	37	55.3
CC K317M RP	TTTTCGTCTTCCATGTCATC	20	40	50.4
CC D305N FP	GTTAAATAGAAACCTCGATGACATG	25	36	52
CC D305N RP	AAATCAATCTCTTCCTTGAGTTTTC	25	32	51.8

* Used in E. coli expression system

** Used in Hi5 insect cell expression system

Appendix III: Space groups



1 x, y, z

2 $\bar{x}, \bar{y}, \frac{1}{2} + z$

3 $\bar{y}, x, \frac{3}{4} + z$

4 $y, \bar{x}, \frac{1}{4} + z$

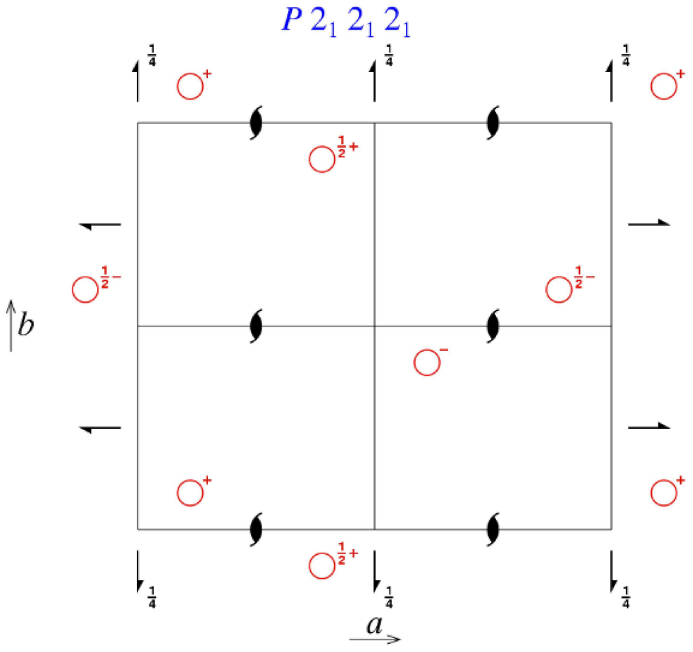
5 $\frac{1}{2} + x, \frac{1}{2} - y, \frac{1}{4} - z$

6 $\frac{1}{2} - x, \frac{1}{2} + y, \frac{3}{4} - z$

7 $\frac{1}{2} - y, \frac{1}{2} - x, \frac{1}{2} - z$

8 $\frac{1}{2} + y, \frac{1}{2} + x, \bar{z}$

Figure I 1: Space group $P4_32_12$



1 x, y, z

2 $\frac{1}{2} + x, \frac{1}{2} - y, \bar{z}$

3 $\bar{x}, \frac{1}{2} + y, \frac{1}{2} - z$

4 $\frac{1}{2} - x, \bar{y}, \frac{1}{2} + z$

Figure I 2 Space group $P2_12_12_1$

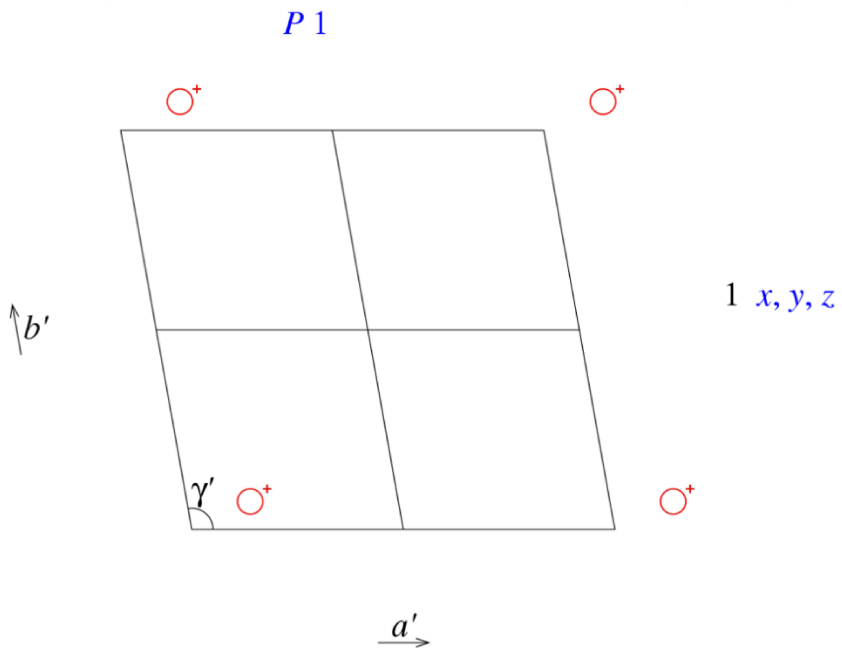


Figure I 3 Space group P1

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Jörg Labahn, for his unwavering support, guidance, and encouragement throughout my PhD journey. His profound knowledge, insightful feedback, and continuous motivation have been invaluable in shaping the course of my research and the successful completion of this dissertation. He has been more than just a supervisor; he has been a mentor and a source of inspiration. His ability to challenge my thinking and push me beyond my perceived limits has greatly enhanced my academic and personal growth. The countless hours he dedicated to discussing my research, providing constructive criticism, and guiding me through complex problems have been crucial in my development as a researcher.

I am profoundly grateful to my co-supervisor, Prof. Dr. Hilal Zaid, for his unwavering support and assistance throughout my studies. His collaboration was instrumental in securing the PhD scholarship, making this achievement possible. I am also grateful to my mentor, Prof. Dr. Oliver H. Weiergräber, for his kind guidance and assistance in protein crystallization and beamline measurements.

I am thankful to the Palestine Academy for Science and Technology and the Palestinian-German Science Bridge (PGSB) for awarding me the scholarship to pursue my PhD.

I would like to thank Dr. Udaya Kumar Tiruttani Subhramanyam for his invaluable assistance during the initial days of my PhD. As my senior, his help in familiarizing me with the lab and his insightful ideas about my project were greatly appreciated. I would also like to thank Dr. Ge Yang for providing various suggestions when I faced difficulties in my research.

I am grateful to my colleagues, Nishika Sabharwal, Yajing Xiao, Abhilasha Kerkmann, Dr. Chengcheng Tao, and Dr. Weihou Guo, for their kindness, support, and every bit of

help. Working with them was a lot of fun, and I could not have accomplished as much without their assistance.

I thank Ian Gering for his help in mass spec measurements and analysis. I would also like to acknowledge technicians Anneliese Cousin and Markus Tusche for being nice and helpful whenever I needed their assistance with lab work.

I am deeply thankful to the secretary's office of IBI-7, especially Dorothea Stobbe, for their unwavering support, exceptional help, and boundless kindness.

Additionally, I am grateful to my friend Mahmoud Shehab for his assistance with my personal life here in Germany. I would like to thank Amin and Pelin for being good friends and for the coffee times we shared.

Last but not least, I owe a special debt of gratitude to my parents for their unconditional love and support. To my beloved wife, Yasmeen, and my little hero, Mahmoud, for their strong support and understanding of my studies in Germany for over four years. I would also like to extend my thanks to all my family members who supported me throughout my career and helped me to stand where I am today.

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und ohne unerlaubte Hilfe verfasst habe. Die vorliegende Arbeit wurde weder in der jetzigen oder in ähnlicher Form bei einer anderen Institution eingereicht. Es wurden zuvor keine Promotionsversuche unternommen.

Jülich

_____

Aziz Tumeh