The molecular basis of central element assembly within the synaptonemal complex during meiosis

Lucy Jennifer Salmon

Thesis presented for the degree of Doctor of Philosophy

Institute for Cell and Molecular Biosciences

Newcastle University

July 2019

Abstract

In the first meiotic division, homologous chromosomes must pair, synapsis, exchange genetic material through crossing over, and then segregate to allow haploid germ cell production. This challenging process is triggered by double-strand break (DSB) induction, which activates repair machinery to generate recombination intermediates between homologues. These physical tethers guide synaptonemal complex (SC) assembly; a supramolecular lattice producing continuous synapsis between homologues. The SC provides the necessary three-dimensional structure for recombination intermediate resolution, including crossover formation, therefore is essential for meiotic division and thereby fertility.

The underlying mammalian SC lattice is provided through self-assembly of SYCP1. However, a further five central element (CE) proteins are essential for full SC structure and function *in vivo*. A direct biochemical interaction was identified between human SYCP1 and CE protein SYCE3, revealing the sole physical link between SYCP1 and the CE. Light and X-ray scattering experiments elucidated conformational re-modelling upon interacting, with SYCE3 modifying the SYCP1 lattice to achieve its regular incorporation. SYCE3 further interacts with, and thereby recruits, other CE proteins thus acting as a molecular glue for SC maturation.

The CE SYCE2-TEX12 complex is recruited via SYCE3 binding, and undergoes fibrous assembly to achieve SC elongation. Through identification of two mutants that separately block SYCE2-TEX12 4:4 complex formation and its fibrous assembly, coupled with detailed biophysical analysis and two X-ray crystal structures, a structural model is proposed for the SYCE2-TEX12 complex and the mechanism of fibrous assembly.

Analysis of mammalian SC assembly *in vivo* is severely hampered by the lack of a genetically tractable system. Whilst SC structure is conserved in yeast, protein constituents show no overt sequence similarity; thus, structural studies of the yeast SC were initiated to understand the underlying conservation. Yeast CE proteins, Ecm11-Gmc2, form a constitutive 2:2 complex with a structural model presented based upon solution scattering studies.

Acknowledgements

Firstly, I would like to sincerely thank my PhD supervisor Dr Owen Davies for his continued support throughout my PhD, and for giving me this opportunity. His wealth of advice has guided me through this experience, for which I am truly grateful. It has been a pleasure to be his first PhD student and to see the lab develop over the years.

This brings me to thank all of the past and present members of the Davies lab who have provided a friendly and humorous working environment. I would like to thank the students I have supervised during this time who have assisted in my research, and lab members I have worked with in collaborative studies, notably James Dunce. In particular I would like to express my gratitude to Dr Orla Dunne, who has been a valued friend and colleague over the past three years and with whom I have thoroughly enjoyed collaborating with. She has provided constant professional and personal support, for which I can't thank her enough. I would also like to thank Dr Ula McClurg for her help and advice along the way.

Next, I would like to acknowledge the extensive support I have gained throughout the university. I really appreciate the technical support provided by Dr Helen Waller for CD experiments and the team at the Newcastle University Electron Microscopy Services for their assistance during TEM experiments. I would particularly like to thank Dr Arnaud Basle for his advice and assistance with crystallography, and specifically for the time he has dedicated to me for training and problem solving. I would also like to thank Prof Rick Lewis for his support throughout my years at Newcastle, and his readiness to lend an ear to discuss any matters.

I would also like to thank our collaborators at Weslyan University, Dr Amy MacQueen and Chandni Ravindan. It has been a pleasure to meet the always enthusiastic Amy on several occasions, and to see my friend and former student Chandni embark on her next step. I also want to acknowledge my gratitude to Prof Juri Rappsilber and Juan Zou for performing mass spectrometry experiments and analysis.

Finally, I would like to express my endless gratitude to my dear family and friends. I am so fortunate to be surrounded by such a supportive network and I can't thank my family enough for their unwavering support. Special thanks go to my big brothers, Rick and Rob, who have always guided me and kept my spirits high. Words cannot describe how grateful I am to my Mum and Dad, who give everything to help their children. They have always encouraged me to pursue my goals and I am eternally grateful for everything they do for me. Lastly I would like to thank my boyfriend Joe for his never-ending support and encouragement. He has kept me sane, cared for me during the most difficult times and has truly been my rock.

Table of contents

List of figures

Chapter One

Chapter Three

Chapter Four

Chapter Five

List of Tables

Abbreviations

Å	Ångstrom
Amp	Ampicillin
APS	Ammonium persulphate
ARLD	Armadillo-repeat-like domain
bp	Base pairs
BS3	Bissulfosuccinimidyl suberate
BSA	Bovine serum albumin
°C	degrees Celsius
CCD	Coiled-coil domain
CD	Circular dichroism
CDK	Cyclin-dependant kinase
C. elegans	*Caenorhabditis elegans*
CE	Central element
CENP	Centromere protein
CO	Crossover
D. melanogaster	*Drosophila melanogaster*
Da	Dalton
dH_2O	Distilled water
Dmax	Maximum dimension
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
dRI	Differential refractive index
DSB	Double-strand break
dsDNA	Double-stranded deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
Ecm11	Extracellular mutant protein 11
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
GAL4	Regulatory protein GAL4
Gmc2	Grand meiotic recombination cluster protein 2
H_2O	Water
HEAT	Protein motif named after huntingtin, elongation factor 3, protein phosphatase 2A, Tor1 kinase
Hop1	Meiosis-specific protein Hop1
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
kb	kilo-base
kDa	kilo Dalton
L	Litre
LB	Luria-Bertani
LDS	Lithium dodecyl sulphate
LE	Lateral element

LS	Light scattering
M	Molar
MALS	Multi-angle light scattering
MBP	Maltose binding protein
MCS	Multiple cloning site
MES	2-(N-morpholino)ethanesulfonic acid
mg	Milligram
ml	Millilitre
mm	Millimetre
mM	Millimolar
MOPS	3-Morpholinopropanesulfonic acid
MPD	2-Methyl-2,4-pentanediol; also known as hexylene glycol
MRE	Mean residue ellipticity
ms	Millisecond
Mw	Molecular weight
MWCO	Molecular weight cut-off
NEB	New England Biolabs
Ni-NTA	Nickel-nitrilotriacetic acid
nl	Nanolitre
nm	Nanometre
OD_{600}	Optical density measured at 600 nm
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
pH	Hydrogen potential
pI	Isoelectric point
PLK	Polo-like kinase
P(r)	Paired-distance distribution function
Q	Scattering vector
Rc	Radius of gyration of the cross-section
Red1	Reductional division protein 1
Rg	Radius of gyration
rpm	Revolutions per minute
s	Seconds
S. cerevisiae	*Saccharomyces cerevisiae*
SAXS	Small-angle X-ray scattering
SC	Synaptonemal complex
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
sfGFP	Superfolder green fluorescent protein
SIC	Synapsis initiation complex
SIX6OS1	Six6 opposite strand transcript 1
SLD	Spt16M-like domain
SLIC	Sequence and ligation independent cloning

Smt3	Ubiquitin-like protein Smt3
SOC broth	Super optimal broth with catabolite repression
S/T P	Serine/Threonine phosphorylation domain
SUMO	Small ubiquitin-like modifier
SYCE1	Synaptonemal complex central element protein 1
SYCE2	Synaptonemal complex central element protein 2
SYCE3	Synaptonemal complex central element protein 3
SYCP1	Synaptonemal complex protein 1
SYCP2	Synaptonemal complex protein 2
SYCP3	Synaptonemal complex protein 3
TBE	Tris/Borate/EDTA
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus
TEX12	Testis-expressed protein 12
TF	Transverse filament
Tm	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
µg	Microgram
µl	Microlitre
µm	Micrometre
µM	Micromolar
UV	Ultraviolet
V	Volts
v/v	volume per volume concentration percentage
w/v	weight per volume concentration percentage
WT	Wild type
x g	Multiplied by gravitational force
Y2H	Yeast 2-hybrid
Zip1	Synaptonemal complex protein Zip1
Zip3	Molecular zipper protein 3
2D	Two dimensional
3D	Three dimensional

Amino acid codes

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

Chapter one.

Introduction

1.1 Meiotic cell division

1.1.1 The cell cycle

The generation of haploid gametes for sexual reproduction is central to eukaryotic life. This is achieved by the cellular division process of meiosis, which involves two rounds of cell division without an intermediate DNA replication stage (Petronczki *et al*., 2003). To ensure each gamete is genetically unique, regions of genetic material are exchanged between maternal and paternal chromosomes during early meiosis (Baudat *et al*., 2013).

The cell cycle contains an initial growth stage (G1 phase), a DNA synthesis step (S phase), another growth stage (G2 phase) and then the mitotic/meiotic phase (M phase) begins (Figure 1.1). G1 phase prepares the cell for DNA duplication by doubling the other cellular contents. DNA is subsequently duplicated in S phase and then G2 phase checks the newly duplicated DNA for errors and prepares for meiotic entry. The collective term for G1, S and G2 stages is named interphase and numerous cell cycle control events monitor the progress of interphase (Alberts, 2002).

In mitosis, the cell containing duplicated DNA undergoes one round of cell division, where one copy of each sister chromatid is carried forward to the diploid daughter cells (Figure 1.1). In meiosis, homologues are paired through synapsis, undergo genetic exchange and are then segregated upon the first meiotic division. Following this, sister chromatids are then segregated upon a second cell division event, without further DNA duplication, to produce four non-identical haploid gametes for reproduction (Figure 1.1) (Handel and Schimenti, 2010).

A whole host of cellular and molecular events are involved in the coordination of meiosis, with tight regulation provided by numerous cell cycle control mechanisms. This is to ensure that the cell cycle does not proceed to the next stage whilst errors, such as DNA damage, are prevalent. The cyclin-dependant kinase (CDK) family of proteins have a crucial role in the coordination of cell cycle progression with their respective cyclin partners. In particular, CDK2 in mammals is required for completion of meiosis I, but is not needed for mitosis (Ortega *et al*., 2003).

1.1.2 Chromosome architecture

DNA molecules are packed into condensed forms called chromatin in order to package all of the DNA in the cell into the nucleus in an organised manner. The chromatin structures are formed by DNA-protein complexes where DNA is wound around protein structures, called histones, to form condensed units (Kornberg, 1974). The histone protein family contains H1, H2A, H2B, H3 and H4 proteins that are positively charged (Thomas and Kornberg, 1975). The

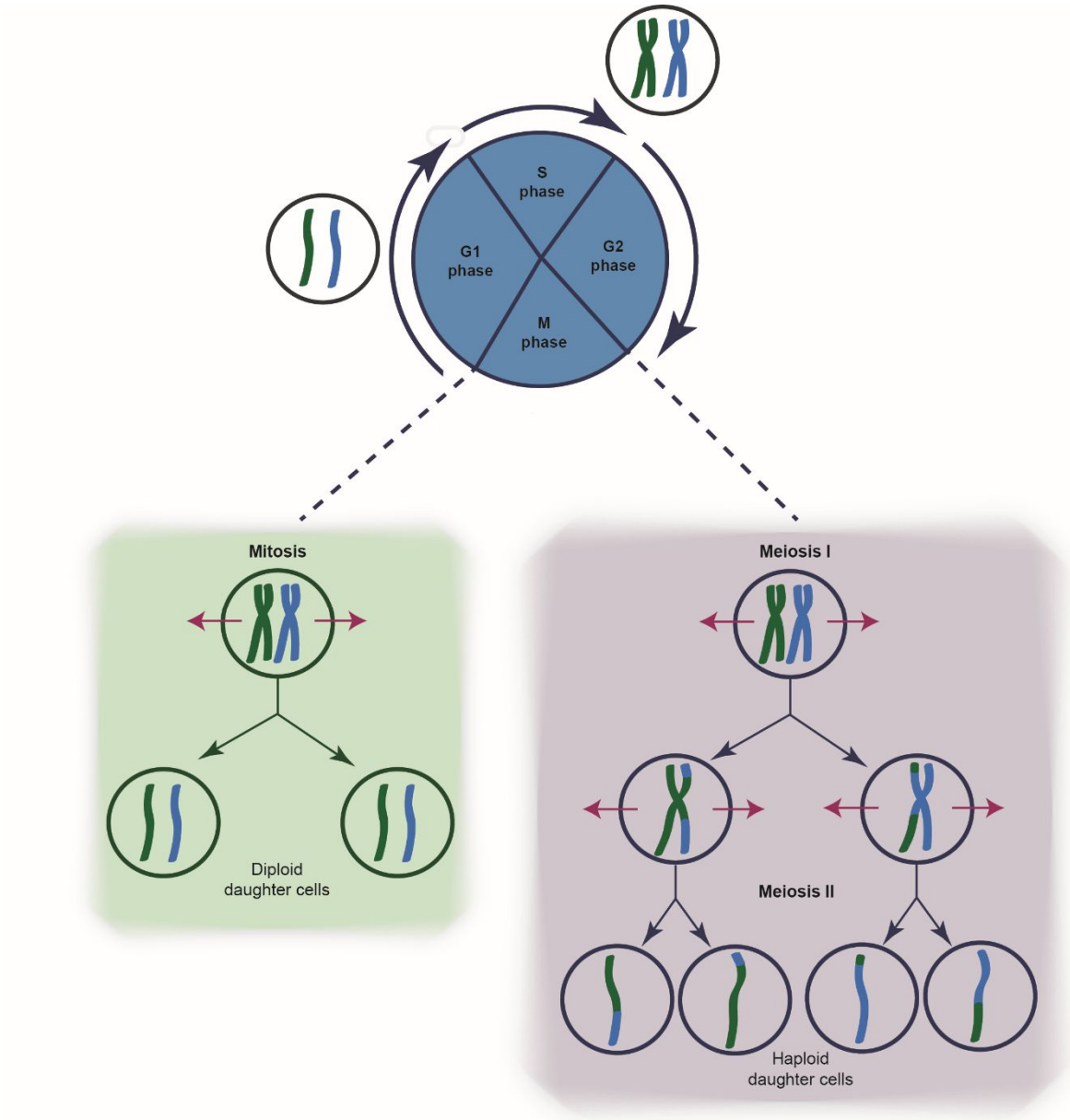


Figure 1.1. Mitotic and meiotic cell division. The cell cycle is comprised of G1, S, G2 and M phases. Diploid cells undergo DNA duplication in S phase, and M phase corresponds to either mitosis or meiosis. Mitotic division involves one cell division event where sister chromatids are segregated to produce two identical diploid daughter cells. Meiotic division has two stages: meiosis I and meiosis II. In meiosis I, homologous chromosomes synapsis and exchange genetic material before segregation of homologues upon the first division. Sister chromatids are then separated in the second division to produce four non-identical haploid cells for sexual reproduction.

positively charged protein units can then form electrostatic interactions with the negatively charged DNA. The DNA is wound around histones and the DNA-histone units are called nucleosomes. Each nucleosome is formed by a hetero-octamer of two copies each of the histone proteins H2A, H2B, H3 and H4 with a stretch of DNA wound around the octamer (Luger *et al*., 1997). The H1 histone then acts as a linker histone that binds at the point in the nucleosome where the DNA enters and leaves the nucleosome (Figure 1.2 A). Nucleosomes are then packaged by continual coiling until the DNA molecule is successfully packaged into a chromatin structure (Kornberg, 1974).

During cell division, further chromatin condensation takes place in order to compact the DNA in preparation for cell division. DNA duplication occurs during S phase and the sister chromatids produced are connected by a structure called a centromere (Clarke and Carbon, 1980) (Figure 1.2 B). The short arms of the chromatids are called the p arms and the long arms are called the q arms and the regions at the end of the chromatids are called the telomeres (Blackburn, 2000) (Figure 1.2 B). Telomeres act as protective caps at the ends of chromosomes and consist of a series of double-stranded repetitive nucleotide units that end with a short single-stranded stretch of DNA (Grandin and Charbonneau, 2008). The DNA forms a telomere loop by invasion of the single stranded stretch to form a displacement loop (Meselson and Radding, 1975) (see Section 1.2.1). The purpose of telomeres is to protect the chromosomes, and therefore vital genetic information, from DNA damage. Telomeres in turn are protected by a protein complex called shelterin that prevents the telomeric DNA cap from being target by double-strand break machinery (Palm and de Lange, 2008).

During cell division, the DNA at the end of chromosomes cannot be duplicated so the copy made has a short DNA truncation. Due to the repetitive nucleotides, only telomeric DNA is truncated. This means that during every cell division event the telomere ends become shorter. In some cell lines, such as germ cells, an enzyme called telomerase can replace the nucleotides that are lost during replication (Blackburn *et al*., 2015). However, in most cell types, the DNA loss occurred at each cell division means that each chromosome has limited number of cell division events before replication results in loss of genetic information. Therefore, telomeres have been implicated in cancer and aging (Blackburn *et al*., 2015).

1.1.3 Meiotic phase

The meiotic phase, similarly to mitosis, is divided into several different stages, though meiosis differs from mitosis as there are two cell division events rather than one (Figure1.1). Meiosis itself is divided into two general stages (Meiosis I and Meiosis II) with each then further split into prophase, metaphase, anaphase, telophase and cytokinesis (Gilbert, 2000a).

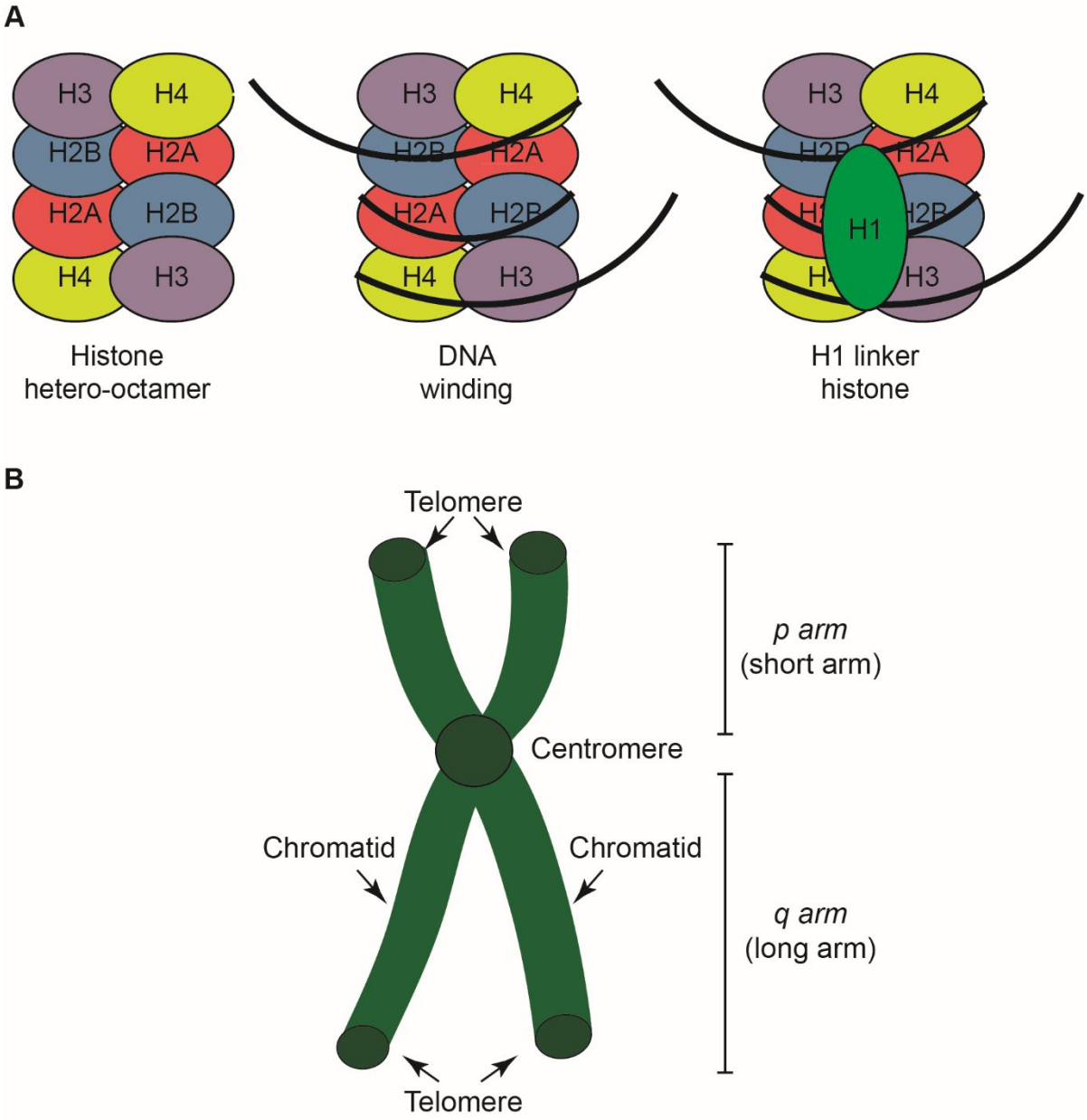


Figure 1.2. DNA compaction and chromosome structure. (A) A hetero-octomer of histone proteins is formed by two copies each of H2A, H2B, H3 and H4. DNA then winds around the histone units to form the nucleosome. The H1 histone protein then binds to the point where DNA enters and leaves nucleosome in order to act as a linker. (B) Nucleosomes are continually compacted to form chromatin and chromatin is further condensed during cell division. Sister chromatids, produced during DNA replication in S phase, are joined by structures called centromeres. The short arms of the chromatids are called the p arms and the long arms are called the q arms. Telomeres cap the ends of chromatids in order to protect genetic information from DNA damage.

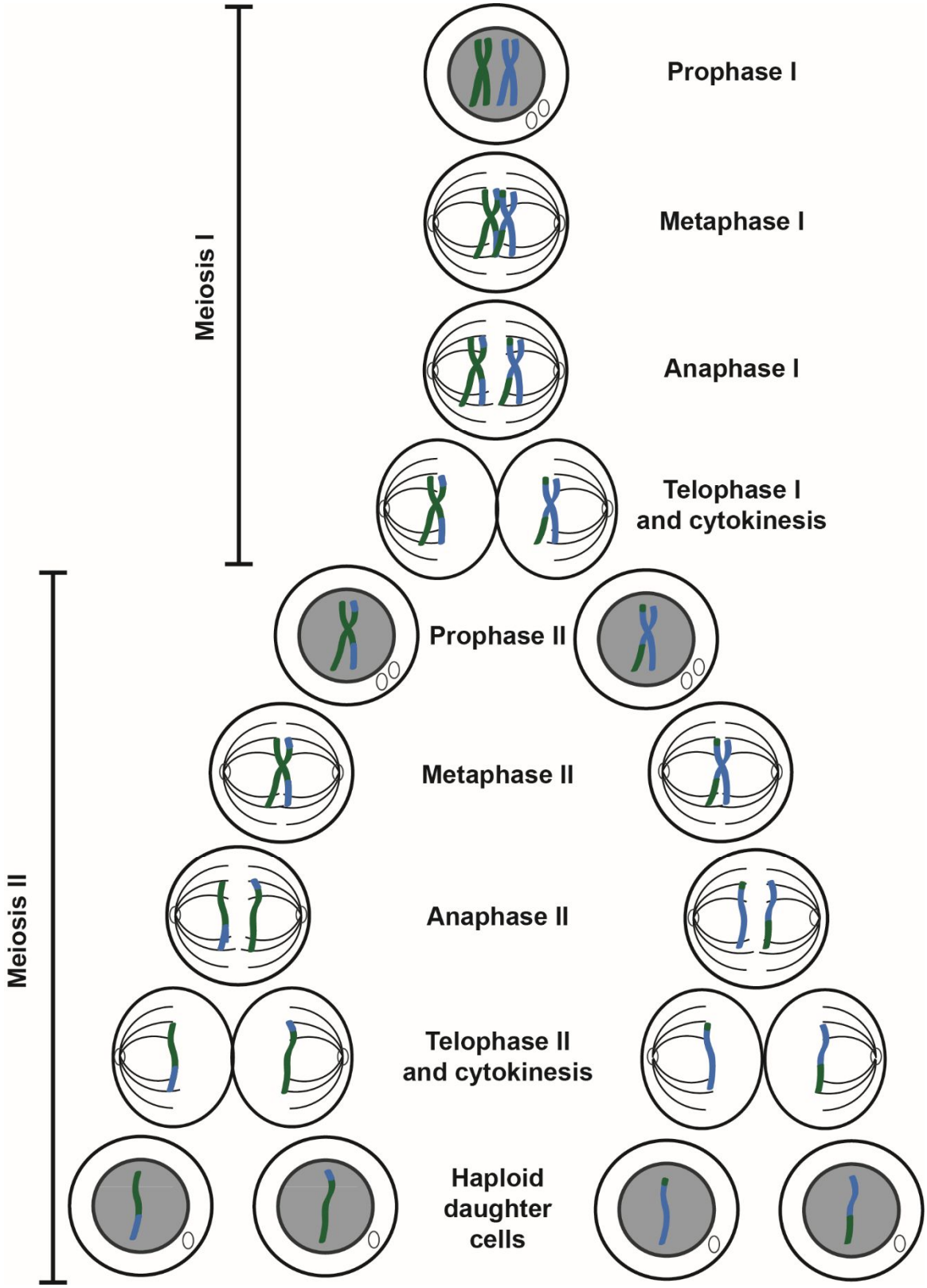


Figure 1.3. Stages of meiotic division. Homologous chromosomes pair in prophase I and undergo crossover formation to exchange genetic material. Homologues align on the spindle in metaphase I and then are separated during anaphase I. The nuclear membrane starts to reform in telophase I and the cells are divided by cytokinesis to produce two diploid daughter cells. Each daughter cells then goes through meiosis II where the sister chromatids are separated to form four haploid daughter cells upon the second meiotic division. The spindle forms from a microtubule organising centre (MTOC).

Homologous chromosomes are synapsed in prophase I, align on the metaphase plate in metaphase I, segregated in anaphase I and then the cell completely divides in telophase I followed by cytokinesis to form the respective cell membranes (Figure 1.3). The second meiotic division is similar to mitosis, where prophase II leads to chromatids aligning on the metaphase plate in metaphase II, with sister chromatid segregation in anaphase II followed by telophase II and cytokinesis (Alberts, 2002) in order to form the four haploid daughter cells (Figure 1.3).

The division of the homologous chromosomes and sister chromatids during meiosis I and meiosis II, respectively, relies on the formation of the meiotic spindle (Figure 1.3). The meiotic spindle protrudes from microtubule organising centres (MTOCs) and binds to the chromosomes during alignment on the metaphase plate (Alberts, 2002). The attachment of the spindle microtubules to the chromosomes occurs via the kinetochore. The kinetochore is a large protein assembly associated to the centromeric region of proteins to facilitate the attachment of microtubules to the chromosomes (Gorbsky *et al*., 1987). Kinetochores have an "inner" region that interacts with the centromere and an "outer" region that connects to the microtubules. When the spindle contracts in order to separate chromosomes, correct attachments of microtubules to kinetochores is essential in order to provide equal forces to segregate the homologues/sister chromatids (Miller *et al*., 2012).

It has been described that microtubules have differing rigidity as they span from the spindle poles towards the equator. The microtubules spanning from the spindles poles are in a parallel arrangement with high rigidity. The parallel microtubules then relax in the middle and then form rigid arrays once more at the equator, where anti-parallel microtubule arrays are formed with the microtubule ends spanning form opposing spindle poles (Takagi and Shimamoto, 2018).

The timeline between male and female meiotic division processes are extremely diverse, with variability between the overall processes due to the difference in the cell types that are made. Spermatogenesis occurs from the onset of puberty and is a continuous process that takes place without interruption in order to produce four haploid gametes for reproduction (Gilbert, 2000a). Conversely, oogenesis is a prolonged process that breaks at multiple meiotic stages for large lengths of time. Oogenesis onset occurs at the embryonic stage, where all of the primary oocytes for an individual's life are created. The cells then pause at prophase I of meiosis, and the oocytes only recommence meiosis upon the onset of puberty. The cells will then continue to metaphase II upon ovulation, and the eggs will then only undergo completion of meiosis by the second meiotic division if fertilised. Therefore, the process of meiosis in females can extend over many decades (Toth and Jessberger, 2016). The long timescale of the female meiotic process has been linked to several impairments of fertility due to age-related defects of eggs.

Notably, the loss of chromosome cohesion has been linked to meiotic errors as the sister chromatids are held together by the same proteins, without replenishment, whilst meiosis is paused. Therefore, as women age the cohesion maintenance can be weakened, which leads to chromosome mis-segregation that gives rise to aneuploidies, such as Down syndrome (Chiang *et al*., 2010).

Another key difference between male and female meiosis is that males form four gametes for reproduction, whereas females only form one egg per meiotic division (Figure 1.4). In spermatogenesis, each cell division is equal to ensure the same amount of genetic material and cell content is divided between gametes. However, although the number of chromosomes divided between the cells is also equal in oogenesis, there is a biased distribution of the cell contents upon division. Upon the first meiotic division, one cell is produced with a much larger cell mass and the other cell is much smaller, termed the polar body (Figure 1.4). In meiosis II,

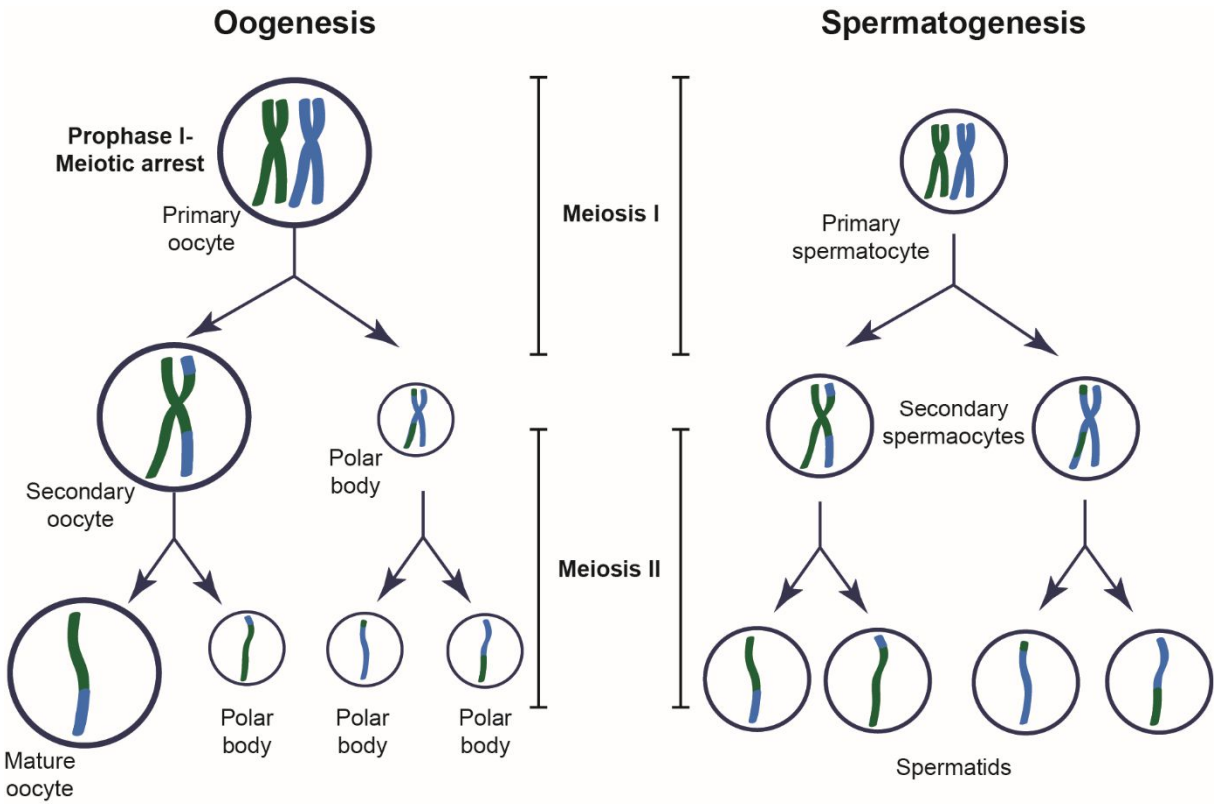


Figure 1.4. Female (oogenesis) and male (spermatogenesis) gametogenesis. During oogenesis, the cell is arrested at prophase I until the onset of puberty where meiosis continues upon ovulation. The first division produces a secondary oocyte and a polar body. The second division results in the formation of one mature oocyte and three polar bodies. During spermatogenesis, meiosis is continual and the cell division events produce 4 mature spermatids.

the large cell divides to produce the functional egg for sexual reproduction and extrudes another polar body. The polar body produced in meiosis I divides again in meiosis II in to produce two more polar bodies (Gilbert, 2000b). Therefore, oogenesis leads to the production of one gamete and three polar bodies (Figure 1.4). This process occurs due to the large cell mass required for the egg cells, and the iterative reshuffling of the cell content allows the formation of the rich egg cell for reproduction. This complements the formation of the much smaller sperm cells, where sperm deliver DNA but the egg provides the cell nutrients (Gilbert, 2000b).

1.2 Homologous recombination in meiosis I

1.2.1 Homologous recombination

The initiation of meiotic recombination has been attributed to the formation of double strand breaks (DSBs) across the genome. These breaks are required for consequent function in early meiosis to pair homologous chromosomes. This observation was initially made in *Saccharomyces cerevisiae* where induction of DSBs lead to recombination in meiosis (Keeney *et al*., 1997). The process of recombination requires initiation, strand exchange and subsequent resolution.

The Spo11 enzyme is a topoisomerase VI-like enzyme that induces DSB formation through its active site. One Spo11 molecule is needed per strand of DNA, so two Spo11 enzymes are necessary to form each DSB. Following DSB induction, the Spo11 enzymes dissociate from the DNA strands and the 5' end of each strand at the break position undergoes end resection to produce 3' overhangs (Keeney, 2008). This is proceeded by strand invasion to the homologous strand, mediated by DNA recombinases, and displacement-loop (D-loop) formation in order to use the homologous DNA as a template for repair (Meselson and Radding, 1975) (Figure 1.5).

Synthesis-dependant strand annealing (SDSA), which occurs in both mitosis and meiosis, results in non-crossover recombination products with short regions of non-reciprocal gene transfer (McMahill *et al*., 2007). The DNA helicase Sgs1 (yeast) or BLM (humans) directs the formation of non-crossover products (Rockmill *et al*., 2003) (Wu and Hickson, 2003). Double-strand break repair (DSBR) results in the formation of a double Holliday junction (dHJ), and subsequent resolution results in the transfer of large regions of genetic material in order to form crossover recombination products (Figure 1.5). The recombination intermediate structure formed by the dHJ is resolved through targeted asymmetric cleavage of DNA by the MLH1-MLH3 hetero-dimer endonuclease (Schwacha and Kleckner, 1995). This asymmetric cleavage results in the formation of chiasmata that form the physical tethers between recombination intermediates (Schwacha and Kleckner, 1995). The stabilisation of dHJ

structures and promotion of crossover formation is achieved in yeast and humans by the MSH4-MSH5 hetero-dimeric complex (Snowden *et al*., 2004), which is homologous to the bacterial MutSγ DNA mismatch repair protein (Blackwell *et al*., 2001). Crossover interference is a process that prevents the formation of multiple crossovers in close proximity, and ensures the formation of at least one crossover per chromosome arm. However, mechanisms governing this regulation are currently uncertain (Martini *et al*., 2006).

1.2.2 Meiotic recombinases

The recombinase proteins form filamentous assemblies with single stranded DNA to produce a nucleoprotein filament that can undergo strand invasion for homology searching (Meselson and Radding, 1975). The bacterial RecA ATPase is a well characterised DNA repair protein that assembles onto single-stranded DNA to produce the nucleoprotein filament. This filament undergoes homology searching through strand invasion, with tighter DNA binding occurring when ATP is bound (Williams and Spengler, 1986).

In eukaryotes, the RecA homologue is RAD51, which is conserved from humans to yeast, and RAD51 functions in both mitosis and meiosis (Benson *et al*., 1994) (Shinohara *et al*., 1992). Another eukaryotic RecA homologue, DMC1, functions specifically in meiosis (Bishop *et al*., 1992). RAD51 and DMC1 both function in meiosis to form nucleoprotein filaments in order to achieve strand exchange for DSB repair (Bishop, 1994). RAD51 has a preference for the repair of DSBs using sister chromatids as the repair template, whereas DMC1 functions to use the homologous chromosomes for repair, reasoning the meiosis-specific function of DMC1.

In meiosis, RAD51 is required to support DMC1 activity, but the activity of RAD51 strand exchange is not essential (Cloud *et al*., 2012). In a recent study, RAD51 and DMC1 filaments have been demonstrated to form distinctly in a self-organising manner, without the need for co-factors for separation, with the DMC1 filaments indicated to stabilise the RAD51 filaments (Crickard *et al*., 2018).

1.2.3 Homologue synapsis in prophase I

Prophase I is divided into five sub-stages: leptotene, zygotene, pachytene, diplotene and diakenesis. In leptotene, chromatin condenses to form compact chromosome structures. During zygotene, the homologous chromosome undergo synapsis, driven by the formation of the synaptonemal complex, and during pachytene crossing over occurs between homologues to exchange genetic material (Alberts, 2002). In diplotene, homologue pairs being to migrate and are now only held together by the crossover intermediate linkages, called chiasmata. Diakinesis

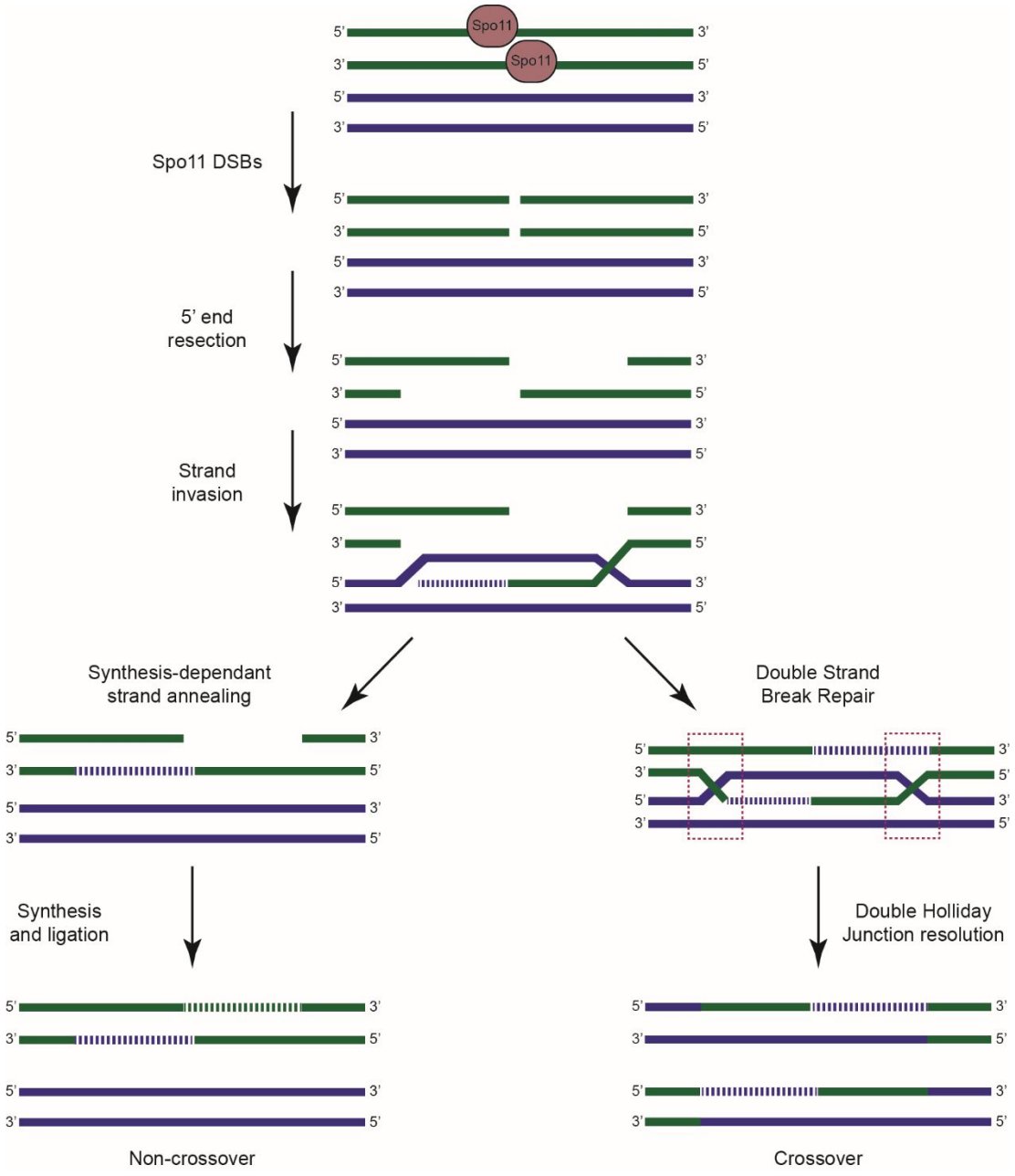


Figure 1.5. Schematic for non-crossover and crossover formations upon Spo11 DSBs. Two Spo11 proteins induce DSBs. The 5' ends are then resected, followed by homologue strand invasion and D-loop formation. Synthesis-dependant strand annealing leads to non-crossover recombination products, with short regions of non-reciprocal DNA transfer. Double strand break repair leads to the formation of a double Holliday junction (pink boxes) that result in crossover products upon resolution.

is the final stage of prophase I and primes the cell for entry into metaphase, which involves further condensation of the chromosomes, dissolution of the nuclear envelope and starting the formation of the meiotic spindle (Alberts, 2002).

Spo11 induced DSBs, and subsequent repair by homologous recombination, are key processes during early meiosis. Programmed DSBs are induced across the genome, with 200-400 breaks induced per cell (Turner, 2007). These break events trigger chromosomes to search for their homologous partner in prophase I through strand invasion in order to form initial homologous chromosome alignments. The homologous chromosomes then need to undergo complete synapsis for the repair of the DSBs and consequent meiotic progression (Zickler and Kleckner, 2015). To achieve homologous chromosome synapsis, the large "zipper-like" multi-protein assembly, the synaptonemal complex (SC), assembles between the homologues to provide a physical tether and structural framework for the repair of DSBs (Moses, 1956). The DSBs are resolved by non-crossovers and crossovers, with at least one crossover occurring per chromosome arm (Martini *et al*., 2006). The SC then disassembles to leave the homologues tethered by the crossover intermediates in order to maintain alignment until the crossovers are fully resolved (Figure 1.6). The chromosomes are segregated upon meiotic progression and the first meiotic division. The failure to form the mature and functional SC structure prevents the repair of the DSBs and therefore, causes meiotic arrest that prevents the production of gametes for reproduction, resulting in infertility (Kouznetsova *et al*., 2011).

1.3 The synaptonemal complex

1.3.1 Architecture and function of the synaptonemal complex

The SC is a highly conserved tripartite assembly that provides the structural and functional framework for homologous chromosome synapsis, repair of DSBs and consequent crossover formation. The SC forms a ladder-like assembly between homologues to impose a three-dimensional architecture for the maintenance of close homologue proximity of approximately 100 nm (Westergaard and von Wettstein, 1972). The formation of the SC is essential to repair the DSBs by homologous recombination, validated by analysing mice lacking SC components that exhibit infertility due to meiotic arrest (Kouznetsova *et al*., 2011).

The SC was first discovered in 1956 through the observation of specific chromosome structures in crayfish spermatocytes (Moses, 1956). The tripartite nature originates from the zipper-like assembly that the SC forms between homologues, with lateral elements coating the chromosomal axes, a central element in the midline and transverse filaments forming the "teeth of zipper" to complete the structure (Figure 1.6). The SC is conserved throughout sexually reproducing organisms, from yeast to humans.

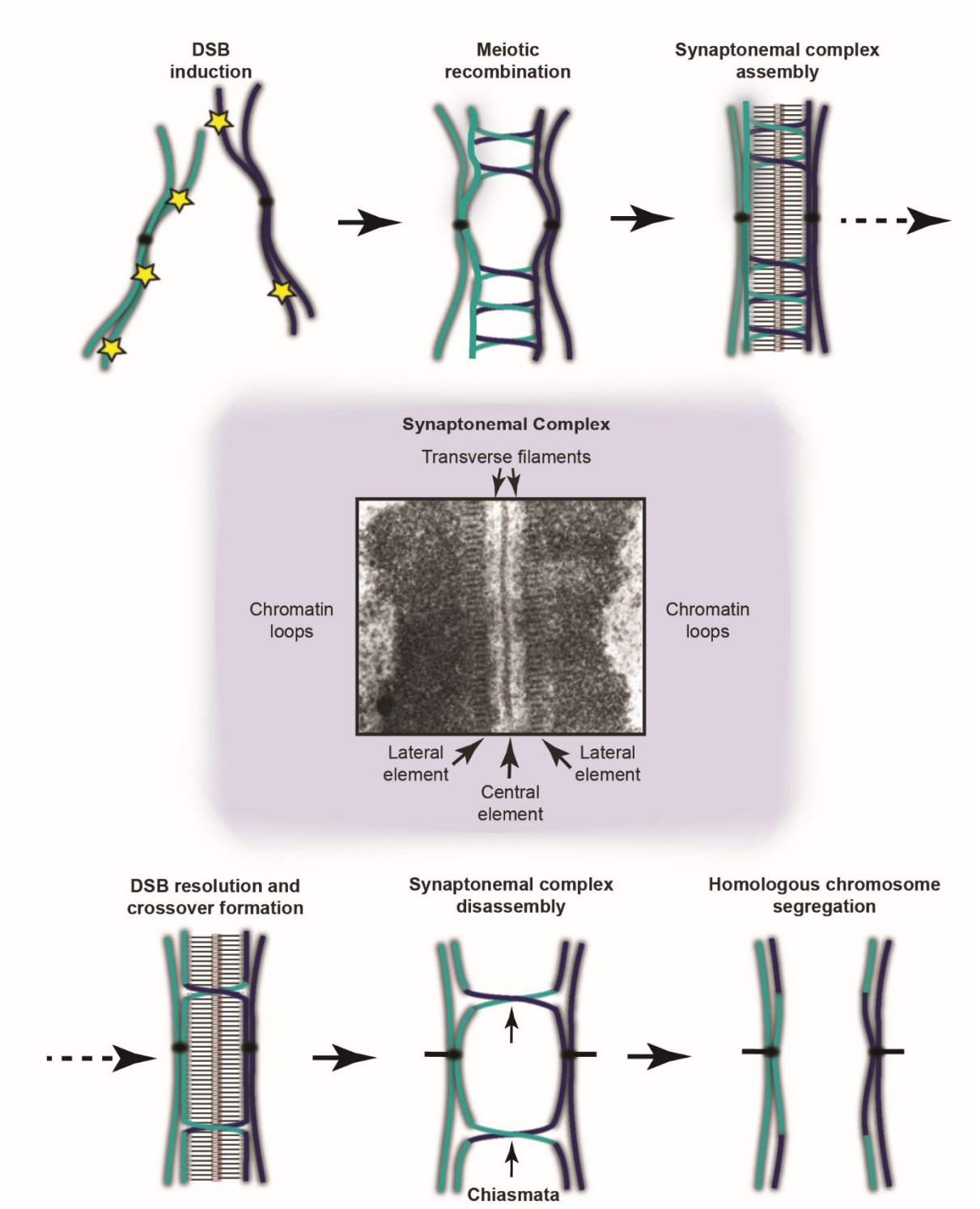


Figure 1.6. Schematic for homologous chromosome synapsis in prophase I. Spo11-induced DSBs (yellow stars) prompt a homology search for the initial pairing and alignment of homologous chromosomes. Full synapsis and crossover formation is achieved through the assembly of the synaptonemal complex (SC). The SC disassembles and the crossovers are resolved before segregation. Transmission electron micrograph of the fungal SC displaying the chromatin loops, the lateral elements at the chromosome axis, the central element in the midline and transverse filaments to connect the structure; image taken from Westergaard and von Wettstein, 1972.

1.3.2 Synaptonemal complex composition

Although the ornate tripartite architecture and overall function of the SC is conserved throughout all meiotically dividing organisms, the individual components differ between species. The mammalian SC is highly conserved on the protein level, but the common model organisms *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster* each have a different protein component composition.

Axial element structures, which are the lateral element precursors, form at the chromosome axis and contain a wide range of proteins (Zickler and Kleckner, 1999). The axial elements become lateral elements upon SC formation, where long transverse filament proteins coat the chromosome axes to form the "teeth" of the zipper structure. The central element then forms a protein-dense region in the midline that completes the structure and extends synapsis along the entire length of the homologous chromosome pairs (Solari and Moses, 1973).

The SC is thought to form through a pattern of recruitment within prophase I, by a network of recursive interactions. However, once fully formed, the SC is an extremely large and complex structure. The mechanisms of SC disassembly are not currently clear, but several indications suggest that modifications, such as phosphorylation, play a role in disassembly, likely in a cell-cycle regulatory manner (Gao and Colaiacovo, 2018). The polo-like kinase, PLK1, has been shown to phosphorylate SC components in mice, with PLK1 inhibition abolishing these events and preventing meiotic exit due to failure of central element disassembly (Jordan *et al*., 2012).

1.4 The mammalian synaptonemal complex

1.4.1 Composition of the mammalian synaptonemal complex

The tripartite synaptonemal complex (SC) is conserved across all meiotically dividing organisms, with its "zipper-like" assembly functioning to tether homologous chromosomes across their entire lengths. This process achieves the resolution of induced DSBs with the production of crossovers, giving rise to genetic diversity between gametes.

In mammals, the lateral elements are comprised of the SYCP2 and SYCP3 proteins. The transverse filaments, which bridge from each lateral element and meet in the midline, are composed by the SYCP1 protein. The central element forms the protein-dense midline assembly that stabilises the SC for full maturation and homologous chromosome synapsis. The mammalian central element proteins are SYCE1, SYCE2, SYCE3, TEX12 and SIX6OS1. (Figure 1.7). To date, these are the currently identified essential SC protein components. However, as the central element proteins in particular have been characterised in more recent

years, with the newly discovered SIX6OS1 component identified in 2016 (Gómez-H *et al*., 2016), it is possible that there are more SC components to be characterised.

The molecular basis of the mammalian SC had, in recent times, been investigated though a limited amount of *in vitro* studies of individual components. However, characterising the array of protein-protein complexes that occur, their respective structures and the oligomeric properties may provide further elucidation into the intricate mechanism of SC assembly.

1.4.2 Lateral elements

The mammalian SC lateral elements are composed of the SYCP2 and SYCP3 proteins, which are recruited to the axial element, the lateral element precursor, at the chromosome axes (Heyting *et al*., 1988). The SYCP2 protein was initially studied in rat spermatocytes and found to be a 173 kDa protein, with a coiled-coil stretch at the C-terminus, and hypothesised to be a DNA-binding protein involved in the organisation of chromosomes in prophase I (Offenberg *et al*., 1998). The amino acid sequence and secondary structure prediction is similar to the yeast axial element component Red1, suggesting these proteins undergo similar functions in the mammalian and yeast SC (Offenberg *et al*., 1998).

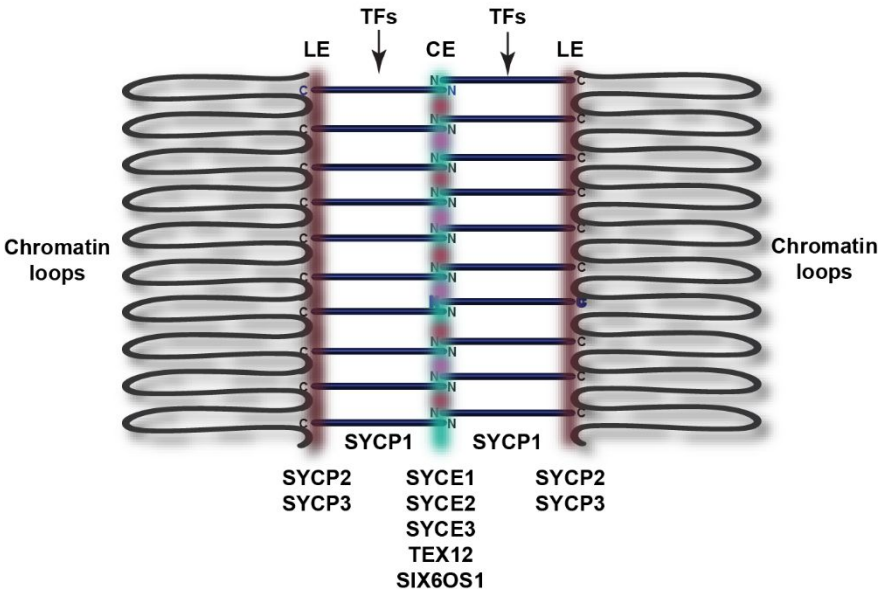


Figure 1.7. Schematic of the mammalian synaptonemal complex. LE = lateral element. TFs = transverse filaments. CE = central element. Schematic depicting the chromatin loops of the homologous chromosomes with the assembly of the lateral elements at the axes (red). The transverse filaments (purple tubes) bridge between the chromosome axis and the central element (orange, blue and pink). The lateral elements are comprised of SYCP2 and SYCP3. SYCP1 forms the transverse filaments and the central element is formed by SYCE1, SYCE2, SYCE3, TEX12 and SIX6OS1.

The removal of the SYCP2 C-terminal coiled-coil domain results in a sexual dimorphic phenotype, where males are infertile and females are sub-fertile. The male infertility is caused by a failure of axial element formation and resultant failure of chromosome synapsis, leading to meiotic arrest, with the sub-fertility of females resulting in a reduced litter size and a higher frequency of aneuploidies (Yang *et al*., 2006). SYCP2 has been found to interact with SYCP3, both in *vitro* and *in vivo*, via the coiled-coil domain in order to form heterodimers (Yang *et al*., 2006). Further to this, the mutant SYCP2 protein (lacking the coiled-coil domain) was found to maintain localisation to the axial elements in both males and females, but the SYCP3 localisation was abolished. Therefore, SYCP2 is required for the recruitment of SYCP3 to the axial elements by the interaction via the SYCP2 coiled-coil region (Yang *et al*., 2006).

SYCP2 has also been found to interact with the SYCP1 transverse filament protein by co-immunoprecipitation and yeast 2-hybrid studies, with the interaction further dissected to take place with the C-terminal region of SYCP1 (Winkel *et al*., 2009). Therefore, it was hypothesised that SYCP2 is not only responsible for recruiting SYCP3 to the axial elements, but also to recruit the transverse filaments to the SC via the C-terminus, thereby providing the link between the lateral elements and transverse filaments.

The X-ray crystal structure of an N-terminal region of SYCP2 was solved, containing the ARLD (armadillo-repeat-like domain) and SLD (Spt16M-like domain) regions (Figure 1.8). The ARLD is comprised by 14 helices and the SLD forms a twisted β-sheet formed by ten strands and two helices (Feng *et al*., 2017). The ARLD is proposed to form a surface for potential interactions, and the SLD is postulated to undergo interactions with chromatin (Feng *et al*., 2017).

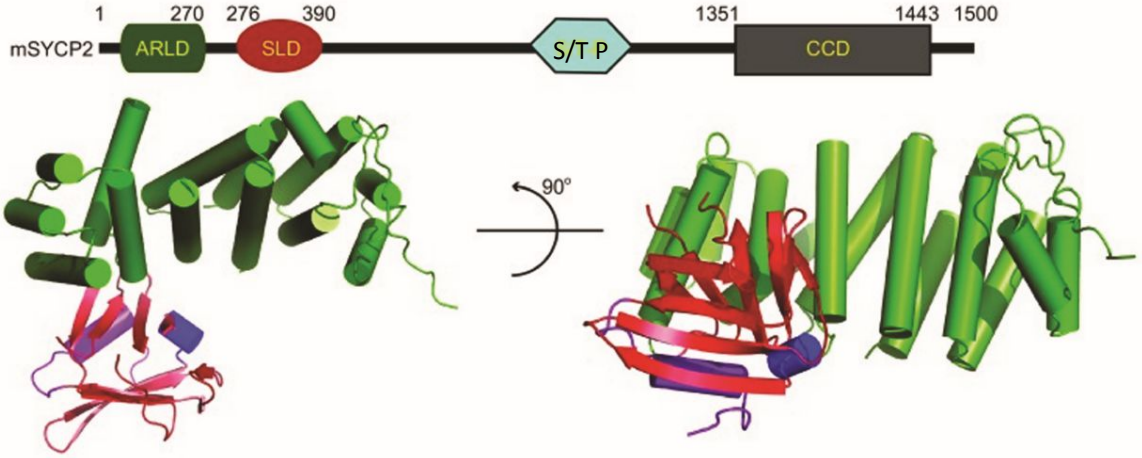


Figure 1.8. Crystal structure of the SYCP2 N-terminal domain. The schematic above represents the linear organisation of SYCP2, containing ARLD (Armadillo-repeat-like domain), SLD (Spt16M-like domain), S/T P (Ser/Thr phosphorylation domain) and the C-terminal CCD (coiled-coil domain). The crystal structure of the ARLD and SLD is formed by 14 alpha helices from the ARLD and the SLD forms a 10 strand β-sheet capped by two helices, shown in two orientations. Taken from Feng *et al*., 2017. PDB code: 5IWZ.

Further investigation of potential SYCP2 interacting partners revealed capacity for the N-terminal region to bind to centromeric proteins through yeast 2-hybrid screening. These interactions were validated by co-immunoprecipitation, whereby SYCP2 was found to strongly bind to CENP J, and also to bind to CENP F to a weaker extent (Feng *et al*., 2017). Therefore, SYCP2 forms interactions with centromeric proteins within the N-terminal region and binds to SYCP1 and SYCP3 through the C-terminal coiled-coil region.

Similarly to SYCP2, the SYPC3 protein is also required for male fertility, where a null mutation of the mouse SYCP3 resulted in the failure of chromosome synapsis, due to the failure of axial element formation, which leads to extensive cell death (Yuan *et al*., 2000). Further to this, deletion of SYCP3 in mice was found to cause oocyte aneuploidy in females, extension of meiotic chromosomes and the failure of SYCP2 recruitment to the chromosomes (Kolas *et al*., 2004). In the mice harbouring the SYCP3 deletion, and therefore absence of SYCP2 at the chromosomes, the chromosome loop attachment was more frequent, thus suggesting a role for SYCP2 and SYCP3 in directing chromatin attachment specificity (Kolas *et al*., 2004).

SYCP3 has been identified to form fibrous assemblies *in vivo* and these striated fibres are formed by SYCP3 self-assembly, which have been characterised by SYCP3 synthesis in somatic cells and by transmission electron microscopy (Yuan *et al*., 1998). The periodicity of the striations were determined as 20 nm, with the pattern of the fibres mimicking the lateral element structures observed within intact SCs (Figure 1.9 B). The polymerisation of the SYCP3 fibres has been shown to display similar patterns from divergent SYCP3 proteins (Baier *et al*., 2007), with the assembly driven by the six C-terminal SYCP3 amino acids (Miyamoto *et al*., 2003).

Structural elucidation of the human SYCP3 protein, harbouring a C-terminal truncation of the six amino acids that mediate assembly, revealed the formation of an anti-parallel tetrameric helical bundle (Figure 1.9 A). The elongated nature of the tetramer provides a length of 20 nm (Syrjanen *et al*., 2014), consistent with the periodicity of the SYCP3 self-assembled fibres (Figure 1.9 B). This study found that the N-terminal regions from each end of the extended tetramer could bind to DNA, and a model was proposed whereby SYCP3 fibres are formed that can bind to DNA between distant regions (Syrjanen *et al*., 2014). Further investigation into the SYCP3 DNA-binding reported that the binding of SYCP3 caused a compaction of DNA, and therefore supports the hypothesis that SYCP3 has a role in chromosomal compaction and organisation (Syrjanen *et al*., 2017).

As well as the role of the axial element proteins, the cohesin complex is required for the organisation of chromosomes in prophase I. The cohesin complexes are responsible for the maintenance of sister chromatid cohesion, by forming a ring around the sister chromatids

(Nasmyth and Haering, 2009). SYCP2 and SYCP3 have a proposed function for maintaining the integrity of the cohesin complex during late prophase, through sex-specific functions. SYCP2 and SYCP3 are recruited to centromeric regions in males, but not females. The deletion of SYCP3, and the subsequent absence of SYCP2, results in the premature dissociation of cohesin complexes during prophase I, observed in females only (Kouznetsova *et al*., 2005). These findings are supported by a previous study that identified interactions between the SMC

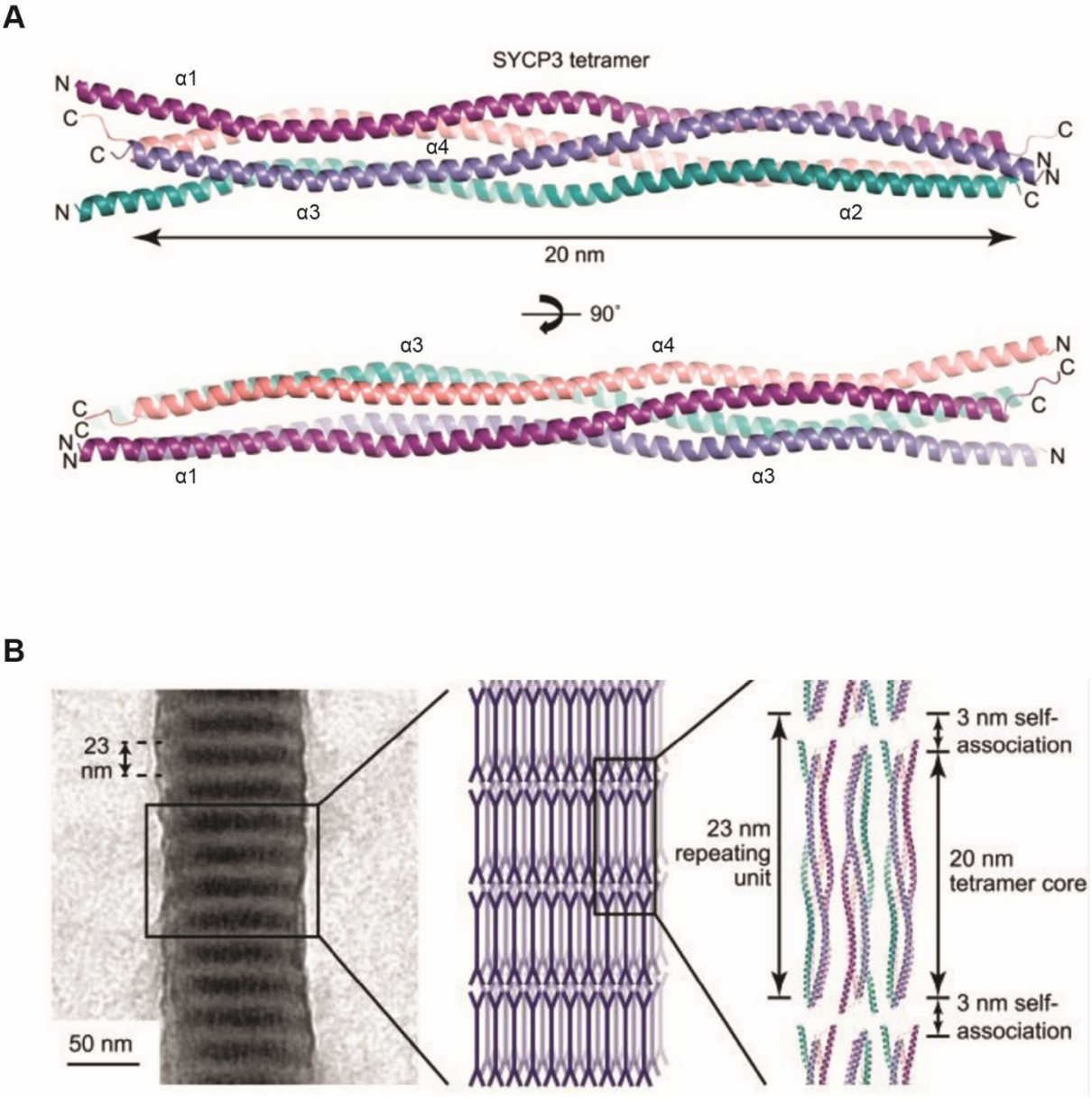


Figure 1.9. The structure and assembly of SYCP3. Images taken from Syrjanen *et al*., 2014. (A) The X-ray crystal structure of the truncated SYCP3 shown in two orientations. A central coiled-coil region and two C-terminal coiled-coil region are separated by four helical bundle. Alpha helix chains labelled as α1-4. PDB code: 4CPC. (B) Transmission electron micrograph of the full length SYCP3 self-assembly, with a model for the assembly based upon the crystal structure.

(Structural Maintenance of Chromosomes) components of cohesin and the axial components SYCP2 and SYCP3 (Eijpe *et al*., 2000). However, SYCP2 and SYCP3 are not required for cohesin stabilisation in metaphase I, and thus, suggests that the SYCP2 and SYCP3 axial proteins have a temporal role in cohesin integrity (Kouznetsova *et al*., 2005).

The repair of DSBs in meiosis differs to the normal mechanism in that the homologous chromosome is used as a template for repair, rather than the inter-sister repair that usually occurs. However, the precise mechanism to encourage inter-homologue repair is unclear, but is thought to be driven by the meiosis-specific DMC1 recombinase (Bishop *et al*., 1992). A study found that strand invasion by RAD51 was substantially suppressed by SYCP3, but does not affect DMC1 function, thus, proposing that homologous recombination by DMC1-mediated strand invasion is promoted by the SYCP3-inhibition of RAD51 (Kobayashi *et al*., 2017).

1.4.3 SYCP1 transverse filament assembly

The mammalian SC transverse filaments (TFs) are comprised of SYCP1, a long 946 molecule with a central helix and non-helical N- and C-termini, with the central helical region proposed to form coiled-coil structures and speculated to have evolved from a nuclear matrix protein through function specialisation (Meuwissen *et al*., 1992). The SYCP1 gene transcription only occurs between zygotene and diplotene, with preliminary data from immunogold labelling of the rat SC indicating the C-terminal region is positioned near the lateral element and the N-termini are within the central region (Meuwissen *et al*., 1992).

Further studies of the rat SYCP1 organisation were performed by raising antibodies to different SYCP1 fragments for immunoelectron microscopy experiments. The C-terminus was found to locate to the lateral element, as previously proposed, and more specifically to the inner half of the lateral elements (Figure 1.10 A). Also corresponding to the previous findings, the N-termini were located in the central region, where the N-termini from opposing lateral elements overlap so that the TFs span from the inner region of the lateral element to the respective far side of the central element (Schmekel *et al*., 1996). A study was published at the same time to show correlating findings in the mouse SC, with further analysis of the N-terminal region detecting self-interactions between N-terminal SYCP1 fragments (Liu *et al*., 1996). These findings suggest that SYCP1 molecules spanning from opposing lateral elements may interact to form rudimentary SC precursor structures.

The N-terminal interaction of SYCP1 provides further insight into the formation of SYCP1 polycomplexes. Polycomplexes are distinct structures from the SC, but still form tripartite assemblies that are only present in meiotic cells. These structures are most frequent following SC formation, with the occurrence of polycomplexes observed prior to pachytene in

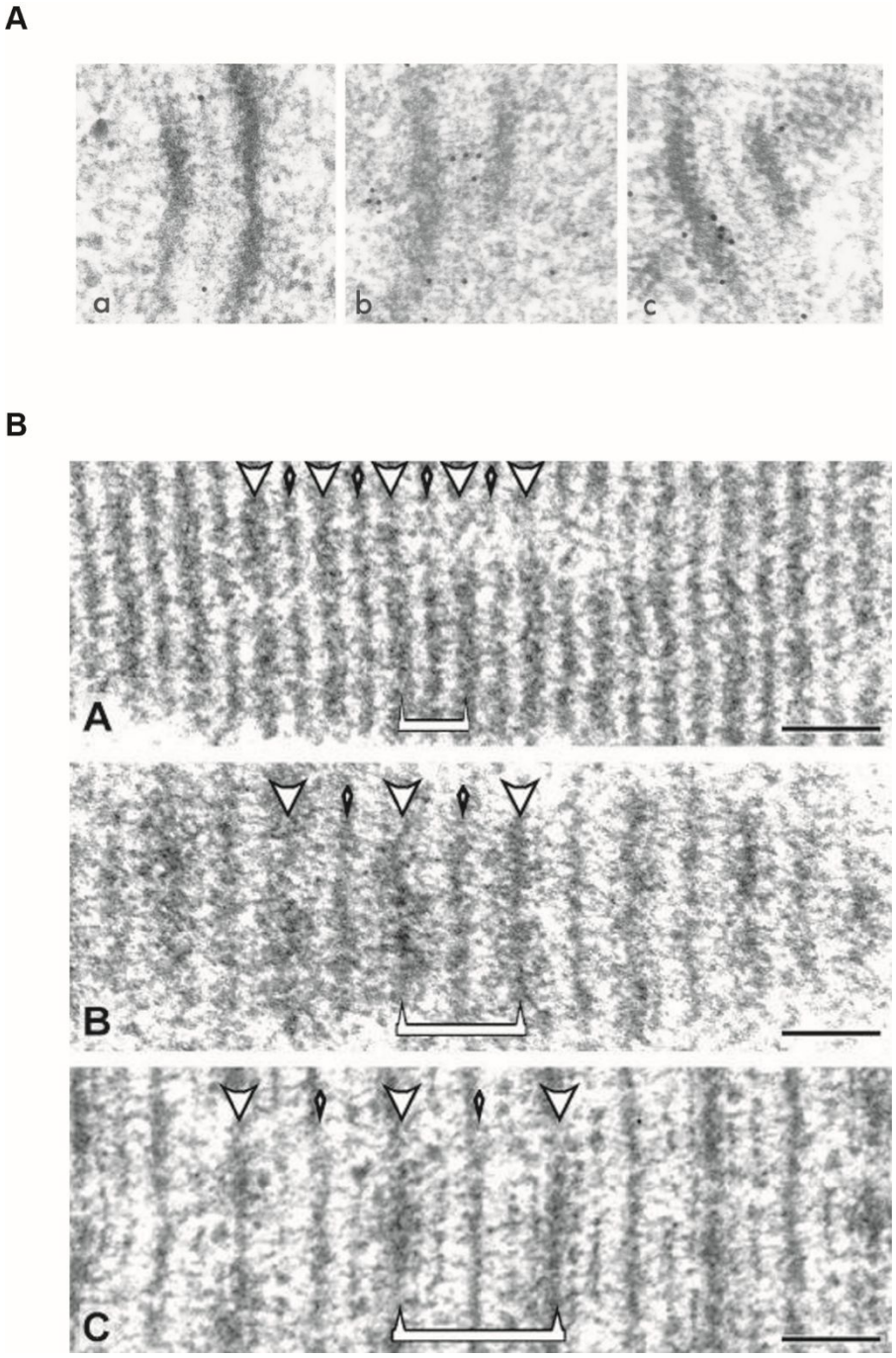


Figure 1.10. The organisation of SYCP1 in the SC and polycomplexes. (A) Immunoelectron microscopy of SCs from sections of rat testes in the frontal view. Immunogold labelling of SYCP1 fragments of the N-terminal region (a), a central region (b) and the C-terminal region (c). Image taken from Schmekel *et al*., 1996. (B) Immunoelectron microscopy of SYCP1 polycomplexes from COS-7 cells. The width of the polycomplexes are altered upn the length of the SYCP1 central region where a ~250 amino acid deletion forms narrower structures (A), and the insertion of ~250 amino acids cause wider structures to form (C) when compared to the wild type protein (B). Scale bars 0.1 μm. Image taken from Ollinger *et al*., 2005.

some organisms (Goldstein, 1987). The SYCP1 polycomplexes have been studied *in vitro* by expressing the rat SYCP1 protein in non-meiotic COS-7 cells. Polycomplexes are formed within the COS-7 cells, and deletion of the N- and C-terminal regions, respectively, disrupt the polycomplex assemblies. Further to this, the central helical domain of SYCP1 was shown to specify the length of the polycomplexes, with insertions or deletions within the central region resulting in an increase or decrease in polycomplex length (Figure 1.10 B), but maintaining the overall architecture (Ollinger *et al*., 2005). Therefore the SYCP1 polycomplexes require both the N- and C-termini for head-to-head and back-to-back self-assembly, with the central helical region providing the structural backbone, with similar mechanisms expected to occur during SYCP1 assembly in the SC.

The deletion of SYCP1 in mice causes infertility and a reduction in testis size. The axial elements are formed and aligned in the same manner as wild type cells. However, chromosome synapsis does not occur, which causes cell arrest in pachytene (de Vries *et al*., 2005). DSBs are found to be normal in leptotene, determined by monitoring the gamma H2AX DNA damage marker, but these foci are maintained throughout pachytene, whereas they are not observed at pachytene in wild type cells due to the repair of DSBs upon mature SC formation (de Vries *et al*., 2005).

Further analysis of the mouse SYCP1 TFs through immunoelectron microscopy demonstrates that the SYCP1 N-termini form a bilayered central element structure that is joined together by the central element proteins. The SYCP1 N-termini are separated by a 100 nm depth, with recruitment of SYCE3 to the bilayers and assembly of the other central element components between the diverged SYCP1-SYCE3 layers (Hernandez-Hernandez *et al*., 2016).

Characterisation of SYCP1 by electron microscopy has elucidated the general organisation of the SYCP1 TFs, however they provide minimal insight into the detailed molecular basis for SYCP1 structure and assembly. The X-ray crystal structure of a SYCP1 C-terminal fragment was solved and reported to form an anti-parallel dimer (Figure 1.11 A), with the interpretation that the SYCP1 chains may in fact be in an anti-parallel arrangement (Seo *et al*., 2016), despite the previous studies. Further clarification was provided through the solution of the X-ray crystal structure of a similar SYCP1 C-terminal fragment, where the same structure of the chains was determined, but formed an anti-parallel tetrameric assembly (Figure 1.8 B). Therefore a model was proposed whereby this anti-parallel four helical assembly occurs at the chromosome axis between lateral parallel dimers to forms a recursive C-terminal SYCP1 assembly mechanism (Dunce *et al*., 2018).

This mechanism was supported by the biochemical analysis of SYCP1 fragments. The region in the crystal structure forms a dimer in solution at neutral pH, but at acidic pH it forms

a tetramer, explaining the possible ambiguity between the crystal structures. As this region is suggested to be localised near the chromosome axis, it was postulated that this dimer-to-tetramer switch may occur due to local pH in the cell, and that it may bind DNA. The dimer (neutral pH) was not found to bind DNA but the tetramer (acidic pH) did bind DNA (Dunce *et al*., 2018). The very C-terminal region was found to bind DNA at neutral pH, providing a mechanism for SYCP1 recruitment to the chromosome axis through initial C-terminal DNA

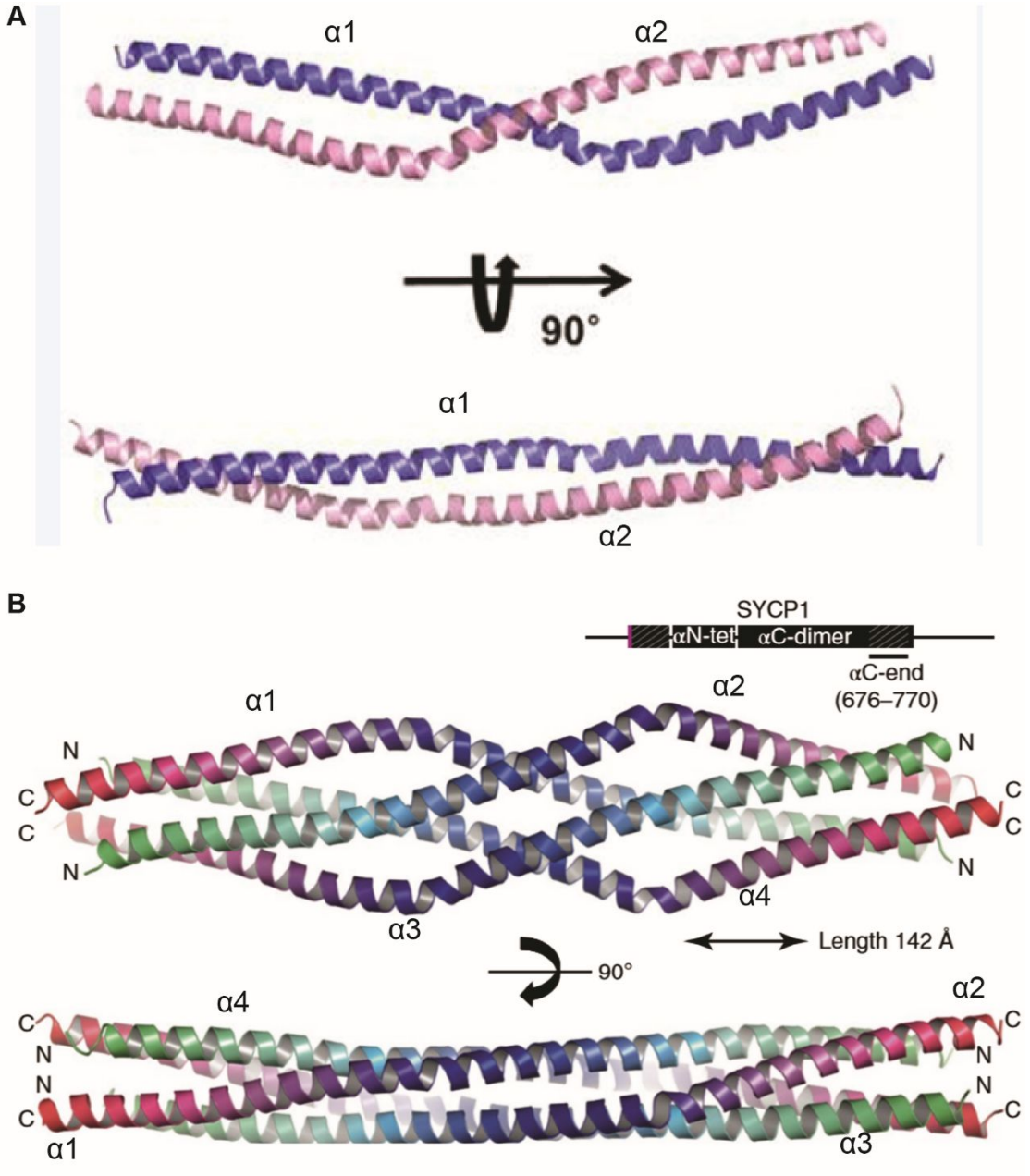


Figure 1.11. The structure of a C-terminal SYCP1 fragment. (A) The X-ray crystal structure of an SYCP1 fragment located at the C-terminus of the central helical region, displayed in two orientations. An anti-parallel dimer is formed with angulation in the SYCP1 chains, labelled α1 and α2, respectively. PDB code: 4YTO. Adapted from Seo *et al*., 2016. (B) The X-ray crystal structure of the same C-terminal SYCP1 region, displayed in two orientations. An anti-parallel tetramer is formed, with each end of the tetramer held together by a four-helix bundle. Alpha helices labelled as α1-4. PDB code: 6F63. Taken from Dunce *et al*., 2018.

interactions, followed by the formation of the back-to-back tetrameric assembly that can further bind and assemble along the chromosome axis. SYCP1 has been shown to bind to the lateral element protein SYCP2, with the proposal that SYCP2 is responsible for the recruitment of SYCP1 C-terminus to the chromosome axis (Winkel *et al*., 2009). However, upon deletion of SYCP3, and therefore loss of SYCP2 axial element localisation, SYCP1 was still detected to assemble at the chromosome axis, although at considerably reduced levels (Yuan *et al*., 2000). This indicates that although recruitment and assembly of the SYCP1 C-terminus is promoted through an interaction with SYCP2, self-assembly of SYCP1 can occur in its absence, which may provide a complementary mechanism for efficient SYCP1 SC assembly.

The SYCP1 N-termini have been shown to self-interact by co-immunoprecipitation studies from mouse testes (Liu *et al*., 1996). These findings were supported by a recent study of the human SYCP1 N-terminus by X-ray crystallography and biophysical characterisation. Crystal structures from an N-terminal region, with varying C-terminal lengths, were solved to reveal the formation of an elongated tetramer formed from a tetrameric head-to-head assembly between opposing dimers (Figure 1.12) (Dunce *et al*., 2018). The SYCP1 region immediately after the crystallised fragment was found to exist as a parallel tetramer through biophysical characterisation, which upon inclusion of the crystallised region, formed large molecular weight species (Dunce *et al*., 2018).

Therefore a model was proposed for the formation of a precursor SC lattice through SYCP1 self-assembly. The SYCP1 C-terminus is initially recruited to the axial elements via DNA-binding, promoted by SYCP2 interactions, followed by the back-to-back tetramer formation between proximate SYCP1 C-terminal dimers, as seen in the crystal structure. The lateral SYCP1 dimers then form parallel N-terminal tetramers in the central region with the N-termini splaying apart and able to form recursive head-to-head associations. This preliminary lattice is then proposed to form the basis for central element component recruitment and assembly for SC maturation (Dunce *et al*., 2018).

1.4.4 The central element initiation complex

The SC central element is thought to consist of two sub-complexes, the initiation complex and the elongation complex, due to the distinction of the respective deletion phenotypes. The first central element proteins to be discovered and characterised were SYCE1 and SYCE2 (previously referred to as CESC1) through profiling of mouse testis expression. These proteins were identified to only localise to the central element of the SC, with their recruitment dependant on the SYCP1 TF (Costa *et al*., 2005).

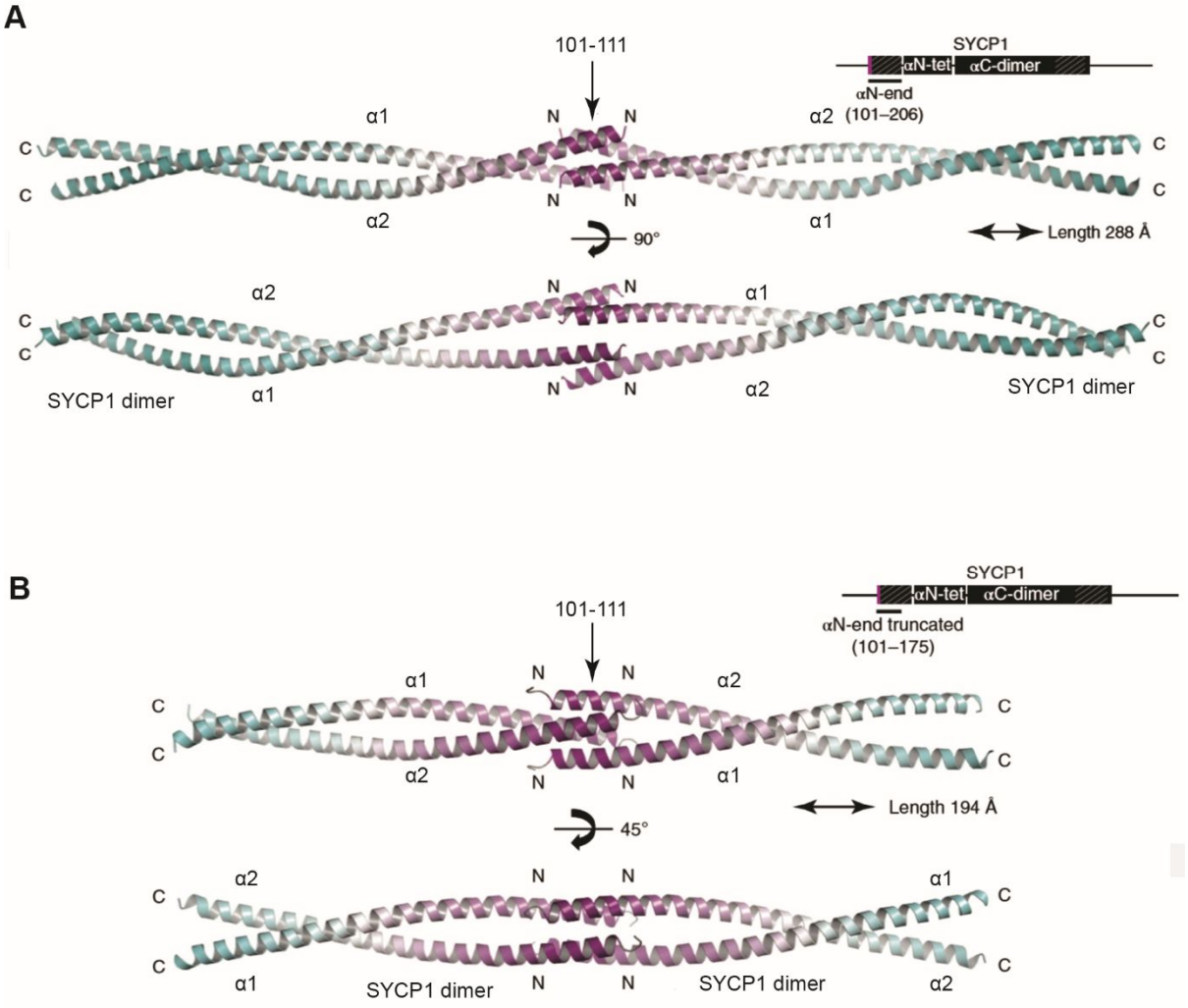


Figure 1.12. The structure of an N-terminal SYCP1 fragment. (A) The X-ray crystal structure of an SYCP1 fragment located at the N-terminus of the central helical region, displayed in two orientations. A head-to-head tetramer is formed from opposing parallel dimers. PDB code: 6F62. Adapted from Dunce *et al*., 2018. (B) The X-ray crystal structure of the same N-terminal SYCP1 region but with ~30 amino acid C-terminal truncation, displayed in two orientations. The same head-to-head tetramer is formed from opposing parallel dimers. PDB code: 6F5X. Adapted from Dunce *et al*., 2018.

The role of SYCE1 was investigated by another study, where a null SYCE1 mutation was characterised in mice. Both sexes were found to be infertile, with an inability to repair DSBs and lack of SC formation. However, alignment of homologous chromosomes was detected, along with the RAD51 and MSH4 recombination markers, but no MLH1 crossover marker foci were detected, indicating that the SC is required for the resolution of DSBs and

crossover formation (Bolcun-Filas *et al*., 2009). At sites of proximal homologue alignment in the SYCE1 null mice, SYCP1 still binds to the axial elements and undergoes limited assembly and at greater homologue distances, non-assembled SYCP1 is detected at the axial elements, which is not observed in wild type cells. These findings suggest that at the closely aligned regions, SYCP1 can self-assemble, but without SYCE1, SYCP1 accumulates at the axial elements and cannot organise into a full SC-like structure. Therefore, SYCE1 is suggested to play a role in guiding, and consequently restricting, the SYCP1 TFs to correctly assemble within the SC. The SYCE2 and TEX12 central element proteins were found to localise at the central element in the SYCE1 null mice, but no SC assembly was detected. SYCE1 was proposed to have a role in stabilising the SYCP1 N-termini within the central region by providing support for the depth of the central element (Bolcun-Filas *et al*., 2009). This correlates to the more recent study where the central region was found to form a bilayered structure, with diverged SYCP1 N-termini providing a 100 nm depth and assembly of central element proteins in between (Hernandez-Hernandez *et al*., 2016).

SYCE1 mutations have been detected and studied in patients from two separate consanguineous families. From genetic screening of a patient with primary ovarian insufficiency (POI), a homozygous non-sense mutation was detected within SYCE1. The parents and several siblings were found to harbour the heterozygous mutation, but were not affected, and one sibling and a group of 90 control participants did not have the mutation. Therefore, it was reasoned that the non-sense mutation caused malfunction of SYCE1 and therefore failure to form the SC, resulting in POI (de Vries *et al*., 2014). Another patient study elucidated a deleterious SYCE1 mutation, with the analysis of the patient testis tissue confirming the absence of SYCE1. The patient had non-obstructive azoospermia, and was consequently infertile (Maor-Sagie *et al*., 2015). These conditions are likely caused by the failure to form a functional SC, due to the absence of SYCE1, and the consequent inability to undergo chromosome synapsis, which leads to meiotic arrest and cell death (Geisinger and Benavente, 2016). These patient mutations provide a rare insight into the consequence of SC malfunction in humans, and corresponds to the findings from studies in model organisms.

Another central element component, SYCE3, was discovered in mice to be a small 88 amino acid protein, extremely well conserved among mammals and was found to localise to the central element. Characterisation of the SYCE3 deletion in mice demonstrated the inability to initiate synapsis, resulting in cell arrest and consequent infertility in both sexes. The initiation of recombination appears normal, but progression into crossovers is abolished according to the lack of MLH1 foci. SYCE3 was therefore found to be necessary for SC assembly at an earlier point than the other central element proteins, but at a later stage than SYCP1 (Schramm *et al*.,

2011). This correlates to the description of the bilayered central element, where SYCE3 is reported to bind to the SYCP1 layers in order for the remaining central element proteins to assemble (Hernandez-Hernandez *et al*., 2016).

In a later study, the recombinant murine SYCE3 protein was purified and the X-ray crystal structure was solved. The structure reveals the formation of a dimeric four helical bundle that is formed through the folding back of the SYCE3 chains in order to inter-lock in an anti-parallel arrangement (Figure 1.13). Analytical ultra-centrifugation was used to verify the dimer formation in solution, and cross-linking demonstrated capacity for multimeric assembly. The homo-dimerisation was clarified by co-immunoprecipitation in COS-7 cells, and further co-immunoprecipitation experiments revealed an interaction between the N-terminal helix of SYCE3 and the SYCE1 central element protein (Lu *et al*., 2014).

The most recent SC component to be discovered was characterised to be a central element protein, named C14ORF39/SIX6OS1, referred to as SIX6OS1. The protein was detected as an anonymous gene variant that affects recombination in humans. A direct interaction was identified between SIX6OS1 and SYCE1 through yeast 2-hybrid screening, but no interactions were detected between SIX6OS1 and any other SC component. A SIX6OS1

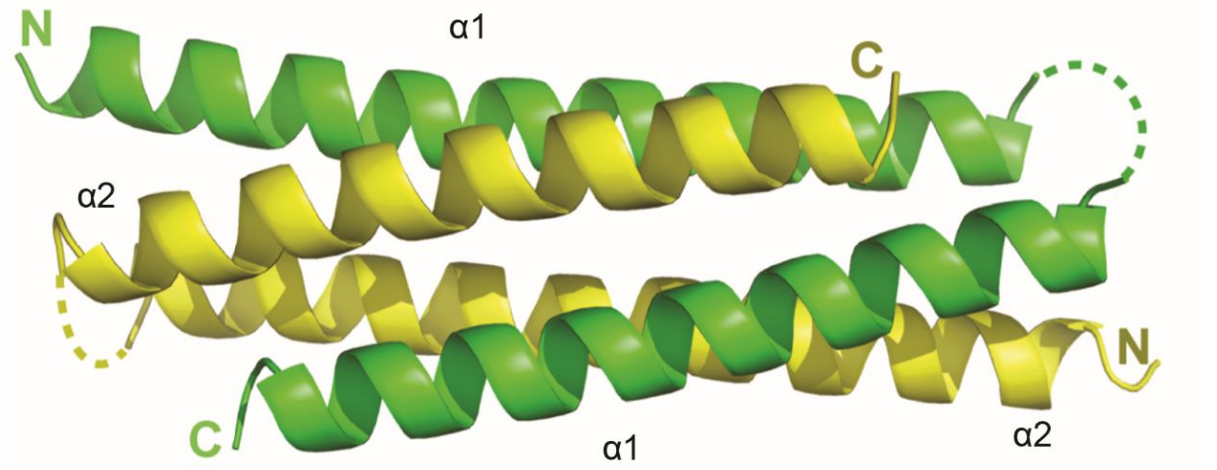


Figure 1.13. The structure of mouse SYCE3. The X-ray crystal structure of the mouse SYCE3 protein. A dimeric four helical bundle is formed through the interlocking of compact anti-parallel monomers. The dotted lines represent unbuilt regions. Alpha helices labelled as α1 (green) and α2 (yellow). PDB code: 4R3Q. Taken from Lu *et al*., 2014.

null mutation was studied in mice, and was found to cause infertility in both sexes, caused by meiotic arrest in prophase I (Gómez-H *et al*., 2016). This is consistent with the phenotypes observed for the other central element proteins. Therefore, SIX6OS1 is speculated to be recruited to the central element by SYCE1, and may form the central element initiation complex, along with SYCE1 and SYCE3.

1.4.5 Central element elongation by the SYCE2-TEX12 complex

As described above, SYCE2 (initially named CESC1) was first discovered in 2005, along with SYCE1, as the only known central element proteins (Costa *et al*., 2005). A subsequent study examined the effect of mouse synapsis upon SYCE2 deletion. Consistent with the phenotypes of the other central element proteins, both sexes were infertile. The chromosomes were aligned, and the axial elements were formed, but chromosomes failed to synapsis (Figure 1.14 A). The induction of DSBs appeared normal, but there was no detection of crossover formation (Bolcun-Filas *et al*., 2007). However, a slightly different phenotype was observed compared to the other central element proteins, in that short regions of synapsis were detected, but these stretches failed to extend into a mature SC (Figure 1.14 A). This finding provides an indication of the role of SYCE2 in SC extension (Bolcun-Filas *et al*., 2007).

TEX12 was characterised as a component of the central element shortly after SYCE2, upon discovery of its meiosis-specific expression and localisation to the central elements in mice (Hamer *et al*., 2006). TEX12 was found to precisely co-localise with SYCE2 in the central element (Figure 1.14 B) and an interaction between SYCE2 and TEX12 was verified through co-immunoprecipitation (Hamer *et al*., 2006). A following study investigated the effects of mouse synapsis upon TEX12 inactivation, with both sexes found to be infertile in the absence of TEX12. As described for the other central element components, DSB formation appeared normal, but there was no indication of crossover formation (Hamer *et al*., 2008). Further to this, the phenotype of TEX12 absence matched the phenotype of the SYCE2 deletion, whereby short stretches of synapsis initiation are observed (Figure 1.14 C) but these sites cannot form an extended SC (Hamer *et al*., 2008).

SYCE2 and TEX12 have also been shown to exist in ancient metazoan species, along with SYCP1 and SYCP3. SYCE1, SYCE2 and TEX12 were found to exist across distant lineages, with further branching of SYCE2 and TEX12 to diverse species (Fraune *et al*., 2013). To examine whether SYCE2 and TEX12 functioned as SC proteins in these diverse species, the *Hydra vulgaris* species was used to probe for SYCE2 and TEX12. Both mRNA transcripts and protein expression were detected for SYCE2 and TEX12. Both proteins were identified within spermatocytes and detected during the pachytene stage along with co-localisation to SYCP1.

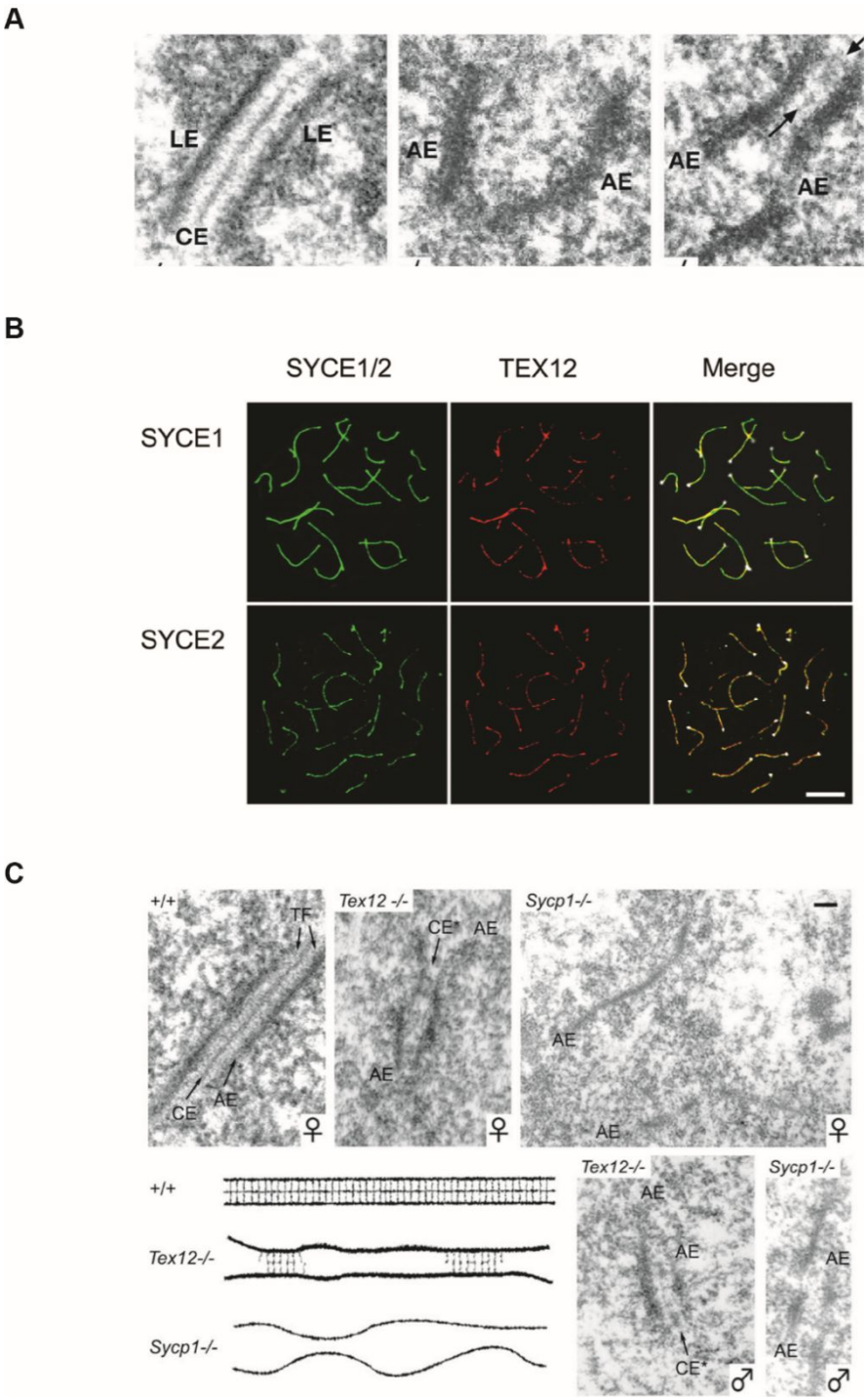


Figure 1.14. The co-localisation and phenotypes of SYCE2 and TEX12. (A) Immunoelectron microscopy of SCs and axial elements from mouse testes spreads in the frontal view. The left image is from wild-type cells and the centre and right hand side images are from SYCE2 deletion mice. A fully formed SC is observed in wild-type, with axial element formed in the mutant with short regions of synapsis indicated by arrows. Image taken from Bolcun-Filas *et al*., 2007. (B) Spermatocyte pachytene spreads and staining of SYCE1 or SYCE2 (green) and TEX12 (red). TEX12 specifically co-localises with SYCE2. Scale bars 5 µm. Image taken from Hamer *et al*., 2006. (C) Immunoelectron microscopy of SCs and axial elements from mouse testes spreads in the frontal view for comparison of wild type to the TEX12 and SYCP1 deletions in both males and females. Panels indicated by the schematic key in the bottom left corner. Axial elements and short regions of synapsis are identified in the TEX12 deletion. Image taken from Hamer *et al*., 2008.

This study indicates that ancient metazoan species form a similar SC-like structure to mammals, with SC formation achieved by SYCP1, SYCE2 and TEX12 in the central region, suggesting that the other central element components have evolved more recently than SYCE2-TEX12 to account for increased complexity within mammals (Fraune *et al*., 2013). This also supports the indications of both SYCP1 preliminary lattice formation and elongation of the SC by SYCE2-TEX12 as these essential functions were performed in the apparent absence of the other central element components in *Hydra vulgaris*.

These findings indicate that both SYCE2 and TEX12 have critical roles for the extension of the central element to span the length of the homologues, with their specific co-localisation and the identification of a SYCE2-TEX12 interaction suggesting that SYCE2-TEX12 form the central element elongation complex in order to achieve SC maturation.

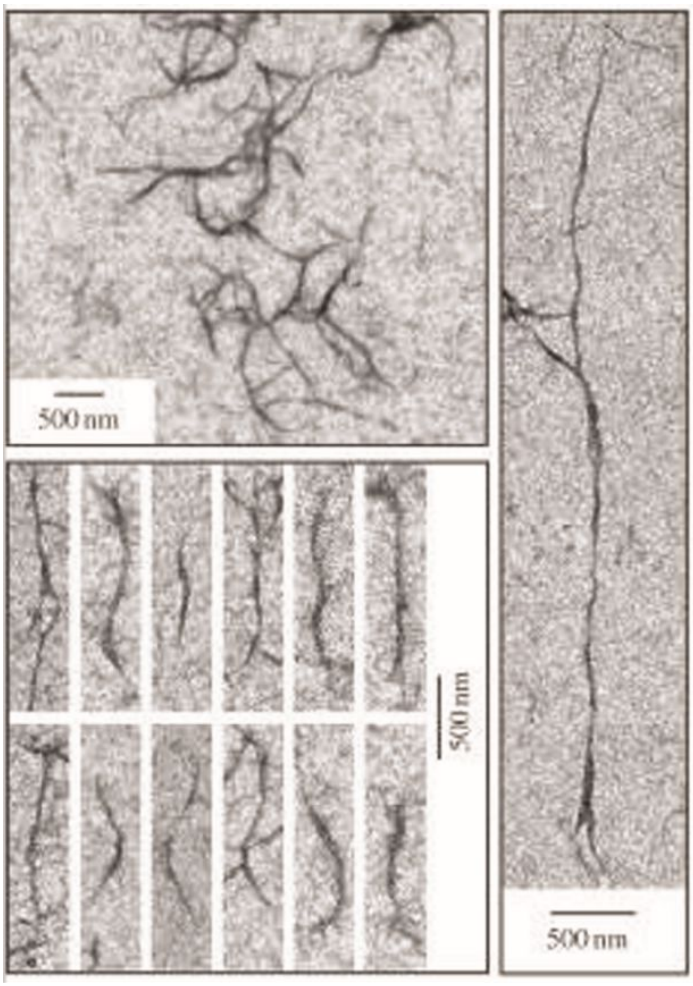


Figure 1.15. Filament formation of the SYCE2-TEX12 complex. Transmission electron micrographs showing various filaments formed by SYCE2-TEX12.Taken from Davies *et al*., 2012.

The proposed SYCE2-TEX12 elongation complex was initially characterised *in vitro* through co-expression and co-purification of the helical core regions of SYCE2 and TEX12. The recombinant SYCE2-TEX12 complex was determined to be a highly stable and constitutive complex, with oligomeric analysis revealing the formation of an equimolar hetero-octameric complex (Davies *et al*., 2012). The SYCE2-TEX12 complex was also found to undergo higher-order assembly in order to produce long filamentous structures (Figure 1.15) that resemble the dense central elements observed in intact SCs (Davies *et al*., 2012). These findings support the hypothesis of SYCE2-TEX12 forming the central element elongation complex, through the higher-order assembly of the 4:4 units to form filaments.

1.5 The yeast synaptonemal complex

1.5.1 The budding yeast synaptonemal complex

The synaptonemal complex in *Saccharomyces cerevisiae* has the same tripartite architecture as the mammalian SC, with lateral elements at the chromosome axes, initially formed by the precursor axial elements (Rockmill and Roeder, 1988, Hollingsworth and Byers, 1989), and a dense central element in the midline. Transverse filaments then interconnect the respective lateral elements though tethering to the central element in order to form the fully assembled "zipper-like" structure to synapse homologous chromosomes for the repair of DSBs and the formation of crossovers (Sym *et al*., 1993).

The components of the yeast SC differ to the individual components of the mammalian SC, but the overall architecture and function remains conserved. Meiosis in yeast has been extensively studied, yet the molecular basis of the protein choreography that takes place to form the ornate and vital SC remains elusive. Therefore, the study of the structures and assembly that occur with the yeast SC components may elucidate the molecular basis for its architecture and may provide further insights into its function. Due to the genetic amenability of yeast, in comparison to mammalian models, relation between yeast SC and mammalian SC assembly mechanisms may allow clarification of *in vitro* characterisation of the human proteins, where the complementary *in vivo* studies are more challenging to achieve.

1.5.2 The yeast lateral elements

Lateral elements are the functional components of the SC at the chromosome axis, but coordinate with a wide range of chromosome axis proteins in order to form the precursor axial elements. The requirement for cohesin loading in budding yeast has been identified in order to form the axial elements. Smc3 and Rec8 are necessary for sister chromatid cohesion, and this allows for axial element formation (Klein *et al*., 1999). The synapsis initiation complex (SIC)

also plays a crucial role at the chromosome axis. The SIC, also known as the ZMM complex, is formed by seven proteins: Zip1, Zip2, Zip3, Zip4, Mer3, Msh4 and Msh5. Zip1 is the SC transverse filament protein, Zip2-4 are thought to govern protein-protein interactions within the axial elements and Mer3, Msh4 and Msh5 promote recombination (Lynn *et al.*, 2007). Zip3 has been shown to co-localise with Zip2, with the deletion of Zip3 resulting in the failure to form the SC, suggesting that Zip3 is a component of both the SIC and the SC lateral elements (Agarwal and Roeder, 2000).

The other lateral element components are Red1 and Hop1 (Figure 1.16). Red1 was identified through mutational screening for viable spores that fail to undergo meiosis. A Red1 mutant was characterised due to its ability to form spores, and analysis of the spores resulted in the detection of a high rate of aneuploidy. The Red1 protein was therefore initially proposed to play a role in homologue pairing during metaphase I (Rockmill and Roeder, 1988). A later study characterising Red1 mutants demonstrated that no SC or precursor structures are formed in the absence of Red1. The functional Red1 was identified to localise to chromosomes during pachytene, specifically to the chromosome cores and not the loops. The recruitment of Red1 to axial elements was maintained in the absence of Zip1, thereby indicating that Red1 is a component of the SC (Smith and Roeder, 1997).

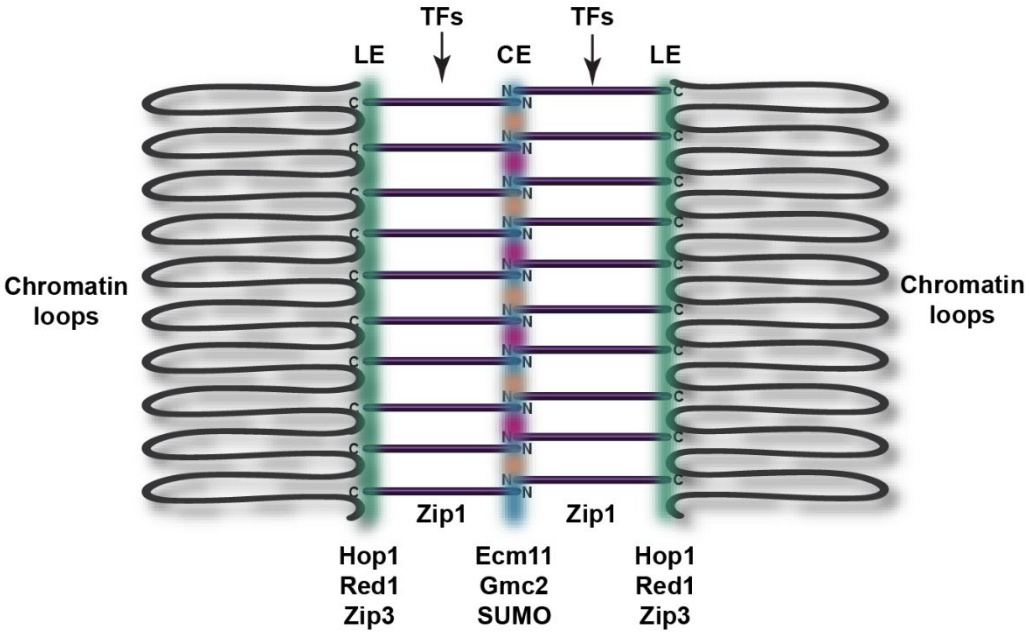


Figure 1.16. Schematic of the yeast synaptonemal complex. LE = lateral element. TFs = transverse filaments. CE = central element. Schematic depicting the chromatin loops of the homologous chromosomes with the assembly of the lateral elements at the axes (green). The transverse filaments (purple tubes) bridge between the chromosome axis and the central element (orange, blue and pink). The lateral elements are comprised of Hop1, Red1 and Zip3. Zip1 forms the transverse filaments and the central element is formed by Ecm11, Gmc2 and SUMO.

Hop1 was characterised through mutational studies searching for mutants that fail to pair homologous chromosomes. The Hop1 mutant was found to have decreased recombination activity and a reduction in crossover frequency (Hollingsworth and Byers, 1989). Hop1 was predicted to function within a multimeric assembly (Friedman *et al*., 1994) and then identified to localise with Red1, with its recruitment to axial elements abolished upon Red1 deletion (Smith and Roeder, 1997).

The assembly of Red1 and Hop1 axial element proteins has been shown to be controlled by Mek1 phosphorylation. Yeast 2-hybrid experiments revealed Red1 self-interactions and also a direct Red1 interaction with Hop1 as predicted. The interaction between Hop1 and Red1, and their recruitment to axial elements, were proposed to be controlled by the putative Mek1 kinase (Hollingsworth and Ponte, 1997a). This was supported by a subsequent study, where Mek1 activity was identified *in vitro*, and Red1 was found to be phosphorylated in a Mek1-dependant manner. Co-immunoprecipitation of Hop1 and Red1 was then found to be enhanced in the presence of Mek1. Therefore, Mek1 was suggested to coordinate the oligomerisation of Hop1 and Red1 for their assembly in the axial elements (de los Santos and Hollingsworth, 1999). The Red1-Hop1 interaction was identified *in vitro* through the purification of full length Red1. Both Red1 and Hop1 have been shown to bind DNA, with the DNA binding capacity of Red1 specified to preferentially bind to 3 or 4-way structures, such as Holliday Junctions. This demonstrates that Red1 may have a direct role in the stabilisation of recombination intermediates (Kshirsagar *et al*., 2017).

1.5.3 Zip1 transverse filament polymerisation

Zip1 was found to be a novel synaptonemal complex component in 1993, where Zip1 null meiotic yeast cells were able to undergo pairing of homologous chromosomes, but exhibited a failure for chromosome synapsis. Therefore, the cells were arrested before the first meiotic division. The Zip1 protein was estimated to be a coiled-coil protein, and was found to localise to synapsed chromosomes, but was not identified at the axial elements, and it was consequently proposed to be the transverse filament in the yeast SC (Sym *et al*., 1993). Following this study, Zip1 mutation did not drastically alter recombination or chromatid cohesion, but the process of crossover interference was severely affected (Sym and Roeder, 1994). To provide further insight into the role of Zip1 in recombination, the meiotic recombinase DMC1 was found to co-localise with Zip1 in recombination modules (Bishop *et al*., 1992).

As Zip1 was suggested to be the transverse filament protein, it was thought that the central coiled-coil region would provide the width between the lateral elements. A study to

investigate this created mutants through insertion or deletion of sequences within this region. These experiments revealed that the width of the SC was altered upon the change in length in Zip1. This was seen for both assembled SC structures and also for polycomplexes (Sym and Roeder, 1995). The polycomplexes have been observed in meiotically dividing organisms, including mammals, and are postulated to be "mini-SC" structures that may act as a storage form for the transverse filament proteins (Goldstein, 1987).

To elucidate the organisation of the Zip1 transverse filaments, several studies sought to decipher the orientation of Zip1 within the SC by immunoelectron microscopy. Through deletion of a C-terminal Zip1 fragment, the Zip1 transverse filaments were no longer recruited to the axial elements, therefore indicating that the C-terminus is required for Zip1 localisation. The deletion of the C-terminal region of the central coiled-coil domain enabled Zip1 localisation, but caused a decrease in SC width, consistent with the previous study (Tung and Roeder, 1998).

Further localisation experiments were performed by immunogold labelling of the Zip1 N- and C-termini. The Zip1 N-termini were found to be located within the central region of the SC, with the C-termini localised at both lateral elements (Dong and Roeder, 2000). These findings suggest that the Zip1 C-termini are recruited to the axial elements, and the transverse filament N-termini stemming from the respective lateral elements meet in the central region to provide the physical tether between homologous chromosomes (Figure 1.17).

A conserved region in the Zip1 C-terminus was found to be phosphorylated, with the Zip1 phosphorylation found to be required for the crossover pathway by the ZMM proteins. These phosphorylation events are promoted by the formation of DNA DSBs, and are found to occur through Mek1 kinase activity. The inhibition of these phosphorylation events results in meiotic arrest and was particularly observed in cells where the non-crossover promotor Sgs1 helicase is absent, as neither repair process occurs. The phenotype could rescued in a strain harbouring the deletion of Spo11, therefore no DNA DSBs are created. This validates both the important role of Spo11 in DSB formation and also demonstrates that the inhibition of Zip1 phosphorylation prevents CO resolution. The phosphorylation of Zip1 was detected at an earlier stage than the ZMM pathway, suggesting that the Zip1 C-terminal phosphorylation directs the ZMM pathway for CO formation (Chen *et al*., 2015).

1.5.4 The yeast central element

The central element of the yeast SC is comprised of Ecm11, Gmc2 and SUMO with the integration of the Zip1 N-terminus. The deletion of Ecm11 in yeast cells caused abrogation of recombination and segregation, and was found to be necessary for crossover formations.

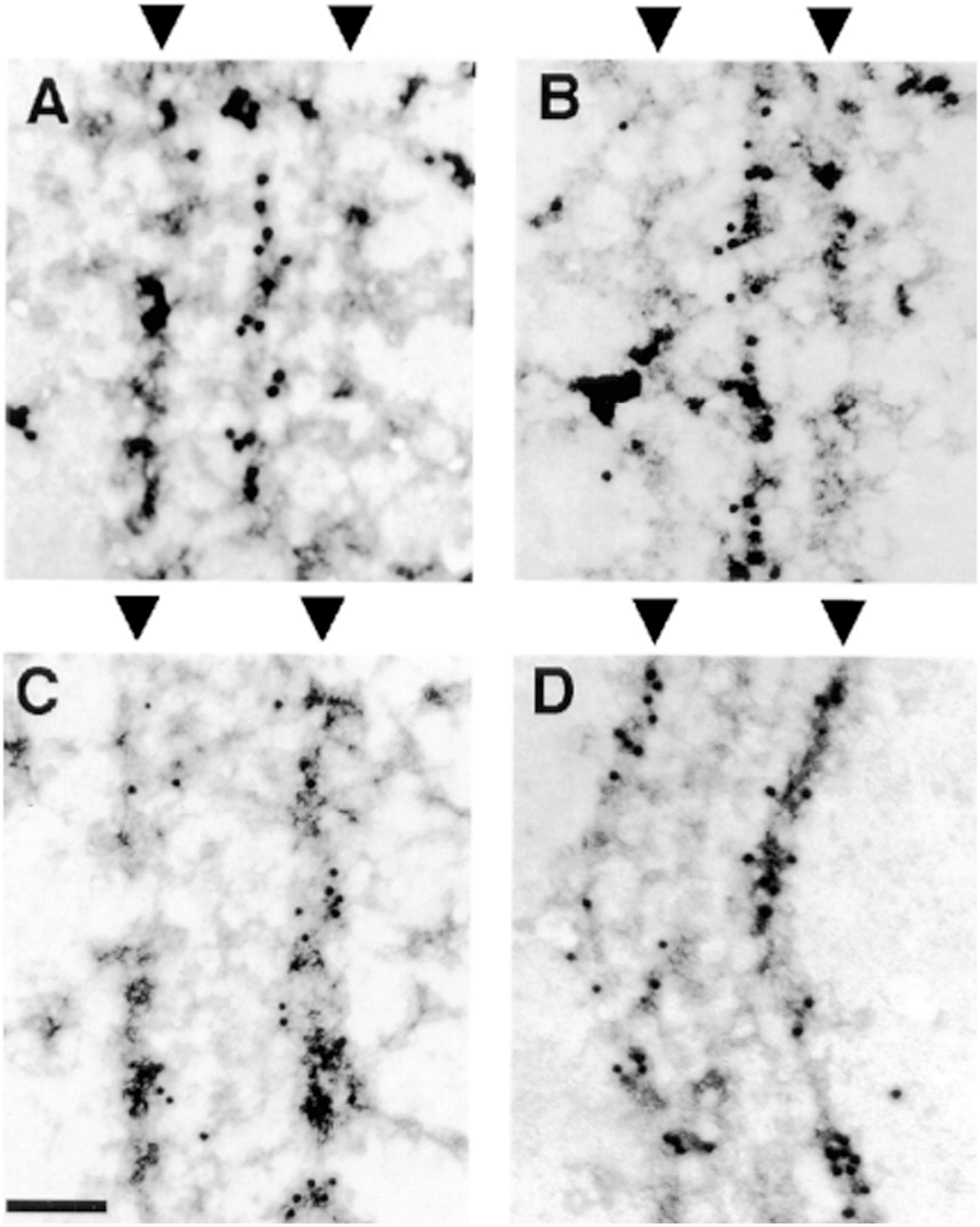


Figure 1.17. The orientation of Zip1 within the yeast SC. Immunoelectron microscopy of SCs in the frontal view, arrows indicate the lateral elements. Zip1 N-termini are labelled in A and B, Zip1 C-termini are labelled in C and D. The Zip1 N-termini localise to the central region of the SC and the C-termini localise to the lateral elements. Scale bar 100 nm. Image taken from Dong *et al*., 2000.

Therefore, it was postulated to be an early meiotic protein with a contributory role for the formation of crossovers (Zavec *et al*., 2004). In a subsequent study, Ecm11 was found to be SUMOylated during meiosis, and these SUMO modifications were found to be required for Ecm11 function in meiosis, and therefore for faithful meiotic progression. The K5 residue in the Ecm11 N-terminus was isolated as the site where SUMO modifications occur (Zavec *et al*., 2008).

Ecm11 was later found to form a complex with Gmc2, with the Ecm11-Gmc2 complex proposed to act as the central element in yeast. The Ecm11-Gmc2 complex is required for the incorporation of Zip1 transverse filaments into the central region. The assembly of Ecm11, Gmc2 and the Zip1 N-terminus within the central element is a recursive process driven by SUMOylation events and a positive feedback loop (Leung *et al*., 2015). Ecm11 and Gmc2 are initially recruited synapsis initiation sites and propagate from these sites upon SC maturation. The previously reported Ecm11 SUMOylation occurs at the K5 site and also at another identified site at the K105 residue. Upon deletion of Ecm11 or Gmc2, Zip1 polymerisation is greatly impaired, both in the context of polycomplex formation and SC incorporation. Upon mutation of the Ecm11 K5 and K105 residues, abolishing Ecm11 SUMOylation, Zip1 formed polycomplexes but was not incorporated into the SC (Humphryes *et al*., 2013).

The Ecm11 SUMOylation events were only found to occur in the presence of Gmc2, suggesting that Ecm11-Gmc2 complex formation may be required for stabilisation of Ecm11. These findings lead to the model whereby the yeast central element is assembled in a recursive manner prompted by the Ecm11-Gmc2 recruitment to sites of synapsis initiation. The Ecm11 SUMO modifications then trigger Zip1 assembly into the SC, with Zip1 SC incorporation further encouraging Ecm11 SUMOylation, thereby creating a positive feedback loop for central element assembly along the entire length of homologous chromosomes to achieve complete synapsis (Humphryes *et al*., 2013).

1.6 Cell cycle regulation and signalling

1.6.1 Regulation of the cell cycle

A major regulatory mechanism to control the cell cycle is the activity of cyclin-dependant kinases (CDKs). These protein kinases require binding of a cyclin regulatory partner in order to activate the kinase to phosphorylate serine and threonine residues of its target substrate (Malumbres, 2014). The CDKs are able to play a crucial role in cell cycle control due to their activation by their specific cyclin activator. The timely transcription of the cyclin binding partners allows for the activation of the necessary CDK at the correct point during the cell cycle and conversely, the targeted degradation of cyclin inactivates the CDK at the correct

time during the cell cycle. Therefore, the CDK is temporarily activated at a specific point during the cell cycle in order to phosphorylate its target substrates (Malumbres, 2014). This mechanism highlights the complex nature of cell control as one of the major regulatory mechanisms, CDK phosphorylation, is itself under tight control.

An example of CDK regulation during meiosis is the role of CDK activity in forming microtubule attachments to kinetochores in meiosis I. Studies in mouse oocytes have shown that a gradual increase in CDK1 activity, regulated by cyclin B, allows for the correct segregation of chromosomes (Davydenko *et al*., 2013). This is achieved by causing a delay in the time it takes to stabilise kinetochore-microtubule attachments as CDK1 has many substrates within the kinetochore that require phosphorylation in order to attach to microtubules. The delay is caused by the timely and gradual expression of cyclin B, which in turn causes the gradual activity of CDK1. This allows for the correct kinetochore-microtubule interactions to form before the cell cycle progresses (Davydenko *et al*., 2013).

Another key regulatory mechanism in the cell cycle is the activity of polo-like kinases (PLKs). There are four PLK proteins, PLK1-4, and all four are expressed upon initiation of spermatogenesis. However, only PLK1 is localised to the SC where it has been shown to phosphorylate SC components SYCP1, SYCE1 and TEX12 (Jordan *et al*., 2012), although the specific phosphorylation sites are not known. Inhibition of PLK1, and consequent inhibition of SC component phosphorylation, resulted in a failure of SC components to disassemble and therefore, prevented meiotic exit (Jordan *et al*., 2012). This suggests that phosphorylation events of the SC plays a role in triggering SC disassembly and that PLK1 is likely responsible for the phosphorylation events.

1.6.2 Cell cycle signalling mechanisms

As well as modifications being required for activity and function of proteins during the cell cycle, cell signalling pathways are also important in order to direct the activities and timings of cellular events. Phosphorylation, as described by CDK and PLK activities above, has a central role in dictating cell cycle progression. Another important signalling mechanism is the role of the small ubiquitin-like modifier (SUMO) and SUMOylation of substrates during meiosis has been shown to have implications in DSB repair and SC assembly (Nottke *et al*., 2017).

The SUMO pathway involves activation of SUMO by a SUMO protease that removes the C-terminal peptide in order to expose a di-glycine motif. This motif then covalently attaches to the E1 activating enzyme by a reactive cysteine. The SUMO is then transferred to the E2 conjugating enzyme and in turn transferred to the E3 ligase. The E3 ligase then transfers the

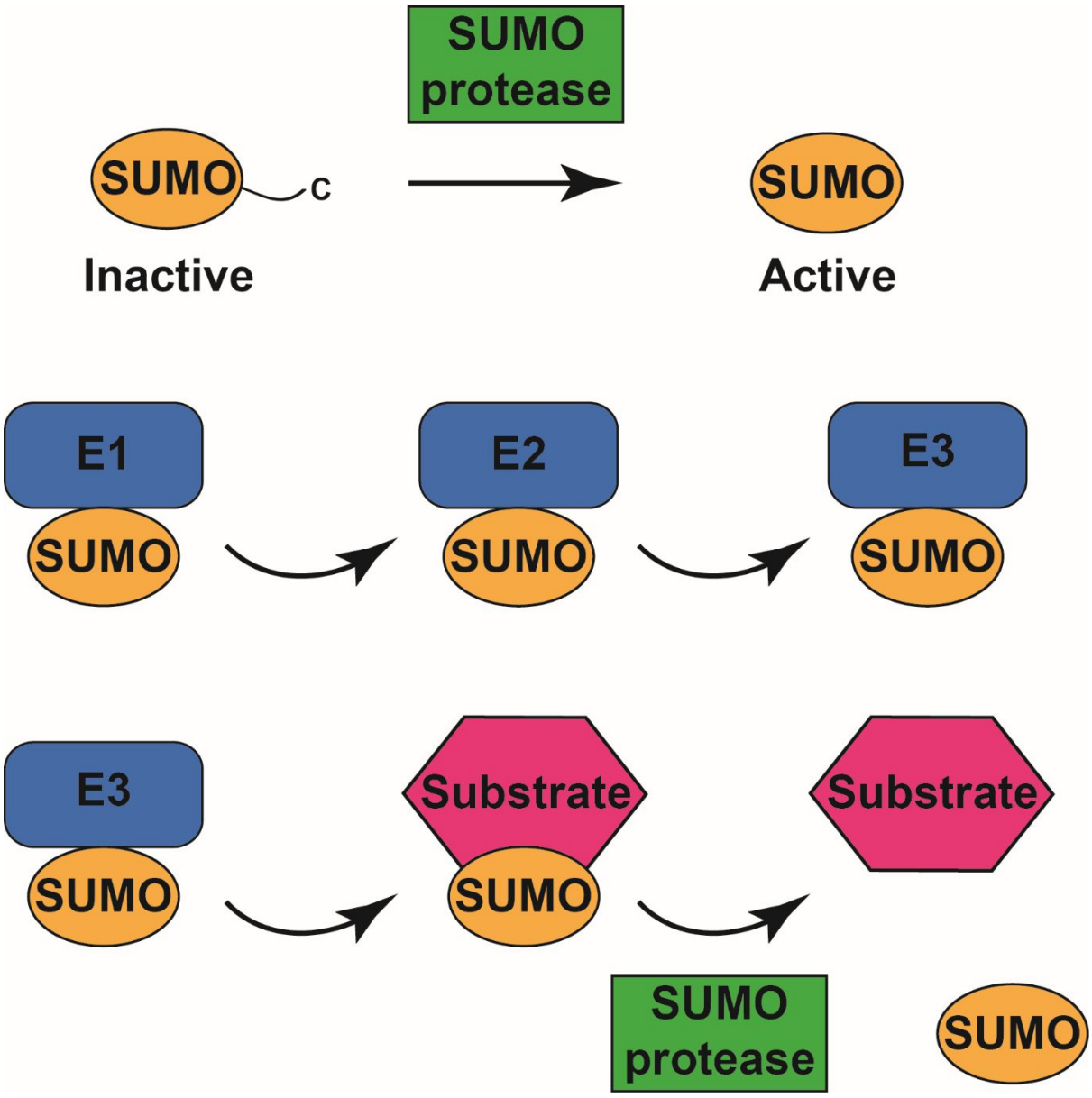


Figure 1.18. SUMO pathway. The SUMO protein is activated by the removal of a C-terminal peptide, by a SUMO protease, in order to expose a di-glycine motif. The activated SUMO then binds, via the di-glycine motif, to the reactive cysteine of the E1 activating enzyme. SUMO is then transferred to the E2 conjugating enzyme and then to an E3 ligase. SUMO is then transferred to the lysine residue on the target substrate. The SUMOylation event can be reversed by removal of SUMO from the substrate by the SUMO protease that initially activates SUMO.

SUMO to the target lysine residue of the substrate (Figure 1.18). This process is reversible and the SUMO medication can be removed by the SUMO protease that initially activates SUMO (Gareau and Lima, 2010).

SUMO has been shown to modify several proteins involved in meiosis, but the specific implications of these events are not entirely clear. In yeast meiosis, two components of the SC have been shown to be SUMOylated and these SUMO modifications are required for SC assembly. The Zip1 transverse filament relies on SUMOylation in order to be recruited to the central region and then a poly-SUMO feedback mechanism of Ecm11 and Zip1 SUMOylation is responsible for polymerisation of the SC (Voelkel-Meiman *et al*., 2013). The lateral element protein Zip3 is known to be an E3 ligase and so may play a role in SUMOylation of the central region (Agarwal and Roeder, 2000). These SUMO events appear to drive SC assembly in yeast, and the reversible nature of the SUMO modifications provides a potential mechanism for SC disassembly.

1.6.3 Chromosome alignment in cell division

The process of cell division involves extensive reorganisation of the genetic material within the cell. In order to control these complicated events, several protein complexes function to support chromosome organisation to ensure the events are correctly performed during cell division and to prevent aberrations such as aneuploidy or infertility.

The structural maintenance of chromosomes (SMC) family of ATPase proteins consist of six SMC proteins that play an integral role in chromosome organisation (Harvey *et al*., 2002). The overall architecture of the SMC proteins are the same, where by N-terminal domain leads to a coiled-coil domain and then to a hinge region in the middle. The hinge region then leads to another coiled coil domain and finally the C-terminal domain. This structure means that each SMC molecule can form inter-molecule anti-parallel coiled coils facilitated by the hinge domain, forming a characteristic V-shape upon SMC dimerisation (Harvey *et al*., 2002).

SMC proteins are known to be constituents of several important complexes, such as condensin and cohesin (Lee, 2013). Condensin is a protein complex that functions to compact DNA during mitosis and meiosis. Condensin has five subunits: SMC2, SMC4, kleisin and two HEAT (Huntingtin, elongation factor 3, protein phosphatase 2A, Tor1 kinase) subunits. SMC2 and SMC4 dimerise to form the V-shaped structure with the coiled coil homo-dimers joining at the hinge regions. The arms of the V-shape provide distance between the hinge region and the opposing end where the kleisin subunit binds, which in turn recruits the HEAT subunits (Lee, 2013). The condensin complex is thought to compact DNA and to produce long chromosome loops that are observed during cell division. The proposed model was that the condensin

complex could clamp DNA between the SMC coiled coils. DNA binding sights have been identified in the kleisin and HEAT subunits, with DNA-protein crystal structures revealing a DNA groove in the kleisin subunit (Kschonsak *et al*., 2017).

The cohesin is another vital complex that functions to hold together sister chromatids during mitosis and meiosis until they are segregated. In female meiosis, this involves holding together sister chromatids during meiotic arrest in prophase I, which can last for decades, and then sustaining the chromatid cohesion during the first meiotic division until the sister chromatids are segregated during meiosis II (Toth and Jessberger, 2016). Due to the long time span that cohesin is required in oogenesis, cohesin malfunction is often linked to female age-related infertility.

Cohesin is comprised of SMC1, SMC3, a kleisin subunit and a HEAT-containing subunit (Lee, 2013). The SMC units form the long v-shaped structure, as in condensin, and the complex is postulated to clamp around sister chromatids at repeating sites in order to provide chromatid cohesion. The exact mechanism for chromatid cohesion is not known, but several models exist. The most common model is that one cohesin ring encapsulates both sister chromatids. However, other models propose that two respective cohesin complexes clamp separate chromatids and then the two cohesin complexes interact in order to tether the sister chromatids (Brooker and Berkowitz, 2014). Cohesins are also implicated in DSB repair. The cohesin complexes have been found to be required for DSB repair in the G2 phase as the cohesin complex tethers a sister chromatid to a damaged chromatid in order for the DSB to repaired by using the sister chromatid as a DNA template (Brooker and Berkowitz, 2014).

1.7 Coiled-coil structures

1.7.1 The properties of coiled-coil proteins

Coiled-coils are alpha helical structures formed by coiling of respective helices. An alpha helix has a periodical turn that is formed by 3.6 residues, and therefore the periodicity of alpha helices forms a gradual drift in side chain positioning. However, due to the geometry of coiled-coils, where two or more helices inter-twine, the periodicity is slightly reduced so that 3.5 amino acids take part in one turn of the helix. This results in a repeating unit for every two turns, where the side chains have a consistent position in relation to the helix. This pattern is referred to as a "heptad" and is the basis of coiled-coil structures (Crick, 1952).

The heptad pattern, where the first to seventh amino acid residues in the heptad are termed the *a*-*g* positions, is formed with the *a* and *d* positions satisfied by hydrophobic residues, with some heptads containing charged residues at the *e* and *g* positions (Figure 1.19). Due to

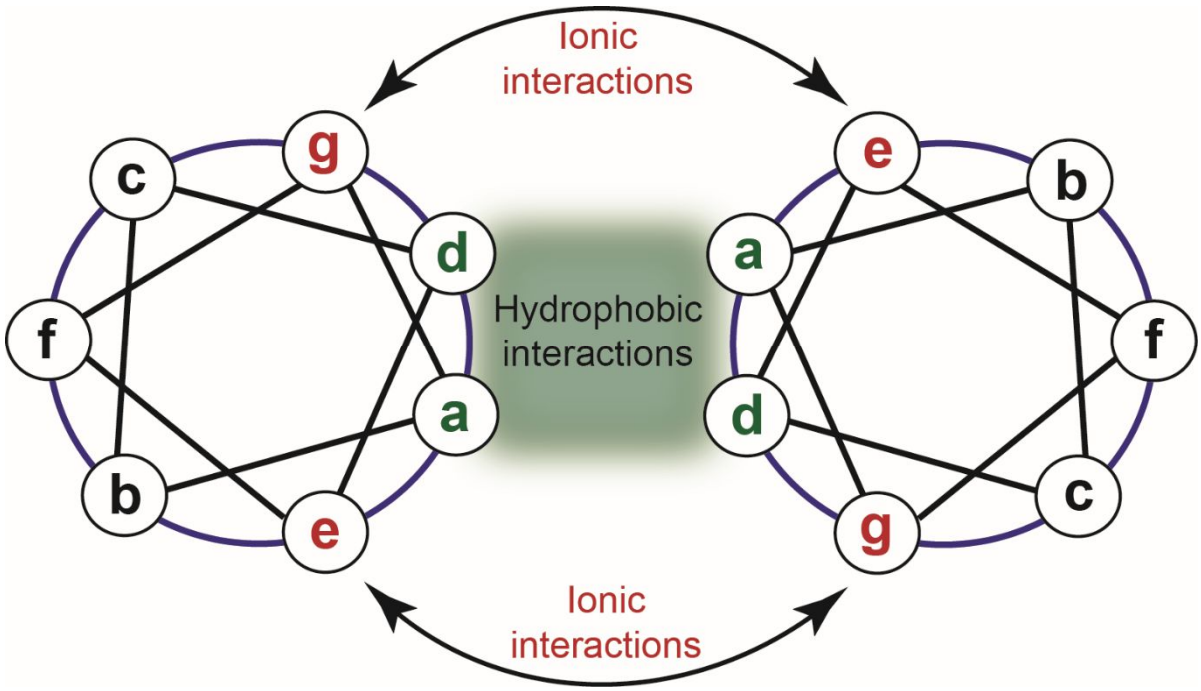


Figure 1.19. Coiled-coil interactions. The helical wheel plot of coiled-coils results in a repeating heptad unit. The "a" and "d" positions are occupied by hydrophobic amino acid residues that drive coiled coil formation through hydrophobic interactions between "a" and "d" positions of alpha helices. The "e" and "g" positions may be occupied by charged residues in order to further stabilise coiled-coil formation with electrostatic interactions.

the periodicity of the coiled-coil heptad, amphipathic helices are formed whereby a continuous hydrophobic surface is achieved. This hydrophobic surface drives the formation of coiled-coils, where multiple helices form the coiled-coil structure through burying the respective hydrophobic surface to form a hydrophobic core (McLachlan, 1975).

1.7.2 The functions of coiled-coiled structures

Coiled-coil proteins perform a wide variety of biological functions. The highly regular nature of the heptad repeats allows the provision of regulated molecular spacing that is exploited for many biological processes. These repetitive scaffolds can be used as molecular spacers that physically separate protein domains, such as providing transmembrane structures in order to connect proteins domains across a membrane (Truebestein and Leonard, 2016).

Further than acting as simple molecular rulers or physical spacers, coiled-coil proteins can undergo sequence specific functions. Notably, the cohesin structures that function to maintain cohesion of sister chromatids through ring formation contain coiled-coil regions that connect a hinge domain to the ATP domain. Although various pairings of the respective coiled coil domains occur, there is a high conservation between the different proteins, suggesting that a sequence specific role is required for the cohesin function (Haering, 2002). The most powerful

coiled-coil structures are those that form motor proteins, in which large coiled-coil assemblies occur to drive highly dynamic processes, such as the kinesin motor protein (Tripet, 1997).

Many proteins of the synaptonemal complex have predicted coiled-coil regions, with some of these regions providing interaction sites to other SC components. Therefore, the SC network likely exploits coiled-coil interactions to form the recursive tethering assembly. In particular, the transverse filaments have a central helical region with predicted coiled-coil formation and the length of this region has been shown to govern the width of the SC (Ollinger *et al*., 2005). Therefore, the elongated coiled-coil provides the physical spacer for the synapsis of homologous chromosomes in meiosis.

1.7.3 The k-m-e-f class of fibrous coiled-coil assemblies

The k-m-e-f class of proteins are a specific alpha helical protein family that were described through the study of fibrous X-ray diffraction from a variety of natural materials. The name was give due to the materials studied: keratin, myosin, epidermin and fibrinogen (Astbury, 1932). The X-ray diffraction observed from these materials displayed fibrous diffraction, with meridional arcs at 5.15 Å and equatorial reflections at 10-20 Å (Figure 1.20). Many subsequent studies of these proteins have characterised these coiled-coil structures that are formed from varying numbers of helices to perform a wide array of biological functions (Parry, 2008).

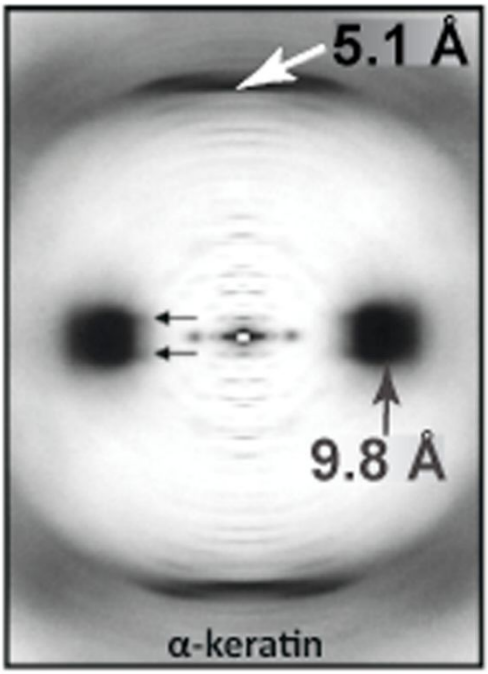


Figure 1.20. X-ray diffraction of α-keratin. The α-keratin diffraction pattern shows horseshoe equatorial reflections at 9.8 Å and a meridional arc at 5.1 Å. Original image collected by T. P. MacRae, taken from Parry *et al*., 2008.

1.8 Aims and objectives

The aim of this thesis is to explore the mechanisms of SC formation, with a primary focus on the organisation and choreography of central element assembly. In particular, he purpose is to address how the transverse filament proteins interact with the central element components, the molecular organisation of the central element and the mechanisms of higher-order assembly that propagate the maturation of the SC.

These areas have been investigated by yeast 2-hybrid studies and biophysical analysis of recombinant proteins with both the human and yeast SC constituents. The major hypotheses that have been tested are the mechanisms of higher-order assemblies that act to trigger SC polymerisation and the network of interactions that occur to form the central element.

In Chapter 3, the connection between the human transverse filament and central element has been investigated and characterisation of sub-complexes has been described. The network of interactions that occur in the central region has been outlined in order to propose a model for human central element formation.

In Chapter 4, the SYCE2-TEX12 complex has been studied to investigate the mechanisms of higher-order assembly that may occur in order to elongate the central element for SC maturation. The TEX12 protein has been characterised in absence of SYCE2 and mutational studies of TEX12 and SYCE2-TEX12 provide a molecular model for SYCE2-TEX12 filament formation.

In Chapter 5, preliminary studies have been carried out in order to characterise the yeast SC central element *in vitro*. Various truncations of the transverse filament protein, Zip1, and the central element proteins Ecm11 and Gmc2 have been characterised in order to provide a preliminary understanding of the organisation of the yeast SC central region.

Chapter two.

Methods

2.1 Materials

2.1.1 Chemicals

Chemicals used in this study were obtained from Sigma Aldrich, Thermo Scientific, Melford and Severn Biotech. Microbial growth media was purchased from Formedium and prepared according to the specified guidelines.

2.1.2 Preparation of antibiotic stocks

Antibiotic stocks for bacterial selection were prepared to be used in a 1:1000 dilution. Ampicillin stocks were made to 100 mg/ml and kanamycin stocks were made to 50 mg/ml by dissolving ampicillin (Melford) or kanamycin (Melford) into distilled water (dH_2O). Stocks were sterilised by filtration and stored at -20°C until use.

2.2 DNA manipulation

2.2.1 Plasmids used in this study

For yeast 2-hybrid experiments, the pGBKT7 and pGADT7 plasmids (Clontech) were used for expression of proteins fused to the GAL4 DNA-binding and activator domains, respectively. The pHAT4 and pMAT11 plasmids were used for single protein expression fused to a His_6 tag and His_6-MBP tag, respectively (Peranen *et al*., 1996). The pRSFDuet1 plasmid (Novagen), containing two multiple cloning sites, was used for dual protein expression. The target sequences, fused to affinity tags, were amplified from the pHAT4 and pMAT11 plasmids and inserted into the pRSFDuet1 multiple cloning sites. The pRSFDuet1 plasmid was adapted to produce a plasmid that expresses a non-cleavable N-terminal MBP tag with no linker between MBP and the inserted sequence, referred to as pLUC7. This plasmid was made by the insertion of the MBP sequence into the first multiple cloning site, with disruption of the pre-existing NcoI site and introduction of a new NcoI site downstream of the MBP sequence. The second multiple cloning site was not disrupted and therefore the pLUC7 plasmid was also used for dual protein expression. Protein expression from the pHAT4, pMAT11, pRSFDuet1 and pLUC7 plasmids was under the control of the IPTG inducible T7 promoter.

All sequences were inserted into these vectors through one-step sequence and ligation independent cloning (SLIC) (Jeong *et al*., 2012). Plasmid linearisation was performed with a single restriction enzyme at the same site each time for target sequence insertion at this site (Table 2.1). Primers used to generate the inserts for each plasmid were designed with the same sequence with homology to the corresponding insertion sites in the plasmid, followed by the coding sequences corresponding to the individual target sequences (Table 2.1). *E. coli* codon optimised sequences used for bacterial protein expression are provided in Appendix I.

Plasmid	Resistance marker	Restriction site	Forward primer arm 5' to 3'	Reverse primer arm 5' to 3'
pGBKT7	Kan	EcoR1	CATGGAGGCCGAATTCATGxxx	GGATCCCCGGGAATTCTTATTAxxx
pGADT7	Amp	EcoR1	GGAGGCCAGTGAATTCATGxxx	CACCCGGGTGGAATTCTTATTAxxx
pHAT4	Amp	NcoI	TTCCAGGGTTCCATGGxxx	AATTCGATATCCATGGTTATTAxxx
pMAT11	Amp	NcoI	TTCCAGGGTTCCATGGxxx	AATTCGATATCCATGGTTATTAxxx
pRSFDuet1	Kan	NcoI (MCS I) NdeI (MCS II)	AGGAGATATACCATGGxxx AAGGAGATATACATATGxxx	TGGCTGCTGCCCATGGTTATTAxx TTGAGATCTGCCATATGTTATTAxxx
pLUC7	Kan	NcoI (MCS I) NdeI (MCS II)	AACGCAGCGGCCATGGxxx AAGGAGATATACATATGxxx	TGGCTGCTGCCCATGGTTATTAxxx TTGAGATCTGCCATATGTTATTAxxx

Table 2.1. Properties of plasmids used in this study. The primer arms correspond to the homology regions to the plasmids, with the "xxx" indicating the region where the target specific coding sequence was inserted.

2.2.2 PCR amplification

PCR amplification was performed to generate DNA fragments for insertion into the expression plasmid using primers synthesised by Sigma Aldrich, with primer design validated by the ExPASy translate tool (http://www.web.expasy.org/translate/). KOD DNA polymerase (Invitrogen) was used for PCR amplification, according to the guidelines specified by the manufacturer. Thermocycling conditions as follows:

Step 1: 95°C for 2 minutes
Step 2: 98°C for 15 seconds (denaturation)
Step 3: 55°C for 5 seconds (annealing)
Step 4: 72°C for 20 seconds (extension)
Step 5: 74°C for 5 minutes
Hold: 4°C

Full PCR amplification was performed through 35 cycles of steps 2-4. For site directed mutagenesis and linker insertions, overlapping PCR was performed. The overlapping fragments were generated by initial PCR amplification by 10 cycles of steps 2-4. The resultant products were mixed and used as the template for a 35 cycle PCR amplification.

2.2.3 Agarose gel electrophoresis

DNA was analysed by 0.8% (w/v) agarose gel electrophoresis. Agarose was added to 1x TBE buffer, diluted from a 10X buffer stock (1M Tris, 0.9M boric acid and 0.01M EDTA pH 8.0). Agarose was dissolved by heating and then cooled (but not set) before addition of 1X SYBR® Safe (Invitrogen) and set in a prepared cassette with appropriate well combs. DNA samples were prepared through addition of 1X gel loading dye (NEB) and loaded into the wells alongside either 1kb or 100bp DNA ladders (NEB). Gel electrophoresis was performed with 1X TBE running buffer at 100 V for approximately 30 minutes, or until appropriate migration was achieved. DNA gels were imaged with the Gel Doc™ XR+ Gel Documentation System (BioRad) or visualised by the Dark Reader transilluminator (Clare Chemical Research) for band excision.

2.2.4 Isolation of DNA

Plasmid DNA was isolated using the GeneJET™ Plasmid Miniprep Kit (Thermo Scientific™), following the instructions supplied by the manufacturer. DNA isolation from enzymatic reactions was performed using a GeneJET™ PCR Purification Kit (Thermo Scientific™), following the manufacturer protocol. DNA bands were excised from agarose gels using a blue light box for visualisation, Dark Reader transilluminator (Clare Chemical Research). The DNA was then extracted and purified using a GeneJET™ Gel Extraction Kit (Thermo Scientific™), according to the supplied instructions. DNA concentrations were determined using a nanodrop UV-Vis spectrophotometer (Thermo Scientific™).

2.2.5 Restriction digestion of DNA

Plasmid linearisation was achieved through single restriction enzyme digestion (Table 2.1). 2 µg of plasmid DNA was used per 50 µl reaction, with 1X enzyme buffer (as specified by the manufacturer) and 2 µl (20 units) restriction enzyme (NEB). Reactions were incubated at 37°C for 2-3 hours and subsequent linearised plasmids were isolated through agarose gel electrophoresis, band excision and gel extraction.

2.2.6 One-step sequence and ligation independent cloning

One-step sequence and ligation independent cloning (SLIC) was performed to insert desired DNA fragments into linearised plasmids (Jeong *et al.*, 2012). DNA fragments were generated by PCR amplification and linearised plasmids were prepared as described above. Inserts and linearised plasmids were mixed at a 3:1 molar ratio (50-100 ng plasmid and mass of insert calculated accordingly) with 1X T4 DNA Polymerase buffer (NEB) and 1X BSA

(NEB) in a 10 µl reaction. 0.3 µl (0.9 units) T4 DNA Polymerase (NEB) was added to the reaction, followed by a precise incubation at 22°C for 2.5 minutes. After 2.5 minutes, reactions were promptly transferred to ice for 10 minutes. 50 µl of DH5α competent cells were transformed with 3 µl cloning reaction (see below) and plasmids were purified from selected colonies. Restriction digests were performed and analysed by agarose gel electrophoresis to detect clones with inserts of the correct size. Plasmids were sequenced by Source Bioscience and data analysed by Chromas (https://technelysium.com.au/wp/chromas/) to confirm the correct sequence of the clone before use. A list of sequenced clone plasmids used in this study are supplied in Tables 2.2, 2.3, 2.4 and 2.5.

Plasmid	Protein	Amino acid boundaries	Name used in text
pGBKT7	SYCP1	1-811	pGBKT7 SYCP1$_{1-811}$
pGBKT7	SYCE1	25-179	pGBKT7 SYCE1$_{25-179}$
pGBKT7	SYCE2	1-218	pGBKT7 SYCE2
pGBKT7	SYCE3	1-88	pGBKT7 SYCE3
pGBKT7	TEX12	1-123	pGBKT7 TEX12
pGBKT7	SIX6OS1	1-262	pGBKT7 SIX6OS1$_{1-262}$
pGBKT7	Ecm11	1-302	pGBKT7 Ecm11
pGBKT7	Ecm11	206-302	pGBKT7 Ecm11$_{206-302}$
pGBKT7	Gmc2	37-188	pGBKT7 Gmc2$_{37-188}$
pGBKT7	Zip1	1-348	pGBKT7 Zip1$_{1-348}$
pGBKT7	Smt3	1-101	pGBKT7 Smt3
pGADT7	SYCP1	1-811	pGADT7 SYCP1$_{1-811}$
pGADT7	SYCE1	1-351	pGADT7 SYCE1
pGADT7	SYCE2	1-218	pGADT7 SYCE2
pGADT7	SYCE3	1-88	pGADT7 SYCE3
pGADT7	TEX12	1-123	pGADT7 TEX12
pGADT7	SIX6OS1	1-587	pGADT7 SIX6OS1
pGADT7	Ecm11	1-302	pGADT7 Ecm11
pGADT7	Ecm11	181-302	pGADT7 Ecm11$_{181-302}$
pGADT7	Ecm11	206-302	pGADT7 Ecm11$_{206-302}$
pGADT7	Gmc2	1-188	pGADT7 Gmc2
pGADT7	Gmc2	37-188	pGADT7 Gmc2$_{37-188}$
pGADT7	Gmc2	96-188	pGADT7 Gmc2$_{96-188}$
pGADT7	Zip1	1-348	pGADT7 Zip1$_{1-348}$
pGADT7	Zip1	175-348	pGADT7 Zip1$_{175-348}$
pGADT7	Smt3	1-101	pGADT7 Smt3

Table 2.2. Yeast 2-hybrid plasmids, corresponding to human and yeast proteins, used in this study.

Plasmid	Protein	Amino acid boundaries	Protein name used in text
pMAT11	SYCP1	112-783	MBP-SYCP1$_{112-783}$
pMAT11	SYCP1	1-362	MBP-SYCP1$_{1-362}$
pMAT11	SYCP1	101-362	MBP-SYCP1$_{101-362}$
pMAT11	SYCP1	101-362 V105E L109E	MBP-SYCP1$_{mut}$
pMAT11	SYCP1	112-362	MBP-SYCP1$_{112-362}$
pMAT11	SYCP1	176-362	MBP-SYCP1$_{176-362}$
pMAT11	SYCP1	206-362	MBP-SYCP1$_{206-362}$
pMAT11	SYCE3	1-88	MBP-SYCE3
pMAT11	SYCE3	1-88 W41E Y44E	MBP-SYCE3 W41E Y44E
pMAT11	SYCE3	9-88	MBP-SYCE3 9-88
pMAT11	SYCE3	9-88 Y11N L15N	MBP-SYCE3 9-88 Y11N L15N
pMAT11	SYCE3	12-88	MBP-SYCE3 12-88
pMAT11	SYCE3	12-88 L15N	MBP-SYCE3 12-88 L15N
pMAT11	SYCE3	1-88 P53Q	MBP-SYCE3 P53Q
pMAT11	SYCE3	1-88 PPP-loop	MBP-SYCE3 PPP-loop
pMAT11	SYCE2	1-218	MBP-SYCE2
pMAT11	SYCE2	57-165	MBP-SYCE2$_{\alpha}$
pMAT11	TEX12	1-123	MBP-TEX12
pMAT11	TEX12	49-113	MBP-TEX12$_{\alpha-\Delta C}$
pMAT11	TEX12	49-123	MBP-TEX12$_{\alpha}$
pMAT11	TEX12	49-123 L110E F114E I117E L121E	MBP-TEX12$_{\alpha\,cc-mut}$
pMAT11	TEX12	49-75-5TGS-76-123	MBP-TEX12$_{\alpha-75}$
pMAT11	TEX12	49-78-5TGS-79-123	MBP-TEX12$_{\alpha-78}$
pHAT4	SYCE3	1-88 P53Q	His-SYCE3 P53Q
pHAT4	SYCE3	1-88 PPP-loop	His-SYCE3 PPP-loop
pHAT4	SYCE3	1-88 Y11A	His-SYCE3 Y11A
pHAT4	SYCE3	1-88 W41A	His-SYCE3 W41A
pHAT4	SYCE3	1-88 Y44A	His-SYCE3 Y44A
pHAT4	SYCE3	1-88 W41A Y44A	His-SYCE3 W41A Y44A
pHAT4	SYCE3	1-52	His-SYCE3 1-52
pHAT4	SYCE3	36-88	His-SYCE3 36-88
pHAT4	SYCE3	45-88	His-SYCE3 45-88
pHAT4	TEX12	49-123 L110E F114E I117E L121E	His-TEX12$_{\alpha\,cc-mut}$
pHAT4	TEX12	1-123 L110E F114E I117E L121E	His-TEX12$_{cc-mut}$

Table 2.3. Single protein expression plasmids, corresponding to human proteins, used in this study.

Plasmid	Protein	Amino acid boundaries	Protein name used in text
pMAT11	Zip1	1-251	MBP-ZIP1$_{1-251}$
pMAT11	Zip1	1-336	MBP-ZIP1$_{1-336}$
pMAT11	Zip1	175-251	MBP-ZIP1$_{175-251}$
pMAT11	Zip1	175-336	MBP-ZIP1$_{175-336}$
pMAT11	Zip1	1-348	MBP-ZIP1$_{1-348}$
pMAT11	Zip1	175-348	MBP-ZIP1$_{175-348}$
pMAT11	Zip1	1-875	MBP-ZIP1$_{1-875}$
pMAT11	Zip1	175-875	MBP-ZIP1$_{175-875}$
pMAT11	Zip1	1-748	MBP-ZIP1$_{1-748}$
pMAT11	Zip1	175-748	MBP-ZIP1$_{175-748}$
pHAT4	Ecm11	TGS3-230-302	His-Ecm11$_{linker}$
pHAT4	Ecm11	206-302	His-Ecm11$_{206-302}$
pHAT4	Ecm11	220-302	His-Ecm11$_{220-302}$
pMAT11	Gmc2	1-188	MBP-Gmc2
pMAT11	Gmc2	37-188	MBP-Gmc2$_{37-188}$
pMAT11	Ecm11	206-302	MBP-Ecm11206$_{-302}$

Table 2.4. Single protein expression plasmids, corresponding to yeast proteins, used in this study.

Plasmid	Site 1	Site 1	Site 2	Site 2	Protein name(s) used in text
	Protein	Amino acid boundaries	Protein	Amino acid boundaries	
pRSFDuet1	-	-	His-SYCE3	1-88	His-SYCE3
pRSFDuet1	His-MBP-SYCP1	206-362	-	-	MBP-SYCP1$_{206-362}$
pRSFDuet1	His-MBP-SYCP1	206-362	His-SYCE3	1-88	MBP-SYCP1$_{206-362}$ His-SYCE3
pLUC7	SYCE3	1-88	-	-	MBP-SYCE3
pRSFDuet1	His-MBP-SYCE1	25-179	-	-	MBP-SYCE1$_{25-179}$
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	49-123	MBP-SYCE2$_{\alpha}$ His-TEX12$_{\alpha}$
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	49-113	MBP-SYCE2$_{\alpha}$ His-TEX12$_{\alpha\text{-}\Delta C}$
pRSFDuet1	-	-	His-TEX12	49-123 F102A F109E V116A	His-TEX12$_{\alpha\text{-}AEA}$
pRSFDuet1	-	-	His-TEX12	49-123 F102A F109A V116A	His-TEX12$_{\alpha\text{-}3A}$
pRSFDuet1	MBP-SYCE2	1-218	His-TEX12	1-123	MBP-SYCE2 His-TEX12
pRSFDuet1	MBP-SYCE2	1-218	His-TEX12	1-113	MBP-SYCE2 His-TEX12$_{\Delta C}$
pRSFDuet1	-	-	His-TEX12	1-123 F102A F109A V116A	His-TEX12$_{AEA}$
pRSFDuet1	MBP-SYCE2	1-218	-	-	MBP-SYCE2
pRSFDuet1	-	-	His-MBP-TEX12	49-123 F102A F109E V116A	MBP-TEX12$_{\alpha\text{-}AEA}$
pRSFDuet1	MBP-Gmc2	1-188	-	-	MBP-Gmc2
pRSFDuet1	-	-	His-Ecm11	230-302	His-Ecm11$_{230-302}$
pRSFDuet1	MBP-Gmc2	37-188	His-Ecm11	206-302	MBP-Gmc2$_{37-188}$ His-Ecm11$_{206-302}$
pLUC7	Ecm11	206-302	-	-	MBP-Ecm11$_{206-302}$
pRSFDuet1	MBP-Gmc2	96-188	-	-	MBP-Gmc2$_{96-188}$

Table 2.5. Dual protein expression plasmids, with either single or double sequence insertions, corresponding to human or yeast proteins, used in this study.

Plasmid	Site 1	Site 1	Site 2	Site 2	Protein name(s) used in text
	Protein	Amino acid boundaries	Protein	Amino acid boundaries	
pRSFDuet1	MBP-SYCE2	64-165	His-TEX12	1-123	MBP-SYCE2$_{64-165}$ His-TEX12
pRSFDuet1	MBP-SYCE2	57-120	His-TEX12	1-123	MBP-SYCE2$_{57-120}$ His-TEX12
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	63-123	MBP-SYCE2$_{\alpha}$ His-TEX12$_{63-123}$
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	76-123	MBP-SYCE2$_{\alpha}$ His-TEX12$_{76-123}$
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	1-113	MBP-SYCE2$_{\alpha}$ His-TEX12$_{\Delta C}$
pRSFDuet1	MBP-SYCE2	57-165	His-TEX12	49-119	MBP-SYCE2$_{\alpha}$ His-TEX12$_{49-119}$

Table 2.6. Dual protein expression plasmids used in the MRes project referred to in Chapter 4.

2.3 Transformation of chemically competent *E. coli* cells

2.3.1 Preparation of competent cells

Chemically competent DH5α and BL21 (DE3) *E. coli* cells were prepared by the same method for cloning and protein expression, respectively. Commercial DH5α (Life Technologies) and BL21 (DE3) (NEB) stocks were plated onto sterile LB agar (Formedium) plates and after 16 hours at 37°C, colonies were used to inoculate 250 ml LB broth (Formedium) supplemented with 20 mM magnesium sulphate (Sigma Aldrich). Cells were cultured at 37°C, 220 rpm for approximately 3 hours, until an optical density at 600 nm (OD_{600}) of 0.6 was obtained. Cells were pelleted by centrifugation at 2000 x g, 4°C for 5 minutes and the cell pellet was maintained at 4°C during resuspension with 100 ml sterilised buffer 1 (30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride and 15% v/v glycerol at pH 5.8). Following a 5 minute incubation on ice, cells were pelleted by centrifugation at 2000 x g, 4°C for 5 minutes, and the cell pellet was resuspended at 4°C with 10 ml of buffer 2 (10 mM MOPS, 75 mM calcium chloride, 10 mM rubidium chloride and 15% v/v glycerol at pH 6.5). The resuspended cells were incubated for 1 hour on ice, before aliquoting and flash freezing in liquid nitrogen for storage at -80°C.

2.3.2 Transformation of competent cells

Chemically competent DH5α *E. coli* cells were transformed with either isolated plasmids to be amplified or SLIC reactions to be recombined, and BL21 (DE3) cells were transformed with plasmids for protein over-expression. DH5α cells were transformed with 300-500 ng of SLIC reaction or plasmid. BL21 (DE3) cells were either transformed by 300-500 ng of a single plasmid or 300 ng of each plasmid for co-transformation and subsequent co-expression. Under sterile conditions and on ice, DNA was added to thawed cell aliquots and incubated for 20 minutes. Cells were heat-shocked at 42°C for 45 seconds and then kept on ice for 10 minutes before the addition of sterile SOC broth (Formedium). The cell culture was incubated at 37°C, 180 rpm for 1 hour, before plating onto sterile LB agar plates supplemented with appropriate antibiotics. Plates were incubated at 37°C for 16 hours.

2.4 Yeast 2-hybrid (Y2H)

2.4.1 Growth of strains

S. cerevisiae strains, Y187 and Y2H Gold (Clontech), were used for yeast 2-hybrid (Y2H) experiments following the Matchmaker Gold protocol (Clontech). Stocks were stored at -80°C in 20% glycerol and for use, were streaked onto sterile YPAD agar (Formedium) plates

adjusted to pH 6.5, followed by incubation at 30°C for 5 days. Strains were maintained by routinely streaking onto fresh YPAD plates.

2.4.2 Transformation of yeast cells

Y187 *S. cerevisiae* cells were transformed with pGBKT7 plasmids and the Y2H Gold *S. cerevisiae* cells were transformed with pGADT7 plasmids using the PEG/lithium acetate/single stranded DNA procedure. Deoxyribonucleic acid sodium salt from salmon testes (Sigma Aldrich) was used as the single stranded carrier DNA and prepared as a 2 mg/ml solution in phosphate buffered saline (PBS), sterilised by filtration and stored at -20°C. Before use, carrier DNA was inactivated through incubation at 95°C for 5 minutes followed by 10 minutes on ice, with the process repeated 2 times. Cells were scraped with a sterilised pipette tip and added to 50 µl of inactivated carrier DNA in a sterilised 1 ml tube. 2 µl of the desired plasmid (1-1.5 µg) was added to the tube, followed by 240 µl of filter sterilised 50% PEG 3350 and 36 µl 1M lithium acetate. Following thorough mixing, the transformation mixture was incubated at 30°C for 30 minutes, followed by a 15 minute incubation at 45°C. Cells were pelleted at 15000 x g for 30 seconds, resuspended with 100 µl dH_2O and plated on synthetically defined (SD) agar plates at pH 5.8, supplemented with appropriate drop out selection (Formedium). Y187-pGBKT7 transformations were plated onto SD agar containing all amino acids apart from tryptophan (SD –Trp) to select for transformants, due to the pGBKT7 plasmid containing a tryptophan biosynthesis gene. Y2H Gold-pGADT7 transformations were plated onto SD agar containing all amino acids apart from leucine (SD –Leu) to select for transformants, due to the pGADT7 plasmid containing a leucine biosynthesis gene. Plates were incubated at 30°C for five days and before use, transformant colonies were streaked onto fresh SD-Trp or SD-Leu plates and incubated for 3-5 days at 30°C.

2.4.3 Yeast mating

Y187-pGBKT7 and Y2H Gold-pGADT7 transformants were mated for Y2H experiments. Stocks of each transformant were made by scraping cells to inoculate 2x YPAD broth (Formedium). 250 µl of each stock for the mating reaction were mixed in a sterilised 1 ml tube and incubated for 24 hours at 30°C with gentle shaking at 50 rpm. After the 24 hour incubation, cells were pelleted by gentle centrifugation at 250 x g and resuspended in 500 µl 0.5x YPAD broth (Formedium). Reactions were plated onto corresponding medium to select for successful mating and protein-protein interactions. Double drop out (DDO) medium, made with SD agar supplemented with all amino acids apart from tryptophan and leucine at pH 5.8, was used to select for cells containing both pGBKT7 and pGADT7 plasmids as a mating

control. Quadruple drop out (QDO) medium was made with SD agar supplemented with all amino acids apart from tryptophan, leucine and histidine and also lacking adenine, adjusted to pH 5.8. Following sterilisation by autoclaving, X-α-gal was added to the QDO medium at a 1:500 dilution from a 20 mg/ml X-α-gal stock dissolved in dimethylformamide (DMF). The QDO-X-α-gal medium was used to select for mated colonies due to the activation of the HIS3, ADE2 and MEL1 reporter genes upon a positive protein-protein interaction and the resultant activation of GAL4. Blue colonies indicate the activation of MEL1 that encodes the X-α-galactosidase enzyme resulting in hydrolysis of the colourless X-α-gal in the medium, to a blue product. 100 µl of the mating reaction was plated onto the DDO agar plates and 150 µl plated onto the QDO-X-α-gal plates and then incubated at 30°C for 5 days. Plates were imaged on a white light box and grids displayed against a black background.

2.5 Recombinant protein expression and purification

2.5.1 Over-expression of recombinant protein

Plasmids containing sequences for protein over-expression were transformed into BL21 (DE3) *E. coli* cells as described above. Transformants were used to inoculate 10-200 ml LB broth starter cultures (Formedium) supplemented with appropriate antibiotics, grown at 37°C, 180 rpm for approximately 2-3 hours until an OD_{600} of 0.6 was reached. 2 L baffled flasks were autoclaved containing 1 L 2xYT (Formedium) media, and the sterilised media was then inoculated with 10 ml of starter culture, with the addition of appropriate antibiotics, and incubated at 37°C, 180 rpm for approximately 4-6 hours, or until an OD_{600} of 0.8 was reached. Protein expression was induced by the addition of 0.5 mM IPTG to the cultures followed by incubation at 25°C, 180 rpm for 16 hours.

2.5.2 Preparation of cell lysate

Following protein over-expression in BL21 (BE3) *E. coli* cells, cells were pelleted by centrifugation at 4000 x g, 4°C for 30 minutes. Pellets were resuspended in 30 ml lysis buffer, consisting of 20 mM Tris pH 8.0, 500 mM KCl and EDTA-free protease inhibitor mini tablets (Thermo Scientific), and lysed through sonication followed by centrifugation at 40000 x g, 4°C for 30 minutes to isolate the soluble protein material.

2.5.3 Nickel affinity chromatography

Clarified cell lysate was loaded onto Ni-NTA resin (Qiagen) packed into an econo-column (Bio-Rad Laboratories) that was pre-equilibrated in lysis buffer using a peristaltic pump, at a rate of 1 ml/min at 4°C. After loading of the cell lysate, the resin was washed with

5 column volumes of lysis buffer (see above), followed by 5 column volumes of lysis buffer supplemented with 20 mM imidazole pH 8.0 to remove any non-specific binding, and elution of bound proteins by lysis buffer supplemented with 200 mM imidazole pH 8.0.

2.5.4 Amylose affinity chromatography

Amylose resin (Qiagen), packed into an econo-column (Bio-Rad Laboratories), was pre-equilibrated in 20 mM Tris pH 8.0, 150 mM KCl, 2 mM DTT before sample loading at 1 ml/min with a peristaltic pump, at 4°C. The resin was washed with 5 column volumes of equilibration buffer to remove non-specific binding, and bound proteins were eluted by 3 column volumes of equilibration buffer supplemented with 30 mM D-maltose.

For recombinant protein pull-down assays, clarified cell lysates were applied directly to pre-equilibrated amylose resin (Qiagen), packed in econo-columns (Bio-Rad Laboratories) and incubated at 4°C for one hour by gentle mixing. Following settling of the resin, the flow-through was collected, followed by washing of the resin and elution of bound proteins as described above. Eluates were normalised for protein concentration and analysed by SDS-PAGE (see below).

2.5.5 Ion exchange chromatography

HiTrap Q or HiTrap SP 5 ml ion exchange chromatography columns (GE Healthcare) were used to perform anion and cation exchange chromatography respectively, using an ÄKTA Start purification system (GE Healthcare). The system and attached column were pre-equilibrated with 20 mM Tris pH 8.0, 100 mM KCl, 2 mM DTT and the sample was adjusted to 100 mM KCl before loading. Samples were loaded at 1 ml/min, followed by washing of the column with 20 mM Tris pH 8.0, 100 mM KCl, 2 mM DTT. Proteins were eluted across a 100% salt gradient in 1 ml fractions through mixing of buffer A (20 mM Tris pH 8.0, 100 mM KCl, 2 mM DTT) and buffer B (20 mM Tris pH 8.0, 1 M KCl, 2 mM DTT) across a 50 ml volume.

2.5.6 Preparative size exclusion chromatography

The S200 HiLoad 16/600 size exclusion chromatography column (GE Healthcare) was used for preparative size exclusion chromatography attached to an ÄKTA Pure purification system (GE Healthcare). The system was pre-equilibrated in 20 mM Tris pH 8.0, 150 mM KCl, 2 mM DTT before injection of 1-5 ml of sample, followed by an isocratic gradient with the output fractionated in 2 ml fractions.

2.5.7 Affinity tag removal by TEV protease

SfGFP-TEV was purified by Ni-NTA affinity and size exclusion chromatography before storage in 40% glycerol at -80°C (prepared by James Dunce). Affinity tags were removed from proteins fused to affinity tags with a TEV protease recognition site between the tag and protein sequence. Fusion protein samples were incubated with sfGFP-TEV protease at a 1:30 molar ratio, through room temperature incubation and gentle mixing for a minimum of 2 hours.

2.5.8 Concentration and buffer exchange of recombinant protein samples

Following purification of recombinant proteins, samples were concentrated by centrifugation. Centrifugal concentrator devices were used to concentrate protein samples by centrifugation at 6000 x g, with suitable concentrators used according to the sample molecular weight. Microsep™ Advance with Omega™ membrane (PALL) spin concentrators were used with either a 3 kDa or 10 kDa MWCO, with samples centrifuged in 10 minute intervals at 4°C and mixed manually by pipetting between centrifugation cycles. After concentration, protein samples were transferred to 0.5 ml tubes and centrifuged at 16000 x g for 30 minutes at 4°C to pellet any aggregated material. The sample was carefully transferred into a fresh tube for the determination of the protein concentration (see below). For storage, samples were flash frozen in liquid nitrogen and kept at -80°C until required. Buffer exchange of protein samples was achieved using NAP-5 columns (GE Healthcare) according to the supplied protocol by the manufacturer.

2.5.9 Protein concentration determination

Protein concentration was calculated using the Beer-Lambert law: $A = \mathcal{E} \times l \times c$. A is the absorbance, $\mathcal{E}$ is the extinction coefficient, l is the cell path length and c is the molar concentration. The absorbance was determined using UV spectroscopy with a Cary 60 UV-Vis spectrophotometer (Agilent) with samples measured in BRAND® micro 15mm UV cuvettes (Sigma Aldrich). The molar concentration was calculated using extinction coefficients determined using the ExPASy Protparam tool (http://www.web.expasy.org/protparam/).

2.5.10 SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to analyse protein samples, by using the NuPAGE Novex Bis-Tris sytem (Invitrogen) with 12% or 15% gels. Resolving gel mixtures and the stacking gel mixture were stored at 4°C and prepared as follows. The resolving mixtures were made with 36 mM Bis-Tris (Sigma Aldrich)

adjusted to pH 6.8 and 12 or 15% v/v acrylamide from a 30% acrylamide, 0.8% bis-acrylamide (37:5:1) solution (Severn Biotech). The stacking gel mixture was prepared with 36 mM Bis-Tris (Sigma Aldrich) adjusted to pH 6.8 and 4% v/v acrylamide from a 40% acrylamide, 2% bis-acrylamide (19:1) solution (Severn Biotech).

Resolving gel mixtures were supplemented with 0.035% v/v ammonium persulphate (APS), from a freshly prepared 10% w/v stock (Melford), and 0.001% v/v TEMED (Melford) to trigger polymerisation. The solution was gently mixed and promptly pipetted into upright Bolt® 1 mm gel cassettes (Invitrogen), followed by a layer of distilled water (dH_2O). Once the resolving gel was completely polymerised, the dH_2O was removed and dried with filter paper. Polymerisation of the stacking gel mixture was initiated by 0.1% v/v APS, from a freshly prepared 10% w/v stock (Melford), and 0.0035% v/v TEMED (Melford). After gentle mixing, the stacking gel solution was pipetted on top of the polymerised resolving gel, and combs were placed into the stacking gel layer before allowing to set. Once set, polymerised gels were wrapped in damp paper towels and cling film to prevent drying out, and stored at 4°C until use.

To perform gel electrophoresis, the combs were carefully removed from set gels and placed into a Bolt® mini gel tank (Invitrogen) with 1x MOPS-SDS running buffer, diluted from a 20x MOPS-SDS stock (Invitrogen). Samples were prepared by addition of 1x LDS loading dye, from a 4x LDS sample buffer stock (Invitrogen), and 1x reducing agent, from a 10x sample reducing agent stock (Invitrogen). Prepared samples were heated at 75°C for 10 minutes before loading into the wells alongside the PageRuler prestained 10-180 kDa protein ladder (Thermo Scientific) for molecular weight standards. Gels were electrophoresed at 180 V for approximately 30 minutes, or until sufficient migration was achieved. Gels were then retrieved from the cassettes and washed with dH_2O 3 times by gentle heating. Protein bands were visualised by the Coomassie Quick Staining method, through heating with the Coomassie solution followed by a 10 minute incubation with gentle rocking. Coomassie stain was prepared with 10% ethanol, 30 mM hydrochloric acid and 0.01% Coomassie Brilliant Blue G (Sigma Aldrich). Following the 10 minute staining, gels were destained by washing with dH_2O several times until sufficient destaining was achieved, followed by imaging by scanning.

2.5.11 Sequence analysis

The ExPASy Protparam tool (http://www.web.expasy.org/protparam/) was used to determine properties of protein sequences such as the extinction coefficient, molecular mass and isoelectric point (pI). Multiple sequence alignments (MSAs) were generated by retrieving protein sequences from UniProtKB (http://www.uniprot.org) and using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/)) to perform the alignments. JNET

(http://www.compbio.dundee.ac.uk/www-jpred/) was used for secondary structure prediction and Consurf (http://consurf.tau.ac.il/) was used to quantify the conservation of amino acid positions from the MSAs. MSAs and JNET secondary structure predictions were visualised in Jalview 2.8.2 (http://www.jalview.org).

2.6 Biochemical and biophysical assays

2.6.1 Circular dichroism (CD)

Far-UV circular dichroism (CD) experiments were performed using a Jasco-810 spectropolarimeter (Institute for Cell and Molecular Biology, Newcastle University). CD wavelength scans were recorded between 260 and 185 nm, with 0.2 nm intervals. Samples were buffer exchanged (as described previously) into disodium phosphate pH 7.5, 150 mM sodium fluoride buffer, and analysed at 0.1-0.5 mg/ml in a 0.2 mm path length quartz cuvette (Hellma) at 4°C. Three accumulations were measured for a buffer sample alone, and nine accumulations were recorded for each protein sample. The nine accumulations were averaged and corrected for buffer before conversion to mean residue ellipticity (MRE), [θ], (x1000 $deg.cm^2.dmol^{-1}.residue^{-1}$) (Kelly *et al*., 2005). The Dichroweb server (http://dichroweb.cryst.bbk.ac.uk) was used to perform deconvolution of the data to estimate the secondary structure composition of the sample using the CDSSTR algorithm (Sreerama and Woody, 2000).

CD thermal denaturation was performed to study the thermal stability of protein samples. Samples were analysed at 0.1 mg/ml in 20 mM Tris pH 8.0, 150 mM KCl, 2mM DTT buffer using a 1 mm path length quartz cuvette (Hellma). The 222 nm signal was recorded between 5°C and 95°C, at 0.5°C intervals, with a ramping rate of 1°C per minute. The data were converted to MRE, [θ], and then plotted as per cent unfolded, assuming that samples are 100% folded at 5°C and 100% unfolded at 95°C $([\theta]_{222,x}-[\theta]_{222,5})/([\theta]_{222,95}-[\theta]_{222,5})$. The melting temperatures were estimated as the temperature at which 50% of the sample is unfolded. In cases where two-state transitions were observed (see Chapter five) multiple melting temperatures were estimated at the midpoint of each unfolding event.

2.6.2 Transmission electron microscopy (TEM)

Transmission electron microscopy (TEM) experiments were performed using a Philips CM100 TEM (Electron Microscopy Research Services, Newcastle University). Samples were analysed at 2-7 mg/ml and centrifuged at 16000 x g, 4°C for 30 minutes prior to use to remove any aggregates. Proteins samples were applied to glow-discharged holey carbon-coated grids for 30 seconds, with excess material carefully removed with filter paper. The grid was washed by application of dH_2O for 5 seconds before removal with filter paper, with the washing process

repeated 2 times to remove excess protein material. The applied protein was then fixed to the grid by negative staining, achieved by application of filtered 0.1% v/v uranyl acetate. Grids were prepared within 30 minutes of glow-discharging and were then analysed, with images taken at several points across the grid at various magnifications.

2.6.3 Analytical size exclusion chromatography (SEC)

Analytical size exclusion chromatography (SEC) experiments were performed with a Superdex 200 increase 10/300 GL column (GE Healthcare) attached to an ÄKTA™ Pure system (GE Healthcare), equilibrated in 20 mM Tris pH 8.0, 150 mM KCl, 2 mM DTT buffer. Samples were incubated for 1 hour on ice before centrifugation at 16000 x g, 4°C for 30 minutes to remove any aggregates. Samples were then loaded onto the column at 0.5 ml/min, with the output fractionated in 1 ml fractions. The fractions corresponding to the protein species were analysed by SDS-PAGE.

2.6.4 Size exclusion chromatography multi-angle light scattering (SEC-MALS)

Size exclusion chromatography multi-angle light scattering (SEC-MALS) was used to determine the absolute molecular mass of protein species following separation according to their size and shape. Superdex 200 or Superose 200 increase 10/300 GL columns (GE Healthcare) were attached to the ÄKTA™ Pure system (GE Healthcare). The output was directed into the DAWN® HELEOS™ II MALS detector (Wyatt Technology), followed by the Optilab® T-rEX™ differential refractometer (Wyatt Technology). The system was equilibrated in 20 mM Tris pH 8.0, 150 mM KCl, 2 mM DTT at 0.5 ml/min overnight before use in order to stabilise the differential refractometer. Samples were centrifuged at 16000 x g, 4°C for 30 minutes before analysing to remove aggregates. Bovine serum albumin (BSA) was analysed as the first sample, for use as the configuration for subsequent samples. Where necessary, the sample output was fractionated for analysis of 1 ml fractions by SDS-PAGE. Data was collected and analysed by the ASTRA® 6 software (Wyatt Technology), using a nominal dn/dc value of 0.185 ml/g for Zimm plot extrapolation to determine the molecular mass of protein species. The molecular weights quoted were calculated across the eluted protein peaks.

2.6.5 Size exclusion chromatography small-angle X-ray scattering (SEC-SAXS)

Size exclusion chromatography small-angle X-ray scattering (SEC-SAXS) was used to study the structural properties of protein species following separation due to their size and shape. The B21 beamline at the Diamond Light Source synchrotron facility (Oxfordshire, UK) was used for SEC-SAXS data collection. A Superdex 200 increase 10/300 GL SEC column

(GE Healthcare) was attached to an Agilent 1200 HPLC system, with the system pre-equilibrated in 20 mM Tris pH 8.0, 150 mM KCl buffer at 0.5 ml/min. Samples were transported in dry ice and before use, were rapidly thawed before centrifugation at 16000 x g, 4°C for 30 minutes to remove any aggregated material. Samples were then pipetted into a 96 well tray that was stored at 4°C in the HPLC system to await injection and analysis. Upon injection, samples were loaded onto the SEC column and the output was then directed to the SAXS experimental cell. SAXS data was collected in 3 s frames at 12.4 keV with a 4.014 m detector distance.

SAXS data were analysed using the ScÅtter 3.0 software (http://www.bioisis.net). The scattering data from protein peaks were subtracted for buffer, followed by averaging of selected buffer-subtracted frames. Guinier analysis was performed to determine the radius of gyration (Rg) and the cross-sectional radius of gyration (Rc). The paired-distance distribution function, P(r), was used for the real space analysis of the inter-atomic distances observed within the sample, determined by using PRIMUS from the ATSAS suite (Konarev *et al*., 2003). The P(r) distributions were normalised for presentation, with the most commonly observed distance normalised to 1.0. *Ab initio* modelling was performed to produce low resolution molecular envelopes from the P(r) distribution analysis. DAMMIF (Franke and Svergun, 2009), from the ATSAS suite, was used for the modelling with 10-20 independent runs performed for each sample. DAMSEL, DAMSUP, DAMAVER and DAMFILT, from the ATSAS suite, were used to compare, filter and average models to produce the final envelope (Kozin and Svergun, 2001). The final models presented are the DAMFILT molecular envelopes with surface representation. Further details about SEC-SAXS data processing provided in appendix II.

2.6.6 Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays (EMSAs) were performed to detect protein-DNA interactions through alterations in band migration. Agarose gels were prepared in the same manner as previously described, but 0.5% agarose gels were cast, rather than 0.8%, using 0.5x TBE buffer, and poured gels were left to set at 4°C. Once fully set, gels were primed by electrophoresis at 20 V for 1 hour before sample loading. Samples were prepared by mixing double-stranded DNA (dsDNA), produced by PCR amplification, with varying concentrations of protein at desired ratios and were incubated at 4°C for 1 hour on ice. After preparative electrophoresis of the agarose gel, glycerol was added to the samples to a final concentration of 4% v/v, and samples were carefully loaded into the well. The agarose gel was electrophoresed at 20 V, 4°C for 4-6 hours, or until sufficient band migration was achieved and imaged using a UV gel doc Gel Doc™ XR+ Gel Documentation System (BioRad).

2.6.7 Cross-linking mass spectrometry

Crosslinking was performed with the bis(sulfosuccinimidyl)suberate (BS3) cross-linker. A 100 mM stock was prepared for serial dilution resulting in the formation of a 10X stock for each reaction, to ensure the volume of cross-linker added to the sample was consistent. Protein sample was buffer exchanged into disodium phosphate pH 7.5, 150 mM sodium fluoride buffer to remove the Tris from the sample buffer. Upon addition of cross-linker and gentle mixing by pipetting, reactions were incubated at room temperature for 30 minutes, followed by quenching of the reaction with 100 mM Tris pH 8.0. Samples were then analysed by SDS-PAGE, with appropriate bands excised for mass spectrometry analysis. The mass spectrometry and subsequent analysis was performed by Juan Zou, Juri Rappsilber group, University of Edinburgh.

2.7 X-ray crystallography

2.7.1 Crystallisation trials

High throughput crystallisation screening was performed as the first stage of crystallographic studies to allow for screening of many different conditions. JSCG+, PACT, Morpheus, Structure (Molecular dimensions) and Index (Hampton Research) commercial 96 well screens were used to set up 96 well crystallisation trays. The Mosquito nanodrop crystallisation robot (X-ray facility, Newcastle University) was used to set up two vapour diffusion sitting drops per condition, with 100 nl of well solution mixed with 100 nl (drop 1) or 200 nl (drop 2) of protein. Protein samples were screened at concentrations between 5 and 75 mg/ml in 20 mM Tris pH 8.0, 150 mM KCl, 2 mM DTT buffer, with centrifugation of samples at 16000 x g, 4°C for 30 minutes prior to use to remove aggregates. Following the formation of crystallisation drops, trays were promptly sealed and incubated at 20°C. Drops were regularly monitored and imaged using a DFC400 digital microscope camera (Leica).

2.7.2 Optimisation screening

Following successful crystallisation during commercial screening, crystals were optimised through screening of numerous variables, such as component concentrations, pH, drop size and protein concentration. Optimisation screens were set up manually in 24 well plates by hanging drop vapour diffusion. Further optimisation techniques were employed when necessary, such as seeding and additive screening. Micro-seeding was performed though manual shattering of crystals using a seed bead tube (Hampton Research) and consequent streaking of the seed stock into pre-equilibrated drops. Cross-seeding was also performed by including the seed stock in initial commercial screens to encourage crystallisation. Additive

screening was performed manually by hanging drop vapour diffusion of 96 wells with consistent protein concentration and well solution, with the inclusion of 10% of each additive from the commercial screen (Hampton Research).

2.7.3 Data collection and structure determination of TEX12$_{\alpha\text{-}\Delta C}$

Commercial crystallisation screening was performed with TEX12$_{\alpha\text{-}\Delta C}$ (TEX12 residues 49-113) at 30 mg/ml with one condition found to form crystals, Structure H7 (Molecular Dimensions) consisting of 35% v/v 1,4-dioxane. Crystallisation was optimised to grow larger, higher quality crystals with 60 mg/ml of protein. However, no crystals were initially detected 24 hours after drop formation at any 1,4-dioxane concentration. Larger crystals were found to form in 25% v/v 1,4-dioxane 3 hours after hanging drop formation and these crystals dissolved 6 hours after drop formation, with deterioration of the crystals detected after 4 hours. Crystals were therefore harvested 3 hours after drop formation, using loops of appropriate size, briefly soaked in 20% ethylene glycol for cryo-protection and then flash frozen in liquid nitrogen.

X-ray diffraction data was collected at the I03 beamline at the Diamond Light Source synchrotron facility (Oxfordshire, UK) on a Pilatus 6M detector as 3600 consecutive 0.10° frames of 0.010 s exposure, at 0.9796 Å, 100 K. XDS (Kabsch, 2010) was used to index and integrate diffraction data, with scaling and merging performed by AutoPROC in Aimless (Vonrhein *et al*., 2011). The TEX12$_{\alpha\text{-}\Delta C}$ asymmetric unit contains two TEX12 chains and the crystals were found to belong to the $P6_522$ hexagonal spacegroup with cell dimensions of a = 47.97 Å, b = 47.97 Å, c = 210.98 Å, $\alpha = 90°$, $\beta = 90°$, $\gamma = 120°$. The UCLA diffraction anisotropy server (https://services.mbi.ucla.edu/anisoscale/) was used to correct data for anisotropy and Quark *ab initio* model decoys were generated (Xu and Zhang, 2012). The CCP4 online server (https://www.ccp4.ac.uk/ccp4online/) was used for structure solution by AMPLE (Bibby *et al*., 2012) that performed molecular replacement with the generated Quark models followed by auto-tracing and rebuilding by SHELX E (Sheldrick, 2008). Coot (Emsley *et al*., 2010) was used for model building and PHENIX refine was used for refinement (Adams *et al*., 2010). The structure was completed through the addition of five ligands of 1,4-dioxane from the well solution and refined against 2.11 Å data that was corrected for anisotropy, to an R value of 0.2291 and R_{free} value of 0.2580. 100% of residues are within the Ramachandran plot favoured regions, with a clashscore and overall MolProbity score of 2.30 and 1.01 respectively (PDB code: 6HK8, see chapter four). Structure solution was performed by Dr Owen Davies.

2.7.4 Data collection and structure determination of TEX12$_{\alpha\text{-AEA}}$

Crystallisation and data collection of TEX12$_{\alpha\text{-AEA}}$ (TEX12 residues 49-123 harbouring F102A, F109E and V116A mutations) was carried out by James Dunce and structure solution was performed by Dr Owen Davies. Crystals were grown in 200 mM calcium acetate, 40% MPD at 20°C in vapour diffusion sitting drops with 100 nl of well solution mixed with 100 nl of protein at 43 mg/ml. Crystals were harvested 24 hours after drop formation and flash frozen in liquid nitrogen.

X-ray diffraction data was collected at the I24 beamline at the Diamond Light Source synchrotron facility (Oxfordshire, UK) on a Pilatus 6M detector as 2000 consecutive 0.10° frames of 0.010 s exposure, at 0.9790 Å, 100 K. XDS (Kabsch, 2010) was used to index and integrate diffraction data, with scaling and merging performed by AutoPROC in Aimless (Vonrhein *et al*., 2011). The TEX12$_{\alpha\text{-AEA}}$ asymmetric unit contains two TEX12 chains and the crystals were found to belong to the C222$_1$ spacegroup with cell dimensions of a = 43.233 Å, b = 219.712 Å, c = 37.501 Å, $\alpha = 90°$, $\beta = 90°$, $\gamma = 90°$. Quark *ab initio* model decoys were generated (Xu and Zhang, 2012) and the CCP4 online server (https://www.ccp4.ac.uk/ccp4online/) was used for structure solution by AMPLE (Bibby *et al*., 2012) that performed molecular replacement with the generated Quark models followed by auto-tracing and rebuilding by SHELX E (Sheldrick, 2008). PHENIX Autobuild (Adams *et al*., 2010) was used for phase improvement by re-building, Coot (Emsley *et al*., 2010) was used for manual model building and PHENIX refine (Adams *et al*., 2010) was used for refinement. The structure was completed through the addition of two acetate ligands, one MPD ligand and seven calcium ions. The structure was refined against 1.45 Å data with an R value of 0.1795 and an R_{free} value of 0.2047. 100% of residues are within the Ramachandran plot favoured, with a clashscore and overall MolProbity score of 2.62 and 1.05 respectively (PDB code: 6HK9, see Chapter four).

2.7.4 Refinement of the mouse SYCE3 structure

The mouse crystal structure of SYCE3 (PDB accession 4R3Q) was previously released (Lu *et al*., 2014). Owen Davies re-built the structure to include the four amino acid loop (residues 50-53) and three amino acid N-terminal extension (residues 10-12) to each chain. This was achieved using PHENIX Autobuild (Adams *et al*., 2010), PHENIX Fit Loops (Adams *et al*., 2010) and manual building in Coot (Emsley *et al*., 2010). PHENIX refine (Adams *et al*., 2010) was used to refine against data to 1.90 Å resolution, to R and R_{free} values of 0.1922 and 0.2167 respectively. 98.65% of residues are within the favoured Ramachandran plot regions (0 outliers), with a clashscore of 3.41 and an overall MolProbity score of 1.13. Rebuilt/re-fined

crystallographic atomic co-ordinates have been deposited in the Protein Data Bank (PDB) under accession number 6H86.

2.7.5 SAXS-directed protein structure modelling

Structural models were fitted to experimental SAXS data using CRYSOL (Svergun D.I., 1995). The N- and C-termini of the full length SYCE3 were modelled onto the re-refined SYCE3 structure using MODELLER (Sali and Blundell, 1993). The SYCE3 domain swap tetramer was modelled using Rosetta FloppyTail (Kleiger *et al*., 2009). The domain swap tetramer and the re-refined crystal structure were used to model the SYCE3 hexamer using RosettaDock (Gray *et al*., 2003) and the SYCE3 dodecamer was modelled from domain swap tetramers in PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. Modelling was performed by Owen Davies.

Chapter three.

Remodelling of the SYCP1 lattice by the molecular adapter SYCE3

3.1 Introduction

3.1.1 SYCP1 forms the mammalian transverse filaments

The SC transverse filaments form arrays that span the distance between the chromosome axis and the central element to tether together homologous chromosomes. The C-termini have been shown to localise at the chromosome axis and the N-termini are located in the central region (Liu *et al*., 1996) (Schmekel *et al*., 1996). The distance of the central region, inclusive of the central element and the transverse filaments, has been determined as approximately 100 nm across a wide range of organisms (Westergaard and von Wettstein, 1972), indicating the requirement for an elongated molecule to bridge the approximate 50 nm distance between the lateral and central elements. The central region has been shown to have a depth of 100 nm, through a bilayer formed by SYCP1 with central element assembly between the layers, demonstrating the 3D nature of the SC (Hernandez-Hernandez *et al*., 2016).

The mammalian transverse filaments are made from the elongated SYCP1 protein. SYCP1 is highly conserved throughout the sequence, with particularly high patches of conservation in some regions (Figure 3.1 A). Through secondary structure prediction of the SYCP1 sequence, a predominant alpha helical content is predicted, particularly in the central region, with the N- and C-termini predicted to be largely unstructured (Figure 3.1 B). The patches of highest conservation correlate to the N- and C-ends of the central helical region and the immediate N-terminus, residues 1-32 (Figure 3.1 A). Through sequence analysis, the central helical region is comprised of residues 101-783 (Figure 3.1 C). Biophysical analysis, performed by another member of the Davies group, has shown that the 101-783 region forms large helical assemblies, with a further N-terminal truncation to 112-783 resulting in tetramer formation. The helical core has two structural regions, 206-362 and 358-783, that drive the opposing assembly mechanisms of the SYCP1 ends in the central element and chromosome axis respectively (Dunce *et al*., 2018).

3.1.2 SYCP1 C-termini self-assemble at the chromosome axis

The mammalian lateral elements are formed from the SYCP2 and SYCP3 proteins that play a role in the structural organisation of the chromosomes for synapsis (Syrjanen *et al*., 2014). SYCP1 C-termini localise at the lateral element, with a direct interaction established between SYCP1 and the SYCP2 lateral element component (Winkel *et al*., 2009). However, SYCP1 is still recruited to meiotic chromosomes in the absence of SYCP2. Although the recruitment is reduced (Yang *et al*., 2006), it does suggest that there is another mechanism for SYCP1 chromosomal assembly. Therefore, the mechanism of SYCP1 C-terminal recruitment and assembly was investigated by another member of the Davies group.

The X-ray crystal structure of the SYCP1 C-terminal region of the central helical core, amino acids 676-770, was solved to reveal an antiparallel alpha helical tetrameric assembly (Dunce *et al*., 2018). This assembly is formed from the interaction between two sets of parallel

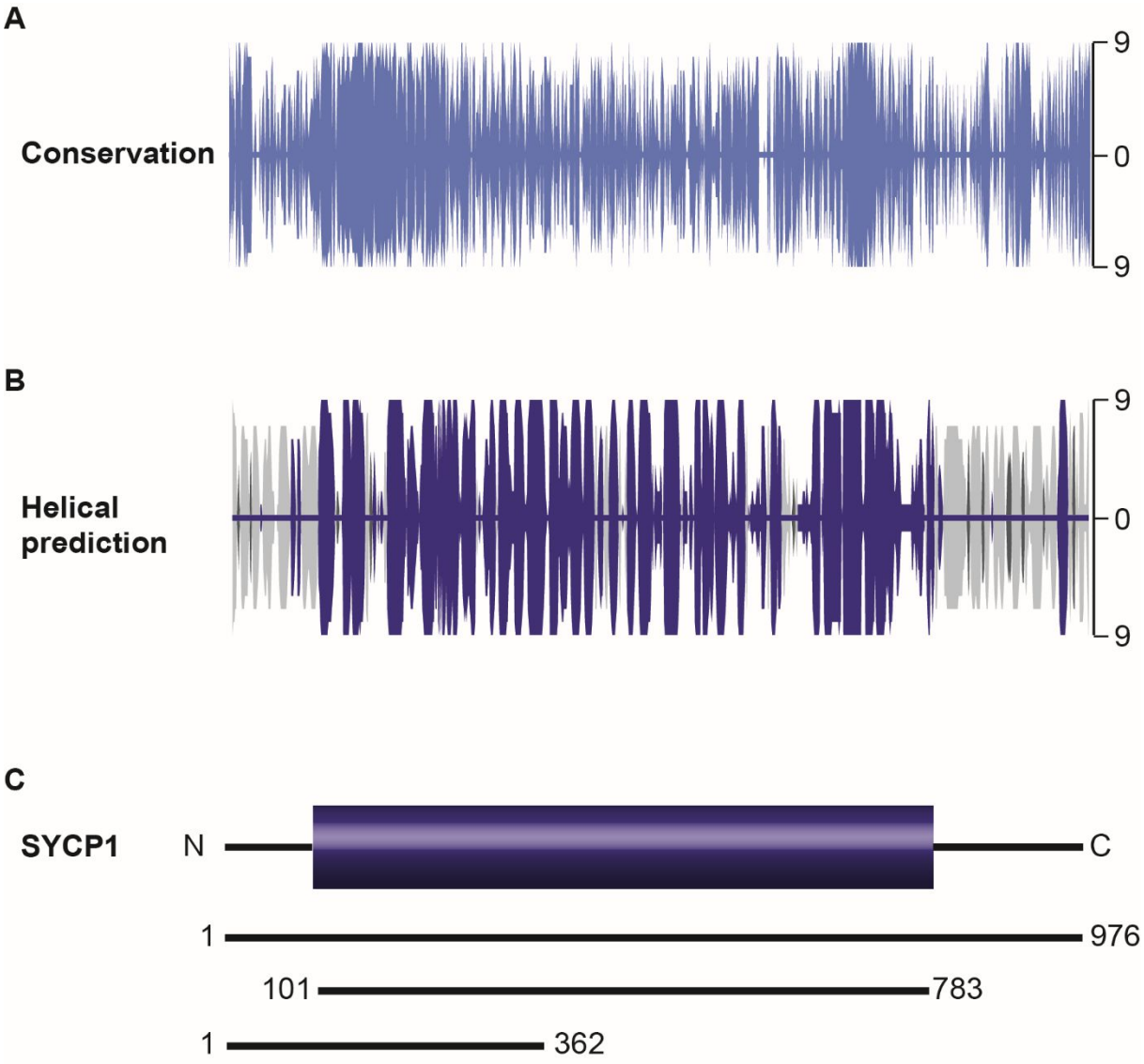


Figure 3.1. Sequence analysis of SYCP1. (A) SYCP1 conservation mirror plot. Clustal multiple sequence alignments were analysed by ConSurf to grade the conservation (0-9) of each amino acid position. (B) Secondary structure prediction of SYCP1 determined by JNET. Each residue is plotted as alpha helix (purple), beta sheet (dark grey) or unstructured (light grey) according to the confidence of the prediction (0-9), presented as a mirror plot. (C) Linear schematic representation of SYCP1. The purple box represents the central helical portion and the black lines represent the unstructured regions. Key regions are displayed below.

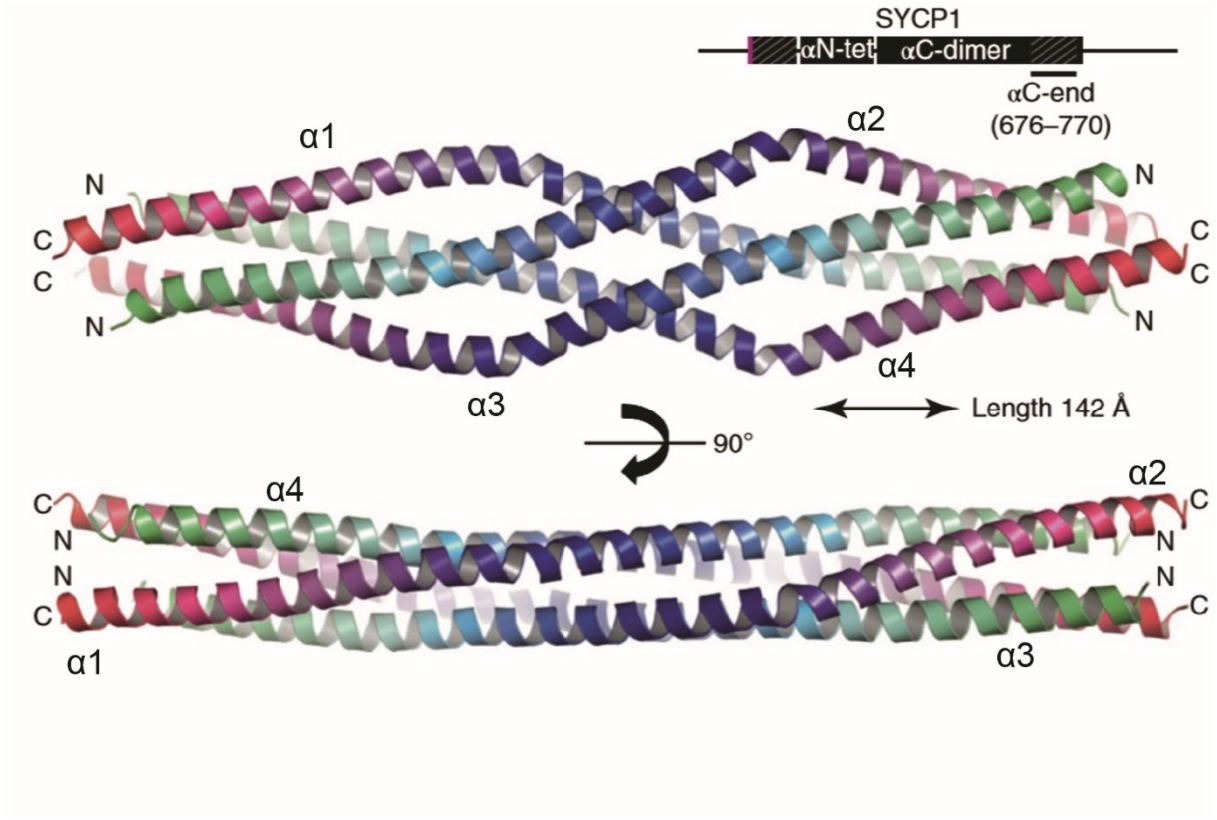


Figure 3.2. X-ray crystal structure of the SYCP1 676-770 region. The SYCP1 C-terminal region forms an alpha helical tetramer from an antiparallel assembly between two parallel dimers (PDB code: 6F63). Alpha-helix chains labelled α1-4 and structures presented at two orientations. Adapted from Dunce *et al.*, 2018.

dimers and this structure has been proposed to form the structural SYCP1 unit that assembles on the chromosomal axes (Figure 3.2). This SYCP1 region was found to exist as a dimer in solution, with tetramer formation detected in acidic conditions (Dunce *et al*., 2018). This region was also found to interact with dsDNA at pH 5.5, but with no DNA interaction detected at pH 8.0, which suggests that this region only interacts with DNA when in the tetrameric form. Further to this, the full length SYCP1 was found to bind dsDNA at pH 8.0, with this binding abolished upon truncation of the unstructured C-terminus (Dunce *et al*., 2018). Therefore, a model has been proposed whereby the unstructured SYCP1 C-terminus is recruited to the chromosome axis through DNA-binding, with the concentration of SYCP1 at the axis inducing the protonation-driven tetramerisation between adjacent SYCP1 dimers. Such mode of interaction enables further SYCP1-DNA associations (Dunce *et al*., 2018), thereby providing a mechanism for SYCP1 C-terminal self-assembly to coat the chromosome axes.

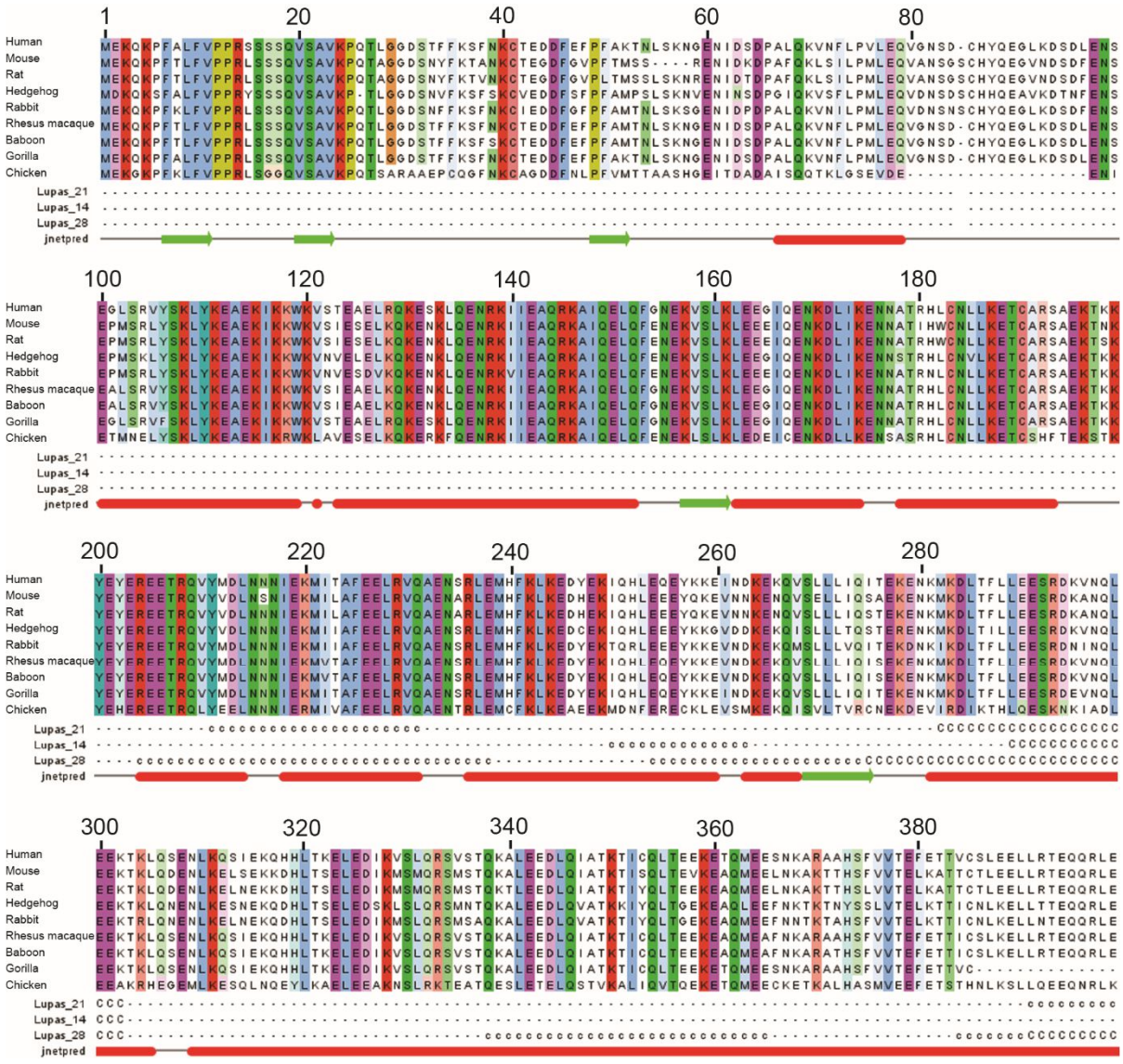


Figure 3.3. Sequence analysis of the SYCP1 N-terminus. SYCP1 multiple sequence alignment was produced by MUSCLE and visualised in Jalview. Secondary structure predictions were made by JNET with the jnetpred results displayed beneath the alignment. Alpha helices are depicted as red rods, beta sheets are indicated as green arrows and grey lines represent unstructured regions. Lupas displays the propensity for coiled-coil formation, calculated by the COILS server.

3.1.3 SYCP1 head-to-head interactions form a recursive preliminary SC lattice

As the N-termini of SYCP1 are located within the central element, it is thought that this region may form interactions with the central element components. The N-terminal third of SYCP1, amino acids 1-362, contains the highly conserved patch at the start of the helical core and the unstructured N-terminus (Figure 3.3). The 101-362 region has been extensively characterised *in vitro* through biophysical and structural analysis of SYCP1 N-terminal fragments by other members of the Davies group.

The X-ray crystal structure of the N-terminal region of the SYCP1 helical core was solved, demonstrating a head-to-head interaction to form an elongated tetramer structure (Dunce *et al*., 2018). Two fragments were crystallised for structure solution, 101-206 (PDB: 6F62) and 101-175 (PDB: 6F5X), with both solutions forming the head-to-head tetramerisation between opposing coiled coil dimers (Figure 3.4 A and B). Through biophysical analysis of the SYCP1 101-362 region in solution, large assemblies were detected. The analysis of this region upon mutation, whereby the V105 and L109 residues in the head-to-head interacting region were mutated to glutamates, found that this region no longer assembles and instead forms a stable tetramer (Dunce *et al*., 2018). Therefore, a mechanism of SYCP1 central region assembly was suggested whereby the head-to-head interactions take place recursively with dimeric SYCP1 N-termini spanning from opposing chromosomal axes.

By combining the analysis of the SYCP1 N- and C-terminal self-assembly mechanisms, a molecular model has been proposed for the formation of a preliminary SC lattice by SYCP1 (Figure 3.5). This model enables the continuous recruitment and iterative self-assembly of SYCP1 at the chromosome axis that enables the formation of a continuous lattice through staggered N-terminal assemblies (Dunce *et al*., 2018). The recruitment and assembly of the central element proteins is required for SC maturation, and therefore function, and so this SYCP1 lattice is likely stabilised by central element assembly.

3.1.4 The central element component SYCE3 forms a compact dimeric helical bundle

SYCP1 forms the preliminary SC lattice, but all central element proteins are required for SC maturation and function. Therefore, further understanding is necessary to map the network of interactions that occur with the central element and to define the association with the transverse filaments. In this study, a direct interaction between SYCE3 and SYCP1 has been extensively characterised, consistent with previous research whereby SYCP1 and SYCE3 were found to interact by co-immunoprecipitation assays (Hernandez-Hernandez *et al*., 2016). Further interactions have also been characterised between SYCE3 and the central element proteins SYCE1 and SYCE2, supporting previous findings (Schramm *et al*., 2011).

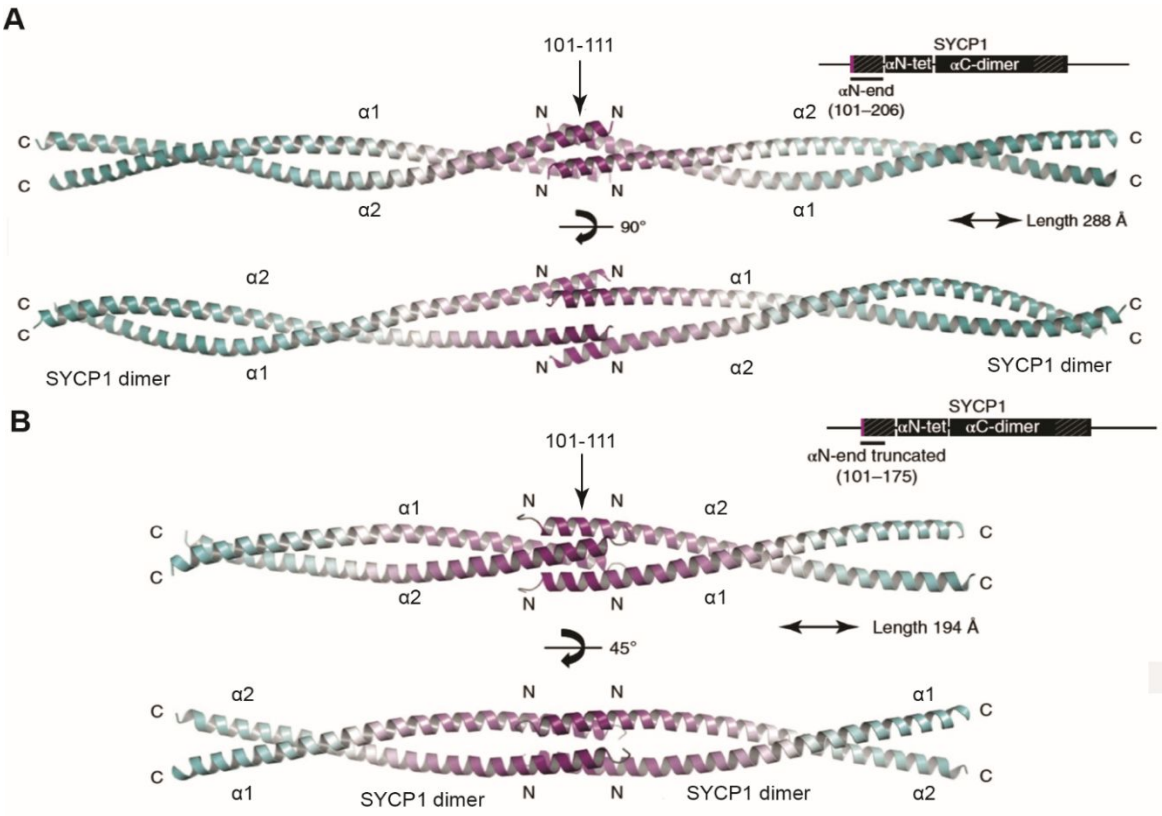


Figure 3.4. SYCP1 N-terminal assembly. (A-B) X-ray crystal structures of SYCP1 N-terminal regions. An alpha helical tetramer is formed through a head-to-head interaction of two parallel dimeric coiled coils, governed by SYCP1 amino acid residues 101-111. Alpha helices comprising the dimer are labelled α1 and α2, respectively. (A) SYCP1 101-206 region (PDB code: 6F62). The C-termini are in an open conformation. (B) SYC1 101-175 region (PDB code: 6F5X). The C-termini are in a closed conformation. Taken from Dunce *et al.*, 2018.

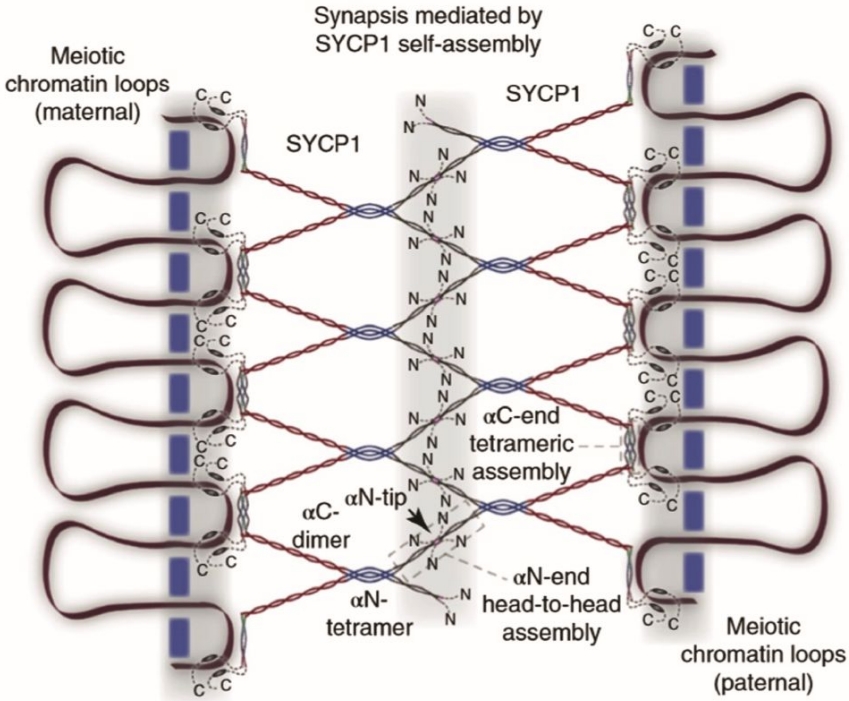


Figure 3.5. SYCP1 self-assembles to form the backbone of the SC. Schematic representing the N-terminal and C-terminal recursive self-assembly within the central element and at chromosome axes respectively. Taken from Dunce *et al,*. 2018.

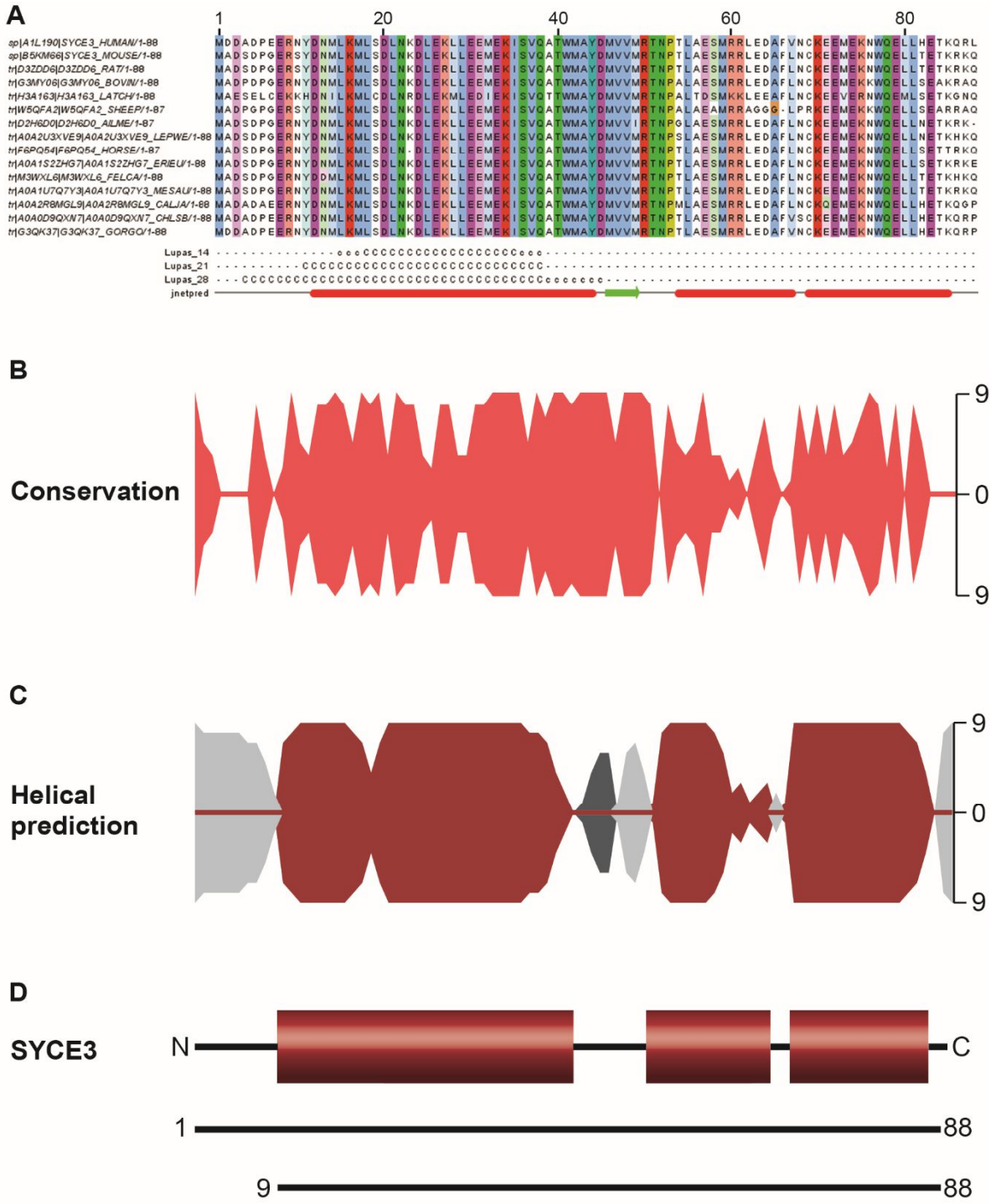


Figure 3.6. Sequence analysis of SYCE3. (A) SYCE3 multiple sequence alignment was produced by MUSCLE and visualised in Jalview. Secondary structure predictions were made by JNET with the jnetpred results displayed beneath the alignment. Alpha helices are depicted as red rods, beta sheets are indicated as green arrows and grey lines represent unstructured regions. Lupas displays the propensity for coiled-coil formation, calculated by the COILS server. Key residues used in this study have been highlighted. (B) SYCE3 conservation mirror plot. Clustal multiple sequence alignments were analysed by ConSurf to grade the conservation (0-9) of each amino acid position.(C) Secondary structure prediction of SYCE3 determined by JNET. Each residue is plotted as alpha helix (red), beta sheet (dark grey) or unstructured (light grey) according to the confidence of the prediction (0-9). (D) Linear schematic representation of SYCE3. The red boxes represent the helical regions and the black lines represent the unstructured regions.

The central element component SYCE3 is a small 88 amino acid protein that is highly conserved, particularly in the N-terminal half of the sequence (Figure 3.6 A and B). SYCE3 is predicted to be predominantly alpha helical, with small unstructured patches, particularly at the immediate N-terminal region (Figure 3.6 C). The SYCE3 molecule appears to have two helical

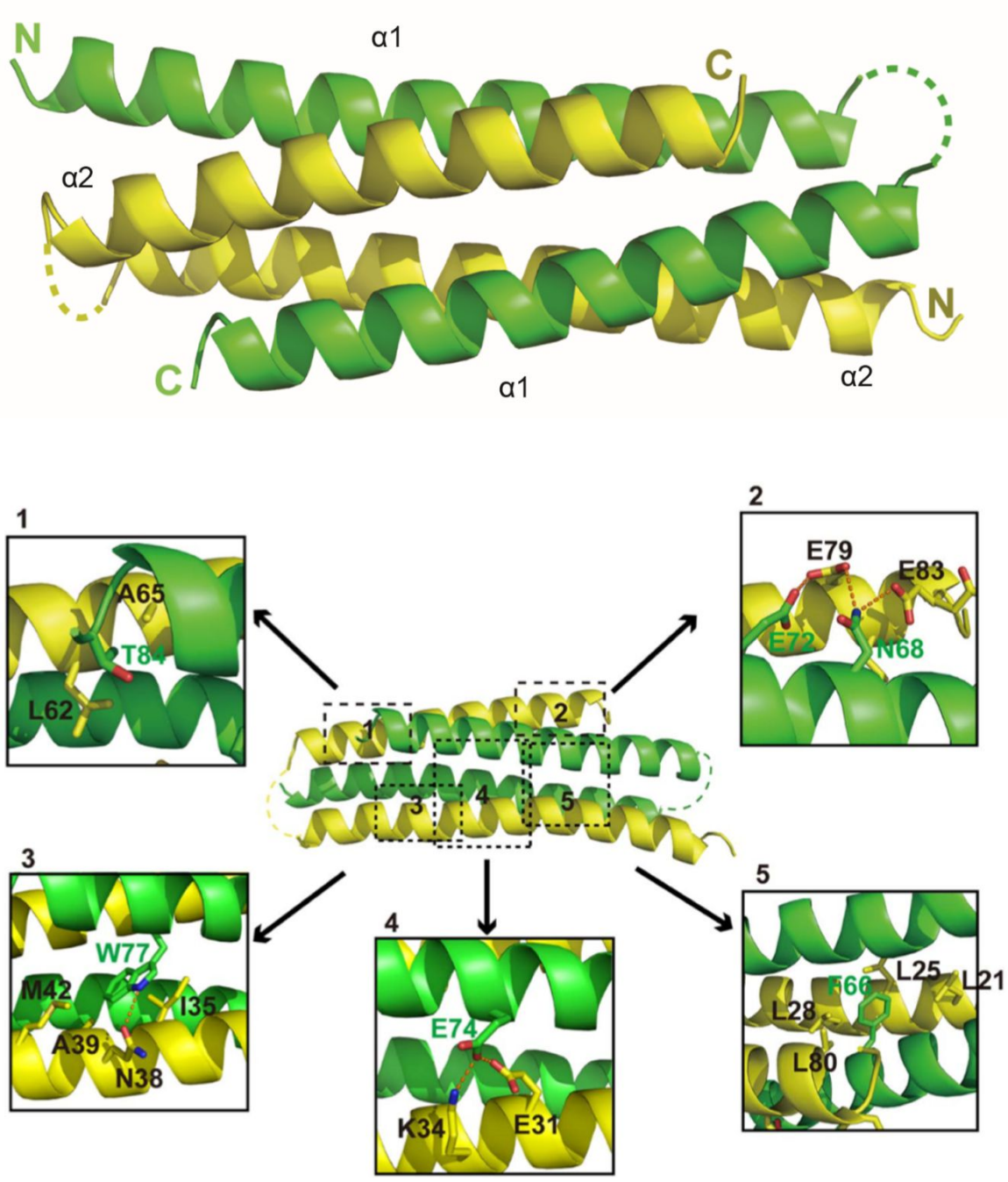


Figure 3.7. X-ray crystal structure of the mouse SYCE3. SYCE3 forms a compact dimeric structure through the folding back of each chain to form a four helical bundle through the interlocking of antiparallel SYCE3 molecules (PDB: 4R3Q). Alpha helices are labelled α1 (green) and α2 (yellow), respectively. Taken from Lu *et al.*, 2014.

patches in the N- and C-terminal halves respectively, with the second patch potentially having a short disruption (Figure 3.6 C and D). Analytical ultra-centrifugation analysis of the mouse SYCE3 recombinant protein showed that SYCE3 forms dimers and crosslinking displayed capacity to form multimers (Lu *et al*., 2014).

The X-ray crystal structure of the mouse SYCE3 protein has been solved to reveal a compact dimeric alpha helical structure (Lu *et al*., 2014). The structure shows that the SYCE3 chain folds back on itself to inter-lock with an antiparallel chain to form a dimeric four helical bundle structure (Figure 3.7).

The compact helical bundle structure is stabilised by a series of interactions between the four chains from the two protomers (Figure 3.7). Hydrophobic interactions are formed between T84 from one molecule and L62 and A65 from the other molecule, and a hydrophobic network is formed between the N68 and E72 residues from one protomer and the E79 and E83 residues from the other protomer. The W77 residue from one molecule is located within a hydrophobic patch from the other molecule comprised of amino acid residues I35, A39, N38 and M42 and the P66 residue from one protomer is buried in a hydrophobic cleft formed by L21, L25, L28 and L80 residues from the other protomer. The central stabilising interaction holding the compact structure together is the formation of a salt bridge between the E74 residue from one molecule and the K34 residue from the other protomer. The E74 residue also forms a hydrogen bond with the E31 residue from the other SYCE3 molecule in order to stabilise the SYCE3 dimeric four helical bundle (Lu *et al*., 2014).

Studies were initiated to investigate the link between the transverse filaments and the central element. SYCE3 was identified as the only central element protein to interact with SYCP1. The characterisation of the SYCP1-SYCE3 complex reveals the remodelling of the SYCP1 N-terminal tetramer and the SYCE3 dimer to result in the formation of a 2:1 complex. This 2:1 complex has been found to undergo head-to-head interactions via the SYCP1 101-111 region. The capacity for SYCE3 to form assemblies has been characterised in solution to be driven by loop opening and stabilisation by a combination of surface exposed residues. These SYCE3 higher-order species are found to bind to the SYCP1-SYCE3 complex to form larger assemblies, which provides a mechanism for SYCE3 central element incorporation and remodelling of the SYCP1 preliminary lattice.

This work has been completed as part of a collaborative project with a post-doctoral research associate, Dr Orla Dunne, with some assistance by a PhD student James Dunce. Several experiments presented in this thesis chapter were performed by Orla Dunne and James Dunce, and have been indicated accordingly throughout.

3.2 Results

3.2.1 SYCE3 is the only central element protein to interact with the transverse filament

SYCP1 has been shown to interact with SYCE3 through immuno-precipitation studies (Hernandez-Hernandez *et al*., 2016). However, it remains unclear whether SYCP1 undergoes further interactions with other SC components for central element incorporation. To investigate this, yeast 2-hybrid screening was performed with SYCP1 and the central element proteins: SYCE1, SYCE2, SYCE3, TEX12 and SIX6OS1. For this analysis, a C-terminal SYCP1 truncation was tested (residues 1-811) as the full length 976 amino acid sequence was too long for the Y2H system.

The Y2H screen demonstrated that SYCP1 and SYCE3 directly interact and that no other central element components were detected to bind to SYCP1 (Figure 3.8 A). The screen also showed that SYCP1 and SYCE3 undergo self-interactions, and that mild interactions were detected between SYCE3 with both SYCE1 and SYCE2 (Figure 3.8 A).

SYCE3 has been identified as the only central element component to directly interact with SYCP1, and therefore, provides the physical link between the transverse filament and the central element. To study this interaction biochemically, a recombinant protein pull-down assay was performed with SYCE3 and various SYCP1 fragments. The SYCP1 fragments were fused to an MBP affinity tag and SYCE3 was fused to a His_6 affinity tag for co-expression in *E. coli*. The clarified bacterial lysates were incubated with amylose resin to bind the MBP tag, and following washing of the resin, any bound protein was eluted by maltose to out-compete the amylose resin for MBP binding. The resultant eluates were standardised for protein concentration and analysed by SDS-PAGE.

As the full length SYCP1 is not soluble in bacteria, an N- and C-terminal truncation was used that comprises the central helical core of SYCP1 but removes the unstructured ends, corresponding to amino acids 112-783. This fragment was indeed found to bind to SYCE3, and suggests that the interaction does not require the unstructured N-terminus of SYCP1 (Figure 3.8 B). As the N-terminal portion of SYCP1 is localised within the central region of the SC, it was reasoned that the SYCE3 binding would occur at the N-terminal end of SYCP1. The SYCP1 N-terminal third, 1-362, was found to bind SYCE3, and further truncations to remove the unstructured N-terminus (101-362) and the head-to-head assembly site (112-362) also maintained SYCE3 binding (Figure 3.8 B). To narrow down the SYCP1 interaction site, two more truncations were tested, 176-362 and 206-362, and both fragments interacted with SYCE3 (Figure 3.8 B). Therefore, the SYCP1-SYCE3 interaction occurs in the N-terminal structured region of SYCP1, suggesting the unstructured end has other functions in the central element.

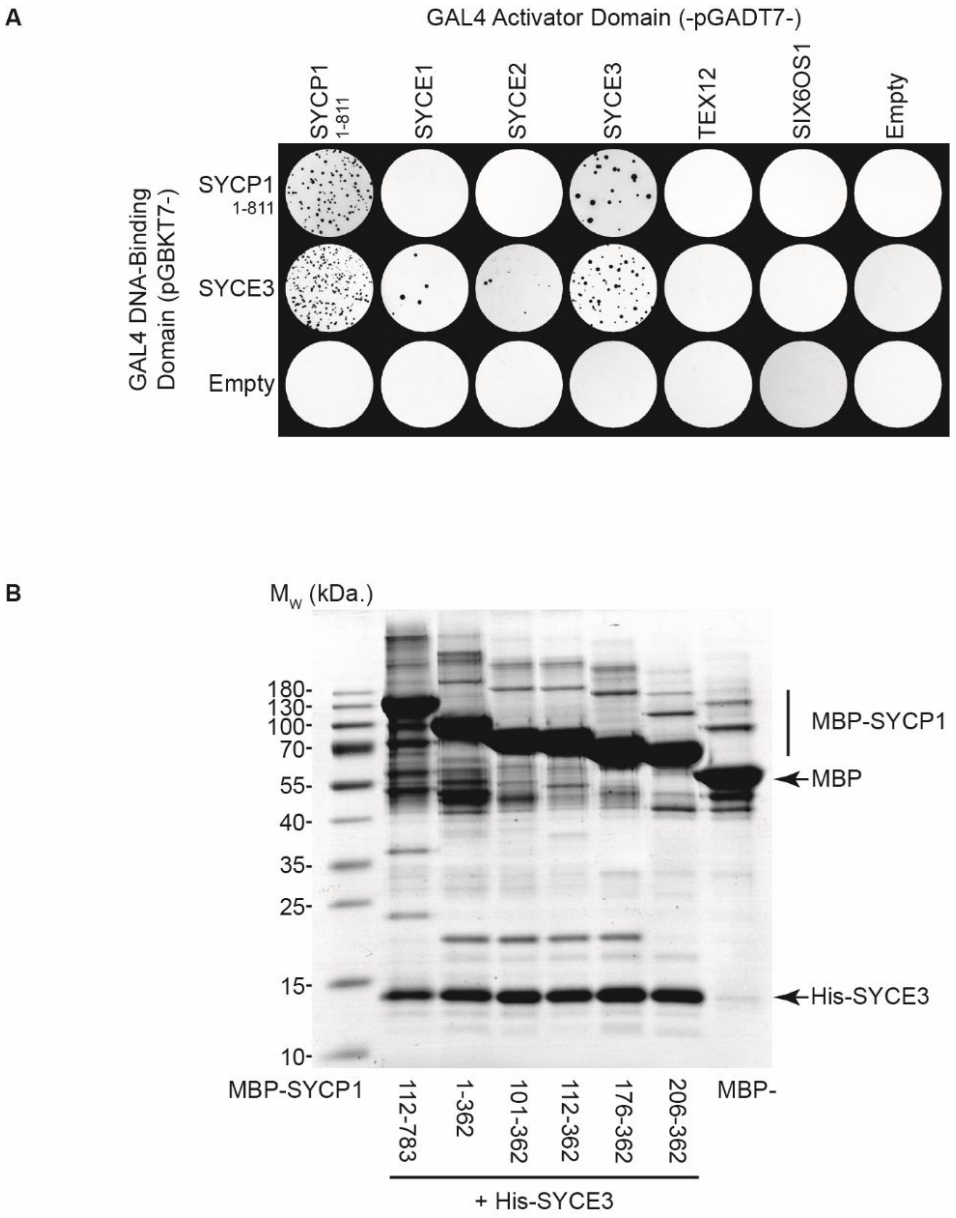


Figure 3.8. Screening for SC central region interactions. (A) Yeast two-hybrid grid with SYCP1 and SYCE3 against SYCP1, SYCE1, SYCE2, SYCE3, TEX12 and SIX6OS1. All sequences correspond to the full-length human proteins, apart from SYCP1 that was tested through a C-terminal truncation (1-811) as the full length 976 amino acid sequence is long for the system. Y187 (pGBKT7-bait) strains were mated with Y2H Gold (pGADT7-target) strains and plated on SD/-Leu/-Trp/-Ade/-His/x-α-gal plates for colony selection, subject to reporter gene activation (ADE1, HIS3 and MEL1) upon positive interactions. (B) Amylose pull-down assay. SYCP1 fragments were fused to a His-MBP affinity tag, including free His-MBP as a control, for bacterial co-expression with His-SYCE3, with clarified lysates incubated with amylose resin. Following washing of the resin, bound protein was eluted and eluates were normalised for protein concentration for analysis by SDS-PAGE. SYCE3 was found to interact with each SYCP1 fragment, but not the free MBP control.

3.2.2 SYCP1 and SYCE3 undergo conformational changes upon binding

The SYCP1-SYCE3 interaction has been isolated to the 206-362 region of SYCP1. This region has been shown to exist as a parallel tetramer *in vitro* (Dunce *et al*., 2018) and the crystal structure of SYCE3 shows that it forms a compact dimeric four helical bundle (PDB code: 4R3Q). To characterise the SYCP1-SYCE3 complex *in vitro*, the $SYCP1_{206-362}$ fragment was fused to a His_6-MBP affinity tag and SYCE3 was fused to a His_6 affinity tag for bacterial co-expression. The fusion protein complex was purified through sequential Ni-NTA and amylose affinity chromatography and subsequent anion exchange chromatography. The affinity tags were removed by TEV protease incubation and the cleaved protein complex was further purified by anion exchange and size exclusion chromatography steps (Figure 3.9 A). The cleaved material appeared to be extremely stable and a high yield of material could be obtained for analysis, at concentrations of 60 mg/ml and higher.

To study the secondary structure composition of the SYCP1-SYCE3 complex CD wavelength scans were recorded and converted to mean residue ellipticity (MRE). The scan shows a predominant helical signal, with deconvolution estimating the helical content to be 92% (Figure 3.9 B). SYCE3 is a small protein that is mostly helical, and the $SYCP1_{206-362}$ region has been previously shown to also be predominantly helical (Dunce *et al*., 2018). Therefore, the estimated 92% helical content of the $SYCP1_{206-362}$-SYCE3 complex demonstrates that the protein folding has not been disrupted upon binding.

In order to measure the thermal stability of $SYCP1_{206-362}$-SYCE3, CD thermal denaturation was performed between 5°C and 95°C by measuring the 222 nm helical signal and plotted as % unfolded. The melting curve shows a sharp cooperative unfolding event with an estimated melting temperature, determined as the point at which 50 % of the sample is unfolded, of 38°C (Figure 3.9 C). At 50°C, a further gradual unfolding event is detected, which may represent residual unwinding of helices following complex dissociation.

To determine the oligomeric state of the $SYCP1_{206-362}$-SYCE3 purified complex, SEC-MALS was utilised to calculate the absolute molecular weight of the sample. From this analysis, a single protein peak was detected, with the molecular mass of the protein species calculated as 47 kDa (Figure 3.9 D and E). The theoretical molecular weight of a 2:1 $SYCP1_{206-362}$:SYCE3 complex is estimated to be 49 kDa, thus indicating that a 2:1 complex is being formed by $SYCP1_{206-362}$-SYCE3. This implies that both $SYCP1_{206-362}$ and SYCE3 undergo conformational changes upon binding, as isolated $SYCP1_{206-362}$ forms a tetramer (Dunce *et al*., 2018) and SYCE3 forms a dimer (Lu *et al*., 2014).

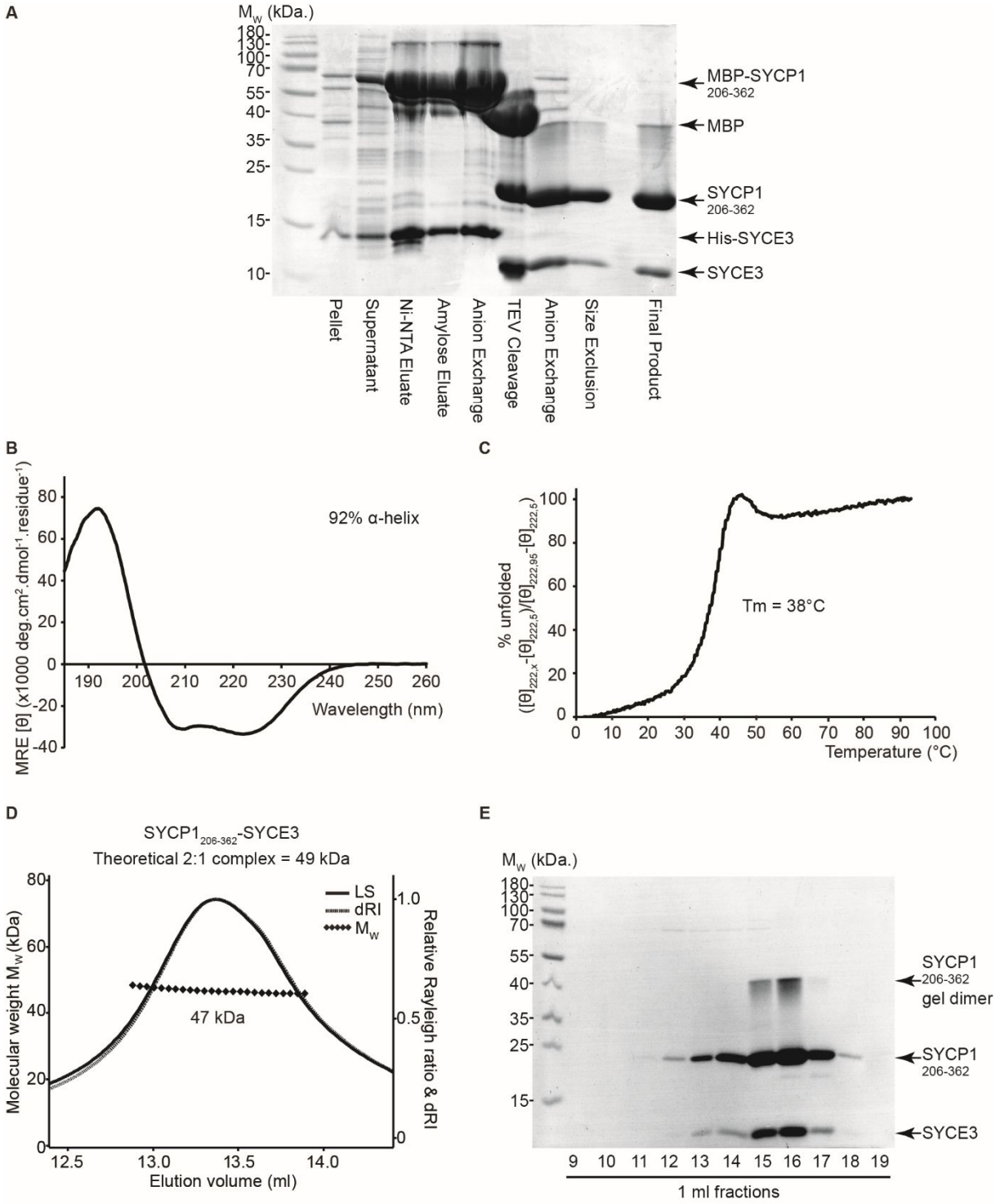


Figure 3.9. Purification and analysis of SYCP1$_{206-362}$-SYCE3. (A) SDS-PAGE summary of SYCP1$_{206-362}$-SYCE3 purification by Ni-NTA, amylose and anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange and size exclusion chromatography. (B) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 92% alpha-helix, 1% β-sheet, 3% turns and 4% unordered. (C) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 38°C, when the sample is 50% unfolded. (D) SEC-MALS analysis. A single 47 kDa species was formed, with the theoretical molecular weight of a 2:1 complex estimated to be 49 kDa. (E) SDS-PAGE analysis of the SEC-MALS output 1ml fractions, demonstrating the co-elution profile of the SYCP1$_{206-362}$-SYCE3 complex. A 40 kDa band elutes with the complex. This band is likely a gel dimer formed by SYCP1 due to the size of the band and the co-elution with the complex.

To study the solution structure of the SYCP1$_{206-362}$-SYCE3 complex, SEC-SAXS experiments were performed. The SEC-SAXS data of the single 2:1 species allowed the calculation of the radius of gyration, Rg, which was determined as 68 Å and so indicating an elongated structure is formed (Figure 3.10 A). The radius of gyration of the cross-section, Rc, was also measured and determined to be 9.6 Å (Figure 3.10 B). The Rc for a dimeric coiled coil is regularly measured as 8 Å, with tetrameric coiled coils having a higher Rc of 10 Å. Therefore, the observed 9.6 Å Rc of SYCP1$_{206-362}$-SYCE3 is consistent with a trimeric or tetrameric coiled coil formation. The real space distance distribution plot indicates that the maximum dimension observed in the sample is 245 Å (Figure 3.10 C). This is consistent with the two SYCP1$_{206-362}$ molecules forming an elongated helical structure, with the distance roughly estimated 1.5x the residue number. *Ab initio* modelling was performed to produce the low resolution molecular envelope of SYCP1$_{206-362}$-SYCE3 from the real space analysis. The modelled envelope demonstrates the extended nature of the complex, with the formation of a long rod shape (Figure 3.10 D).

From the analysis of the cleaved SYCP1$_{206-362}$-SYCE3 complex, it has been demonstrated that an elongated hetero-trimeric helical structure is formed. However, the orientation of the SYCP1 chains is unclear, and the conformation and orientation of SYCE3 in the complex is also unknown. To address this, a series of MBP-tagged SYCP1$_{206-362}$-SYCE3 complexes were purified and analysed to probe the location of the MBP tags within the complex, in order to elucidate the organisation of the molecules.

The MBP-SYCP1$_{206-362}$-His-SYCE3 fusion protein complex was purified, as described earlier, for solution scattering experiments. SEC-MALS analysis produced a single protein species with a calculated molecular mass of 136 kDa (Figure 3.11 A and B). The theoretical molecular weight of MBP-SYCP1$_{206-362}$-His-SYCE3 is estimated to be 142 kDa, further indicating the formation of a 2:1 complex. SEC-SAXS analysis determines the radius of gyration to be 63 Å and real space analysis estimates the maximum dimension of the sample to be 275 Å, with the extra 30 Å due to the addition of the MBP tags (Figure 3.11 C and D). Due to the globular nature of the MBP affinity tags, the paired distance distribution plot highlights the distance between the MBP tags at 68 Å (Figure 3.11 D). This observed distance indicates that the N-termini are within close proximity, and therefore, suggests that the SYCP1$_{206-362}$ chains are in a parallel nature.

SYCE3 is a small protein with an estimated molecular weight of 10 kDa and so SYCP1$_{206-362}$ was analysed with SYCE3 fused to an MBP affinity tag. The molecular mass of the single species formed is measured as 88 kDa by SEC-MALS, with the estimated molecular

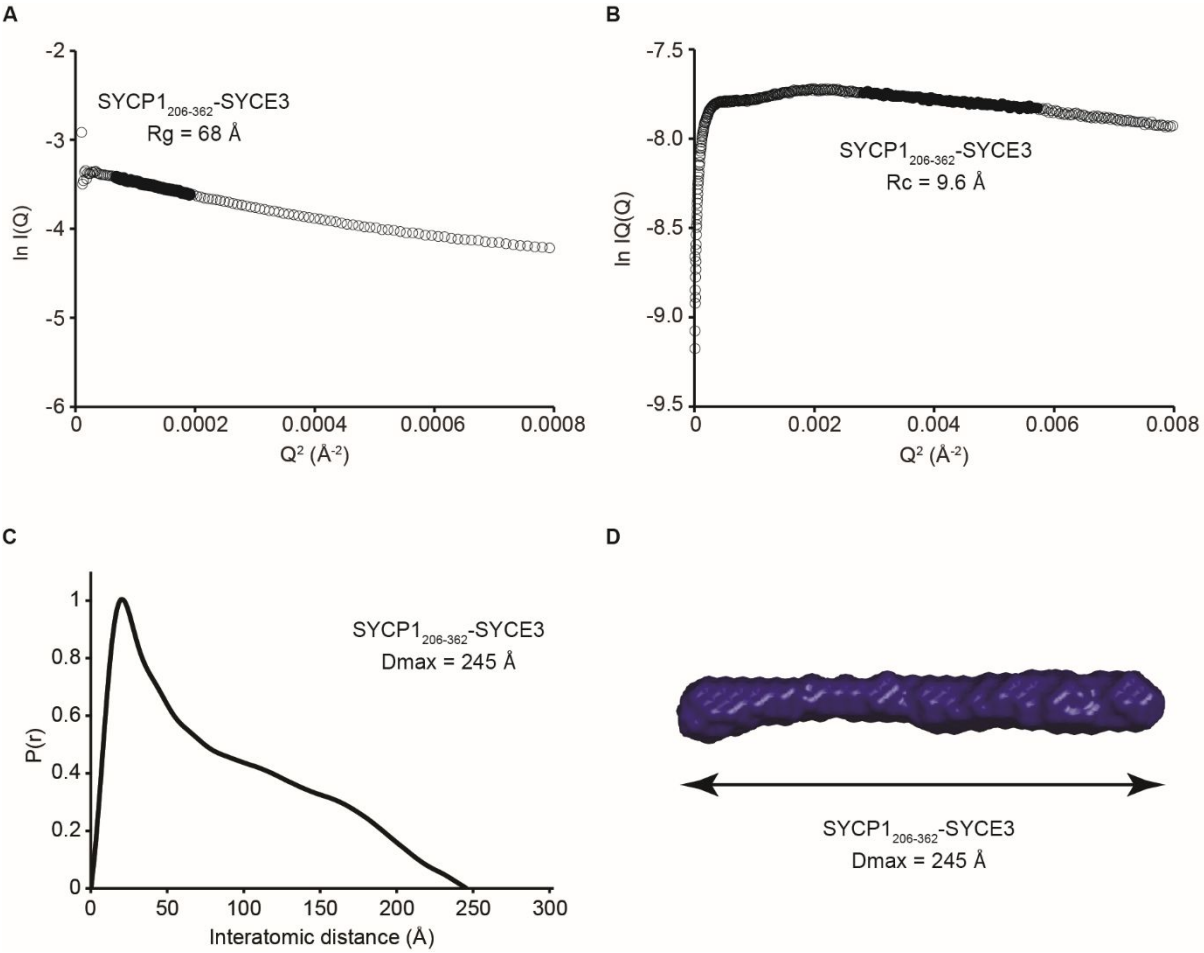


Figure 3.10. SEC-SAXS analysis of SYCP1$_{206-362}$-SYCE3. (A) Guinier analysis to determine the radius of gyration (Rg) as 68 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 9.6 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. SYCP1$_{206-362}$-SYCE3 has a maximum dimension of 245 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

weight of 90 kDa for a 2:1 complex and therefore, further validating that there is one SYCE3 molecule binding to two SYCP1$_{206-362}$ chains (Figure 3.12 A and B). SEC-SAXS analysis determined the Rg to be 61 Å and the real space analysis estimates the maximum dimension to be 280 Å (Figure 3.12 C and D). These values are similar to those measured for MBP-SYCP1$_{206-362}$-His-SYCE3, but with the absence of the inter-MBP peak that further iterates that only one SYCE3 molecule is present in the complex.

To gain understanding about the orientation of SYCE3 in relation to the two SYCP1$_{206-362}$ parallel chains, the complex was analysed with an MBP-tag fused to both SYCP1$_{206-362}$ and SYCE3. The molecular weight of the single protein species was calculated as 164 kDa by SEC-MALS, with the theoretical 2:1 complex mass of 179 kDa, again indicating the formation of a

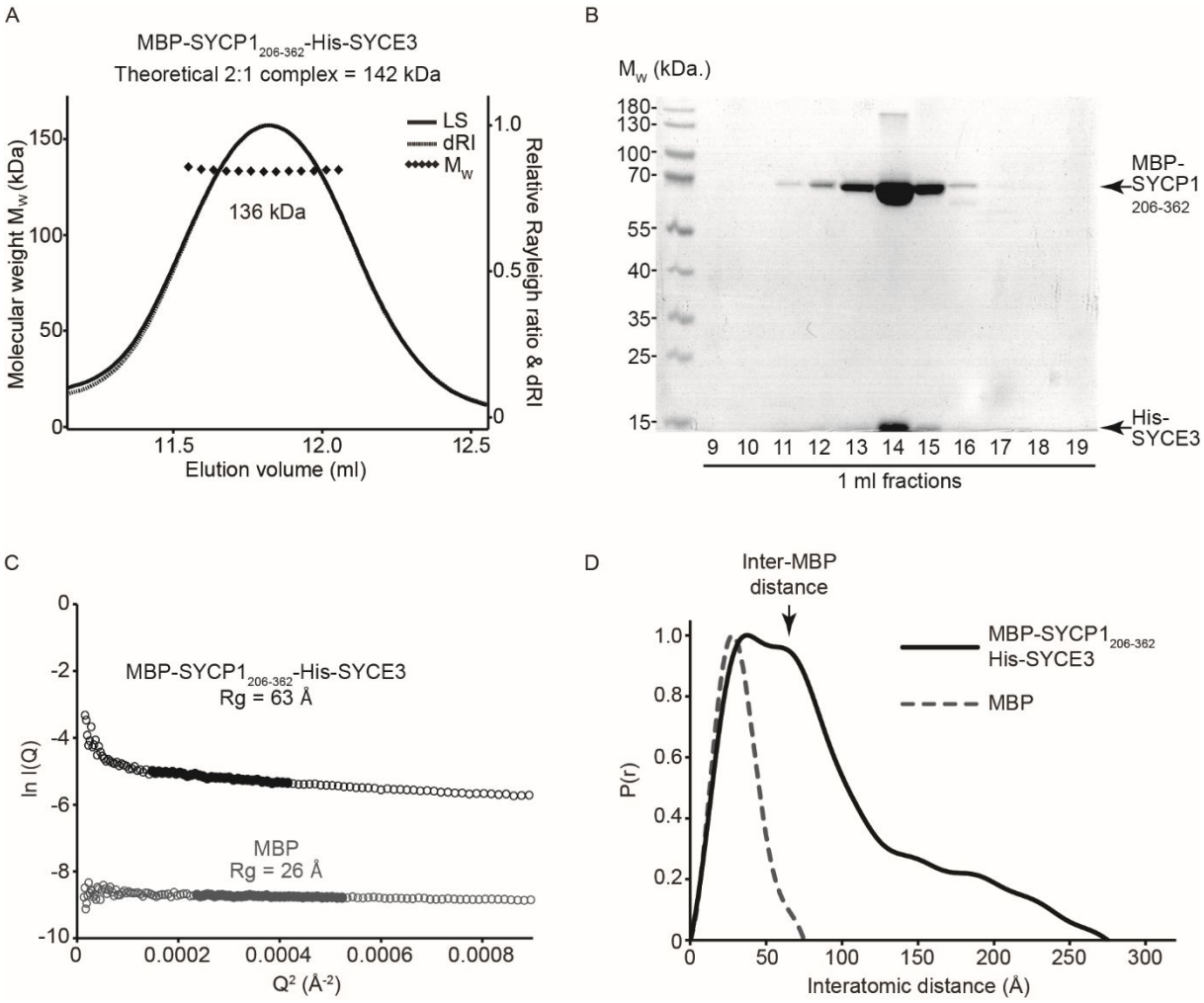


Figure 3.11. Structural analysis of MBP-SYCP1$_{206-362}$-His-SYCE3. (A) SEC-MALS analysis showing a single 136 kDa peak. A theoretical 2:1 complex has an estimated 142 kDa molecular weight suggesting a 2:1 complex formation. (B) SDS-PAGE analysis of the SEC-MALS output 1ml fractions, demonstrating the co-elution profile of the complex. (C) Guinier analysis to determine the radius of gyration (Rg) as 63 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. MBP-SYCP1$_{206-362}$-His-SYCE3 has a maximum dimension of 275 Å, with an inter-MBP peak at 68 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

2:1 complex that can be observed through SDS-PAGE analysis of the SEC-MALS elution fractions (Figure 3.13 A and B). SEC-SAXS determined the Rg of MBP-SYCP1$_{206-362}$-MBP-SYCE3 to be 67 Å and the distance distribution plot estimates the maximum dimension within the sample to be 300 Å, with the extra distance due to the inclusion of an MBP tag on all three molecules (Figure 3.13 C and D). The paired real space analysis also reveals a peak that correlates to the distance between the globular MBP tags, suggesting that all N-termini from the three molecules are in close proximity, and therefore, suggests that SYCE3 is in a parallel arrangement to the SYCP1$_{206-362}$ chains (Figure 3.13 D).

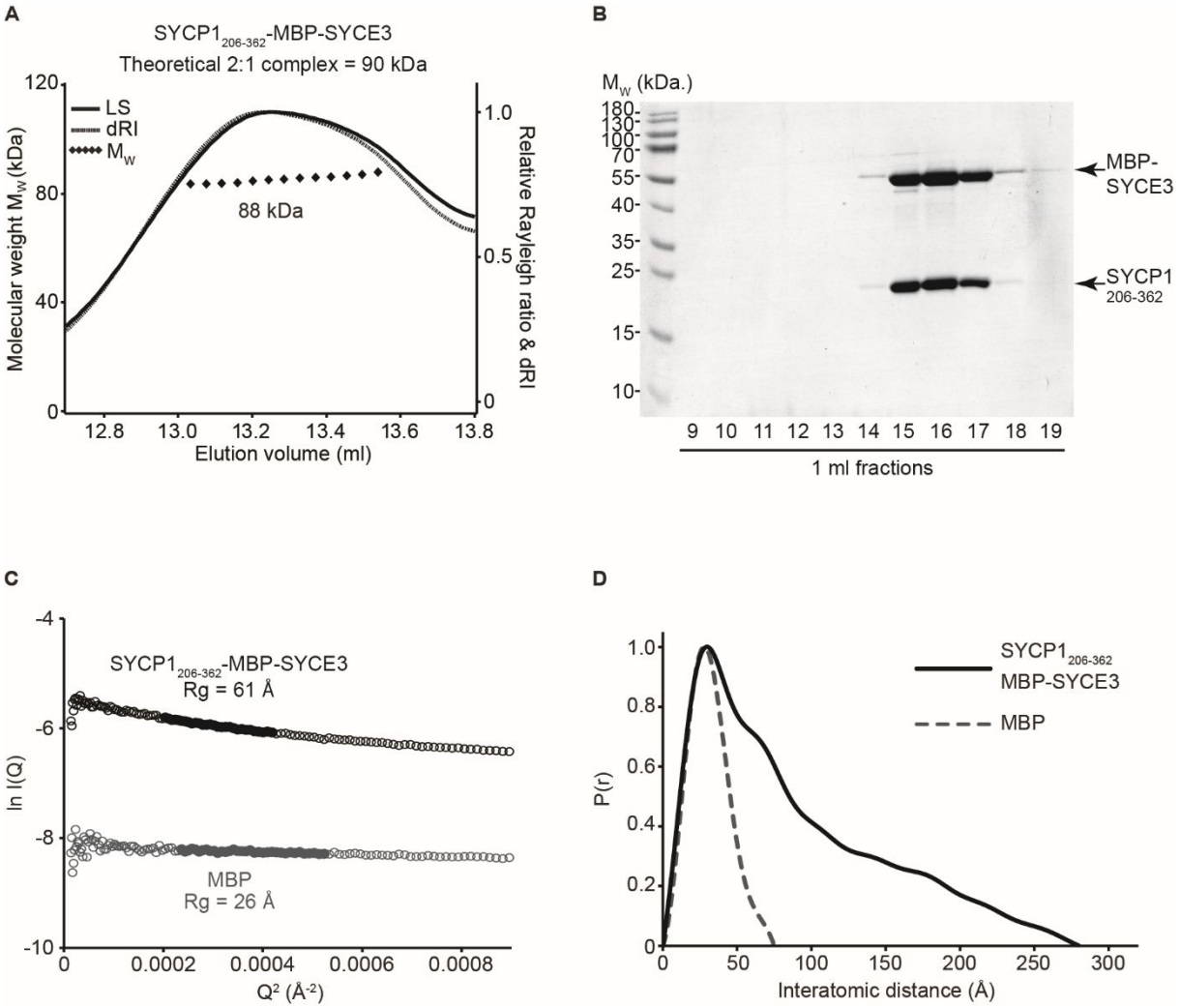


Figure 3.12. Structural analysis of SYCP1$_{206-362}$-MBP-SYCE3. (A) SEC-MALS analysis showing a single 88 kDa peak. A theoretical 2:1 complex has an estimated 90 kDa molecular weight suggesting a 2:1 complex formation. (B) SDS-PAGE analysis of the SEC-MALS output 1ml fractions, demonstrating the co-elution profile of the 2:1 complex. (C) Guinier analysis to determine the radius of gyration (Rg) as 61 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. SYCP1$_{206-362}$-MBP-SYCE3 has a maximum dimension of 280 Å (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

The analysis of the recombinant SYCP1$_{206-362}$-SYCE3 complex in solution has elucidated the formation of an elongated helical hetero-trimeric complex with the chains organised in a parallel arrangement. However, in order to study the precise molecular basis of the SYCP1$_{206-362}$-SYCE3 complex, a high resolution structure is required. To address this, crystallographic studies were initiated with SYCP1$_{206-362}$-SYCE3, with the aim of solving the X-ray crystal structure of the complex.

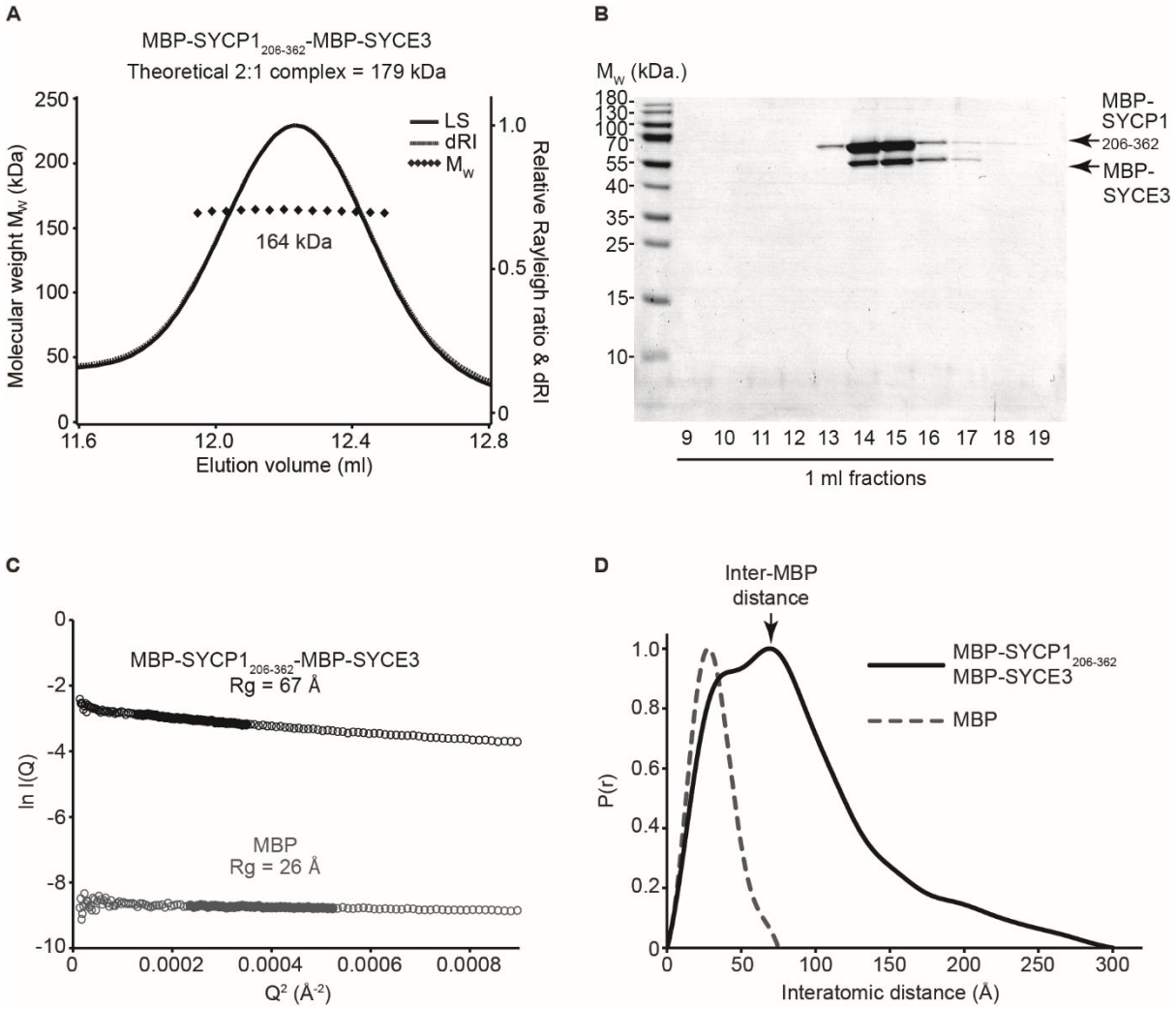


Figure 3.13. Structural analysis of MBP-SYCP1$_{206-362}$-MBP-SYCE3. (A) SEC-MALS analysis showing a single 164 kDa peak. A theoretical 2:1 complex has an estimated 179 kDa molecular weight suggesting a 2:1 complex formation. (B) SDS-PAGE analysis of the SEC-MALS output 1ml fractions, demonstrating the co-elution profile of the 2:1 complex. (C) Guinier analysis to determine the radius of gyration (Rg) as 67 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. MBP-SYCP1$_{206-362}$-MBP-SYCE3 has a maximum dimension of 300 Å, with an inter-MBP at 69 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

As large concentrations of the complex were obtained, a wide range of protein concentrations and commercial screen conditions were tested for crystal growth. One condition gave rise to crystalline needle clusters with no other hits detected (Figure 3.14 A). As this was the only condition that exhibited any crystal growth, hanging drop optimisation screens were performed to attempt to improve crystal growth and quality. Screening involved varying the buffer component concentrations, protein concentration and pH. The optimisation screening

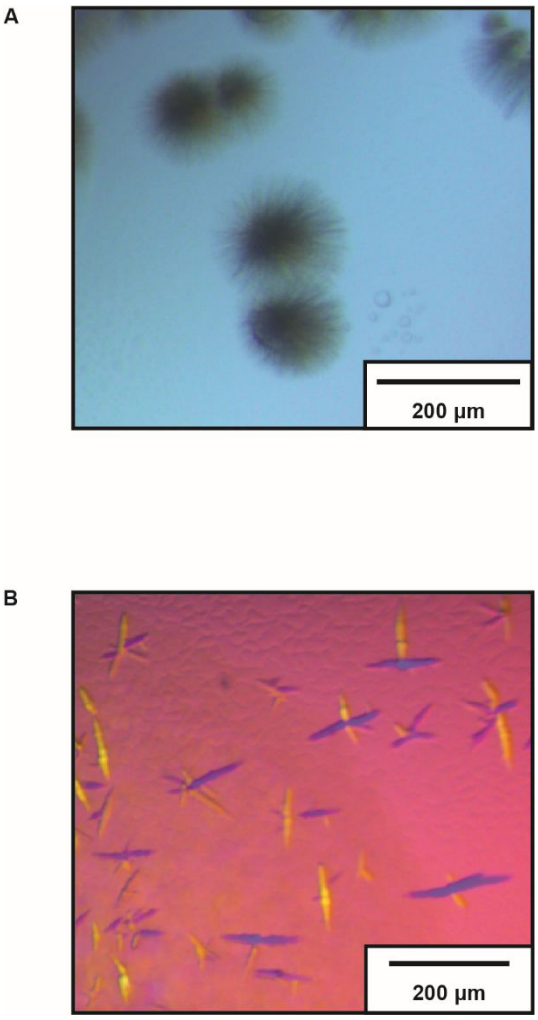


Figure 3.14. Crystallisation of SYCP1$_{206-362}$-SYCE3. (A) Crystalline needle clusters grown in 0.1 M sodium cacodylate pH 6.5, 40% v/v MBP, 5% w/v PEG 8000, from the commercial JCSG+ screen. Crystallisation was achieved through vapour diffusion sitting drops with protein at 37 mg/ml. (B) Optimisation screening to form larger needles. Optimisation was performed by manual screening of vapour diffusion hanging drops in 0.1 M sodium cacodylate pH 7.0, 44% v/v MBP, 6% w/v PEG 8000, with protein at 37 mg/ml.

yielded needle crystal growth within a fine MPD range (Figure 3.14 B). However, upon crystal harvesting to test for X-ray diffraction, the crystals were found to be adhered to the cover slip and heavily agglutinated.

In order to address these issues, further optimisation attempts were performed. Firstly, an additive screen was performed whereby the optimised condition was used for crystallisation, but with the addition of additives from a commercial screen. The additive screening did not yield any preferential crystal growth, and therefore, crystals had no detectable improvement in quality. Further optimisation attempts were made by macro-seeding and micro-seeding. The macro-seeding involved the transfer of the smaller un-adhered needles from a heavily

crystalline drop to a pre-equilibrated drop with fewer crystals, with the aim to continue growth of the needles. This method did not improve crystal growth and therefore, micro-seeding was performed that involved the formation of a crystal seed stock from crushing existing crystals. These crystals were used to set up fresh crystallisation drops, with the aim to trigger crystal formation from the seeds to grow larger and better quality crystals. This method also did not improve crystal growth, which was perhaps the consequence of the low quality of the original crystals. A final attempt to aid crystallisation was made by altering the pH of the protein stock before commercial screening. The only previous crystal growth occurred at pH 6.5, and therefore, the protein stock was altered to pH 6.5 for subsequent commercial screening. No crystallisation was detected from the screening with the lower pH protein stock. After extensive crystallisation efforts, it is possible that construct optimisation may be required to isolate a suitable unit for crystallisation.

3.2.3 The SYCP1 head-to-head interaction is maintained in the SYCP1-SYCE3 complex

The study of the $SYCP1_{101-175}$ fragment revealed that, although found to be monomeric in solution, forms a tetrameric structure in the X-ray crystal structure (Dunce *et al.*, 2018). The tetramer interface is formed through a head-to-head interaction from two parallel dimers, mediated by a short stretch of amino acids, 101-112. As this head-to-head interaction provides a mechanistic insight into the self-assembly of SYCP1 in the SC central region to form the preliminary lattice, it was postulated whether this interaction is maintained when SYCP1 is bound to SYCE3.

To investigate this, the $SYCP1_{101-362}$ fragment, which encompasses the head-to-head interaction site and the SYCE3 binding site, was co-expressed with SYCE3. The SYCP1 fragment was fused to an MBP affinity tag and SYCE3 was fused to a His_6 tag. The MBP-$SYCP1_{101-362}$-His-SYCE3 complex was co-expressed in *E. coli* and co-purified through sequential Ni-NTA and amylose affinity chromatography followed by anion exchange chromatography. TEV protease incubation was performed to remove the affinity tags, with the reaction applied over Ni-NTA to remove the His_6 tag and His_6-tagged TEV protease, with further purification achieved through anion exchange chromatography (Figure 3.15 A). The final product has a specific degradation band that roughly correlates to the size of $SYCP1_{206-362}$, suggesting that $SYCP1_{206-362}$ is a stable region and that $SYCP1_{101-362}$ undergoes some degradation (Figure 3.15 A).

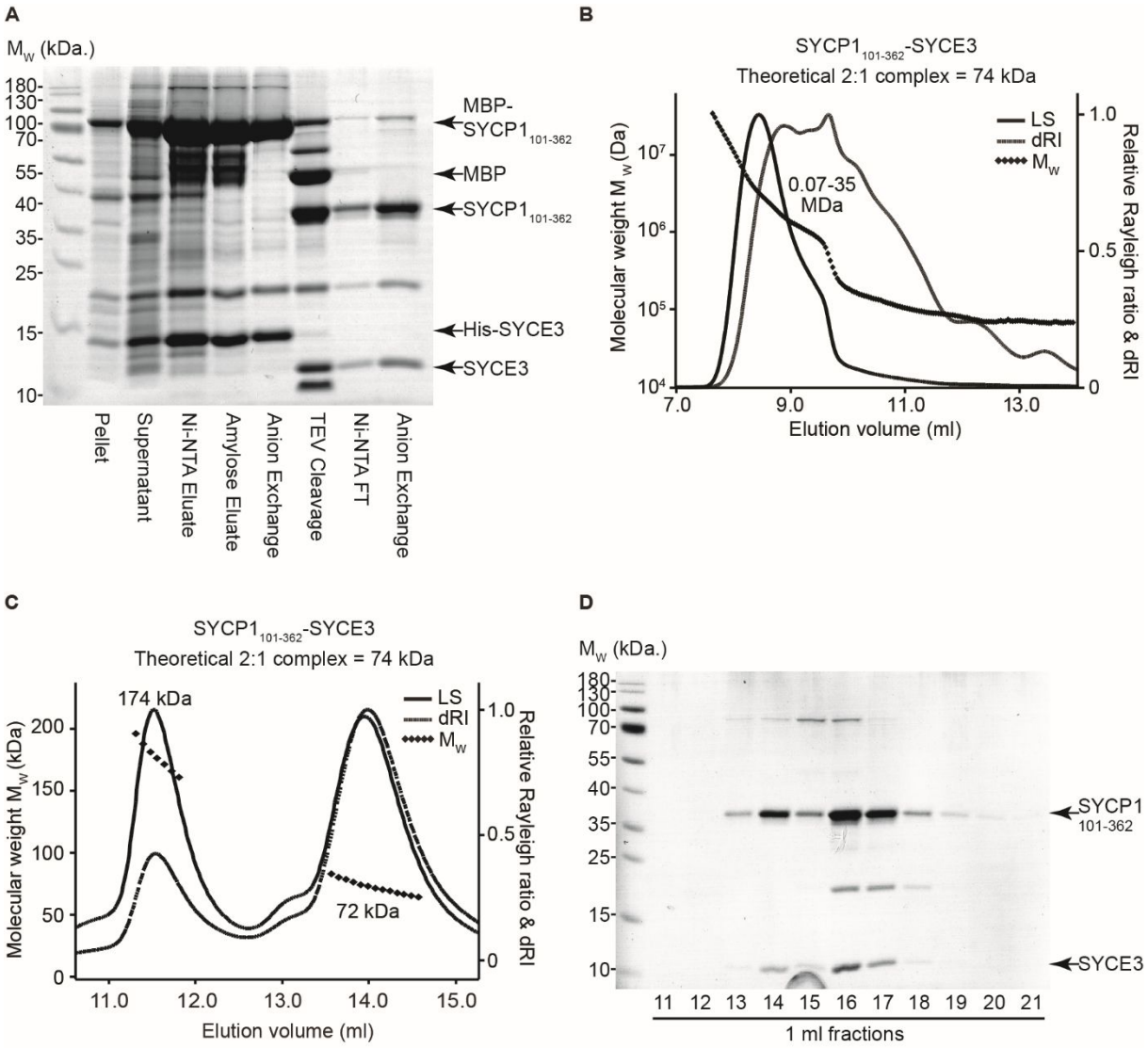


Figure 3.15. Purification and structural analysis of SYCP1$_{101-362}$-SYCE3. (A) SDS-PAGE showing the purification summary of the complex through Ni-NTA and amylose affinity chromatography, anion exchange chromatography, tag removal by TEV protease incubation and subsequent Ni-NTA and anion exchange chromatography. (B) SEC-MALS analysis with sample at 20 mg/ml showing the formation of large molecular weight species. (C) SEC-MALS analysis with sample at 2 mg/ml. The predominant species is calculated as 72 kDa with some of the material forming a 174 kDa species. The theoretical molecular weight for a 2:1 complex is 74 kDa, suggesting the formation of a 2:1 and 4:2 complex, respectively. (D) SDS-PAGE analysis of SEC-MALS output 1 ml fractions showing the co-elution profile that highlights the two species formed. The SYCP1 degradation product is only identified in the species corresponding to the 2:1 complex.

SEC-MALS was utilised to study the oligomeric status of the SYCP1$_{101-362}$-SYCE3 complex, in order to examine the difference in oligomerisation upon the inclusion of the SYCP1 101-205 residues. As a high yield of material was obtained, initial analysis was performed at 20 mg/ml. No specific oligomeric species were detected, but large molecular weight species

were formed (Figure 3.15 B). This suggests that the SYCP1 head-to-head interaction is maintained when bound to SYCE3. To test whether specific species can be detected, the SEC-MALS experiment was repeated at a lower concentration, 2 mg/ml. From this analysis, the propensity for assembly was drastically reduced and smaller protein species were isolated. The two specific protein peaks correlate to a 2:1 complex, as identified with SYCP1$_{206-362}$-SYCE3, and also a 4:2 complex (Figure 3.15 C). These species are also observed through SDS-PAGE analysis of the output fractions, which demonstrate that both SYCP1 and SYCE3 are present in both protein peaks at a 2:1 ratio (Figure 3.15 D). An indication that the 4:2 complex is driven through the SYCP1 101-205 residues is that the degradation product, which roughly correlates to the SYCP1$_{206-362}$ fragment, is not present in the 4:2 peak (Figure 3.15 D).

To study the structural properties of the 2:1 and 4:2 complexes in solution, SEC-SAXS experiments were performed. The Rg for SYCP1$_{101-362}$-SYCE3 2:1 complex was measured as 82 Å, which is larger than the Rg for SYCP1$_{206-362}$-SYCE3 due to the inclusion of the N-terminal residues, and suggests the formation of an extended structure (Figure 3.16 A). The cross-sectional radius was determined as 17 Å, which suggests a bulkier structure is formed with the additional residues (Figure 3.16 B). The real space distance distribution analysis estimates the maximum dimension within the sample to be 310 Å, with the shape of the curve indicating an elongated structure (Figure 3.16 C). *Ab initio* modelling of the SYCP1$_{101-362}$-SYCE3 complex was performed to produce the low resolution molecular envelope. This analysis demonstrates that the complex forms an elongated rod shape (Figure 3.16 D).

In order to probe the basis of the 4:2 complex formation, SEC-SAXS data was compared to the 2:1 species. The Rg of the 4:2 complex was measured as 113 Å, which is significantly larger than the Rg for the 2:1 complex and hence insinuates that the 4:2 complex has an even more extended shape than the 2:1 complex, rather than a lateral association that would be more globular (Figure 3.17 A). The cross-sectional radius for the 4:2 complex was established as 20 Å, which is approximately the same value as the 2:1 complex (Figure 3.17 B). The Rg and Rc values therefore suggest that the 4:2 complex may be formed through the SYCP1 head-to-head interactions as the cross-section is a similar size to the 2:1 complex but is potentially more elongated.

The real space distance distribution plot estimates the maximum dimension of the 4:2 complex to be 500 Å, which is considerably longer than the 310 Å maximum distance of the 2:1 complex (Figure 3.17 C). *Ab initio* modelling was carried out to generate the molecular envelope and the 4:2 complex is an elongated rod shape, which resembles a more extended

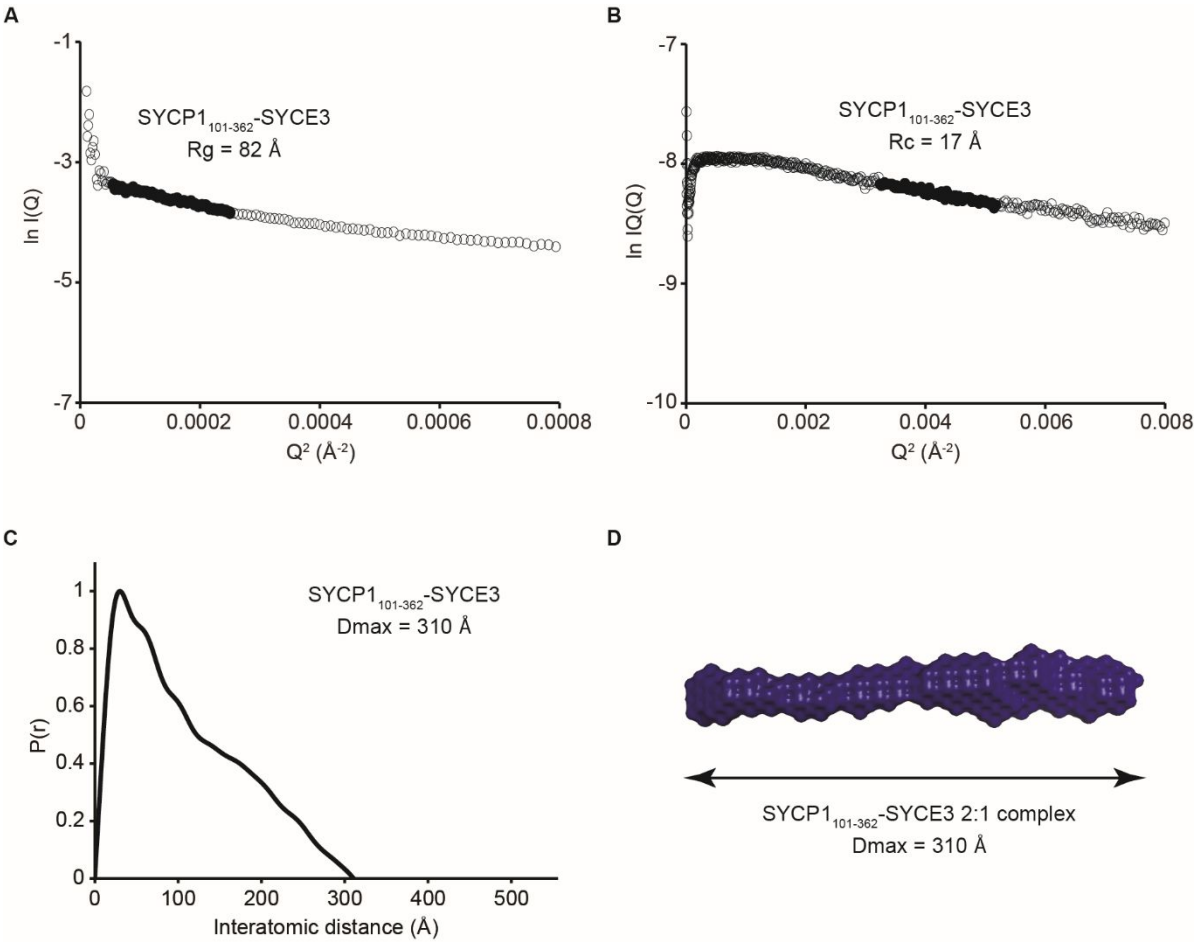


Figure 3.16. SEC-SAXS analysis of the SYCP1$_{101-362}$-SYCE3 2:1 complex. (A) Guinier analysis to determine the radius of gyration (Rg) as 82 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 17 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. The 2:1 complex has a maximum dimension of 310 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

structure than the 2:1 complex (Figure 3.17 D). The comparison between the 2:1 and 4:2 complexes supports the hypothesis that the SYCP1 head-to-head interaction does occur when bound to SYCE3, and that these head-to-head interactions may form recursively in a concentration-dependant manner to form large assemblies.

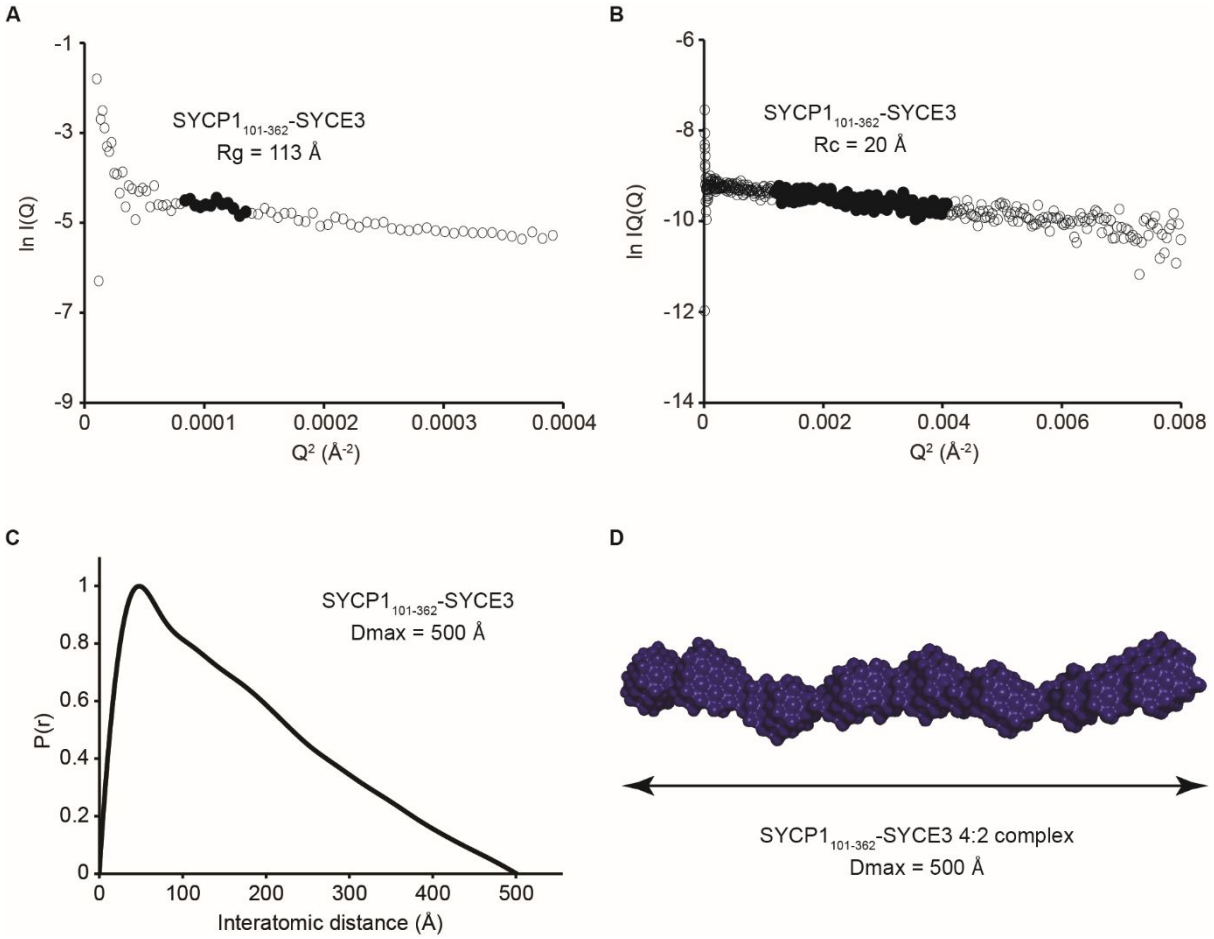


Figure 3.17. SEC-SAXS analysis of the SYCP1$_{101-362}$-SYCE3 4:2 complex. (A) Guinier analysis to determine the radius of gyration (Rg) as 113 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 20 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. The 4:2 complex has a maximum dimension of 500 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

To examine whether the same head-to-head interaction is occurring within the SYCP1-SYCE3 complex as identified in the SYCP1$_{101-175}$ crystal structure, the SYCP1 101-111 residues, which form the anti-parallel four helical assembly, were truncated. The SYCP1$_{112-362}$-SYCE3 complex was purified by consecutive Ni-NTA and amylose affinity chromatography, followed by TEV protease incubation to remove affinity tags. The cleaved protein complex was further purified by anion exchange chromatography, with the SYCP1 degradation product also observed, and a high yield of material was obtained (Figure 3.18 A).

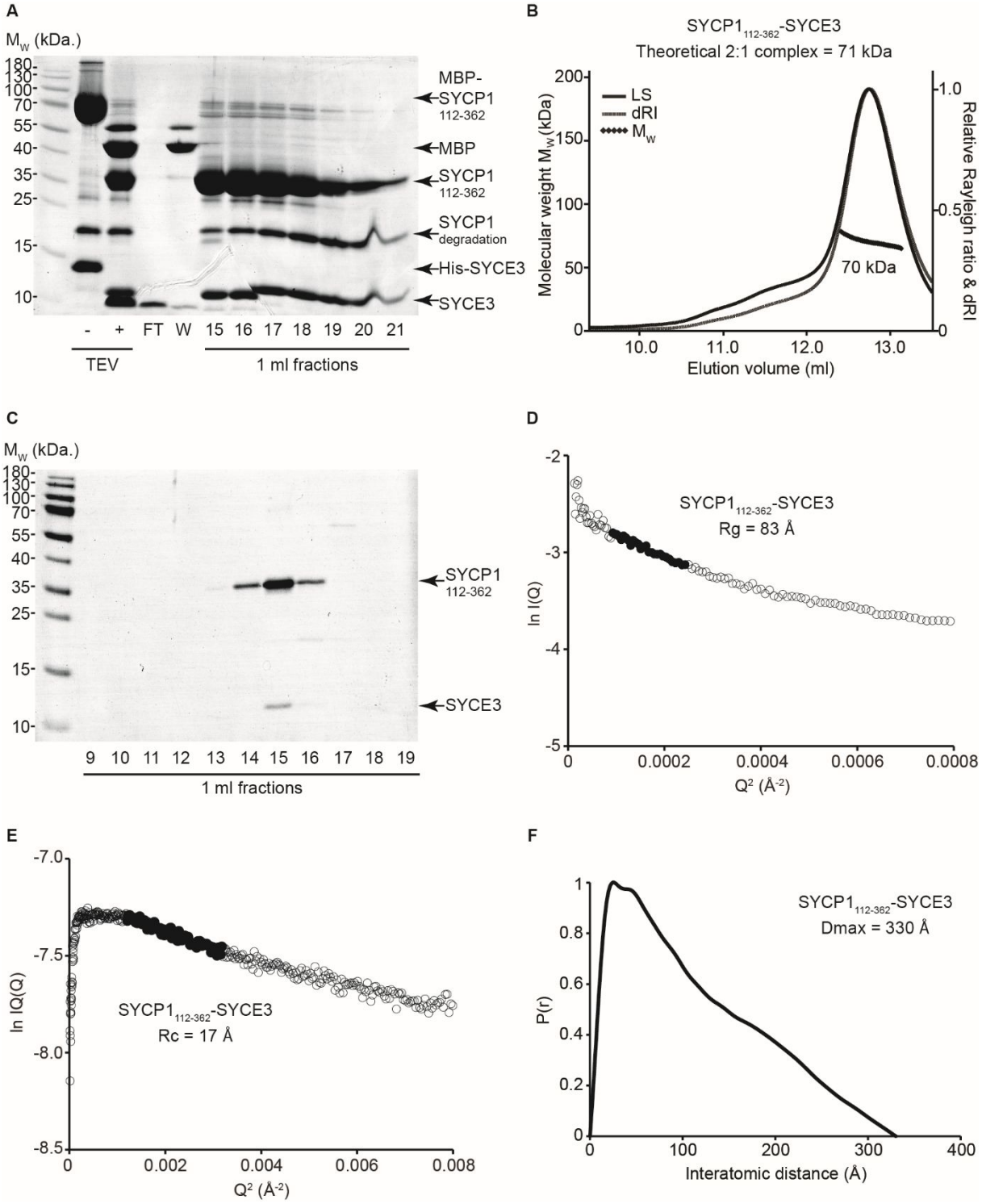


Figure 3.18. Purification and structural analysis of SYCP1$_{112-362}$-SYCE3. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. (B) SEC-MALS analysis showing the single 70 kDa peak. A theoretical 2:1 complex has an estimated molecular weight of 71 kDa, indicating a 2:1 complex formation. (C) SDS-PAGE analysis of the SEC-MALS output 1 ml fractions that demonstrates co-elution of the 2:1 complex. (D) Guinier fit to determine the radius of gyration (Rg) as 83 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (E) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 17 Å. Clear circles are the data, solid circles represent the region used for the fit. (F) Paired real-space distribution plot showing a maximum dimension of 330 Å.

SEC-MALS was used to characterise the oligomeric state of the complex upon the removal of the SYCP1 101-111 residues. The SYCP1$_{112-362}$-SYCE3 complex was analysed at 40 mg/ml and a single species was detected. The theoretical molecular weight of a 2:1 complex is 71 kDa, and the calculated mass of the single protein peak was determined as 70 kDa (Figure 3.18 B and C). This analysis was performed at a higher concentration than SYCP1$_{101-362}$-SYCE3, where large concentration-dependant assemblies were formed. This indicates that the SYCP1 101-111 residues are responsible for the head-to-head recursive interactions as their removal produces a single 2:1 complex, irrespective of the high concentration analysed.

The SYCP1$_{112-362}$-SYCE3 complex was characterised by SEC-SAXS to study the structural parameters for comparison to the SYCP1$_{101-362}$-SYCE3 2:1 complex. The Rg for SYCP1$_{112-362}$-SYCE3 was determined as 83 Å, and the Rc was measured as 17 Å, which correlates to the 82 Å and 17 Å values of the SYCP1$_{101-362}$-SYCE3 2:1 complex (Figure 3.18 D and E). The distance distribution plot highlights the extended shape of the complex with the maximum distance estimated as 330 Å (Figure 3.18 F). The SEC-MALS and SEC-SAXS data indicate that the 2:1 complexes formed by SYCP1$_{101-362}$-SYCE3 and SYCP1$_{112-362}$-SYCE3 have the same structure, with the propensity for higher order assemblies driven by the SYCP1 101-111 residues.

To identify the critical residues required for the head-to-head SYCP1 interaction, a mutant was designed whereby two point mutations were made within the SYCP1 101-111 region. The V105 and L109 amino acids form hydrophobic interactions in the SYCP1$_{101-175}$ crystal structure between the four chains to form the head-to-head four helical bundle (Dunce *et al*., 2018). V105 and L109 residues within the SYCP1$_{101-362}$ fragment were therefore both mutated to glutamates, in order to disrupt these hydrophobic interactions, and co-expressed with SYCE3 for analysis, hereby referred to as SYCP1$_{mut}$-SYCE3.

The SYCP1$_{mut}$-SYCE3 complex was purified in the same manner as SYCP1$_{112-362}$-SYCE3, with the cleaved complex purified by anion exchange chromatography (Figure 3.19 A). The SYCP1 degradation product, approximately the 206-362 region, was also detected. SEC-MALS was performed to analyse the oligomeric state of SYCP1$_{mut}$-SYCE3 to investigate the capacity for the head-to-head interactions to occur, with the protein sample analysed at 40 mg/ml to ensure the detection of any assemblies, if formed. The analysis showed the formation of a single protein species with a calculated mass of 87 kDa (Figure 3.19 B). A theoretical 2:1 complex is estimated to be 74 kDa, and therefore, implies that a 2:1 complex is formed, with the slightly high molecular weight resulting from the minor MBP-SYCP1$_{mut}$ contamination that can be observed by SDS-PAGE analysis of the SEC-MALS output fractionation (Figure 3.19

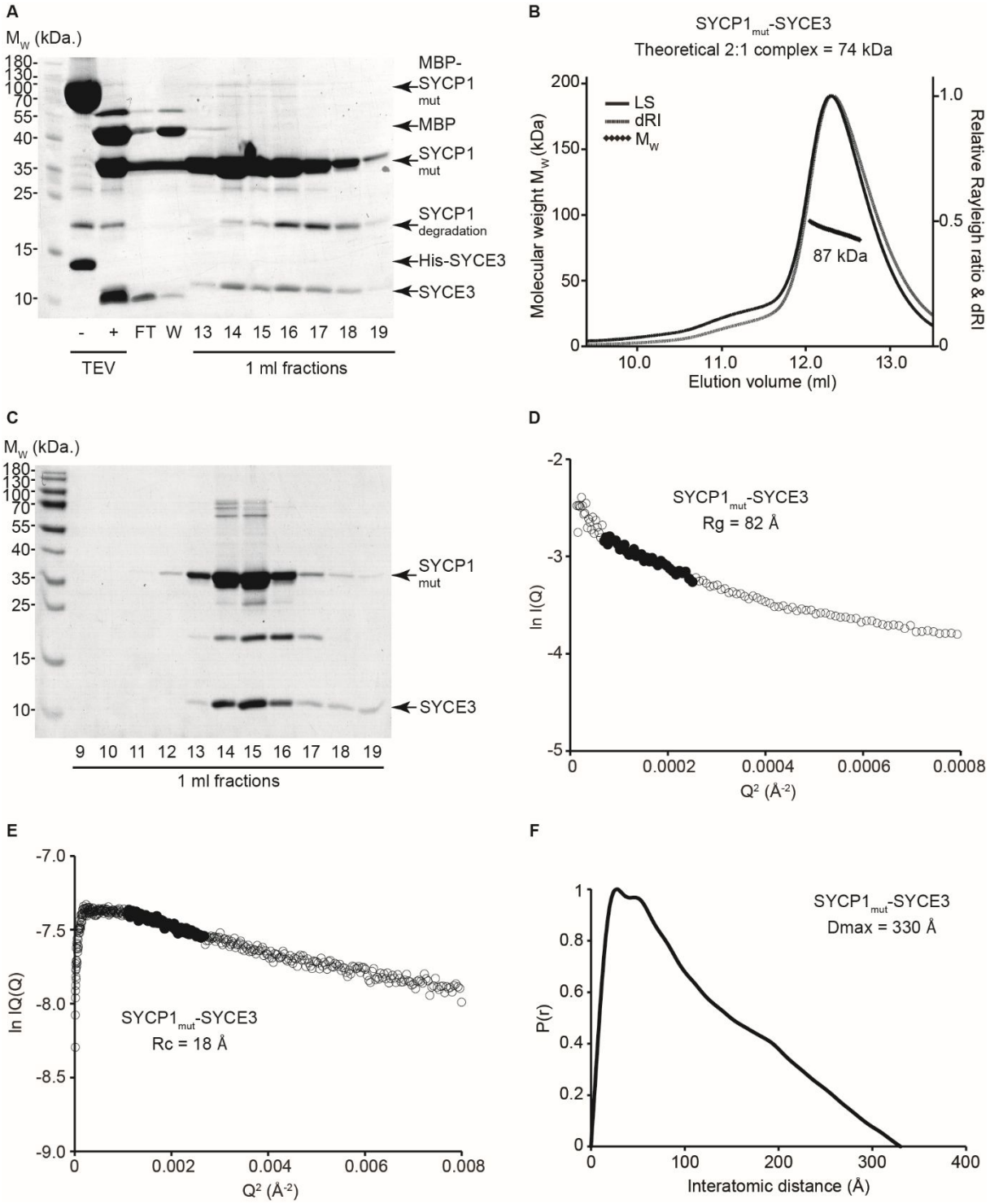


Figure 3.19. Purification and structural analysis of SYCP1$_{mut}$-SYCE3. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. (B) SEC-MALS analysis showing the single 87 kDa peak. A theoretical 2:1 complex has an estimated molecular weight of 74 kDa, indicating a 2:1 complex is likely formed, with the slightly higher molecular weight owing to the presence of uncleaved MBP-SYCP1$_{mut}$ contamination. (C) SDS-PAGE analysis of the SEC-MALS output 1 ml fractions showing the co-elution profile of the complex, with the MBP-SYCP1$_{mut}$ and SYCP1 degradation contamination products. (D) Guinier fit to determine the radius of gyration (Rg) as 82 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (E) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 18 Å. Clear circles are the data, solid circles represent the region used for the fit. (F) Paired real-space distribution plot showing a maximum dimension of 330 Å.

C). Therefore, it has been demonstrated that the mutation of the V105 and L109 SYCP1 residues has blocked the head-to-head association, and any consequent higher-order assembly. To validate the structural integrity of the SYCP1$_{mut}$-SYCE3 complex, upon the introduction of the mutations, SEC-SAXS was performed. The radius of gyration was calculated as 82 Å and the cross-sectional radius was determined as 18 Å (Figure 3.19 D and E). The real space analysis demonstrates the extended shape of the complex, with the maximum dimension observed established as 330 Å (Figure 3.19 F). These parameters for the SYCP1$_{mut}$-SYCE3 2:1 complex are comparable to the parameters determined for SYCP1$_{112-362}$-SYCE3 and the SYCP1$_{101-362}$-SYCE3 complex. This suggests that the 2:1 complex is consistent between the three samples, which is driven by the SYCP1$_{206-362}$ region. The capacity for SYCP1 to undergo head-to-head associations is therefore dependent on the hydrophobic interactions that are formed in the SYCP1$_{101-175}$ crystal structure (Dunce *et al*., 2018).

3.2.4 SYCE3 undergoes conformational changes to form self-assemblies

The X-ray crystal structure of the mouse SYCE3 showed that it forms a compact dimeric four helical bundle in the crystal lattice (Lu *et al*., 2014). In the SYCE3 structure, the N-terminal 1-11 residues are not fitted, likely due to flexibility in this region, and the loop region is also not fitted. In order to elucidate the missing regions, further refinement efforts were made by Owen Davies, and this resulted in the building of the loop region and some further N-terminal residues (Figure 3.20) (Dunne *et al.,* manuscript in preparation).

One SYCE3 molecule interacts with two SYCP1 chains to form the SYCP1-SYCE3 complex, suggesting that SYCE3 undergoes conformational adaptability, and therefore, the structural properties of SYCE3 were studied in solution to investigate this.

SYCE3 was fused to an MBP affinity tag for bacterial expression and purification by Ni-NTA and amylose affinity chromatography. The affinity tag was removed by TEV protease incubation and the cleaved protein was purified though subsequent Ni-NTA and anion exchange chromatography to isolate SYCE3 to near homogeneity for analysis (Figure 3.21 A).

To investigate the oligomeric properties of SYCE3 in solution, SEC-MALS was used to separate protein species for molecular mass determination. The analysis showed the formation of multiple species, with the majority of the material forming a dimer, in agreement with the crystal structure (Figure 3.21 B). The larger species correspond to the size of two, three and six dimers respectively, with the SDS-PAGE analysis of the output fractions showing the broad elution profile of SYCE3, which omits the possibility of contamination products forming the larger species (Figure 3.21 B and C).

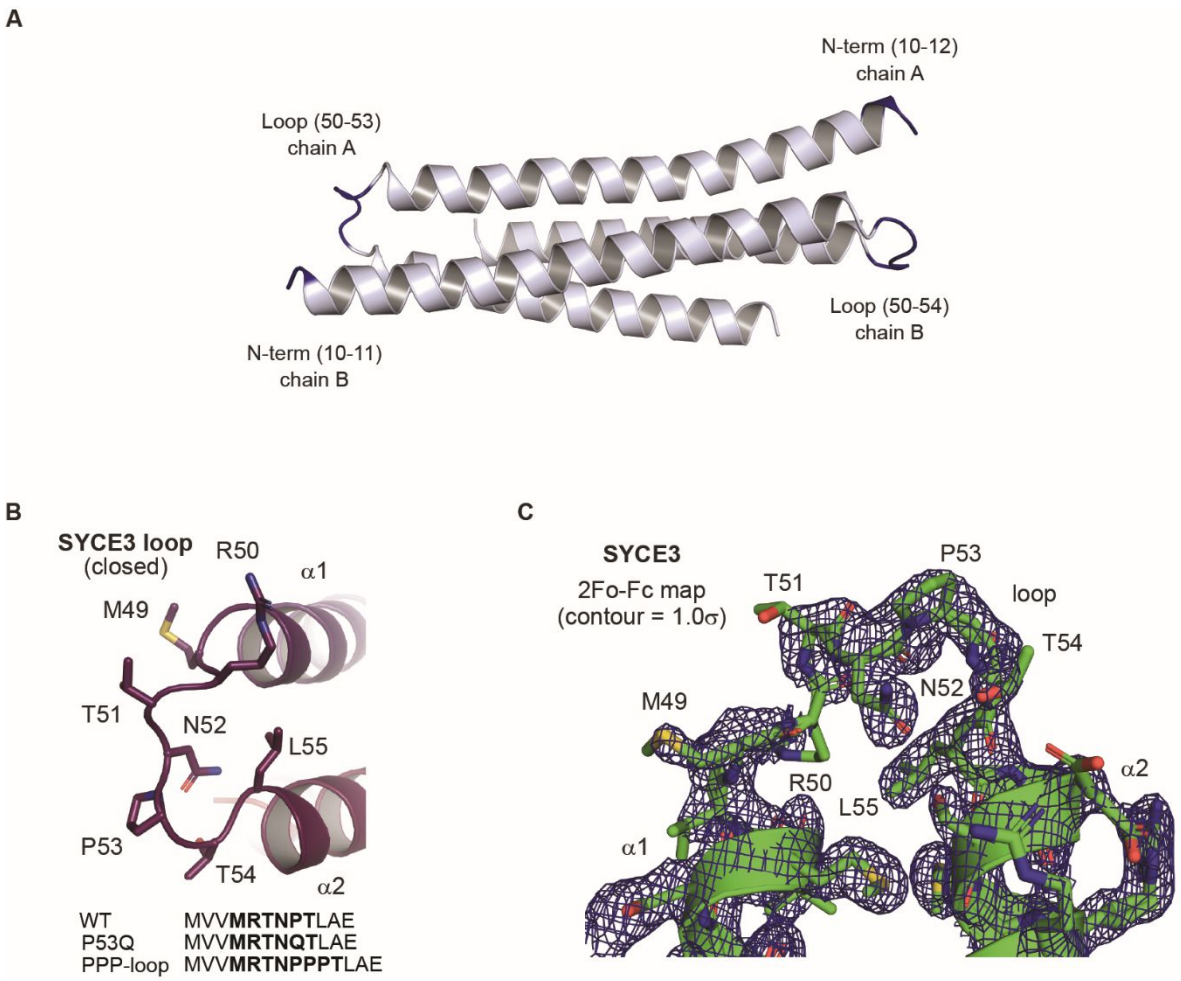


Figure 3.20. X-ray crystal structure of the re-refined mouse SYCE3. SYCE3 forms a compact dimeric structure through the folding back of each chain to form a four helical bundle through the interlocking of antiparallel SYCE3 molecules. Further refinement allowed the building of the loop regions, which allow the chains to fold back, and further N-terminal residues (rebuilt regions coloured in blue. (B) Focussed view of the re-built loop region. The sequences below correlate to the loop region and two mutants characterised below. (C) 2Fo-Fc electron density map showing the rebuilt loop region. PDB code: 6H86. Figures prepared by Owen Davies (Dunne *et al*., manuscript in preparation).

To study the structural properties of the dimeric SYCE3 in solution, SEC-SAXS experiments were performed to analyse the isolated dimer species for comparison to the crystal structure. The radius of gyration was measured as 24 Å, indicating a compact structure (Figure 3.21 D). The cross-sectional radius was determined as 14 Å, consistent with a four helical structure (Figure 3.21 E). The real space distance distribution analysis also indicates a highly globular shape with a slight tail that may represent protrusion of the N-terminus of SYCE3 that is not present in the crystal structure, likely due to flexibility (Figure 3.21 F). The maximum dimension of the sample is 92 Å, which correlates to the length of the crystal dimer (Figure 3.21 F).

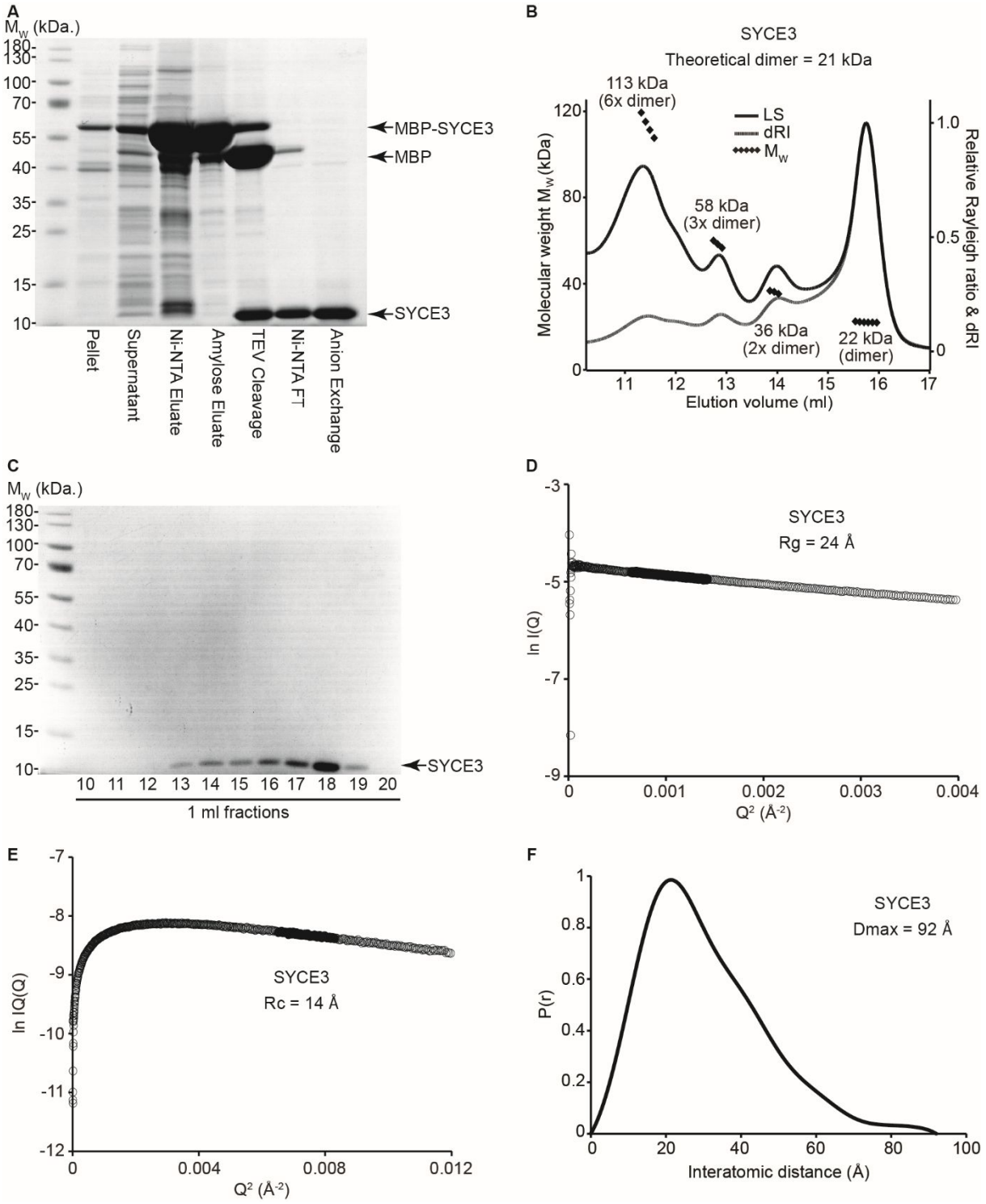


Figure 3.21. Purification and structural analysis of SYCE3. (A) SDS-PAGE showing the purification through Ni-NTA and amylose affinity chromatography, tag removal by TEV protease incubation and subsequent Ni-NTA and anion exchange chromatography. (B) SEC-MALS analysis showing a predominant 22 kDa peak corresponding to a dimer. Larger assemblies were also detected, correlating to 2x, 3x and 6x dimer species. (C) SDS-PAGE analysis of the SEC-MALS output 1 ml fractions showing the broad elution profile of the SYCE3 species. (D) Guinier fit to determine the radius of gyration (Rg) as 24 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (E) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 14 Å. Clear circles are the data, solid circles represent the region used for the fit. (F) Paired real-space distribution plot showing a maximum dimension of 92 Å.

To verify that the solution dimer is the same conformation as the crystal structure, *ab initio* modelling was carried out and the mouse SYCE3 crystal structure was docked into the resultant envelope. The crystal structure does indeed fit into the SAXS molecular envelope, and therefore, suggests that the solution dimer has the same structure (Figure 3.22). The larger species observed by the SEC-MALS analysis were also analysed by SEC-SAXS. The real-space distance distribution plot for the incremental species demonstrates a gradual increase in length (Figure 3.23 A).

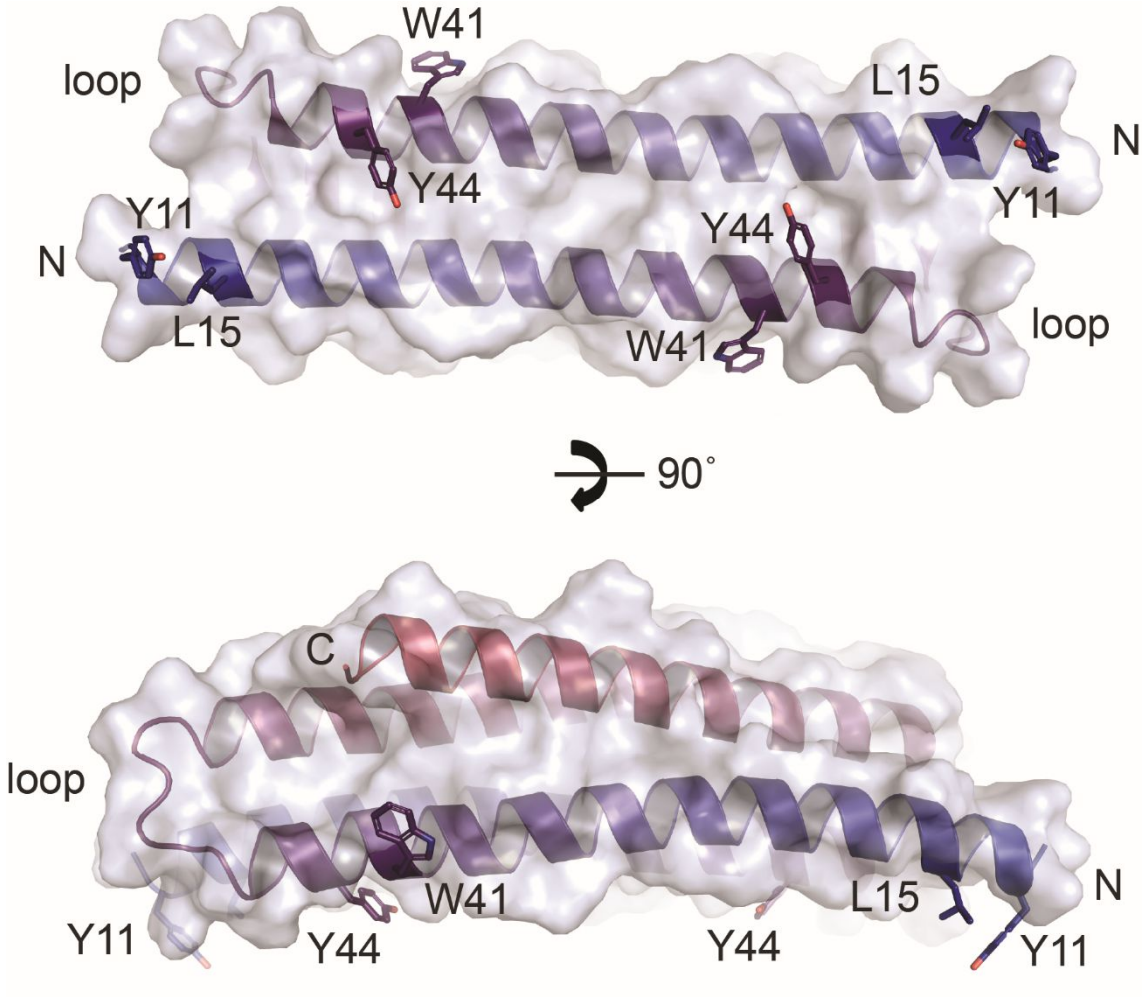


Figure 3.22. Validation of the solution SYCE3 dimer. The refined mouse SYCE3 crystal structure (PDB code: 6H86) was docked into the human dimeric SYCE3 SAXS envelope, which was generated by DAMMIF *ab initio* modelling, and displayed in two orientations. The investigated surface exposed residues and loop region are indicated. Figure prepared by Owen Davies (Dunne *et al.,* manuscript in preparation).

SYCE3 is found to form larger species with increase length, but the mechanism of the formation of these larger assemblies is unclear. By examining the crystal SYCE3 dimer, several surface exposed residues propose candidate amino acids for interactions between dimers. The aromatic surface exposed W41 and Y44 residues are positioned to form a possible interaction site for self-assembly (Figure 3.22). The Y11 and L15 residues at the N-terminus of the structure are also surface exposed and positioned close-by to provide another possible interaction site (Figure 3.22). Finally, the dimeric SYCE3 is formed from the SYCE3 chains folding back on themselves to form a four helical bundle, and it was postulated that the SYCE3 loop may open to form a linear conformation in order to assemble.

To test the role of the W41 and Y44 residues in SYCE3 assembly, two mutants were designed whereby W41 and Y44 were simultaneously mutated to glutamates or alanines in order to remove the potential aromatic binding patch. The SYCE3 W41E Y44E and W41A Y44A mutants were purified in the same manner as the wild type SYCE3. SEC-MALS analysis was performed for oligomeric comparison to the wild type profile. The SYCE3 W41E Y44E still forms a predominant dimeric species, suggesting that the mutations have not disrupted the folding of the SYCE3 dimer (Figure 3.23 B). However, these mutations have disrupted the formation of the larger species, but maintain the propensity to form the 2x dimer species (Figure 3.23 B). This suggests that there may be multiple mechanisms of assembly.

Further to this, the W41A Y44A mutant also formed the dimer and tetramer species, with the majority of the material forming a tetramer (Figure 3.23 D), suggesting that the W41 and Y44 residues are required for the 3x dimer and larger species. SEC-SAXS distance distribution analysis of the dimer and tetramer species shows that the tetramer is almost double the length of the dimer, with the P(r) profile indicating an extended conformation of the tetramer compared to the more globular shape of the dimer (Figure 3.23 C). The Guinier analysis of the W41A Y441A tetramer and dimers species measure the Rg values to be 37 Å and 22 Å respectively (Figure 3.23 E) and the Rc values to be 12 Å and 13 Å respectively (Figure 3.23 F). The cross-sectional radius is very similar for both the dimer and tetramer, but the real space analysis and Rg values indicate that a more elongated structure is formed, suggesting a longitudinal association of the dimers to form the tetramer.

To assess whether the Y11 and L15 N-terminal surface exposed residues contribute to SYCE3 higher-order assembly, a series of mutants were designed to dissect the role of the N-terminus. As the N-terminal residues 1-9 are absent from the rebuilt SYCE3 crystal structure, and therefore from the core helical bundle structure, it was reasoned that these amino acids are flexible and so a truncation was made to analyse SYCE3 without the flexible N-terminus. The SEC-MALS analysis of $SYCE3_{9\text{-}88}$ demonstrates the dimer structure is still

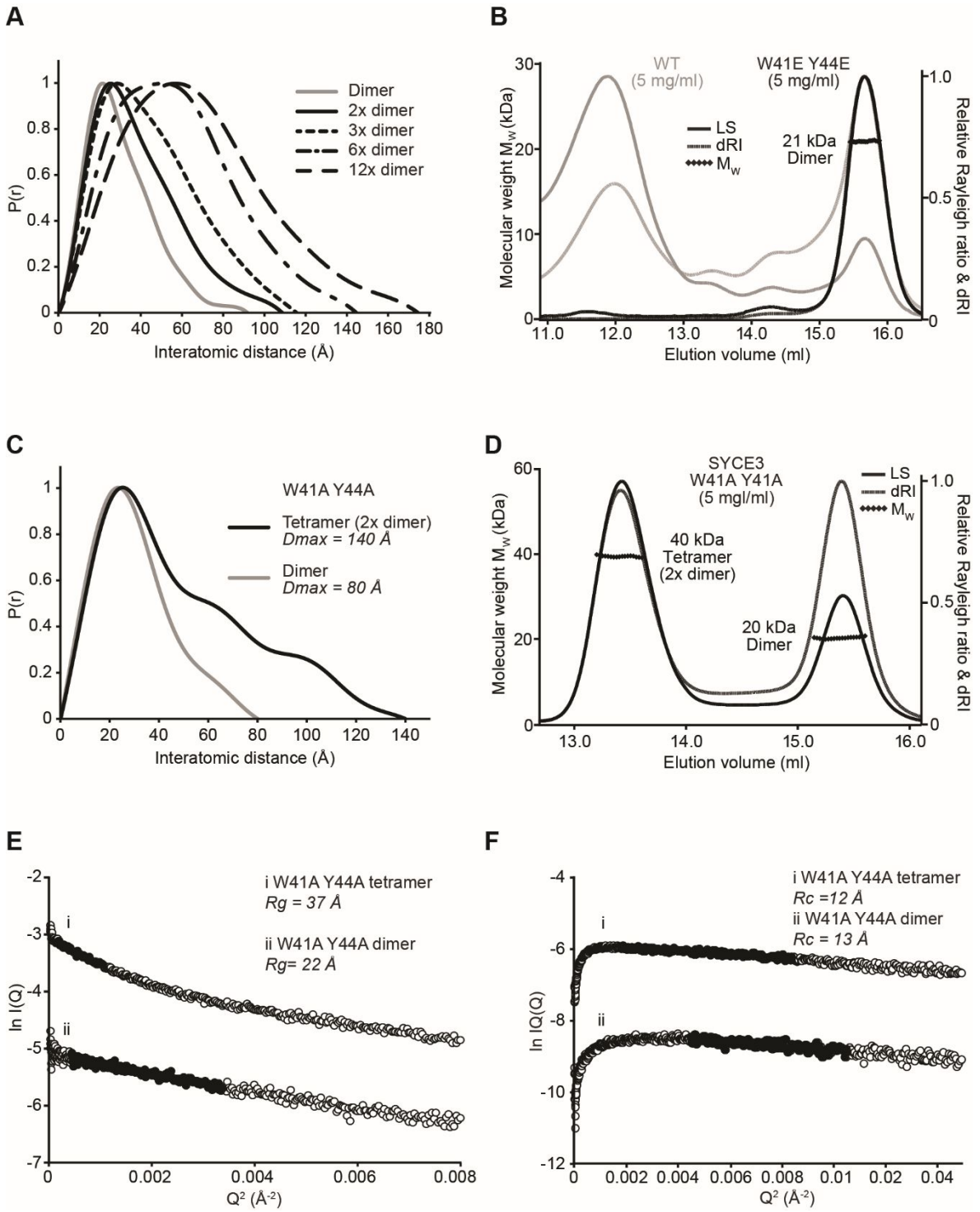


Figure 3.23. Structural analysis of SYCE3 W41 Y44 mutants. Protein samples purified and analysed by Orla Dunne. (A) SEC-SAXS P(r) curves for the various assemblies of the wild type SYCE3 protein. (B) SEC-MALS analysis of the W41E Y44E mutant (black) that forms a predominantly dimeric species, with a small peak corresponding to a 2x dimer species. WT data plotted for comparison (grey). (C) SEC-SAXS P(r) curves for the W41A Y44A dimer and tetramer species. The maximum lengths of the dimer and tetramer are 80 Å and 140 Å, respectively. (D) SEC-MALS analysis of the W41A Y44A mutant showing 20 kDa dimer and a 40 kDa tetramer formation. (E) Overlaid Guinier fits to determine the Rg of the W41A Y44A dimer and tetramer to be 22 Å and 37 Å, respectively. Clear circles are the data, solid circles represent the region used for the Guinier fit. (F) Guinier analysis to determine the Rc of the W41A Y44A dimer and tetramer to be 13 Å and 12 Å, respectively. Clear circles are the data, solid circles represent the region used for the fit.

formed but also that the high-order species are still formed and are more prominent (Figure 3.24 A). This implies that the SYCE3 N-terminus does not mediate the formation of the larger species, and may in fact play a role in inhibiting the higher-order assemblies to some extent.
In order to evaluate whether the Y11 and L15 residues have an impact on higher-order assembly formation, the two residues were mutated to asparagine and analysed in the $SYCE3_{9-88}$ fragment. The $SYCE3_{9-88}$ Y11N L15N mutant was purified in the same manner as the wild type SYCE3. SEC-MALS analysis displays a drastic reduction in the formation of larger species in comparison to $SYCE3_{9-88}$ (Figure 3.24 A).

To further dissect the role of the N-terminal residues, another N-terminal truncation was made for the analysis of the SYCE3 12-88 amino acids, whereby the Y11 residue is absent. The purified $SYCE3_{12-88}$ was analysed by SEC-MALS, with the capacity to form the larger species almost entirely abolished (Figure 3.24 A). Finally, the $SYCE3_{12-88}$ was analysed with the L15N mutation, and this mutant was found to solely exist in the dimeric form (Figure 3.22 B). Therefore, Y11 and L15 appear to be crucial to form higher-order species formation.

SEC-SAXS experiments were performed with the SYCE3 12-88 L15N mutant to verify that the dimer formation matches the wild type SYCE3. Guinier analysis determined the Rg and Rc values of the mutant to be 21 Å and 12.6 Å respectively, which correlate to the values determined for the wild type protein (Figure 3.24 B and C). The real space analysis shows that the 12-88 L15N mutant has a similar P(r) profile to the wild type, with maximum lengths of 70 Å and 92 Å respectively (Figure 3.24 D). The wild type protein has a slight tailing of the peak and a longer length due to the N-terminal 1-11 residues. *Ab initio* modelling was performed with the mutant to verify the molecular shape of the dimer for comparison to wild type. The molecular envelope of the mutant matches that of the wild type, although with the removal of the bulk ends due to the truncation. The refined mouse structure was successfully docked into the mutated SYCE3, suggested that the dimer species of the mutant is the same as the wild type (Figure 3.24 E).

The candidate surface exposed residues have been found to contribute to the self-assembly of SYCE3. However, it remains unclear whether the SYCE3 compact dimer structure is maintained upon assembly or whether the SYCE3 chains form an open conformation to self-assemble. To address this, the SYCE3 loop region, towards the centre of the SYCE3 sequence, was investigated as it forms the link between the two halves of SYCE3 to allow the folding back in order to form the four helical bundle (Figure 3.35 A). This region has been refined in the crystal structure (Dunne *et al.*, manuscript in preparation) and shows that the proline residue, as predicted by the rigid geometry features of proline, provides the helical break to allow the folding back of the SYCE3 molecule.

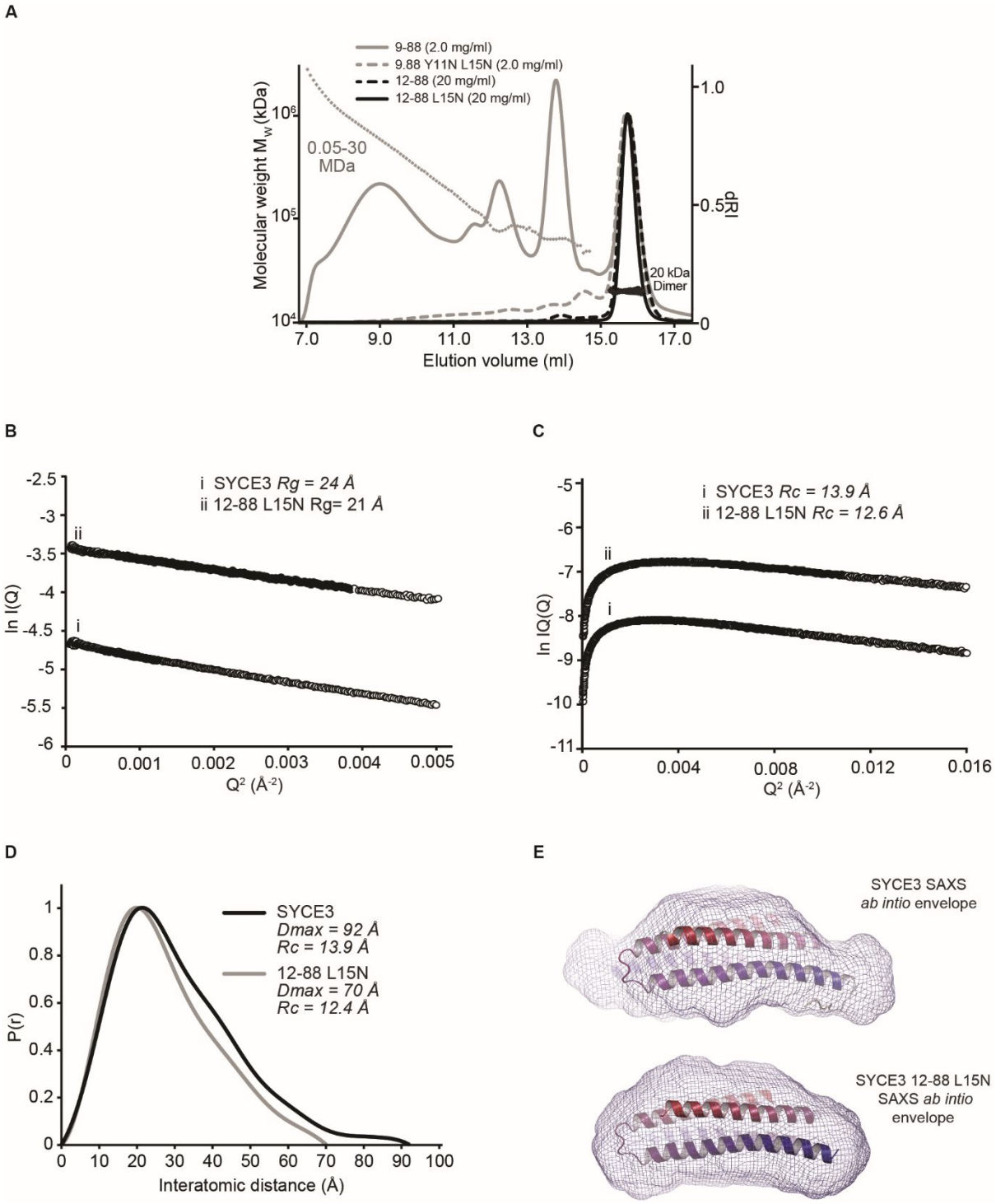


Figure 3.24. Structural analysis of SYCE3 N-terminal mutants. Protein samples purified and analysed by Orla Dunne. (A) 9-88 (grey) forms a dimer and the higher order species as WT, though with much larger assemblies detected. 9-88 Y11N L15N mutant mostly forms a dimer, with a mild propensity to form the 2x dimer species (grey dashed). 12-88 predominantly forms a dimer with a small 3x dimer peak (black dashed). The 12-88 L15N mutant solely forms a dimer (black). (B) Overlaid Guinier fits to determine the Rg of the 12-88 L15N and wild type dimers to be 21 Å and 24 Å respectively. Clear circles are the data, solid circles represent the region used for the Guinier fit. (C) Guinier analysis to determine the Rc of 12-88 L15N and wild type dimers to be 12.6 Å and 13.9 Å, respectively Clear circles are the data, solid circles represent the region used for the fit. (D) SEC-SAXS P(r) curves for the 12-88 L15N and wild type dimers, with the maximum lengths of 70 Å and 92 Å, respectively. (E) DAMMIF *ab initio* models of the wild type and 12-88 L15N dimers, with docking of the refined structure.

The loop stretch allows the formation of the dimeric helical bundle, therefore to probe the conformation of the assembled SYCE3, two mutants were designed. One mutant was designed whereby the P53 residue, that provides the loop rigidity to form the compact dimer structure, was mutated to a glutamine in order to prevent the folding back of the SYCE3 chain and to encourage SYCE3 to form an extended linear conformation. When the SYCE3 P53Q mutant was analysed by SEC-MALS, for comparison to the wild type SYCE3, large molecular weight species were formed, with the majority of the material forming the large assemblies and no dimeric material detected (Figure 3.25 B). This implies that the linear conformation of SYCE3 promotes the self-assembly mechanism.

To validate that the extenuated higher-order assemblies are a result of an open conformation of the SYCE3 chains, the converse experiment was performed, whereby another mutant was designed with the inclusion of two additional proline residues within the SYCE3 flexible loop (Figure 3.25 A). This mutant, hereby referred to as the SYCE3 PPP-loop, was created to ensure the compact folding back form of SYCE3 to probe whether the compact SYCE3 conformation can form the larger species. This mutant was purified in the same manner as the wild type and the purified material was analysed by SEC-MALS to examine the oligomeric properties. The molecular mass analysis shows that the SYCE3 PPP-loop does not form any larger species and solely exist as a dimer (Figure 3.25 B).

These mutants were studied by SEC-SAXS experiments, where the P53Q individual assemblies show the gradual increase in species length as observed for the wild type, but without any dimer formation (Figure 3.25 C). The real space analysis of the SYCE3 PPP-loop shows the globular formation, with *ab initio* modelling and docking of the refined SYCE3 structure verifying that the SYCE3 PPP-loop dimer matches the wild type dimer (Figure 3.25 D and E).

As the wild type SYCE3 forms both dimers and larder species, these findings suggest that the SYCE3 chain requires an open conformation for the higher-order assemblies, further driven by the Y11/L15 and W41/Y44 residues, but that the intermediate species require capping by the compact SYCE3 form. Modelling of the refined crystal structure was performed to create the theoretical model of the tetramer (2x dimer) species, whereby two open loop chains replicate a side-by-side dimers through capping of each end by folded SYCE3 chains, where the respective four helical bundles mimic the folding of the dimer. This theoretical model was fitted into the solution structure of the W41A Y44A tetramer to demonstrate that the theoretical model fits experimental data (Figure 3.26 A). The W41 and Y44 residues are not required for the tetramer formation, and therefore, it was proposed that they function to form staggered lateral

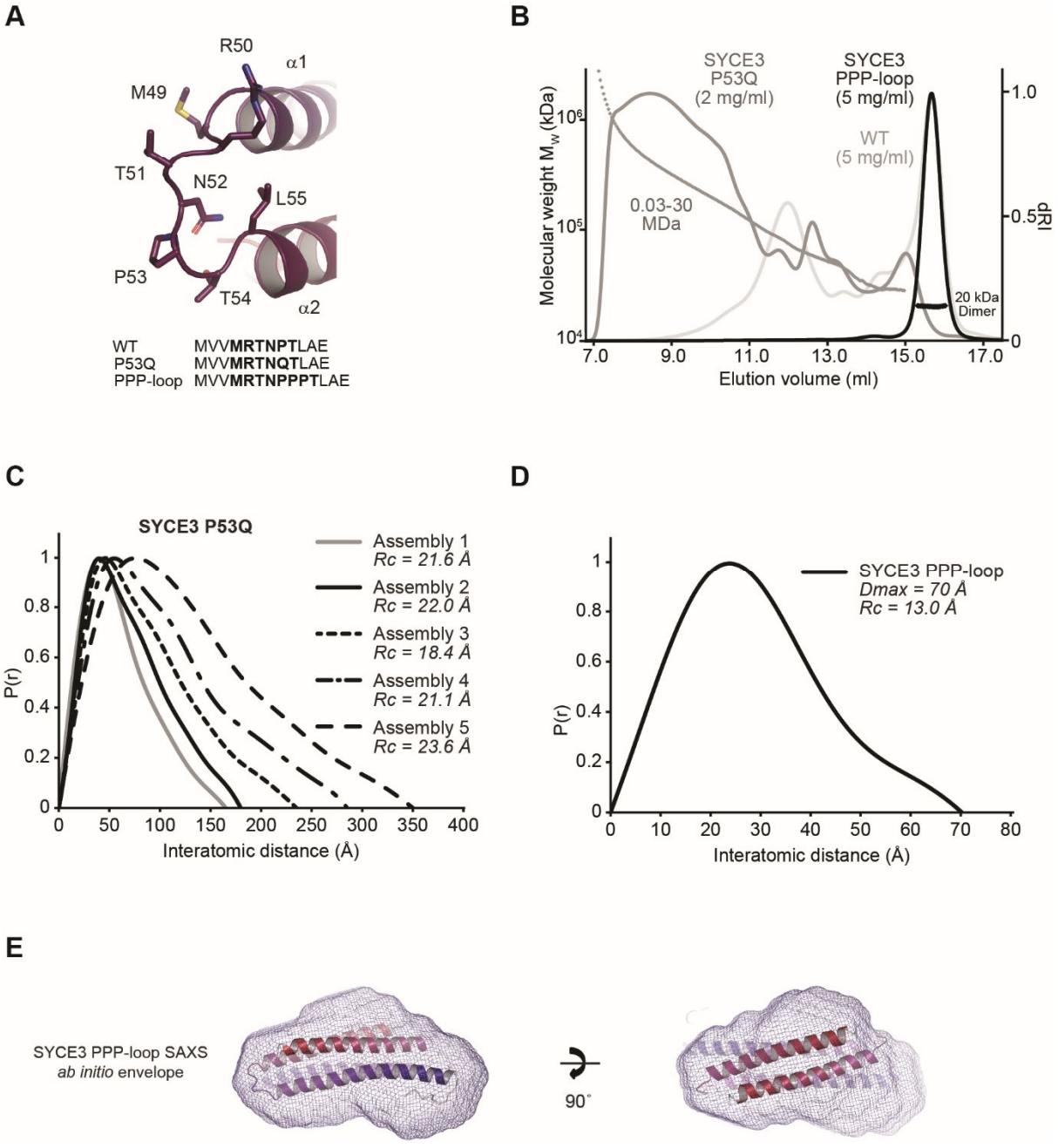


Figure 3.25. Structural analysis of the SYCE3 open and closed conformations. Protein samples were purified and analysed by Orla Dunne. (A) Focussed view of the re-built loop region. The sequences below correlate to the loop region and two mutants characterised. (B) SEC-MALS analysis with the WT trace displaying the dimer and larger species (light grey). The P53Q mutant forms increased large molecular species with no dimer species detected (dark grey). The PPP-loop mutant only forms dimeric species (black). (C) SEC-SAXS P(r) curves for the various assemblies of the P53Q SYCE3 protein showing an increasing in length upon the larger assemblies. (D) SEC-SAXS P(r) curve for the SYCE3 PPP-loop dimer, with a maximum length of 70 Å. (E) DAMMIF *ab initio* model of the SYCE3 PPP-loop dimer, with docking of the refined mouse SYCE3 structure to validate the solution dimer.

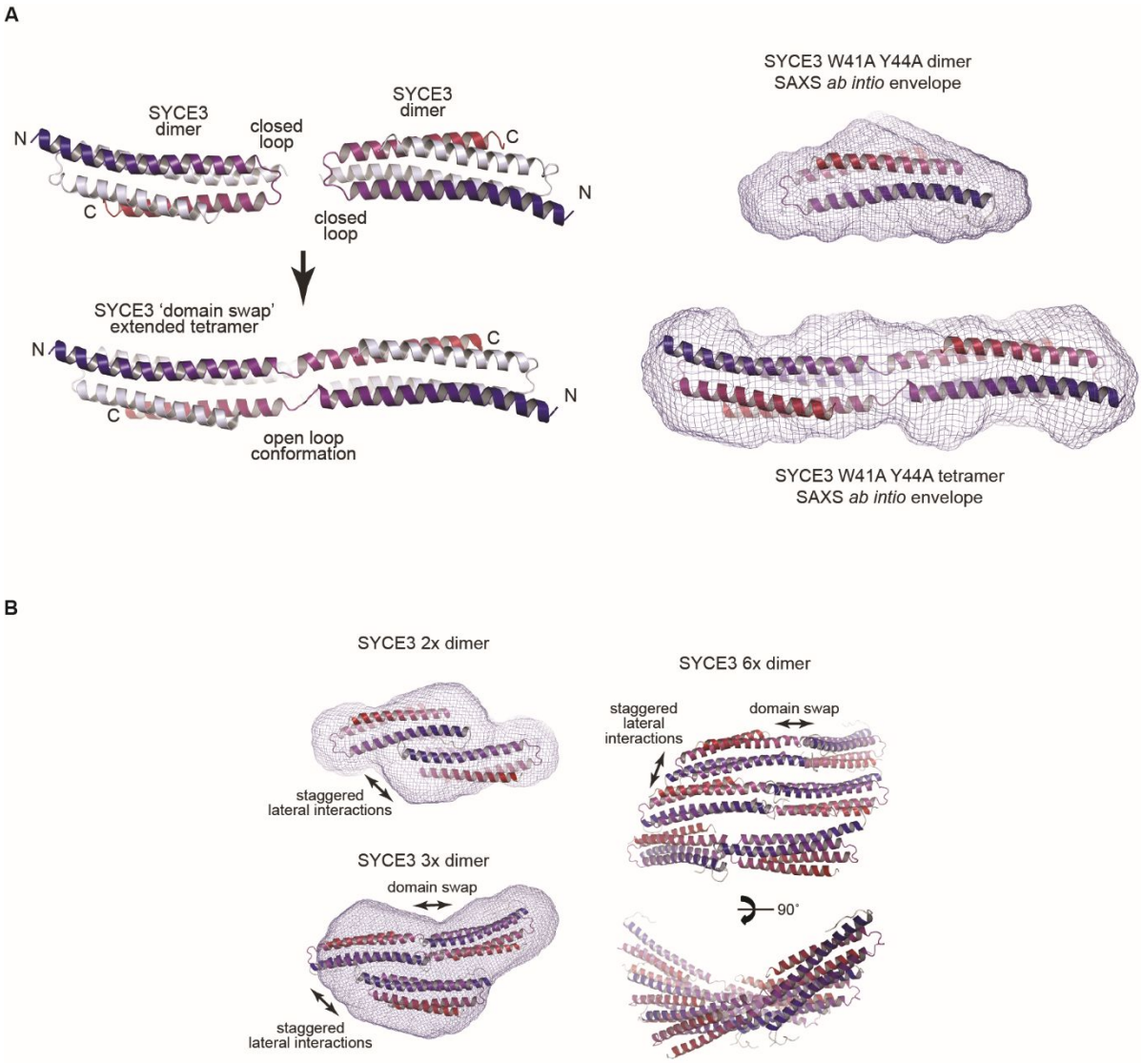


Figure 3.26. Theoretical modelling of the SYCE3 species. Analysis and figure preparation were performed by Owen Davies. (A) Theoretical modelling of the SYCE3 tetramer from the SYCE3 crystal dimer through two chains in the open conformation capped by two closed loop chains. The modelled tetramer was docked into the experimental DAMMIF *ab initio* model of the SYCE3 W41A Y44A tetramer. (B) Further modelling was performed to generate models for the staggered lateral interactions between dimer species or between tetramer and dimer species. Models were fitted into DAMMIF *ab initio* experimental models. The 6x dimers species of formed through a combination of the "domain swap" tetramers and staggered lateral interactions.

interactions between dimer and tetramer species due to their surface exposed position in the structure. The combination of the open and compact loop conformations, coupled with the staggered lateral interactions, allows the formation of the higher-order species (Figure 3.26 B).

3.2.5 SYCE3 undergoes self-assembly to tether lateral SYCP1-SYCE3 units

Although SYCE3 has been characterised as a compact dimer that undergoes larger assemblies, through loop opening and staggered lateral interactions stabilised by surface exposed residues, the function of both the dimer and larger species remains uncertain. One SYCE3 molecule has been shown to interact with two SYCP1 chains to incorporate the transverse filament into the central element, and therefore, the biological relevance of the SYCE3 higher-order assembly is unclear.

As the surface exposed residues and loop opening contribute to the formation of the larger SYCE3 species, it was postulated whether these regions have an impact on the SYCP1 interaction. To investigate this, recombinant protein pull-down assays were performed between SYCP1 and the various SYCE3 mutants to assess the binding capacity.

An initial pull-down assay was performed with the non-assembling $SYCP1_{112-362}$ fragment fused to an MBP affinity tag that was co-expressed with SYCE3 fused to a His_6 affinity tag. Four SYCE3 mutants were tested, in which candidate aromatic surface-exposed residues were mutated to alanine: Y11A, W41A, Y44A and the double W41A Y44A mutant. After application of clarified bacterial lysate to amylose resin and subsequent washing, bound protein was eluted from the resin and eluates were analysed by SDS-PAGE. This analysis demonstrated that all of the mutants sustained binding to SYCP1 (Figure 3.27 A). Therefore, the SYCE3 surface exposed residues are redundant for SYCP1 binding.

However, as one SYCE3 molecule is present in the SYCP1-SYCE3, it is ambiguous whether the SYCE3 binds to SYCP1 in the closed conformation, whereby it may form a four helical assembly with the two extended SYCP1 chains, or whether SYCE3 binds to SYCP1 in its open loop state, where the extended SYCE3 chain may form a trimeric helical bundle with the two SYCP1 molecules. To address this ambiguity, a recombinant amylose pull-down assay was performed with the $SYCP1_{206-362}$ fragment, fused to an MBP affinity tag, co-expressed with the P53Q and PPP-loop His_6 tagged SYCE3 mutants for comparison to wild type, in order to assess the binding capacity upon SYCE3 conformation. The experiment revealed that the SYCE3 P53Q interacted with SYCP1 to the same extent as the wild type protein, with the SYCE3 PPP-loop binding almost entirely abolished (Figure 3.27 B). This indicates that the open loop conformation of SYCE3 is necessary for binding to the SYCP1 chains to form a trimeric helical bundle.

A non-assembling extended SYCE3 chain is required for the SYCP1-SYCE3 complex. However, it is unknown whether the whole SYCE3 chain forms the SYCP1-binding surface, or whether only part of the chain is involved in the complex formation. To decipher the interacting

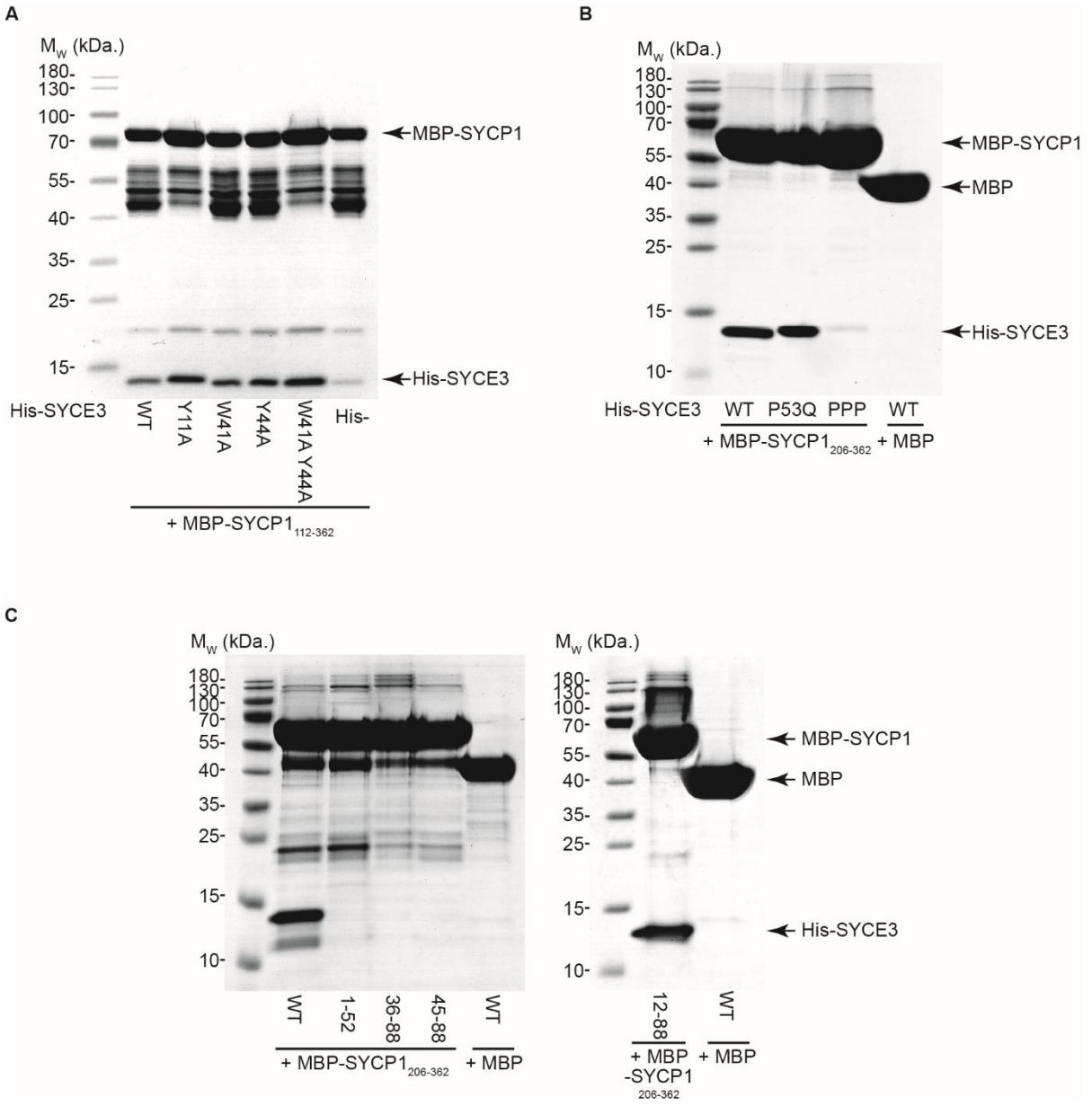


Figure 3.27. Screening for SYCP1 binding with SYCE3 mutations and truncations. (A-C) Recombinant protein pull-down amylose assays. SYCP1 fragments were fused to a His-MBP affinity tag for bacterial co-expression with His-SYCE3 constructs. Clarified lysates were incubated with amylose resin, followed by washing of the resin and elution of bound proteins by maltose buffer. Eluates were normalised for protein concentration for analysis by SDS-PAGE. (A) SYCP1$_{112-362}$ was found to interact with SYCE3 WT, Y11A, W41A, Y44A and the double W41A Y44A mutant. (B-C) Experiments performed by Orla Dunne. (B) SYCP1$_{206-362}$ binds to SYCE3 WT and the P53Q mutant. No binding was detected between SYCP1$_{206-362}$ and the SYCE3 PPP-loop mutant. Free MBP- was co-expressed with WT SYCE3 as a control. (C) SYCP1$_{206-362}$ binds to the SYCE3 full length sequence and the 12-88 region. No binding was detected with the SYCE3 1-52, 36-88 or 45-88 fragments. Free MBP- was co-expressed with WT SYCE3 as a control.

region, an amylose pull-down assay was performed with the $SYCP1_{206-362}$ fragment, fused to an MBP affinity tag, with a series of His_6 tagged SYCE3 truncations in order to identify the minimum binding region. The SYCE3 crystal structure highlights the two helices of SYCE3, amino acids 10-48 and 55-85, separated by the loop region. Therefore, it was postulated whether one of the helices interacts with SYCP1, leaving the other helix to form additional associations in the central element.

To probe this hypothesis, the SYCE3 1-52 fragment was tested, which encompasses the first helix, and this fragment did not bind SYCP1 and thus, the first helix is not sufficient for complex formation (Figure 3.27 C). To test whether the second helix maintains SYCP1 binding, the SYCE3 36-88 and 45-88 fragments were tested, which contain the second helix, and neither of these SYCE3 fragments interacted with SYCP1 (Figure 3.27 C). These findings indicate that both helices are likely required for the SYCP1 interaction. As the flexible SYCE3 N-terminus is not present in the crystal structure, this region might not be required for the SYCP1 binding. An N-terminal truncation was made to test the SYCE3 12-88 region, lacking the flexible N-terminal residues, for SYCP1 binding. The SYCE3 12-88 fragment maintains binding to SYCP1 in the amylose pull-down experiment (Figure 3.27 C). This therefore suggests that the predominant structural portion of SYCE3 forms the SYCP1-SYCE3 complex, and that the flexible N-terminus may perform other biological functions within the central element.

The identification of the $SYCE3_{12-88}$ region forming the structural core of the SYCP1-SYCE3 complex provides an optimised target for crystallisation, as the presence of the flexible N-terminus may hinder crystal growth. The $SYCP1_{206-362}$ and $SYCE3_{12-88}$ fragments were individually fused to His_6-MBP tags for co-expression and purification of the complex through Ni-NTA and amylose affinity chromatography steps. The fusion complex was subsequently incubated with TEV protease to remove the affinity tags. The tag was successfully cleaved from $SYCP1_{206-362}$, but the tag from $SYCE3_{12-88}$ was only partially cleaved to result in a mixture of cleaved and uncleaved $SYCE3_{12-88}$ (Figure 3.28 A). Although a mixture of species was formed, SEC-MALS analysis was performed to assess whether the species could be isolated for molecular mass determination. Some of the cleaved complex was indeed isolated, a calculated molecular weight of 48 kDa, which corresponds to the theoretical mass of a 2:1 complex (Figure 3.28 B). Therefore, the SYCE3 truncation has not disrupted the SYCP1-SYCE3 2:1 complex structure, and the partial affinity tag cleavage from $SYCE3_{12-88}$ suggests that the TEV recognition site may be occluded due to the tight structural nature of the $SYCP1_{206-362}$-$SYCE3_{12-88}$ complex at the N-termini.

Although the $SYCP1_{206-362}$-$SYCE3_{12-88}$ complex appears to be a favourable structural unit for crystallisation studies, the issue of partial tag cleavage inhibits the purification of the

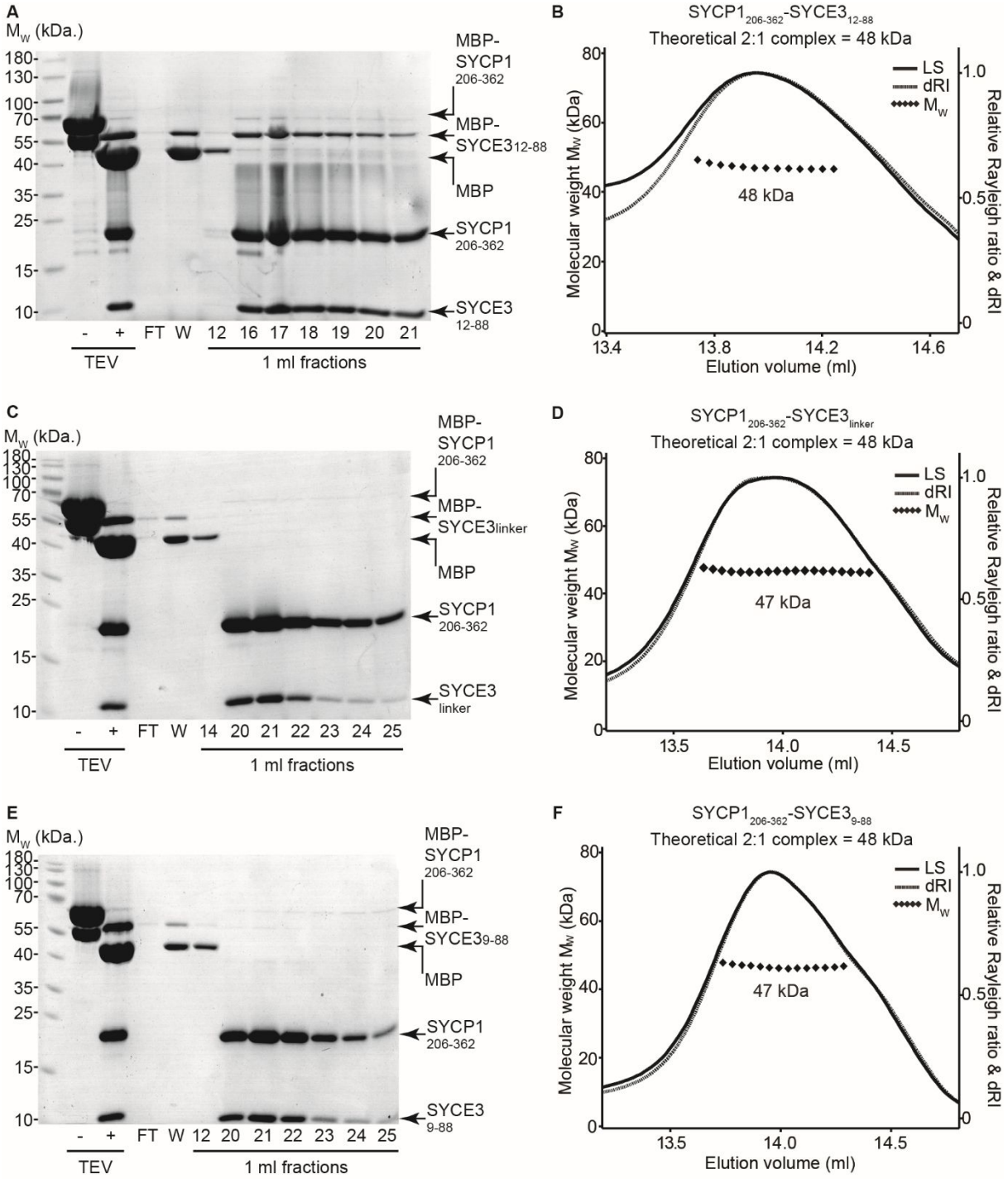


Figure 3.28. Co-purification and oligomer analysis of SYCP1$_{206-362}$ and SYCE3 truncations. (A, C, E) SDS-PAGE showing the purification of the cleaved complexes by anion exchange chromatography, following removal of the MBP tags by TEV protease incubation. (B, D, F) SEC-MALS analysis. Each complex has a theoretical molecular weight of 48 kDa for a 2:1 complex. (A) SYCP1$_{206-362}$-SYCE3$_{12-88}$. The MBP tag was only partially cleaved from SYCE3$_{12-88}$, resulting in a mixed species. (B) SYCP1$_{206-362}$-SYCE3$_{12-88}$ forms a 48 kDa species (as shown) corresponding to a 2:1 complex. A larger species was also formed with the MBP-SYCE3$_{12-88}$ contamination (not shown). (C) SYCP1$_{206-362}$-SYCE3$_{linker}$. The insertion of the linker between the TEV protease recognition site and the SYCE3$_{12-88}$ region resulted in complete MBP tag removal and purification of the cleaved complex. (D) SYCP1$_{206-362}$-SYCE3$_{linker}$ forms a single 47 kDa species, corresponding to a 2:1 complex. (E) SYCP1$_{206-362}$-SYCE3$_{9-88}$ cleaved complex purification following tag removal. (F) SYCP1$_{206-362}$-SYCE3$_{9-88}$ forms a single 47 kDa species, corresponding to a 2:1 complex.

cleaved complex to obtain a high yield of material. To address this, a linker was inserted between the TEV protease recognition site and the SYCE3$_{12-88}$ fragment to aid tag removal. The linker sequence was generated from three threonine-glycine-serine repeats in order to protrude the TEV protease site away from the structural core of the complex, hereby referred to as SYCP1$_{206-362}$-SYCE3$_{linker}$. The MBP-SYCP1$_{206-362}$-MBP-SYCE3$_{linker}$ complex was purified by Ni-NTA and amylose affinity chromatography and incubated with TEV protease to remove the affinity tags. The insertion of the linker resulted in complete cleavage of the MBP tags from both SYCP1 and SYCE3, with the cleaved complex purified by anion exchange chromatography for analysis (Figure 3.28 C). SEC-MALS was used to study the oligomeric state of the SYCP1$_{206-362}$-SYCE3$_{linker}$ complex, with a single 47 kDa species formed that corresponds to a 2:1 complex (Figure 3.28 D).

SYCE3$_{12-88}$ forms the SYCP1 interaction surface and is therefore a candidate construct for structural studies. However, the insertion of the linker is required for purification of the SYCP1$_{206-362}$-SYCE3$_{12-88}$ complex, which results in a 9 amino acid stretch overhang that may hamper crystallisation. Therefore, the SYCP1$_{206-362}$-SYCE3$_{9-88}$ complex was purified to assess whether the affinity tag could be fully removed in order to isolate a core structural SYCP1-SYCE3 complex. The complex was purified in the same way as SYCP1$_{206-362}$-SYCE3$_{12-88}$ with full affinity tag cleavage achieved upon TEV protease incubation. The cleaved SYCP1$_{206-362}$-SYCE3$_{9-88}$ was purified by anion exchange chromatography to obtain pure material for analysis (Figure 3.28 E). SEC-MALS was performed to examine the oligomeric state of the SYCP1$_{206-362}$-SYCE3$_{9-88}$ complex. Although the SYCE3$_{9-88}$ fragment in isolation forms large assemblies, with a higher propensity than the full length protein, when in complex with SYCP1, a 2:1 complex of 47 kDa was formed (Figure 3.28 F). This further highlights the remodelling event that both SYCP1 and SYCE3 undergo upon complex formation, as only a single 2:1 species is formed, even with the SYCE3$_{9-88}$ truncation that accentuates SYCE3 self-assembly.

Although a single 2:1 species is formed with the SYCP1$_{206-362}$-SYCE3$_{9-88}$ complex, to ensure that no larger assemblies are formed, the SYCP1$_{206-362}$-SYCE3$_{9-88}$ complex was purified with the introduction of two point mutants that abolish the SYCE3 self-assembly; Y11N and L15N. The mutated complex, hereby referred to as SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$, was purified in the same manner as previous complexes, with both proteins fused to a His$_6$-MBP tag for sequential Ni-NTA and amylose affinity chromatography and further purified by anion exchange chromatography (Figure 3.29 A). The fusion protein was analysed by SEC-MALS and found to form a single species with a calculated mass of 168 kDa, with the predicted molecular weight of a 2:1 complex of 182 kDa, and therefore, indicating the formation of a 2:1

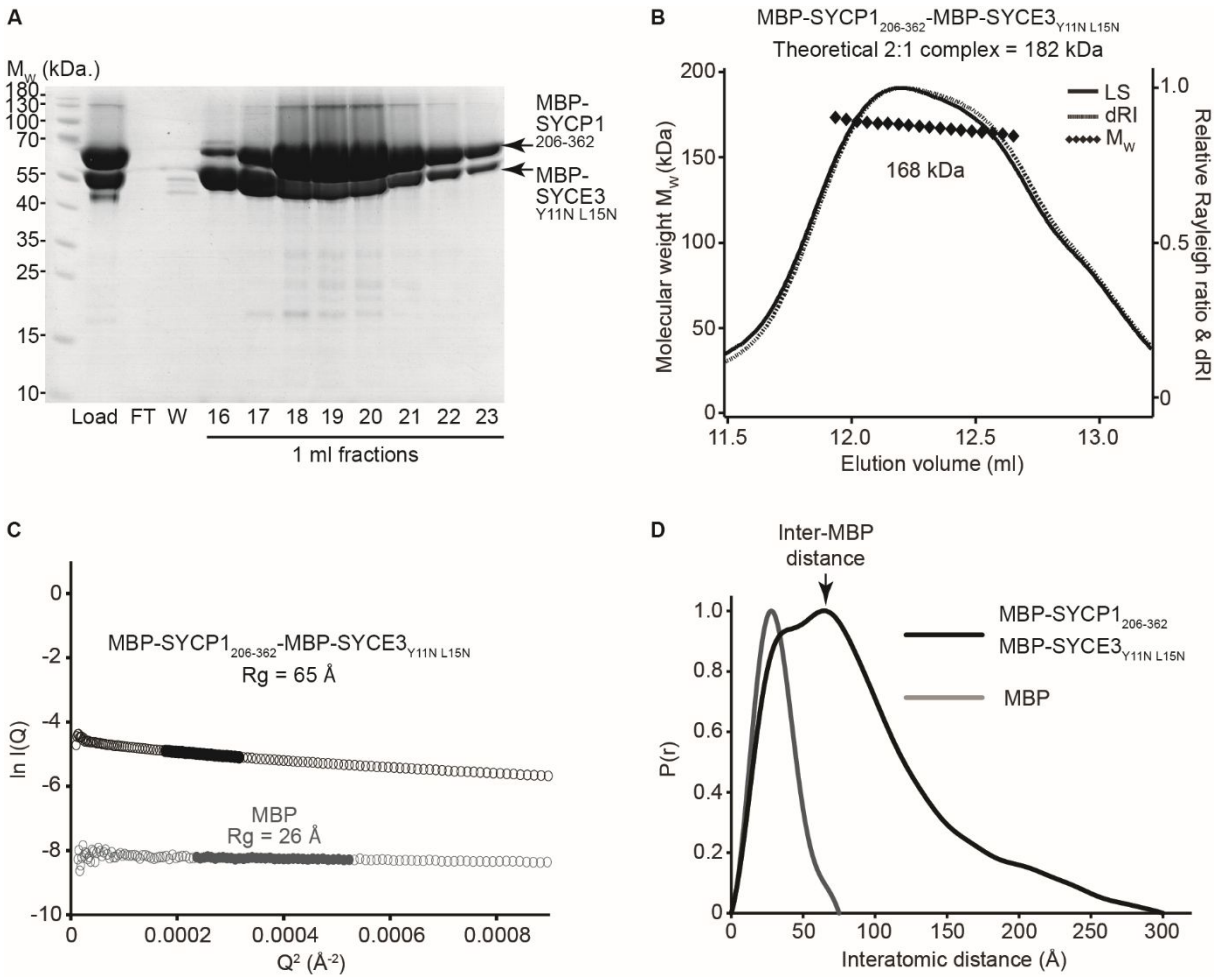


Figure 3.29. Purification and structural analysis of MBP-SYCP1$_{206-362}$-MBP-SYCE3$_{Y11NL15N}$. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (B) SEC-MALS analysis showing a single 168 kDa peak. A theoretical 2:1 complex has an estimated 182 kDa molecular weight suggesting a 2:1 complex formation. (C) Guinier analysis to determine the radius of gyration (Rg) as 65 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. MBP-SYCP1$_{206-362}$-MBP-SYCE3$_{Y11NL15N}$ has a maximum dimension of 300 Å, with an inter-MBP peak at 65 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey).

complex (Figure 3.29 B). The fusion complex was analysed by SEC-SAXS to assess the structural properties in comparison to the MBP-SYCP1$_{206-362}$-MBP-SYCE3 complex. The radius of gyration of the fusion complex was determined as 65 Å, which is comparable to MBP-SYCP1$_{206-362}$-MBP-SYCE3 (Figure 3.29 C).The distance distribution plot determines the maximum dimension in the MBP-SYCP1$_{206-362}$-MBP-SYCE3$_{Y11NL15N}$ sample to be 300 Å, with a prominent peak correlating to the distance between the three MBP tags (Figure 3.29 D). These structural features relate to the MBP-SYCP1$_{206-362}$-MBP-SYCE3 complex and indicates that the SYCE3 truncation and mutations have not affected the folding of the SYCP1-SYCE3 complex.

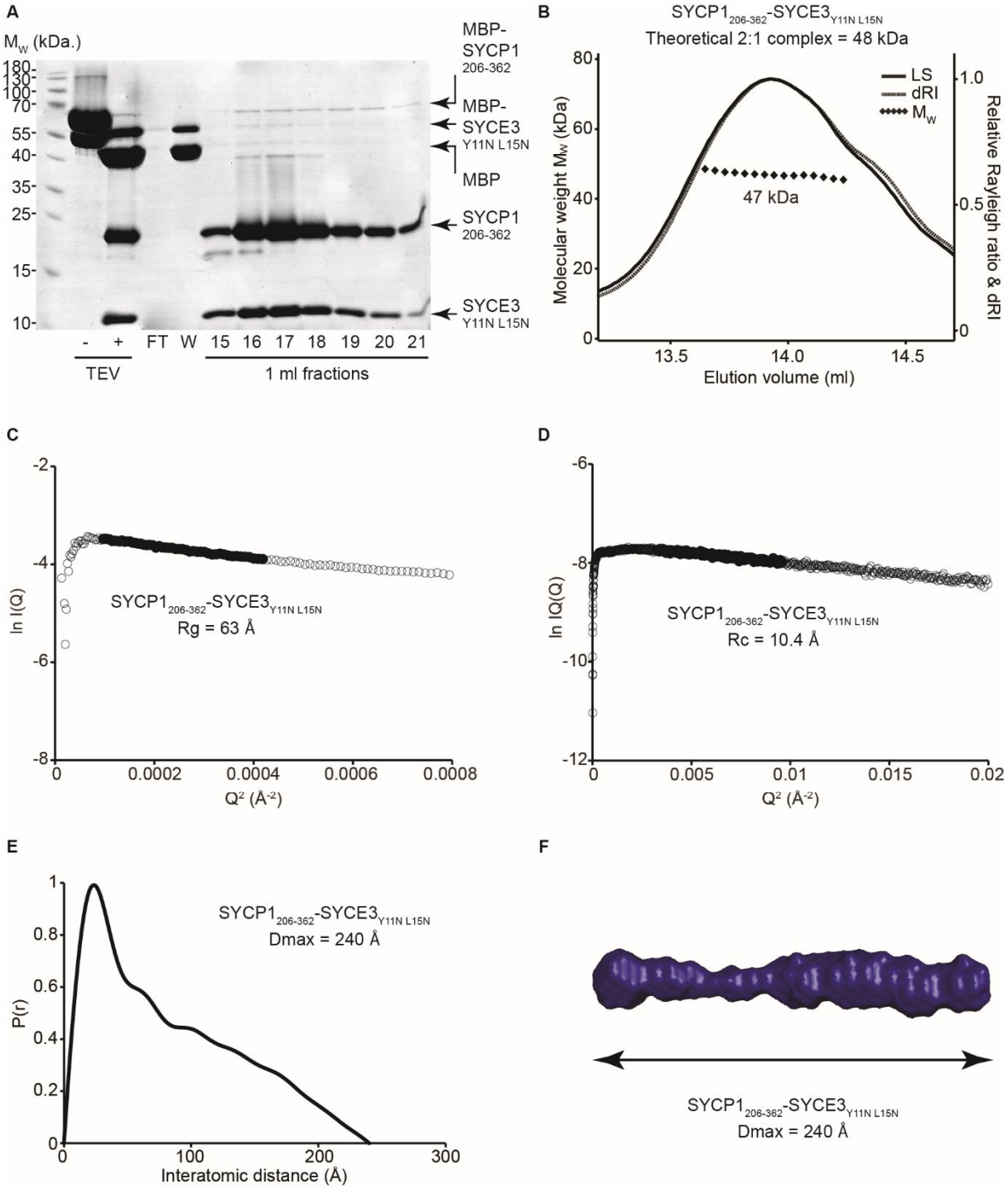


Figure 3.30. Structural analysis of SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$. (A) SDS-PAGE analysis showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. (B) SEC-MALS analysis showing the single 47 kDa peak. A theoretical 2:1 complex has an estimated molecular weight of 48 kDa, indicating a 2:1 complex formation. (C) Guinier analysis to determine the radius of gyration (Rg) as 63 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (D) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 10.4 Å. Clear circles are the data, solid circles represent the region used for the fit. (E) Paired real-space distribution plot. SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$ has a maximum dimension of 240 Å. (F) *Ab initio* DAMMIF model presented as the molecular envelope.

The cleaved SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$ complex was purified following TEV protease incubation and subsequent anion exchange chromatography (Figure 3.30 A). SEC-MALS analysis revealed the formation of a single protein species with a calculated molecular mass of 47 kDa, which corresponds to a 2:1 complex (Figure 3.30 B). SEC-SAXS analysis was carried out to determine the structural properties of the SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$ in solution. The radius of gyration was measured as 63 Å and the cross-sectional radius was determined as 10.4 Å (Figure 3.30 C and D). The real space analysis from the distance distribution plot demonstrates the elongated nature of the complex, with the maximum dimension of the sample estimated as 240 Å (Figure 3.30 E). *Ab initio* modelling was performed to visualise the molecular envelope of SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$, which demonstrates the extended rod shaped structure that mimics a similar shape when SYCP1 is bound to the full length SYCE3 molecule (Figure 3.30 F). The parameters determined by the SEC-SAXS experiment demonstrates that the SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$ complex has comparable structural properties to SYCP1$_{206-362}$-SYCE3 and is therefore a suitable candidate for crystallographic studies. However, upon commercial screening of the characterised SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$ complex, no crystals were detected.

The characterisation of the various SYCE3 mutations and truncations when in complex with SYCP1 have all resulted in the formation of a 2:1 SYCP1-SYCE3 complex, reiterating that both SYCP1 and SYCE3 are remodelled upon interacting. The various assemblies of SYCE3 are abolished upon SYCP1 binding, and thus, the biological context of the SYCE3 higher-order species is unknown. Therefore, a hypothesis for the formation of the larger species is that they may assemble to tether neighbouring SYCP1-SYCE3 units.

To address this hypothesis, the SYCP1$_{206-362}$-SYCE3 complex was mixed with free SYCE3, to investigate whether the SYCE3 assemblies would interact with the SYCP1-SYCE3 units. The pre-formed SYCP1$_{206-362}$-SYCE3 complex was incubated with separately purified SYCE3 before analysis by SEC, with the pre-formed complex and free SYCE3 also incubated in the same manner and analysed by SEC at the same concentration as a control. The SDS-PAGE analysis of the SEC elution profiles demonstrate that the SYCP1$_{206-362}$-SYCE3 complex does form a larger species with an excess of SYCE3 when compared to the control experiments (Figure 3.31). This implies that the free excess SYCE3 assembles with the preformed SYCP1$_{206-362}$-SYCE3 complex to form the larger species. This finding proposes that the biological context of the characterised SYCE3 higher-order assemblies are to function to tether lateral SYCP1-SYCE3 units in order to recapitulate the preliminary SYCP1 SC lattice for the recruitment of further central element proteins for SC maturation.

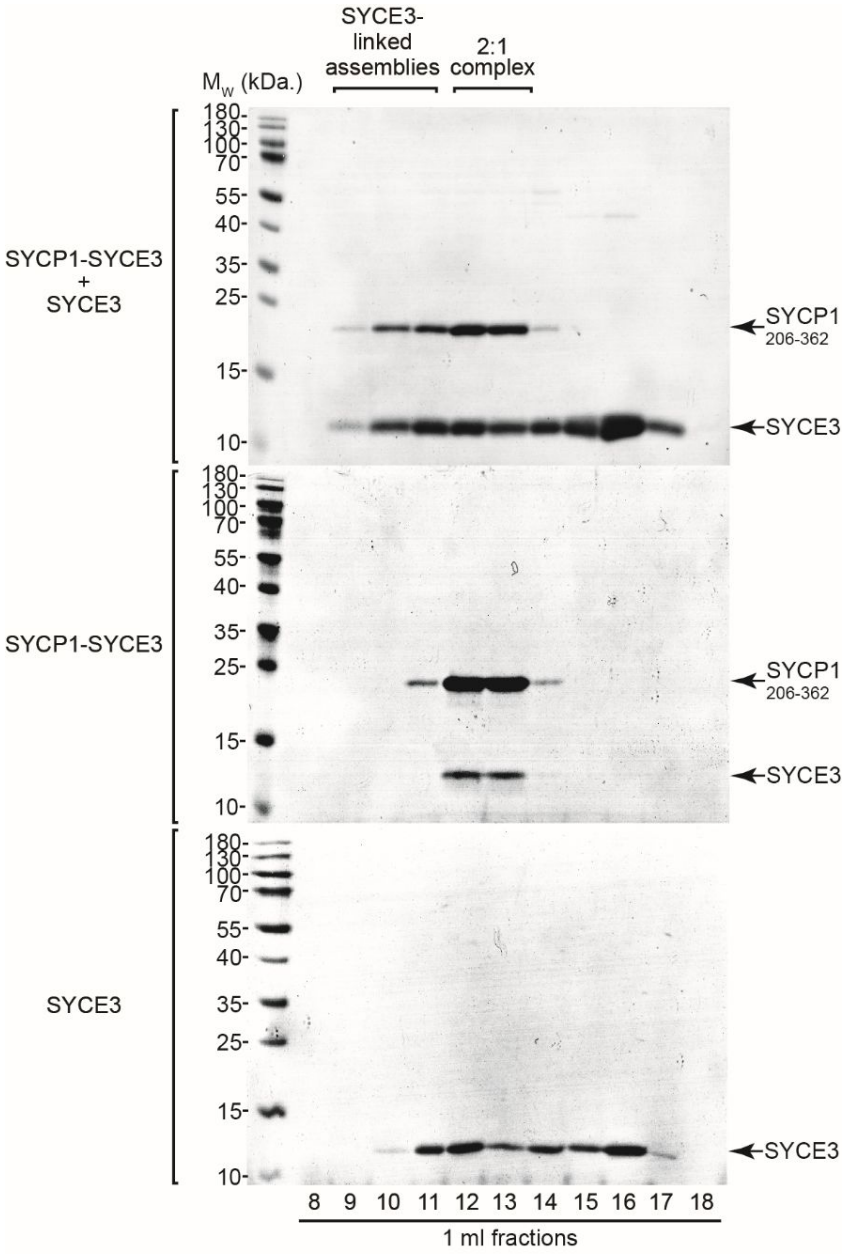


Figure 3.31. The SYCP1-SYCE3 complex forms larger species through tethering by SYCE3 assemblies. Analytical size exclusion chromatography was performed with the SYCP1$_{206-362}$-SYCE3 complex after incubation with free SYCE3. The SYCP1$_{206-362}$-SYCE3 complex and free SYCE3 were analysed individually as controls. The incubation with free SYCE3 resulted in a partial shift of both the SYCP1$_{206-362}$-SYCE3 complex and the free SYCE3, suggesting that the SYCE3 assemblies tether together the SYCP1$_{206-362}$-SYCE3 units. Experiments performed by Orla Dunne.

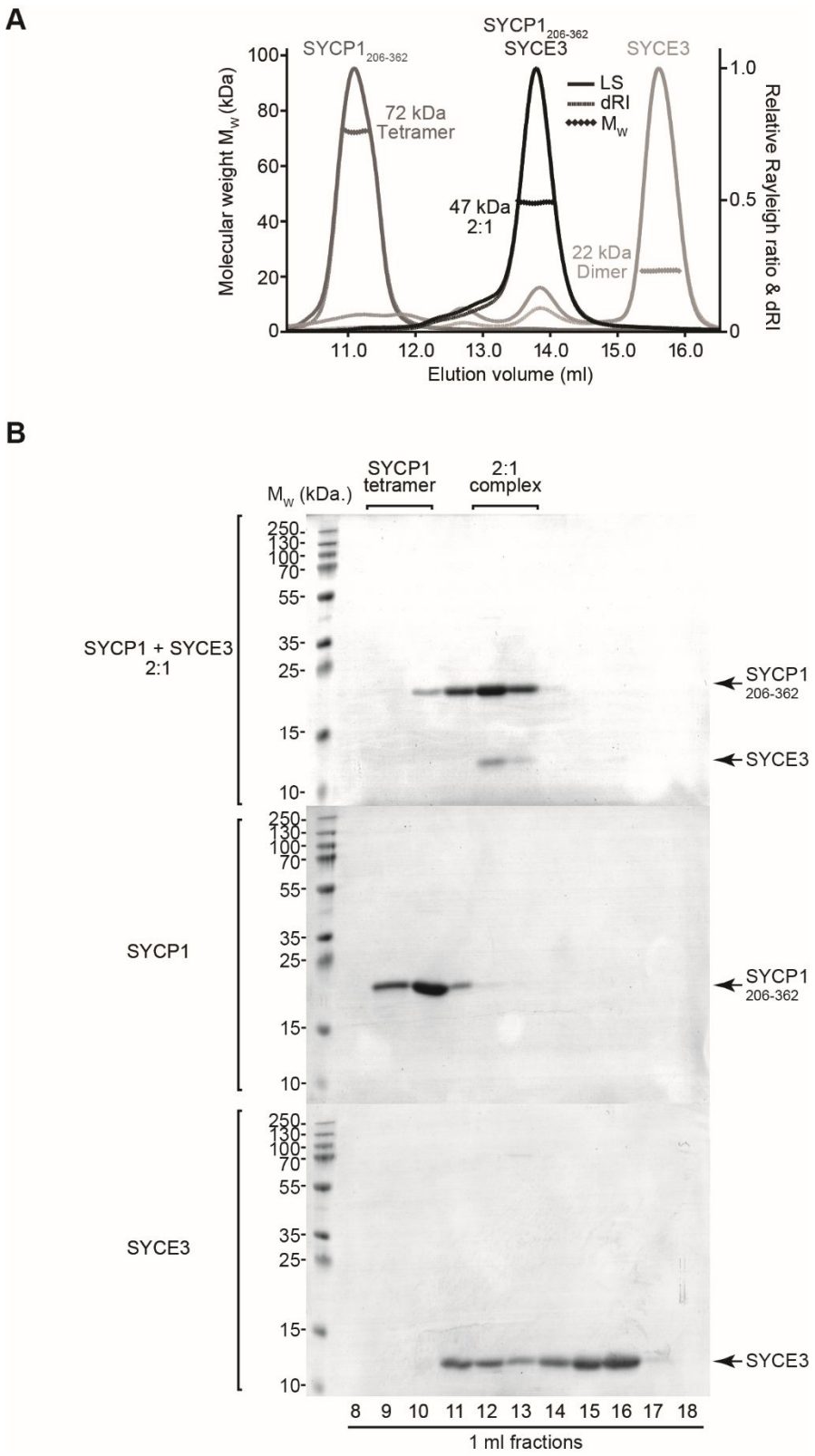


Figure 3.32. The SYCP1-SYCE3 complex is formed by remodelling of SYCP1 and SYCE3. (A) SEC-MALS analysis showing the overlaid profiles of the SYCP1 72 kDa tetramer, the 47 SYCP1-SYCE3 2:1 complex and the 22 kDa SYCE3 dimer with assembly. (B) Analytical size exclusion chromatography was performed after the incubation of the isolated $SYCP1_{206-362}$ homo-tetramer and assembling SYCE3 dimer at a 2:1 molar ratio, with the two species analysed separately as controls. The $SYCP1_{206-362}$-SYCE3 complex was formed by incubation of the isolated proteins. Experiment performed by James Dunce.

Previous analysis of the SYCP1-SYCE3 has been performed following co-purification of the complex. The analysis of the co-purified complex against the two proteins in isolation demonstrates that the tetrameric SYCP1 and the assembling SYCE3 dimer form a 2:1 complex (Figure 3.32 A). To validate that the proteins truly undergo conformational remodelling, the two proteins were purified separately and then mixed at a 2:1 SYCP-SYCE3 molar ratio for analysis by SEC. The SDS-PAGE analysis of the SEC profile, compared to the matching analysis of the proteins in isolation, demonstrates that the preformed SYCP1 tetramer and the preformed SYCE3 assembling dimer are indeed remodelled to form the 2:1 complex (Figure 3.32 B).

This analysis was also performed at a 2:10 ratio of SYCP1-SYCE3, where the 2:1 species is again formed with excess SYCE3 observed. Interestingly, an indication of the SYCE3-linked assemblies were also observed, suggesting that some of the excess SYCE3 was able to tether the remodelled 2:1 complexes (Figure 3.33).

To confirm that the larger assemblies are specifically formed through free SYCE3 assembly between SYCP1-SYCE3 units, the mixing experiment using the preformed SYCP1-SYCE3 complex was carried out with the non-assembling SYCE3 W41E Y44E mutant. The SEC profile demonstrates that the 2:1 complex does form the SYCE3-linked assemblies with addition of the non-assembling free SYCE3 (Figure 3.33). These findings indicate that the biological function of the various SYCE3 assemblies may be to tether the lateral SYCP1-SYCE3 complexes to stabilise the parallel central element cables (Hernandez-Hernandez *et al*., 2016).

3.2.6 SYCE3 acts as molecular glue to recruit other central element components

SYCP1 has been shown to self-assemble into a preliminary SC lattice that forms the initial SC chromosome tethering structure, with remodelling of this lattice triggered by the binding and incorporation of SYCE3. The remodelled lattice then requires the recruitment of the further four central element proteins in order to form the mature and fully functional SC tripartite structure.

To probe the network of central element interactions that may occur to form the complete central element structure, a combination of yeast 2-hybrid screening and recombinant amylose pulldown assays were performed to map these interactions. The transverse filament protein SYCP1 has been shown to only interact with SYCE3 in the central element, and therefore, SYCE3 is likely responsible for the incorporation of the other components.

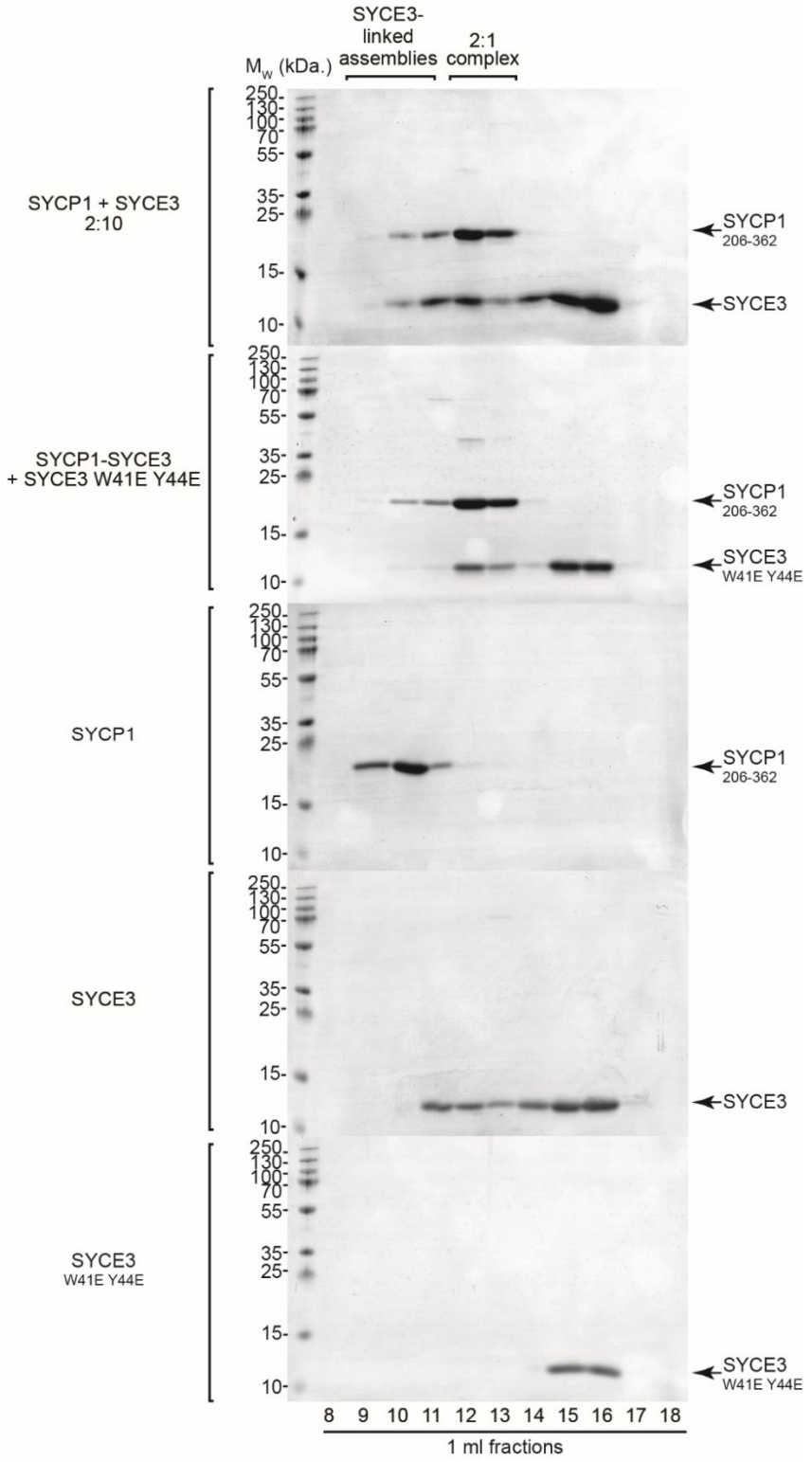


Figure 3.33. SYCP1-SYCE3 complexes are tethered by SYCE3 assembly. Analytical size exclusion chromatography was performed after the incubation of the isolated $SYCP1_{206-362}$ homo-tetramer and assembling SYCE3 dimer at a 2:10 molar ratio, with the two species analysed separately as controls. SYCE3 was remodelled into the SYCP1 2:1 complex, with free SYCE3 existing both in its smaller species but also some SYCE3 formed the linked SYCP1-SYCE3 assemblies. The $SYCP1_{206-362}$-SYCE3 complex was incubated with the isolated SYCE3 W41E Y44E non-assembling mutant, with the mutant analysed in isolation as a control. The 2:1 complex did not form larger assemblies and the mutant solely existed in the dimer/tetramer forms. Experiment performed by James Dunce.

A yeast 2-hybrid screen was performed with SYCP1, SYCE3 and the central element proteins SYCE1 and SIX6OS1. SIX6OS1 is the most recently identified central element protein, and thus far, has been shown to interact solely with SYCE1 in the central element (Gómez-H *et al.*, 2016). The screen highlights the characterised SYCP1-SYCE3 interaction and their respective self-associations, as previously identified. Self-associations were also detected for SYCE1 along with the identification of the direct interaction between SYCE1 and SIX6OS1 and a weak interaction was also observed between SYCE1 and SYCE3 (Figure 3.34 A).

To explore whether SYCE1 and SYCE3 do directly interact, an amylose pulldown assay was performed, with the eluates analysed by SDS-PAGE. The pull-down was performed with the structural dimeric helical core of SYCE1, corresponding to residues 25-179 (Dunne, unpublished findings), fused to an MBP affinity tag that was co-expressed with His_6 tagged SYCE3 fragments. The full length SYCE3 was found to bind the structural SYCE1 core, thus confirming that the two central element proteins directly interact (Figure 3.34 B). To probe whether the interaction occurs with a specific SYCE3 helix, the binding was tested with each half of the SYCE3 chain, 1-52 and 54-88. This analysis demonstrates that binding only takes place with the first half of SYCE3, 1-52, indicating that the first helix may be involved in SYCE1 binding (Figure 3.34 B). To further dissect the interacting region, the N-terminal region of SYCE3 was truncated. The SYCE3 12-88 amino acids were found to no longer bind to SYCE1 (Figure 3.34 B). This SYCE3 12-88 region is indicated as the binding surface to SYCP1, and therefore, the SYCE3-SYCE1 interaction may occur through the SYCE3 1-11 N-terminal region, to allow for binding to both SYCP1 and SYCE1 simultaneously. SIX6OS1 has, thus far, only been found to interact with SYCE1 in the SC, and perhaps the SYCE3 interaction with SYCE1 also recruits SIX6OS1-SYCE1.

A constitutive complex has been identified between the central element proteins SYCE2-TEX12, and this complex is thought to drive SC extension through the propensity for filament formation (Davies *et al.*, 2012). To further elucidate the central element network, yeast 2-hybrid and recombinant pull-down assay experiments were carried out to probe whether SYCE3 binds to SYCE2-TEX12. The Y2H screen demonstrates the established interaction between SYCE2 and TEX12, and a direct interaction between SYCE2 and SYCE3 that has previously been identified by co-immunoprecipitation (Schramm *et al.*, 2011) (Figure 3.35 A). To study whether the SYCE2-SYCE3 interaction occurs when SYCE2 is bound to TEX12, a recombinant amylose pull-down assay was performed through co-expression of MBP-tagged SYCE2-TEX12 with His-tagged SYCE3, with further experiments performed with MBP-SYCE2, MBP-TEX12 and free MBP as controls. The SDS-PAGE analysis of the amylose eluates showed that SYCE3 binding was detected to the SYCE2-TEX12 complex

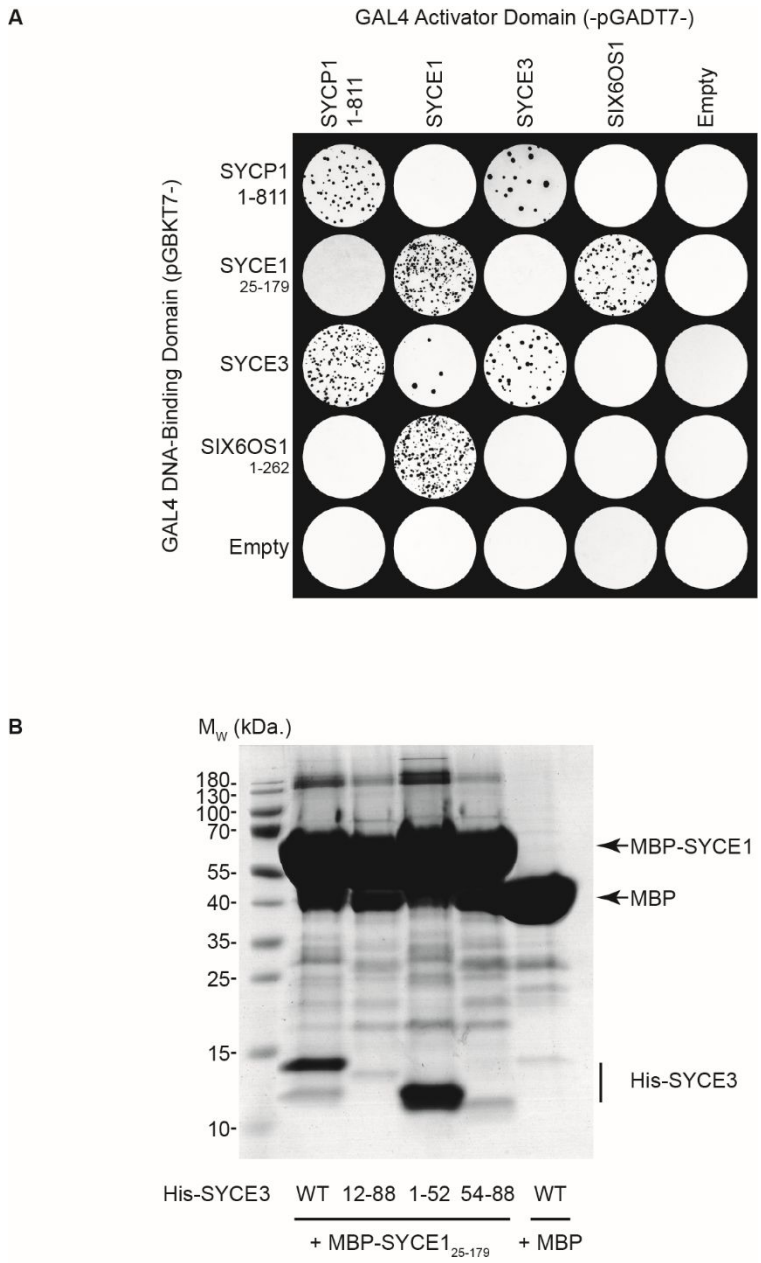


Figure 3.34. SYCE3 directly interacts with SYCE1. (A) Yeast two-hybrid grid with SYCP1, SYCE1, SYCE3 and SIX6OS1. Sequences correspond to the full-length human proteins, unless boundaries are specified. SYCP1 was tested through a C-terminal truncation (1-811) as the full length 976 amino acid sequence is long for the system. The SYCE1$_{25\text{-}179}$ and SIX6OS1$_{1\text{-}262}$ fragments were used in the bait strain to prevent auto-activation. Y187 (pGBKT7-bait) strains were mated with Y2H Gold (pGADT7-target) strains and plated on SD/-Leu/-Trp/-Ade/-His/x-α-gal plates for colony selection, subject to reporter gene activation (ADE1, HIS3 and MEL1) upon positive interactions. (B) Amylose pull-down assay, performed by Orla Dunne. SYCE1$_{25\text{-}179}$ was fused to a His-MBP affinity tag, including free His-MBP as a control, for bacterial co-expression with His-SYCE3 fragments. Clarified lysates were incubated with amylose resin, followed by washing of the resin, and elution of bound protein. Samples were normalised for protein concentration for analysis by SDS-PAGE. SYCE1$_{25\text{-}179}$ was found to interact with the full length SYCE3 and the 1-52 region, but not the 12-88 and 54-88 fragments. The free MBP control did not bind to His-SYCE3.

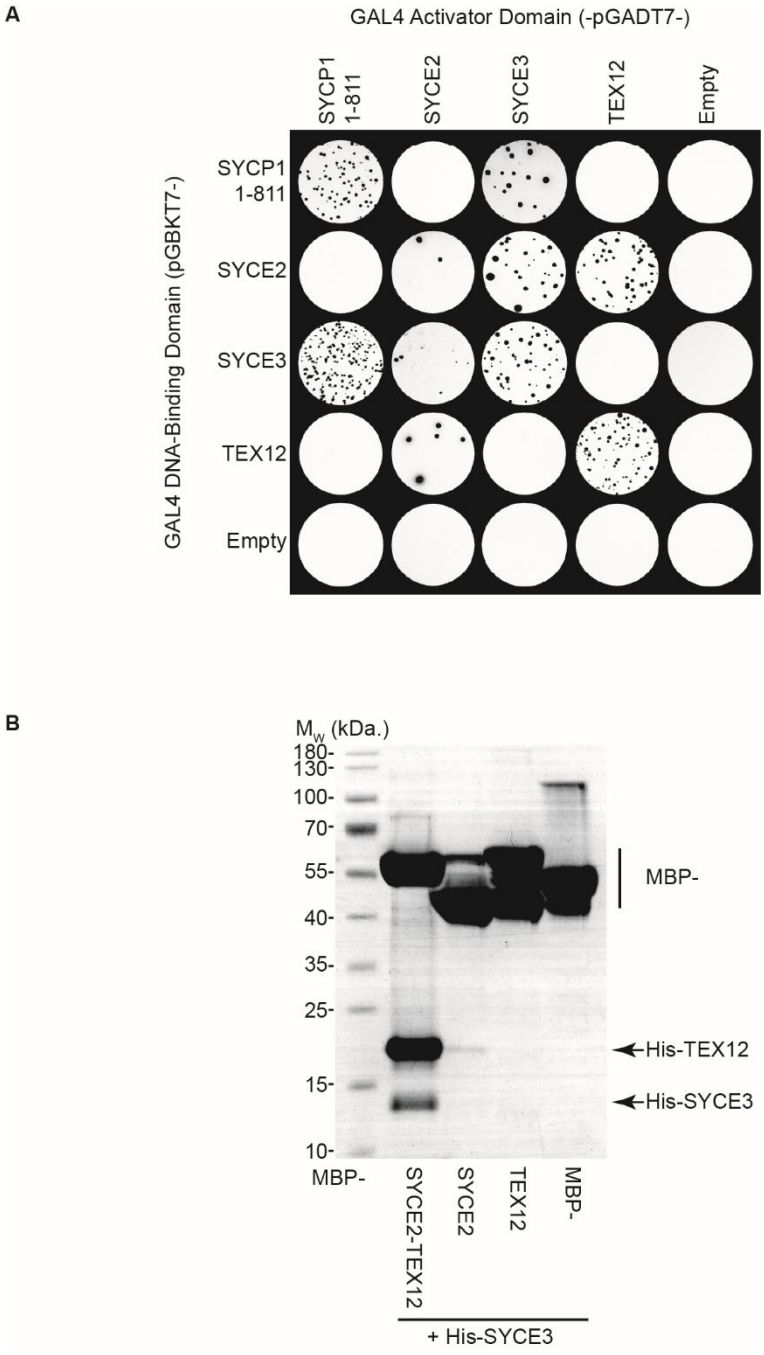


Figure 3.35. SYCE3 directly interacts with SYCE2 in the SYCE2-TEX12 complex. (A) Yeast two-hybrid grid with SYCP1, SYCE2, SYCE3 and TEX12. Sequences correspond to the full-length human proteins, except SYCP1 was tested through a C-terminal truncation (1-811) as the full length 976 amino acid sequence is long for the system. Y187 (pGBKT7-bait) strains were mated with Y2H Gold (pGADT7-target) strains and plated on SD/-Leu/-Trp/-Ade/-His/x-α-gal plates for colony selection, subject to reporter gene activation (ADE1, HIS3 and MEL1) upon positive interactions. (B) Amylose pull-down assay. MBP-SYCE2$_{57-165}$-His-TEX12, MBP-SYCE2$_{57-165}$, MBP-TEX12 and free MBP were co-expressed with His-SYCE3. Clarified lysates were incubated with amylose resin, followed by washing of the resin, and elution of bound protein. Samples were normalised for protein concentration for analysis by SDS-PAGE. The MBP-SYCE2$_{57-165}$-His-TEX12 complex was found to interact with His-SYCE3. His-SYCE3 didn't interact with MBP-SYCE2$_{57-165}$, MBP-TEX12 or free MBP. Heavy degradation of MBP-SYCE2$_{57-165}$ was observed due to the instability of SYCE2 without TEX12 bound.

(Figure 3.35 B). SYCE3 was not detected to interact with SYCE2 or TEX12 in isolation, although severe degradation was observed with SYCE2 due to its instability in the absence of TEX12 (Figure 3.35 B). Therefore, by combing the Y2H and pull-down assay findings, SYCE3 was found to bind to the SYCE2-TEX12 complex, most likely via interacting with SYCE2, which may lead to the incorporation of the SYCE2-TEX12 into the central element for elongation in order to form the fully mature SC structure.

Transmission EM experiments (TEM) have demonstrated the filamentous assembly of the SYCE2-TEX12 complex (Davies *et al*., 2012). To investigate whether SYCE3 binding to SYCE2-TEX12 would affect filament formation, further TEM experiments were performed with SYCE3 present. SYCE3 was incubated with the SYCE2-TEX12 complex, at a 2:1 molar ratio, before fixing to the TEM grids for analysis, with SYCE3 and SYCE2-TEX12 control samples prepared by the same handling and protein concentrations. The SYCE2-TEX12-SYCE3 sample revealed extensive filamentous networks forming mesh-like assemblies (Figure 3.36 A). The SYCE2-TEX12 control formed filamentous assemblies, as previously characterised (Davies *et al*., 2012), and no species were detected for SYCE3 (Figure 3.36 B and C). Without further experimentation, it is not known whether SYCE3 within the SYCE2-TEX12-SYCE3 sample is definitely bound to SYCE2-TEX12, but it does indicate that the presence of SYCE3 does not disrupt the filament formation. With the assumption that SYCE3 is bound to SYCE2-TEX12, these findings imply that the SYCE3 binding to the SYCE2-TEX12 complex may propagate and encourage the SYCE2-TEX12 filaments, perhaps occurring to form continuous interactions along the central element as SYCE2-TEX12 drives elongation.

3.3 Discussion

3.3.1 SYCP1 and SYCE3 are remodelled upon complex formation

The SYCP1 N-terminus has been shown to form a tetramer with the capacity to undergo higher-order assembly through recursive head-to-head interactions (Dunce *et al*., 2018). SYCE3 exists as a compact dimer with the propensity to form larger species. A direct interaction has been identified between the SYCP1 N-terminus and SYCE3, with analysis of the SYCP1-SYCE3 complex revealing the formation of a 2:1 complex. Therefore, both SYCP1 and SYCE3 undergo conformational remodelling upon binding (Figure 3.37 A).

Through mutation studies of SYCE3, the opening of the SYCE3 loop has been found to promote higher-order assembly in combination with staggered lateral interactions, with the compact structure solely existing as a dimer. SYCP1 has been shown to only bind to the open loop SYCE3 form, in order to form a three helical bundle. The N-termini of the three chains

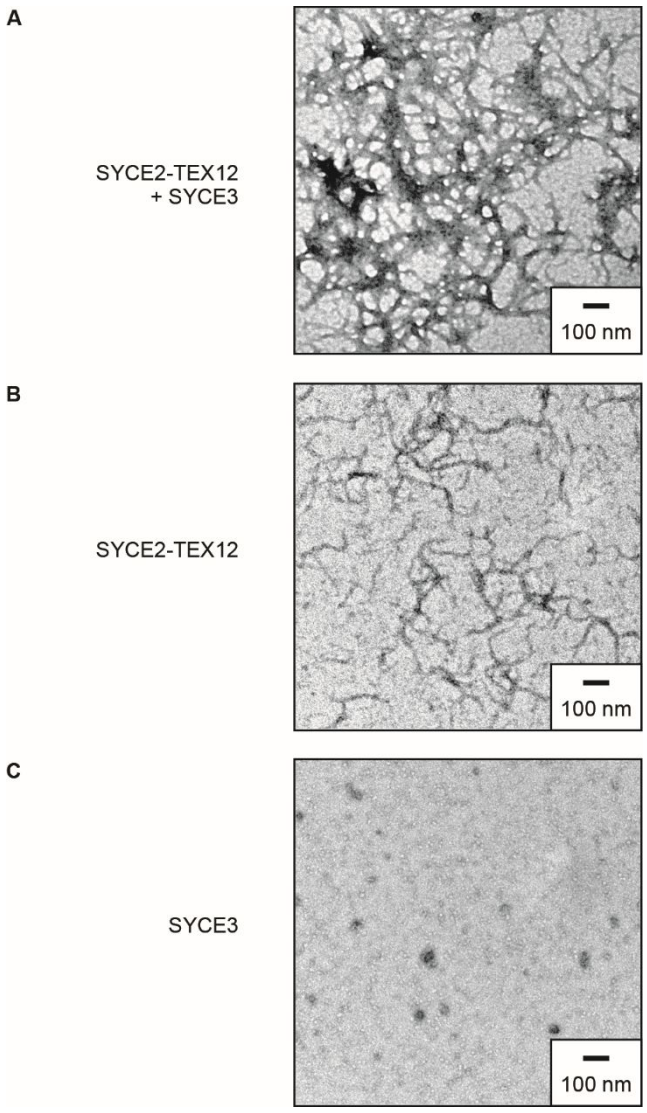


Figure 3.36. Filamentous assembly of SYCE2-TEX12-SYCE3. (A-C) Transmission electron microscopy experiments following fixation of samples to carbon grids and negative staining by uranyl acetate for imaging. (A) The $SYCE2_{57-165}$-$TEX12_{49-123}$ complex was incubated with free SYCE3 to form large filamentous networks. (B) The $SYCE2_{57-165}$-$TEX12_{49-123}$ complex was imaged without SYCE3 at the same concentration as a control for filament formation. (C) Free SYCE3 was imaged without the $SYCE2_{57-165}$-$TEX12_{49-123}$ complex to demonstrate no filament formation in its absence.

have been indicated to be in close proximity, and therefore, a parallel heterotrimeric complex is likely formed (Figure 3.37 A). Through the inclusion of further SYCP1 N-terminal residues, the SYCP1-SYCE3 2:1 complex has been demonstrated to undergo head-to-head interactions through the SYCP1 101-111 region to form an extended 4:2 complex (Figure 3.37 B). At high concentrations, large molecular weight species are formed, indicating that the head-to-head interactions may occur recursively between staggered 2:1 complexes. This corresponds to the reports that SYCE3 is required to stabilise SYCP1 in the central element bilayer, in order for central element progression (Hernandez-Hernandez *et al*., 2016).

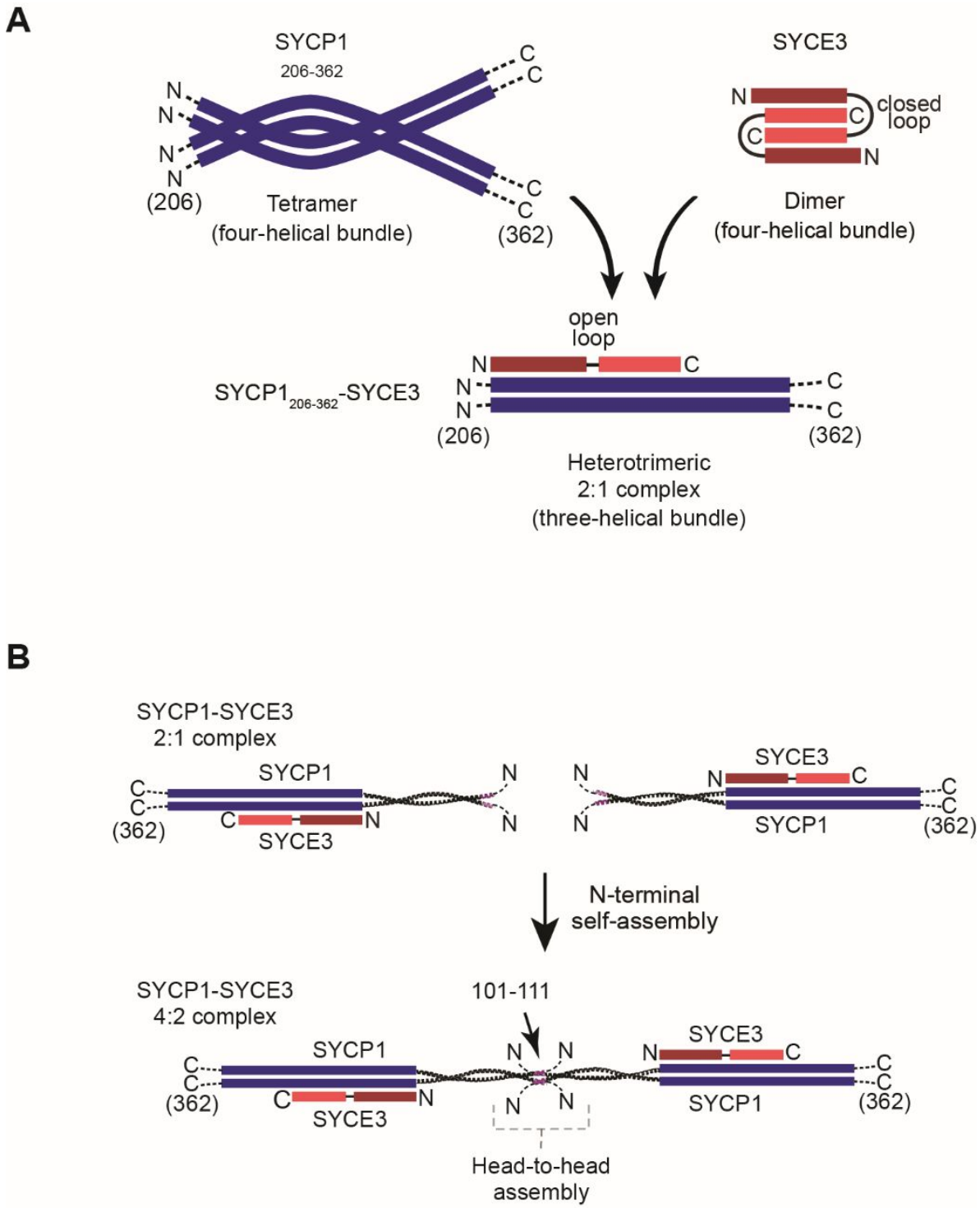


Figure 3.37. Models for the formation of the SYCP1-SYCE3 complex. (A) Schematic to show the remodelling of the respective SYCP1 tetramer and SYCE3 dimer to form the elongated three helical 2:1 SYCP1-SYCE3 complex. (B) Schematic to demonstrate the head-to-head assembly that occurs between opposing 2:1 complexes to form an extended 4:2 complex. The head-to-head assembly is driven by the SYCP1 101-111 residues.

3.3.2 SYCE3 forms larger species that tether SYCP1-SYCE3 complexes

The self-assembly of the SYCE3 dimer has been characterised, with the larger species requiring opening of the SYCE3 loop. When SYCE3 is forced into the open conformation, only large species form, whereas when the loop is forced closed, only dimers are formed. Therefore, the formation of the intermediate species requires both SYCE3 conformations and it is proposed that open loop linear SYCE3 chains are capped by the closed loop structures (Figure 3.38 A). This proposed organisation of SYCE3 assembly allows the maintenance of the four helical anti-parallel bundle that is formed in the crystal structure.

The 2x dimer structure formed through the open and closed chains is seemingly enhanced by the Y11 and L15 residues (Figure 3.38 A), with mutation of these residues abolishing SYCE3 assembly. The higher-order assemblies are then thought to be formed by staggered lateral interactions between SYCE3 dimers or 2x dimers in an iterative manner (Figure 3.38 A). These staggered lateral interactions rely on the SYCE3 amino acid residues W41 and Y44 that are surface exposed in the dimeric SYCE3 structure (see Section 3.2.4). The mutation of these residues abolishes the higher-order assemblies but the 2x dimer species is still formed. Therefore, a model of SYCE3 assembly is proposed through coordination of loop opening, stabilisation by capping of closed loop structures and staggered lateral interactions mediated by the P53 loop residue, N-terminal Y11 and L15 residues and the surface exposed W41 and Y44 residues, respectively (Figure 3.38 A).

These SYCE3 assemblies were then tentatively shown to interact with the SYCP1-SYCE3 complex, suggesting that the function of these assemblies may be to tether the neighbouring SYCP1-SYCE3 units in order to progress the maturation of the SC central element. A model is proposed whereby the preliminary SYCP1 lattice is disrupted at the SYCP1 tetramerisation site through SYCE3 binding. This SYCE3 binding enables the formation of the 2:1 elongated structures in the central element that would occur on either side of the SC midline with SYCP1 molecules spanning from opposing chromosomal axes (Figure 3.38 B). The disrupted lattice may then be stabilised by assembly of free SYCE3 that can then form between lateral SYCP1-SYCE3 units. It remains unclear exactly how the free SYCE3 assembly would interact with the SYCP1-SYCE3 units. The open loop extended SYCE3 chain is proposed to interact with SYCP1 in the 2:1 complex from the SYCE3 mutational studies, which means that the antiparallel four helical bundle that forms in all SYCE3 conformations proposed does not occur upon interaction with SYCP1. Therefore, this may provide an interaction surface for the SYCE3 assemblies to initiate from and provides a mechanism for capping of the assemblies from recursive SYCP1-SYCE3 units in the central element (Figure 3.38 B).

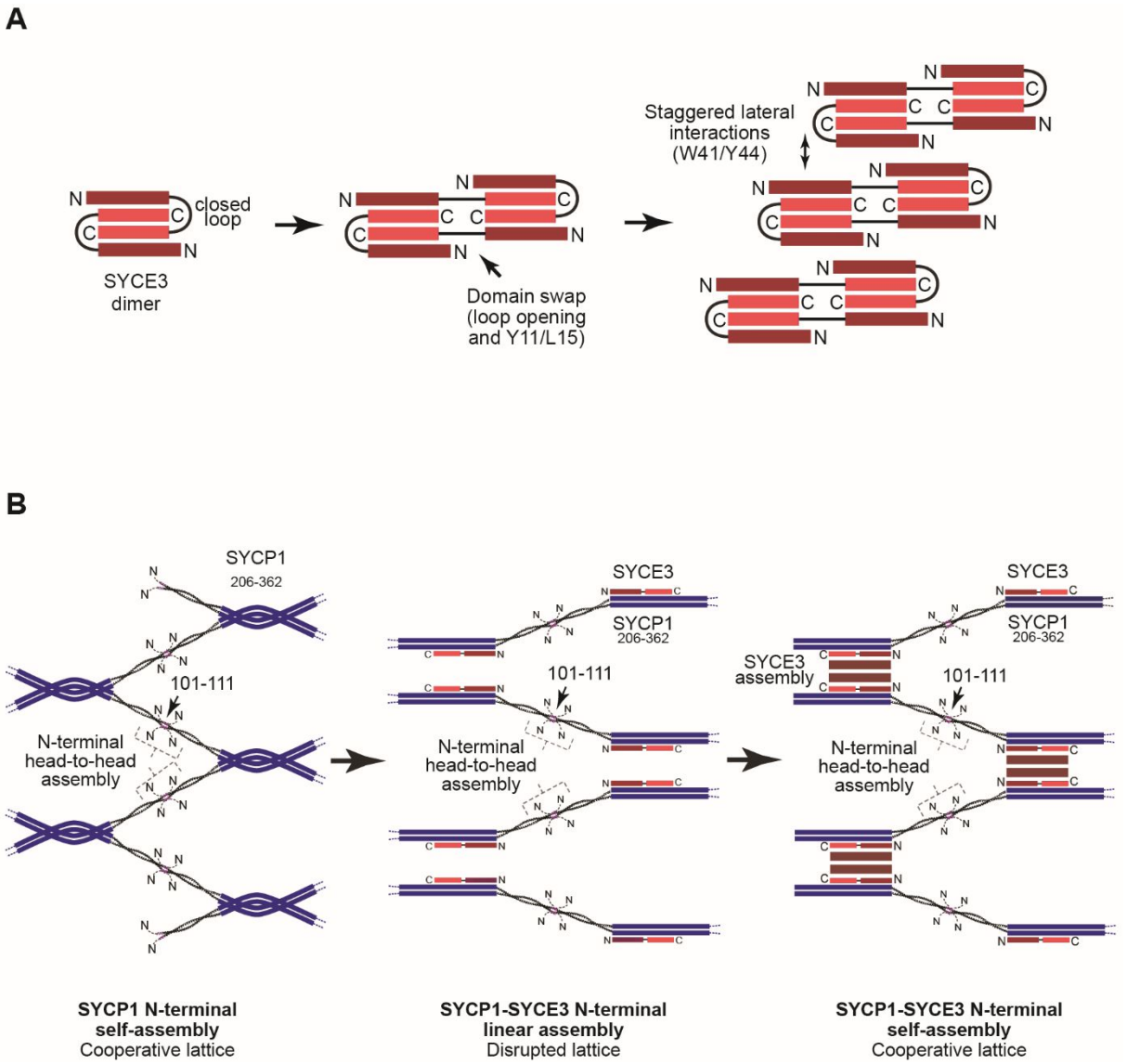


Figure 3.38. SYCE3 assembles between lateral SYCP1-SYCE3 units. (A) Schematic to show the model for SYCE3 assembly upon domain swapping, driven through SYCE3 loop opening (stabilised by the Y11 and L15 residues) and staggered lateral interactions (driven by W41 and Y44). (B) Schematic to show the model for SC lattice maturation upon SYCE3 incorporation. The cooperative SYCP1 lattice is disrupted at the SYCP1 tetramer sites upon SYCE3 binding to form the 2:1 complexes. The SYCP1-SYCE3 units are then tethered through self-assembly of SYCE3. SYCE3 assemblies depicted as maroon rectangles.

The SYCP1 lattice is therefore suggested to be remodelled through SYCP1-SYCE3 complex formation, and subsequent tethering by SYCE3, to produce an intermediate lattice where the SYCP1 N- and C-terminal self-assembly interactions are maintained, but with stabilisation of the central region (Figure 3.38 B). This model is consistent with the elucidation of the central element existing as two parallel cables (Schücker *et al.*, 2015).

It must be highlighted that the formation of these SYCE3 higher-order assemblies may not necessarily be representative of the behaviour of SYCE3 *in vivo*. The *in vitro* analysis presented in this thesis was performed with material produced by over-expression in bacterial cells. It is possible that the formation of the assemblies may be as a result of the expression method, which may cause some protein aggregation. However, even if it is the case that the assemblies are an artefact of bacterial over-expression, the dissection of the assembly mechanism and identification of key residues involved does suggest that the assemblies are specific structures, rather than non-specific aggregation, which may indeed occur within the SC to stabilise the central element.

3.3.3 SYCE3 acts as molecular glue to form the central element network

SYCE3 acts as the bridging molecule between the transverse filament and the central element. Further direct interactions have also been identified between SYCE3 and the central element proteins SYCE1 and SYCE2. SYCE1 in turn has been shown to bind to the most recently discovered SC component SIX6OS1, as previously reported (Gómez-H *et al*., 2016). Due to the recruitment patterns of the central element proteins, it is postulated that SYCE3-SYCE1-SIX6OS1 forms the central element initiation complex. A direct interaction was detected between SYCE3 and SYCE2, with the interaction also detected between SYCE3 and the SYCE2-TEX12 complex. SYCE2-TEX12 form filamentous structures (Davies *et al*., 2012) and are thought to drive elongation and maturation of the central element, and consequently are required to form the fully assembled SC (see Chapter four). Therefore, SYCE3 may bind periodically along the SYCE2-TEX12 filaments to guide the assemblies for incorporation into the central element. The recruitment of the further central element proteins, and the resultant sub-structures formed, enables the maturation of the SYCP1-SYCE3 lattice to form the mature and fully functional SC.

A model for SC central element maturation is proposed through the network of interactions primarily mediated by SYCE3 (Figure 3.39). The SYCP1-SYCE3 complexes, as described above, may be stabilised by tethering of SYCE3 assemblies. The SYCE1-SIX6OS1 proteins, which have been shown to interact *in vivo* (Gomez *et al*., 2016), may then bind to SYCE3 in the central element to stabilise the central element architecture. As the SC has a 3D organisation, the central element requires not only longitudinal progression, but also vertical support to stabilise the depth of the complex. The SYCE2 and TEX12 central element proteins, which form filamentous structures upon interacting (Davies *et al*., 2012), have been shown to be required for maturation of the SC to elongate the central element along the length of the homologous chromosomes (Bolcun-Filas *et al*., 2007) (Hamer *et al*., 2006). Therefore, it is

proposed that the SYCE1-SIX6OS1 complex may provide the vertical support in the central element through interactions with SYCE3 to form the central element initiation complex (Figure 3.39). The SYCE2-TEX12 complex is then proposed to form the central element elongation complex that may be recruited to the central element by SYCE3 binding. The SYCE2-TEX12 complex may then assemble in a longitudinal manner in order to propagate central element assembly for SC maturation (Figure 3.39).

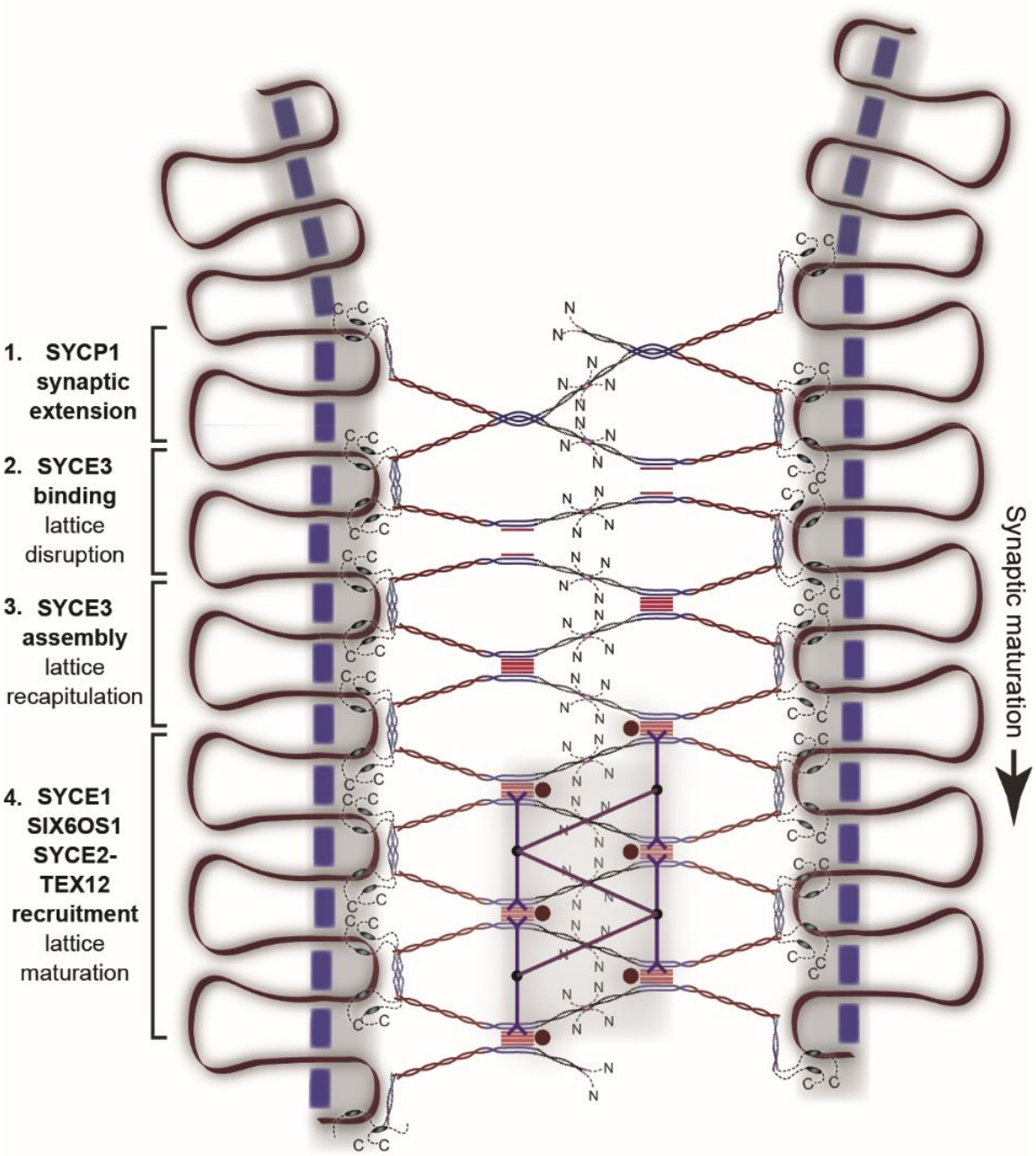


Figure 3.39. Model for synaptic progression upon SC maturation. Schematic to show the model of SYCP1 self-assembly, disruption and recapitulation of the SYCP1 lattice through SYCE3 binding, and recruitment of the further central element proteins to provide the depth and length for central element assembly in order to drive homologous chromosome synapsis. SYCE3 assemblies depicted as maroon rectangles. Brown circles represent SYCE1-SIX6OS1 vertical support and purple lines represent SYCE2-TEX12 assembly.

Chapter four.

Fibrous assembly of SYCE2-TEX12 mediated through the TEX12 C-terminus

4.1 Introduction

4.1.1 Recruitment of central element components

The SC lateral element components are involved in the organisation and compaction of DNA to prime the chromosome axes for synapsis (Syrjanen *et al*., 2014). The transverse filaments then self-assemble along the chromosome axes to form a preliminary lattice through recursive head-to-head interactions in the central element (Dunce *et al*., 2018). The central element components are then required to form the mature and functional SC (Figure 4.1). The mammalian central element proteins can be further grouped into two sub-complexes; the initiation complex and the elongation complex. SYCE3 is the only central element protein to interact with the SYCP1 transverse filament (Hernandez-Hernandez *et al*., 2016) and this interaction disrupts the preliminary lattice (see Chapter three). SYCE3 then interacts with SYCE1 (Lu *et al*., 2014), which in turn binds to SIX6OS1 (Gómez-H *et al*., 2016). As the SYCE3, SYCE1 and SIX6OS1 proteins exhibit similar phenotypes upon their absence, whereby SYCP1 polymerisation occurs at short regions but without propagation and assembly of the central element along the chromosomes, they perhaps form an initiation complex to alter the SYCP1 lattice and to form the depth of the complex (Schramm *et al*., 2011) (Bolcun-Filas *et al*., 2009) (Gómez-H *et al*., 2016).

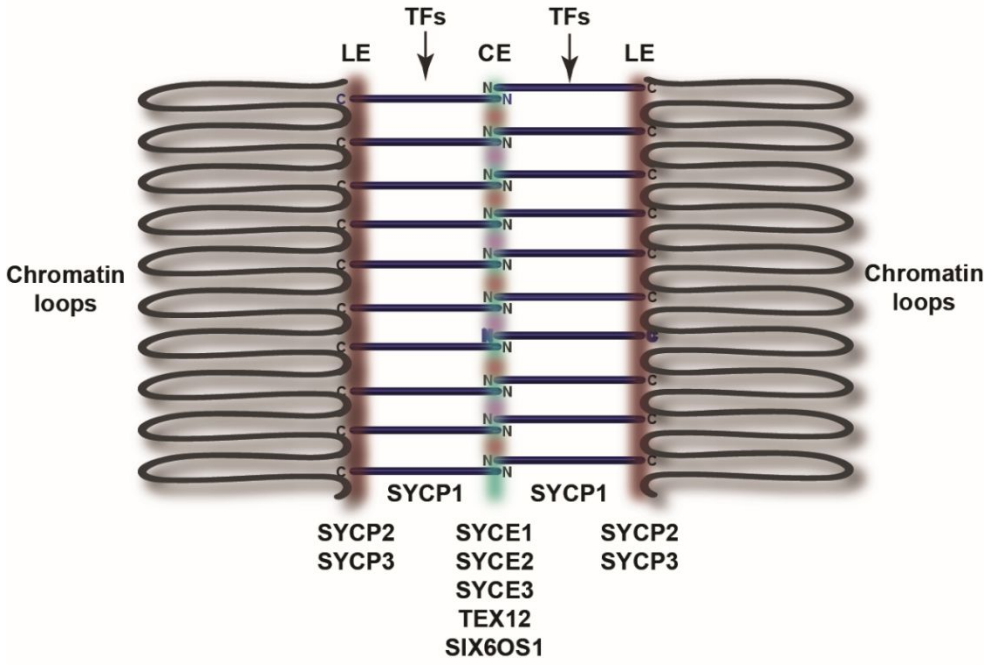


Figure 4.1. Schematic of the mammalian synaptonemal complex. LE = lateral element. TFs = transverse filaments. CE = central element. Schematic depicting the chromatin loops of the homologous chromosomes with the assembly of the lateral elements at the axes (red). The transverse filaments (purple tubes) bridge between the chromosome axis and the central element (orange, green and pink). The lateral elements are comprised of SYCP2 and SYCP3. SYCP1 forms the transverse filaments and the central element is formed by SYCE1, SYCE2, SYCE3, TEX12 and SIX6OS1.

The SYCE2 and TEX12 central element proteins have been shown to co-localise with the SC central element (Hamer *et al*., 2006), with both proteins giving rise to an intermediate phenotype upon their absence, where short regions of synapsis are observed, but these regions fail to extend in order to propagate the SC assembly between the entire length of chromosomes (Bolcun-Filas *et al*., 2007) (Hamer *et al*., 2006). Therefore, the SYCE2-TEX12 complex is thought to be responsible for the assembly and maturation of the SC.

4.1.2 SYCE2-TEX12 form an equimolar complex that undergoes higher-order assembly

The human SYCE2 protein is 218 amino acids in length, with high conservation in the central region of the protein and low conservation at the N- and C-termini (Figure 4.2 A and B). The conserved central region of the protein is predicted to be highly alpha helical, with some coiled-coil prediction, with the poorly conserved N- and C-termini predicted to be largely unstructured (Figure 4.2 A and C). Therefore, the centre of the SYCE2 protein corresponding to amino acids 57-165 is predicted to comprise the conserved helical core, hereby referred to as $SYCE2_{\alpha}$ (Figure 4.2 D).

TEX12 is a smaller protein of 123 amino acids in humans (Figure 4.3 A). TEX12 is highly conserved throughout the sequence, particularly in the C-terminal half (Figure 4.3 A and B). The N-terminal region of TEX12 is predicted to be unstructured, with the highly conserved C-terminal half of the protein predicted to be almost entirely alpha helical (Figure 4.3 A and C). Therefore, the C-terminal region corresponding to amino acids 49-123 is predicted to comprise the structured helical core of TEX12, hereby referred to as $TEX12_{\alpha}$ (Figure 4.3 D).
The proposed SYCE2-TEX12 complex from the *in vivo* co-localisation was validated by co-immunoprecipitation assays (Hamer *et al*., 2006) and by the identification of a direct interaction through yeast 2-hybrid and recombinant protein studies (Davies *et al*., 2012). Co-purification and subsequent analysis of the core helical regions of SYCE2 and TEX12 reveals the formation of a 4:4 complex. The proteins were preliminarily characterised separately by analysis of SYCE2 and TEX12 as fusion proteins to an MBP affinity tag, with MBP-SYCE2 found to form a tetramer and MBP-TEX12 forming a dimer. As the proteins were stated to be unstable in isolation, the equimolar unit was described as a constitutive complex formed from a SYCE2 tetramer interacting with two TEX12 dimers (Davies *et al*., 2012).

Further analysis of the 4:4 complex demonstrated the capacity for SYCE2-TEX12 to undergo self-assembly. Large molecular weight species were also detected in solution from the core helical complex, with electron microscopy experiments revealing the formation of filamentous structures that mimic the electron dense central element as observed from electron micrographs of intact mice SCs (Figure 4.4 A and B) (Davies *et al*., 2012).

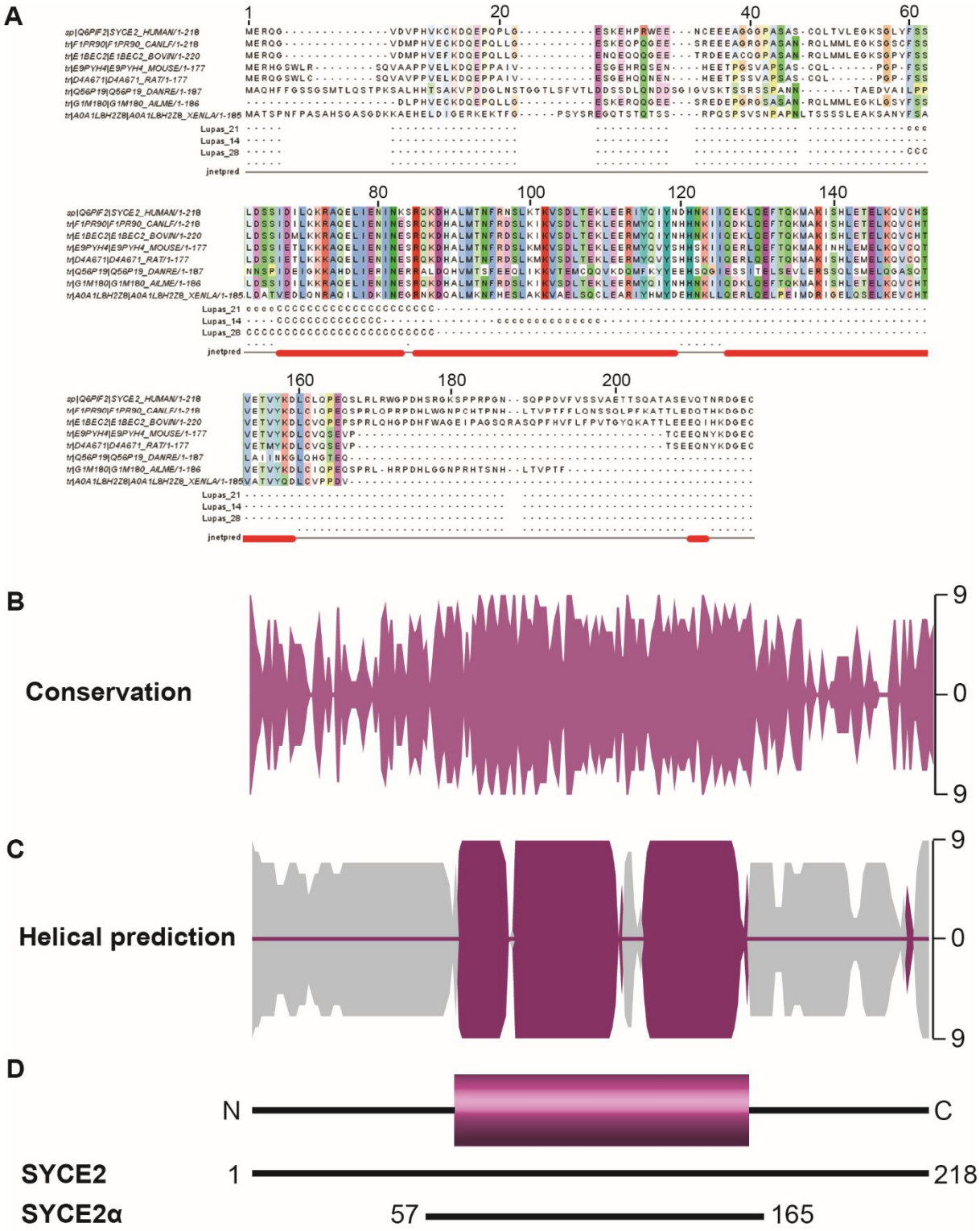


Figure 4.2. Sequence analysis of SYCE2. (A) SYCE2 multiple sequence alignment generated by MUSCLE and visualised in Jalview. JNET predictions are plotted below. Alpha helices (red rods), beta sheets (green arrow) and unstructured regions (gray line). Lupas displays coiled-coil formation as predicted by COILS. Numbered positions correspond to the human sequence. (B) SYCE2 conservation mirror plot. MUSCLE multiple sequence alignments were analysed by ConSurf to grade the conservation of each amino acid position (0-9). (C) Secondary structure prediction plot for SYCE2 from the JNET prediction. Alpha helix (dark pink) and unstructured (light gray) were plotted for each amino acid according to the confidence of the prediction (0-9). (D) Linear schematic representation of SYCE2. The pink box represents the main helical region (57-165) and the black lines represent the unstructured regions.

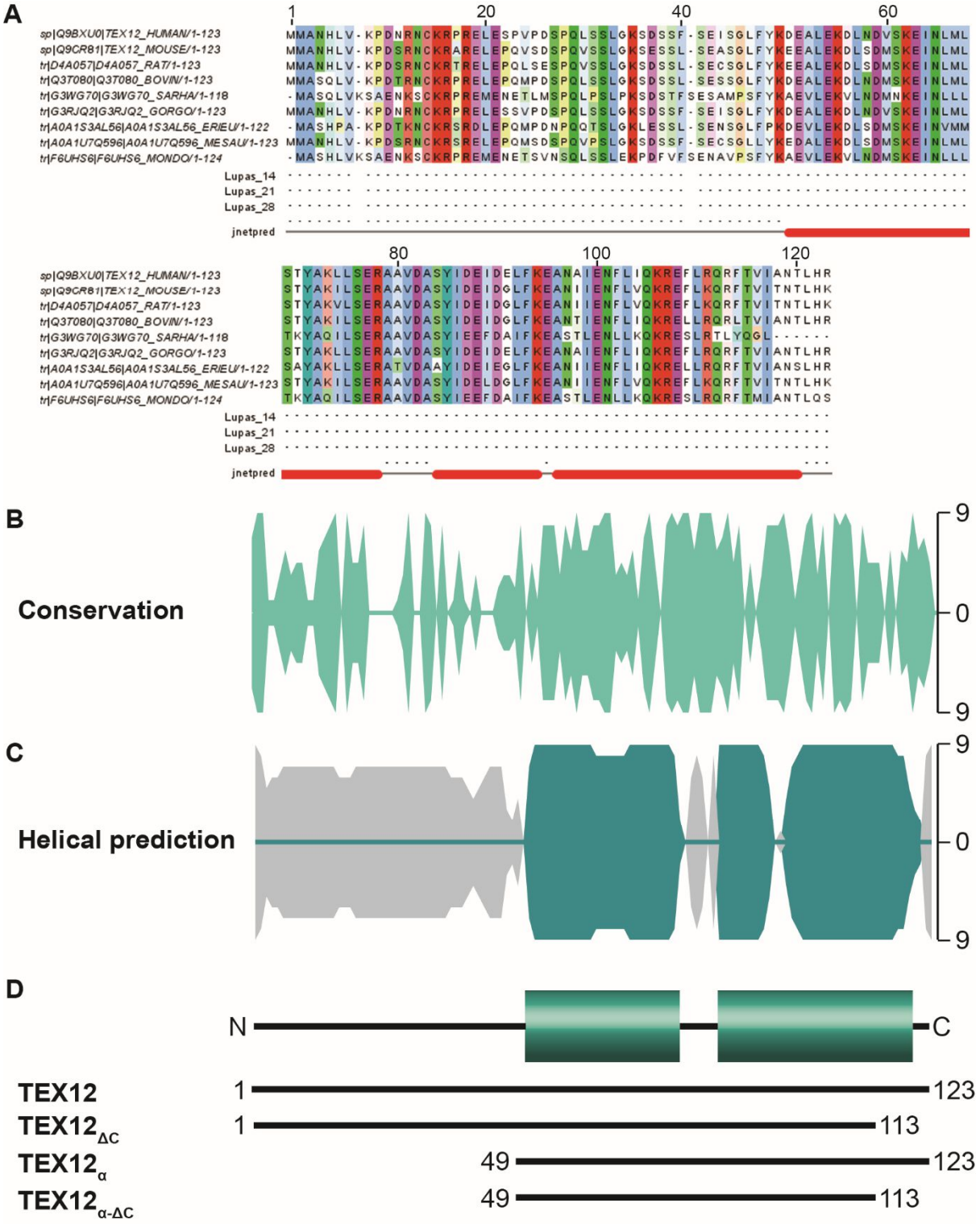


Figure 4.3. Sequence analysis of TEX12. (A) TEX12 multiple sequence alignment generated by MUSCLE and visualised in Jalview. JNET predictions are plotted below. Alpha helices (red rods), beta sheets (green arrow) and unstructured regions (gray line). Lupas displays coiled-coil formation as predicted by COILS. Numbered positions correspond to the human sequence (B) TEX12 conservation mirror plot. MUSCLE multiple sequence alignments were analysed by ConSurf to grade the conservation of each amino acid position (0-9). (C) Secondary structure prediction mirror plot for TEX12 from the JNET prediction. Alpha helix (teal) and unstructured (light gray) were plotted for each amino acid according to the confidence of the prediction (0-9). (D) Linear schematic representation of TEX12. The green boxes represent the main helical region (49-123) and the black lines represent the unstructured regions. The constructs labelled below correspond to key regions used in this study.

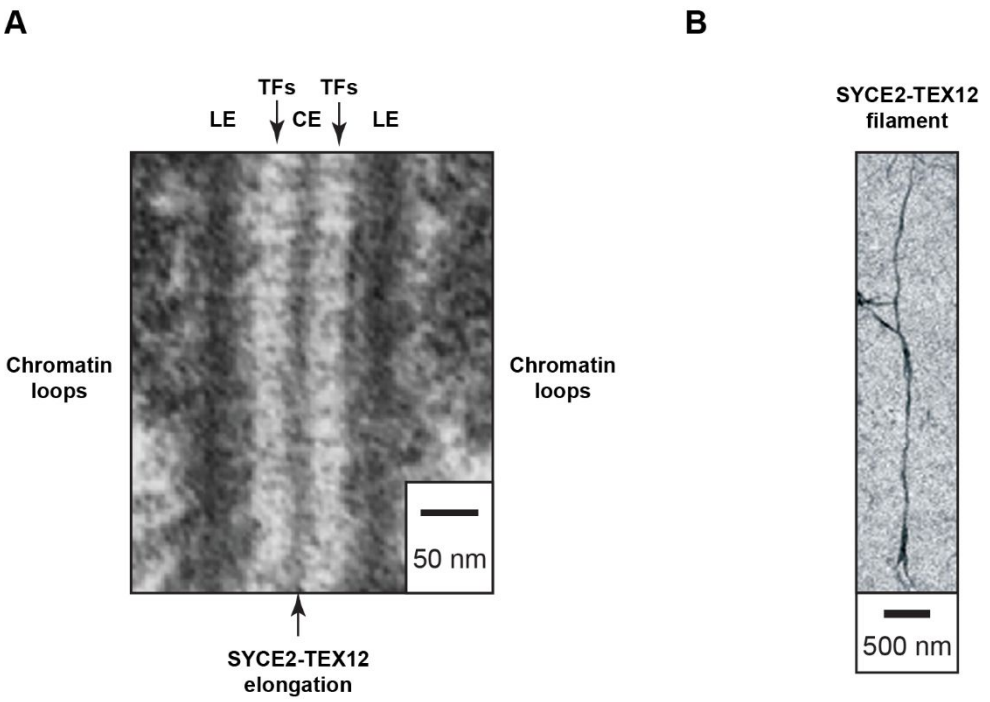


Figure 4.4. SYCE2-TEX12 filamentous assembly. (A) Transmission electron microscopy (TEM) of the mouse SC. The tripartite structure can be observed with lateral elements (LE) at the chromosomal axes, the central element (CE) in the midline and transverse filaments (TFs). TEM image taken and adapted from Kouznetsova *et al*., 2011. (B) TEM of a filament formed by the SYCE2-TEX12 core helical complex. Taken and adapted from Davies *et al.*, 2012.

The depth of the SC central region has been shown to form a bilayered organisation through separation of the SYCP1 N-termini, with recruitment of the central element proteins within the diverged SYCP1 N-termini. SYCE2 and TEX12 are required for the stabilisation of SYCP1 transverse filaments for extension (Hernandez-Hernandez *et al*., 2016). Therefore, the SYCE2-TEX12 assemblies are thought to organise and stabilise the central region for its subsequent assembly.

In this study, the molecular basis for the SYCE2-TEX12 higher-order assembly was investigated. The TEX12 C-terminal region was found to be vital for these assemblies through two modalities and the SYCE2-TEX12 filaments were found to form fibrous assemblies that are consistent with the k-m-e-f class of protein fibres (Fraser *et al*., 1964). The X-ray crystal structures presented here were solved by Owen Davies, with crystallisation of $TEX12_{\alpha\text{-}AEA}$ performed by James Dunce. A select amount of preliminary data presented in this chapter has been previously presented in my MRes dissertation and has been indicated accordingly throughout the chapter:

Lucy Salmon, MRes dissertation, 2015 "*Uncovering the molecular basis of SYCE2-TEX12 assembly within the human synaptonemal complex*".

4.2 Results

4.2.1 SYCE2-TEX12 form filamentous structures driven by the TEX12 C-terminus

Previous analysis of the SYCE2-TEX12 complex was performed with the structural core of SYCE2, owing to the unstructured, and therefore unstable, nature of the SYCE2 N- and C-termini (Davies *et al*., 2012). In order to access the function of these regions, optimisation attempts were made to purify the full length SYCE2-TEX12 complex for characterisation, despite the instability of SYCE2.

The full length SYCE2 (Figure 4.5 A) was fused to a MBP affinity tag and co-expressed with the full length TEX12 (Figure 4.5 A) fused to a His_6 tag. The fusion complex was purified through Ni-NTA and amylose affinity chromatography steps followed by anion exchange chromatography. The affinity tags were removed by incubation with TEV protease, and the cleaved SYCE2-TEX12 complex was consequently purified by a further anion exchange chromatography step (Figure 4.5 B). As suspected, the full SYCE2 protein exhibited degradation, but a large proportion of the full length SYCE2 was purified (Figure 4.5 B).

CD spectroscopy was used to study the secondary structure composition of the SYCE2-TEX12 complex. CD wavelength scans were performed between 260 and 185 nm to reveal an alpha helical profile (Figure 4.5 D), corresponding to the structure prediction and helical nature of the previously characterised core of the complex (Davies *et al*., 2012). Deconvolution, by the CDDSTR algorithm (Sreerama and Woody, 2000), estimates the helical content to be 64% and the rest of the content is largely unstructured due to the SYCE2 flexible N- and C-termini. To assess the thermal stability of SYCE2-TEX12, CD thermal denaturation was carried out to monitor the 222 nm helical signal across a temperature gradient from 5°C to 95°C. The melting curve shows an initial steady progression in protein unfolding, likely corresponding to the gradual unwinding of alpha helices (Figure 4.5 E). The gradual unfolding is followed by sharper cooperative unfolding event, with an estimated melting temperature of 70°C (Figure 4.5 E).

To investigate the oligomeric properties of the full length SYCE2-TEX12 complex, SEC-MALS was used to separate protein species according to their shape and size for subsequent calculation of the absolute molecular mass. The SEC profile demonstrates the majority of the material was eluted immediately after the void of the column, indicating the formation of large species or aggregation, with a small proportion of the material existing in smaller eluted peaks (Figure 4.5 C). The MALS analysis confirms that the majority of the material forms a range of large molecular weight species (Figure 4.5 C). The smaller peaks have molecular weights of 130 kDa and 78 kDa, perhaps suggesting the formation of distinct species. However, it is probable that these species correspond to the degradation products of SYCE2 (Figure 4.5 C). The formation of the soluble higher-order species corresponds to the

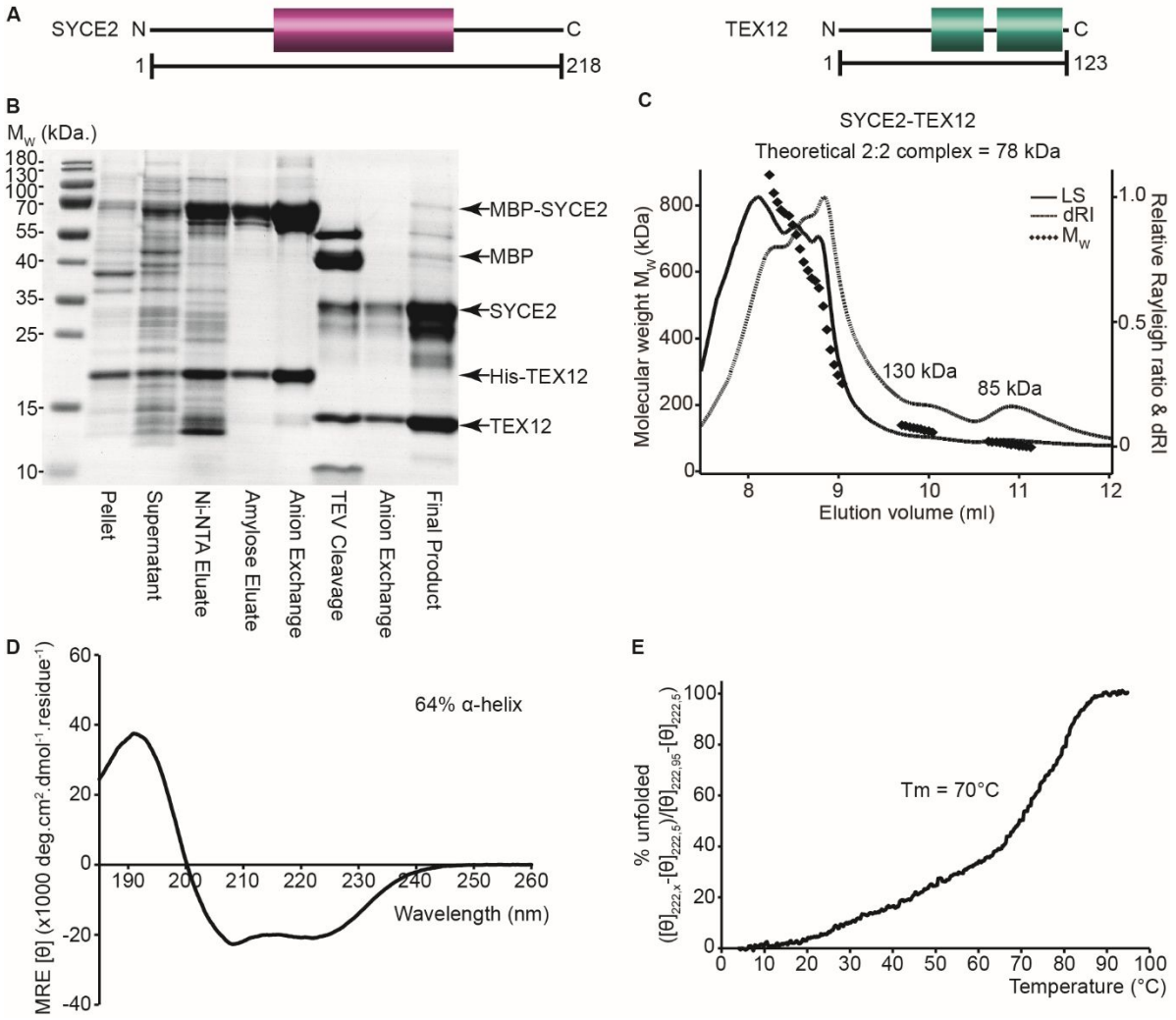


Figure 4.5. Purification and analysis of SYCE2-TEX12. (A) Schematic of SYCE2 and TEX12**.** (B) SDS-PAGE summary of SYCE2-TEX12 purification by Ni-NTA, amylose and anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange chromatography. (C) SEC-MALS analysis. The majority of the material forms large molecular weight assemblies with smaller peaks of 130 kDa and 85 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 78 kDa. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 64% alpha-helix, 4% β-sheet, 10% turns and 22% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 70°C, when the sample is 50% unfolded.

suggested function of the SYCE2-TEX12 complex to extend the SC central element through self-assembly.

To further investigate the assembly of SYCE2-TEX12, transmission electron microscopy (TEM) experiments were performed. These experiments demonstrate the capacity for the full length SYCE2-TEX12 complex to undergo filament formation, with the detection of large single filaments and branched filamentous networks (Figure 4.6 A). The analysis of the SYCE2-TEX12 larger filaments observed show remarkable similarity to the dense central

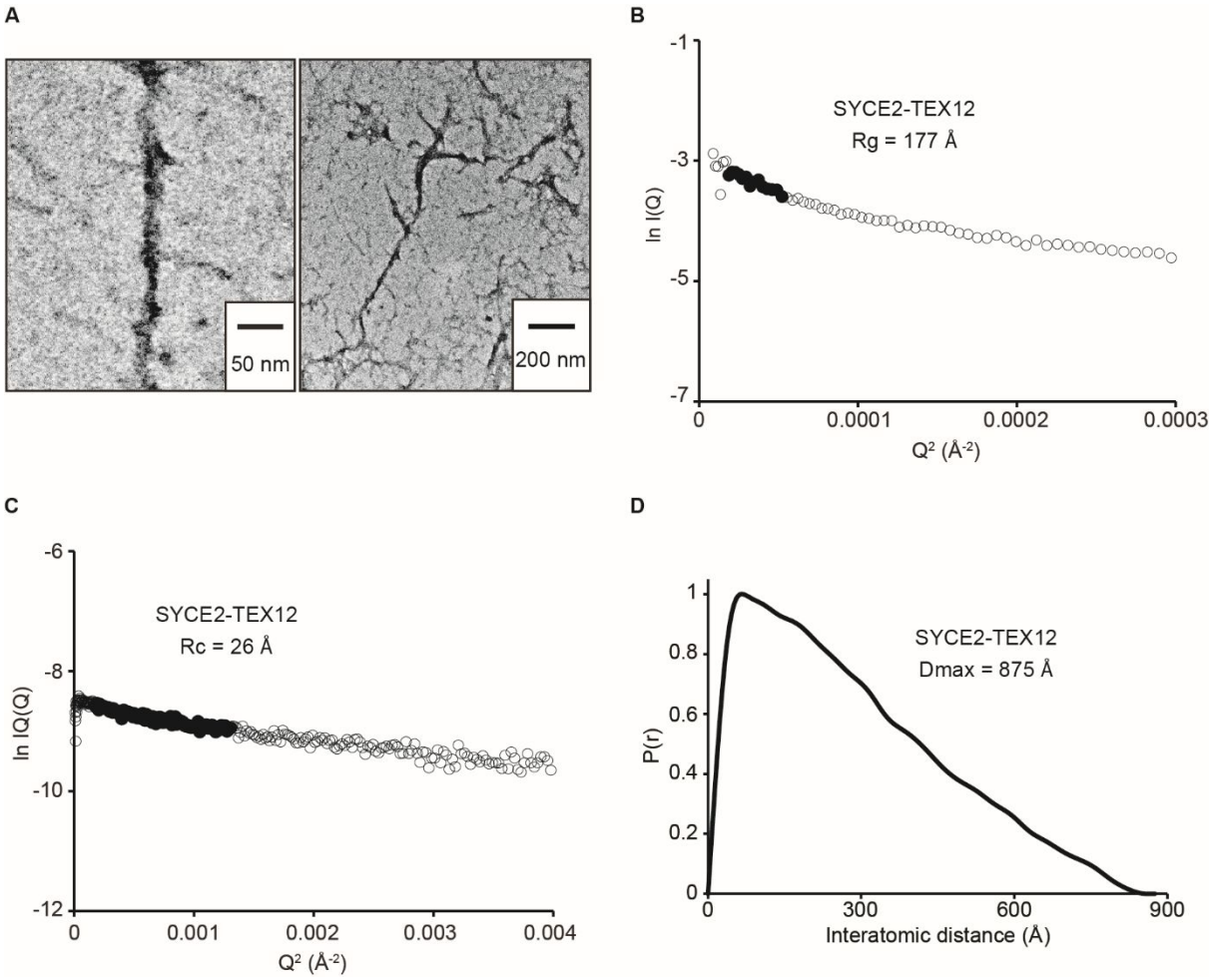


Figure 4.6. Analysis of SYCE2-TEX12 assemblies. (A) TEM analysis of SYCE2-TEX12 filaments. (B-D) SEC-SAXS analysis of SYCE2-TEX12 assembly. (B) Guinier analysis to determine the radius of gyration (Rg) as 177 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (C) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 26 Å. Clear circles are the data, solid circles represent the region used for the fit. (D) Paired real-space distribution plot. The SYCE2-TEX12 assembly has a maximum dimension of 875 Å.

element structures identified from EM experiments of intact SC structures (Kouznetsova *et al.*, 2011). This suggests that these assemblies may be biologically relevant to SC maturation.

Although no distinct species were detected by SEC-MALS analysis, size exclusion chromatography small-angle X-ray scattering (SEC-SAXS) experiments were performed to attempt to determine the structural properties of SYCE2-TEX12 in solution. From the broad high molecular weight species formed, a fine portion of the data were analysed that corresponds to an intermediate higher-order assembly species, as identified at approximately 9 ml of the SEC-MALS elution profile (Figure 4.5 B). Through averaging of this region and subsequent Guinier analysis, the radius of gyration (Rg) was determined to be 177 Å, potentially suggesting an elongated structure is formed (Figure 4.6 B). The radius of gyration of the cross-section (Rc) was determined to be 26 Å, further indicating an elongated structure rather than a globular one

(Figure 4.6 C). Real space analysis was performed by the paired-distance distribution function, P(r), to plot the inter-atomic distances observed within the sample. The P(r) curve demonstrates the extended nature of the complex, with a shape corresponding to a characteristic P(r) profile of an elongated rod structure (Figure 4.6 D). The maximum dimension of the species was determined to be 875 Å, validating that a long structure is formed (Figure 4.6 D). The species therefore has a maximum length of 875 Å and a cross-sectional radius of 26 Å and thus, suggests that the assemblies formed in solution are filamentous in nature.

Self-assembly formation has been detected for other SC components, particularly the lateral element component SYCP3 (Baier *et al*., 2007). The full length recombinant SYCP3 protein was found to form large assemblies that assemble on DNA and are abolished by the removal of six C-terminal amino acids (Syrjanen *et al*., 2014). Therefore, to test whether these SYCE2-TEX12 filamentous assemblies could be blocked in a similar manner, a truncated complex was characterised. As the core helical regions of SYCE2 and TEX12 ($SYCE2_{\alpha}$-$TEX12_{\alpha}$) were still found to form filamentous structures (Davies *et al*., 2012), it was reasoned that although they may encourage higher-order assembly, the TEX12 N-terminus and SYCE2 N- and C-termini are not vital for the assemblies. Thus, the TEX12 C-terminus was truncated by 10 amino acids (Figure 4.7 A) to investigate the effect of this region on filament formation.

The SYCE2-$TEX12_{\Delta C}$ complex was purified with an MBP tag on SYCE2 and a His_6 tag on the truncated TEX12. The fusion complex was purified by Ni-NTA and amylose affinity chromatography steps followed by anion exchange chromatography. The tags were removed by TEV protease and the cleaved complex was purified by anion exchange and size exclusion chromatography (Figure 4.7 B). The final product was found to have a prominent degradation protein that was co-eluted with the SYCE2-$TEX12_{\Delta C}$ complex through every purification stage, with the size of this product roughly corresponding to the size of the SYCE2 helical core ($SYCE2_{\alpha}$) suggesting the degradation is due to the instability of the unstructured SYCE2 N- and C-termini (Figure 4.7 B).

To validate the folding of SYCE2-TEX12 upon truncation of TEX12, CD experiments were performed. A far-UV CD scan was recorded between 260 and 185 nm to reveal an alpha helical trace (Figure 4.7 D), similar to the full length SYCE2-TEX12. Through deconvolution of the data, the alpha helical content was estimated as 62%, which is comparable to the 64% helical content as estimated for the full length complex (see above). The large proportion of unstructured residues again corresponds to the TEX12 N-terminus and SYCE2 N- and C-termini, and also suggests that the TEX12 truncation has not disrupted the folding of SYCE2-TEX12. Thermal denaturation was performed between 5°C and 95°C to assess the thermal

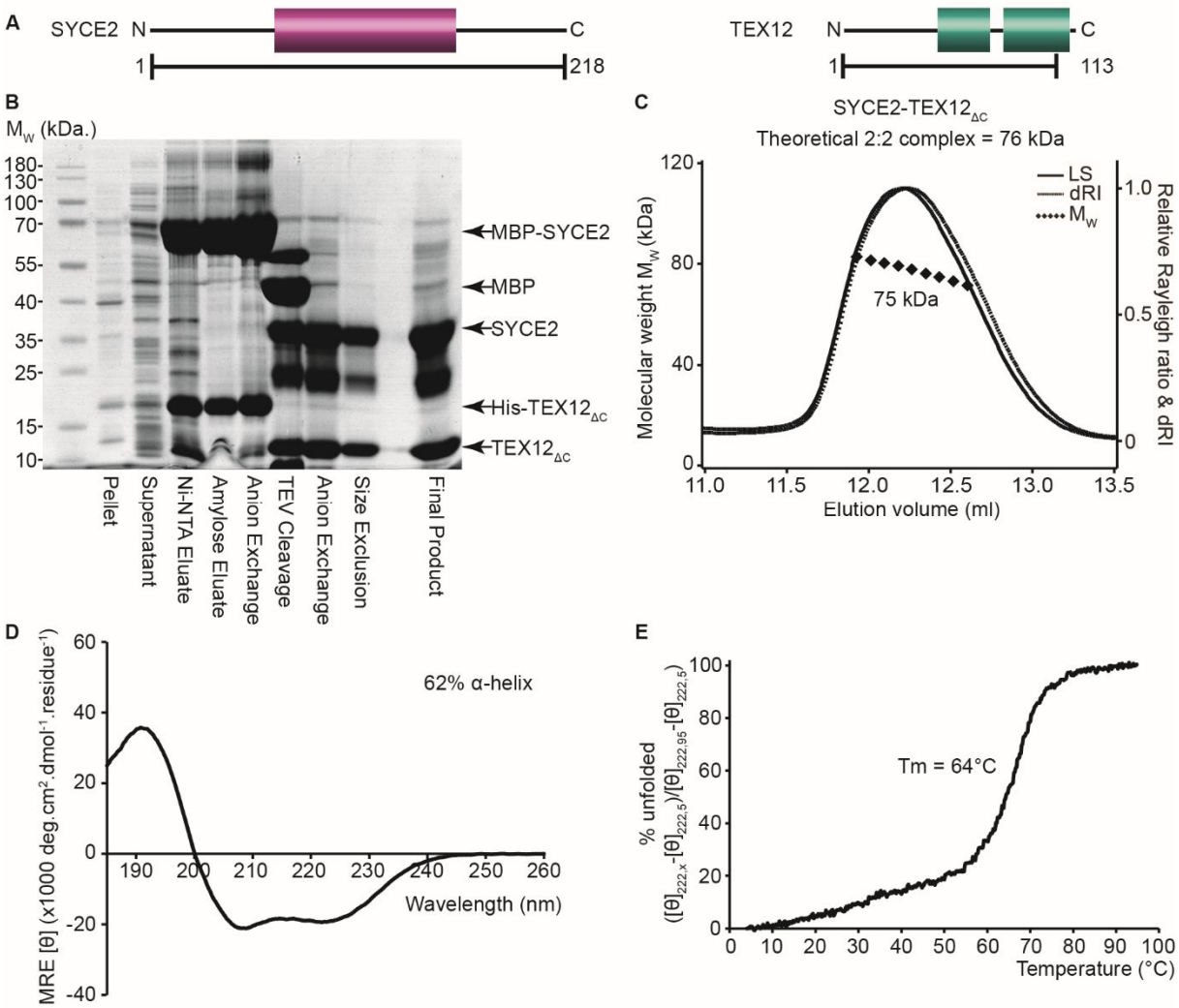


Figure 4.7. Purification and analysis of SYCE2-TEX12$_{\Delta C}$. (A) Schematic of SYCE2 and TEX12 boundaries. (B) SDS-PAGE summary of SYCE2-TEX12$_{\Delta C}$ purification by Ni-NTA, amylose and anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange and size exclusion chromatography. (D) SEC-MALS analysis. A single peak is formed with a calculated molecular mass of 75 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 76 kDa. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 62% alpha-helix, 5% β-sheet, 11% turns and 22% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 64°C, when the sample is 50% unfolded.

stability of SYCE2-TEX12$_{\Delta C}$ by CD. The resultant melting curve demonstrates an initial shallow unfolding, corresponding to gradual unwinding of helices, followed by a sharp cooperative unfolding event (Figure 4.7 E). The melting temperature was determined to be 64°C, comparable to the 70°C estimation for the full length complex, and therefore, further validates that the TEX12 truncation has not disrupted the folding of the SYCE2-TEX12 complex (Figure 4.7 E).

The consequence of the TEX12 truncation on ability to assemble was assessed by SEC-MALS to calculate the molecular mass of the separated protein species. The SEC-MALS profile for SYCE2-TEX12$_{\Delta C}$ showed the formation of a single species at a significantly later elution point than the SYCE2-TEX12 assemblies, with no larger species observed. The molecular mass of the single species was determined as 75 kDa, which corresponds to the formation of a 2:2 complex (Figure 4.7 C). This analysis demonstrates that the truncation of the TEX12 C-terminus has indeed blocked the formation of higher order assemblies in solution, and forms a different oligomeric unit than the previously characterised helical core complex, SYCE2$_{\alpha}$-TEX12$_{\alpha}$, which forms a 4:4 complex with filamentous propensity (Davies *et al*., 2012). To validate that the difference in oligomerisation was not simply caused by protein concentration, the SEC-MALS analysis for SYCE2-TEX12 was performed at 4 mg/ml and the analysis for SYCE2-TEX12$_{\Delta C}$ was performed at 40 mg/ml, showing that a 10-fold increase in protein concentration did not induce concentration-dependant assembly and therefore, that the TEX12 C-terminus is specifically required for higher-order assembly, but not for the interaction between SYCE2 and TEX12.

As the SYCE2-TEX12 higher-order assembly was found to be abolished by the removal of the TEX12 C-terminus in solution, TEM experiments were performed to assess whether any filamentous structures of SYCE2-TEX12$_{\Delta C}$ could be detected. The TEM analysis was performed with the SYCE2-TEX12$_{\Delta C}$ complex at 4 mg/ml for comparison to the full length complex and no filamentous structures were found (Figure 4.8 A).

As the higher-order assembly was blocked both in solution and upon fixation, the SYCE2-TEX12$_{\Delta C}$ 2:2 species was analysed by SEC-SAXS to compare the structural parameters to the assembled SYCE2-TEX12 complex. Guinier analysis measured the Rg of the 2:2 complex to be 69 Å (Figure 4.8 B), which is a smaller value than the 177 Å Rg determined for the assembled complex. The Rc was measured to be 27 Å (Figure 4.8 C), which is very close to the 26 Å Rc value measured for the assembled complex. Real space analysis was carried out to show the P(r) curve of the 2:2 complex. The distance distribution indicates the formation of an extended structure, due to the shape of the curve, with a maximum dimension observed in the sample determined to be 260 Å (Figure 4.8 D). Comparison of the structural parameters, as determined by SEC-SAXS, for the SYCE2-TEX12 assembly and the SYCE2-TEX12$_{\Delta C}$ 2:2 complex reveal that although the assembled complex has a much larger maximum dimension and Rg, the two species have almost the same Rc value, suggesting that the increase in complex length upon assembly occurs in a longitudinal manner. This suggests that the SYCE2-TEX12$_{\Delta C}$ 2:2 complex may resemble the core building block of these assemblies and that the TEX12 C-terminus is responsible for self-assembly, and therefore filament formation, of these core units.

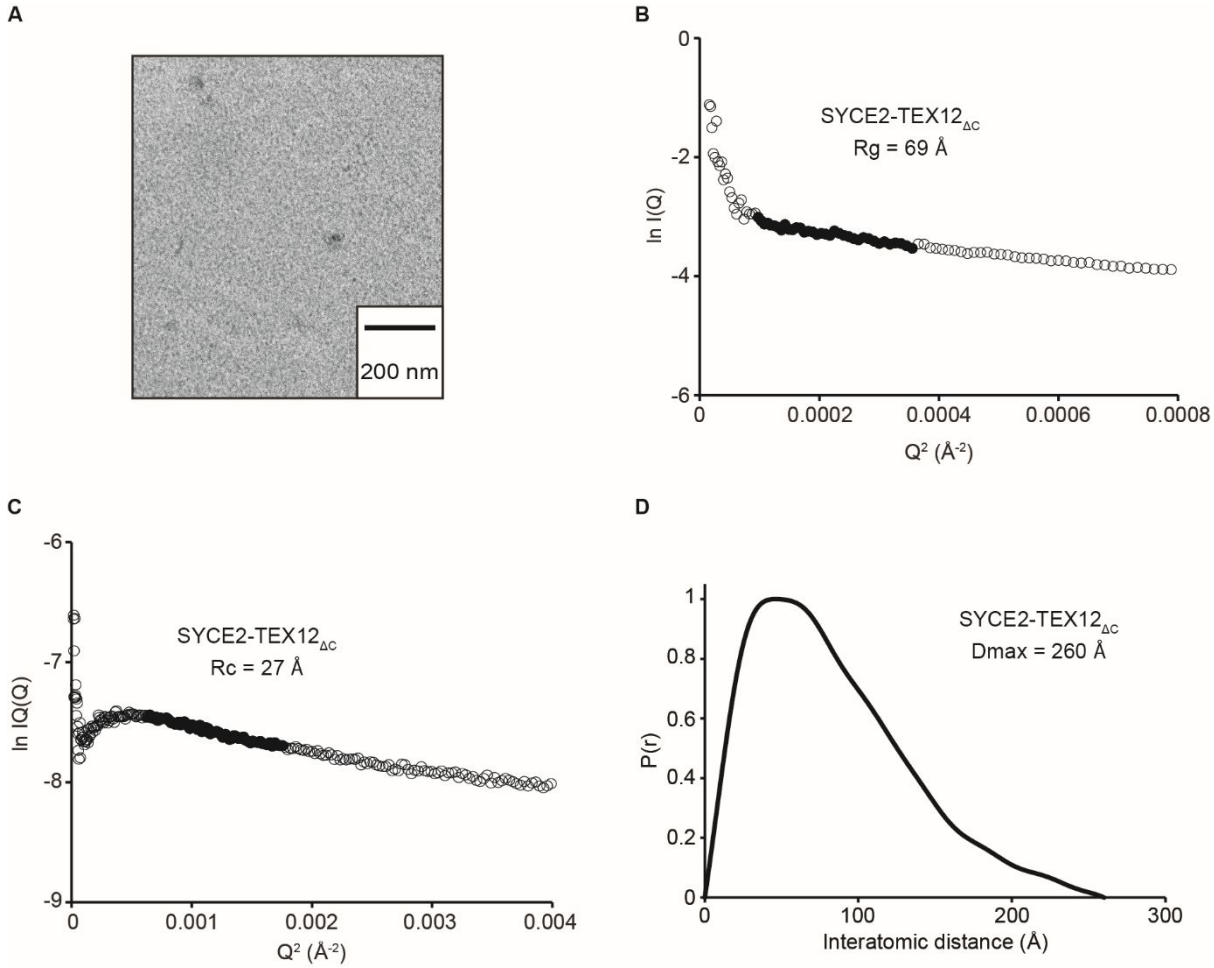


Figure 4.8. Analysis of SYCE2-TEX12$_{\Delta C}$. (A) TEM analysis of SYCE2-TEX12$_{\Delta C}$, with no filaments observed. (B-D) SEC-SAXS analysis of the SYCE2-TEX12$_{\Delta C}$ 2:2 complex. (B) Guinier analysis to determine the radius of gyration (Rg) as 69 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (C) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 27 Å. Clear circles are the data, solid circles represent the region used for the fit. (D) Paired real-space distribution plot. The SYCE2-TEX12$_{\Delta C}$ 2:2 complex has a maximum dimension of 260 Å.

4.2.2 Dissecting the role of the TEX12 C-terminus in SYCE2-TEX12 filament formation

As the higher-order assembly of SYCE2-TEX12 has been attributed to the C-terminus of TEX12, the TEX12 C-terminal sequence was further analysed to select candidate residues that may be implicated in the assemblies. From aligning several mammalian TEX12 sequences, two groups of candidate residues were selected for further investigation (Figure 4.9 A).

The potential for coiled-coil formation at the C-terminus was investigated by plotting potential coiled-coil positions (*a*-*g*), with hydrophobic residues falling in the *a* and *d* positions (Figure 4.9 A). Two coiled coil heptad patterns were satisfied at the TEX12 C-terminus, and therefore, four candidate hydrophobic residues were selected: L110, F114, I117 and L121. The

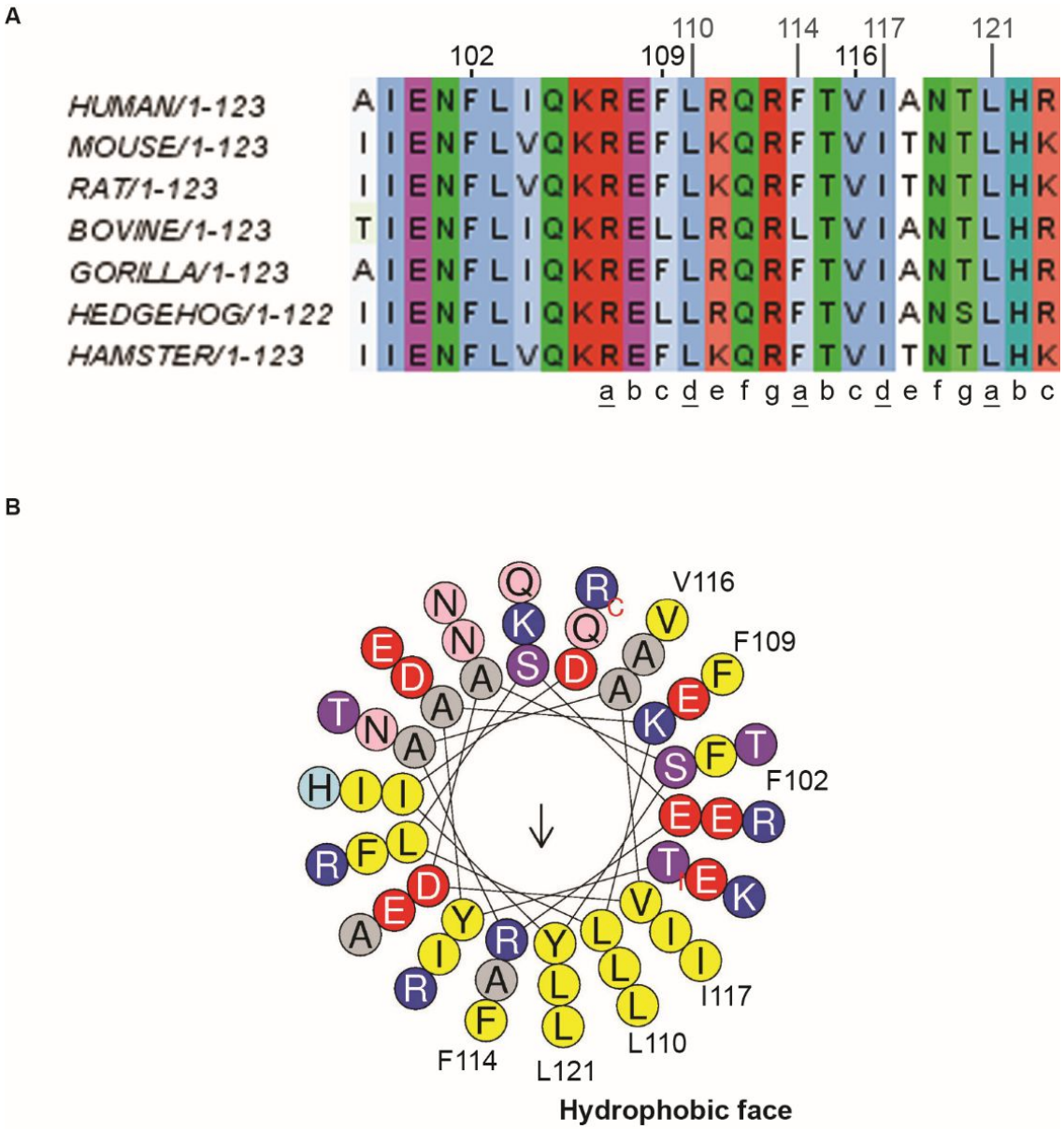


Figure 4.9. Sequence analysis of the TEX12 C-terminus. (A) The mammalian multiple sequence alignment of the TEX12 C-terminus was produced by MUSCLE and visualised in Jalview. The a-g labels beneath relate to coiled-coil heptad positions, with underlined positions corresponding to hydrophobic residues. The F102, F109 and V116 residues (black) and L110, F114, I117 and L121 residues (gray) are indicated. (B) HeliQuest plot of the TEX12 C-terminal helical wheel. The L110, F114, I117 and L121 residues form a hydrophobic face on one side of the plotted helix, with F102, F109 and V116 residues forming a patch away from the hydrophobic face.

TEX12 sequence was also plotted as a helical wheel to map any patterns that may arise upon helix formation. This analysis highlighted that the four candidate hydrophobic coiled-coil residues fall on the same side of the plotted helical map, implying the formation of a hydrophobic interface by these residues and upstream residues (Figure 4.9 B).

As the SYCE2-TEX12 filaments are thought to be formed by iterative self-assembly of SYCE2-TEX12 structural units, it was postulated whether assembly may be supported by

surface exposed residues from the SYCE2-TEX12 unit. Therefore, three potential surface exposed residues were proposed: F102, F109 and V116 (Figure 4.9 A). The V116 residue is the only one that falls within the TEX12 region that was removed to block assembly. The two upstream phenylalanine residues form a patch with V116, away from the hydrophobic face, that may form a surface exposed stretch for higher-order assembly (Figure 4.9 B).

To examine the role of the candidate surface exposed residues (F102, F109 and V116) in SYCE2-TEX12 assembly, the three residues were simultaneously mutated to alanines (Figure 4.10 A), with the mutated complex hereby referred to as SYCE2-TEX12$_{3A}$. The SYCE2-TEX12$_{3A}$ complex was purified in the same manner as the wild type (WT) complex, with the cleaved complex purified by anion exchange chromatography following affinity tag removal (Figure 4.10 B).

SYCE2-TEX12$_{3A}$ was analysed by SEC-MALS to investigate the oligomeric state of the complex upon introducing the triple mutation, with the experiment performed in the same manner, and at the same protein concentration, as the WT complex. The SEC elution showed a similar profile to WT, with the MALS analysis indicating the formation of large molecular weight species with some smaller 161 kDa and 81 kDa species, which may correspond to 4:4 and 2:2 complexes respectively (Figure 4.10 C). This suggests that the candidate residues do not have a critical role in SYCE2-TEX12 assembly.

To further probe the capacity for filament formation, TEM experiments were performed with SYCE2-TEX12$_{3A}$. The TEM analysis showed that, as indicated by the SEC-MALS analysis, filament formation is retained (Figure 4.10 D). However, the filaments appear to be limited in length compared to WT, with no extensive networks identified (Figure 4.10 E). As SYCE2-TEX12$_{3A}$ maintains capacity for higher-order assembly, SEC-SAXS experiments were performed to analyse these filaments. From the scattering data, a small portion of the assembly peak, was selected for averaging and analysis. The Rg was measured as 110 Å (Figure 4.10 E) and the Rc was measured as 27 Å (Figure 4.10 F). These values are similar to those determined for the WT complex, and indicates an extended structure. The real space analysis reveals a distance distribution indicative of an elongated species, with the maximum dimension observed in the sample estimated to be 550 Å (Figure 4.10 G).

The SYCE2-TEX12$_{3A}$ complex maintained the capacity for filament formation, therefore suggesting the F102, F109 and V116 TEX12 residues are not vital for higher-order assembly. To study the role of the four candidate hydrophobic residues, and consequently a potential coiled-coil structure at the TEX12 C-terminus, the L110, F114, I117 and L121 were simultaneously mutated to glutamic acid in order to remove the hydrophobic face at the C-terminus (Figure 4.11 A), with the mutated complex hereby referred to as SYCE2-TEX12$_{cc\text{-}mut}$.

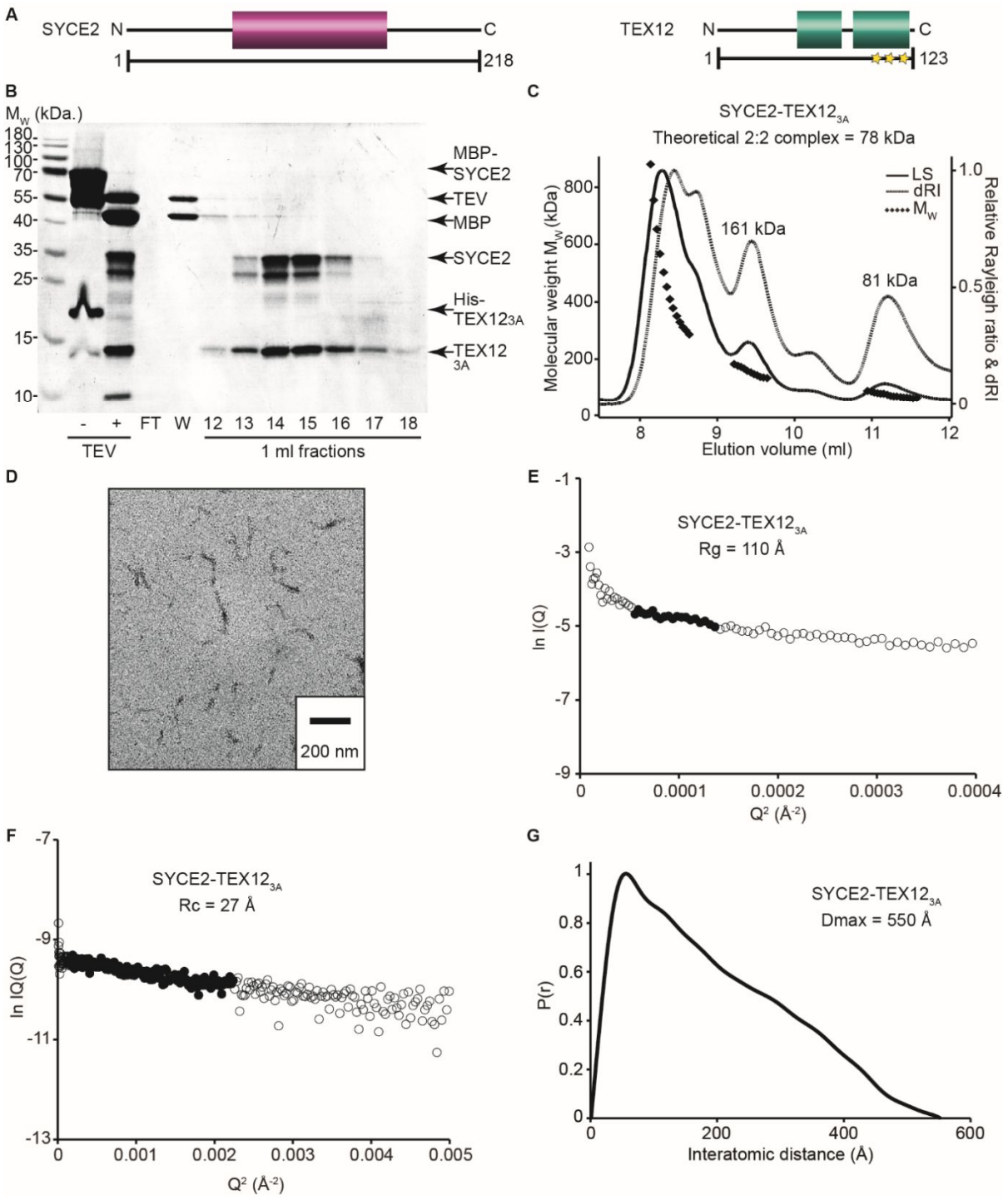


Figure 4.10. Purification and analysis of SYCE2-TEX12$_{3A}$. (A) Schematic of SYCE2 and TEX12 boundaries. Stars represent point mutations. (B) SDS-PAGE of SYCE2-TEX12$_{3A}$ purification by anion exchange chromatography, following tag removal by TEV protease incubation. (C) SEC-MALS analysis. The majority of the material forms large molecular weight assemblies with smaller peaks of 161 kDa and 81 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 78 kDa. (D) TEM analysis of SYCE2-TEX12$_{3A}$ filaments. (E-G) SEC-SAXS analysis of SYCE2-TEX12$_{3A}$ assembly. (E) Guinier analysis to determine the radius of gyration (Rg) as 110 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (F) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 27 Å. Clear circles are the data, solid circles represent the region used for the fit. (G) Paired real-space distribution plot. The SYCE2-TEX12$_{3A}$ assembly has a maximum dimension of 550 Å.

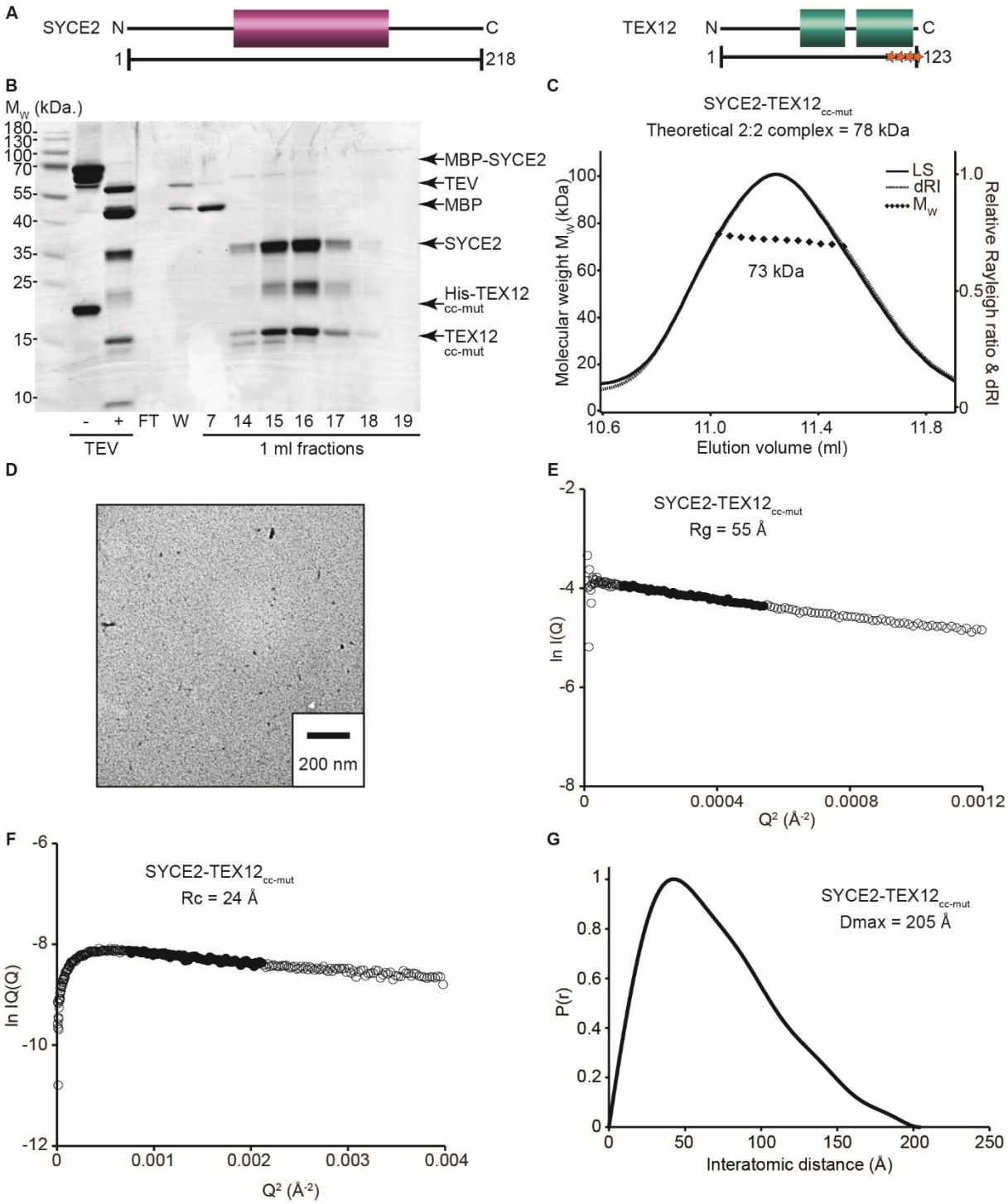


Figure 4.11. Purification and analysis of SYCE2-TEX12$_{cc\text{-}mut}$. (A) Schematic of SYCE2 and TEX12 boundaries. Stars represent point mutations. (B) SDS-PAGE of SYCE2-TEX12$_{cc\text{-}mut}$ purification by anion exchange chromatography, following tag removal by TEV protease incubation. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 73 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 78 kDa. (D) TEM analysis of SYCE2-TEX12$_{cc\text{-}mut}$ revealing no filament formation. (E-G) SEC-SAXS analysis of the SYCE2-TEX12$_{cc\text{-}mut}$ 2:2complex. (E) Guinier analysis to determine the radius of gyration (Rg) as 55 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (F) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 24 Å. Clear circles are the data, solid circles represent the region used for the fit. (G) Paired real-space distribution plot. The SYCE2-TEX12$_{cc\text{-}mut}$ 2:2 complex has a maximum dimension of 205 Å.

The SYCE2-TEX12$_{cc\text{-}mut}$ complex was purified by the same procedure as the WT complex, with the cleaved complex purified by anion exchange chromatography subsequent to TEV protease incubation for affinity tag removal (Figure 4.11 B). SYCE2-TEX12$_{cc\text{-}mut}$ was analysed by SEC-MALS to compare the oligomeric properties to the WT complex. The SEC profile reveals a single protein peak, with the MALS analysis determining the molecular weight of the species to be 73 kDa, corresponding to a 2:2 complex (Figure 4.11 C). This suggests that a coiled-coil structure at the C-terminus of TEX12 is indeed critical for the higher-order assembly of SYCE2-TEX12. The SEC-MALS analysis was performed at 8 mg/ml, which is a 2-fold higher protein concentration then the WT complex was analysed, further indicating that the higher-order assemblies have been abolished.

To further validate the inability of filament formation, TEM experiments were performed with SYCE2-TEX12$_{cc\text{-}mut}$ at both 3 mg/ml and 7 mg/ml. These experiments show that at both concentrations, no filaments were observed and thus, confirms that the SYCE2-TEX12$_{cc\text{-}mut}$ complex can no longer assemble (Figure 4.11 D).

The non-assembling 2:2 SYCE2-TEX12$_{cc\text{-}mut}$ complex was studied by SEC-SAXS to investigate the structural parameters in solution. The Rg was determined to be 55 Å (Figure 4.11 E) and the Rc was measured to be 24 Å (Figure 4.11 F). The real space analysis estimates that the maximum dimension in the 2:2 complex is 205 Å, with the distance distribution curve indicating an elongated shape (Figure 4.11 G). These values suggest that the SYCE2-TEX12$_{cc\text{-}mut}$ behaves in the same way as the SYCE2-TEX12$_{\Delta c}$ complex, with disruption of the TEX12 C-terminal coiled-coil preventing higher-order assemblies.

Construct	SYCE2-TEX12	SYCE2-TEX12$_{\Delta C}$	SYCE2-TEX12$_{3A}$	SYCE2-TEX12$_{cc\text{-}mut}$
SYCE2 region	1-218	1-218	1-218	1-218
TEX12 region	1-123	1-113	1-123	1-123
Mutations	N/A	N/A	F102A F109A V116A	L110E F114E I117E L121E
% α-helix	64	62	N.D.	N.D.
Tm (°C)	70	64	N.D.	N.D.
Oligomer state	Large species	2:2 complex	Large species	2:2 complex
Assembly	Yes	No	Yes	No
Rg (Å)	177	69	110	55
Rc (Å)	26	27	27	24
Dmax (Å)	875	260	550	205
EM filaments	Yes	No	Yes	No

Table 4.1. Biophysical analysis summary of SYCE2-TEX12 constructs

4.2.3 The structural SYCE2-TEX12 core forms discrete units with the capacity for assembly

Sequence analysis of the SYCE2 and TEX12 proteins indicates that both proteins contain an alpha helical core region, with the N- and C-termini of SYCE2 and the N-termini of TEX12 predicted to be unstructured (Figure 4.12 A). Core helical regions (SYCE2 residues 57-165 and TEX12 residues 49-123) were found to form a 4:4 complex, with the propensity for filament formation (Davies *et al*., 2012).

Several constructs were tested in the MRes project to explore favourable amino acid boundaries that may optimise the stability of these core regions (Table 4.2). This analysis suggested that truncation of the N-termini of both SYCE2 and TEX12 destabilises complex formation and that truncation at the C-terminus of both proteins results in the formation of a 2:2 complex (Table 4.2). Truncation at the C-terminus of SYCE2, to amino acid residue 120, formed a 2:2 complex with full length TEX12. Thermal denaturation of this complex showed non-cooperative unfolding suggesting that a structured unit is not formed and that the SYCE2 truncation has perhaps affected SYCE2 folding (Table 4.2). Therefore, these truncations did not lead to identification of more stable SYCE2 amino acid boundaries than residues 57-165 and the C-terminal truncation of TEX12, to amino acid residue 1-113, formed a stable 2:2 complex. The TEX12 amino acid residues 49-123 comprise the helical core of the protein with the C-terminal residues (114-123) appearing to have a role in higher order assembly, as seen with the full length proteins (Table 4.1).

Construct	$SYCE2_{64-165}$-TEX12	$SYCE2_{57-120}$-TEX12	$SYCE2_{\alpha}$-$TEX12_{63-123}$	$SYCE2_{\alpha}$-$TEX12_{76-123}$	$SYCE2_{\alpha}$-$TEX12_{\Delta C}$	$SYCE2_{\alpha}$-$TEX12_{49-119}$
SYCE2 region	64-165	57-120	57-165	57-165	57-165	57-165
TEX12 region	1-123	1-123	63-123	76-123	1-113	49-119
Stable complex	N.D.	Yes	N.D.	N.D.	Yes	Yes
% α-helix	N.D.	55	N.D.	N.D.	67	N.D.
Tm (°C)	N.D.	Non-cooperative	N.D.	N.D.	70	N.D.
Oligomer state	N.D.	2:2 complex	N.D.	N.D.	2:2 complex	4:4 complex
Assembly	N.D.	No	N.D.	N.D.	No	Yes
EM filaments	N.D.	No	N.D.	N.D.	No	N.D.

Table 4.2. Biophysical analysis summary of SYCE2-TEX12 constructs tested in MRes project (Lucy Salmon, MRes dissertation, 2015)

The helical regions (SYCE2 residues 57-165 and TEX12 residues 49-123) were co-expressed for *in vitro* studies and the complex, hereby referred to as $SYCE2_{\alpha}$-$TEX12_{\alpha}$, was purified in this study for further characterisation. $SYCE2_{\alpha}$ and $TEX12_{\alpha}$ were co-expressed as fusion proteins to MBP and His_6 affinity tags, respectively. The fusion complex was purified by Ni-NTA, amylose and anion exchange chromatography, the tags were removed by incubation with TEV protease and the cleaved complex was subsequently purified by anion exchange chromatography (Figure 4.12 B).

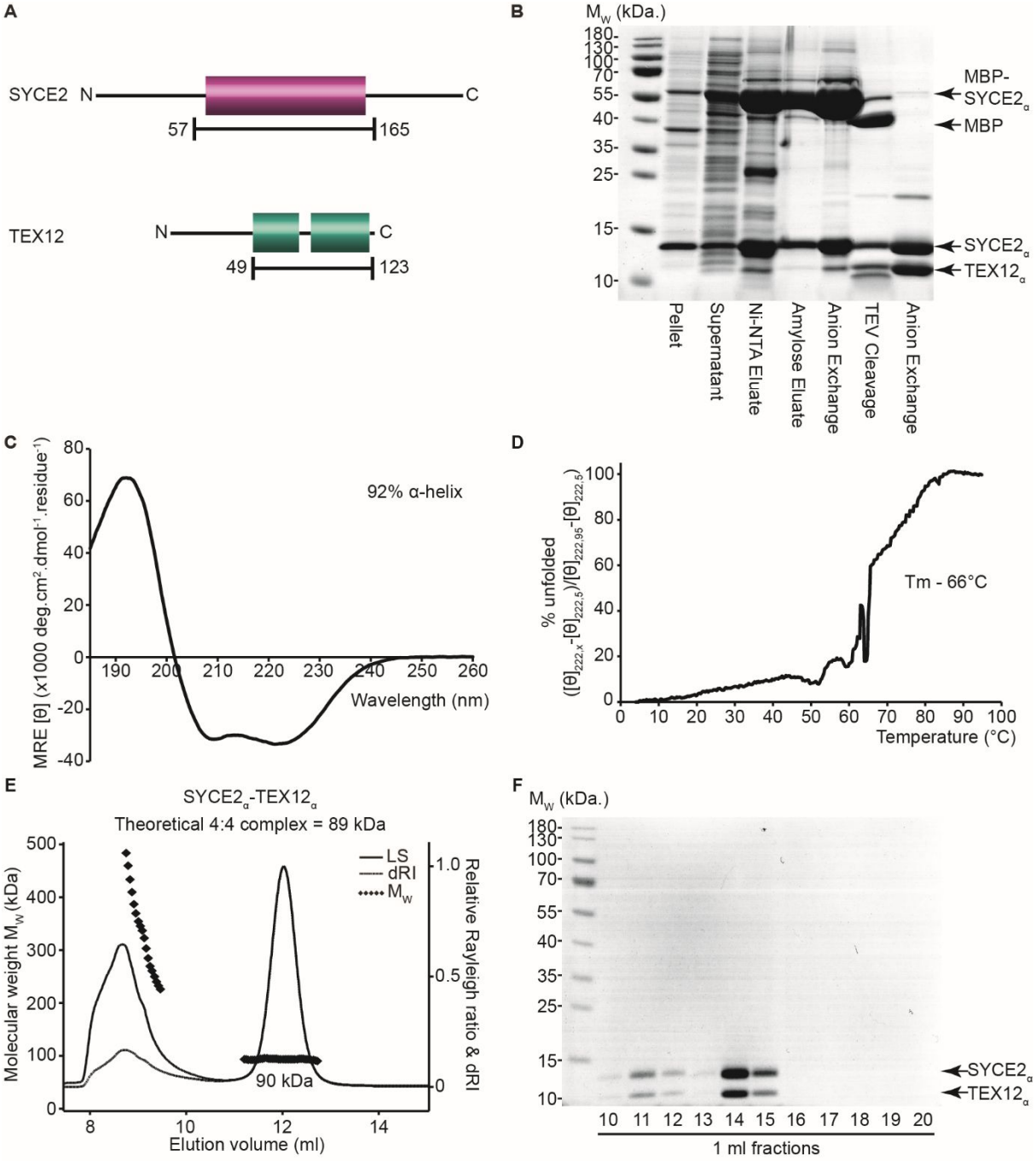


Figure 4.12. Purification and analysis of SYCE2$_\alpha$-TEX12$_\alpha$. This data was presented in my MRes dissertation. (A) Schematic of SYCE2 and TEX12 boundaries. (B) SDS-PAGE summary of SYCE2$_\alpha$-TEX12$_\alpha$ purification by Ni-NTA, amylose and anion exchange chromatography steps, followed by TEV protease incubation and anion exchange chromatography. (C) CD wavelength scan between 260-185 nm plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 92% alpha-helix, 3% β-sheet, 1% turns and 4% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 66°C, when the sample is 50% unfolded. (E) SEC-MALS analysis. A protein peak is formed with a calculated molecular mass of 90 kDa, with some material forming large molecular weight species. The theoretical molecular weight of a 4:4 complex is estimated to be 89 kDa. (F) SDS-PAGE analysis of SEC-MALS fractions.

The SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex was analysed by CD to validate the folding of the complex. A far-UV CD scan was recorded between 260 and 185 nm, to reveal an alpha helical profile with an estimated helical content of 92% (Figure 4.12 C). This is a slightly higher helical content than previously published (Davies *et al*., 2012), with the difference potentially due to expression protocol, sample preparation or the analysis buffer. However, in this experiment, the 92% helical content validates that the structural core has been isolated and that the proteins are properly folded. CD thermal denaturation was performed by recording the 222 nm helical signal across a temperature range, with the data plotted as per cent unfolded. This analysis showed that SYCE2$_{\alpha}$-TEX12$_{\alpha}$ undergoes cooperative unfolding with an estimated melting temperature of 66°C (Figure 4.12 D). However, a sharp dip at approximately 65°C is observed, with the experiment repeated several times to confirm the feature was not due to practical errors. This feature is also observed with the previously published data (Davies *et al*., 2012) and thus, indicates that this event is a feature of the complex, potentially due to the occurrence of a mixture of unfolding events from the 4:4 complex and any higher order assembly.

To compare the oligomeric state of the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex to the published data, SEC-MALS experiments were performed. This analysis showed that the majority of the material exists as a 4:4 species with the formation of larger species (Figure 4.12 E), which was in agreement with the published data (Davies *et al*., 2012). To validate that the higher-order assemblies were formed by the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex, and not due to contamination products, SDS-PAGE analysis was carried out with the output fractions from the SEC-MALS elution profile. This analysis demonstrates the two separate species of the pure SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex that corresponds with the elution pattern (Figure 4.12 F).

As the SEC-MALS data shows the capacity for SYCE2$_{\alpha}$-TEX12$_{\alpha}$ to form larger molecular weight species in solution, TEM experiments were performed to test the formation of filamentous structures, as previously published (Davies *et al*., 2012). The TEM analysis demonstrates the capacity of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ to assemble into filaments, with both long single filaments and branched networks detected (Figure 4.13 A). The analysis of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ agrees with the previously published data, and the comparison of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ to the full length SYCE2-TEX12 complex indicates that, whilst the core helical region is sufficient for filamentous assembly, these assemblies are encouraged by the unstructured ends of SYCE2 and TEX12.

The assembled SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex was characterised by SEC-SAXS to compare the parameters of the assemblies between the core and full length complexes. The Rg for SYCE2$_{\alpha}$-TEX12$_{\alpha}$ was measured as 142 Å (Figure 4.13 B) and the Rc was determined as 15 Å (Figure 4.13 C). The real space distance distribution analysis indicates an extended rod shape

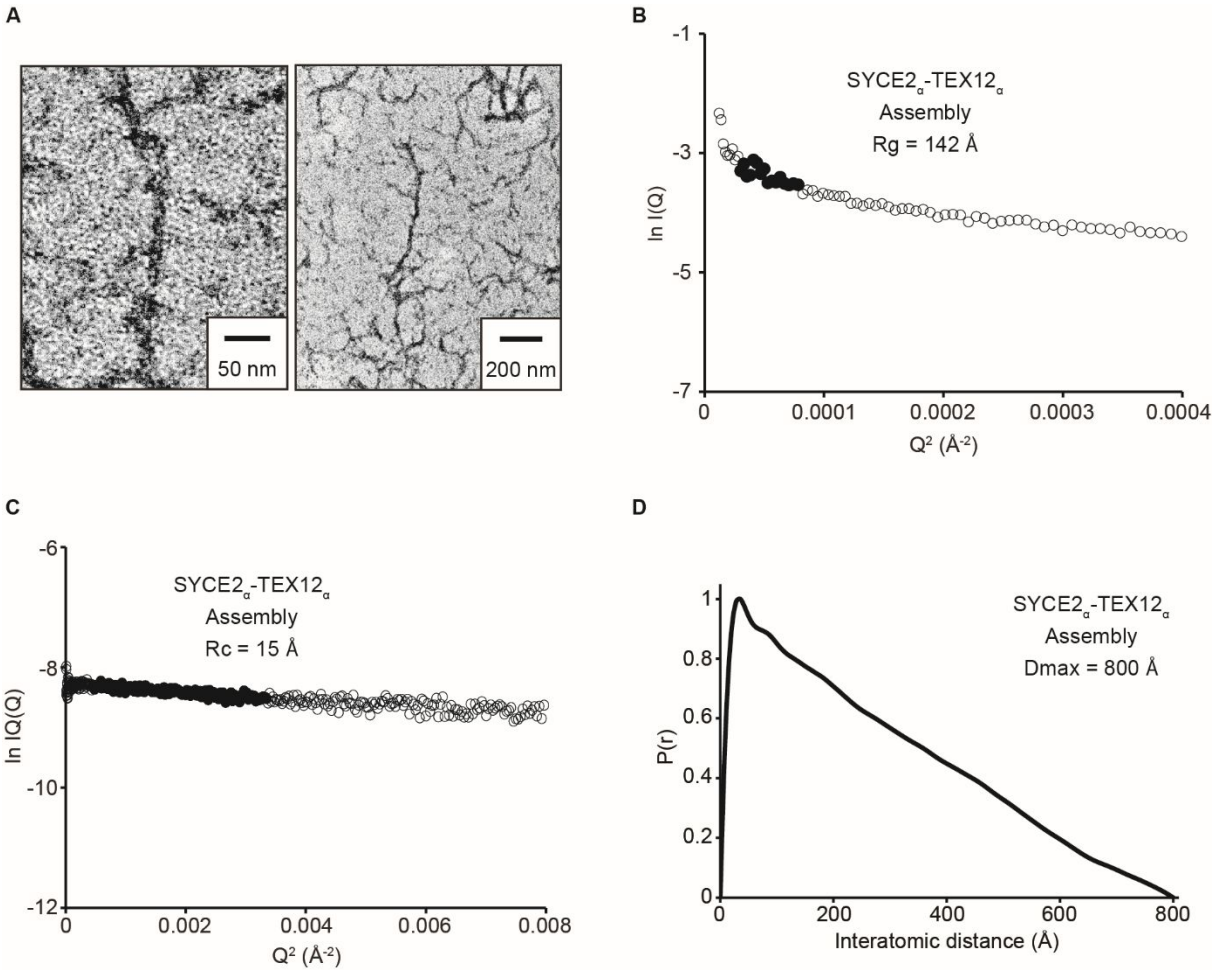


Figure 4.13. Analysis of SYCE2$_\alpha$-TEX12$_\alpha$ assembly. (A) TEM analysis of SYCE2$_\alpha$-TEX12$_\alpha$ filaments. (B-D) SEC-SAXS analysis of SYCE2$_\alpha$-TEX12$_\alpha$ assembly. (B) Guinier analysis to determine the radius of gyration (Rg) as 142 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (C) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 15 Å. Clear circles are the data, solid circles represent the region used for the fit. (D) Paired real-space distribution plot. The SYCE2$_\alpha$-TEX12$_\alpha$ assembly has a maximum dimension of 800 Å.

with a maximum distance of 800 Å (Figure 4.13 D). These values correspond to an elongated filament structure, as characterised for the full length SYCE2-TEX12 complex. However, the Rg value of core complex is larger than the full length complex, but the Rc value is roughly half the value of the full length complex, suggesting that the full length SYCE2-TEX12 assemblies are a bulkier filament structure, due to the presence of the unstructured regions.

To study the structural properties of the core 4:4 unit, SEC-SAXS analysis was performed with the 4:4 SYCE2$_\alpha$-TEX12$_\alpha$ species. The Rg for the 4:4 unit was determined to be 62 Å (Figure 4.14 A) and the Rc was measured as 21 Å (Figure 4.14 B). As predicted this suggests that a more compact structure is formed, although the complex is still elongated in nature. The distance distribution curve highlights the extended nature of the 4:4 complex, and estimates the maximum distance observed within the sample to be 210 Å (Figure 4.14 C). To

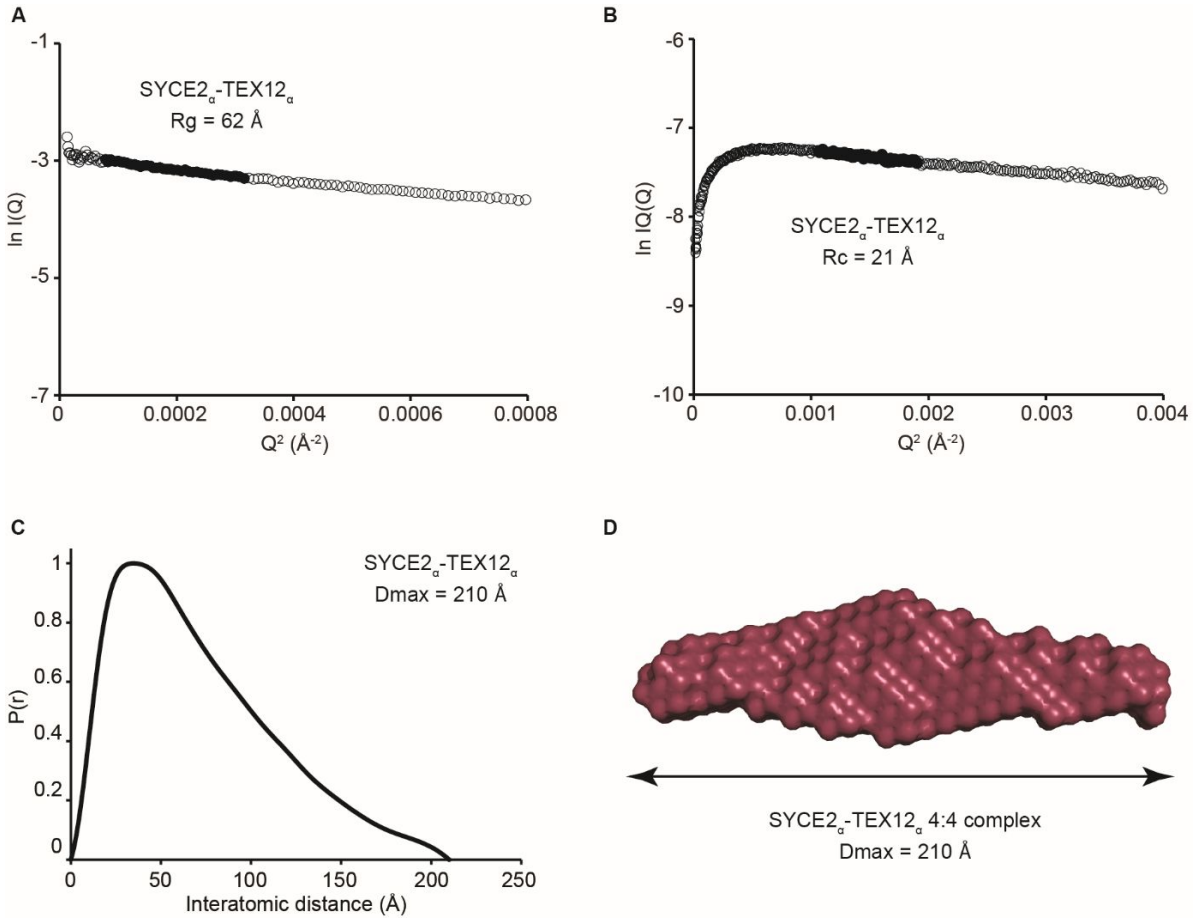


Figure 4.14. SEC-SAXS analysis of the 4:4 SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex. (A) Guinier analysis to determine the radius of gyration (Rg) as 62 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 21 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. The SYCE2$_{\alpha}$-TEX12$_{\alpha}$ 4:4 complex has a maximum dimension of 210 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

further study the shape of the 4:4 complex, *ab initio* modelling was performed to generate a low resolution molecular envelope. The produced model verified the elongated nature of the 4:4 complex and although the modelling was performed without an imposed symmetry, the resultant envelope is reasonably symmetrical (Figure 4.14 D).

The low resolution characterisation of the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ structure in solution has revealed that an elongated 4:4 helical complex is assembled with the capacity to form filamentous structures. However, to fully understand the molecular basis of SYCE2-TEX12 complex formation, a high resolution structure is required. Therefore, crystallographic studies were initiated with the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex, using material following SEC to isolate the 4:4 species and to remove larger assembles that may inhibit crystallisation.

High throughput commercial screening was performed, with some drops containing crystalline precipitate and two conditions containing small crystals (Figure 4.15 A and B).

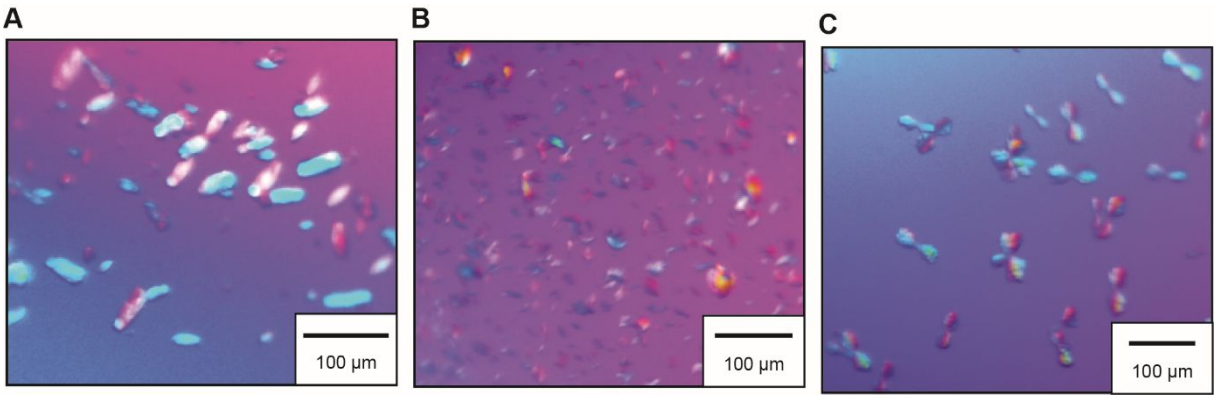


Figure 4.15. Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha}$. (A) Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ at 12 mg/ml in 0.2 M sodium chloride, 0.1 M sodium acetate pH 4.6, 30 % v/v MPD. (B) Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ at 12 mg/ml in 0.02 M calcium chloride dehydrate, 0.1 M sodium acetate pH 4.6, 30 % v/v MPD. (C) Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ with 20% glycerol at 12 mg/ml in 0.2 M lithium chloride, 40% (v/v) MPD.

These crystallisation conditions were used for optimisation screening. However, the crystals were not reproduced following extensive screening of well solution component concentrations, pH and protein concentration. Although the 4:4 species was isolated for crystallisation, the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex has the capacity to assemble, and it is possible that assembly is triggered upon the high protein concentration imposed during crystallisation. Therefore, glycerol was added to the 4:4 complex material before crystallisation to attempt to stabilise the 4:4 species. Some crystals were formed from this screening, although they formed elongated micro-crystal clusters (Figure 4.15 C). These crystals were not improved upon optimisation screening. Therefore, no crystals were obtained from extensive screening of SYCE2$_{\alpha}$-TEX12$_{\alpha}$, and suggested that an optimised SYCE2-TEX12 complex would be required to fully block assembly before crystallisation.

As the removal of the 10 TEX12 C-terminal residues abolished assembly of the full length SYCE2-TEX12 complex, this truncation was introduced into the core complex (Figure 4.16 A), hereby referred to as SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$. The SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex was purified in the same manner as SYCE2$_{\alpha}$-TEX12$_{\alpha}$ (Figure 4.16 B). CD experiments were performed to verify that the folding of the core complex was not disrupted upon the TEX12 truncation. The far-UV CD scan recorded a highly helical signal, with an estimated helical content of 91% (Figure 4.16 C).

This corresponds to the 92% helical content determined for SYCE2$_{\alpha}$-TEX12$_{\alpha}$, validating that the truncation did not alter the secondary structure composition. CD thermal denaturation was carried out to study the thermal stability of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$. A single cooperative unfolding event was observed, with an estimated melting temperature of 67°C

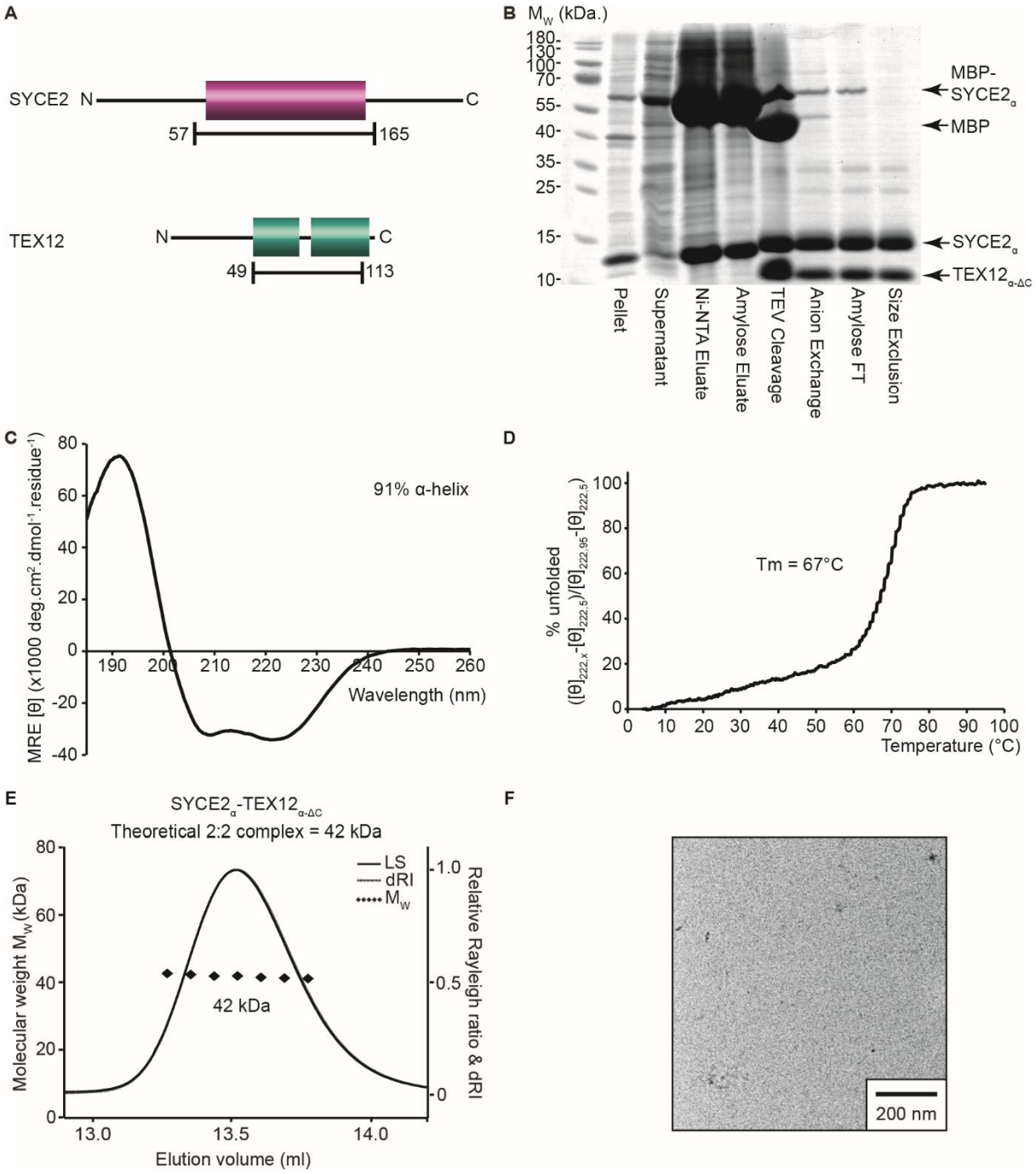


Figure 4.16. Purification and analysis of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$. This data was presented in my MRes thesis. (A) Schematics of SYCE2 and TEX12 showing construct boundaries. (B) SDS-PAGE summary of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ purification by Ni-NTA and amylose affinity chromatography steps, followed by TEV protease incubation and subsequent anion exchange and size exclusion chromatography. (C) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 91% alpha-helix, 1% β-sheet, 6% turns and 2% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 67°C, when the sample is 50% unfolded. (E) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 42 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 42 kDa. (F) TEM analysis of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ revealing no filament formation.

(Figure 4.16 D). This is comparable to the 66°C melting temperature of SYCE2$_{\alpha}$-TEX12$_{\alpha}$, although no dip in the curve was observed with the truncated complex.

The oligomeric state of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ was studied by SEC-MALS, and was found to form a single species with a molecular mass of 42 kDa (Figure 4.16 E). This corresponds to a 2:2 complex, as was formed with the SYCE2-TEX12$_{\Delta C}$ complex. TEM was carried out with SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ and no filamentous structures were identified, also agreeing with the findings for the SYCE2-TEX12$_{\Delta C}$ complex (Figure 4.16 F).

The structural properties of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex were studied by SEC-SAXS. The Rg was measured as 48 Å (Figure 4.17 A) and the Rc was measured as 13 Å (Figure 4.17 B). The real space distance distribution curve indicates an extended shape, with the estimation of the maximum distance within the sample to be 195 Å (Figure 4.17 C). *Ab initio* modelling was performed to produce the low resolution molecular envelope of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex (Figure 4.17 D). The resultant model demonstrates the extended rod shape of the complex, with the maximum dimension similar to the 210 Å distance as measured for the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ 4:4 complex. The Rc value for the 2:2 complex is almost half the value of the 4:4 complex, and the two complexes have similar maximum dimensions. This implies that the 4:4 complex is formed by a lateral organisation of two 2:2 complexes, rather than a longitudinal association.

As the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex is highly structured and does not undergo filamentous assembly, this complex was used for crystallographic studies to attempt to crystallise the complex for the determination of a high resolution structure by X-ray diffraction. High throughput commercial crystallisation screening was performed with SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$, and from screening of two drop sizes in 480 different conditions, only one drop was found to contain crystal growth. The needle shaped crystals were roughly 50 µm in length but only the larger crystals reached 10 µm in width (Figure 4.18 A and B). The crystals were harvested in order to test for X-ray diffraction, with rapid deterioration of the crystals upon opening of the well. The largest crystal in the drop was successfully harvested and tested for diffraction at a micro-focus X-ray beamline (Figure 4.18 C). From testing the diffraction of the crystal, low resolution diffraction spots were detected (Figure 4.18 D). Although this suggests that the crystal contains internal order, the crystals required optimisation to improve crystal quality and size to obtain diffraction data for structure solution. As the crystallisation solution contained 20% 2-propanol, optimisation of these crystals was difficult, due to the volatility of the propanol. From extensive screening of both sitting and hanging drop vapour diffusion drops, with a large variety of screening attempts altering both the well solution and protein

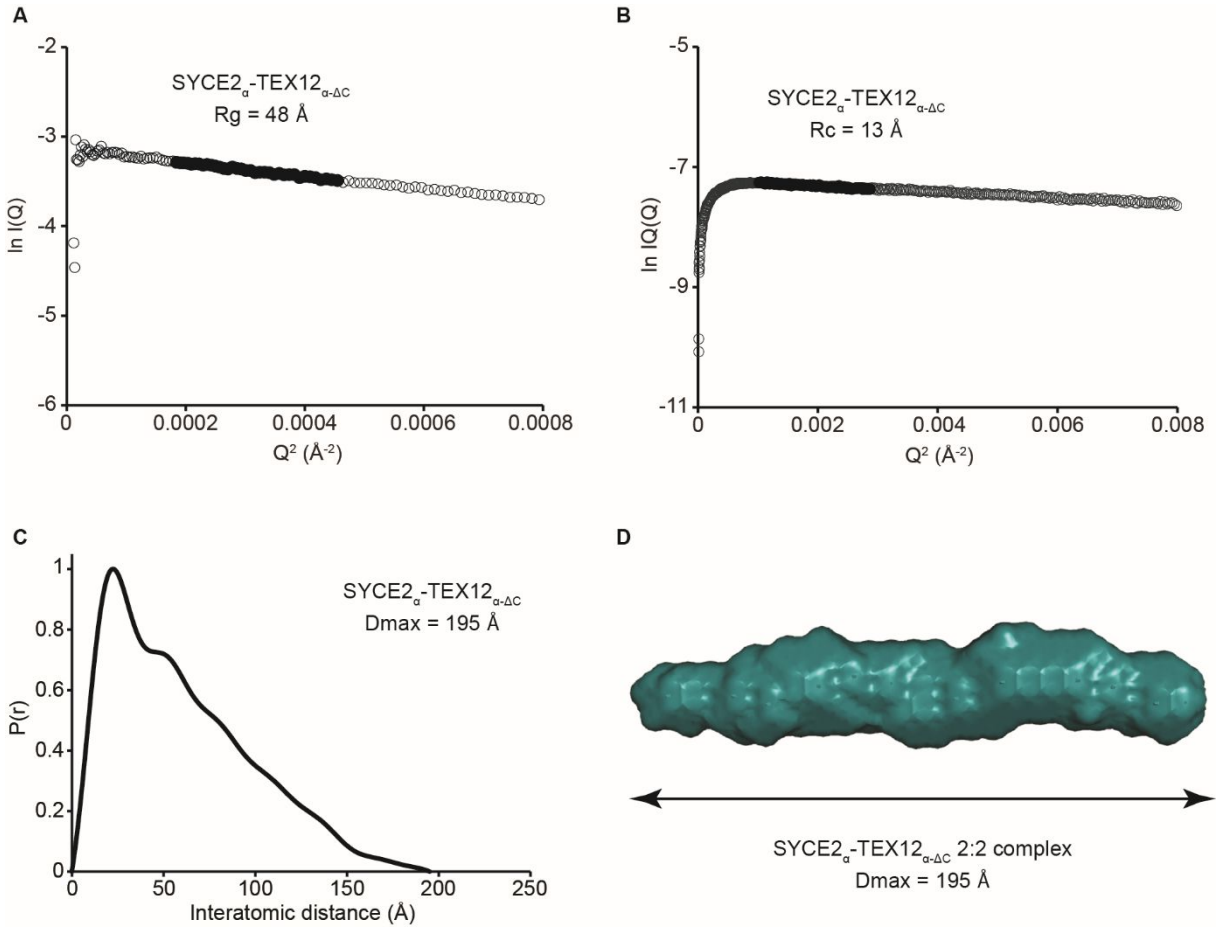


Figure 4.17. SEC-SAXS analysis of the 2:2 SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex. (A) Guinier analysis to determine the radius of gyration (Rg) as 48 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 13 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. The SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex has a maximum dimension of 195 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

concentration, only smaller and lower quality crystals were formed, even upon repeating the initial commercial screening.

Further commercial screening was carried out with the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex through the addition of 20% glycerol to the protein stock prior to crystallisation screening in an attempt to stabilise the complex to improve crystal growth. From these screens, more drops were found to contain micro-crystals, with one condition containing crystals up to 100 µm in length and 20-50 µm in width (Figure 4.19 A). These crystals were harvested in order to test for X-ray diffraction. The diffraction pattern obtained did not contain suitable diffraction data for structure solution, but rather exhibited a fibrous diffraction pattern (Figure 4.19 B). The

pattern consisted of a 5.1 Å meridional arc, and equatorial patterns at 12 Å (Figure 4.19 B). The observed fibre diffraction corresponds to the specific k-m-e-f class of coiled-coil fibres, where a specific 5.1 Å meridional arc is observed with 9 Å equatorial planes for dimeric coiled-coil fibres (Fraser and MacRae, 1961). Although the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex does not form filamentous structures by TEM or in solution, residual affinity for higher-order assembly is maintained in order to form fibrous crystals. From further analysis of the equatorial diffraction, d-spacing values could be determined, corresponding to the distances between unit cells, with a length of 150 Å and a width of 50 Å (Figure 4.19 B). SEC-SAXS experiments have determined the cross-sectional radius of the filaments to be 26-27 Å, correlating to the 50 Å width, and the length of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex to be 195 Å, roughly corresponding to the 150 Å d-spacing.

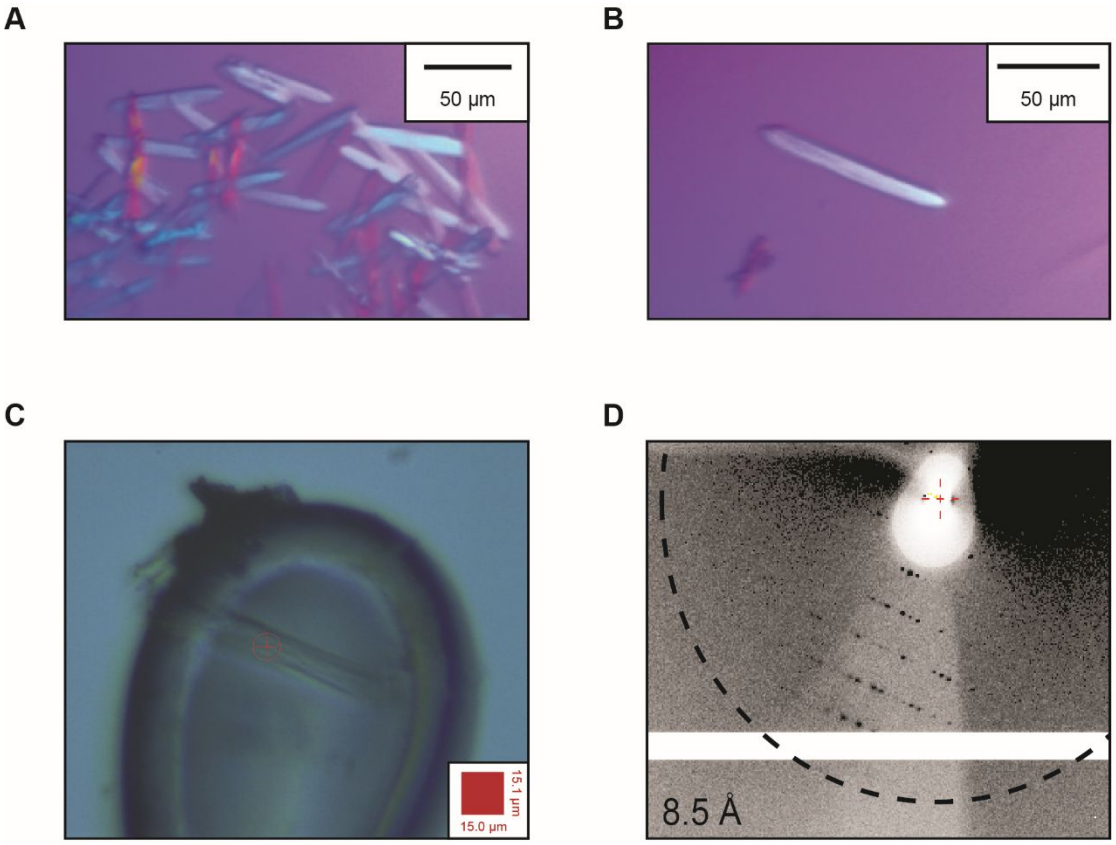


Figure 4.18. Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$. (A) and (B) Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ at 17 mg/ml in 0.14 M calcium chloride, 0.07 M sodium acetate pH 4.6, 30 % glycerol and 14% v/v 2-propanol. (C) Image of the crystal in the loop at the I24 micro focus beamline (Diamond Light Source). (D) Diffraction image collected from the crystal at the I24 beamline, with low resolution protein diffraction.

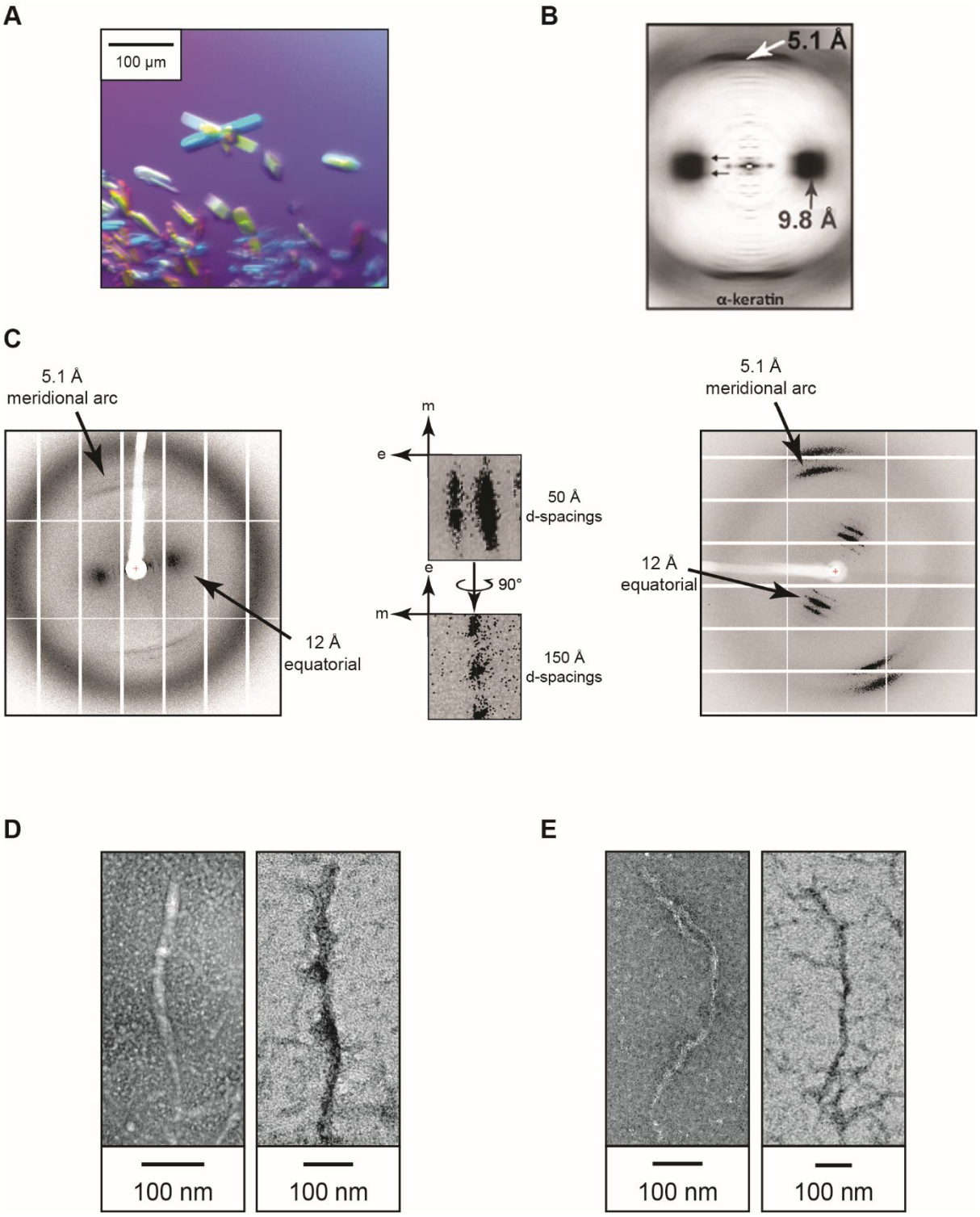


Figure 4.19. Fibrous assembly of SYCE2-TEX12. (A) Crystallisation of SYCE2$_{\alpha}$-TEX12$_{\alpha-\Delta C}$ with 20% glycerol at 17 mg/ml in 0.05 M caesium chloride, 0.1 M MES pH 6.5, 30 % v/v Jeffamine M-600. (B) The α-keratin diffraction pattern shows horseshoe equatorial reflections at 9.8 Å and a meridional arc at 5.1 Å. Original image collected by T. P. MacRae, taken from Parry *et al*., 2008. (C) X-ray diffraction images collected at the I03 beamline, showing a fibre diffraction pattern from two crystals at different orientations. Analysis of the equatorial planes determines d-spacings of 50 Å and 150 Å. (D) TEM analysis of SYCE2-TEX12 showing both positively and negatively stained filaments by uranyl acetate fixation. (E) TEM analysis of SYCE2$_{\alpha}$-TEX12$_{\alpha}$ showing both positively and negatively stained filaments by uranyl acetate fixation.

Therefore, the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex assembles into fibrous crystals, correlating to the k-m-e-f class of coiled-coil fibres. These fibres may mimic the filamentous assembly detected by TEM, whereby extensive filamentous structures are formed by the full length SYCE2-TEX12 complex and the helical core SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex (Figure 4.19 C and D).

4.2.4 The dual modality of the TEX12 C-terminus in SYCE2-TEX12 assembly

The C-terminus of TEX12 has been implicated in higher-order assembly of the SYCE2-TEX12 complex, but the C-terminal truncation retained propensity for fibrous assembly upon crystallisation. In order to dissect the molecular basis of the higher-order assembly, the two groups of candidate C-terminal residues, as described for the full length complex (see above), were characterised in the context of the core complex.

As the TEX12 C-terminus contains potential coiled-coil heptads, which were shown to be required for the higher-order assembly of the full length SYCE2-TEX12 complex, the four candidate hydrophobic residues were mutated to glutamic acid for analysis within the core complex, hereby referred to as SYCE2$_{\alpha}$-TEX12$_{\alpha\ cc\text{-}mut}$ (Figure 4.20 A). This mutated complex was purified with an MBP affinity tag fused to SYCE2 and a His$_6$-MBP tag fused to TEX12. The proteins were co-purified by affinity and anion exchange chromatography steps, with the tags removed by TEV protease and subsequent purification of the cleaved complex by anion exchange chromatography (Figure 4.20 B).

To compare the oligomeric state of the mutated complex compared to SYCE2$_{\alpha}$-TEX12$_{\alpha}$, SEC-MALS analysis was carried out. A single protein species was detected with a calculated molecular weight of 44 kDa (Figure 4.20 C). This corresponds to the formation of a 2:2 complex, as observed for the full length SYCE2-TEX12$_{cc\text{-}mut}$ complex. To confirm that higher-order assembly was abolished, TEM experiments were performed. Through this analysis, no filaments were observed (Figure 4.20 D).

The solution structure of the SYCE2$_{\alpha}$-TEX12$_{\alpha\ cc\text{-}mut}$ 2:2 complex was investigated by SEC-SAXS. The Rg was measured to be 48 Å (Figure 4.20 E) and the Rc was determined to be 14 Å (Figure 4.20 F). Real space analysis of the interatomic distances observed within the sample revealed a P(r) profile indicative of an extended rod shape, with an estimated maximum length of 190 Å (Figure 4.20 G). These structural parameters correlate to those determined for the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ complex, suggesting the same 2:2 complex is formed, and that the coiled-coil formation at the TEX12 C-terminal is required to form the 4:4 complex and consequent filamentous structures.

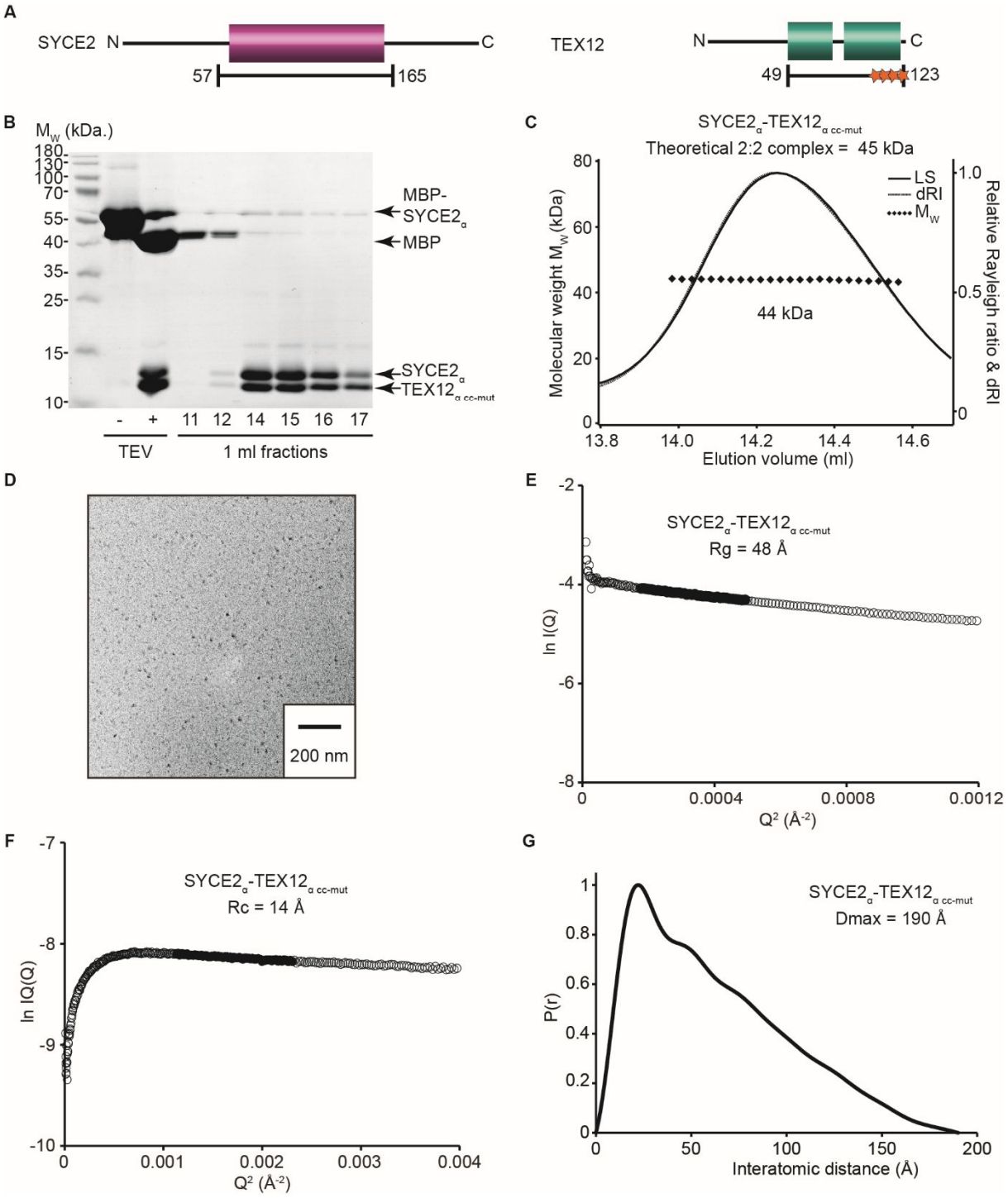


Figure 4.20. Purification and analysis of SYCE2$_\alpha$-TEX12$_{\alpha\ cc\text{-}mut}$. (A) Schematic of SYCE2 and TEX12 boundaries. Stars represent point mutations. (B) SDS-PAGE of SYCE2$_\alpha$-TEX12$_{\alpha\ cc\text{-}mut}$ purification by anion exchange chromatography, following tag removal by TEV protease incubation. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 44 kDa, with the theoretical molecular weight of a 2:2 complex estimated to be 45 kDa. (D) TEM analysis of SYCE2$_\alpha$-TEX12$_{\alpha\ cc\text{-}mut}$ revealing no filament formation. (E-G) SEC-SAXS analysis of the SYCE2$_\alpha$-TEX12$_{\alpha\ cc\text{-}mut}$ 2:2 complex. (E) Guinier analysis to determine the radius of gyration (Rg) as 48 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (F) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 14 Å. Clear circles are the data, solid circles represent the region used for the fit. (G) Paired real-space distribution plot. The SYCE2$_\alpha$-TEX12$_{\alpha\ cc\text{-}mut}$ 2:2 complex has a maximum dimension of 190 Å.

The mutation of the potential surface exposed patch, composed of TEX12 F102, F109 and V116 residues (as described above), exhibited a mild decrease in assembly, but did not abolish filamentous propensity. In order to probe the significance of these residues for assembly, the three candidate residues were simultaneously mutated for characterisation in the core helical complex. Two constructs were characterised, in which the F102 and V116 residues were both mutated to alanine, with the F109 residue mutated to alanine or glutamic acid. The complexes are hereby referred to as SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-3A}}$ and SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ respectively.

The SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-3A}}$ complex (Figure 4.21 A) was purified in the same manner as SYCE2$_{\alpha}$-TEX12$_{\alpha}$, with the cleaved complex purified by anion exchange and preparative size exclusion chromatography steps (Figure 4.21 B). The oligomeric state of the purified SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-3A}}$ complex was characterised by SEC-MALS. This analysis showed the formation of a single species with a molecular weight of 85 kDa (Figure 4.21 C). This corresponds to a 4:4 complex, with no higher-order assembly observed. The capacity for assembly was further investigated by TEM, and again no filaments were detected (Figure 4.21 D). In the context of the full length complex, this mutation did not disrupt assembly. Therefore, these residues may drive the assembly of the core complex, but the unstructured arms can overcome their mutation to assemble. This consequently suggests that the 4:4 complex is required as the building block for filament formation.

The SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ complex (Figure 4.22 A) was purified by the same procedure as SYCE2$_{\alpha}$-TEX12$_{\alpha}$, with the cleaved complex purified by anion exchange chromatography (Figure 4.22 B). SEC-MALS analysis showed the formation of a single species of 85 kDa, corresponding to a 4:4 complex (Figure 4.22 C) and from TEM analysis, no filaments were observed (Figure 4.22 D). SEC-SAXS experiments were performed to study the structural properties of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ 4:4 complex in order to compare to the 4:4 complex formed by SYCE2$_{\alpha}$-TEX12$_{\alpha}$. The Rg was measured to be 54 Å (Figure 4.22 E) and the Rc was determined to be 20 Å (Figure 4.22 F). The interatomic distance distribution analysis estimates the maximum length within the sample is 200 Å, with a P(r) profile indicative of an extended shape (Figure 4.22 G). *Ab initio* modelling was performed to produce the molecular envelope of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ 4:4 complex (Figure 4.23 A). These structural parameters correlate to the SAXS analysis of SYCE2$_{\alpha}$-TEX12$_{\alpha}$, suggesting the 4:4 units form the same structure.
As the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ 4:4 complex could be purified to a much larger yield than the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ complex and did not exhibit any propensity for high-order assembly. Therefore, this complex was used for crystallographic studies to attempt to elucidate the high resolution structure of the SYCE2-TEX12 4:4 unit.

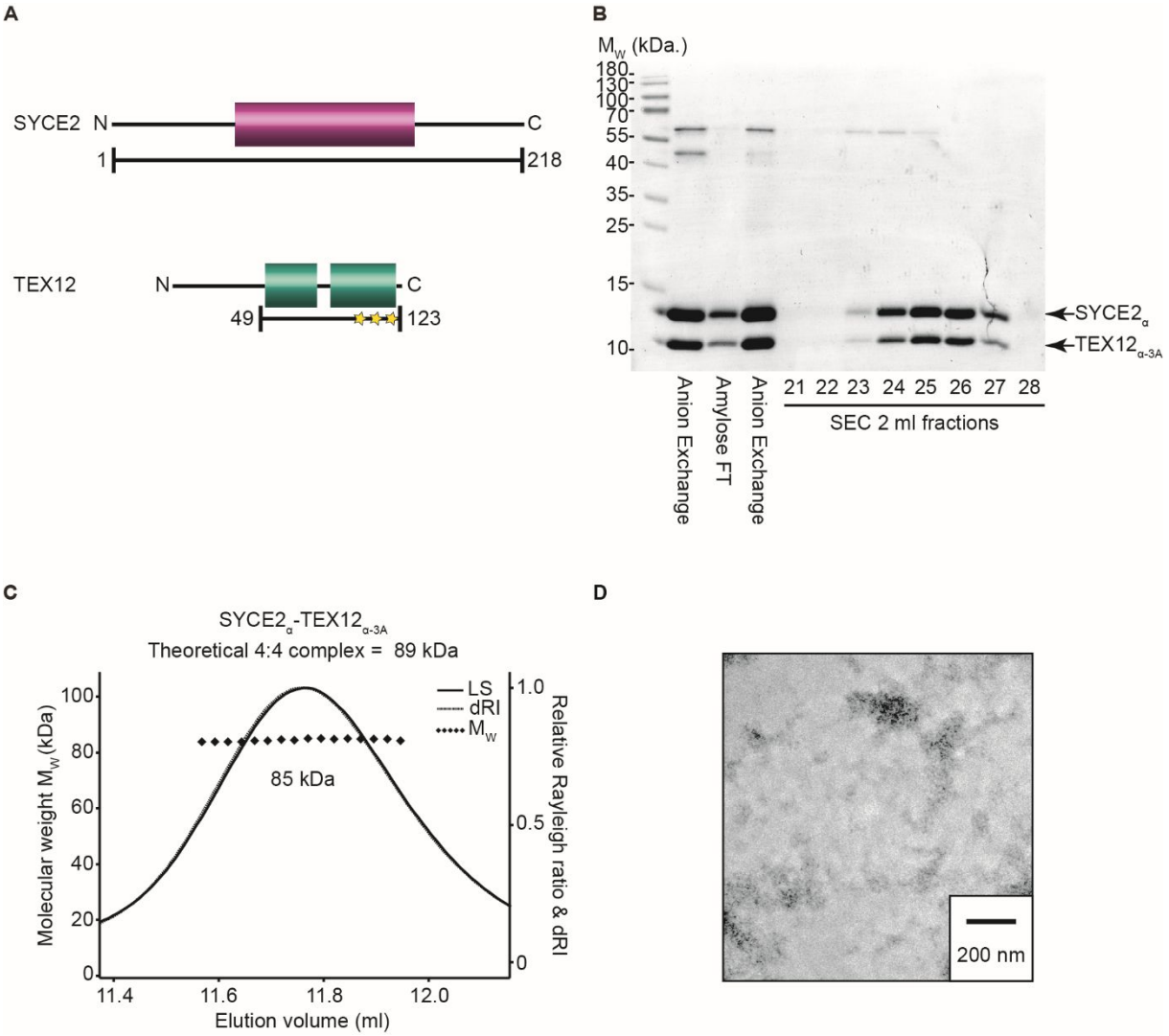


Figure 4.21. Purification and analysis of SYCE2$_\alpha$-TEX12$_{\alpha\text{-}3A}$. (A) Schematic of SYCE2 and TEX12 boundaries. Stars represent point mutations. (B) SDS-PAGE of SYCE2$_\alpha$-TEX12$_{\alpha\text{-}3A}$ purification by size exclusion chromatography, following anion exchange chromatography subsequent to tag removal by TEV protease incubation. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 85 kDa, with the theoretical molecular weight of a 4:4 complex estimated to be 89 kDa. (D) TEM analysis of SYCE2$_\alpha$-TEX12$_{\alpha\text{-}3A}$ revealing no filament formation.

High throughput commercial screening was carried out with the SYCE2$_\alpha$-TEX12$_{\alpha\text{-}AEA}$ complex, which led to the formation of crystals of varying sizes in multiple conditions (Figure 4.23 B and C). Optimisation screens were performed to improve crystal quality and size (Figure 4.23 D and E), with many crystals from both the initial and optimisation screening harvested to test for X-ray diffraction. From screening of the harvested crystals, no diffraction was observed, suggesting that, despite appearance, there is disorder within the crystal lattice. Further optimisation screening was performed, including micro-seeding and additive screening, with resultant crystals forming the same size and shape with no X-ray diffraction observed.

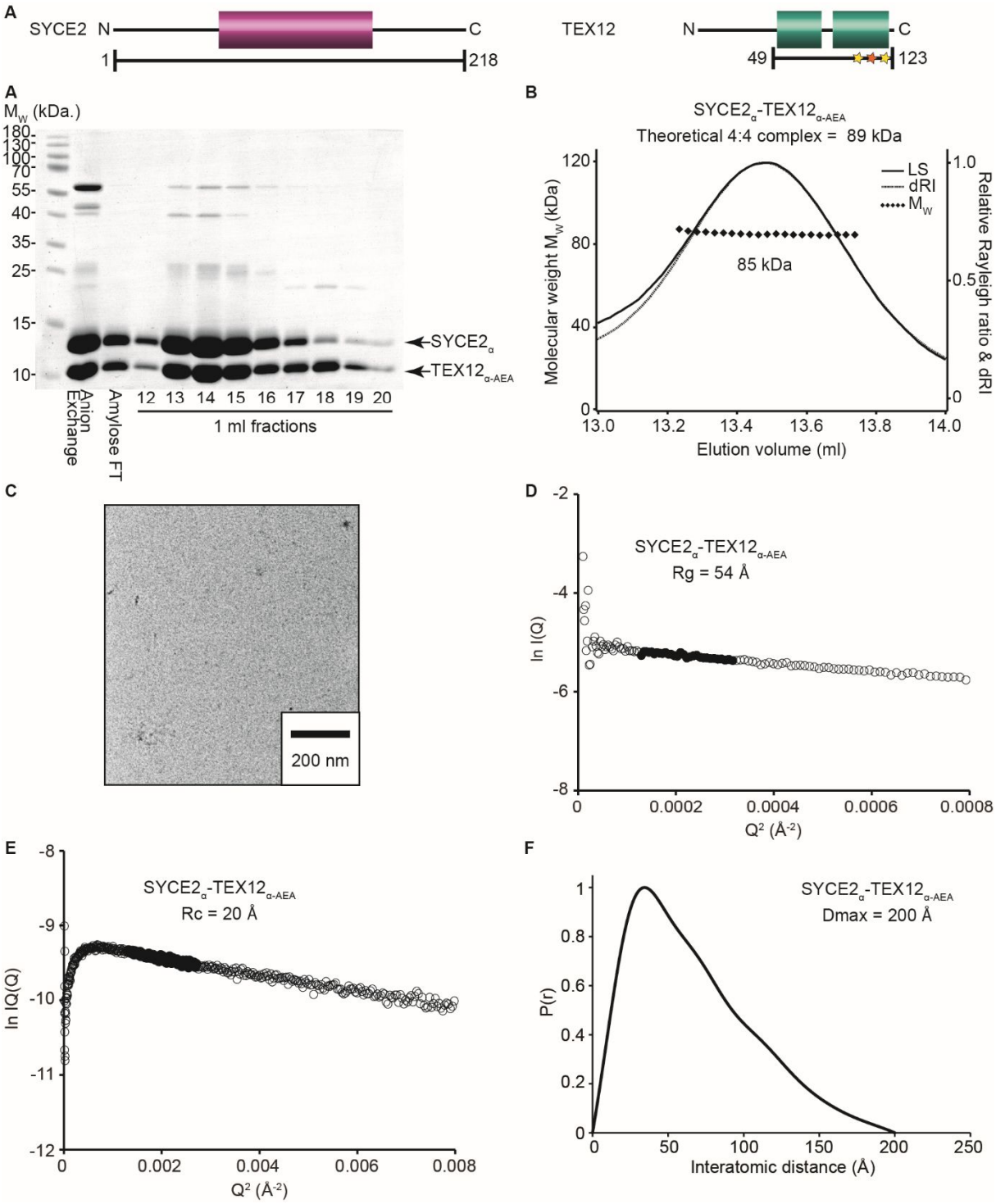


Figure 4.22. Purification and analysis of SYCE2$_\alpha$-TEX12$_{\alpha\text{-AEA}}$. (A) Schematic of SYCE2 and TEX12 boundaries. Stars represent point mutations (yellow- alanine, orange- glutamic acid). (B) SDS-PAGE of SYCE2$_\alpha$-TEX12$_{\alpha\text{-AEA}}$ purification by anion exchange chromatography, following tag removal by TEV protease incubation. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 85 kDa, with the theoretical molecular weight of a 4:4 complex estimated to be 89 kDa. (D) TEM analysis of SYCE2$_\alpha$-TEX12$_{\alpha\text{-AEA}}$ revealing no filament formation. (E-G) SEC-SAXS analysis of the SYCE2$_\alpha$-TEX12$_{\alpha\text{-AEA}}$ 4:4 complex. (E) Guinier analysis to determine the radius of gyration (Rg) as 54 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (F) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 20 Å. Clear circles are the data, solid circles represent the region used for the fit. (G) Paired real-space distribution plot. The SYCE2$_\alpha$-TEX12$_{\alpha\text{-AEA}}$ 4:4 complex has a maximum dimension of 200 Å.

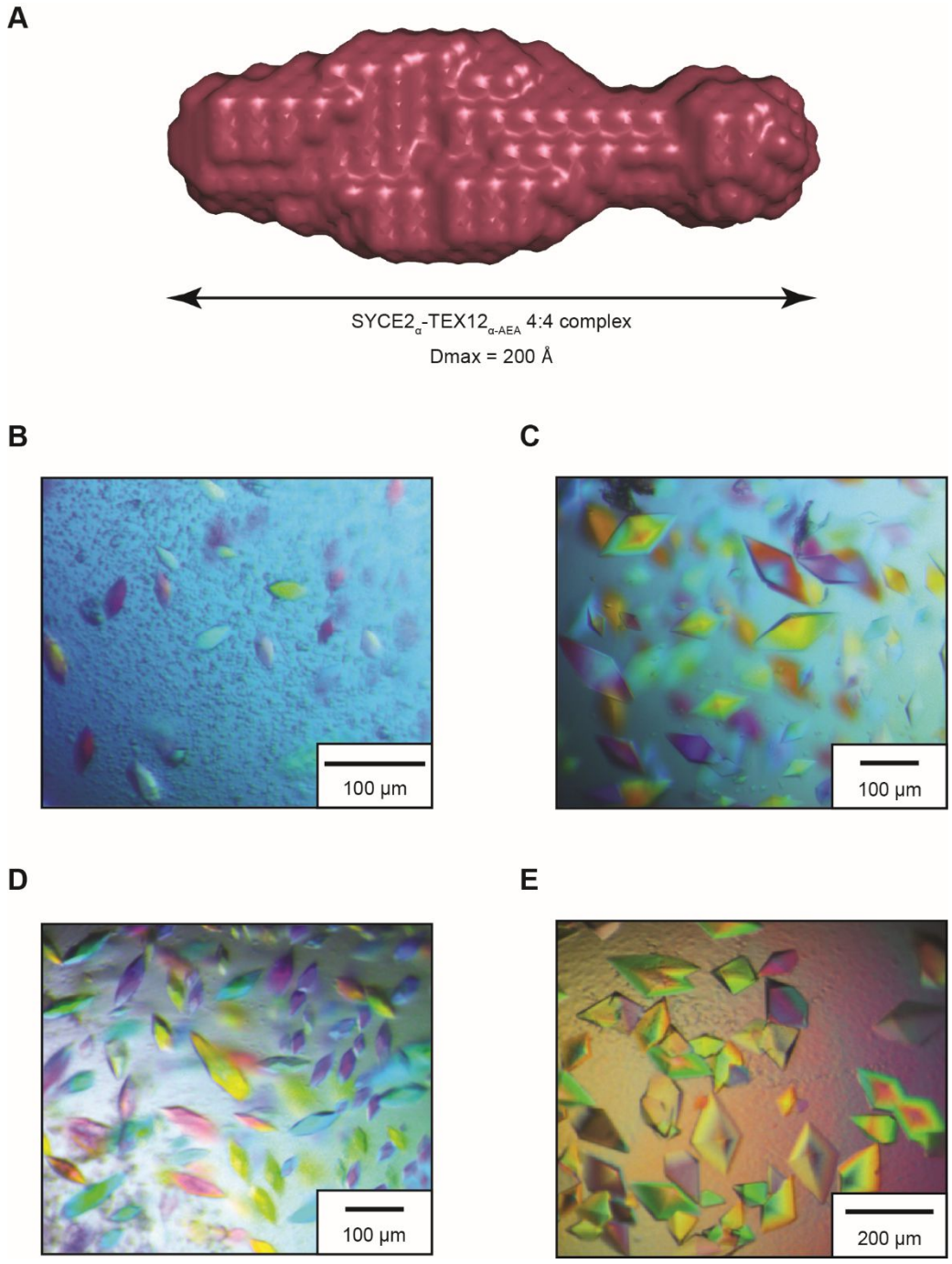


Figure 4.23. Structural analysis of the 4:4 SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ complex. (A) *Ab initio* DAMMIF model presented as the molecular envelope, with a maximum dimension of 200 Å. (B-C) SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ crystals grown from initial commercial screening. (B) Structure D10: 0.05 M potassium dihydrogen phosphate, 20 % w/v PEG 8000. (C) Morpheus G2: 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M sodium citrate tribasic dehydrate, 0.02 M sodium potassium tartrate tetrahydrate, 0.02 M sodium oxamate, 0.05 M imidazole, 0.05 M MES monohydrate acid, 20% v/v ethylene glycol and 10% w/v PEG 8000. (D-E) Optimisation of SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA}}$ crystals. (D) 0.05 M potassium dihydrogen phosphate, 10 % w/v PEG 8000. (E) 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M sodium citrate tribasic dehydrate, 0.02 M sodium potassium tartrate tetrahydrate, 0.02 M sodium oxamate, 0.05 M imidazole, 0.05 M MES monohydrate acid, 16% v/v ethylene glycol and 8% w/v PEG 8000.

Construct	$SYCE2_{\alpha}$-$TEX12_{\alpha}$	$SYCE2_{\alpha}$-$TEX12_{\alpha\text{-}\Delta C}$	$SYCE2_{\alpha}$-$TEX12_{\alpha\ cc\text{-}mut}$	$SYCE2_{\alpha}$-$TEX12_{\alpha\text{-}3A}$	$SYCE2_{\alpha}$-$TEX12_{\alpha\text{-}AEA}$
SYCE2 region	57-165	57-165	57-165	57-165	57-165
TEX12 region	49-123	49-113	49-123	49-123	49-123
Mutations	N/A	N/A	L110E F114E I117E L121E	F102A F109A V116A	F102A F109E V116A
% α-helix	92	91	N.D.	N.D.	N.D.
Tm (°C)	66	67	N.D.	N.D.	N.D.
Oligomer state	4:4 complex	2:2 complex	2:2 complex	4:4 complex	4:4 complex
Assembly	Yes	No	No	No	No
Rg (Å)	62	48	48	N.D.	54
Rc (Å)	21	13	14	N.D.	20
Dmax (Å)	210	195	190	N.D.	200
EM filaments	Yes	No	No	No	No

Table 4.3. Biophysical analysis summary of $SYCE2_{\alpha}$-$TEX12_{\alpha}$ constructs

4.2.5 Characterisation of TEX12 in isolation

The SYCE2-TEX12 complex has previously been reported as a highly constitutive assembly, with preliminary analysis of SYCE2 and TEX12 separately as fusion proteins to an MBP affinity tag revealing the formation of a tetramer and dimer, respectively (Davies *et al*., 2012). In this study, TEX12 was found to exist as a highly stable structured unit in the absence of SYCE2. Further to the stability of TEX12 in isolation, a colleague has identified the expression of SC proteins in cancer cells, contradictory to the belief that they are exclusively expressed in meiotic cells, with the high expression of TEX12 in particular correlating to a predicted poor survival prognosis in cancer patients. TEX12 has been identified to localise to centrosomes in cancer cells, with silencing of TEX12 expression in these cells resulting in abrogation of cancer cell growth (Salmon *et al.*, manuscript in preparation). This suggests that TEX12 has a biological implication away from its role in central assembly when in complex with SYCE2, and therefore, the structural properties of TEX12 were investigated.

The full length TEX12 protein (Figure 4.24 B) was over-expressed in *E. coli* as a fusion protein to a His_6-MBP affinity tag. The fusion protein was purified by sequential Ni-NTA and amylose chromatography steps, followed by anion exchange chromatography with some degradation of the affinity tag observed during these stages. The tag was removed by incubation with TEV protease, and the cleaved protein was purified by anion exchange chromatography (Figure 4.24 A). SEC-MALS was used to study the oligomeric properties of the isolated TEX12 protein. This analysis demonstrated a single protein species with a molecular weight of 27 kDa, corresponding to the formation of a dimer (Figure 4.24 C). This correlates to previous analysis of MBP-TEX12, which also forms a dimer (Davies *et al*., 2012).

The secondary structure composition of TEX12 was studied by CD experiments. A CD wavelength scan was recorded between 260 and 185 nm and displayed a typical alpha helical trace, with deconvolution estimating a helical content of 45% (Figure 4.24 D. This corresponds to the predicted unstructured nature of the TEX12 N-terminus. CD thermal denaturation was

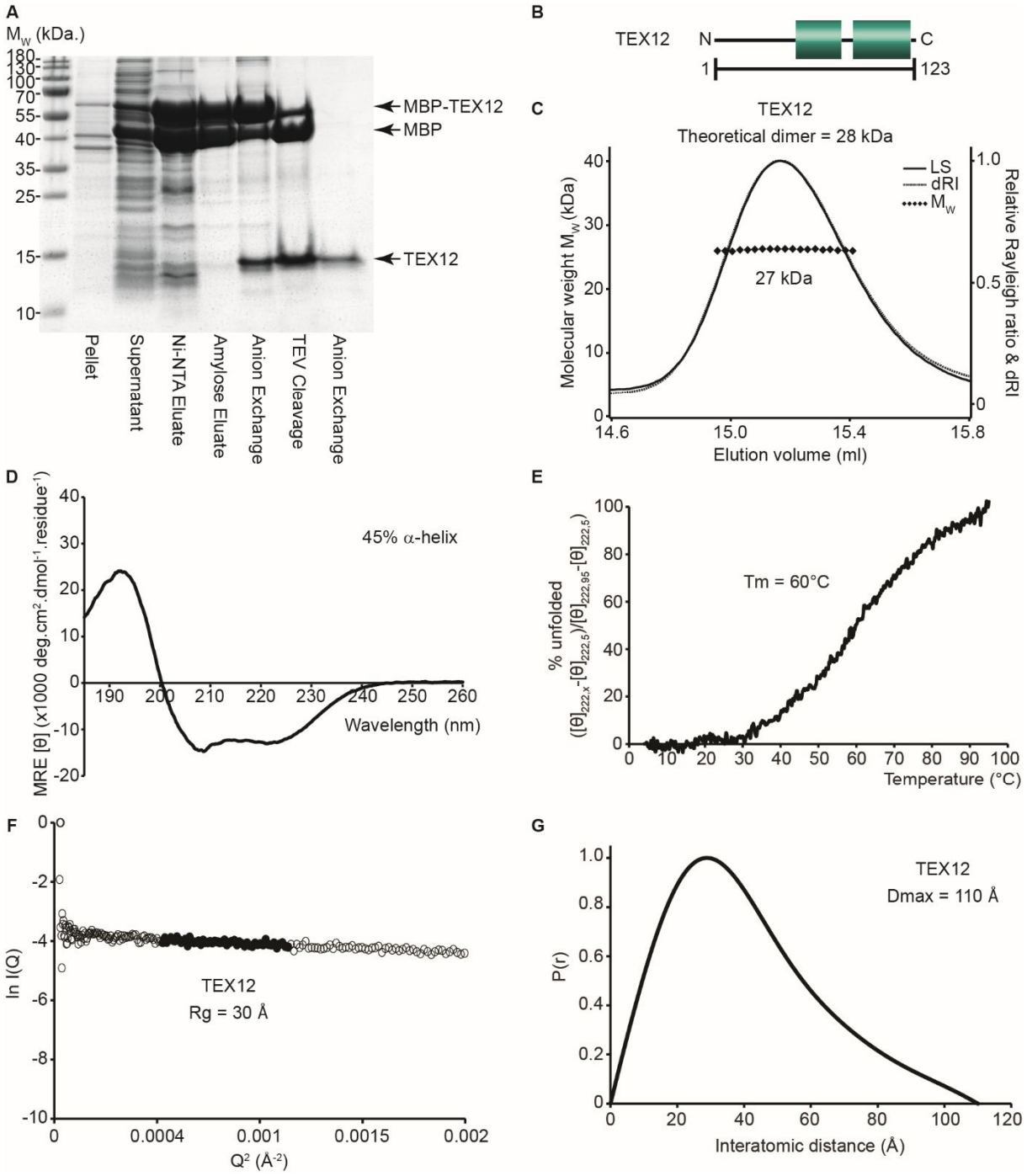


Figure 4.24. Purification and analysis of TEX12. (A) SDS-PAGE summary of the TEX12 purification by Ni-NTA, amylose, anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange chromatography. (B) Schematic of TEX12. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 27 kDa, with the theoretical molecular weight of a dimer estimated to be 28 kDa. (D) CD wavelength scan between 260-185 nm plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 45% alpha-helix, 9% β-sheet, 16% turns and 30% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 60°C, when the sample is 50% unfolded. (F-G) SEC-SAXS analysis of TEX12. (F) Guinier analysis to determine the radius of gyration (Rg) as 30 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (G) Paired real-space distribution plot. The TEX12 dimer has a maximum dimension of 110 Å.

carried out in order to study the thermal stability of the TEX12 dimer by recording the 222 nm helical signal across a temperature gradient. The resultant melting curve demonstrates a cooperative unfolding event, with an estimated melting temperature of 60°C (Figure 4.24 E). SEC-SAXS was performed with TEX12 to study the structural parameters of the TEX12 dimer. The Rg was measured as 30 Å (Figure 4.24 F) and the real space analysis determined the maximum length of the sample to be 110 Å (Figure 4.24 G). The distance distribution plot indicates a partially globular structure is formed (Figure 4.24 G).

As the TEX12 N-terminus is largely unstructured, the core helical region of TEX12 was characterised, hereby referred to as TEX12$_\alpha$ (Figure 4.25 B). The TEX12$_\alpha$ region was purified in the same manner as the full length TEX12 protein (Figure 4.25 A). SEC-MALS analysis revealed a single protein species with an estimated molecular mass of 17 kDa,

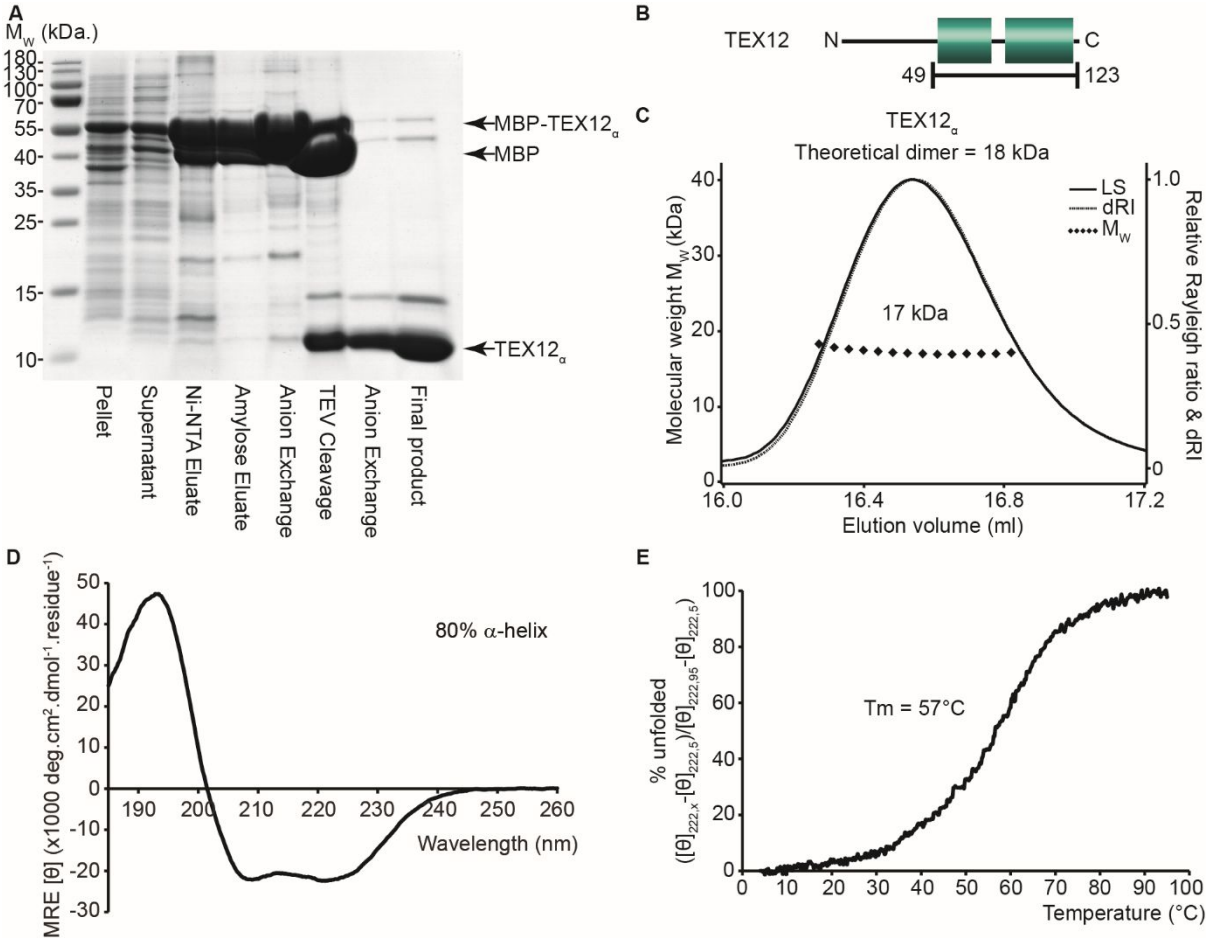


Figure 4.25. Purification and analysis of TEX12$_\alpha$. (A) SDS-PAGE summary of the TEX12 purification by Ni-NTA, amylose, anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange chromatography. (B) Schematic of TEX12 boundaries. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 17 kDa, with the theoretical molecular weight of a dimer estimated to be 18 kDa. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 80% alpha-helix, 4% β-sheet, 6% turns and 10% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 57°C, when the sample is 50% unfolded.

corresponding to a dimer (Figure 4.25 C). This demonstrates that the truncation of the N-terminus has not disrupted the oligomerisation of TEX12.

CD experiments were performed to study the structural composition of TEX12$_{\alpha}$ for comparison to the full length protein. CD wavelength scans were recorded, and the resultant trace displayed an alpha helical profile, with an estimated helical content of 80% determined by deconvolution (Figure 4.25 D). CD thermal denaturation was performed to study the unfolding of TEX12$_{\alpha}$ across a temperature gradient by monitoring the 222 nm signal. The melting curve shows a single cooperative unfolding event with an estimated melting temperature of 57°C (Figure 4.25 E). TEX12$_{\alpha}$ has a similar thermal stability to the full length protein, but with a higher helical content, indicating that the structural oligomerisation region has been isolated upon the N-terminal truncation.

The structural properties of TEX12$_{\alpha}$ were investigated by SEC-SAXS. The Rg was determined to be 21 Å (Figure 4.26 A) and the Rc was measured as 12 Å (Figure 4.26 B). The real space analysis revealed that the TEX12$_{\alpha}$ interatomic distance distribution profile is indicative of a globular structure, with the maximum distance observed in the sample determined to be 65 Å (Figure 4.26 C). *Ab initio* modelling was performed to produce the low resolution molecular envelope of TEX12$_{\alpha}$, which highlights the globular nature of the TEX12 helical core (Figure 4.26 D). These findings demonstrate that TEX12 forms a helical dimeric core that may undergo folding back of helical chains to form a compact structure, with the unstructured N-termini protruding from the structural core. In order to study the high resolution structure of the TEX12 dimer, crystallisation studies were initiated with TEX12$_{\alpha}$, in order to solve the X-ray crystal structure. However, no crystal growth was observed from commercial crystallisation screening. The concentration of TEX12$_{\alpha}$ following purification was 4 mg/ml, with an inability to improve the obtained yield despite upscaling the purification process due to precipitation of the material at higher concentrations. Therefore, a higher concentration of material may be required for successful crystallisation and consequent structure solution.

As the TEX12 C-terminus has been implicated in the higher-order assembly of the SYCE2-TEX12 complex, the role of the TEX12 C-terminal region in absence of SYCE2 was investigated by characterisation of TEX12$_{\alpha}$ with the 10 C-terminal residues truncated, hereby referred to as TEX12$_{\alpha\text{-}\Delta C}$ (Figure 4.27 B). TEX12$_{\alpha\text{-}\Delta C}$ was purified in the same manner as TEX12$_{\alpha}$ (Figure 4.27 A), with the purified cleaved TEX12$_{\alpha\text{-}\Delta C}$ reaching protein concentrations of 60 mg/ml, and thus, indicates that the truncation has stabilised TEX12.

To verify that the truncation did not disrupted the folding of the TEX12 structural core, CD experiments were performed. A CD wavelength scan of TEX12$_{\alpha\text{-}\Delta C}$ displayed an alpha

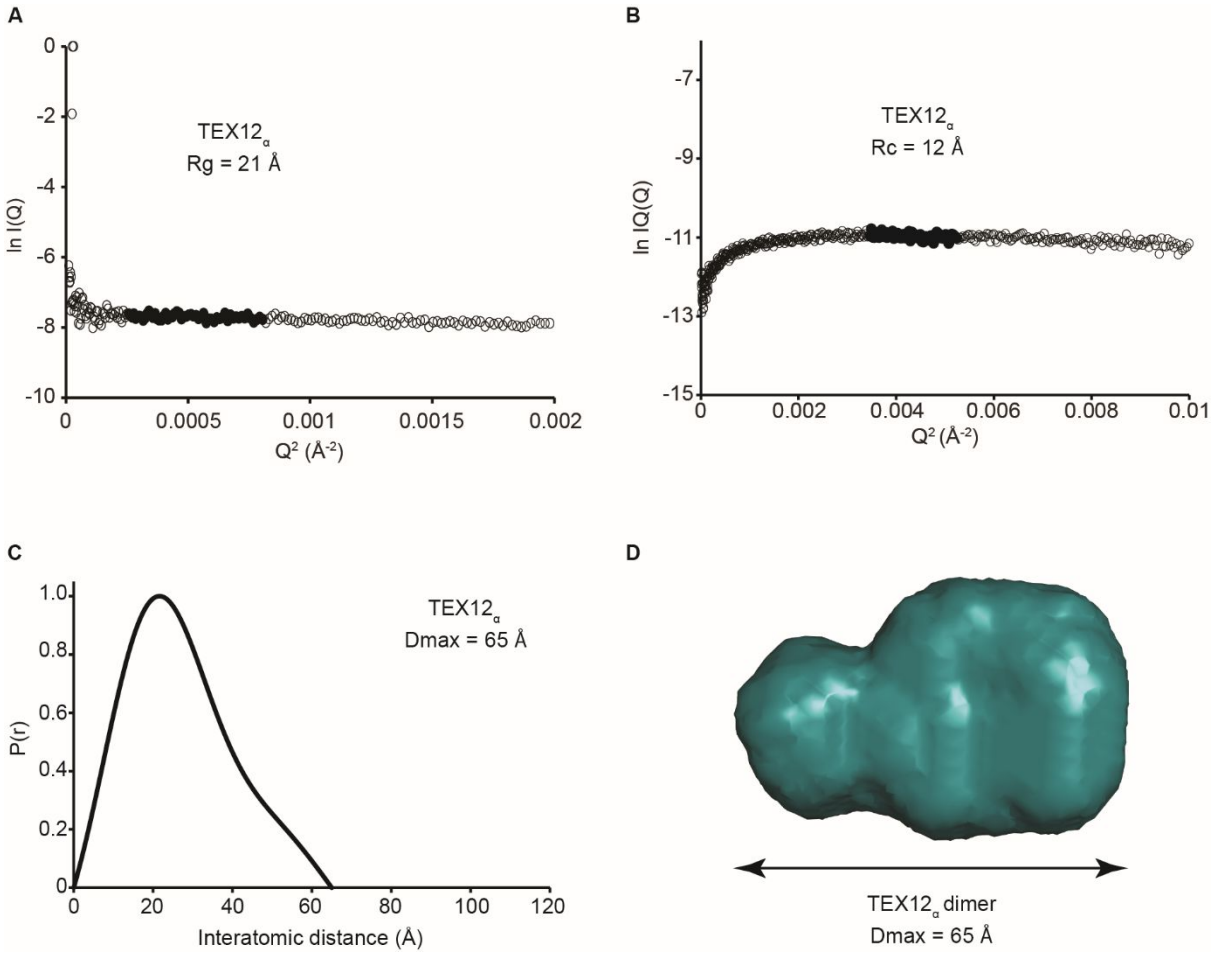


Figure 4.26. SEC-SAXS analysis of the TEX12$_{\alpha}$ dimer. (A) Guinier analysis to determine the radius of gyration (Rg) as 21 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 12 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. The TEX12$_{\alpha}$ dimer has a maximum dimension of 65 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

helical trace and deconvolution of the data determined the helical content to be 90% (Figure 4.27 D). The helical content of TEX12$_{\alpha}$ was determined as 80%, which corresponds to 62 helical residues, with the 90% helical content of TEX12$_{\alpha\text{-}\Delta C}$ relating to 61 residues. Therefore, the truncation of the TEX12 C-terminal does not disrupt the folding of TEX12. CD thermal denaturation was performed to study the thermal stability of TEX12$_{\alpha\text{-}\Delta C}$, by recording the rate of protein unfolding across a temperature gradient. The melting curve demonstrates a low thermal stability of TEX12$_{\alpha\text{-}\Delta C}$, with a rapid unfolding event with an estimated melting temperature of 27°C (Figure 4.27 E). This demonstrates that TEX12$_{\alpha\text{-}\Delta C}$ has a much lower thermal stability than TEX12$_{\alpha}$, and therefore, implies that the TEX12 C-terminus is required to stabilise the structural core of TEX12.

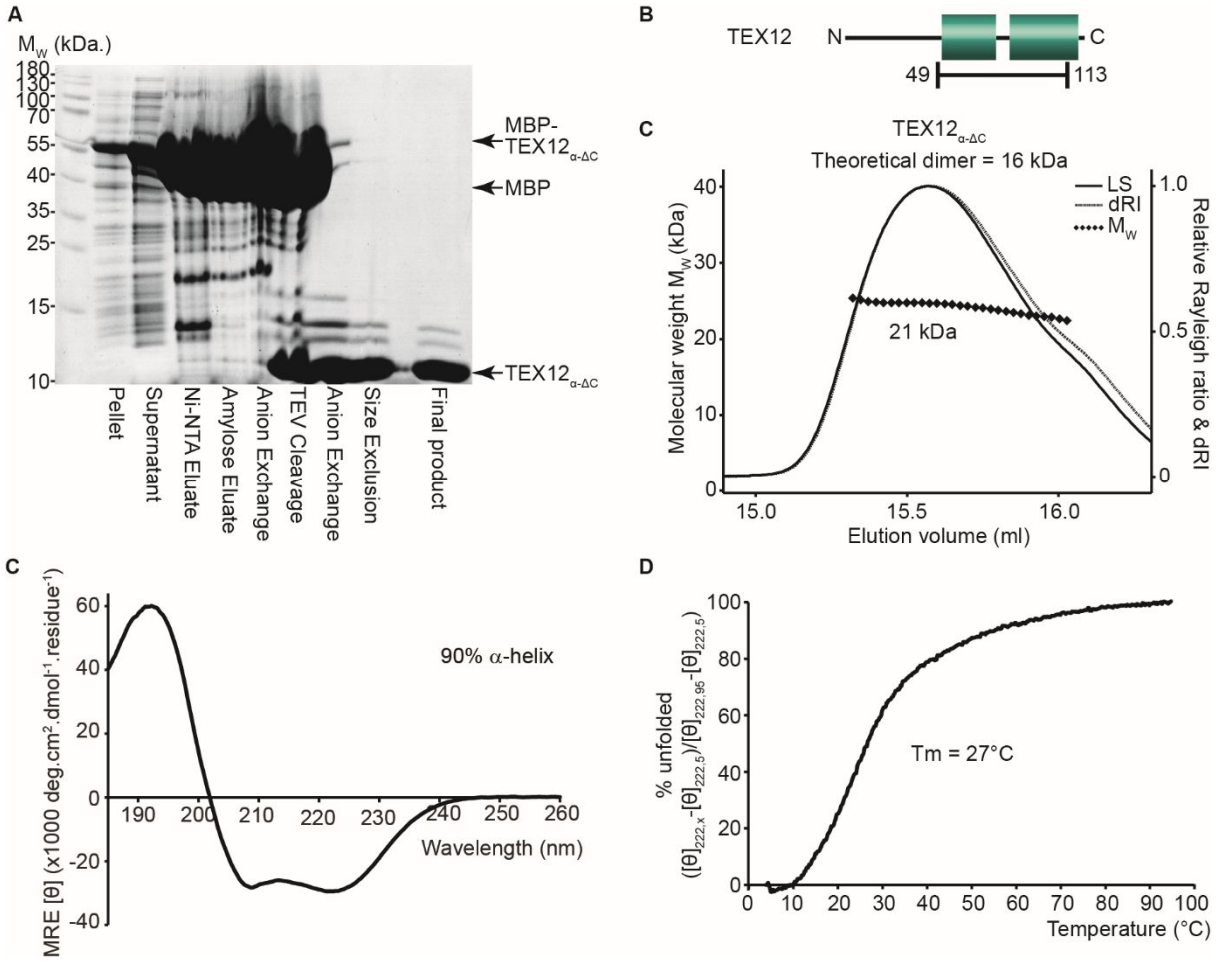


Figure 4.27. Purification and analysis of TEX12$_{\alpha\text{-}\Delta C}$. (A) SDS-PAGE summary of the TEX12$_{\alpha\text{-}\Delta C}$ purification by Ni-NTA, amylose, anion exchange chromatography steps, followed by TEV protease incubation and subsequent anion exchange and size exclusion chromatography. (B) Schematic of TEX12 boundaries. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 21 kDa, with the theoretical molecular weight of a dimer estimated to be 16 kDa. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 90% alpha-helix, 0% β-sheet, 2% turns and 8% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 27°C, when the sample is 50% unfolded.

To investigate the oligomeric state of the TEX12 core upon truncation, SEC-MALS was used to study TEX12$_{\alpha\text{-}\Delta C}$. This analysis revealed the formation of a single peak with a molecular mass of 21 kDa (Figure 4.27 C). The oligomeric state of TEX12$_{\alpha\text{-}\Delta C}$ is uncertain, with dissociation across the peak observed and therefore, a mixed species may be formed (Figure 4.27 C). TEX12$_{\alpha\text{-}\Delta C}$ was studied by SEC-SAXS to determine the structural parameters in comparison to TEX12$_{\alpha}$. The Rg was measured as 21 Å (Figure 4.28 A), which is larger than the 21 Å Rg determined for the dimeric TEX12$_{\alpha}$ and thus, indicating a less globular structure. The Rc was measured as 13 Å (Figure 4.28 B), which is very similar to the 12 Å Rc value of TEX12$_{\alpha}$.

The real space analysis indicates that TEX12$_{\alpha\text{-}\Delta C}$ forms an extended rod shape, with a maximum dimension of 110 Å (Figure 4.28 C). *Ab initio* modelling was performed to generate a low resolution molecular envelope of TEX12$_{\alpha\text{-}\Delta C}$, which highlights the elongated nature of TEX12$_{\alpha\text{-}\Delta C}$ (Figure 4.28 D). As the cross-sections of TEX12$_{\alpha}$ and TEX12$_{\alpha\text{-}\Delta C}$ are determined to be roughly the same size, but the maximum dimension of TEX12$_{\alpha\text{-}\Delta C}$ is approximately double than TEX12$_{\alpha}$, these data indicate that TEX12$_{\alpha\text{-}\Delta C}$ forms a larger oligomer than a dimer, potentially suggesting the formation of a tetramer that dissociates over the SEC column and thus, giving rise to the oligomeric ambiguity.

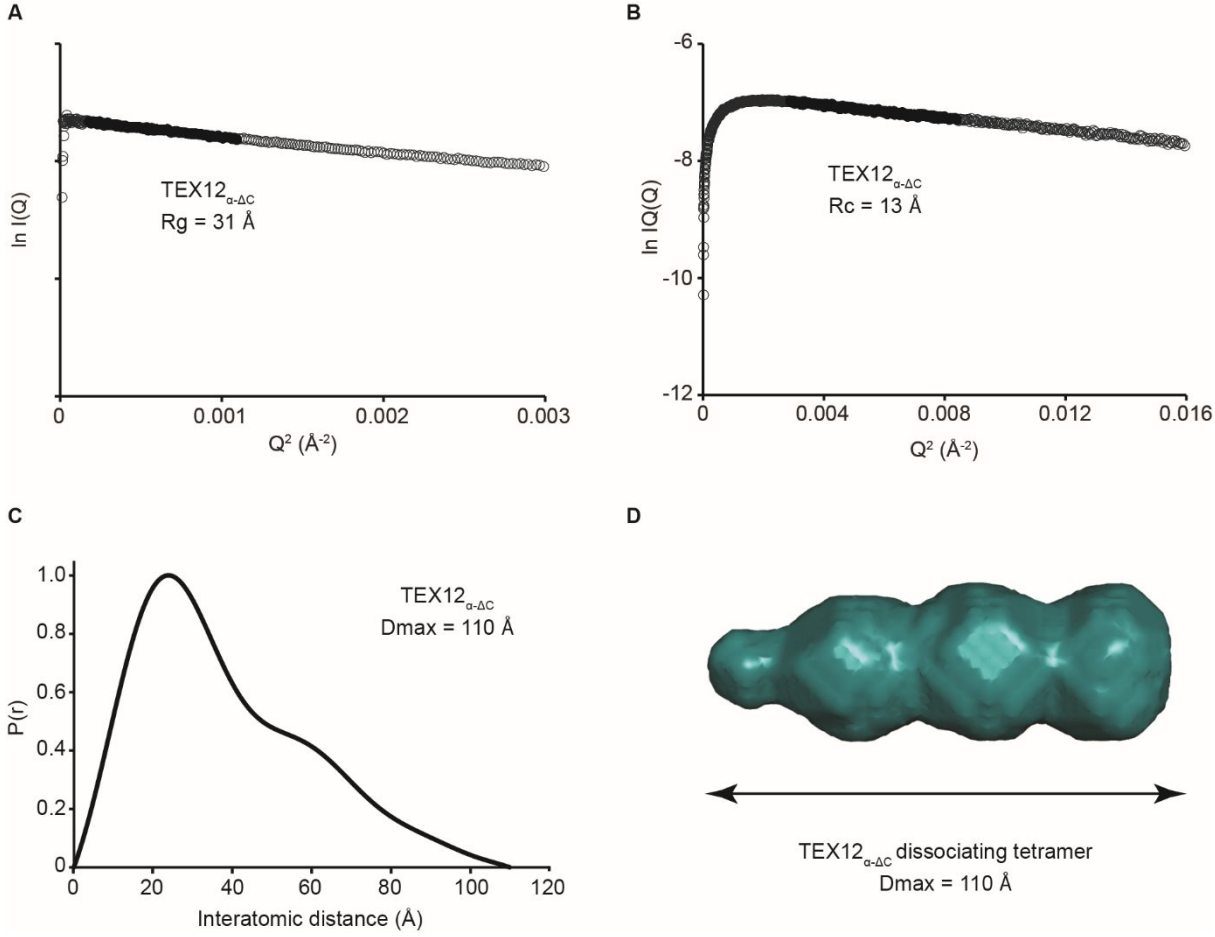


Figure 4.28. SEC-SAXS analysis of TEX12$_{\alpha\text{-}\Delta C}$. (A) Guinier analysis to determine the radius of gyration (Rg) as 31 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 13 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. TEX12$_{\alpha\text{-}\Delta C}$ has a maximum dimension of 110 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

To investigate the structural function of the TEX12 C-terminal residues, the TEX12 core was characterised with the introduction of the mutations as described previously within the SYCE2-TEX12 complex. To probe whether the C-terminal coiled-coil heptads are implicated in the formation of the compact dimer structure, the four candidate coiled-coil hydrophobic residues (L110, F114, I117 and L121) were simultaneously mutated to glutamic acid for analysis within the TEX12 core (Figure 4.29 A), hereby referred to as $TEX12_{\alpha\ cc\text{-}mut}$.

$TEX12_{\alpha\ cc\text{-}mut}$ was purified by the same method as $TEX12_{\alpha}$, with the cleaved complex purified by anion exchange chromatography following affinity tag removal by TEV protease (Figure 4.29 B). SEC-MALS analysis revealed a single protein peak, with a calculated molecular mass across the peak of 27 kDa (Figure 4.29 C). Similarly to $TEX12_{\alpha\text{-}\Delta C}$, dissociation was observed across the peak, with a molecular weight corresponding to a trimer.

$TEX12_{\alpha\ cc\text{-}mut}$ was analysed by SEC-SAXS to compare the structural parameters to $TEX12_{\alpha}$ and $TEX12_{\alpha\text{-}\Delta C}$. The $TEX12_{\alpha\ cc\text{-}mut}$ Rg was measured as 36 Å (Figure 4.29 D) and the Rc was determined to be 14 Å (Figure 4.29 E). Real space analysis was performed to produce the P(r) curve depicting the interatomic distances observed within the sample. The curve is indicative of an extended shape, with a maximum dimension of 125 Å (Figure 4.29 F). These values indicate that, similarly to $TEX12_{\alpha\text{-}\Delta C}$, an extended structure is formed. The cross-sectional radius values for $TEX12_{\alpha}$ and $TEX12_{\alpha\ cc\text{-}mut}$ are 12 Å and 14 Å, with the maximum lengths of 65 Å and 125 Å respectively. These parameters demonstrate that the elongated $TEX12_{\alpha\ cc\text{-}mut}$ is twice the length of the compact $TEX12_{\alpha}$ but with the same cross-section, which suggests that $TEX12_{\alpha\ cc\text{-}mut}$ may form a tetramer that mimics the overall structure of two $TEX12_{\alpha}$ dimers. Therefore, coiled-coil formation at the TEX12 C-terminus is seemingly required for the compact dimer structure, perhaps required for stabilisation of the folding back of the chains.

As mutations of the candidate surface exposed residues were found to stabilise the $SYCE2_{\alpha}$-$TEX12_{\alpha}$ complex, these mutations were analysed with the TEX12 core in the absence of SYCE2, hereby referred to as $TEX12_{\alpha\text{-}AEA}$ (Figure 4.30 A). This mutant was purified in the same way as $TEX12_{\alpha}$, with the cleaved $TEX12_{\alpha\text{-}AEA}$ material purified by anion exchange chromatography following tag removal (Figure 4.30 B). SEC-MALS analysis of $TEX12_{\alpha\text{-}AEA}$ revealed a single protein species with a calculated molecular mass of 33 kDa, which corresponds to a tetramer (Figure 4.30 C). In order to further dissect the role of these residues in altering the oligomeric properties of TEX12, a further series of mutants were analysed with a combination of the three mutations (Figure 4.30 A). SEC-MALS analysis revealed that upon introduction of the F109E and V116A mutations, the TEX12 core exists as a dimer, but with the single F102A point mutation, the TEX12 core forms the tetramer structure (Figure 4.30 E).

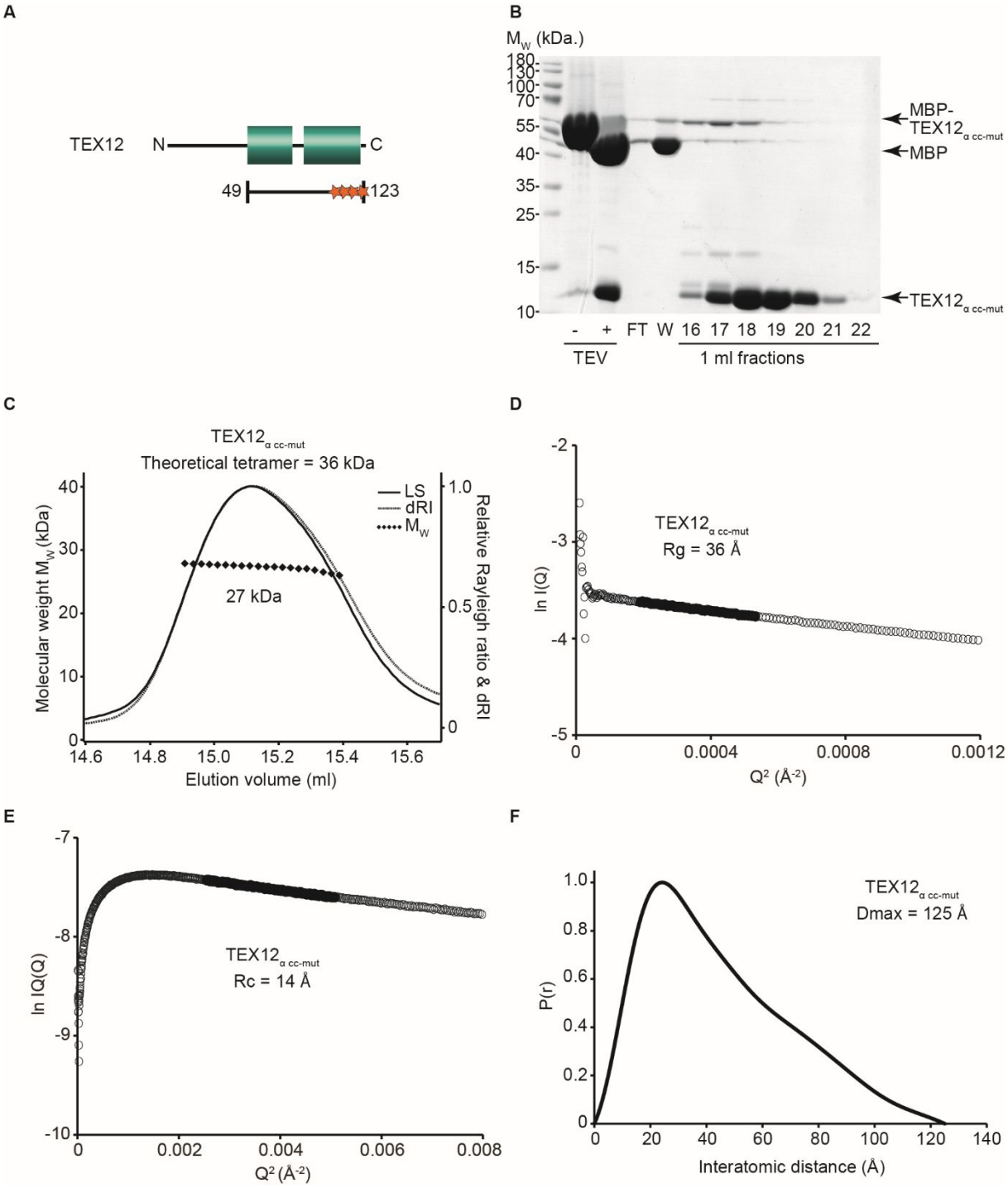


Figure 4.29. Purification and analysis of TEX12$_{\alpha\ cc\text{-}mut}$. (A) Schematic of TEX12 boundaries. Stars represent point mutations. (B) SDS-PAGE analysis of the TEX12$_{\alpha\ cc\text{-}mut}$ purification by anion exchange chromatography, following TEV protease incubation for tag removal. (C) SEC-MALS analysis. A single protein peak is formed with a calculated molecular mass of 27 kDa, with the theoretical molecular weight of a tetramer estimated to be 36 kDa. (D) Guinier analysis to determine the radius of gyration (Rg) as 36 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (E) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 14 Å. Clear circles are the data, solid circles represent the region used for the fit. (F) Paired real-space distribution plot. TEX12$_{\alpha\text{-}\Delta C}$ has a maximum dimension of 125 Å.

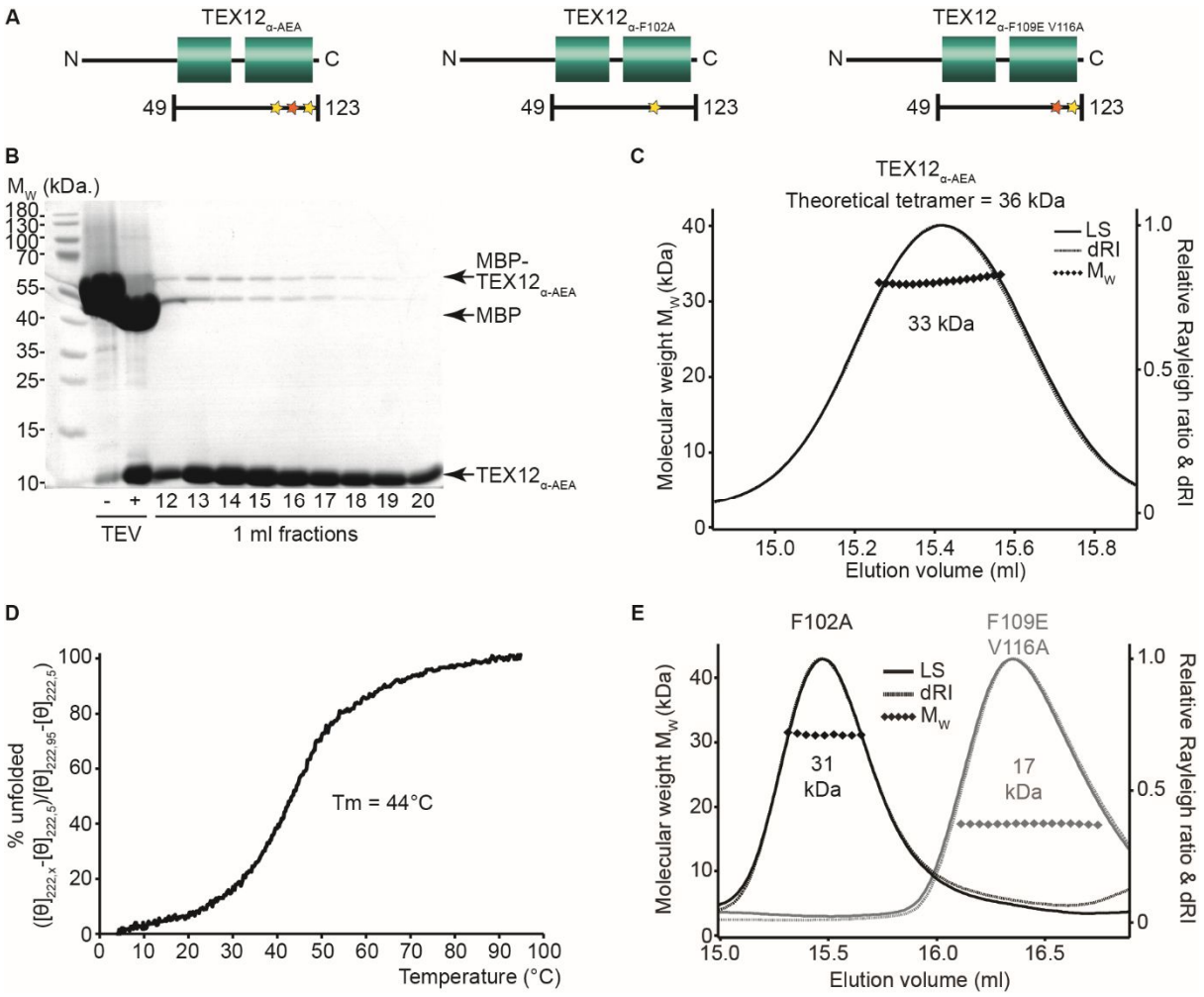


Figure 4.30. Purification and analysis of TEX12α mutants. (A) Schematics of TEX12 mutants. Stars represent point mutations. (B) SDS-PAGE analysis of the TEX12$_{\alpha\text{-AEA}}$ purification by anion exchange chromatography, following TEV protease incubation for tag removal. (C) SEC-MALS analysis of TEX12$_{\alpha\text{-AEA}}$. A single protein peak is formed with a calculated molecular mass of 33 kDa, with the theoretical molecular weight of a tetramer estimated to be 36 kDa. (D) CD thermal denaturation of TEX12$_{\alpha\text{-AEA}}$ between 5°C and 95°C, plotted as % unfolded. Tm estimated as 44°C, when the sample is 50% unfolded. (E) Overlaid MALS analysis of TEX12α harbouring the F102A point mutation or the double F109E V116A mutation. Single protein peaks were formed with calculated molecular masses of 31 kDa and 17 kDa respectively, with the theoretical molecular weight of 18 kDa for a dimer.

Therefore, the F102 residue is critical for the compact dimer form. Although the F102A mutation is solely responsible for the oligomeric switch, the F109E and V116A mutations fortuitously stabilise the F102A tetramer to obtain a higher protein yield for analysis.

The thermal stability of the TEX12$_{\alpha\text{-AEA}}$ tetramer was investigated by CD thermal denaturation by recording protein unfolding across a temperature gradient, using the 222 nm helical signal. The melting curve shows a single cooperative unfolding event, with an estimated melting temperature of 44°C (Figure 4.30 D). The melting temperature for TEX12$_{\alpha\text{-AEA}}$ lies between the 57°C and 27°C melting temperatures determined for TEX12α and TEX12$_{\alpha\text{-}\Delta C}$.

SEC-SAXS experiments were performed with TEX12$_{\alpha\text{-AEA}}$ to compare the structural parameters of the tetramer to the TEX12$_{\alpha}$ dimer. The TEX12$_{\alpha\text{-AEA}}$ Rg was determined as 32 Å (Figure 4.31 A), and the Rc was measured as 13 Å (Figure 4.31 B). Real space analysis was performed to produce the P(r) curve to reveal the interatomic distances observed within the sample. The distance distribution was indicative of an extended shape with the maximum dimension estimated as 120 Å (Figure 4.31 C). Ab initio modelling was performed to generate the TEX12$_{\alpha\text{-AEA}}$ molecular envelope, and the resultant model indeed showed an extended rod structure (Figure 4.31 D). Similarly to TEX12$_{\alpha\text{ cc-mut}}$, the cross-sectional radius of the TEX12$_{\alpha\text{-AEA}}$ is approximately the same size as the value for the TEX12$_{\alpha}$ dimer, with the tetramer length double that of the dimer length.

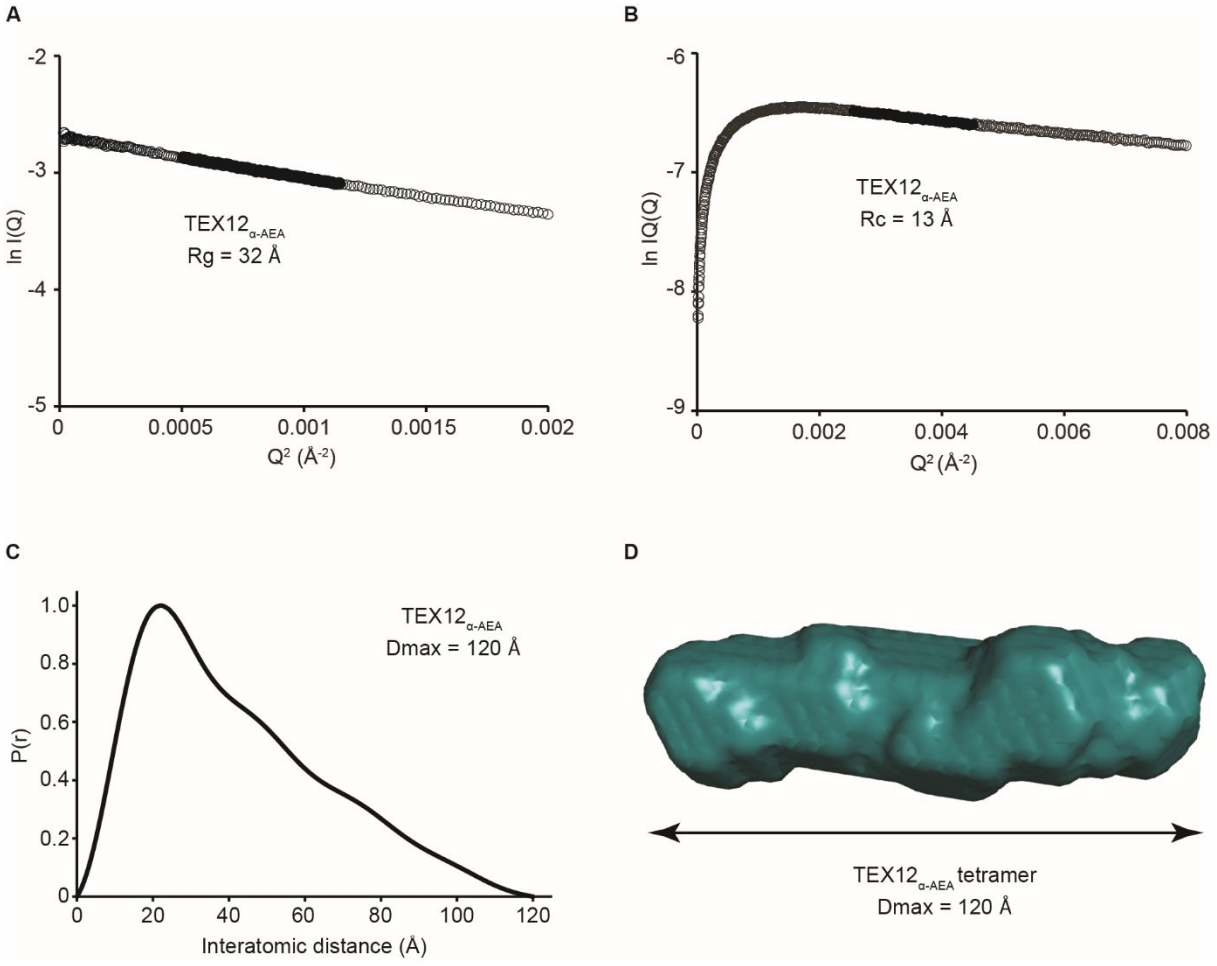


Figure 4.31. SEC-SAXS analysis of the TEX12$_{\alpha\text{-AEA}}$ tetramer. (A) Guinier analysis to determine the radius of gyration (Rg) as 32 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 13 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. TEX12$_{\alpha\text{-AEA}}$ has a maximum dimension of 120 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

Construct	TEX12	$TEX12_{\alpha}$	$TEX12_{\alpha\text{-}\Delta C}$	$TEX12_{\alpha\ cc\text{-}mut}$	$TEX12_{\alpha\text{-}AEA}$
TEX12 region	1-123	49-123	49-113	49-123	49-123
Mutations	N/A	N/A	N/A	L110E F114E I117E L121E	F102A F109E V116A
% α-helix	45	80	90	N.D.	N.D.
Tm (°C)	60	57	27	N.D.	44
Oligomer state	Dimer	Dimer	Dissociating tetramer	Dissociating tetramer (trimer)	Tetramer
Rg (Å)	30	21	31	36	32
Rc (Å)	N.D.	12	13	14	13
Dmax (Å)	110	65	110	125	120

Table 4.4. Biophysical analysis summary of TEX12 constructs

As the dimeric structure seemingly requires folding back of the TEX12 chains, it is likely that a helical break occurs in the central region of $TEX12_{\alpha}$, in order to provide a turn region to form the compact arrangement. To test this, two mutants were designed whereby flexible linker regions were inserted in between TEX12 residues to characterise whether the dimer structure is formed upon the imposed helical breaks (Figure 4.32 A). Five stretches of threonine-serine-glycine repeats were inserted between the L75 and S76 residues to provide the flexible linker, hereby referred to as $TEX12_{\alpha\text{-}75}$, and between the R78 and A79 residues, hereby referred to as $TEX12_{\alpha\text{-}78}$, and expressed as fusion proteins to a His_6-MBP affinity tag.

Both constructs were purified by Ni-NTA and amylose affinity chromatography steps, followed by anion exchange chromatography. The His-MBP-$TEX12_{\alpha\text{-}75}$ fusion protein exhibited severe degradation of the either the full $TEX12_{\alpha\text{-}75}$ sequence or partial degradation from the C-terminal end, with only a small proportion of the material existing as the full length sequence (Figure 4.32 B). SEC-MALS analysis of the full length material produced multiple species, likely due to presence of degradation products, with the majority of the material forming a 55 kDa species that corresponds to a monomer (Figure 4.32 C).

The purification of the His-MBP-$TEX12_{\alpha\text{-}78}$ fusion protein revealed that the majority of the material corresponded to the full length sequence (Figure 4.32 D), demonstrating a more stable fusion protein than the $TEX12_{\alpha\text{-}75}$ linker. SEC-MALS analysis of His-MBP-$TEX12_{\alpha\text{-}78}$ revealed a single protein species with a molecular weight of 92 kDa, correlating to a dimer (Figure 4.32 E). These findings indicate that the linker insertion after the L75 residue destabilises the TEX12 chain and prevents dimerisation, whereas the linker insertion after the R78 residue allows the folding back of the TEX12 and subsequent dimerisation.

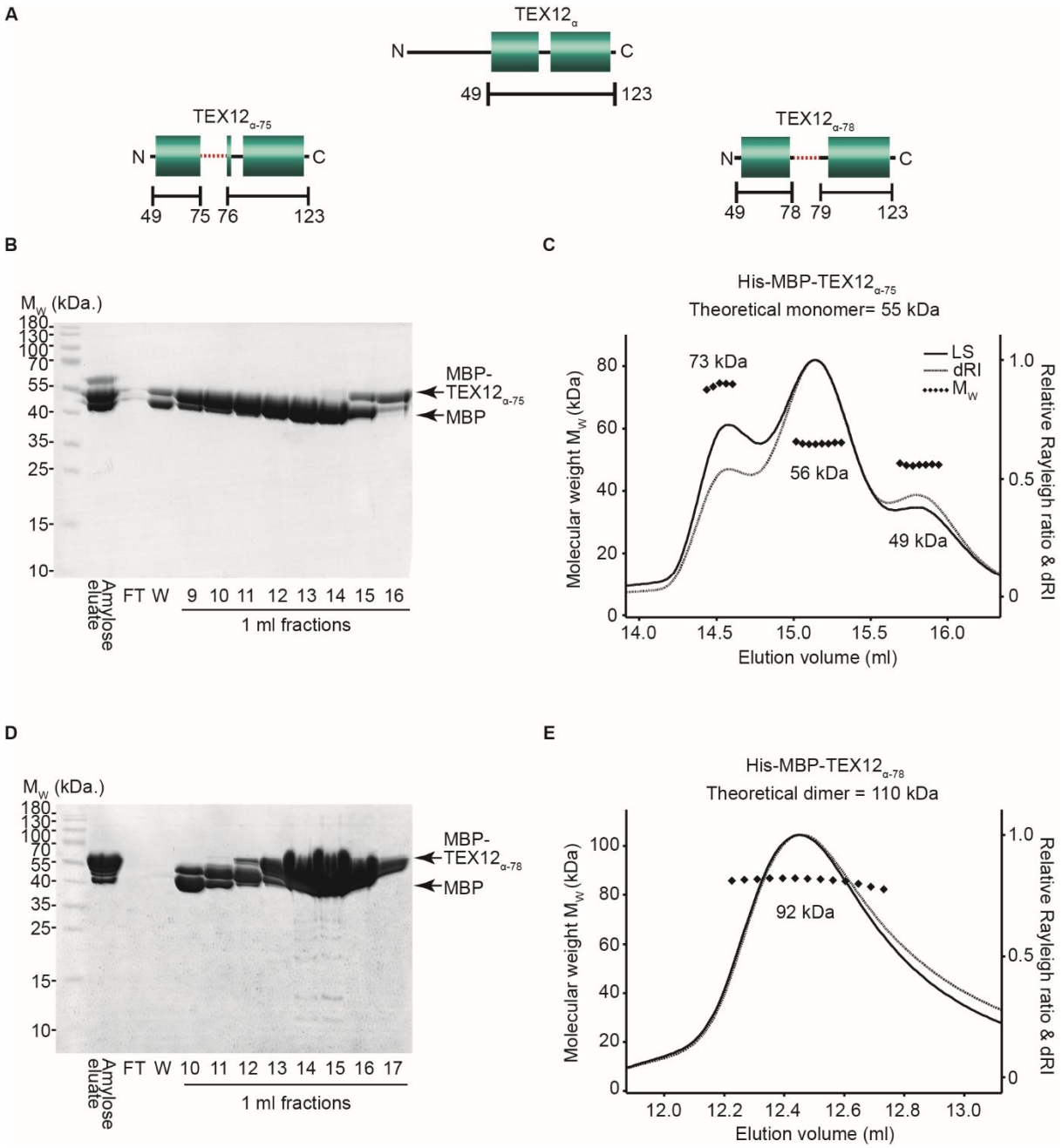


Figure 4.32. Purification and analysis of TEX12$_{\alpha}$ linker insertions. (A) Schematics of TEX12 constructs. Red dotted line represents the linker insertion. (B) SDS-PAGE analysis of the TEX12$_{\alpha\text{-}75}$ purification by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (C) SEC-MALS analysis of TEX12$_{\alpha\text{-}75}$. Multiple protein species are formed with calculated molecular masses of 73 kDa, 56 kDa and 49 kDa, with the theoretical molecular weight of a monomer estimated to be 55 kDa. (D) SDS-PAGE analysis of the TEX12$_{\alpha\text{-}78}$ purification by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (E) SEC-MALS analysis of TEX12$_{\alpha\text{-}78}$. A single protein peak is formed with a calculated molecular masses of 92 kDa, with the theoretical molecular weight of a dimer estimated to be 110 kDa.

The biophysical analysis of TEX12, and the various mutants, revealed that TEX12 forms a compact dimer in the absence of SYCE2, and that coiled-coil formation is likely required to stabilise the folding back of the TEX12 chains. The removal of these potential coiled-coil interactions, either by truncation or mutation, results in the formation of an unstable extended structure, with mutation of the candidate surface exposed patch resulting in the formation of a stable elongated tetramer (Figure 4.33). Therefore, the coiled-coil interactions are likely required for both the dimer and tetramer structure, with the F102A mutation causing the switch between the oligomeric states to stabilise the tetramer structure. From the preliminary analysis of the TEX12 linker insertions, it is likely that the folding back of the TEX12 chains in order to form the compact dimer structure occurs at a helical break after the TEX12 R78 residue.

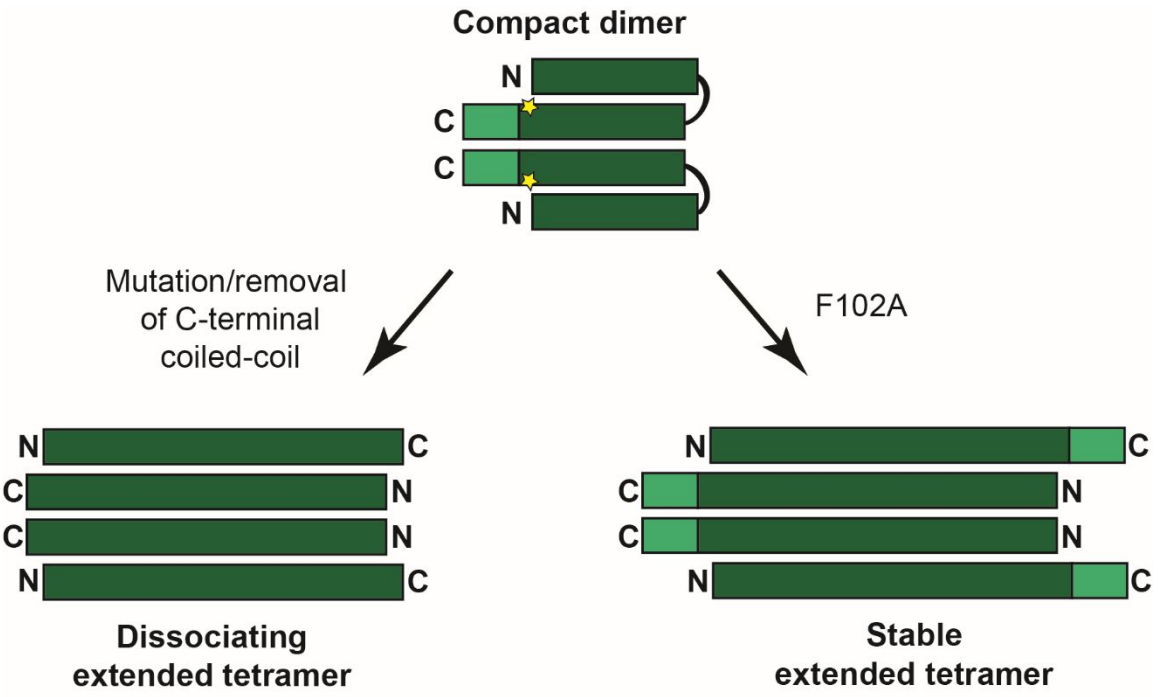


Figure 4.33. TEX12 schematic models to summarise biophysical analysis. The WT protein forms a compact dimer through folding back of the chains. The truncation or mutation of the C-terminal coiled-coil heptads forms an unstable extended tetramer. The mutation of the F102 residue (yellow star) forms a stable anti-parallel extended tetramer in solution.

4.2.6 Crystal structures of TEX12

As the TEX12$_{\alpha}$ dimer did not crystallise, further crystallisation trials were performed with the TEX12 mutants. Commercial screening was performed with TEX12$_{\alpha\text{-}\Delta C}$, with one condition enabling crystal growth (Figure 4.34 A). These crystals were harvested to test for X-ray diffraction, whereby protein diffraction was detected. However, the quality of the diffraction was not sufficient for structure solution and so, crystal optimisation was carried out in order to improve crystal quality and size. From a wide screening attempt, the crystals were not reproduced. Further monitoring revealed that crystallisation occurred two hours subsequent to drop formation, with rapid deterioration beginning four hours after drop formation, and the crystals were completely dissolved after 24 hours. After this finding, larger optimised crystals were harvested 3 hours after drop formation (Figure 4.34 B), with a variety of cryo-protectants tested for screening of the crystals for improvement in X-ray diffraction. The X-ray diffraction pattern from an optimised crystal was improved to allow for structure solution. Therefore, the X-ray crystal structure of TEX12$_{\alpha\text{-}\Delta C}$ was solved by a fragment-based *ab initio* molecular replacement approach (structure solution performed by Owen Davies).

The structure reveals the formation of an anti-parallel tetrameric bundle (Figure 4.34 C). Further dissection of the structure shows the formation of anti-parallel dimeric coiled-coils, with a break in the coiled-coil formation at the central region of the structure (Figure 4.34 D). This region corresponds to the potential helical break position that takes place in the solution dimer, predicted to occur at the R78 residue that is located in the central region (Figure 4.34 D). Therefore, the dimeric solution structure potentially mimics the helical arrangement as observed in each half of the crystal structure, with the dimeric compact structure formed through folding back of the chains by the six amino acid "insertion" region.

Crystallisation screening and data collection was also performed with TEX12$_{\alpha\text{-}AEA}$, by another member of the Davies group, James Dunce. The X-ray structure was solved by the fragment-based *ab initio* molecular replacement method (structure solution performed by Owen Davies). The TEX12$_{\alpha\text{-}AEA}$ crystal structure revealed the formation of an anti-parallel dimer, in which the C-termini protrude to produce an overhang at each end of the central dimeric region (Figure 4.35 A). The protruding C-termini were found to form crystallographic interactions to form a "dimer of dimers" (Figure 4.35 B). The C-terminal region also undergoes further crystallographic associations to form tetrameric assemblies, in order to produce a lattice whereby a longitudinal "dimer of dimers" and lateral tetrameric associations occur through the C-terminal protrusions (Figure 4.35 C). The tetrameric assembly is driven by the previously predicted hydrophobic coiled-coil residues (see above), where two interfaces are each formed from anti-parallel chains, with the formation of a continuous interface and an

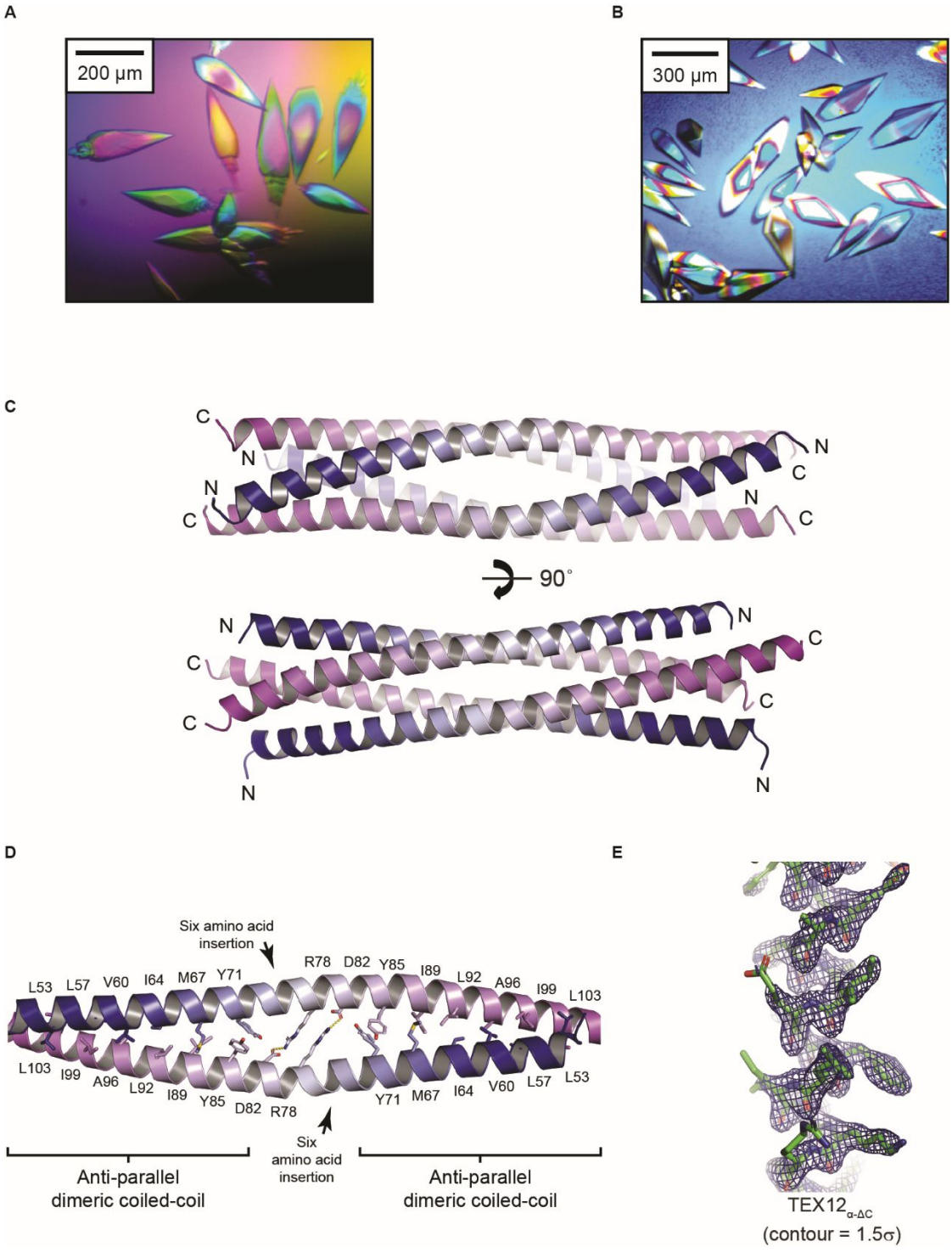


Figure 4.34. Crystal structure of TEX12$_{\alpha\text{-}\Delta C}$. (A) TEX12$_{\alpha\text{-}\Delta C}$ crystals grown in 35% v/v 1,4-dioxane from initial commercial screening. (B) Optimised crystals grown in 25% v/v 1,4-dioxane. (C-E_ Figures prepared by Owen Davies. (C) The X-ray crystal structure of TEX12$_{\alpha\text{-}\Delta C}$ showing the formation of an anti-parallel tetrameric helical bundle in two views. (D) The central region of the structure, showing electrostatic interactions between the R78 and D82 residues from opposing chains. Each half of the structure contains anti-parallel dimeric coiled-coils with the central region breaking the coiled-coil with a six amino acid insertion. (E) The electron density map of the C-terminal portion of TEX12$_{\alpha\text{-}\Delta C}$.

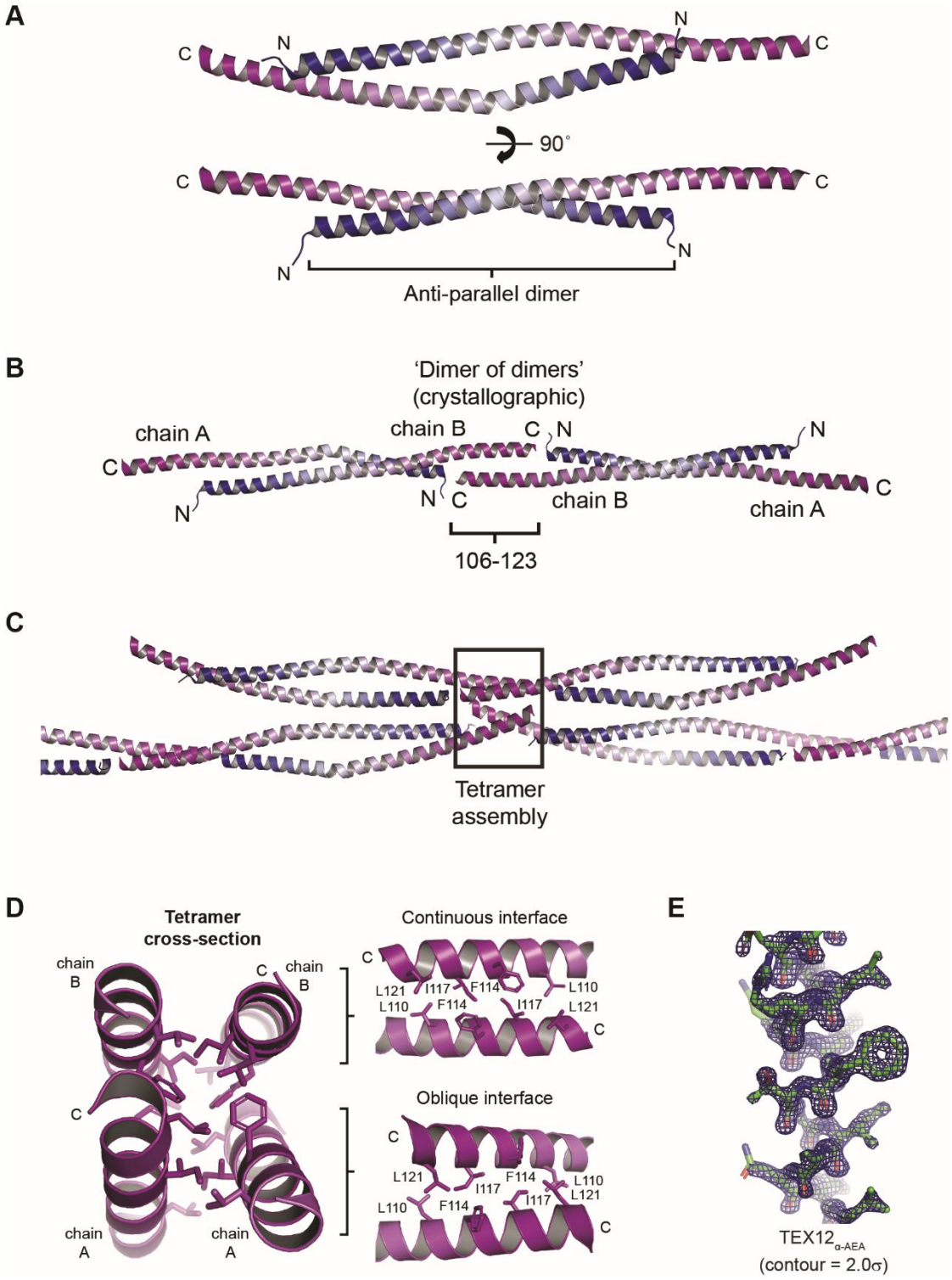


Figure 4.35. Crystal structure of TEX12$_{\alpha\text{-AEA}}$. Figures prepared by Owen Davies. (A) The X-ray crystal structure of TEX12$_{\alpha\text{-AEA}}$ showing the formation of an anti-parallel dimer with the C-terminal region protruding from the central dimer. Displayed in two views. (B) The C-terminal protrusion, residues 106-123, forms a crystallographic "dimer of dimers" association from neighbouring dimers. (C) The 106-123 amino acid protrusion also forms tetrameric assemblies in the crystal lattice. (D) Cross-section of the C-terminus. The hydrophobic coiled-coil residues, L110, F114, I117 and L121, drive the tetrameric assemblies with a continuous interface formed by two chains and an oblique interface formed by the other two chains. (E) The electron density map of the C-terminal portion of TEX12$_{\alpha\text{-AEA}}$.

	TEX12$_{\alpha\text{-}\Delta C}$ (PDB 6HK8)	TEX12$_{\alpha\text{-}AEA}$ (PDB 6HK9)
Data collection		
Space group	$P6_522$	$C222_1$
Cell dimensions		
a, *b*, *c* (Å)	47.97, 47.97, 210.98	43.23, 219.71, 37.50
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 90.00
Resolution (Å)	52.75 – 2.11 (2.15 – 2.11)*	54.93 – 1.45 (1.48 – 1.45)*
R_{meas}	0.082 (2.330)	0.044 (0.472)
R_{pim}	0.014 (0.374)	0.017 (0.170)
$I / \sigma(I)$	28.1 (2.4)	20.3 (2.6)
$CC_{1/2}$	1.000 (0.890)	0.999 (0.993)
Completeness (%)	92.7 (100.0)	98.1 (96.6)
Redundancy	34.5 (37.5)	7.1 (7.4)
Refinement		
Resolution (Å)	41.54 – 2.11	30.97 – 1.45
UCLA anisotropy (Å)	2.2, 2.2, 2.1	-
No. reflections	7701	59176
R_{work} / R_{free}	0.2291/0.2580	0.1795/0.2047
No. atoms	1146	1536
Protein	1056	1313
Ligand/ion	30	23
Water	60	200
B-factors	55.63	37.67
Protein	54.99	35.55
Ligand/ion	73.83	59.85
Water	57.82	49.06
R.m.s. deviations		
Bond lengths (Å)	0.002	0.008
Bond angles (°)	0.316	0.856

*Values in parentheses are for highest-resolution shell.

Table 4.5. Data collection, phasing and refinement statistics of the TEX12 structures.

oblique interface through angulated anti-parallel chains (Figure 4.35 D). The two TEX12 structures reveal that the oligomeric region is formed by anti-parallel chains, spanning residues 49-105, to form dimeric and tetrameric structures. By modelling the $TEX12_{\alpha\text{-}AEA}$ dimer structure, a tetramer model was built that mimics the same structure as the $TEX12_{\alpha\text{-}\Delta C}$ four-helical bundle, with protrusion of the C-terminal residues forming the coiled-coil interactions between parallel chains (Figure 4.36 A). This modelled tetramer forms approximately the same shape as its solution tetramer (Figure 4.36 C). As the $TEX12_{\alpha}$ helical core forms a dimer in solution, it was postulated that the solution dimer is formed by folding back of the TEX12 chains to mimic the four-helical bundle with the same organisation as observed in each half of the tetramer structure. The six amino acid insertion region likely forms a loop to allow the formation of the compact dimer (Figure 4.36 B), which correlates to the experimental solution dimer (Figure 4.36 D). The $TEX12_{\alpha\text{-}AEA}$ tetramer model and the $TEX12_{\alpha}$ compact dimer model are stabilised by the four hydrophobic coiled-coil residues that are observed in the $TEX12_{\alpha\text{-}AEA}$ crystal structure (Figure 4.36 E and F). The mutation of these residues in solution destabilised both the dimer and tetramer structures, and therefore, the C-terminal coiled-coil heptads have been shown to be critical for TEX12 structure.

In the $TEX12_{\alpha\text{-}\Delta C}$ structure, the F102 residue is located at the end of the helical bundle, with its orientation potentially forming an interaction site for the folding back of the chain in the dimer structure. The F102A mutation therefore removes this site, which destabilises the compact dimer structure and encourages the tetramer formation. When in complex with SYCE2, $TEX12_{\alpha\text{-}\Delta C}$ forms a 2:2 complex and $TEX12_{\alpha\text{-}AEA}$ forms a 4:4 complex. Therefore, it suggests that TEX12 may form a tetrameric assembly through the C-terminal coiled-coil in order to form the 4:4 complex.

4.2.7 Structural organisation of SYCE2-TEX12 units and fibrous assemblies

The study of TEX12 in the absence of SYCE2 has revealed the formation of an anti-parallel structure that is stabilised by the TEX12 C-terminus and provides a potential mechanism for larger assemblies through longitudinal and lateral interactions. However, it remains unclear as to the conformation of TEX12 when bound to SYCE2, with little information known about the conformation of SYCE2 within the complex. In the absence of a high resolution structure of the SYCE2-TEX12 complex, crosslinking was performed with the $SYCE2_{\alpha}$-$TEX12_{\alpha\text{-}\Delta C}$ 2:2 complex in order to elucidate potential interactions between the chains through mass spectrometry. The $SYCE2_{\alpha}$-$TEX12_{\alpha\text{-}\Delta C}$ complex was crosslinked with bissulfosuccinimidyl suberate (BS3) at increasing concentrations of cross-linker, followed by

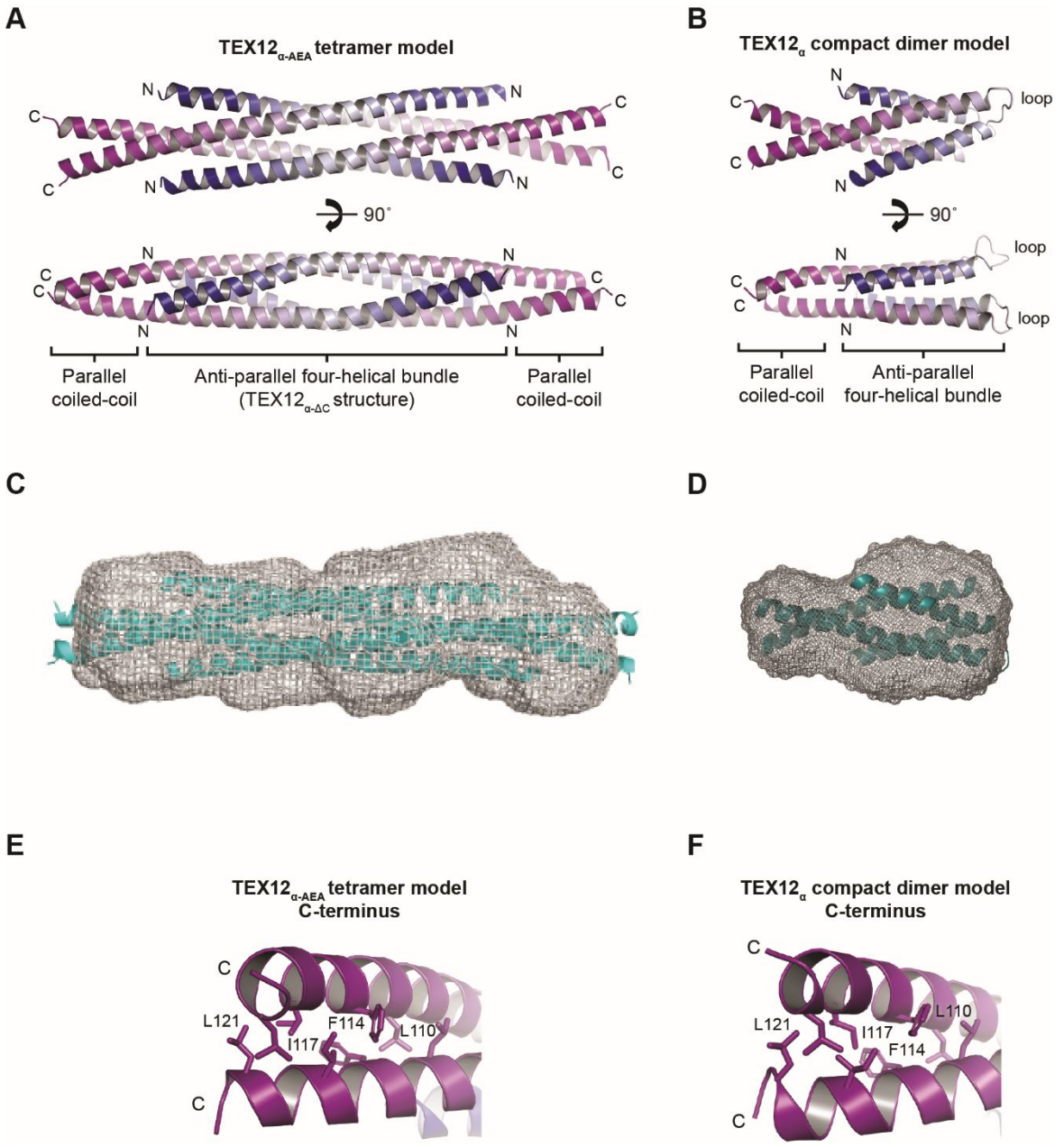


Figure 4.36. Modelling of the TEX12 crystal structures. (A) Modelling of the TEX12$_{\alpha\text{-AEA}}$ structure to form a four-helical bundle that mimics the TEX12$_{\alpha\text{-}\Delta C}$ tetramer structure, with the inclusion of the C-terminal protrusions forming parallel coiled-coils. Displayed in two views. (B) Modelling of the solution TEX12 compact dimer from the TEX12$_{\alpha\text{-AEA}}$ structure. An anti-parallel four-helical bundle may be formed that mimics the organisation of the four-helical bundle observed in each half of the tetramer structure. The protruding C-termini then form coiled-coil interactions to stabilise the compact dimer. (C) The modelled tetramer docked into the experimental TEX12$_{\alpha\text{-AEA}}$ tetramer SAXS *ab initio* model. (D) The modelled dimer docked into the experimental TEX12$_{\alpha}$ dimer SAXS *ab initio* model. (E) The TEX12 C-termini from the modelling of the TEX12$_{\alpha\text{-AEA}}$ modelled tetramer. (F) The TEX12 C-termini from the modelling of the TEX12$_{\alpha}$ modelled compact dimer.

SDS-PAGE in order to separate the cross-linked species (Figure 4.37 A). Bands were then excised from the gel for trypsin digestion and subsequent mass spectrometry analysis (performed by Juan Zou, Juri Rappsilber group, University of Edinburgh). Through analysis of the 1:1 SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ band, a prominent cross-link was identified within TEX12 between residues K55 and K106 (Figure 4.37 B). This link correlates to the TEX12 crystal structures, whereby the K55 and K106 residues from anti-parallel chains are located in close proximity at the ends of the tetrameric bundles. As this occurs in the 1:1 species, this insinuates that the conformation of TEX12$_{\alpha\text{-}\Delta C}$ within the 2:2 complex is in the compact form. Analysis of the SYCE2 dimer species and the 2:1 SYCE2-TEX12 species demonstrated an array of cross-links, with no cross-links detected within the SYCE2 N-terminal region and several cross-links detected between the glycine at the N-terminus, which is residual from the cleavage of the affinity tag for purification, and the latter half of the SYCE2 chain (Figure 4.38 A and B). This suggests that the SYCE2 dimer is in an anti-parallel arrangement within the complex, as well as the TEX12 chains. Therefore, the cross-linking analysis suggests that the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex is formed from anti-parallel SYCE2 and TEX12 dimers to form a four-helical bundle.

The crystal structures of TEX12 demonstrate the anti-parallel assembly of the TEX12 chains. However, the orientation of SYCE2 was unclear through solution experiments, with the cross-linking data suggesting SYCE2 forms an anti-parallel dimer in the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex. To test the anti-parallel nature of the SYCE2 chains in solution, a construct was designed whereby a stretch of five threonine-glycine-serine repeats were used to link two SYCE2 chains, with the C-terminus of one SYCE2 molecule fused to the N-terminus of a second molecule via the linker. The tethered SYCE2 dimer was co-expressed with TEX12$_{\alpha}$, hereby referred to as SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha,}$ and the complex was purified by the same procedure as SYCE2$_{\alpha}$-TEX12$_{\alpha}$ (see above), whereby anion exchange chromatography was used to purify the cleaved complex subsequent to affinity tag removal by TEV protease incubation (Figure 4.39 A). The oligomeric state of the SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$ complex was studied by SEC-MALS for comparison to SYCE2$_{\alpha}$-TEX12$_{\alpha}$. This analysis showed that the majority of the material formed a peak with a calculated molecular weight of 94 kDa (Figure 4.39 B), corresponding to a 2:4 complex, which mimics the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ 4:4 complex. Some of the material was also found to undergo higher-order assembly, suggesting that the SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$ complex behaves in the same manner as SYCE2$_{\alpha}$-TEX12$_{\alpha}$, and therefore, that the SYCE2 chains are in an anti-parallel arrangement, as suggested by the crosslinking mass spectrometry analysis.

To validate that the 2:4 SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$ forms the same structure as the SYCE2$_{\alpha}$-TEX12$_{\alpha}$ 4:4 complex, SEC-SAXS experiments were performed. The Rg was measured as 53 Å (Figure 4.39 C) and the Rc was determined as 19 Å (Figure 4.39 D). Real space analysis of

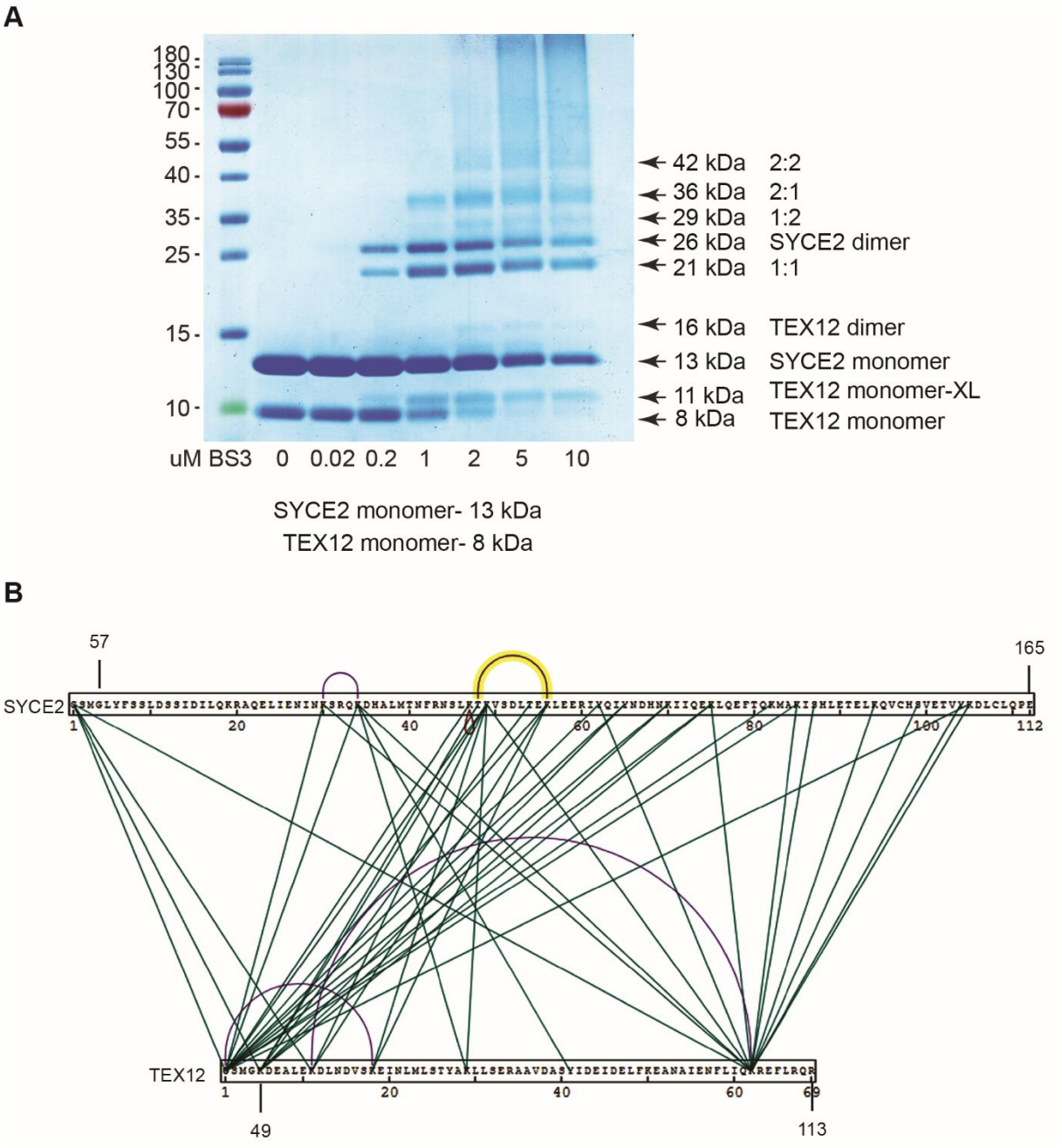


Figure 4.37. Crosslinking of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex. (A) SDS-PAGE analysis of crosslinked SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex, at 1.7 mg/ml, by bissulfosuccinimidyl suberate (BS3) at increasing concentrations. The crosslinking reaction at 1 µM BS3 was used for analysis, with the bands corresponding to a 1:1 complex, an SYCE2 dimer and a 2:1 SYCE2-TEX12 complex excised for analysis. (B) Crosslinking data for the 1:1 complex band. The data shows identified crosslinks within the sample, with both intra- (purple) and inter-(green) crosslinks. Data provided by Juan Zou, Juri Rappsilber group, University of Edinburgh.

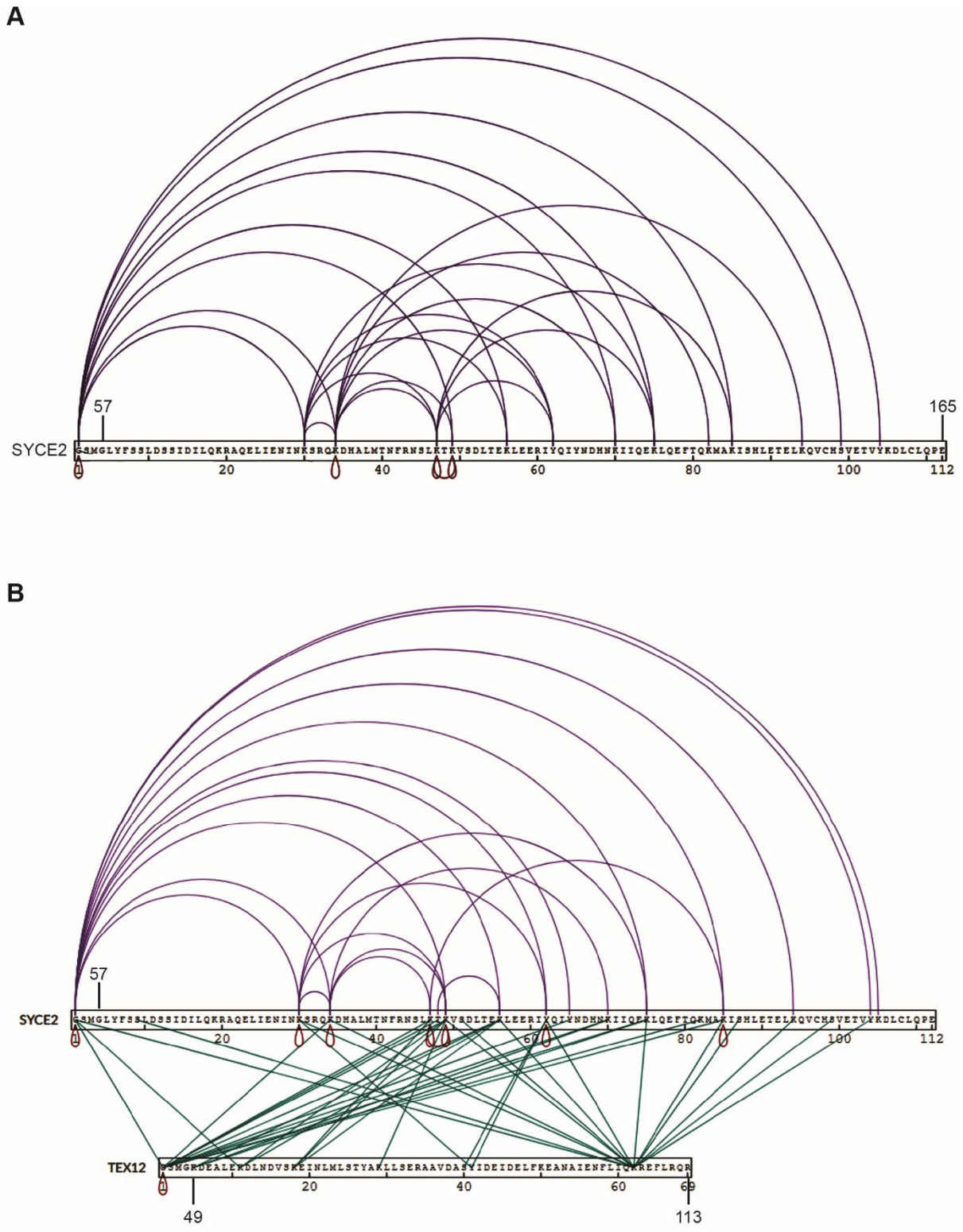


Figure 4.38. Crosslinking of the SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C}$ 2:2 complex- band analysis. (A) Crosslinking data for the SYCE2 dimer band. The data shows identified intra-SYCE2 crosslinks (purple). (B) Crosslinking data for the 2:1 complex band. The data shows identified crosslinks within the sample, with both intra- (purple) and inter-(green) crosslinks. Data provided by Juan Zou, Juri Rappsilber group, University of Edinburgh.

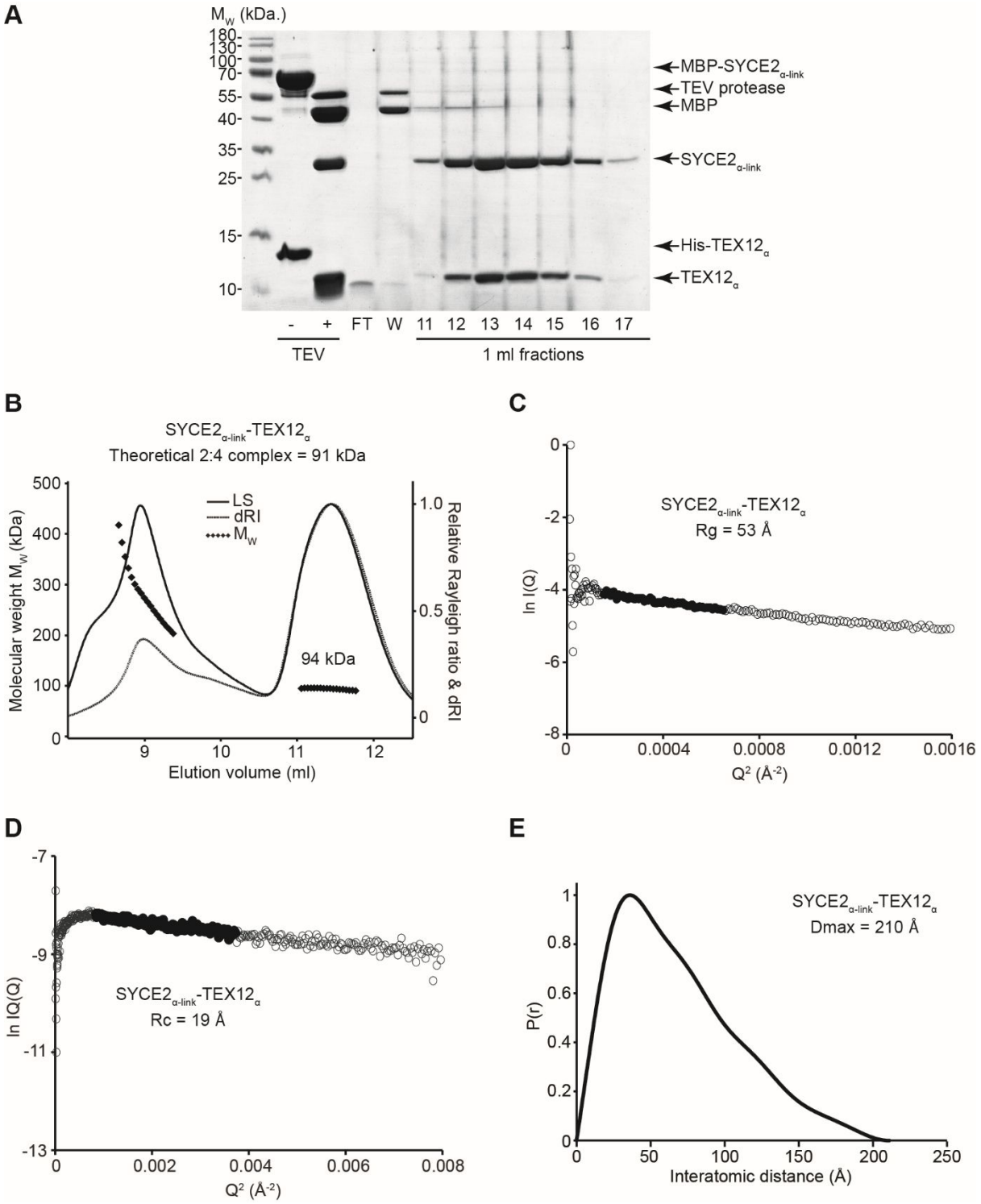


Figure 4.39. Purification and analysis of SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$. (A) SDS-PAGE analysis of the SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$ purification by anion exchange chromatography, following TEV protease incubation for tag removal. (B) SEC-MALS analysis. A predominant protein peak is formed with a calculated molecular mass of 94 kDa, with the theoretical molecular weight of a 4:4 complex estimated to be 94 kDa. Large molecular weight species are also formed. (C) Guinier analysis to determine the radius of gyration (Rg) as 53 Å. Clear circles represent the data, solid circles represent the region used for the Guinier fit. (D) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 19 Å. Clear circles are the data, solid circles represent the region used for the fit. (E) Paired real-space distribution plot. SYCE2$_{\alpha\text{-link}}$-TEX12$_{\alpha}$ has a maximum dimension of 210 Å.

the interatomic distances within the sample demonstrate an extended structure with an estimated maximum dimension of 210 Å (Figure 4.39 E). These structural parameters match those of the $SYCE2_{\alpha}$-$TEX12_{\alpha}$ 4:4 complex, and therefore, suggest that the same 4:4 unit is formed and subsequently that the SYCE2 chains are in an antiparallel orientation.

4.3 Discussion

4.3.1 TEX12 is stable in the absence of SYCE2

SYCE2 and TEX12 were proposed to exist in an equimolar constitutive complex that undergoes higher-order assembly (Davies *et al*., 2012). In this study, TEX12 was found to be stable in the absence of SYCE2. The full length protein, and the structural core, were found to exist as a compact dimeric structure, with a proposed model for its organisation based upon solution scattering and modelling based upon the mutant TEX12 structures. The model suggests that parallel TEX12 chains fold back to form a four helical bundle structure, with the C-termini protruding from the four helical assembly (Figure 4.40 A).

The truncation or mutation of the protruding C-terminal region caused TEX12 to form an extended tetramer structure, which is found to dissociate in solution. The X-ray crystal structure of the truncated TEX12 revealed the formation of an anti-parallel four helical bundle. Therefore, the TEX12 C-terminus is predicted to form coiled-coil interactions that are required to form the compact dimer structure (Figure 4.40 A).

The analysis of further TEX12 C-terminal mutations revealed that the F102 residue is required for the compact dimer structure, potentially by forming a key interaction site within the folded back chains, and so a point mutation of this residue resulted in the formation of an extended structure. The mutation formed an elongated antiparallel dimeric structure in the X-ray crystal structure, and formed an extended tetramer in solution (Figure 4.40 A). Therefore, two modalities of the TEX12 C-terminus were characterised, in which coiled-coil reactions are required to stabilise both the dimer and tetramer structure, with the F102 residue responsible for stabilising the compact dimer.

The core oligomerisation region from both mutant TEX12 structures were formed by the same region, with the C-termini from the extended dimer structure protruding from this region. From analysis of the crystal structures, each half of the central oligomerisation region forms anti-parallel coiled-coils, with a break in the centre of the oligomeristion region formed by a six amino acid insertion. This insertion region is proposed to form the loop region in the compact dimer structure to allow for folding back of the chains, with preliminary solution analysis supporting this.

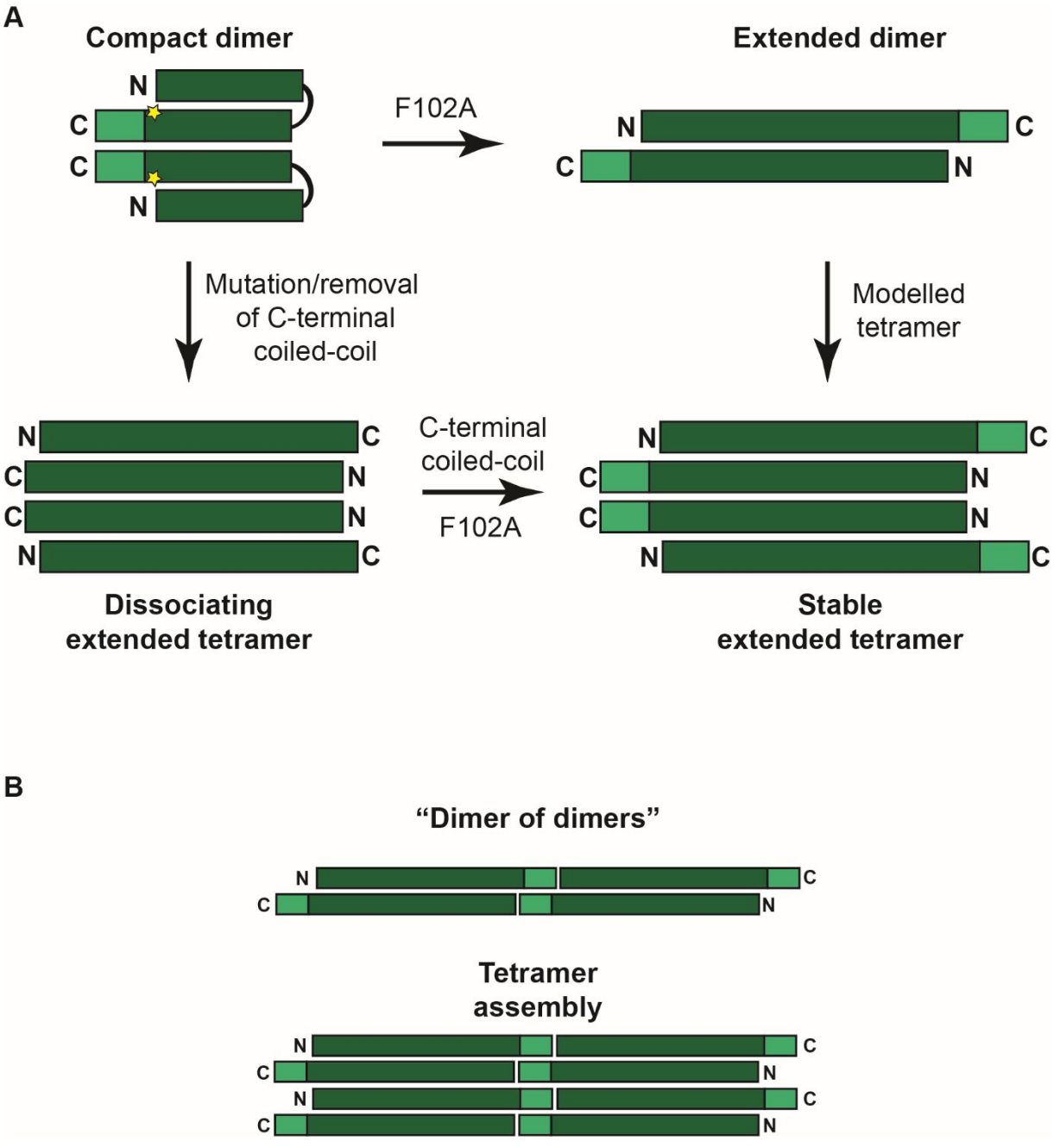


Figure 4.40. Schematic models for TEX12 conformations. (A) Schematic to show the various conformations of TEX12. The WT protein forms a compact dimer through folding back of the chains. The truncation or mutation of the C-terminal coiled-coil heptads forms an unstable extended tetramer. The mutation of the F102 residue (yellow star) forms a stable anti-parallel extended dimer with C-terminal protrusions. The stable tetramer, as characterised in solution, can be modelled from the extended dimer structure. (B) Schematic to show the crystallographic associations that occur in the crystal lattice. A "dimer of dimers" assembly is formed between anti-parallel C-termini. A tetrameric assembly is also formed by the C-terminal coiled-coils.

The extended dimeric TEX12 mutant crystal structure showed crystallographic oligomers, where the protruding C-termini form anti-parallel coiled-coil interactions produce a "dimer of dimers", which then forms a recursive assembly (Figure 4.40 B). The protruding C-termini also form tetrameric assemblies through continuous and oblique coiled-coil interfaces (Figure 4.40 B). Therefore, the crystallographic assemblies provide insight into the mechanisms of TEX12 interactions through the critical C-terminal coiled-coil region, whereby a continuous assembly of TEX12 can be formed. Although TEX12 assembles do not occur in solution, these assembly mechanisms provide molecular insight into how TEX12 may drive the higher-order assembly of the SYCE2-TEX12 complex (Figure 4.40 B).

4.3.2 The TEX12 C-terminus is critical for the formation of SYCE2-TEX12 filaments

The crystal structures of the TEX12 mutants, combined with extensive biophysical characterisation, has demonstrated the capacity for TEX12 to undergo several conformations due to the associations in the C-terminal region. The study of TEX12 in isolation has consequently provided the molecular basis for assembly of TEX12, which may form the backbone of the SYCE2-TEX12 filaments.

Through the truncation or mutation of the hydrophobic coiled-coil residues, the previously described 4:4 complex was restricted to a 2:2 complex. This 2:2 complex could no longer undergo self-assembly in order to form filamentous structures. Therefore, the coiled-coil associations observed in the TEX12 crystal structure are seemingly required for 4:4 complex formation and consequent assembly, with the assembly of the hetero-tetrameric complex likely forming a similar helical assembly as the TEX12 homo-tetramer (Figure 4.41 A). The 4:4 complex is only formed with the inclusion of the TEX12 C-terminal coiled-coil, suggesting that the coiled-coil interaction drives the interaction between two 2:2 complexes. Therefore, the 4:4 complex may be formed through a tetrameric assembly of TEX12 sandwiched by SYCE2 (Figure 4.41 A).

The cross-linking of the 2:2 complex indicates that both SYCE2 and TEX12 form respective anti-parallel arrangements of chains. It is unclear whether TEX12 exists in the compact dimer form or the extended tetramer form when bound to SYCE2. The wild type TEX12 protein in isolation forms a compact dimer structure by folding back of the chains mediated by a break in the helical core region. This could suggest that TEX12 may also adopt this structure when in complex with SYCE2. In this case, two sets of compact TEX12 dimers would likely interact with four extended SYCE2 chains in order to form the elongated 4:4 complex (Figure 4.41 A). Although the wild type protein forms the compact structure in solution, the TEX12 mutants have been found to exist in an open, extended conformation. The

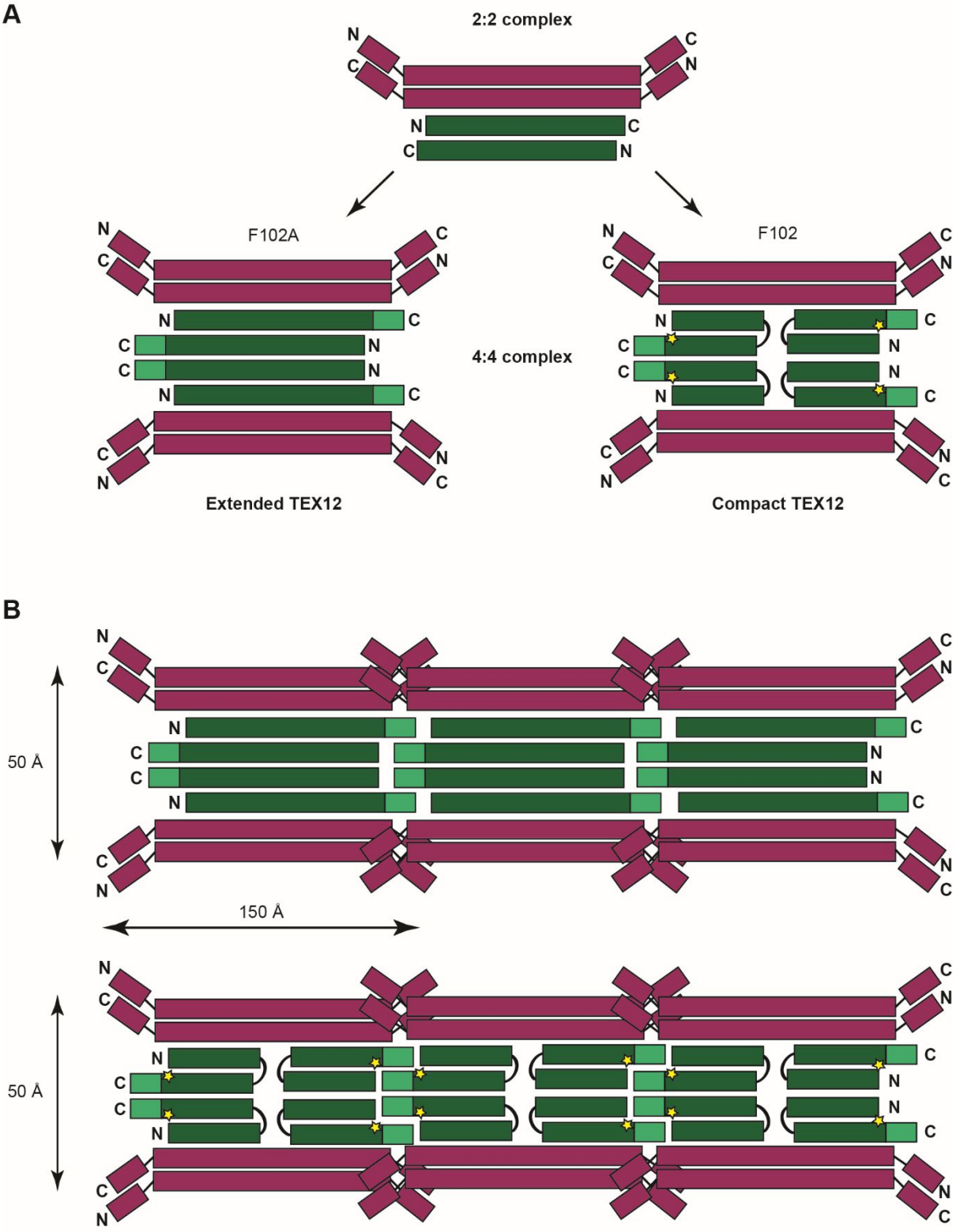


Figure 4.41. 2D schematic models for SYCE2-TEX12 assembly. (A) Schematic to show the formation of the SYCE2-TEX12 2:2 complex upon TEX12 coiled-coil disruption. The 4:4 complex is formed through associations of 2:2 complexes by the TEX12 C-terminus, with TEX12 either in an extended or a compact form. The F102 residue (yellow star) imposes the TEX12 compact structure in the absence of SYCE2. (B) Schematic to show the higher-order assembly of SYCE2-TEX12 4:4 complexes driven by the TEX12 C-terminal coiled-coil interactions. The labelled dimensions correspond to the parameters determined by the fibre X-ray diffraction pattern.

mutation or removal of the TEX12 C-terminal coiled-coil residues results in the formation of a 2:2 complex. The solution structure of these complexes show an elongated shape with an Rc value suggesting the formation of a four-helical bundle. This implies that the TEX12 chains would be in the open, extended conformation in the 2:2 complex with SYCE2 (Figure 4.41A). The elongated 2:2 complexes could then interact in a lateral manner to form the 4:4 complex that could then form the higher-order assemblies by longitudinal interactions between neighbouring 4:4 units (Figure 4.41 A and B). In the case of either of these two models the TEX12 chains, and consequential interaction surface with the SYCE2 chains, would have a very similar overall organisation and both possibilities result in the formation of an eight helical assembly with protruding TEX12 C-termini well positioned for higher-order assembly (Figure 4.41 B).

The TEX12 helical core F102A mutation, stabilised by F109E and V116A mutations, forms a 4:4 complex with the SYCE2 core. However, the complex does not form higher-order assemblies, unlike the WT core complex. Upon inclusion of the unstructured SYCE2 and TEX12 termini, the F102A complex does undergo filament formation. This suggests that the F102, F109 and V116 residues contribute to the assembly of the SYCE2-TEX12 core, but that the unstructured termini can compensate for this.

SYCE2-TEX12 crystals were found to produce a fibrous X-ray diffraction pattern that correlates to the k-m-e-f class of proteins (Fraser and MacRae, 1961). The fibrous crystals were formed by the core complex harbouring the 10 amino acid C-terminal truncation. Therefore, this suggests that when subject to the high protein concentration in a crystal, the truncated complex has residual affinity for higher-order assembly. The truncated complex lacks three of the four characterised hydrophobic coiled-coil residues, suggesting the coiled-coil formation does not occur, but does contain the crucial F102 residue, along with the F109 residue, indicating that the F102 residue indeed has a contributory role for the fibrous assembly. The d-spacings determined from the fibre diffraction reveals a unit cell length of 150 Å and a width of 50 Å, which approximates to the dimensions estimated by SAXS for the 4:4 complex. Therefore, the SYCE2-TEX12 filamentous structures are likely formed from longitudinal interactions between 4:4 units, driven by the TEX12 C-terminus, with the basis of assembly maintained when TEX12 is in the compact or extended form (Figure 4.41 B).

4.3.3 Model for SC maturation through SYCE2-TEX12 assembly

The SYCE2-TEX12 higher-order assemblies have been characterised as coiled-coil fibres formed from SYCE2-TEX12 units, which are connected through coiled-coil interactions at the TEX12 C-terminus. These filaments are thought to drive elongation of the SYCE2 by

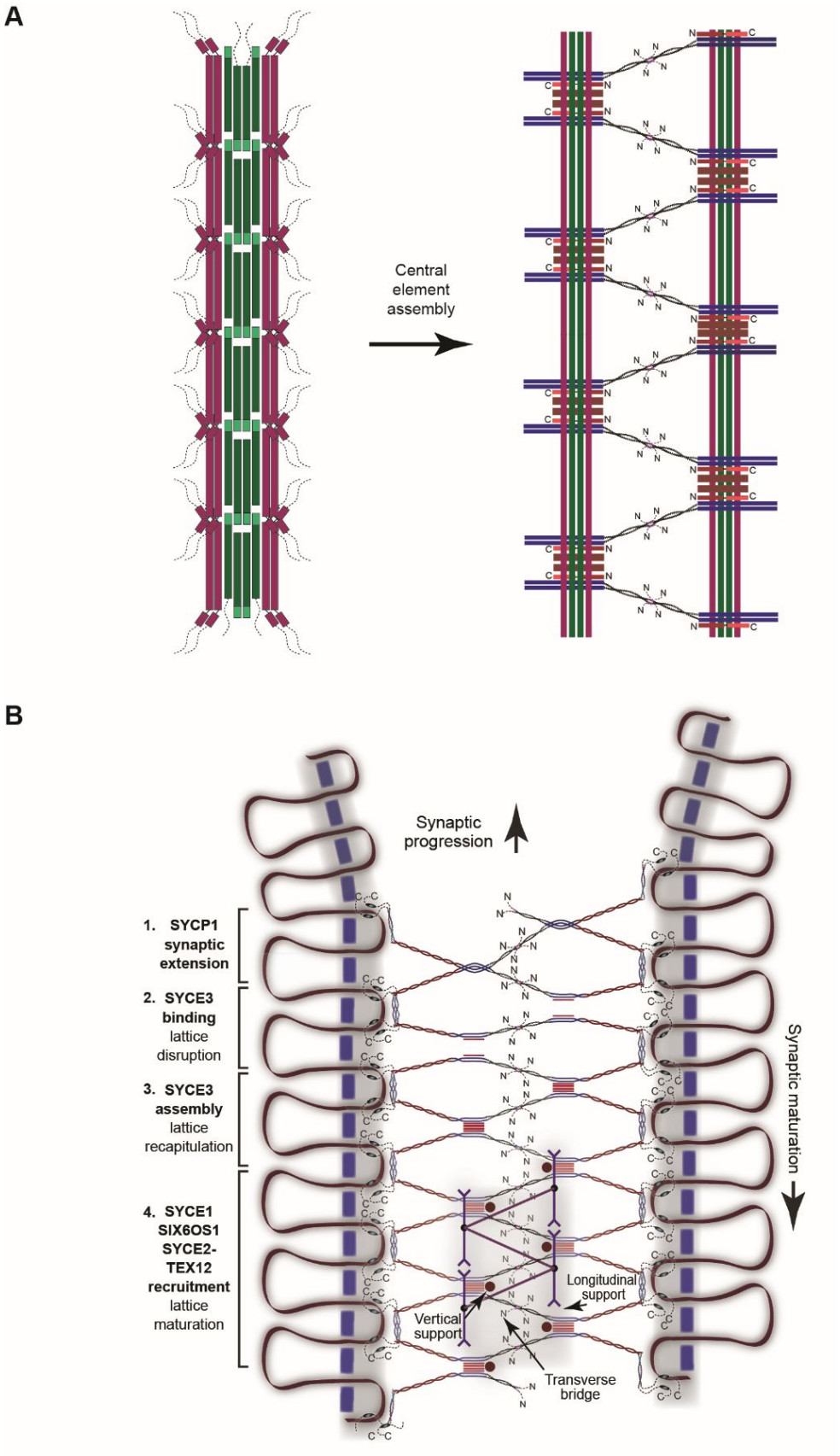


Figure 4.42. Schematic models for SYCE2-TEX12 assembly within the SC. (A) Schematic to show the formation of the SYCE2-TEX12 fibres that bind to SYCP1-bound SYCE3 at the two central element parallel cables. (B) Progressive model for SC assembly upon component recruitment and gradual re-organisation of the SC lattice.

stabilising the central element (Hernandez-Hernandez *et al*., 2016). An interaction between SYCE2-TEX12 and the central element protein SYCE3 has been established, and therefore, SYCE3 is thought to recruit the SYCE2-TEX12 complex (see Chapter three). Super-resolution microscopy experiments have revealed that the central element exists as two parallel cables (Schücker *et al*., 2015). Therefore, SYCE2-TEX12 may self-assemble along the two cables though periodical interactions with SYCE3 and consequent organisation of the SYCP1 transverse filaments (Figure 4.42 A), in order to drive elongation of the SC along the entire lengths of chromosomes (Figure 4.42 B).

The assembly of the SC occurs through the iterative recruitment of SC proteins and the gradual re-organisation of the lattice by each component until SYCE2-TEX12 can drive the recursive assembly to achieve mature SC formation and chromosome synapsis (Figure 4.42 B). However, the mechanism of SC disassembly remains uncertain. It has been postulated that modification events, such as phosphorylation, occur in order to trigger disassembly. As the TEX12 C-terminus is critical for high-order assembly of the SYCE2-TEX12 complex, it could be reasoned that modification at this site may disrupt the critical coiled-coil formation to lead to disassembly of the filamentous structures, which may then cause the dissociation of the rest of the SC when required. Analysis of the TEX12 sequence for potential phosphorylation sites reveals two potential PKC phosphorylation sites within the TEX12 C-terminus at the T115 and T120 positions. The PKC family of proteins have been shown to function at several stages within meiosis, and therefore, may provide a potential mechanism for SC disassembly (Viveiros *et al*., 2003).

Chapter five.

Assembly of the yeast SC central region

5.1 Introduction

5.1.1 The yeast synaptonemal complex

The study of mammalian meiosis *in vivo* is challenging due to the lack of a genetically tractable system for cellular experimentation. The overall tripartite architecture and function of the SC is conserved throughout meiotically dividing organisms, including the unicellular budding yeast (von Wettstein *et al*., 1984). However, the SC components can vary throughout evolution, with differing protein constituents having no discernible sequence homology. The common model organisms, *S. cerevisiae*, *C. elegans* and *D. melanogaster*, are such species that have a differing SC composition to mammals, although they retain the zipper-like structure with lateral elements at the axes, a central element in the midline and transverse filaments to complete the chromosome tethering assembly (Westergaard and von Wettstein, 1972).

The study of cellular processes in yeast is genetically amenable and affordable, with the rapid timing of the cell cycle providing an efficient system to investigate (Slater *et al*., 1977). The *in vitro* study of the yeast SC central region, through biochemical and structural biology techniques, can efficiently lead to the corresponding experimentation *in vivo* to test new hypotheses. This complementation of *in vivo* and *in vitro* data may provide the detailed molecular basis for SC assembly in yeast, and therefore, may elucidate mechanisms of the mammalian SC. In particular, extensive characterisation of the analogous mammalian and yeast transverse filament proteins could facilitate the design of a hybrid molecule, where the yeast C-terminus could be fused to the mammalian N-terminal region. This may enable the study of the mammalian central region through genetic manipulation, whilst maintaining the chromosomal recruitment of the lateral elements and transverse filament C-terminus to the homologue axes.

The yeast SC lateral element consists of Hop1, Red1 and Zip3 (Hollingsworth and Ponte, 1997b). These proteins assemble at the chromosome axes. The transverse filaments are formed by Zip1 and interdigitate to form the backbone lattice of SC assembly (Sym and Roeder, 1995). This preliminary lattice is matured by central element assembly to form the fully functional structure. The main components of the yeast SC are the Ecm11 and Gmc2 proteins (Humphryes *et al*., 2013), along with the yeast SUMO protein Smt3 (Voelkel-Meiman *et al*., 2013) (Figure 5.1). SUMOylation of the yeast SC components, by Smt3, has been implicated in SC assembly (Leung *et al*., 2015) and the lateral element component Zip3 has been shown to be a SUMO E3 ligase (Cheng *et al*., 2006). Two candidate Ecm11 SUMO sites have been located in the N-terminus, at K5 (Zavec *et al*., 2008) and K101, and these sites are SUMOylated in a Gmc2-dependant manner (Humphryes *et al*., 2013). This Ecm11 SUMOylation promotes Zip1 assembly within the SC, and has been suggested to cause the switch between Zip1 poly-complex formation and Zip1 SC assembly (Humphryes *et al*., 2013).

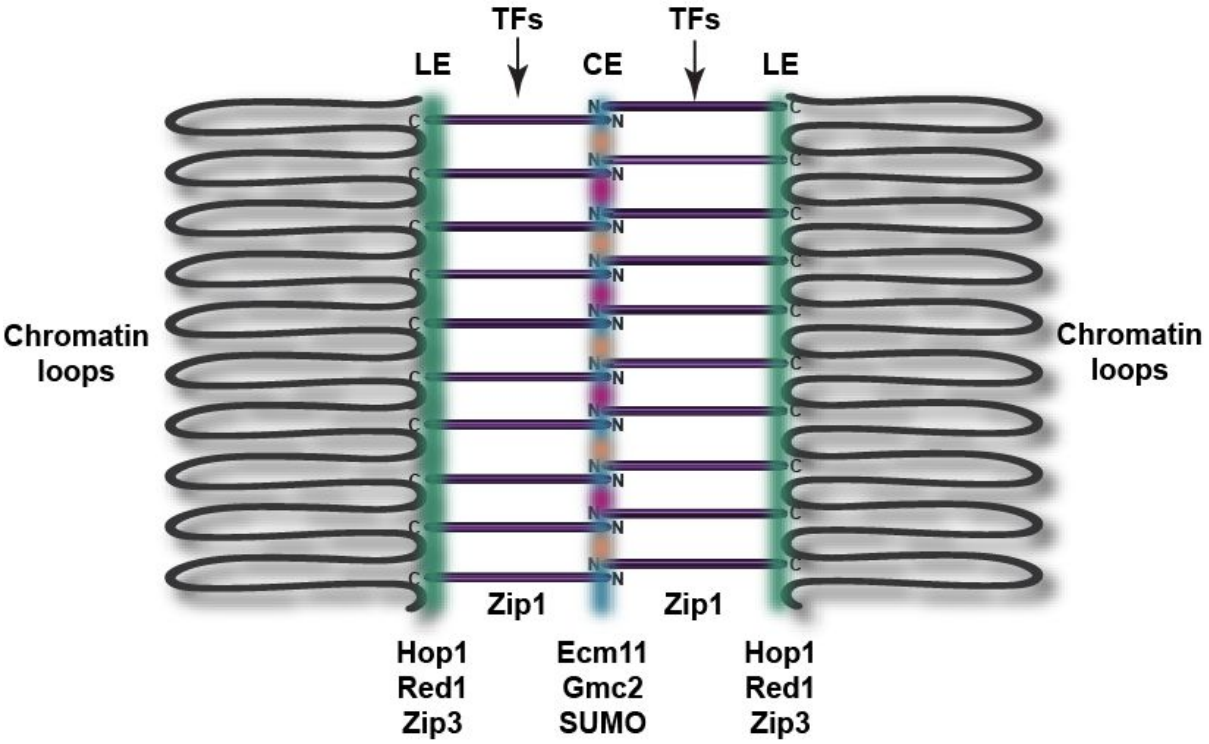


Figure 5.1. Schematic of the yeast synaptonemal complex. LE = lateral element. TFs = transverse filaments. CE = central element. Schematic depicting the chromatin loops of the homologous chromosomes with the assembly of the lateral elements at the axes (green). The transverse filaments (purple tubes) bridge between the chromosome axis and the central element (orange, blue and pink). The lateral elements are comprised of Hop1, Red1 and Zip3. Zip1 forms the transverse filaments and the central element is formed by Ecm11, Gmc2 and SUMO.

SC components	*Human*	*Yeast*
Lateral element	SYCP2 SYCP3	Red1 Hop1 Zip3
Transverse filament	SYCP1	Zip1
Central element	SYCE1 SYCE2 SYCE3 SIXSOS1 TEX12	Ecm11 Gmc2 Smt3

Table 5.1. Comparison of human and yeast SC constituents

5.1.2 Zip1 forms the SC backbone and undergoes poly-complex formation

The yeast transverse filaments are formed by Zip1, which is a long molecule that spans the distance between the chromosome axis and the central element. It has been shown through localisation studies that the transverse filament C-termini are located at the lateral elements and the N-termini meet at the central element (Liu *et al*., 1996). A detailed mechanism for the human transverse filament protein SYCP1 has been proposed whereby recursive self-assembly interactions propagate the formation of a preliminary SC lattice, with the model based upon biochemical and structural analysis of SYCP1 fragments (Dunce *et al*., 2018). Although SYCP1 shows no sequence homology to Zip1, an overall pattern of conservation and secondary structure prediction is conserved, implying that the sequences have diverged throughout evolution but have a preserved structure and function.

An interesting feature of both Zip1 and SYCP1 is the ability to form poly-complexes, which are back-to-back assembled units that are distinct from the SC (Sym and Roeder, 1995). The function of these polycomplexes is unclear, but the Ecm11 SUMOylation drives the incorporation of Zip1 into the SC preferentially from the poly-complexes (Leung *et al*., 2015). This potentially suggests that these structures may provide a storage unit for the transverse filaments until the correct time in the cell cycle for SC maturation.

Zip1 has 875 amino acids with patches of conservation found throughout (Figure 5.2A). Through secondary structure prediction, Zip1 is proposed to be largely alpha helical. In particular the central region of the molecule, corresponding to residues 175-748, forms the helical core of the molecule with predominantly unstructured N- and C-termini (Figure 5.2 B). Interestingly, the very N-terminal Zip1 region (amino acids 1-34) is highly conserved, corresponding to SYCP1 residues 1-32, and therefore possibly implicates a crucial role for central element incorporation (Figure 5.3). The SYCP1 region spanning amino acids 1-101 is largely unstructured, comparable to the unstructured Zip1 1-174 residues. The short N-terminal SYCP1 fragment (amino acids 101-111) form head-to-head interactions to allow self-assembly of SYCP1 in the central region, with the 206-362 region forming tetramers that undergo assembly with the inclusion of the 101-205 residues (Dunce *et al*., 2018). The analogous C-terminal boundary in Zip1 is predicted to be approximately amino acid 348 (Figure 5.3).

Therefore, biochemical studies of Zip1 were initiated based upon the SYCP1 findings to investigate whether the molecular properties are comparable, with the initial focus on the N-terminal region to probe central region assembly. As the proposed unstructured N-terminus contains the conserved 1-34 patch and a potential site for head-to-head assembly, fragments were characterised with this region to probe for the potential function, and also without this region in order to isolate a structured biological unit for biophysical and structural analysis.

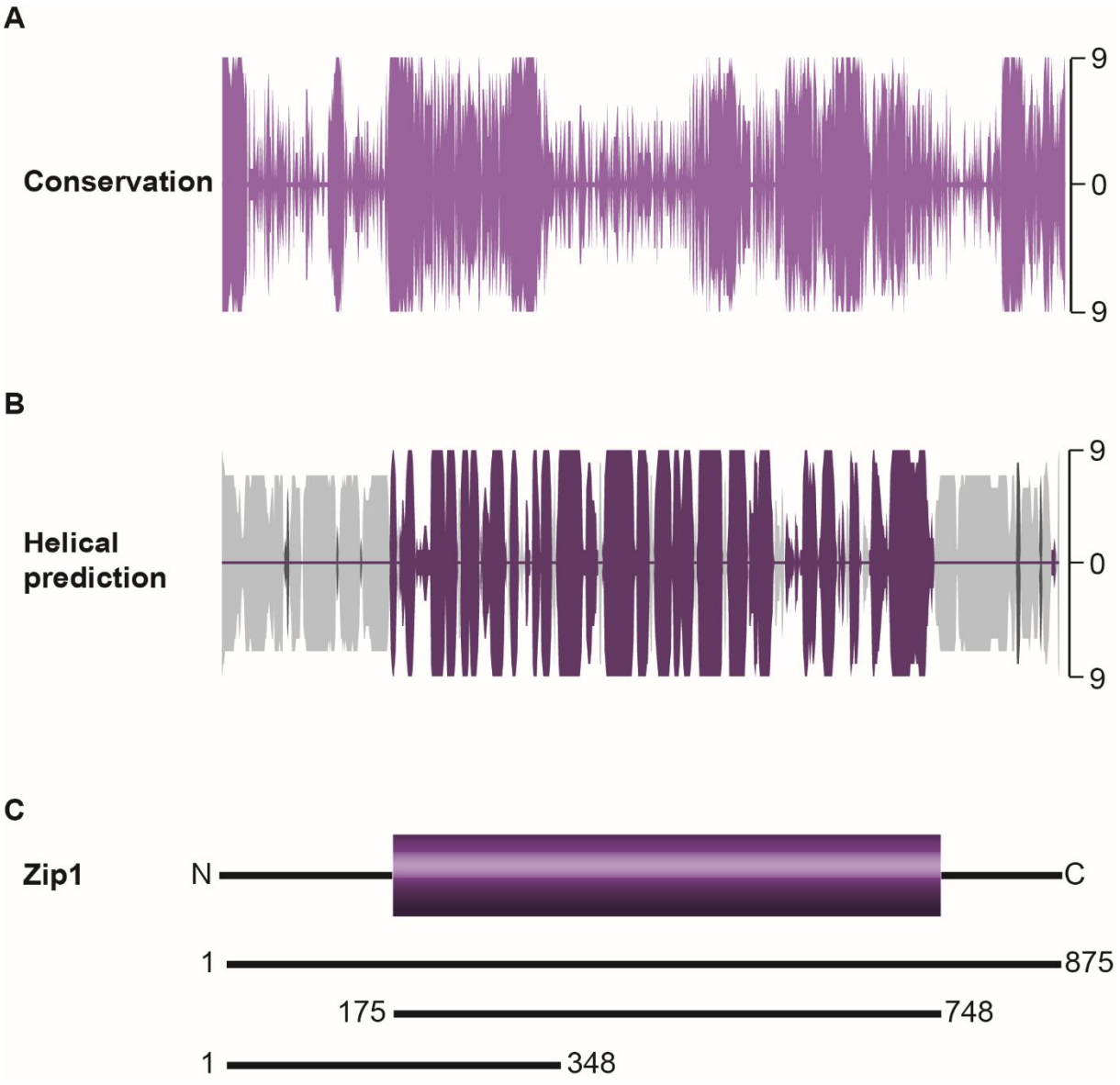


Figure 5.2. Sequence analysis of Zip1. (A) Zip1 conservation mirror plot. Multiple sequence alignments were generated by MUSCLE and analysed by ConSurf to grade the conservation of each amino acid position (0-9). (B) Secondary structure prediction plot for Zip1 from JNET prediction. Alpha helix (purple) beta sheet (dark grey) and unstructured (light grey) were plotted for each amino acid according to the confidence of the prediction (0-9). (C) Linear schematic representation of Zip1. The purple box represents the central helical core and the black lines represent the unstructured N-and C-termini. Key Zip1 regions indicated below: full length (1-875), the helical core (175-748) and the N-terminus (1-348).

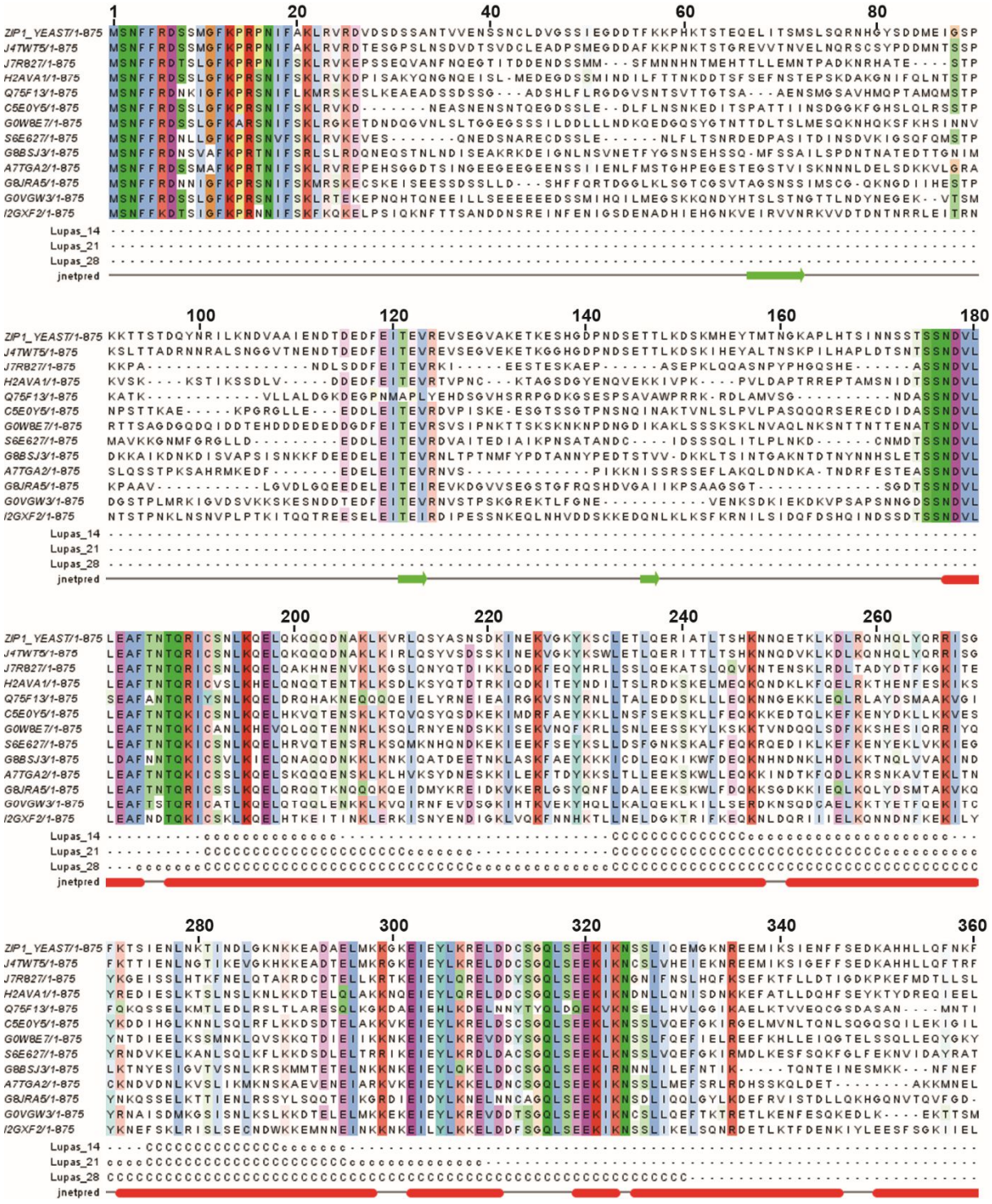


Figure 5.3. Sequence analysis of the Zip1 N-terminus. Zip1 multiple sequence alignments were generated by MUSCLE and visualised by Jalview. JNET predictions are plotted below. Alpha helices (red rods), beta sheets (green arrow) and unstructured regions (grey line). Lupas predictions demonstrate the prediction for coiled-coil formation.

5.1.3 Ecm11 and Gmc2 form the yeast central element

The yeast central element is comprised of Ecm11 and Gmc2. Although these proteins show no sequence similarity to any of the mammalian central element proteins, they have been shown to co-localise within the midline of the SC and are essential to form a mature and functional SC (Humphryes *et al*., 2013). Ecm11 and Gmc2 are thought to be analogous to the mammalian SYCE2-TEX12 complex that undergoes higher-order assembly to achieve SC elongation and maturation (Davies *et al*., 2012).

This hypothesis is supported by the evolution of the mammalian SC components, where the transverse filament SYCP1 and central element SYCE2 and TEX12 proteins are found to exist earlier in evolution than the other central element components (Fraune *et al*., 2016). This could suggest that SYCP1, SYCE2 and TEX12 were sufficient for SC formation, until development and recruitment of other central element components, facilitating the increase in complexity of meiotic division throughout evolution to humans. This suggests that the yeast SC may function as an early mammalian SC, with Ecm11 and Gmc2 performing both the initiation and elongation of central element assembly. An additional assembly mechanism may be provided by the SUMOylation events of Smt3.

Ecm11 has 302 amino acids and is predicted to contain alpha-helical stretches (Figure 5.4 A and C). The Ecm11 C-terminal region, spanning amino acids 206-302, is highly conserved (Figure 5.4 A and B) and has a high helical content (Figure 5.4 C). Therefore, this region is thought to represent the structural core. The rest of the molecule is predicted to be largely unstructured with patches of helical prediction, suggesting that this region may form interactions within the central element.

Gmc2 is a smaller protein comprised of 188 amino acids. Patches of amino acid residue conservation exist throughout the molecule, with the C-terminal region in particular having high conservation (Figure 5.5 A and B). Gmc2 is predicted to be predominantly alpha-helical, with four main helical portions (Figure 5.5 C). Therefore, the full Gmc2 sequence may be required for proper protein folding. However, the N-terminus has relatively low conservation and helical prediction and so removal of the 1-36 residues may stabilise Gmc2 *in vitro* for characterisation. The C-terminal half of Gmc2 (96-188) contains the highest amino acid residue conservation and predicted alpha-helical formation, suggesting this region may comprise the structural core of Gmc2 (Figure 5.5 D).

The Ecm11-Gmc2 complex has not previously been characterised *in vitro*, and further interactions have not been identified between Zip1 and the central element. The study of the

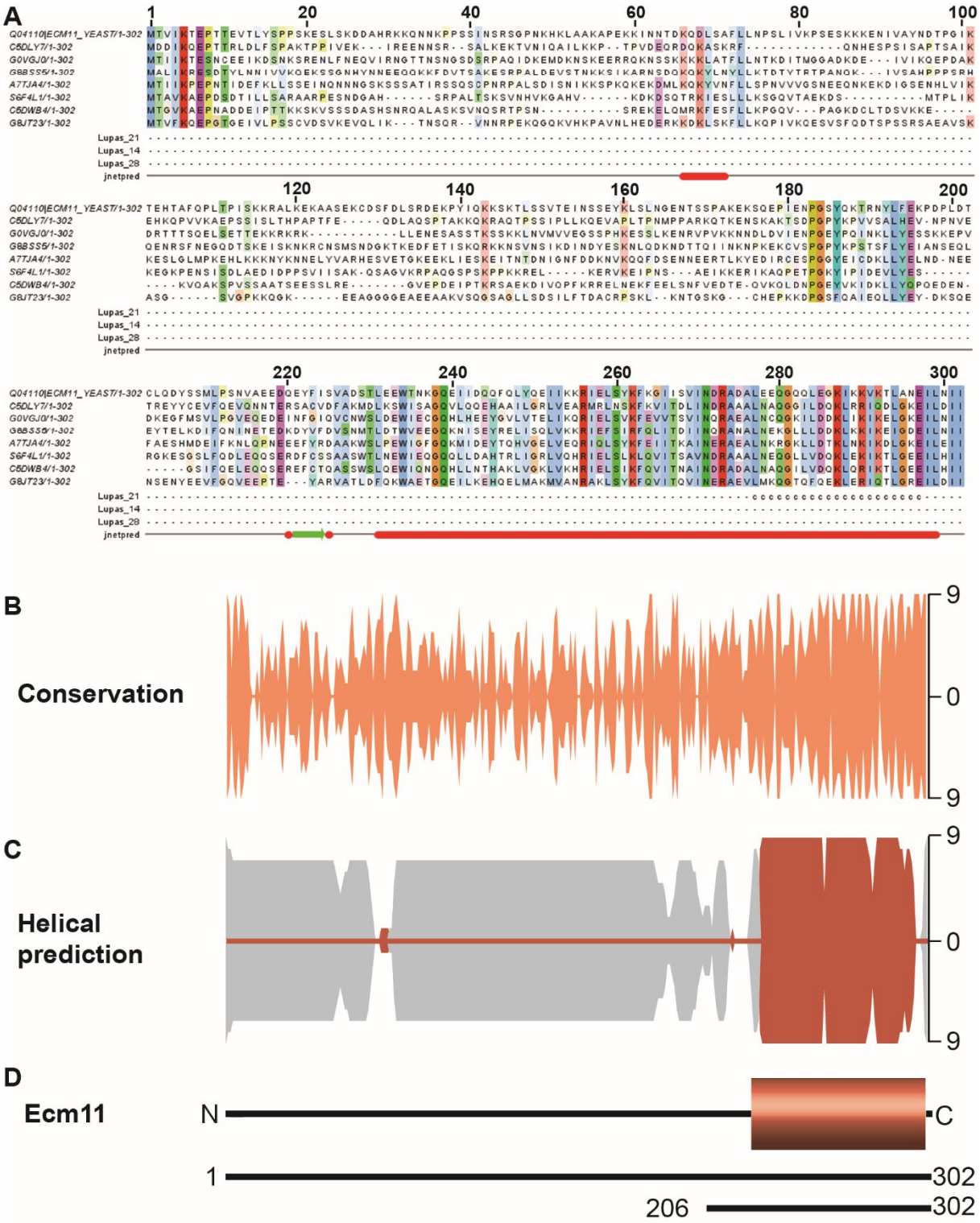


Figure 5.4. Sequence analysis of Ecm11. (A) Ecm11 multiple sequence alignment generated by MUSCLE and visualised in Jalview. JNET predictions are plotted below. Alpha helices (red rods), beta sheets (green arrow) and unstructured regions (grey line). Lupas predicts coiled-coil formation. (B) Ecm11 conservation mirror plot. MUSCLE multiple sequence alignments were analysed by ConSurf to grade the conservation of each amino acid position (0-9). (C) Secondary structure prediction plot for Ecm11 from the JNET prediction. Alpha helix (dark orange) beta sheet (dark grey) and unstructured (light grey) were plotted for each amino acid according to the confidence of the prediction (0-9). (D) Linear schematic representation of Ecm11. The orange boxes represent the main helical region (206-302) and the black lines represent the unstructured regions.

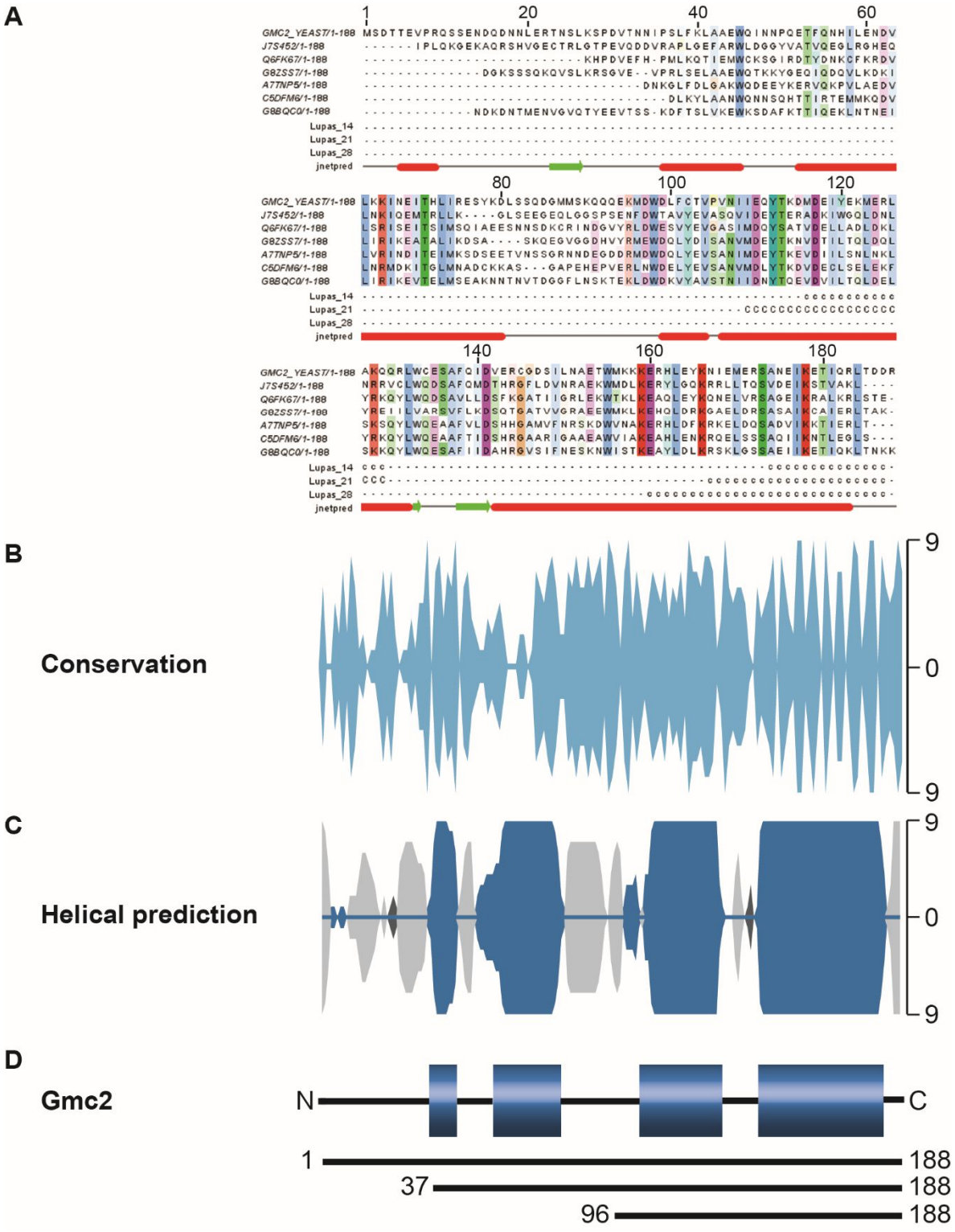


Figure 5.5. Sequence analysis of Gmc2. (A) Gmc2 multiple sequence alignment generated by MUSCLE and visualised in Jalview. JNET predictions are plotted below. Alpha helices (red rods), beta sheets (green arrow) and unstructured regions (grey line). Lupas predicts coiled-coil formation. (B) Gmc2 conservation mirror plot. MUSCLE multiple sequence alignments were analysed by ConSurf to grade the conservation of each amino acid position (0-9). (C) Secondary structure prediction plot for Gmc2 from the JNET prediction. Alpha helix (blue) beta sheet (dark grey) and unstructured (light grey) were plotted for each amino acid according to the confidence of the prediction (0-9). (D) Linear schematic representation of Gmc2. The blue boxes represent the main helical regions and the black lines represent the unstructured regions. The 1-88, 37-88 and 96-88 regions are used in this study.

recombinant Ecm11 and Gmc2 proteins may permit the structural characterisation of both proteins in isolation and when in the complex. This could allow for the investigation of the molecular basis of central element assembly, for comparison to the human SYCE2-TEX12 complex, and to understand the implication of Ecm11 SUMOylation by Smt3.

5.1.4 SUMO pathway in Saccharomyces cerevisiae

SUMOylation events within the cell have been reported to have a function in regulating cellular processes such as DNA repair and maintenance of chromosome architecture. The SUMO pathway involves the transfer of SUMO between E1, E2 and E3 enzymes in order to conjugate SUMO to the target substrate (Nottke *et al*., 2017).

In *S. cerevisiae*, there is only one form of SUMO; Smt3. The inactive Smt3 is activated by the cleavage of the C-terminal peptide by a SUMO protease (Figure 5.6). The SUMO proteases in yeast are Ubiquitin-like proteases (ULPs), ULP 1 and ULP2, and they activate SUMO by exposing a di-glycine motif upon removal of the C-terminal peptide. The activated SUMO is then linked to the E1 activating enzyme by the di-glycine motif. The E1 activating enzyme in *S. cerevisiae* is comprised by a heterodimer of Aos1 and Uba2 and Smt3 is linked covalently by the cysteine residue of Uba2 (Gareau and Lima, 2010).

Smt3 is then transferred to the E2 conjugating enzyme, which is Ubc9 in *S. cerevisiae* (Figure 5.6). The E3 ligase then transfers Smt3 to the lysine residue of the target substrate. *In vitro*, SUMOylation of target substrates has been shown to take place without an E3 ligase suggesting that the E2 conjugating enzyme is sufficient. However, *in vivo* SUMOylation has greater specificity by an E3 ligase, which in *S. cerevisiae* are Siz1, Siz2, Zip3 and Mms21. The process of SUMOylation is reversible and the Smt3 modifications to target substrates can be removed by ULP1 and ULP2 (Gareau and Lima, 2010), which are the proteases that initially active Smt3 (Figure 5.6).

5.1.5 SUMOylation drives yeast SC assembly

SUMOylation of Ecm11 has been detected *in vivo* by the yeast SUMO protein Smt3 (Voelkel-Meiman *et al*., 2013). The SUMOylation of the Ecm11 K5 residue has been shown to promote Zip1 assembly within the SC. This has been shown through mutation of the lysine residue to an arginine, in order to prevent the SUMOylation event (Zavec *et al*., 2008). A second proposed Ecm11 SUMO site, K101, was also investigated. It was found that mutation of K5 and K101 separately did not completely prevent Ecm11 SUMOylation, but the combination of these mutants abolished these modifications (Voelkel-Meiman *et al*., 2013).

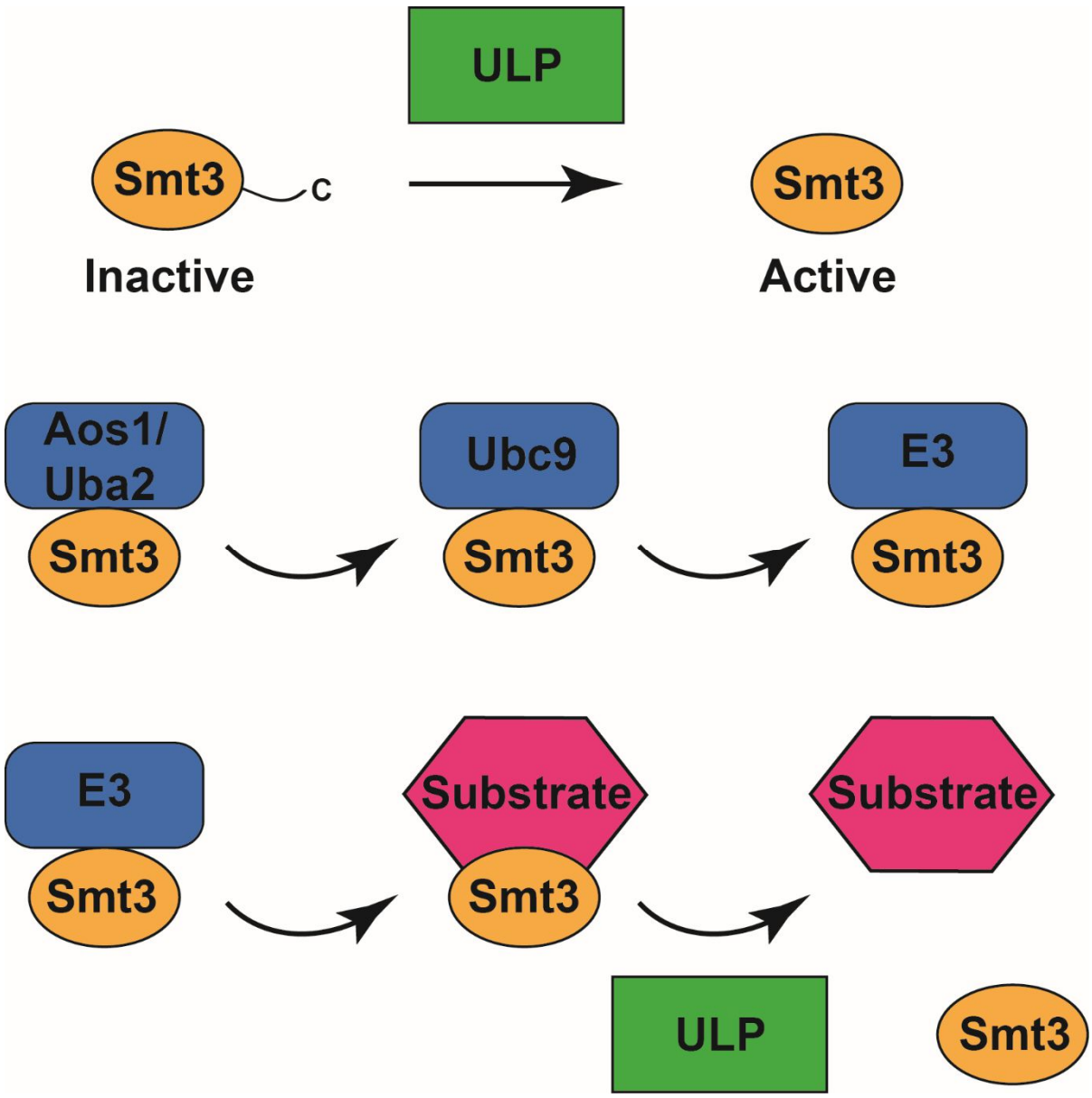


Figure 5.6. SUMO pathway in *S. cerevisiae*. The SUMO protein, Smt3, is activated by the removal of a C-terminal peptide by a ubiquitin-like protease (ULP). The activated Smt3 then binds to the E1 activating enzyme, comprised of Aos1 and Uba2, via a cysteine from Uba2. Smt3 is then transferred to the E2 conjugating enzyme Ubc9 and then to an E3 ligase. E3 ligases in *S. cerevisiae* are Siz1, Siz2, Zip3 and Mms21. Smt3 is then transferred to the lysine residue on the target substrate. The SUMOylation event can be reversed by removal of Smt3 from the substrate by the ULP enzymes that initially activate Smt3.

This implies that both of these residues are SUMOylated and that without these events Zip1 assembly in the SC is compromised, without altering the poly-complex formation. It has also been shown that the presence of Gmc2 is required for the Ecm11 SUMOylation. The deletion of Ecm11 alters both SC and poly-complex assembly of Zip1, suggesting a specific role for SC recruitment by the Ecm11 SUMOylation (Voelkel-Meiman *et al*., 2013).

Further to these findings, it has been identified that the Zip1 N-terminus is required for Ecm11 SUMOylation. Therefore, a mechanism for SC assembly has been proposed through a positive feedback loop, whereby Zip1 and Gmc2 are required for Ecm11 SUMOylation, and these events drive Zip1 SC assembly, with this iterative process allowing the recruitment and assembly of further transverse filaments (Leung *et al*., 2015).

Statement of contribution

I initiated the yeast SC project with the assistance of several undergraduate and master's students. I performed all purification and analysis of the Zip1 1-348 and 175-348 fragments, with the purification and analyses of further Zip1 truncations carried out alongside Jack Hughes (MSci) under my supervision. The Ecm11-Gmc2 studies were initiated by Chandni Ravindan (MRes) under my supervision. Further Ecm11-Gmc2 purifications were performed by Mumen Chowhury (MRes) with my supervision, and I performed all biophysical assays and analyses. All figures and analysis presented in this thesis have been prepared by myself and any data produced directly by students have been indicated. Some of the data presented in this chapter, including analysis and figures made by myself, have been presented by these students in their respective dissertations:

Chandni Ravindan, MRes dissertation, 2016 "*A constitutive complex between Ecm11 and Gmc2 within the yeast synaptonemal complex*".

Jack Hughes, MSci dissertation, 2017 "*Biophysical characterisation of the Zip1 N-terminal domain*".

Mumen Chowdhury, MRes dissertation, 2017 "*The molecular basis of chromosome synapsis and genetic exchange in meiosis*".

5.2 Results

5.2.1 Yeast 2-Hybrid studies of Zip1, Ecm11 and Gmc2

Ecm11-Gmc2 are shown to co-localise *in vivo* (Humphryes *et al*., 2013) and thought to act analogously to SYCE2-TEX12 in mammals. As they are the main CE components, it is thought that they may interact with an N-terminal region of the transverse filament protein Zip1. The Matchmaker Gold Y2H system was used to investigate any direct interactions between the central region proteins Ecm11, Gmc2 and the Zip1 N-terminus. Full length Ecm11, full length Gmc2 and the Zip1 N-terminus (1-348) were fused to the GAL4 activator domain with the full length Ecm11, Gmc2 37-188 (N-terminus truncated to abolish auto-activation) and the Zip1 N-terminus (1-348) were fused to the GAL4 DNA-binding domain. Interactions were determined through the detection of activated reporter genes (ADE1, HIS3 and MEL1) upon the two GAL4 domains binding due to the proximity imposed by positive interactions. As predicted, an interaction between Ecm11 and Gmc2 was observed, with the interaction observed to be maintained when the Gmc2 N-terminus is truncated (Figure 5.7 A). Ecm11 is also shown to self-interact, which indicates that Ecm11 can oligomerise by itself (Figure 5.7 A). A self-interaction is not observed with Gmc2, suggesting that Gmc2 may exist as a monomer in isolation with the capacity for Ecm11 binding. No direct interactions were found between the Zip1 N-terminus and either Ecm11 or Gmc2. This could suggest that an interaction may only be observed between Zip1 and a pre-formed Ecm11-Gmc2 complex, or that the full Zip1 sequence is required. However, a self-interaction is detected for the Zip1 N-terminus suggesting that this region can form oligomers (Figure 5.7 A).

To further investigate the Zip1 self-interaction, an additional Y2H screen was performed with Zip1 1-348 and a truncated N-terminal region corresponding to residues 175-348, which is predicted to be the structured helical portion of the N-terminus. The 175-348 amino acids are sufficient for the interaction, suggesting this predicted helical region is responsible for oligomerisation of the Zip1 N-terminus (Figure 5.7 B).

A final Y2H screen was then performed with truncations of Ecm11 and Gmc2 to identify the binding regions of the proteins, with constructs designed according to structure prediction and conservation (as described in Figure 5.4 and 5.5). The Ecm11 self-interaction is maintained with residues 206-302, which is the C-terminal structured region, and these residues are also sufficient for Gmc2 binding (Figure 5.7 C). The C-terminal structured region of Gmc2, residues 96-188, are observed to maintain the interaction with Ecm11, again with no Gmc2 self-interactions detected (Figure 5.7 C). These findings suggest that the structural cores of these proteins are driving the interactions, and also indicates suitable constructs for biochemical studies.

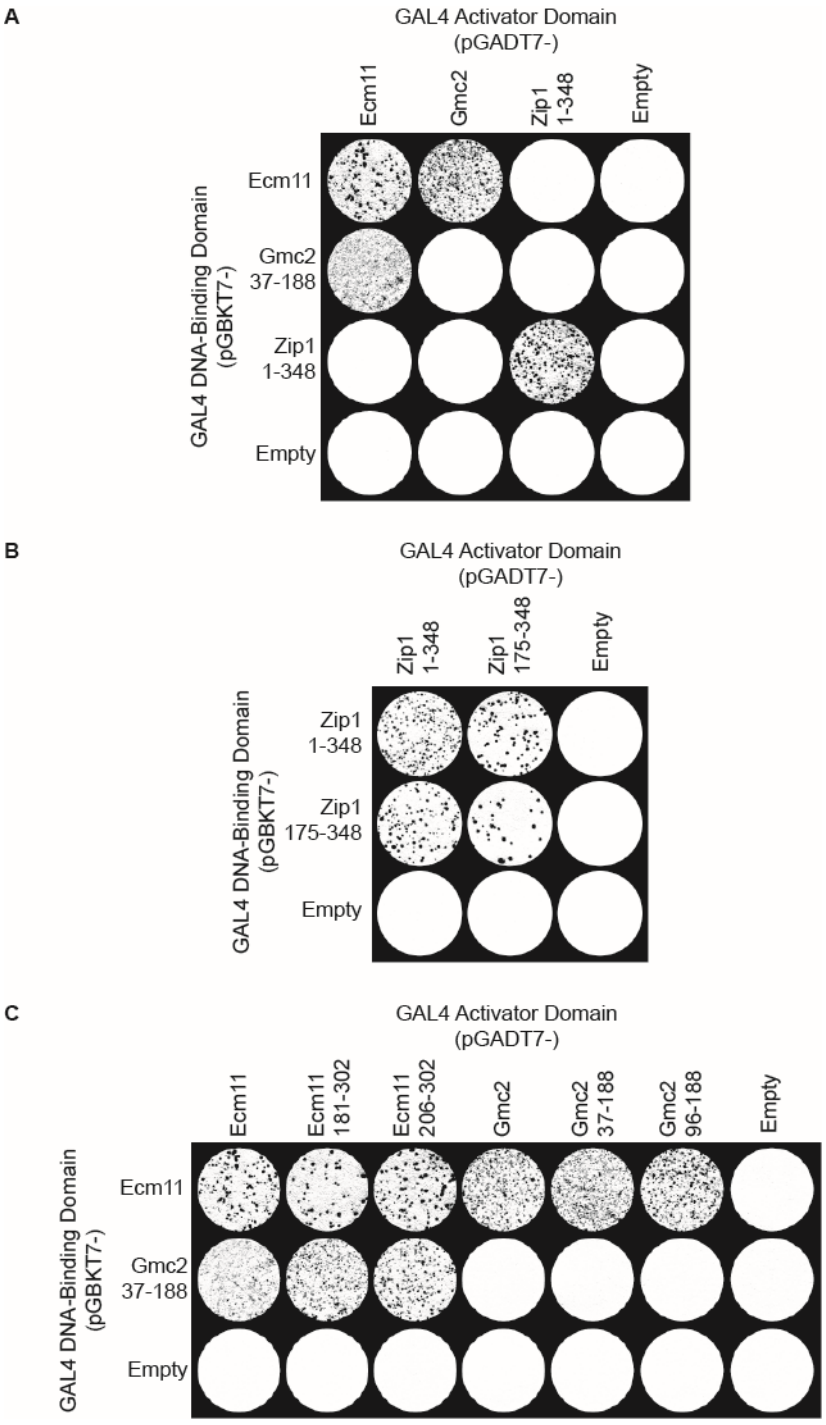


Figure 5.7. Yeast two-hybrid analysis of yeast SC proteins. Y187 (pGBKT7-bait) strains were mated with Y2H Gold (pGADT7-target) strains and plated on SD/-Leu/-Trp/-Ade/-His/x-α-gal plates for colony selection, subject to reporter gene activation (ADE1, HIS3 and MEL1) upon positive interactions. Experiments were performed by Chandni Ravindan under my supervision. (A) Y2H grid with Ecm11, Gmc2 and the Zip1 N-terminus showing self-interactions for Ecm11 and Zip1 and an interaction between Ecm11 and Gmc2. (B) $Zip1_{1\text{-}348}$ and $Zip1_{175\text{-}348}$ highlighting the Zip1 self-interaction through residues 175-348. (C) Ecm11 and Gmc2 truncations. Ecm11 interacts with itself and Gmc2 through amino acids 206-302 and Gmc2 interacts with Ecm11 through the 96-188 region.

5.2.2 Characterisation of the Zip1 N-terminus

To investigate whether Zip1 behaves in a similar manner to SYCP1 as hypothesised, N-terminal Zip1 fragments that correlate to the characterised SYCP1 constructs, were expressed in *E. coli* for purification and analysis. Zip1 residues 1-348 were found to self-interact by Y2H and from sequence analysis relate to the SYCP1 1-362 construct that oligomerises and interacts with the central element. $Zip1_{1-348}$ was expressed as a fusion protein to a His_6-MBP tag to aid solubility and stability. The recombinant protein was stable and the fusion protein was purified through sequential nickel-NTA and amylose affinity chromatography followed by anion exchange chromatography. The His_6-MBP tag was then removed by incubation with TEV protease and following cleavage, Zip1 was purified from the cleaved tag and TEV protease by cation exchange and then by SEC, where all fractions containing pure $Zip1_{1-348}$ were pooled and concentrated for analysis (Figure 5.8 A).

Zip1 was prone to heavy precipitation upon cleavage, suggesting either the protein is unstable without the tag or that the buffer conditions are not suitable. To determine whether KCl concentration in the buffer was unfavourable, a salt precipitation assay was carried out whereby protein aliquots were adjusted across a range of 400 to 100 mM KCl. The samples were incubated for one hour before centrifugation and subsequent analysis of the soluble and insoluble material. The assay revealed that at 100 mM KCl a large proportion of Zip1 was insoluble, indicating that low salt concentrations were encouraging precipitation (Figure 5.8 B). To test whether this precipitation could be reversed, a salt precipitation rescue assay was performed where Zip1 aliquots were incubated for one hour in 100 mM KCl to precipitate. The pelleted Zip1 was then incubated for a further hour in KCl concentrations of 400 to 100 mM for analysis of the soluble and insoluble material. This showed that all concentrations of KCl could re-solubilise Zip1 (Figure 5.8 C). These assays demonstrate that the precipitation is reversible and that higher KCl concentrations stabilise Zip1. As the protein is both precipitated and rescued at 100 mM KCl, it could suggest that protein concentration is also a factor. From these findings, the purification was optimised to keep the KCl concentration at 200 mM and above, and it was determined that the $Zip1_{1-348}$ at 200 mM KCl could still bind strongly to a cation exchange chromatography column. Therefore, through purification optimisation, $Zip1_{1-348}$ was isolated for analysis (Figure 5.8 A).

To study the secondary structure composition of $Zip1_{1-348}$, CD was used to measure far-UV spectra between 260 and 185 nm. The averaged data were plotted as MRE for direct comparison between samples. This analysis showed that, as predicted, $Zip1_{1-348}$ has an alpha helical composition and through deconvolution, the estimated helical content is 54%, corresponding to 190 helical residues out of 352 (Figure 5.8 D). CD was also utilised to perform

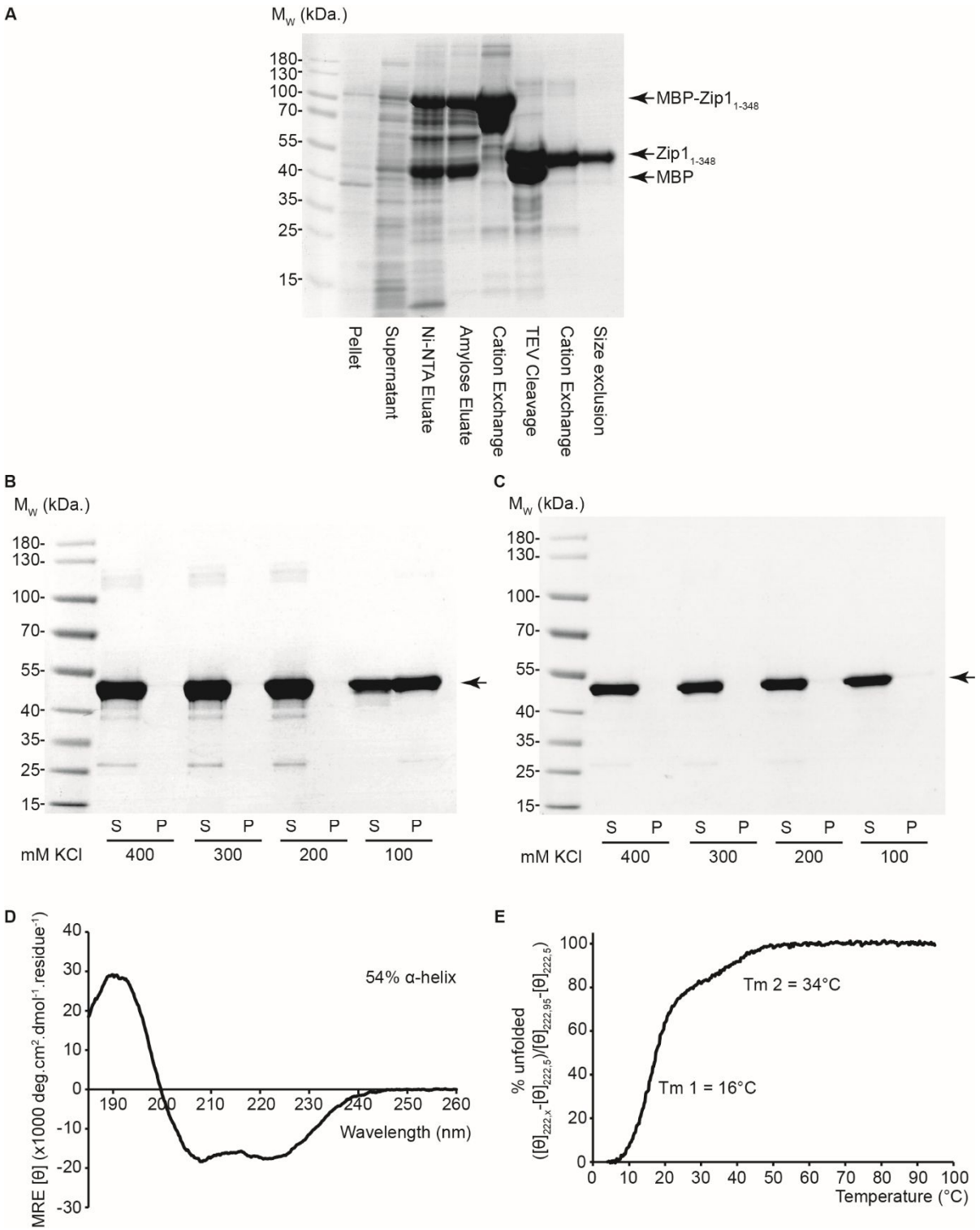


Figure 5.8. Purification and structural analysis of Zip1$_{1-348}$. (A) SDS-PAGE summary of Zip1$_{1-348}$ purification by Ni-NTA, amylose and cation exchange chromatography steps, followed by TEV protease incubation and subsequent cation exchange and size exclusion chromatography. (B-C) Salt precipitation assays. Samples were incubated at 400 to 100 mM KCl with soluble and insoluble material analysed by SDS-PAGE. Arrow indicates Zip1$_{1-348}$. (B) Precipitation was induced at 100 mM KCl. (C) Rescue assay. Precipitated Zip1$_{1-348}$ material was incubated with 400 to 100 mM KCl. All material was re-solubilised. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 54% alpha-helix, 8% β-sheet, 13% turns and 25% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm values estimated at the midpoint of the two respective events: 16°C and 34°C.

thermal denaturation, from 5°C to 95°C, in order to investigate the stability of Zip1$_{1-348}$ and to determine its melting temperature. The data, plotted as % unfolded, shows that Zip1$_{1-348}$ is relatively unstable with two unfolding events. The estimated melting temperatures determined at the midpoint of the unfolding events are 16°C and 34°C, respectively (Figure 5.8 E).

To investigate the oligomeric status of Zip1$_{1-348}$, SEC-MALS was used to determine the molecular mass of the protein species, using both the tagged and cleaved protein. This analysis was carried out in 300 mM KCl to prevent protein precipitation and it showed that the majority of the Zip1$_{1-348}$ material (demonstrated by the dRI trace) is monomeric, with a small proportion forming a dimer and higher molecular weight species (Figure 5.9 A). The tagged protein, His-MBP-Zip1$_{1-348}$, was predominantly monomeric, but also formed dimeric species, constituting ~30% of the material (Figure 5.9 B).

As Zip1$_{1-348}$ contains a basic patch, iterated by its ability to strongly bind to a cation exchange column, it was postulated whether this Zip1 region could bind to DNA. An electrophoretic mobility shift assay (EMSA) was performed to test this by incubating double-stranded DNA (dsDNA) with increasing concentrations of Zip1$_{1-348}$. These samples were then analysed on an agarose gel, where a shift of the DNA band would indicate binding to Zip1$_{1-348}$. As the instability of Zip1$_{1-348}$ has been determined by low salt concentrations, this experiment was performed at two concentrations of KCl. One experiment was carried out at 150 mM KCl to maintain near-physiological concentrations, but has the potential for Zip1$_{1-348}$ to precipitate. The other experiment was performed at 500 mM KCl to ensure Zip1$_{1-348}$ stability, but could affect DNA binding. A DNA shift was observed at 3.2 µM Zip1$_{1-348}$ at both salt concentrations, demonstrating the formation of protein-DNA interactions. The binding at 500 mM KCl suggests a strong affinity between Zip1$_{1-348}$ and dsDNA (Figure 5.9 C-D).

As suggested by sequence analysis and the CD spectra of Zip1$_{1-348}$ estimating 54% helical content, the very N-terminal region is predicted to be largely unstructured. Therefore, an N-terminal truncation, corresponding to residues 175-348 that was found to self-interact by Y2H, was expressed for characterisation. Through sequence analysis and isoelectric point (pI) calculation, Zip1$_{175-348}$ was found to contain the basic region as described for Zip1$_{1-348}$, with the pI for Zip1$_{175-348}$ and Zip1$_{1-348}$ determined as 9.00 and 5.69, respectively.

Expression and purification of Zip1$_{175-348}$ was carried out as described for Zip1$_{1-348}$, with the same issue of precipitation arising. To determine if the sensitivity to KCl concentration was occurring, a salt precipitation assay was performed from 400 to 100 mM KCl. As found for Zip1$_{1-348}$, 100 mM KCl caused Zip1$_{175-348}$ precipitation, with the majority of the material found to be insoluble (Figure 5.10 B). To establish whether the precipitation could be rescued, a

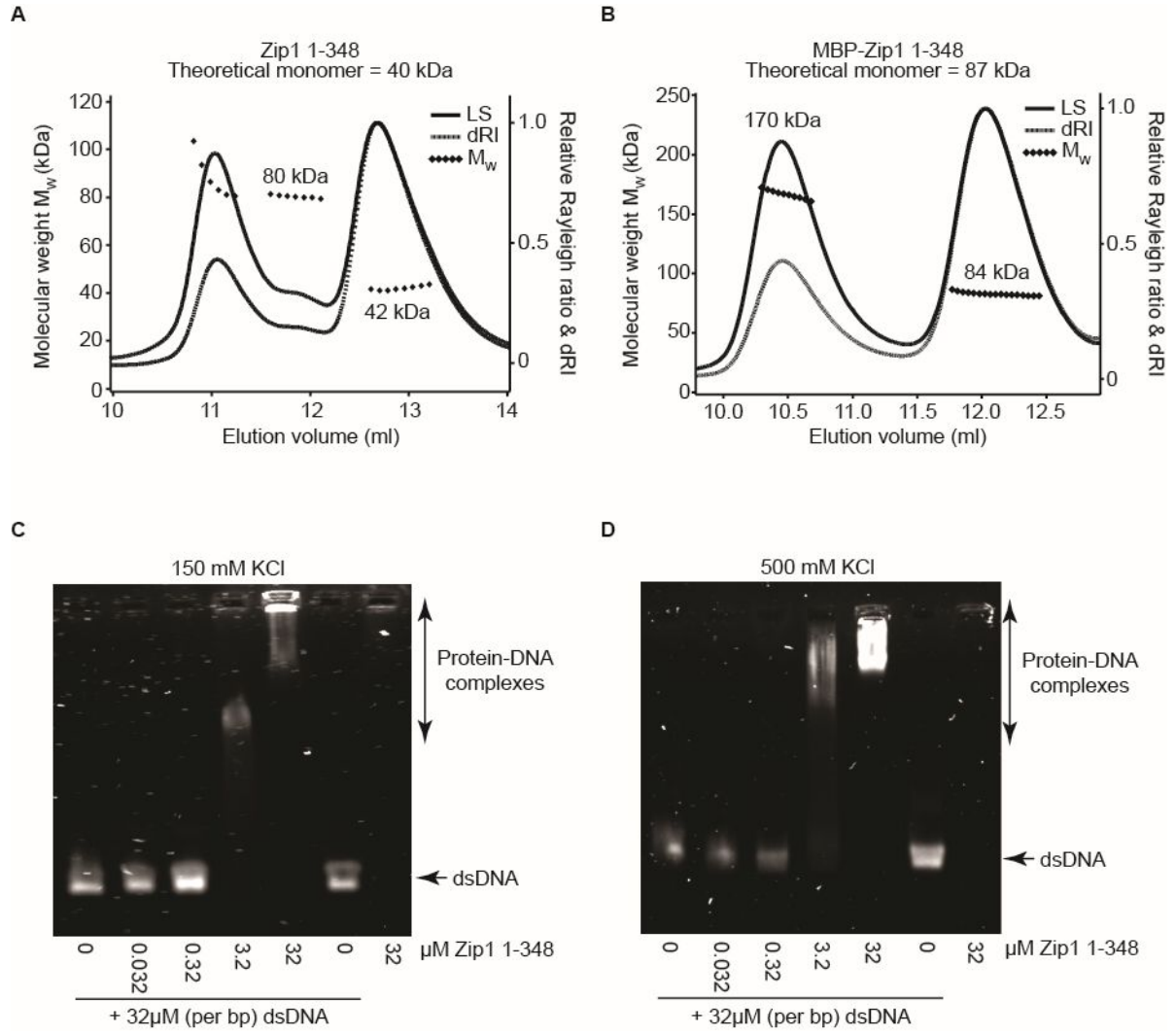


Figure 5.9. Oligomer determination and DNA-binding analysis of Zip1$_{1\text{-}348}$. (A) SEC-MALS analysis of cleaved Zip1$_{1\text{-}348}$. The predominant species is monomeric, with some larger species formed. (B) SEC-MALS analysis of the affinity tagged Zip1$_{1\text{-}348}$. His$_6$-MBP-Zip1$_{1\text{-}348}$ exists mostly as a monomer with some dimer species formed. (C-D) EMSA's with Zip1$_{1\text{-}348}$ and double-stranded DNA (dsDNA). Each lane contains 32 µM per base-pair dsDNA with indicated concentrations of Zip1$_{1\text{-}348}$. Band shifts occur at 3.2 µM Zip1$_{1\text{-}348}$, signifying DNA-binding. (C) Experiment performed in 150 mM KCl. (D) Experiment performed in 500 mM KCl.

salt precipitation rescue assay was carried out with pelleted Zip1$_{175\text{-}348}$ incubated at 400 to 100 mM KCl. Unlike Zip1$_{1\text{-}348}$, only a small proportion of the precipitated material was re-solubilised after incubation with 100 mM KCl (Figure 5.10 C). However, all material incubated at 200 to 400 mM KCl was fully rescued. Therefore, during purification KCl concentration was kept at 200 mM and above to enable the isolation of cleaved Zip1$_{175\text{-}348}$ for analysis (Figure 5.10 A).

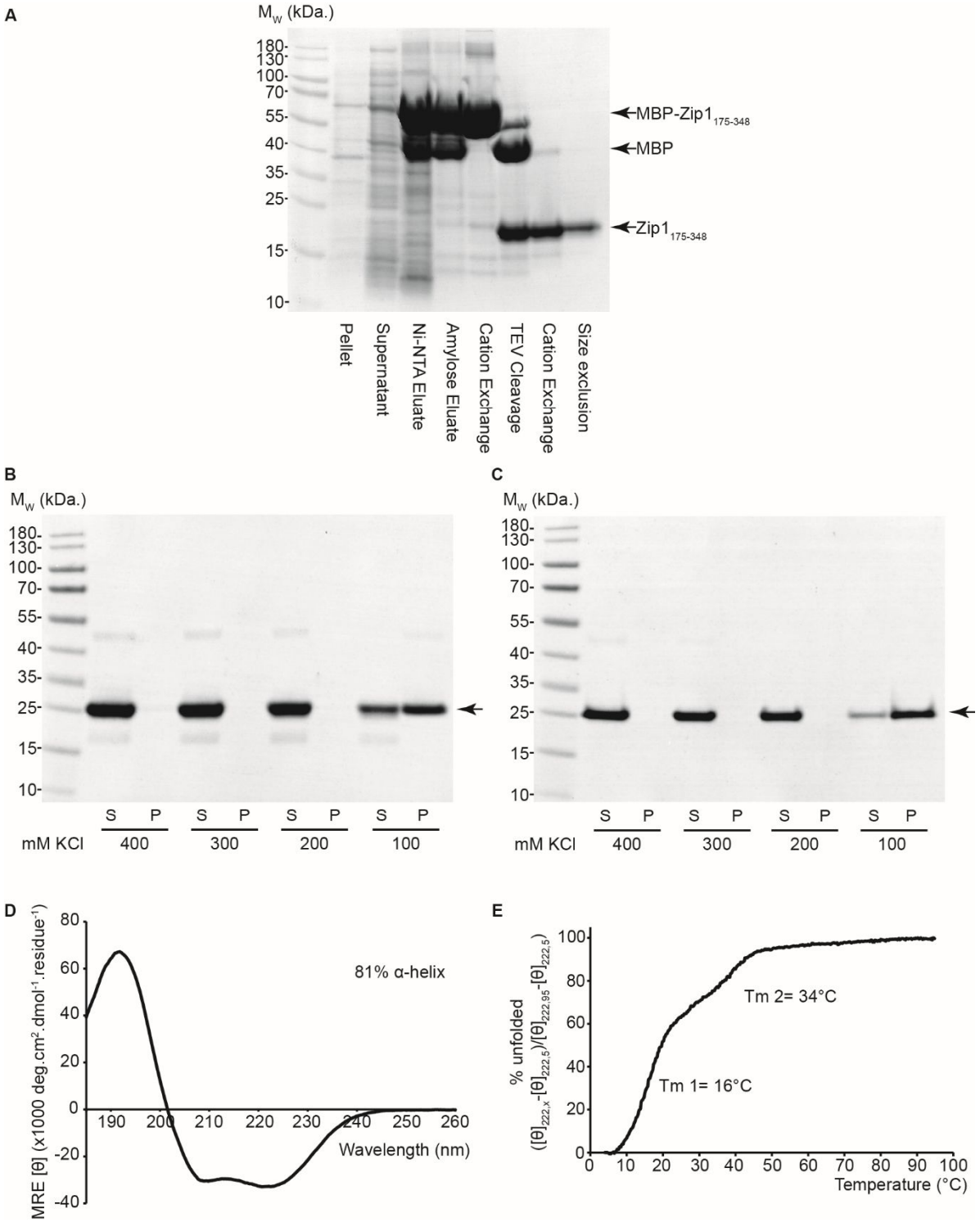


Figure 5.10. Purification and structural analysis of Zip1$_{175-348}$. (A) SDS-PAGE summary of Zip1$_{175-348}$ purification by Ni-NTA, amylose and cation exchange chromatography steps, followed by TEV protease incubation and subsequent cation exchange and size exclusion chromatography. (B-C) Salt precipitation assays. Samples were incubated at 400 to 100 mM KCl with soluble and insoluble material analysed by SDS-PAGE. Arrows indicates Zip1$_{175-348}$. (B) Precipitation was induced at 100 mM KCl. (C) Rescue assay. Precipitated Zip1$_{175-348}$ material was incubated with 400 to 100 mM KCl. Material was re-solubilised at 400-200 mM KCl but remained largely insoluble at 100 mM KCl. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 81% alpha-helix, 5% β-sheet, 7% turns and 7% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm values estimated at the midpoint of the two respective events: 16°C and 34°C.

A far-UV CD wavelength scan was collected to determine the secondary structure content of Zip1$_{175\text{-}348}$. The data show, as with Zip1$_{1\text{-}348}$, that Zip1$_{175\text{-}348}$ has a high helical signal with deconvolution estimating the helical content as 81%, correlating to 144 residues out of 178 total residues (Figure 5.10 D). The comparison of helicity between Zip1$_{1\text{-}348}$ and Zip1$_{175\text{-}348}$ reveals that residues 175-348 constitutes the helical core of the Zip1 N-terminus and highlights the unstructured nature of the region encompassing residues 1-174. This correlates with the secondary structure prediction but also relates to the structure of SYCP1, where the very N-terminus is unstructured and then followed by a highly helical region. CD thermal denaturation was used to determine the stability of Zip1$_{175\text{-}348}$ from 5°C to 95°C. The curve, plotted as % unfolded, shows a two-state transition. The estimated melting temperatures determined at the midpoint of the unfolding events are 16°C and 34°C, respectively (Figure 5.10 E).

To determine whether this structured region undergoes oligomerisation, SEC-MALS was used to calculate the molecular mass of the protein species, for both the tagged and cleaved Zip1$_{175\text{-}348}$. The cleaved Zip1$_{175\text{-}348}$ is predominantly monomeric, with small amounts of dimeric and higher-molecular weight species (Figure 5.11 A). His-MBP-Zip1$_{175\text{-}348}$ also exists as a monomer with 22% of the material forming dimers (Figure 5.11 B).

The basic region in Zip1$_{175\text{-}348}$ is likely responsible for DNA binding due to the electrostatic nature of protein-DNA interactions. To test this, EMSAs were performed with dsDNA and increasing concentrations of Zip1$_{175\text{-}348}$, in both low (150 mM) and high (500 mM) KCl concentrations, as described for Zip1$_{1\text{-}348}$. In 150 mM KCl, dsDNA was observed at 3.2 μM (Figure 5.11 C) and the same was seen at 500 mM KCl (Figure 5.11 D).

With these findings combined, the Zip1 N-terminal region is characterised as a helical DNA binding region with an unstructured N-terminal tail. The oligomeric status infers a predominantly monomeric structure with the ability to form dimers, although analysis has been carried out at a higher than physiological KCl concentration, potentially disrupting oligomerisation. The tendency of Zip1$_{1\text{-}348}$ and Zip1$_{175\text{-}348}$ to precipitate suggests that either Zip1 is not stable by itself, or that the construct boundary at the C-terminus is not optimal. To address this, further C-terminal boundaries were designed to optimise the N-terminal constructs to test whether Zip1 can be stable without a binding partner.

5.2.3 Optimisation of the Zip1 N-terminus for structural characterisation

The N-terminal fragment of SYCP1, corresponding to amino acids 101-175, was found to be a helical monomer in solution that forms an anti-parallel tetramer in the crystal structure, providing crucial insight for SYCP1 head-to-head assembly and therefore SYCP1 assembly *in*

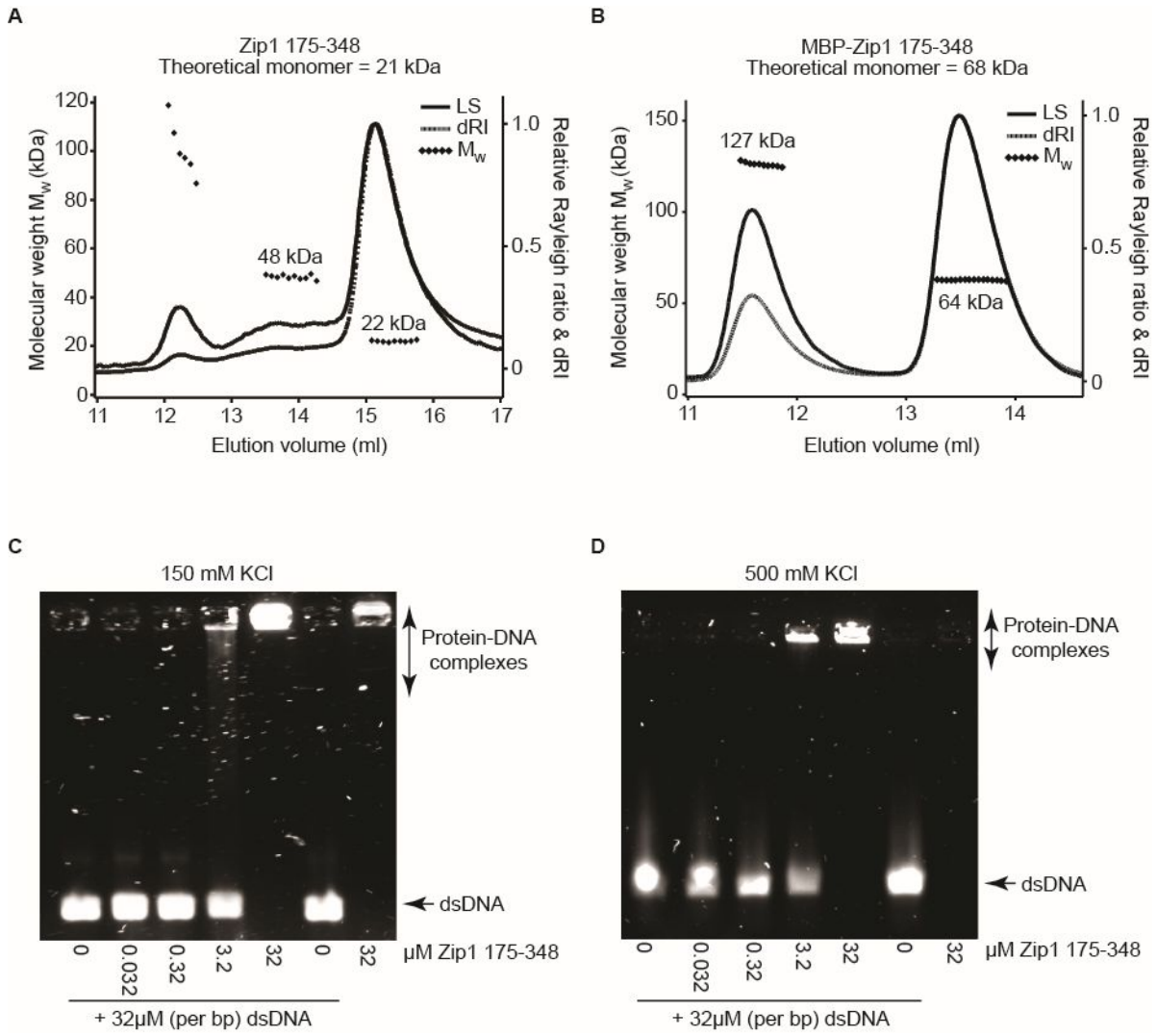


Figure 5.11. Oligomer determination and DNA-binding analysis of Zip1$_{175-348}$. (A) SEC-MALS analysis of cleaved Zip1$_{175-348}$. The predominant species corresponds to monomers, with some minor larger species formed. (B) SEC-MALS analysis of the affinity tagged Zip1$_{175-348}$. His-MBP-Zip1$_{175-348}$ exists predominantly as a monomer with some dimer formation. (C-D) EMSA experiments with Zip1$_{175-348}$ and double-stranded DNA (dsDNA). Each lane contains 32 µM per base-pair dsDNA with indicated concentrations of Zip1$_{175-348}$. Drastic band shifts occur at 3.2 µM Zip1$_{175-348}$, demonstrating DNA-binding. (C) Experiment performed in 150 mM KCl. (D) Experiment performed in 500 mM KCl.

vivo (Dunce *et al*., 2018). It was postulated whether an analogous region of Zip1 could be identified in order to determine whether SYCP1 and Zip1 behave similarly. The corresponding Zip1 region was predicted to span residues 175-251, therefore Zip1$_{1-251}$ and Zip1$_{175-251}$ were expressed and purified for analysis.

Zip1$_{1-251}$ was purified by the same method as Zip1$_{1-348}$, but precipitation was not observed at lower salt concentrations, including 100 mM KCl, indicating that the truncation has increased the stability of Zip1 (Figure 5.12 A). To investigate whether the truncation has

disrupted the folding of Zip1, CD far-UV spectra were measured. The data shows that the prominent secondary structure feature is alpha helix, but due to the reduction of the 222 nm minima peak, and the increase in the 208 nm minima peak, the spectra reveals a decrease in alpha helical content percentage and indicates an increase of random coil and unstructured residue percentage (Figure 5.12 C). Deconvolution of this data estimates a helical content of 21%, corresponding to 54 residues out of 255. This again indicates the unstructured nature of the Zip1 1-175 region, and demonstrates that the truncated amino acids, 252-348, are also highly alpha-helical. CD was used to measure thermal denaturation of $Zip1_{1-251}$ between 5°C and 95°C by recording the helical signal at 222 nm. The thermal melt, plotted as % unfolded, shows a sharp unfolding event, with an estimated melting temperature of 15°C (Figure 5.12 D). A single unfolding event was observed, unlike the two-stage unfolding of $Zip1_{1-348}$, suggesting that residues 252-348 may contain a more stable fold that denatures in the second stage.

Zip1 amino acids 175-348 have an overall basic charge, with a pI of 9.00, so the truncation of residues 252-348 has removed a large basic portion of the Zip1 N-terminus. Amino acids 175-251 still have an overall basic charge and maintain the ability to bind to a cation exchange column, although the affinity is reduced and requires loading at a lower salt concentration of 50 mM KCl. Therefore, the DNA-binding capacity of $Zip1_{1-251}$ may be altered. An EMSA was performed with dsDNA and increasing concentrations of $Zip1_{1-251}$. $Zip1_{1-251}$ showed limited DNA binding with no distinct species formed and smears were observed at the higher protein concentrations (Figure 5.12 B). This data shows a reduction of DNA binding due to the removal of Zip1 residues 252-348, compared to the distinct and larger species formed by $Zip1_{1-348}$ at 3.2 and 32 µM protein concentrations (see Figure 5.9 C-D). This suggests that DNA interactions are driven by a region within amino acids 252-348, or that these residues are required for a stable interaction. Furthermore, it indicates a specific DNA binding event taking place, rather than a non-specific interaction due to the basic nature of the Zip1 N-terminus. As the majority of this Zip1 fragment is comprised of the unstructured N-terminal region, comparably to the 1-101 region of human SYCP1, a Zip1 fragment was expressed without the unstructured region for characterisation and comparison to the SYCP1 101-175 fragment.

$Zip1_{175-251}$ was purified in the same manner as $Zip1_{175-348}$, but without the propensity for precipitation observed, to obtain near homogenous material for analysis (Figure 5.13 A). To verify that this construct is indeed stable at lower salt concentrations, a salt precipitation assay was performed at 150 mM and 200 mM KCl concentrations to decipher suitable buffer conditions. In both cases, $Zip1_{175-251}$ remained soluble with no material detected in the insoluble samples (Figure 5.13 B). This suggests that the C-terminal truncation has stabilised the Zip1 N-terminal region, perhaps to a break in alpha-helical structure, whereas amino acid 348 may have

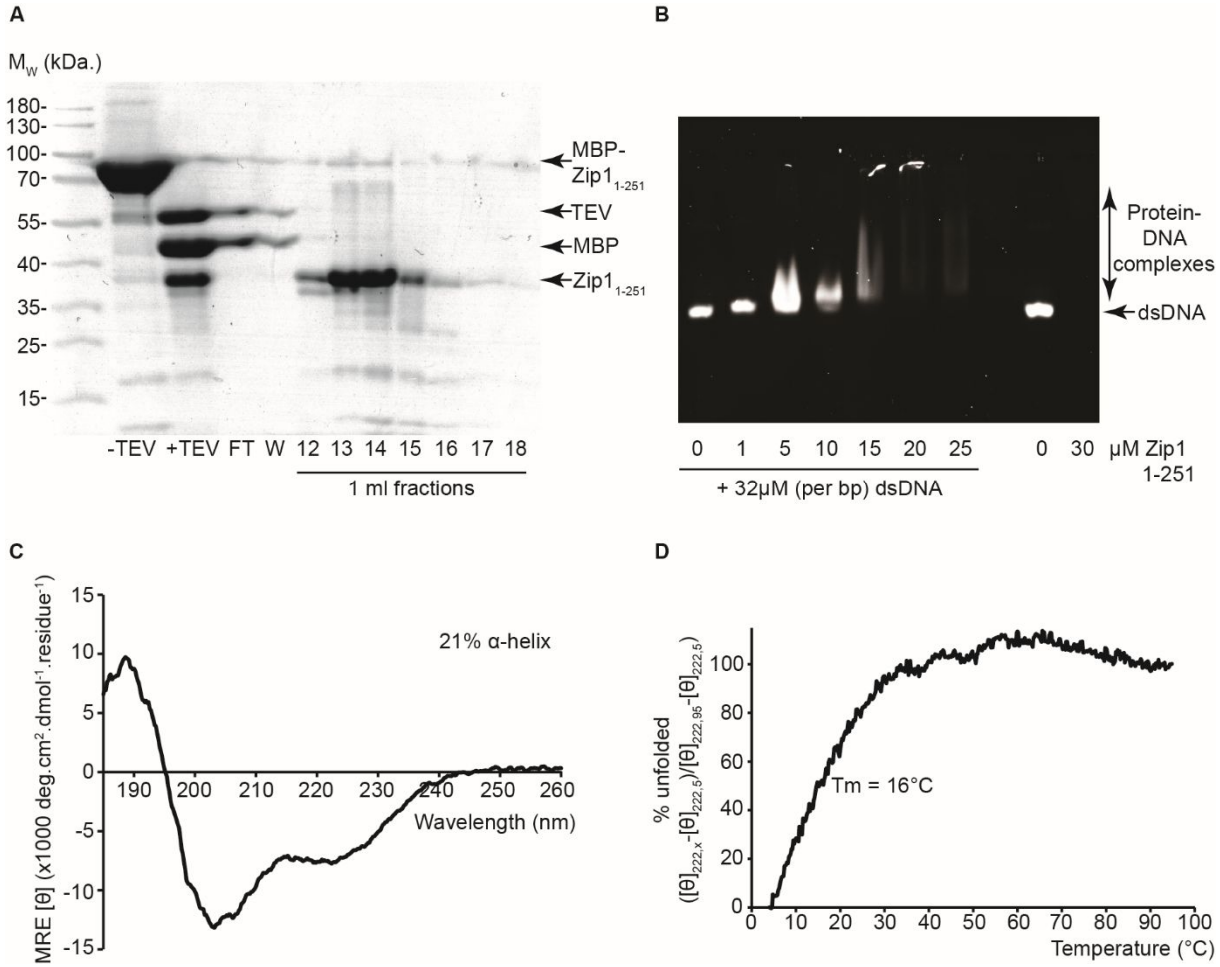


Figure 5.12. Purification and analysis of Zip1$_{1-251}$. Data collected by Jack Hughes under my supervision. (A) SDS-PAGE analysis of Zip1$_{1-251}$ purification by cation exchange chromatography following TEV protease incubation for tag removal. TEV protease labelled as TEV. (B) EMSA experiment with Zip1$_{1-251}$ and double-stranded DNA (dsDNA). Each lane contains 32 μM per base-pair dsDNA with indicated concentrations of Zip1$_{1-251}$. Limited band shifts occur at 25 μM Zip1$_{1-251}$, demonstrating restricted DNA-binding. (D) CD wavelength scan between 260-185 nm and plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 21% alpha-helix, 21% β-sheet, 25% turns and 33% unordered. (E) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 16°C, when the sample is 50% unfolded.

been at a critical structural point. This is iterated by the proportional yield of Zip1 material, where Zip1$_{175-348}$ was only stable up to 1 mg/ml but Zip1$_{175-251}$ was stable up to 13 mg/ml, with protein concentration limited due to material yield rather than stability.

To determine whether this truncated construct has altered oligomeric properties, SEC-MALS was used to calculate the molecular mass of Zip1$_{175-251}$. Due to the stability of this construct at near physiological KCl concentration, SEC-MALS experiments could be performed at 150 mM KCl. This analysis revealed that Zip1$_{175-251}$ formed a single species with a molecular mass corresponding to a monomer (Figure 5.13 C). This suggests that removal of the helical 252-348 residues has abolished the limited dimer formation as previously described.

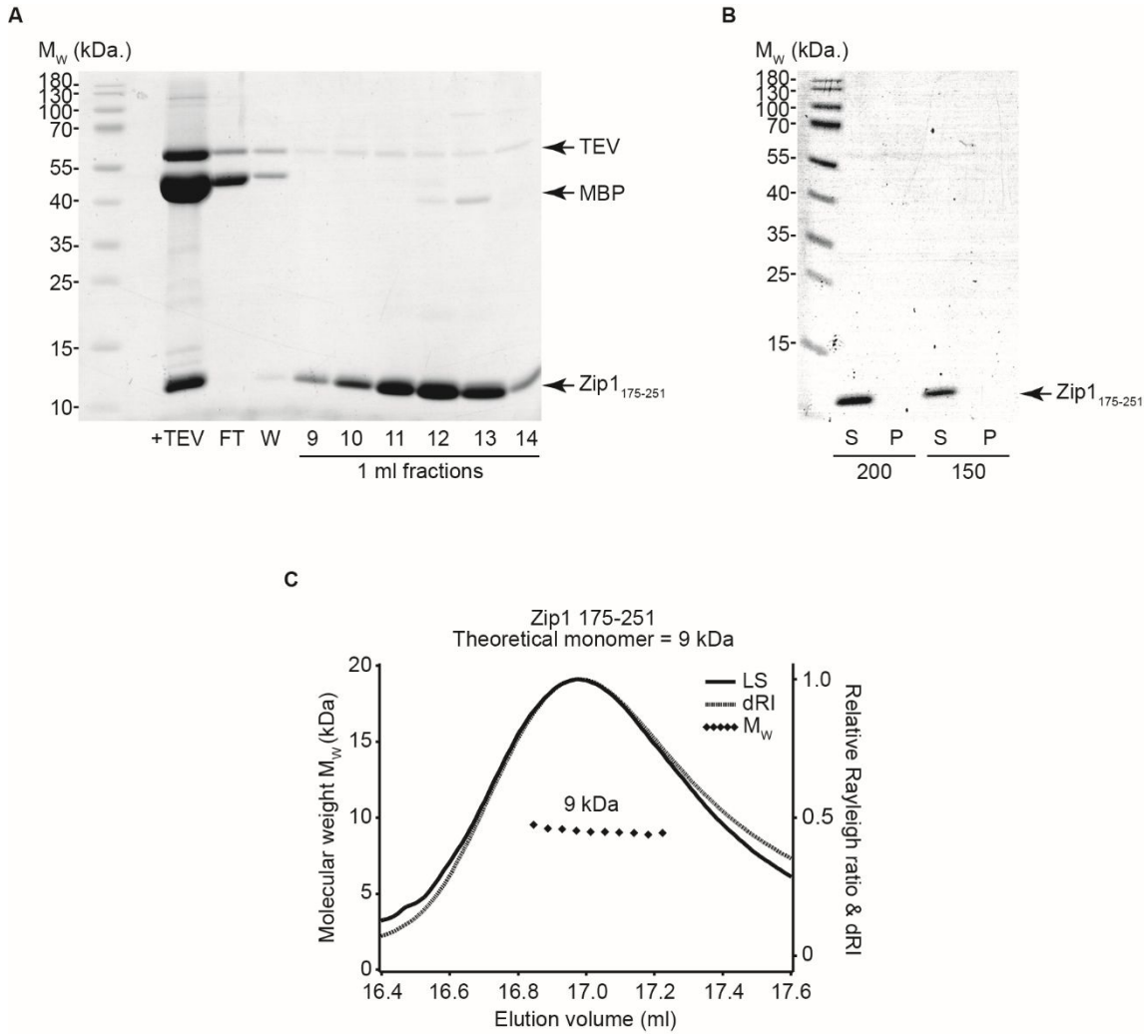


Figure 5.13. Purification and analysis of $Zip1_{175-251}$. Data collected by Jack Hughes under my supervision. (A) SDS-PAGE analysis of $Zip1_{175-251}$ purification by cation exchange chromatography, following TEV protease incubation for tag removal. TEV protease labelled as TEV. (B) Salt precipitation assay. Samples were incubated at 200 and 150 mM KCl with soluble and insoluble material analysed by SDS-PAGE. Arrow indicates $Zip1_{175-251}$. No precipitation was observed. (C) SEC-MALS analysis of $Zip1_{175-251}$. A single species of 9 kDa was formed, corresponding to a monomer.

However, this correlates to the behaviour of SYCP1 101-175, which exists as a monomer in solution but crystallised as an antiparallel tetramer (Dunce *et al.*, 2018). This, together with the fact that a more stable and higher yield of Zip1 was obtained with $Zip1_{175-251}$, crystallisation studies were initiated with this fragment. From the commercial crystallisation trials, no conditions resulted in crystal growth. Many factors, such as protein concentration and pH, may need to be optimised to encourage crystallisation, although it is likely that $Zip1_{1-251}$ is not the optimal fragment for crystallisation and therefore, further constructs were expressed for characterisation.

5.2.4 Stabilising the Zip1 N-terminus through a modest truncation

The Zip1 residues 175-348 comprise a highly helical fragment, with amino acids 252-348 required to stabilise the capacity for the Zip1 N-terminus to bind DNA. Therefore, the instability of the $Zip1_{175-348}$ region was proposed to be due to the construct boundaries, perhaps ending abruptly to disturb a helical turn. To address this, the C-terminal boundary was truncated by 12 residues and characterised both with and without the flexible N-terminus.

The $Zip1_{1-336}$ construct was purified in the same manner as previous constructs, with cleaved material successfully purified for analysis (Figure 5.14 A). Unlike $Zip1_{1-348}$, no precipitation was observed at low salt buffer concentrations, suggesting that this fragment is more stable. Further to this, a much higher yield of material was obtained, reaching a final concentration of 7 mg/ml compared to the 1.3 mg/ml maximum yield of $Zip1_{1-348}$.

To validate that this truncation doesn't alter the secondary structural properties of Zip1, CD wavelength scans were recorded. The CD trace shows high helical content with an estimation of 53% alpha helix, similar to the 54% estimated for $Zip1_{1-348}$ (Figure 5.14 B). This verifies that the truncation maintains the same structural fold as $Zip1_{1-348}$, with the unstructured N-terminus. CD thermal denaturation shows a two stage unfolding profile with an estimated melting temperature of 33°C (Figure 5.14 C). This two-step process is observed with other Zip1 fragments, such as $Zip1_{1-348}$, though for $Zip1_{1-336}$ the first unfolding stage is less severe with a more prominent second unfolding stage. This shift in the melting curve has resulted in a higher melting temperature. Combining this increase in thermal stability, maintenance of secondary structure that adheres to sequence-based prediction and the lack of precipitation at lower KCl concentrations proposes that the C-terminal truncation has stabilised the Zip1 N-terminus.

To test whether the truncation has altered the oligomeric properties of the Zip1 N-terminus, SEC-MALS was used to determine the molecular mass of $Zip1_{1-336}$. This analysis shows that the majority of the material forms a dimer, with a smaller peak likely corresponding to dimeric degradation products (Figure 5.14 D). This further suggests that this fragment is a stable unit and also proposes that the Zip1 N-terminus does form dimers, with the lower stability of other fragments demonstrating oligomeric uncertainty.

As $Zip1_{1-348}$ is shown to bind double-stranded DNA, an EMSA was performed with $Zip1_{1-336}$ to test whether the DNA binding property is retained. $Zip1_{1-336}$ was incubated with dsDNA at increasing protein concentrations and analysed by agarose gel electrophoresis to identify any protein-DNA interactions through a shift in the DNA band. DNA band shifts were observed at the lowest protein concentration tested, 0.5 μM $Zip1_{1-336}$, with a specific protein-DNA species identified at 1 μM $Zip1_{1-336}$ (Figure 5.14 E). This reveals a strong affinity for dsDNA and $Zip1_{1-336}$, particularly when compared to $Zip1_{1-251}$ that displays limited DNA

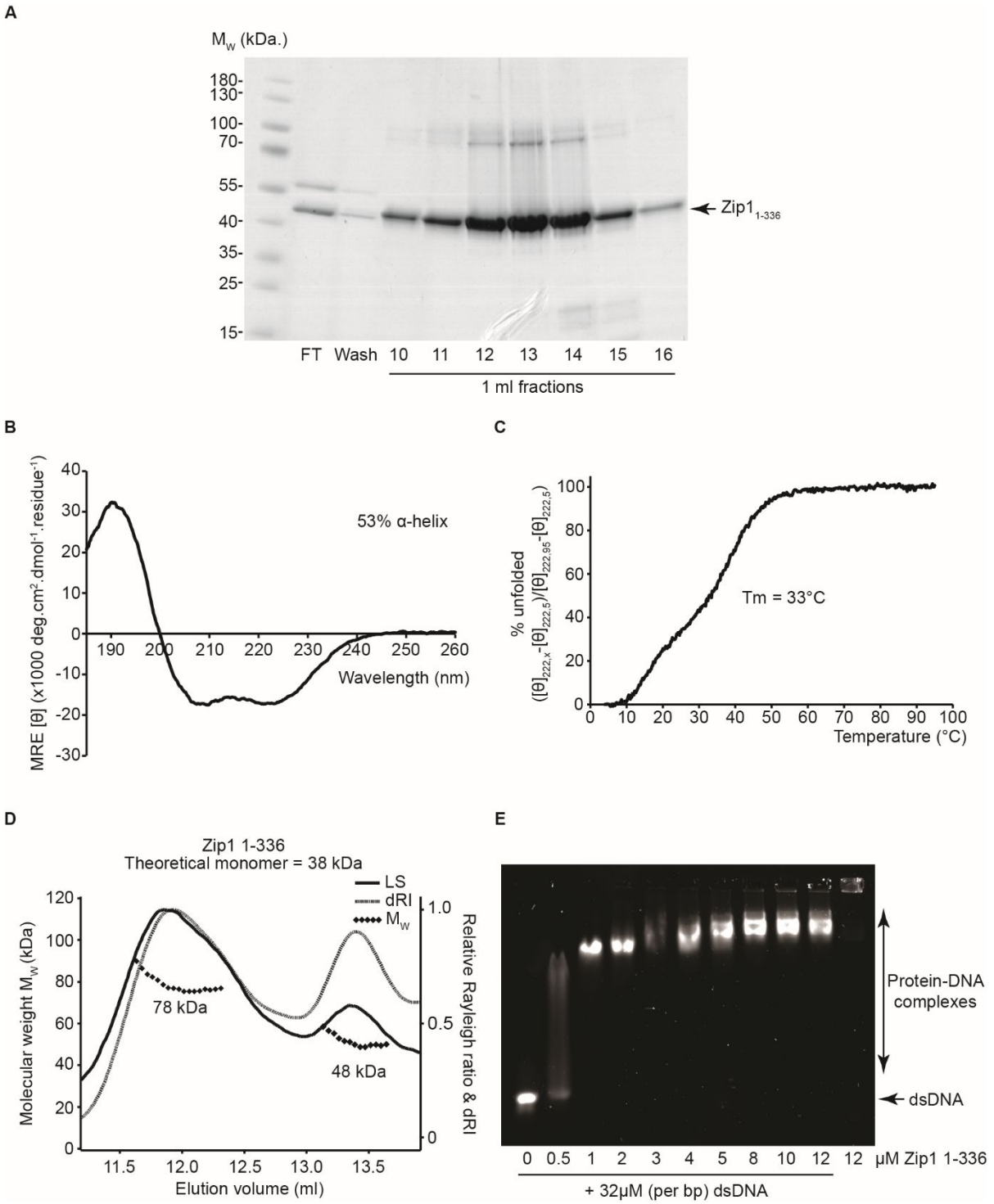


Figure 5.14. Purification and analysis of Zip1$_{1\text{-}336}$. Data collected by Jack Hughes under my supervision. (A) SDS-PAGE analysis of the Zip1$_{1\text{-}336}$ purification by cation exchange chromatography following TEV protease incubation for tag removal. Arrow indicates Zip1$_{1\text{-}336}$. (B) CD scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 53% alpha-helix, 8% β-sheet, 10% turns and 29% unordered. (C) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 33°C, taken as the point at which the sample is 50% unfolded. (D) SEC-MALS analysis of Zip1$_{1\text{-}336}$. A predominant species of 78 kDa was formed, corresponding to dimer formation. A second 48 kDa species likely corresponds to a dimer of degradation products. (E) EMSA experiment with Zip1$_{1\text{-}336}$ and double-stranded DNA (dsDNA). Each lane contains 32 µM per base-pair dsDNA with indicated concentrations of Zip1$_{1\text{-}336}$. A band shift occurs at 0.5 µM Zip1$_{1\text{-}336}$, with specific species formed at 1 µM indicating robust DNA-binding.

binding at 15 μM $Zip1_{1-251}$. The increase in DNA binding between $Zip1_{1-336}$ and $Zip1_{1-251}$ suggests that the critical Zip1 residues to form protein-DNA interactions are within the 252-336 region. This region also comprises the alpha-helical core of the Zip1 N-terminus and appears to be responsible for oligomerisation.

The C-terminal boundary at Zip1 amino acid 336 appears to stabilise the Zip1 N-terminus, and so this region was characterised without the unstructured N-terminus. Recombinant $Zip1_{175-336}$ was purified in the same manner as $Zip1_{1-336}$ to obtain homogenous material for analysis (Figure 5.15 A). Since $Zip1_{175-348}$ is significantly affected by low KCl concentrations, a salt precipitation assay was performed by incubating $Zip1_{175-336}$ with a range of buffer KCl concentrations, 500-100 mM KCl, and then analysing the soluble and insoluble material. The assay shows that $Zip1_{175-336}$ does not precipitate at any KCl concentration tested (Figure 5.15 B). In comparison, $Zip1_{175-348}$ was found to heavily precipitate, thus further indicating that the C-terminal boundary at amino 336 has stabilised the Zip1 N-terminal region for *in vitro* characterisation.

To validate the folding of this region, CD spectra were recorded between 260 and 185 nm to determine the secondary structure composition of the $Zip1_{175-336}$ fragment. A strong helical signal was detected, with deconvolution estimating an alpha-helical content of 84% (Figure 5.15 C). This is similar to the estimated 81% helical content of $Zip1_{175-348}$, demonstrating that the 12 amino acid truncation has not disrupted the folding of Zip1. The Zip1 N-terminal region 1-174 has been demonstrated to be largely unstructured through CD analysis of Zip1 fragments, and this is observed by comparing $Zip1_{1-336}$ and $Zip1_{175-336}$, where the helical contents are determined as 53% and 84% respectively (Table 5.2). This highlights the unstructured nature of the 1-174 region. However, the estimated number of helical residues of these fragments are 180 and 139, respectively. This suggests that the 1-175 region contains some helical propensity or that the N-terminal truncation to amino acid 175 has disrupted some helical formation that occurs in the $Zip1_{1-336}$ fragment. CD thermal denaturation was performed between 5°C and 95°C to study the thermal stability of $Zip1_{175-336}$ and a two-stage unfolding profile was observed. The estimated melting temperatures determined at the midpoint of the unfolding events are 17°C and 34°C, respectively (Figure 5.15 D). $Zip1_{1-336}$ does not have a clear two-state transition suggesting that although the unstructured N-terminal region is present, $Zip1_{1-336}$ appears to be a more stable unit than $Zip1_{175-336}$.

SEC-MALS was used to determine the molecular mass of $Zip1_{175-336}$ and consequently the oligomeric status. This analysis shows that $Zip1_{175-336}$ predominantly exists as a monomer with some of the material forming a dimer, according to approximately 24% of the total sample

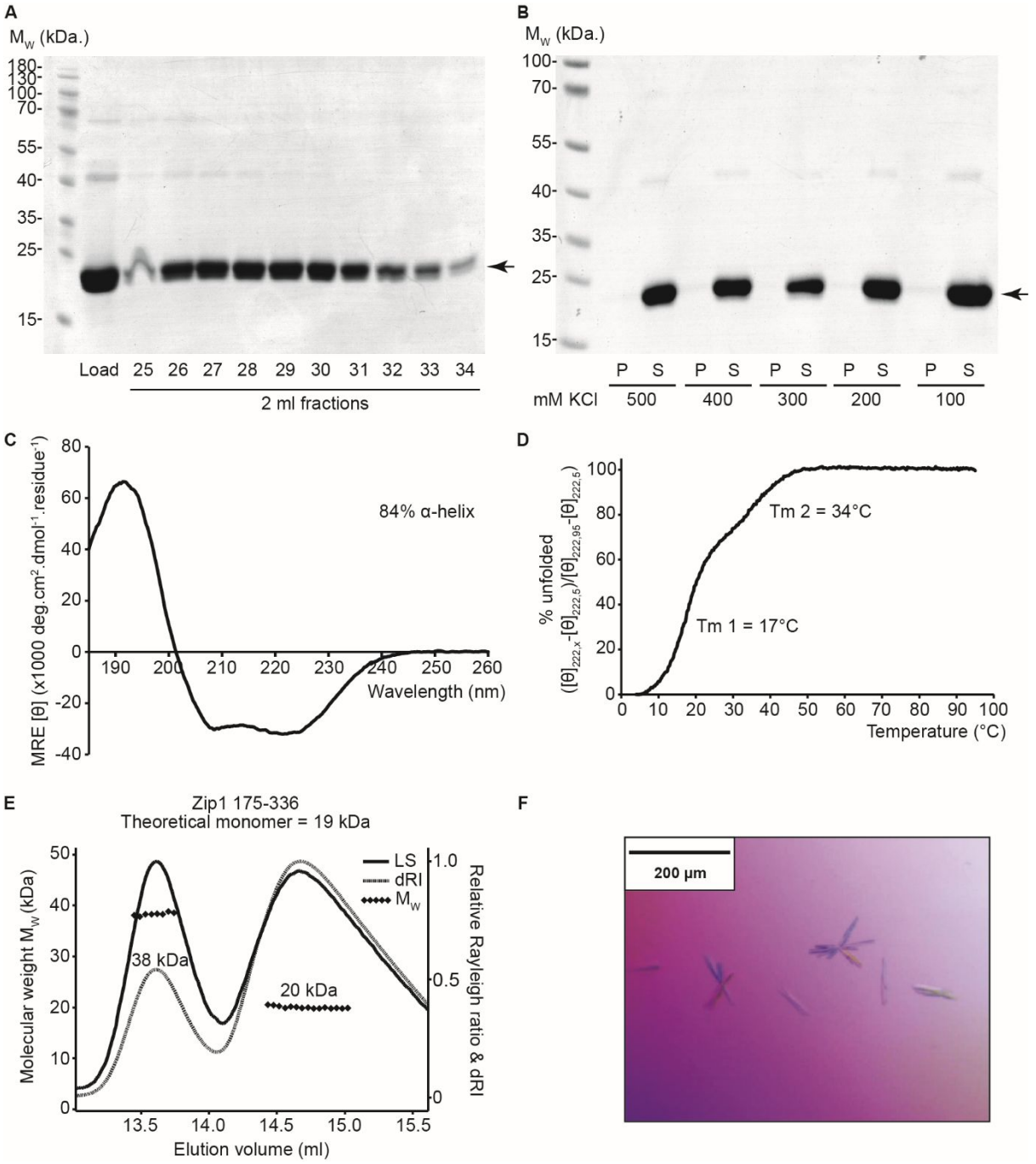


Figure 5.15. Purification and analysis of Zip1$_{175\text{-}336}$. Data collected by Jack Hughes under my supervision. (A) SDS-PAGE analysis of Zip1$_{175\text{-}336}$ purification by size exclusion chromatography, following TEV protease incubation for tag removal. Arrow indicates Zip1$_{175\text{-}336}$. (B) Salt precipitation assay. Samples were incubated at 500 to 100 mM KCl with soluble and insoluble material analysed by SDS-PAGE, showing no observed precipitation. Arrow indicates Zip1$_{175\text{-}336}$. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 84% alpha-helix, 4% β-sheet, 6% turns and 6% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm values estimated at the midpoint of the two respective events: 17°C and 34°C. (E) SEC-MALS analysis of Zip1$_{175\text{-}336}$. A predominant species of 20 kDa was formed, corresponding to a monomer. A larger 38 kDa species corresponds to a dimer. (F) Zip1$_{175\text{-}336}$ crystal hit, in 0.1 M CAPS pH10.5, 40 % v/v MPD.

(Figure 5.15 E). The Zip1$_{1-336}$ fragment exists as a dimer and appears to be a more stable unit due to the higher melting temperature. The Zip1$_{175-336}$ fragment maintains a propensity for dimerisation, but the lower thermal stability may suggest a disruption in oligomerisation.

As Zip1$_{175-336}$ comprises the alpha-helical core of the Zip1 N-terminus with the capacity for dimerisation, crystallisation studies were initiated with this fragment in order to determine the high resolution structure. Only one condition of those tested appeared to have any crystal growth, with the formation of small needles (Figure 5.15 F). As these crystals were too small to harvest, optimisation screens were performed to try and increase crystal size and quality by varying pH, MPD concentration and drop sizes. However in these optimisation screens, no crystal growth was observed. These crystals in the same condition were also observed in crystallisation trials for different proteins, suggesting that these may be salt crystals. In the absence of robust Zip1 crystals, the construct boundaries may need to be optimised further to isolate a more stable structural region, perhaps through the inclusion of further N-terminal residues due to the apparent increased stability of Zip1$_{1-336}$. Another approach would be to perform further crystallisation trials at 4°C due to the low melting temperature determined for Zip1$_{175-336}$.

Zip1 region	1-348	1-336	1-251	175-348	175-336	175-251
Number of residues	352	340	255	178	166	81
Theoretical pI	5.69	5.93	5.26	9	9.25	9.01
% α-helix	54	53	21	81	84	n.d.
% β-sheet	8	8	21	5	4	n.d.
% turns	13	10	25	7	6	n.d.
% unordered	25	29	33	7	6	n.d.
Number of helical residues	190	180	54	144	139	n.d.
Tm (°C)	16 and 34	33	16	16 and 34	17 and 34	n.d.
Oligomer state	Monomer- Dimer	Dimer	n.d.	Monomer	Monomer- Dimer	Monomer

Table 5.2. Summary of Zip1 constructs and biophysical analysis.

5.2.5 Expression of longer Zip1 constructs

The central region of Zip1 has high alpha helical prediction that spans the majority of the molecule, and the role of Zip1 as the yeast SC transverse filament protein implies it exists as an elongated helical structure to span the distance between the CE and the chromosome axis. Zip1 is thought to behave similarly to SYCP1, due to their comparative roles within yeast and humans, and it has been shown that the large helical SYCP1 region forms a dimeric coiled-coil that forms recursive tetramer structures between neighbouring SYCP1 dimers to form the preliminary SC lattice (Dunce *et al*., 2018). Therefore, longer portions of Zip1 were designed for bacterial expression and characterisation to investigate the mechanistic assembly of Zip1.
Four Zip1 constructs were tested: the full length sequence (1-875), the alpha-helical central region (175-748) and the central region with either unstructured termini (1-748 and 175-875). These sequences were expressed with an N-terminal His_6-MBP tag to encourage solubility and bacterial expression of these constructs was tested by comparing the bacterial cultures before and after protein expression induction. All four tagged sequences were successfully expressed upon induction, which was determined by the appearance of the appropriate bands in the induced sample compared to the un-induced sample (Figure 5.16 A).

Due to the successful expression of the full length Zip1 sequence in bacteria, the cells were lysed by sonication and the lysate was clarified through centrifugation for analysis of the soluble and insoluble material. By comparison of the pelleted material and the supernatant following lysis and centrifugation, the majority of the tagged $Zip1_{1\text{-}875}$ was found to be in the soluble fraction, with a small proportion of the material observed in the pellet (Figure 5.16 B). As the full sequence was found to be largely soluble, the supernatant was applied to sequential Ni-NTA and amylose affinity chromatography resins to purify the tagged Zip1. However, during these chromatography steps, the $Zip1_{1\text{-}875}$ was found to be heavily degraded from the affinity tag and then several cleaved Zip1 bands were identified (Figure 5.16 B). This suggests that the full length Zip1 may be need optimisation for bacterial expression and purification, or that it may require eukaryotic expression for post-translational modifications.

As the N-terminal Zip1 region 1-174 has experimentally been found to be largely unstructured, this region was removed to see if the Zip1 region spanning residues 175-875 could be purified. By comparison of the insoluble and soluble fractions following cell lysis, it was observed that a large proportion of the tagged $Zip1_{175\text{-}875}$ was soluble, with a small amount in the insoluble fraction (Figure 5.16 C). The clarified lysate was applied to Ni-NTA and amylose affinity resins and likewise to the full length Zip1, heavy degradation was observed during the chromatography stages, with the cleavage of the affinity tag and multiple $Zip1_{175\text{-}875}$ bands identified (Figure 5.16 C).

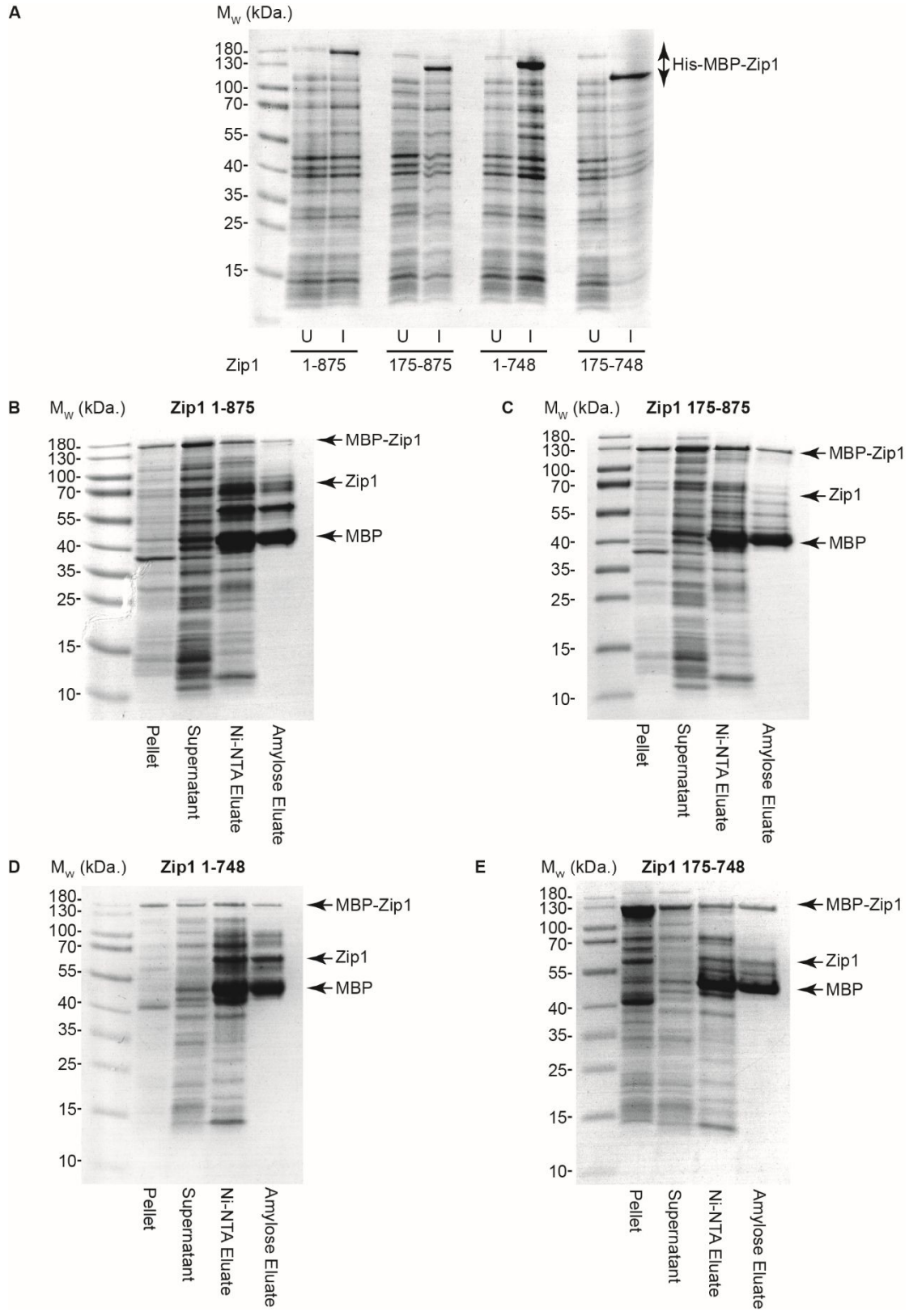


Figure 5.16. SDS-PAGE analysis of long Zip1 constructs. (A) Bacterial expression test of His_6-MBP-Zip1 longer constructs: 1-875, 175-875, 1-748 and 175-748. U= un-induced sample, I= induced sample after IPTG induction to initiate protein expression. (B-E) Analysis of protein solubility and affinity chromatography steps. Pellet and supernatant samples were taken after cell lysis and centrifugation. (B) His_6-MBP-$Zip1_{1-875}$. (C) His_6-MBP-$Zip1_{175-875}$. (D) His_6-MBP-$Zip1_{1-748}$. (E) His_6-MBP-$Zip1_{175-748}$.

Although residues 1-174 are determined to be predominantly unstructured, the Zip1$_{1-336}$ fragment was determined to be more stable than Zip1$_{175-336,}$ proposing that the inclusion of this region may increase stability and that the removal of the predicted unstructured C-terminal amino acids 749-875 may stabilise the longer Zip1 fragment. Therefore, the Zip1$_{1-748}$ region was expressed and the solubility of the fusion material analysed. From comparison of the insoluble and soluble material, relatively equal amounts of His-MBP-Zip1$_{1-748}$ were observed in the pelleted material and soluble supernatant, suggesting that this construct is less soluble without the C-terminus (Figure 5.16 D). Some soluble material was observed, therefore the affinity chromatography steps were performed to test whether any non-degraded material could be purified. As found with the Zip1$_{1-875}$ and Zip1$_{175-875}$, degradation of the affinity tag from Zip1 was observed (Figure 5.16 D).

As the Zip1 residues 175-748 are predicted to form the alpha helical core without unstructured N- and C-terminal regions, Zip1$_{175-748}$ was expressed to determine if any intact material could be obtained. Following cell lysis and centrifugation, the majority of the Zip1 material was found to be in the insoluble fraction (Figure 5.16 E). However, by comparing the background bands of the insoluble and soluble fractions, the position and intensity of the bacterial bands are similar, suggesting that only partial cell lysis occurred. As some of the material was soluble, the Ni-NTA and amylose affinity chromatography steps were performed. Similar to the other long Zip1 fragment tested, the affinity tag was cleaved during the purification steps with degradation of the Zip1 sequence (Figure 5.16 E).

From the expression tests and initial purification steps of the longer Zip1 regions, it was found that these regions were successfully expressed in *E. coli* but were heavily degraded upon purification. This could suggest that expression conditions, such as temperature, and buffer conditions for lysis may need optimisation to increase solubility and decrease degradation of these Zip1 portions. Another consideration is that an eukaryotic expression system may be required to obtain properly folded and stable material due to post-translational modifications that cannot be provided by a prokaryotic system. As Zip1 is a yeast protein, it stands to reason that a yeast expression system could be used. Another option would be to utilise another eukaryotic system, such as insect cell expression.

Another consideration is that Zip1 may require a binding partner for stabilisation. No direct interaction was detected between Zip1 and the central element Ecm11 and Gmc2 proteins through Y2H screening. However, this has not been tested biochemically and it is known that the N-terminal region of Zip1 is localised at the central element, suggesting that Zip1 may bind directly or indirectly to the central element (Humphryes *et al.*, 2013). This hypothesis is reinforced by the identification of the human SYCP1 transverse filament interaction with the

SYCE3 central element component (Hernandez-Hernandez *et al*., 2016) (see Chapter 3). Therefore, studies were initiated to characterise the central element proteins Ecm11 and Gmc2.

5.2.6 Expression and characterisation of Ecm11 and Gmc2 central element proteins

The yeast central element is comprised of the Ecm11 and Gmc2 proteins that are shown to co-localise *in vivo* (Humphryes *et al*., 2013) and are found to interact by yeast 2-hybrid screening by the C-terminal regions of both proteins. Sequence analysis of Ecm11 shows that the latter half of the molecule is highly conserved with high alpha helical prediction. The Y2H screening showed that the Ecm11 amino acids 206-302, corresponding to the conserved helical region, maintained the interaction with Gmc2 and the self-interaction. Therefore, this region was fused to a His_6-MBP affinity tag for bacterial expression and purification.

Following bacterial expression and cell lysis, Ni-NTA and amylose affinity chromatography was used to purify the His-MBP-$Ecm11_{206-302}$ fragment. Some degradation of the affinity tag was observed, but this was purified away from the fusion protein by anion exchange chromatography (Figure 5.17 A). The $Ecm11_{206-302}$ was unstable upon tag cleavage, therefore the fusion protein was used for analysis.

The Ecm11 self-interaction observed by Y2H suggests that Ecm11 can form oligomers. Therefore, SEC-MALS was used to determine the oligomeric state of the purified His-MBP-$Ecm11_{206-302}$. This analysis reveals that a small amount of material exists as a monomer, with the rest of the material forming higher-molecular weight species (Figure 5.17 B). As the SC central element assembles between the entire lengths of homologous chromosome pairs, it could be hypothesised that the central element proteins undergo higher-order assembly to achieve this. It is also known that human central element proteins SYCE2-TEX12 undergo filament formation (Davies *et al*., 2012) and that Ecm11 and Gmc2 are hypothesised to act analogously to SYCE2-TEX12 (see Chapter 4). Another possibility is that the large molecular weight species observed are due to aggregation of Ecm11 suggesting that Ecm11 is unstable without a binding partner, further implied by the instability of Ecm11 upon tag cleavage.

Through sequence analysis and structural prediction Gmc2 is found to be highly conserved with high alpha helical prediction and short stretches of coiled-coil prediction towards the C-terminus. The full length Gmc2 sequence was fused to a His_6-MBP affinity tag for bacterial expression and purification. His-MBP-$Gmc2_{1-188}$ was well expressed, with a large yield of fusion protein obtained through sequential Ni-NTA and amylose affinity chromatography followed by anion exchange chromatography (Figure 5.18 A).

The fusion material was analysed by SEC-MALS to determine the oligomeric status based upon molecular mass calculation. The predominant species is monomeric with a

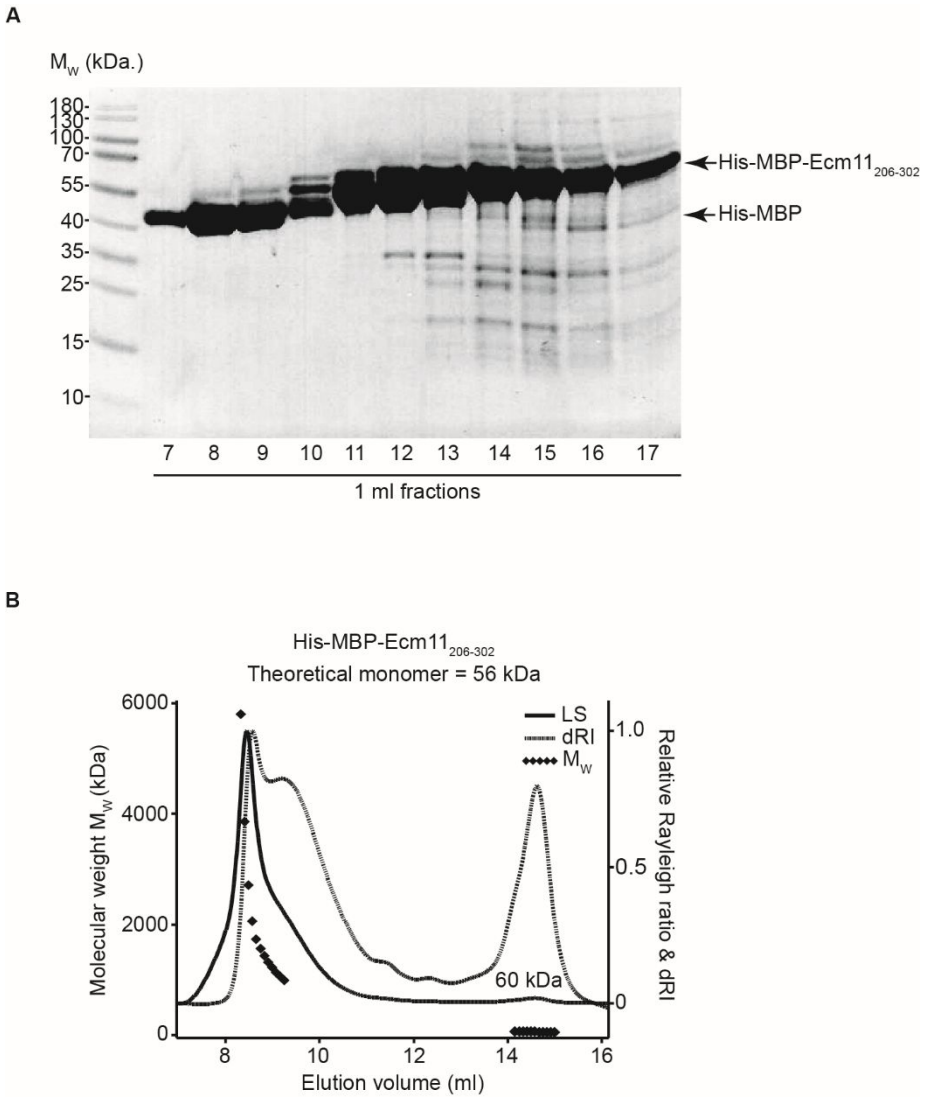


Figure 5.17. Purification and analysis of His-MBP-ECM11$_{206-302}$. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE analysis of His-MBP-ECM11$_{206-302}$ purification by anion exchange chromatography. (B) SEC-MALS analysis of His-MBP-ECM11$_{206-302}$. The 60 kDa peak corresponds to a monomer, with the higher molecular weight species forming large assemblies.

propensity for dimer species observed (Figure 5.18 B). This corresponds to the Y2H analysis whereby no self-interaction was detected between Gmc2.

However, as some dimeric species are formed, the orientation of the Gmc2 chains were investigated by SEC-SAXS analysis to determine the inter-atomic distances between the MBP affinity tags. Guinier analysis of the His-MBP-Gmc2 dimer peak determines the radius of gyration to be 82 Å, with the radius of gyration of an MBP monomer to be 26 Å (Figure 5.18 C). The probability of interatomic distances within the dimer peak were analysed by the distance distribution plot, allowing the determination of the maximum dimension within the sample and the observation of structural features. The maximum distance determined within

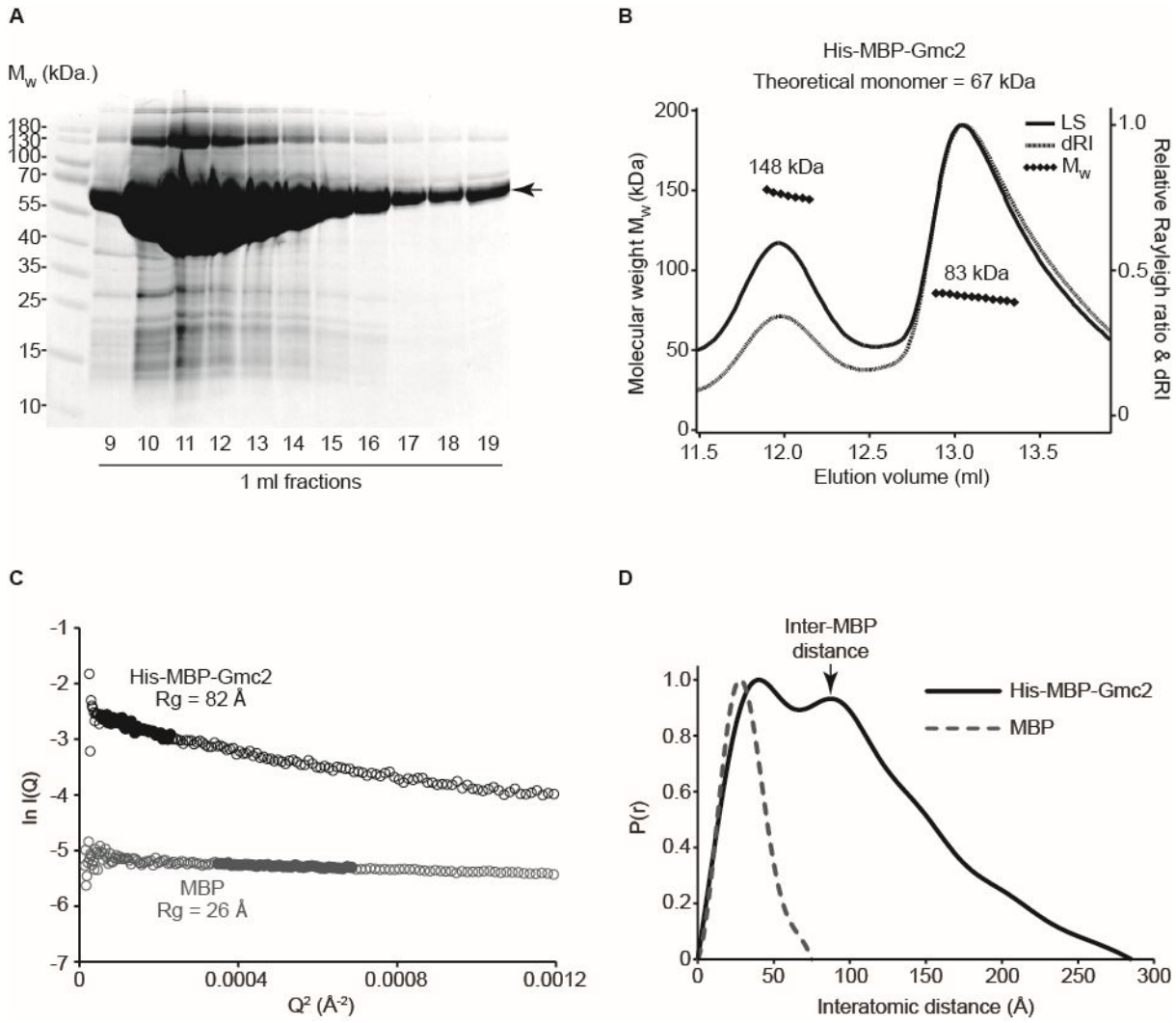


Figure 5.18. Purification and structural analysis of His-MBP-Gmc2. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the His-MBP-Gmc2 purification by anion exchange chromatography. Arrow indicates His-MBP-Gmc2. (B) SEC-MALS analysis. The 83 kDa peak corresponds to a monomer, with the 148 kDa peak corresponding to a dimer. (C-D) SEC-SAXS analysis. (C) Guinier fit to determine the radius of gyration (Rg) as 82 Å (black) with the control MBP fit as 26 Å (grey). Clear circles are the data, solid circles represent the region used for the Guinier fit. (D) Paired real-space distribution plot. His-MBP-Gmc2 has a maximum dimension of 285 Å, with an inter-MBP peak at 88 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

the dimer His-MBP-Gmc2 species is 285 Å, compared to the 75 Å maximum distance observed for the MBP monomer (Figure 5.18 D). Combining the maximum distance observed and the tailing shape of the peak implies the dimer species forms an elongated structure. Another feature of the curve is the large peak at 88 Å, correlating to the distance between MBP molecules (Figure 5.18 D). This distance correlates to the inter-MBP peak observed from parallel helices, indicating that the dimeric Gmc2 chains are in a parallel arrangement with the N-termini in close proximity.

Unlike Ecm11$_{206-302}$, Gmc2 was stable upon tag removal, allowing the purification of cleaved Gmc2 by anion exchange chromatography for analysis (Figure 5.19 A). To analyse the secondary structure of Gmc2, CD wavelength scans were measured. These measurements indicate a helical composition of Gmc2, with deconvolution of the data estimating an alpha helical content of 51% (Figure 5.19 C). To assess Gmc2 thermal stability, CD was used to perform thermal denaturation between 5°C and 95°C with the measurement of the 222 nm helical signal. The melting curve displays non-cooperative unfolding of Gmc2, indicating the gradual unwinding of alpha helices, with an estimated melting temperature of 45°C (Figure 5.19 D).

The oligomer state of Gmc2 was investigated through molecular mass determination by SEC-MALS. This analysis shows that the majority of the material exists as a monomer, with a small amount of Gmc2 forming dimeric species (Figure 5.19 B). This correlates with the oligomeric status of the His-MBP-Gmc2, whereby the sample is predominantly monomeric with the capacity for dimerisation. This further relates to the Y2H screening, where no self-interactions were detected for Gmc2.

As Ecm11 is not stable without the affinity tag, and forms large molecular species that may be caused by aggregation as a fusion protein, it is reasonable to propose that Ecm11 is not stable in the absence of its binding partner Gmc2, as identified through Y2H mating.

5.2.7 In vitro characterisation of the Ecm11-Gmc2 complex

The Y2H screening between Ecm11 and Gmc2 showed that Ecm11 self-interacts and also interacts with Gmc2. These interactions are sustained with the C-terminal region of Ecm11, corresponding to amino acids 206-302, which is a highly conserved and predicted alpha-helical region. The interaction between Ecm11, and also Ecm11$_{206-302}$, still take place with the C-terminal portion of Gmc2 (residues 96-188). The Ecm11$_{206-302}$ and Gmc2$_{96-188}$ fragments were co-expressed in bacteria to investigate whether a direct interaction would be identified between the central element proteins as hypothesised.

The Ecm11$_{206-302}$ was fused to a His$_6$-tag and Gmc2$_{96-188}$ was fused to an MBP-tag for expression and purification by sequential Ni-NTA and amylose chromatography resins, which would also only result in the purification of both proteins if they interacted. His-Ecm11$_{206-302}$ and MBP-Gmc2$_{96-188}$ were indeed co-purified by affinity chromatography, indicating that the fusion proteins interact *in vitro*. The fusion protein complex was further purified by anion

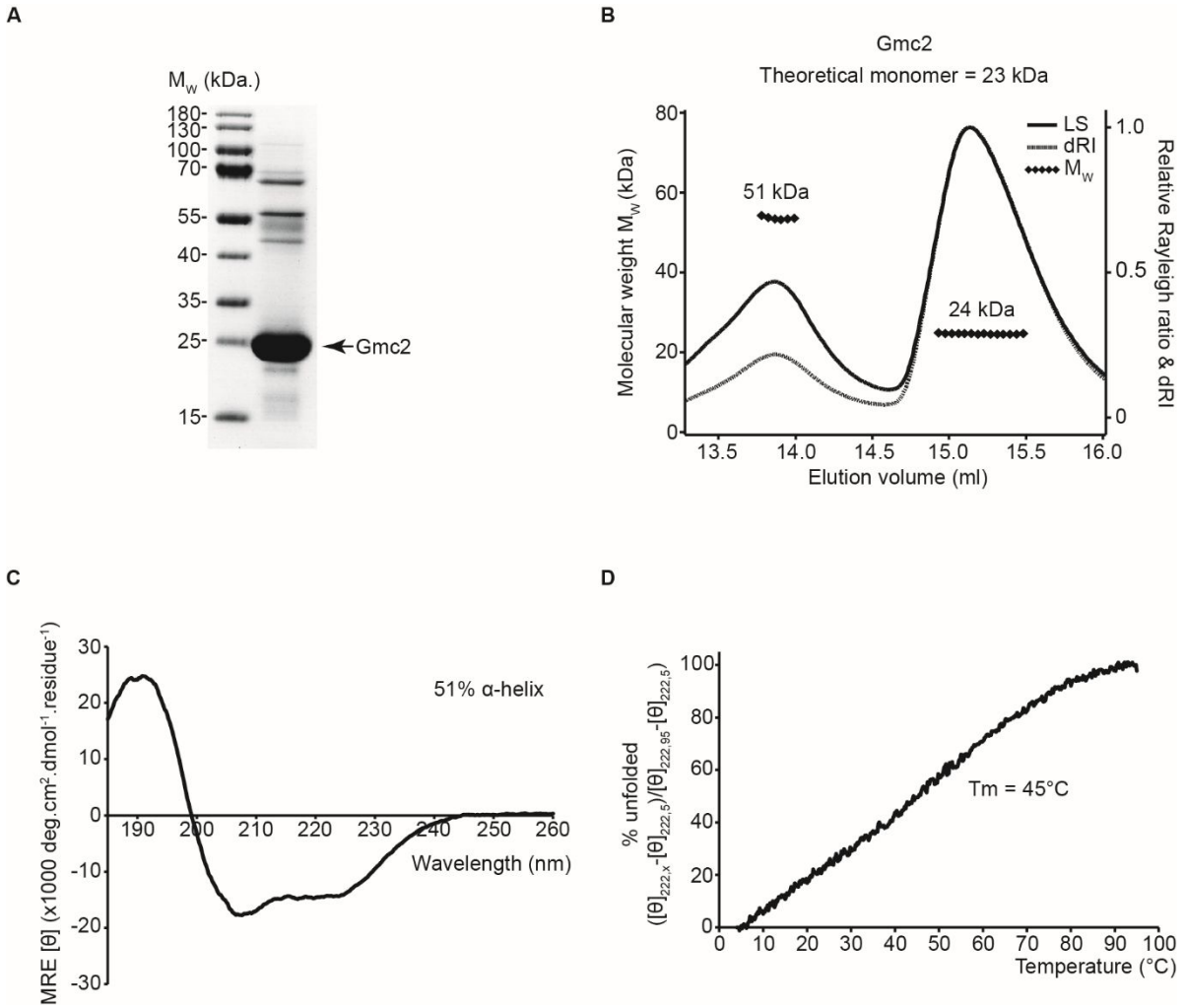


Figure 5.19. Purification and structural analysis of Gmc2. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the final Gmc2 purified product. Arrow indicates Gmc2. (B) SEC-MALS analysis. The 24 kDa peak corresponds to a monomer, with the 51 kDa peak corresponding to a dimer. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 51% alpha-helix, 16% β-sheet, 12% turns and 21% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 45°C, taken as the point at which the sample is 50% unfolded.

exchange chromatography, where the two fusion proteins were found to co-elute across the salt elution gradient (Figure 5.20 A).

The fusion protein complex, His-Ecm11$_{206-302}$-MBP-Gmc2$_{96-188}$, was analysed by SEC-MALS to test whether the proteins would remain as a complex by size-exclusion chromatography and to determine the molecular mass of the protein species identified. This analysis shows that the Ecm11 and Gmc2 fusion proteins were found to co-elute over the isocratic gradient, indicating that they form a robust protein complex. The His-Ecm11$_{206-302}$-

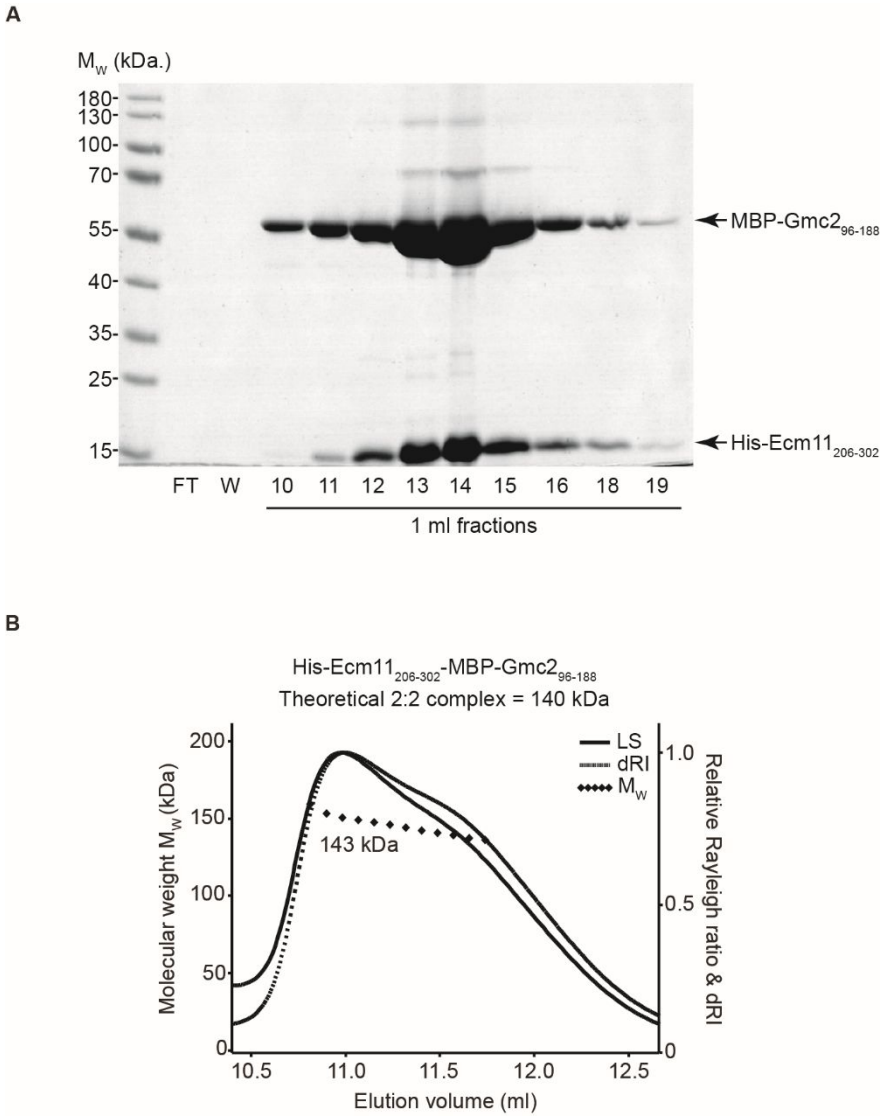


Figure 5.20. Purification and oligomer analysis of His-Ecm11$_{206-302}$-MBP-Gmc2$_{96-188}$. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the purification by anion exchange chromatography. (B) SEC-MALS analysis. The 143 kDa peak corresponds to a 2:2 complex.

MBP-Gmc2$_{96-188}$ peak had a calculated molecular mass of 143 kDa (Figure 5.20 B). The theoretical molecular mass of a 2:2 complex of His$_6$-Ecm11$_{206-302}$ and MBP-Gmc2$_{96-188}$ is estimated to be 140 kDa, therefore indicating that the Ecm11 and Gmc2 proteins interact to form a 2:2 complex. The His-Ecm11$_{206-302}$-MBP-Gmc2$_{96-188}$ complex was incubated with TEV protease to remove the affinity tags. However, upon tag removal the proteins were de-stabilised and heavily aggregated on subsequent chromatography columns. Therefore, the cleaved material could not be obtained for analysis and suggests that these fragments are unsuitable for biophysical characterisation.

The Gmc2 amino acids 37-188 were also found to interact with Ecm11$_{206-302}$ by Y2H and includes more of the predicted helical residues than the 96-188 fragment, which may be required to stabilise the Ecm11-Gmc2 complex. The His-Ecm11$_{206-302}$ was co-expressed with MBP-Gmc2$_{37-188}$ and then fusion proteins were co-purified by Ni-NTA and amylose affinity chromatography, demonstrating that these fragments do interact *in vitro*. The fusion protein complex was also maintained throughout anion exchange chromatography, and the affinity tags were then removed by TEV protease incubation. The cleaved material was purified by further chromatography steps to obtain purified cleaved material for analysis (Figure 5.21 A). The purification of the cleaved Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex implies that the inclusion of Gmc2 residues 37-95 has stabilised the complex, although these residues are not necessary for the Ecm11-Gmc2 interaction.

To examine the secondary structure composition of the Ecm11$_{206-302}$-Gmc2$_{37-188,}$ far-UV CD spectra were recorded. The data shows a high helical signal with an estimated alpha helical content of 83% (Figure 5.21 C). CD thermal denaturation was performed to produce a thermal melting curve between 5°C and 95°C for the Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex. The curve shows a cooperative two-stage unfolding profile with an estimated melting temperature of 36°C (Figure 5.21 D).

To determine the oligomeric state of the Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex, SEC-MALS was performed to calculate the molecular mass of the protein species. The data shows that the complex forms a 53 kDa species (Figure 5.21 B). The theoretical molecular weight for a 2:2 Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex is 60 kDa, suggesting the formation of either a 1:2 (49 kDa) or 2:2 (60 kDa) complex, with uncertainty arising due to the low molecular weight of Ecm11$_{206-302}$ and possible dissociation of the complex across the peak.

To study the low resolution structural properties of Ecm11$_{206-302}$-Gmc2$_{37-188}$ in solution, SEC-SAXS was performed. Through analysis of the Ecm11$_{206-302}$-Gmc2$_{37-188}$ ambiguous complex peak (1:2 or 2:2), the radius of gyration (Rg) was determined as 63 Å, indicating an elongated conformation due to the size of the species (Figure 5.22 A). Due to the elongated nature of the complex, the radius of gyration of the cross-section (Rc) could be determined and was found to be 19 Å (Figure 5.22 B). A four helical bundle has an approximate Rc of 10 Å, so this suggests that Ecm11$_{206-302}$-Gmc2$_{37-188}$ forms a bulkier structure, potentially through folding back of chains rather than a completely linear structure. The real space distance distribution curve shows an extended shape and estimates the maximum distance for Ecm11$_{206-302}$-Gmc2$_{37-188}$ to be 275 Å (Figure 5.22 C). This further demonstrates that Ecm11$_{206-302}$-Gmc2$_{37-188}$ forms an elongated structure.

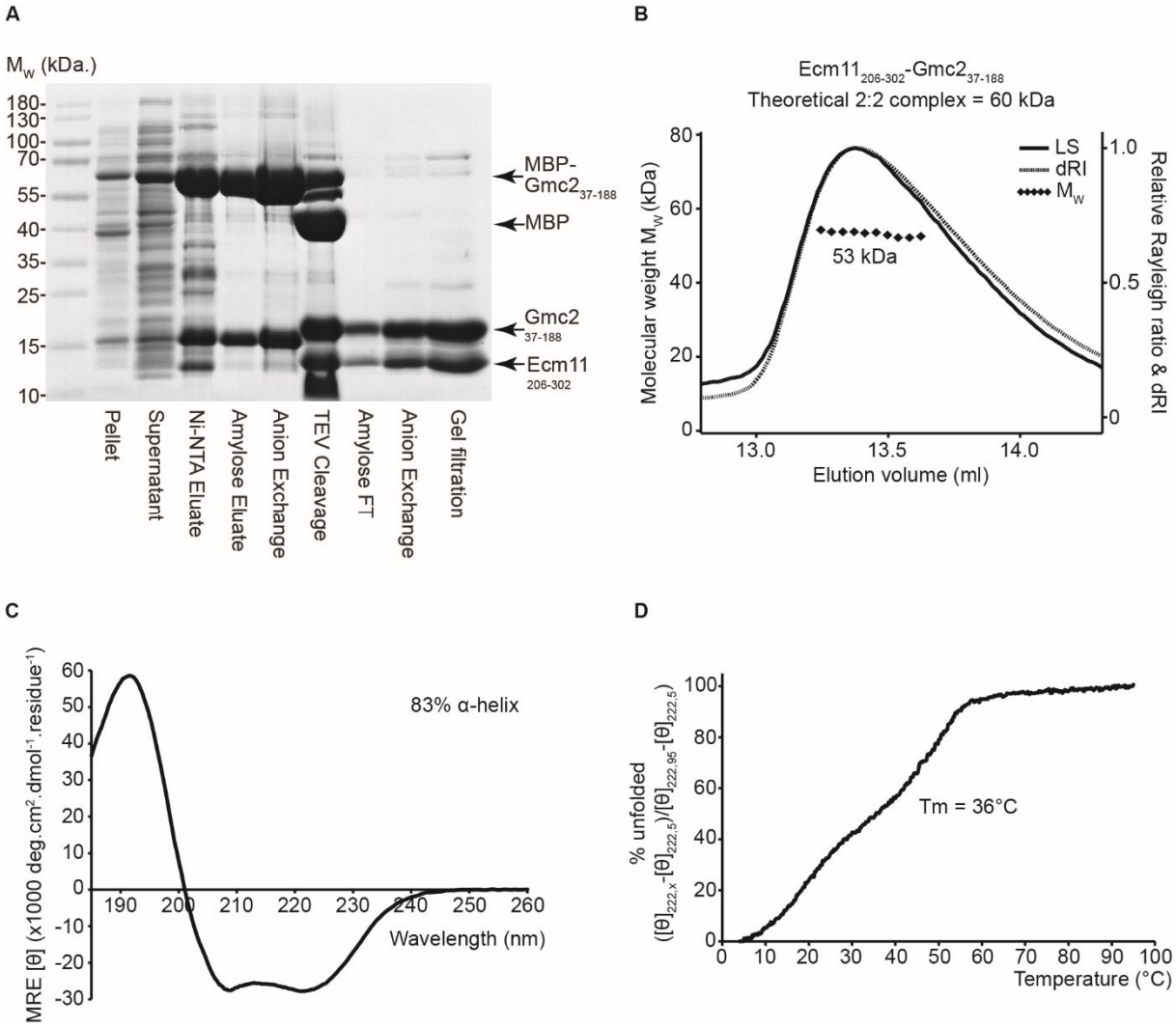


Figure 5.21. Purification and structural analysis of Ecm11$_{206-302}$-Gmc2$_{37-188}$. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the purification summary of the complex through Ni-NTA and amylose affinity chromatography, anion exchange chromatography, tag removal by TEV protease incubation and subsequent anion exchange and size exclusion chromatography. (B) SEC-MALS analysis showing a single 53 kDa peak. A theoretical 2:2 complex has an estimated 60 kDa molecular weight suggesting a 1:2 or 2:2 complex formation. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 83% alpha-helix, 4% β-sheet, 3% turns and 10% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 36°C, taken as the point at which the sample is 50% unfolded.

To verify the number of Gmc2$_{37-188}$ molecules within Ecm11$_{206-302}$-Gmc2$_{37-188}$, the fusion complex (Figure 5.23 A) was analysed by SEC-MALS. The theoretical molecular weight of a 2:2 His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ complex is 154 kDa, and the experimental molecular mass was determined to be 134 kDa (Figure 5.23 B). This confirms that two Gmc2 molecules exist in the complex, with the theoretical MBP-Gmc2$_{37-188}$ molecular weight estimated to be 62 kDa. However, the number of Ecm11 molecules remains uncertain.

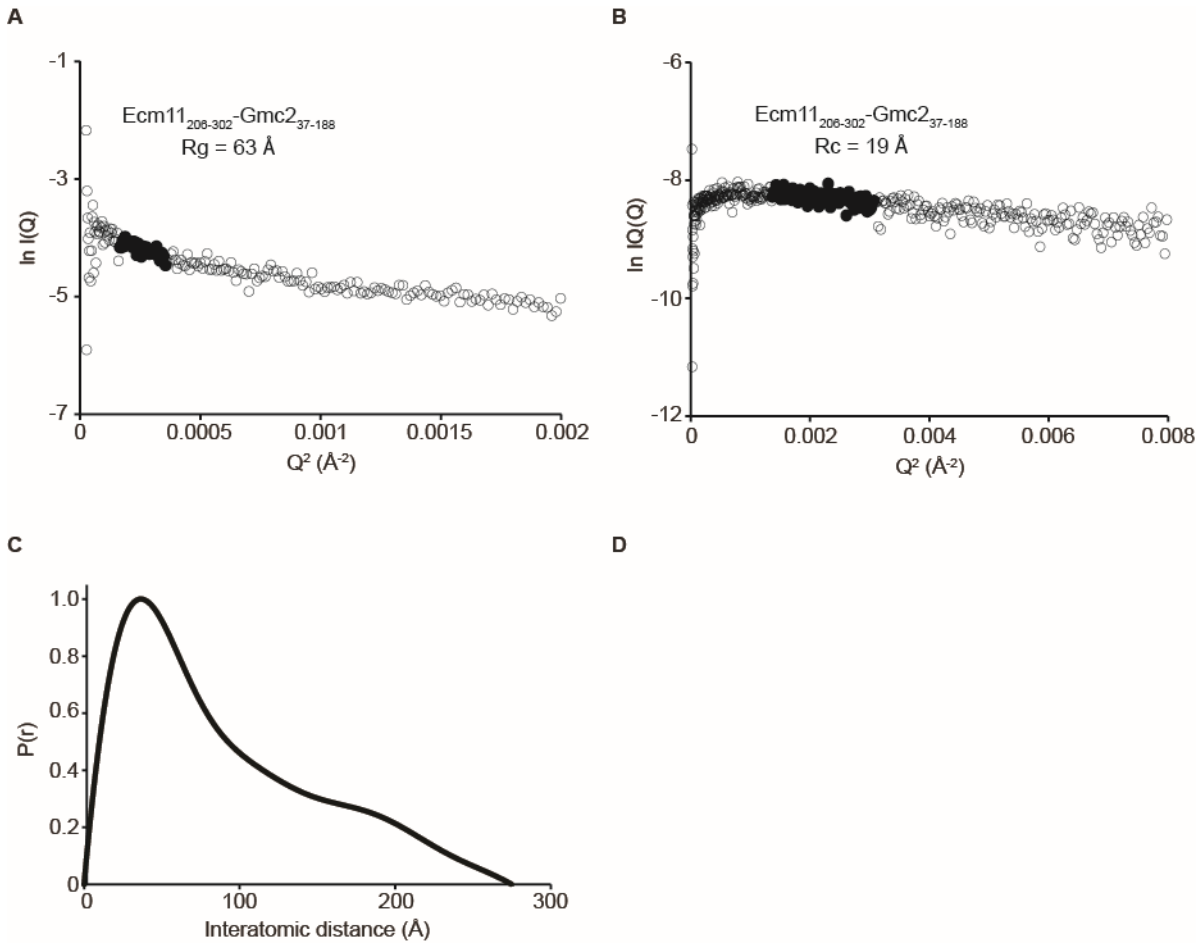


Figure 5.22. SEC-SAXS analysis of Ecm11$_{206-302}$-Gmc2$_{37-188}$. (A) Guinier fit to determine the radius of gyration (Rg) as 63 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 19 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. Ecm11$_{206-302}$-Gmc2$_{37-188}$ has a maximum dimension of 275 Å.

SEC-SAXS was used to study the structural properties of His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$. Guinier analysis determined the Rg to be 104 Å, which is larger than the cleaved complex due to the addition of the two MBP molecules (Figure 5.23 C). The paired distance distribution curve shows the same extended shape, with the maximum distance estimated to be 470 Å (Figure 5.23 D). The increase in length, compared to the cleaved complex, is due to the addition of the two MBP molecules. Another peak is observed at 69 Å, which correlates to the distance between the MBP molecules. This distance suggests that the Gmc2$_{37-188}$ chains are in a parallel arrangement due to the proximity of the N-terminal globular affinity tags, and this also further validates that two MBP-Gmc2$_{37-188}$ molecules are present within the complex.

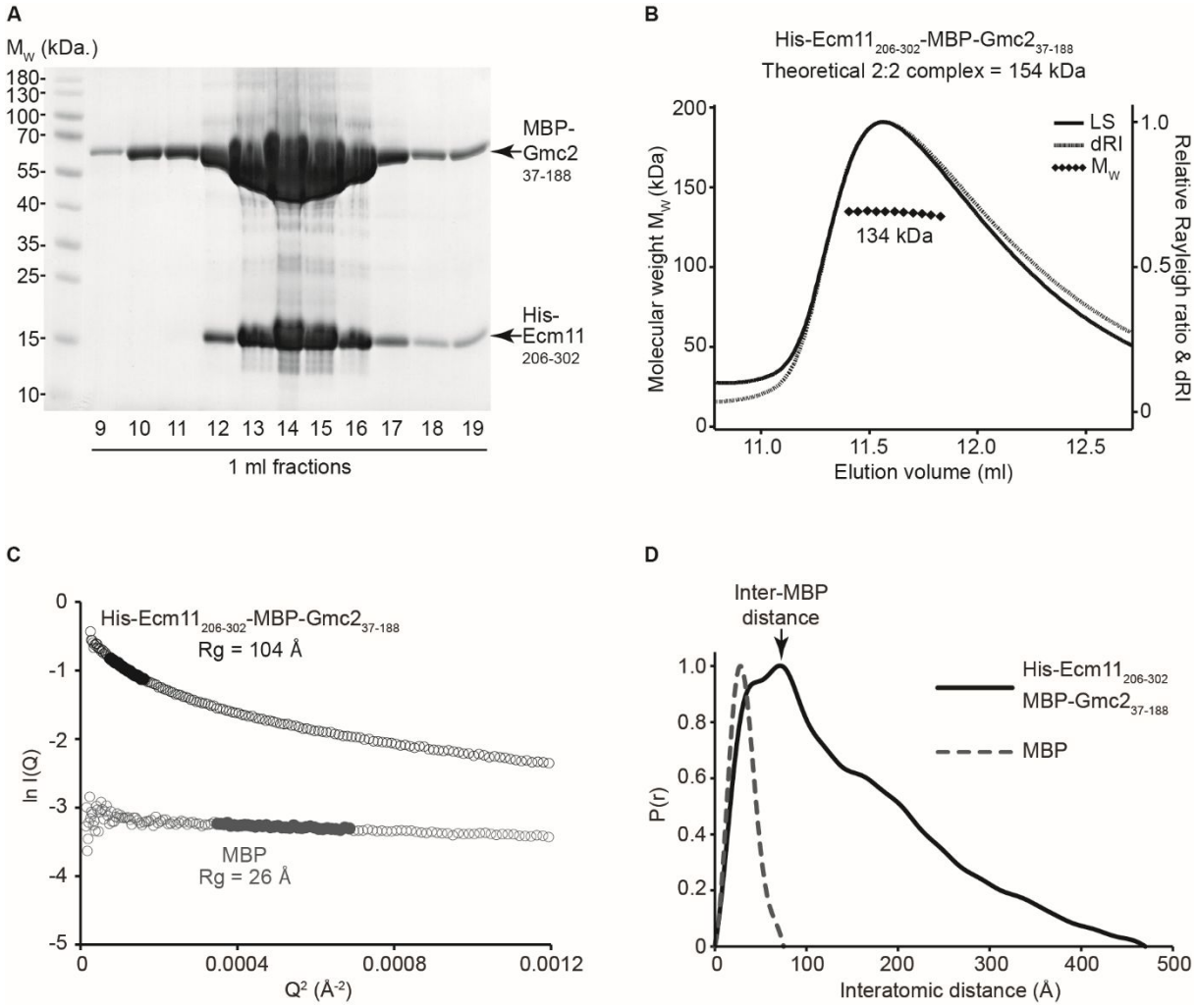


Figure 5.23. Purification and structural analysis of His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (B) SEC-MALS analysis showing a single 134 kDa peak. A theoretical 2:2 complex has an estimated 154 kDa molecular weight suggesting a 1:2 or 2:2 complex formation. (B) Guinier fit to determine the radius of gyration (Rg) as 104 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ has a maximum dimension of 470 Å, with an inter-MBP peak at 69 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

The number of Ecm11 molecules with the Ecm11-Gmc2 remains unclear. To address this, the Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex was expressed with both proteins fused to an MBP affinity tag to increase the mass of Ecm11, allowing for a less ambiguous oligomer determination. The MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ complex was co-purified through Ni-NTA and amylose

affinity chromatography followed by anion exchange chromatography (Figure 5.24 A). This demonstrates that the presence of the large globular MBP tag on both proteins has not disrupted the interaction. The purified fusion complex was analysed by SEC-MALS to calculate the molecular weight of the MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ complex.

This analysis determines the complex to be 209 kDa (Figure 5.24 B), with the estimated 2:2 MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ complex molecular weight to be 232 kDa. The theoretical molecular mass for a 1:2 and a 2:2 complex is 180 kDa and 232 kDa respectively and therefore, although the experimental mass is closer to a theoretical 2:2 than a 1:2, the number of Ecm11 molecules with the Ecm11-Gmc2 complex is still uncertain.

MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ was analysed by SEC-SAXS to further characterise the structural properties in solution. Guinier analysis determines the Rg to be 80 Å (Figure 5.24 C), which is smaller than the Rg determined when only Gmc2 was fused to an MBP-tag suggesting that a more compact structure is formed. The distance distribution curve shows an extended shape with a maximum dimension of 348 Å (Figure 5.24 D). This is a shorter length than determined for His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$, again suggesting that a more compact structure is formed. A large peak is identified at 80 Å, corresponding to the distances between the MBP tags. This demonstrates the previously determined parallel nature of the Gmc2 chains, but also suggests that the N-terminus of Ecm11$_{206-302}$ is in close proximity to the Gmc2 N-termini. This proposes a model whereby all Gmc2 and Ecm11 chains are in a parallel arrangement. However, the oligomer status of Ecm11 within the complex is still ambiguous.

A non-cleavable MBP tag was fused to Ecm11$_{206-302}$ for analysis with the cleaved Gmc2$_{37-188}$. The two proteins were co-purified by anion exchange chromatography, following Gmc2 tag-cleavage, for analysis (Figure 5.25 A). SEC-MALS calculated the molecular weight of the MBP-Ecm11$_{206-30}$-Gmc2$_{37-188}$ complex peak to be 172 kDa (Figure 5.25 B). The theoretical molecular mass for a 2:2 complex is estimated to be 142 kDa. This suggests that a 2:2 complex is being formed, although the experimental molecular weight is slightly higher than the estimated 144 kDa theoretical 2:2 complex, but this may be due to the presence of some uncleaved MBP-Gmc2$_{37-188}$ material that may be incorporated in some of the complexes. However, this does imply the presence of two Ecm11 molecules within the complex.

MBP-Ecm11$_{206-302}$-Gmc2$_{37-188}$ was analysed by SEC-SAXS and the Rg was determined to be 83 Å (Figure 5.25 C). The distribution plot of the interatomic distances within the sample show that MBP-Ecm11$_{206-302}$-Gmc2$_{37-188}$ forms an extended shape with a maximum dimension observed to be 387 Å (Figure 5.25 D). A large peak was observed at 73 Å, indicating the distance between the globular MBP affinity tags on the Ecm11 chains (Figure 5.25 D). This distance suggests that the Ecm11 molecules are parallel to one another.

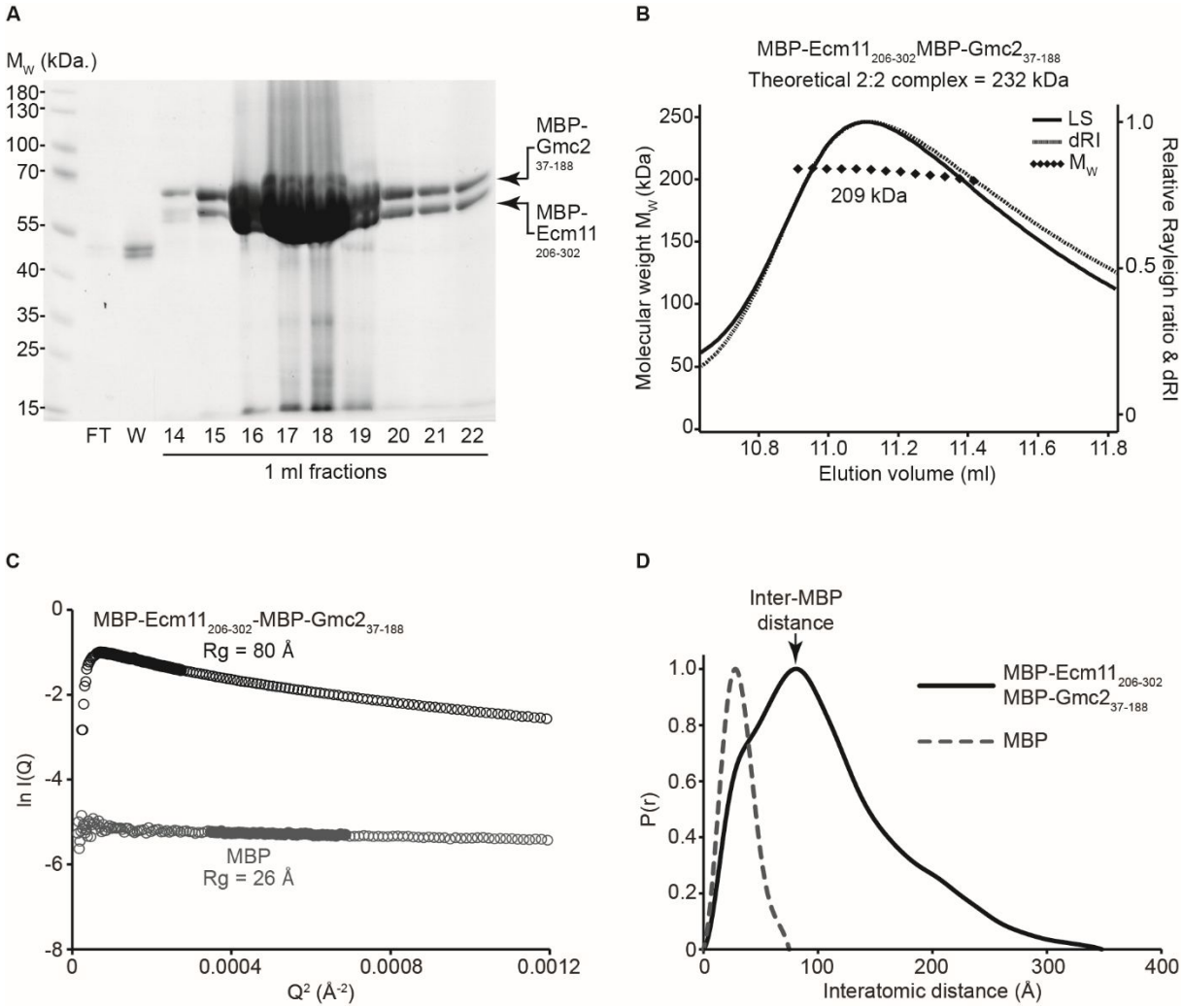


Figure 5.24. Purification and structural analysis of MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (B) SEC-MALS analysis showing a single 209 kDa peak. A theoretical 2:2 complex has an estimated 232 kDa molecular weight suggesting a 1:2 or 2:2 complex formation. (B) Guinier fit to determine the radius of gyration (Rg) as 80 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$ has a maximum dimension of 348 Å, with an inter-MBP peak at 80 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

Through combining the data from the cleaved and fusion Ecm11$_{206-302}$-Gmc2$_{37-188}$ complexes, an elongated 2:2 complex has been characterised with all four chains likely in a parallel arrangement. However, to fully study the structure of Ecm11-Gmc2, a high resolution structure is required. Ecm11$_{206-302}$-Gmc2$_{37-188}$ was prone to aggregate during chromatography steps for purification and therefore enough material was not obtained for crystallographic studies. In order to determine a stable structural Ecm11-Gmc2 unit for crystallography, further constructs were designed for biophysical characterisation.

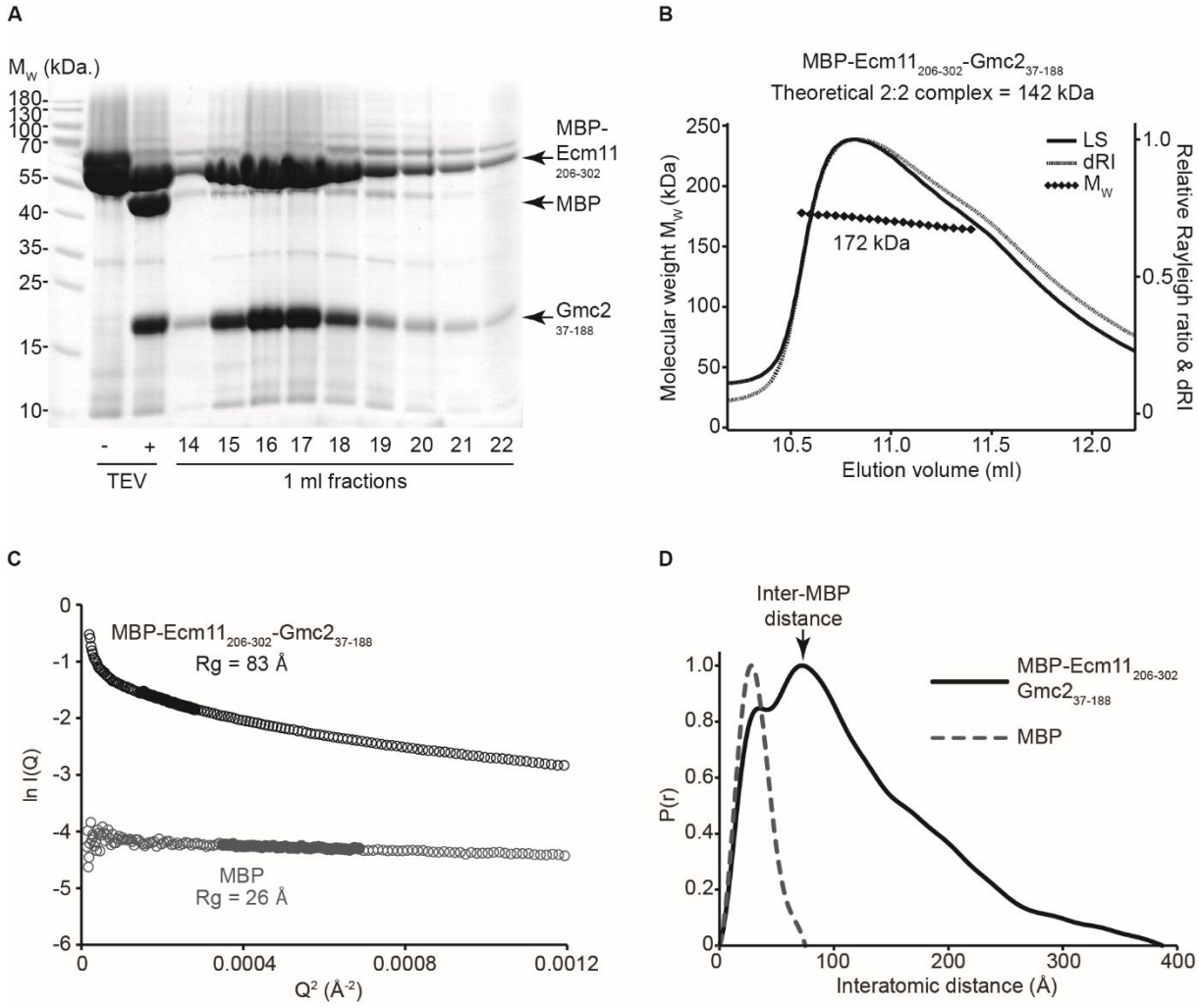


Figure 5.25. Purification and structural analysis of MBP-Ecm11$_{206-302}$-Gmc2$_{37-188}$. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. (B) SEC-MALS analysis showing a single 172 kDa peak. A theoretical 2:2 complex has an estimated 142 kDa molecular weight suggesting a 2:2 complex formation, with the molecular weight higher than expected due to the presence of uncleaved MBP-Gmc2$_{37-188}$ contamination. (C) Guinier fit to determine the radius of gyration (Rg) as 83 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. MBP-Ecm11$_{206-302}$-Gmc2$_{37-188}$ has a maximum dimension of 387 Å, with an inter-MBP peak at 73 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

5.2.8 Optimisation of the Ecm11 structured C-terminus in the Ecm11-Gmc2 complex

Although the C-terminal half of Gmc2 is more conserved than the N-terminal half, the secondary structure prediction stretches across the whole protein, suggesting that there may not be a specific structural core of Gmc2. This is supported by the stability of the full length Gmc2 in the absence of Ecm11. Therefore, the full length Gmc2 protein was co-expressed with $Ecm11_{206-302}$ for characterisation.

His-$Ecm11_{206-302}$ and MBP-Gmc2 were co-expressed and purified by sequential Ni-NTA and amylose affinity chromatography steps, followed by anion exchange chromatography (Figure 5.26 A). The fusion complex was analysed by SEC-MALS and the molecular mass of His-$Ecm11_{206-302}$-MBP-Gmc2 was determined to be 161 kDa (Figure 5.26 B). The theoretical mass of a 2:2 complex is 162 kDa, therefore a 2:2 complex is formed.

The His-$Ecm11_{206-302}$-MBP-Gmc2 complex was analysed by SEC-SAXS. The Guinier analysis determined the Rg to be 107 Å (Figure 5.26 C), suggesting an extended structure. The paired distance distribution plot estimated the maximum dimension to be 450 Å (Figure 5.26 D), further indicating an elongated structure. An inter-MBP peak was observed at 83 Å (Figure 5.26 D), demonstrating the close proximity of the Gmc2 N-termini and therefore indicating the parallel arrangement of the Gmc2 chains.

To study the cleaved complex, the affinity tags were removed by TEV protease incubation and $Ecm11_{206-302}$-Gmc2 was purified by anion exchange chromatography (Figure 5.27 A). The secondary structure of the $Ecm11_{206-302}$-Gmc2 complex was measured by CD wavelength scans. A strong helical signal was recorded and deconvolution estimated an alpha helical content of 80% (Figure 5.27 C), which correlates to the predicted secondary structure for this complex and is also comparable to the 83% helical content for $Ecm11_{206-302}$-$Gmc2_{37-188}$. The estimated number of helical residues is 232 and 210 respectively, highlighting that the N-terminal Gmc2 region (amino acids 1-36) does contain alpha helical elements. The thermal stability of $Ecm11_{206-302}$-Gmc2 was tested by CD thermal denaturation. The melting curve showed a two-stage cooperative unfolding with an estimated melting temperature of 36°C (Figure 5.27 D).

SEC-MALS calculated the molecular mass of the $Ecm11_{206-302}$-Gmc2 complex peak to be 64 kDa (Figure 5.27 B). The theoretical 2:2 complex mass is estimated to be 68 kDa, therefore the inferred structure is the formation of a 2:2 complex, consistent with the study of $Ecm11_{206-302}$-$Gmc2_{37-188}$.

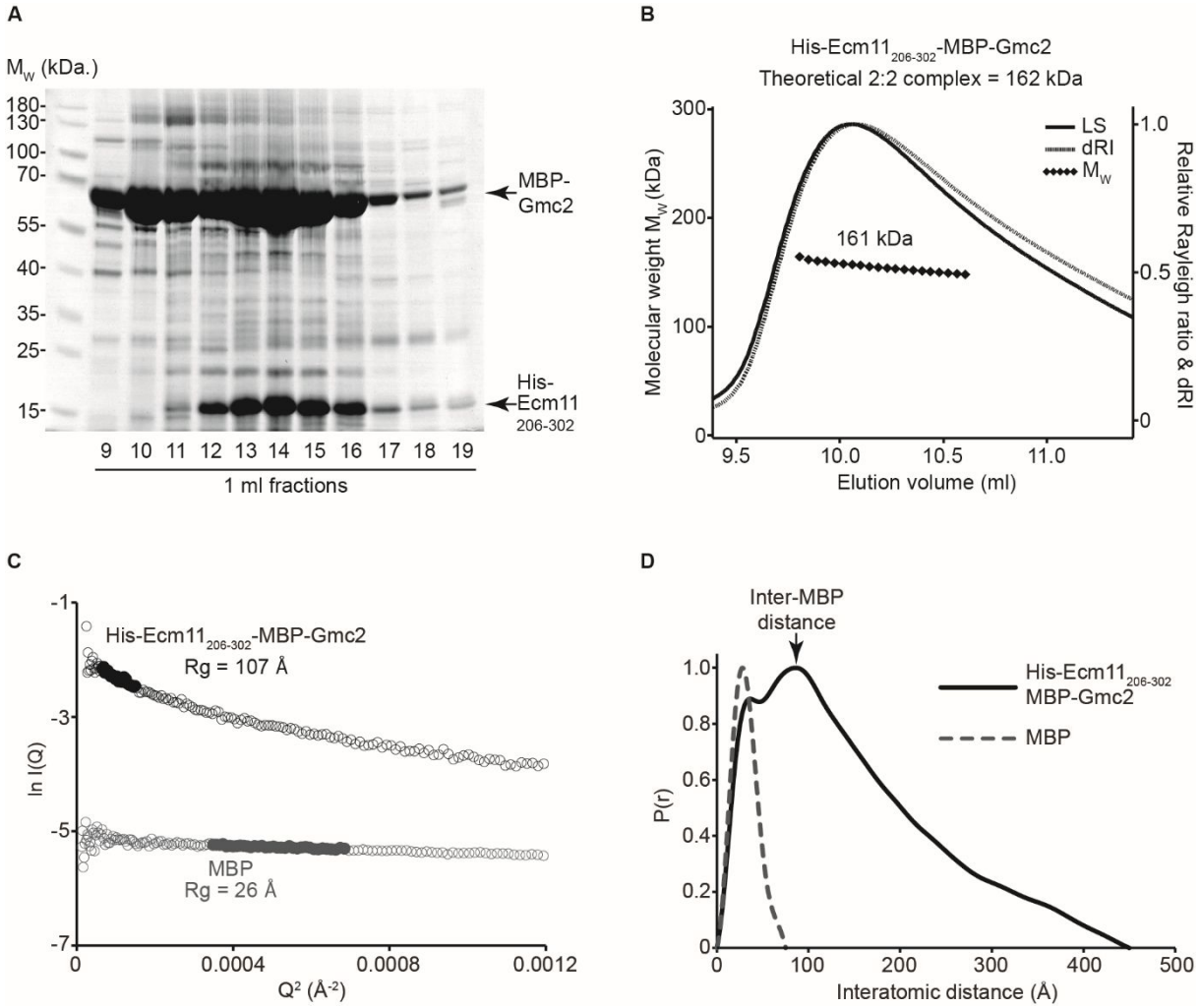


Figure 5.26. Purification and structural analysis of His-Ecm11$_{206-302}$-MBP-Gmc2. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. Material was prepared by Chandni Ravindan. (B) SEC-MALS analysis showing a single 161 kDa peak. A theoretical 2:2 complex has an estimated 162 kDa molecular weight indicating a 2:2 complex formation. (C) Guinier fit to determine the radius of gyration (Rg) as 107 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. His-Ecm11$_{206-302}$-MBP-Gmc2 has a maximum dimension of 450 Å, with an inter-MBP peak at 86 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

Although the Ecm11$_{206-302}$-Gmc2 complex was more stable than the Ecm11$_{206-302}$-Gmc2$_{37-188}$ complex, some aggregation was still observed during purification and a high enough yield for crystallisation was still not obtained. Therefore, it was speculated whether the Ecm11$_{206-302}$ fragment is contributing to the tendency for aggregation. The Ecm11 amino acids 176-302 are predicted to be highly alpha helical, with a stretch at the very C-terminus predicted

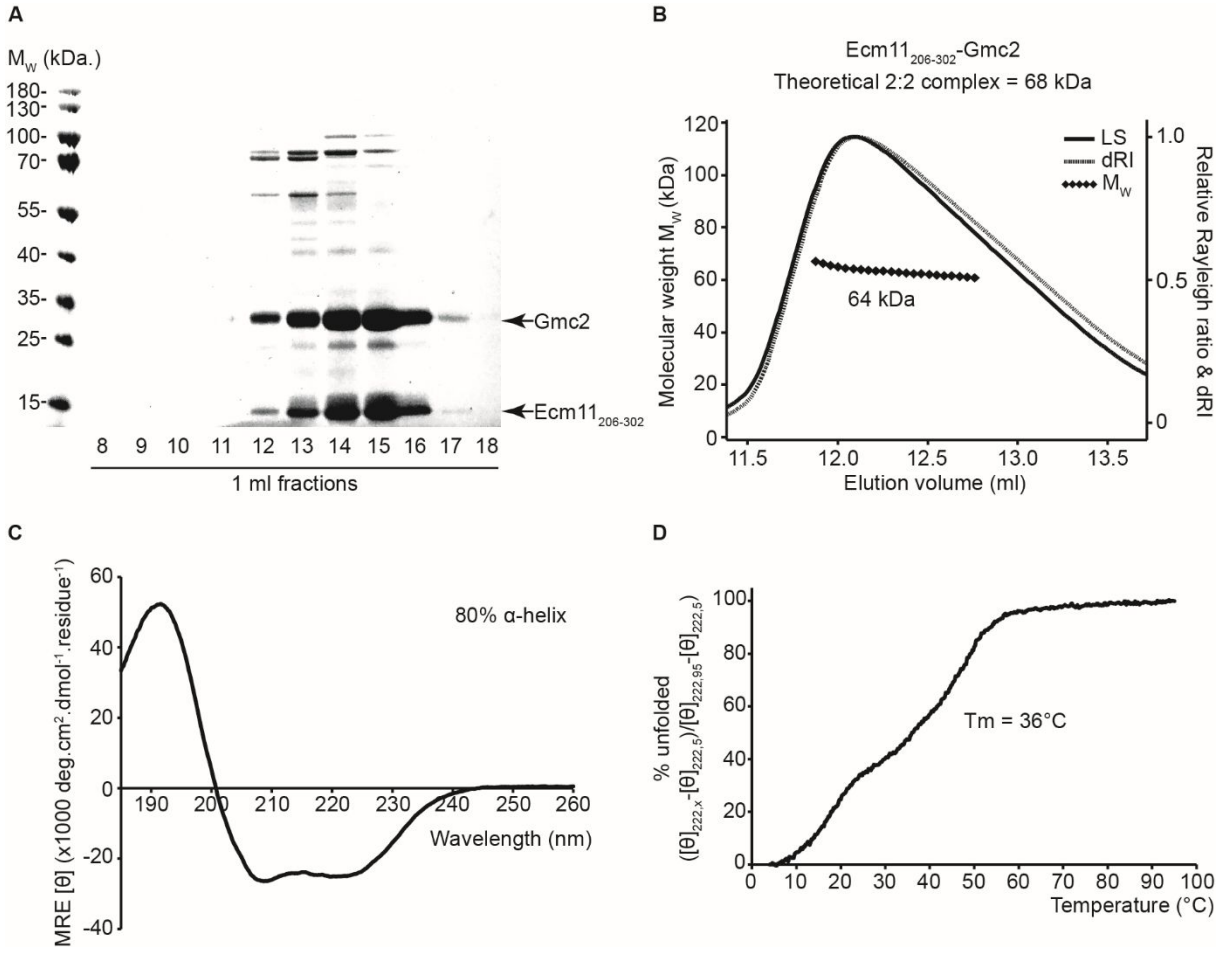


Figure 5.27. Purification and structural analysis of Ecm11$_{206-302}$-Gmc2. Data collected by Chandni Ravindan under my supervision. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. (B) SEC-MALS analysis showing a single 64 kDa peak. A theoretical 2:2 complex has an estimated 68 kDa molecular weight indicating a 2:2 complex formation. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 80% alpha-helix, 6% β-sheet, 4% turns and 10% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 36°C, taken as the point at which the sample is 50% unfolded.

to form a coiled-coil. However, the alpha helical prediction is fragmented across this region, which may suggest a more intricate structure than simple elongated chains. This also implies that achieving a stable N-terminal construct boundary may require identifying a residue that falls in a helical break position. Therefore, further Ecm11 constructs were designed by selecting N-terminal boundaries within the predicted helical breaks for co-expression with full length Gmc2.

MBP-Gmc2 was co-expressed with His-Ecm11$_{220-302}$ for purification. However, upon sonication and lysate clarification by centrifugation, both the soluble and insoluble material had a strikingly viscous consistency, and this was observed with several independent cultures for these particular sequences when co-expressed. To obtain suitable lysate for purification, several centrifugation steps were required, followed by sequential Ni-NTA and amylose affinity chromatography and then His-Ecm11$_{220-302}$-MBP-Gmc2 was further purified by anion exchange chromatography (Figure 5.28 A). The fusion protein was analysed by SEC-MALS, where the majority of the material exists as a 1:1 complex, with some of the sample forming large molecular weight species (Figure 5.28 B). Combining the SEC-MALS data with the observed agglutination of the bacterial lysate, suggests that the His-Ecm11$_{220-302}$-MBP-Gmc2 complex is prone to aggregation.

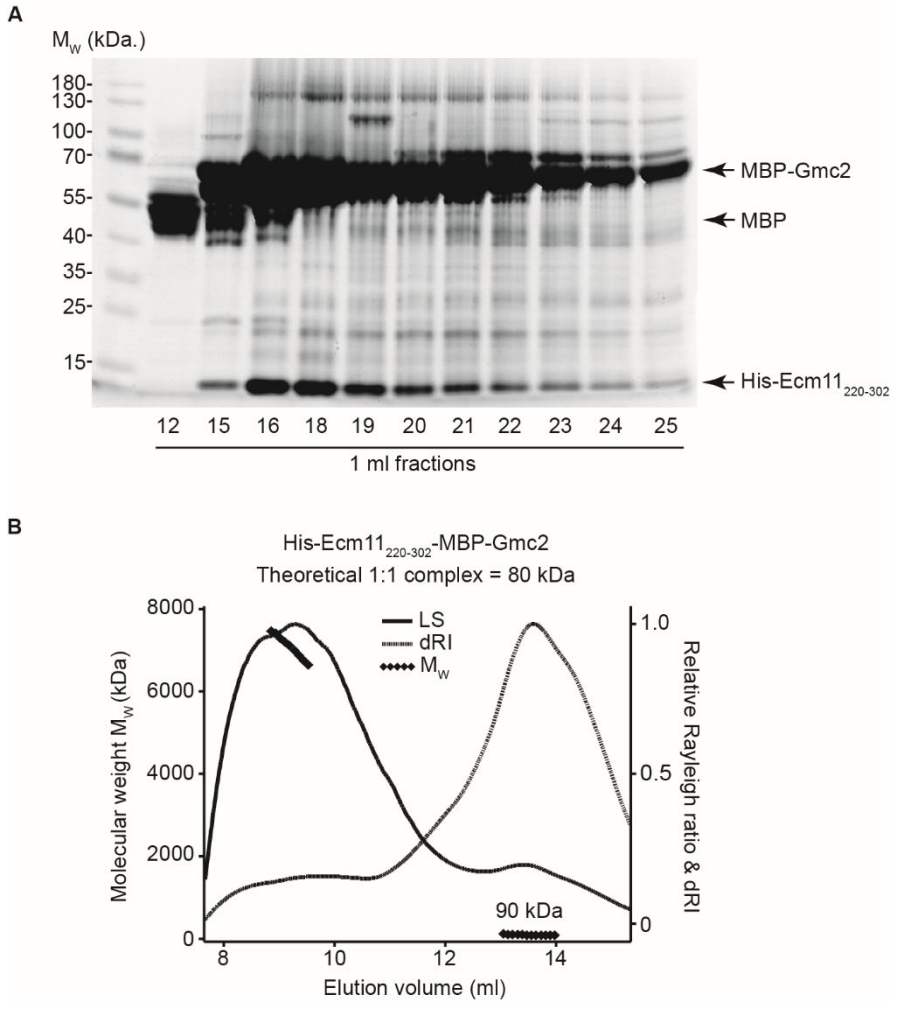


Figure 5.28. Purification and structural analysis of His-Ecm11$_{220-302}$-MBP-Gmc2. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. Material was prepared by Mumen Chowhury under my supervision. (B) SEC-MALS analysis showing the majority of the material is monomeric (according to the dRI trace) with the rest of the sample forming higher-order assemblies.

Although the fusion protein complex undergoes aggregation, the affinity tags could be removed by TEV protease incubation and the cleaved complex was purified by anion exchange chromatography (Figure 5.29 A). Following the purification of the cleaved Ecm11$_{220-302}$-Gmc2 complex, no further aggregation was observed. Therefore, the secondary structure could be measured by CD wavelength scans. The CD scan reveals a strong alpha helical signal for Ecm11$_{220-302}$-Gmc2, with the helical content estimated to be 77% by deconvolution (Figure 5.29 C). This correlates to an estimation of 213 helical residues, which is less than the 232 helical residues as determined for the Ecm11$_{206-302}$-Gmc2 complex. This shows that although only 14 amino acids have been removed, there are 19 fewer helical residues, suggesting that the N-terminal region may not have fallen in a helical break position, which could explain the propensity for aggregation. CD thermal denaturation was performed with Ecm11$_{220-302}$-Gmc2, and the melting curve shows a two-stage unfolding curve with an estimated melting temperature of 40°C (Figure 5.29 D). To investigate the oligomeric state of Ecm11$_{220-302}$-Gmc2, SEC-MALS was performed. The molecular mass for the Ecm11$_{220-302}$-Gmc2 complex peak was calculated to be 62 kDa (Figure 5.29 B). The estimated theoretical weight of the 2:2 complex is 64 kDa, therefore Ecm11$_{220-302}$-Gmc2 was found to form a 2:2 complex, as determined by previous constructs. As the Ecm11 N-terminal boundary was found to cause excessive aggregation, the Ecm11$_{220-302}$-Gmc2 was not suitable for further structural analysis.

A further Ecm11 truncation was designed to express residues 230-302. His-Ecm11$_{230-302}$ was co-expressed with MBP-Gmc2 and the fusion protein was purified for analysis (Figure 5.30 A). Molecular mass determination of the His-Ecm11$_{230-302}$-MBP-Gmc2 complex was carried out by SEC-MALS. The molecular weight of the fusion complex was calculated to be 141 kDa (Figure 5.30 B). The theoretical 2:2 complex molecular weight is estimated as 158 kDa, suggesting a 2:2 complex is being formed.

SEC-SAXS Guinier analysis measures the His-Ecm11$_{230-302}$-MBP-Gmc2 Rg to be 61 Å (Figure 5.30 C). The paired distribution plot estimates a 296 Å maximum dimension, correlating to an extended structure (Figure 5.30 D). A peak at 72 Å indicates the distance between the globular MBP tags and, consistent with other SAXS experiments, implies a parallel orientation of Gmc2 chains.

To study the cleaved Ecm11$_{230-302}$-Gmc2 complex, the fusion proteins were incubated with TEV protease. However, although the MBP was successfully removed from Gmc2, the His-tag cleavage was partial, resulting in a mixture of cleaved and tagged Ecm11$_{230-302}$. This suggests that the TEV recognition site between the His-tag and the Ecm11 sequence has been occluded, further implying that the amino acid 230 may in fact fall at the start of a structured helical region and is an appropriate construct to characterise for structural studies.

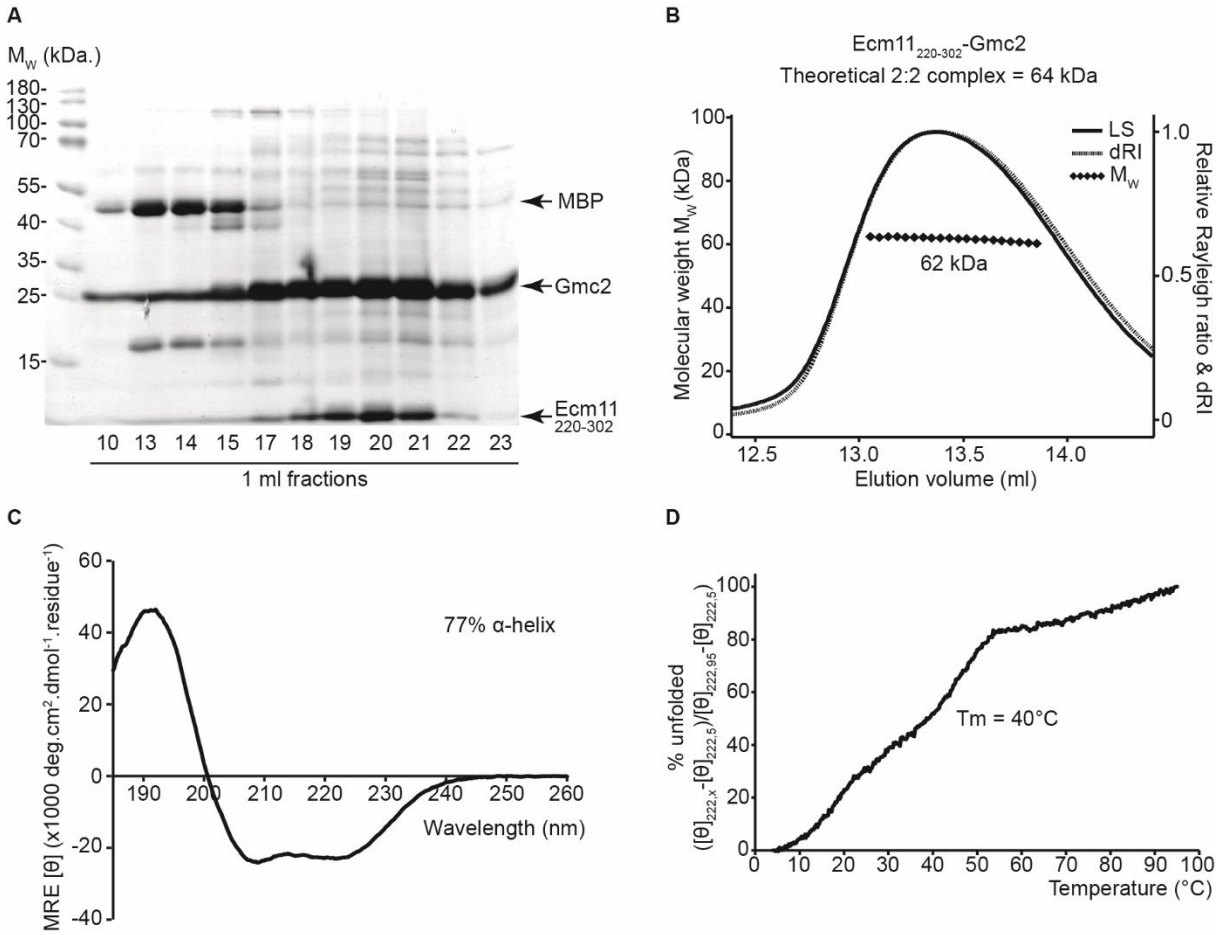


Figure 5.29. Purification and structural analysis of Ecm11$_{220-302}$-Gmc2. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. Material was prepared by Mumen Chowhury under my supervision. (B) SEC-MALS analysis showing a single 62 kDa peak. A theoretical 2:2 complex has an estimated molecular weight of 64 kDa indicating the formation of a 2:2 complex. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 77% alpha-helix, 8% β-sheet, 5% turns and 10% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 40°C, taken as the point at which the sample is 50% unfolded.

To address the issue of TEV site occlusion, two further constructs were designed. One such construct was designed with a stretch of three threonine-glycine-serine repeats (3TGS) between the TEV protease cleavage site and the Ecm11 sequence in order to protrude the site away from the Ecm11-Gmc2 structure, hereafter referred to as Ecm11$_{linker}$. The second approach was to design a non-cleavable His-tag so that upon TEV incubation, the MBP-tag could still be removed from Gmc2, but would prevent the mixture of species that are formed upon partial His-tag cleavage.

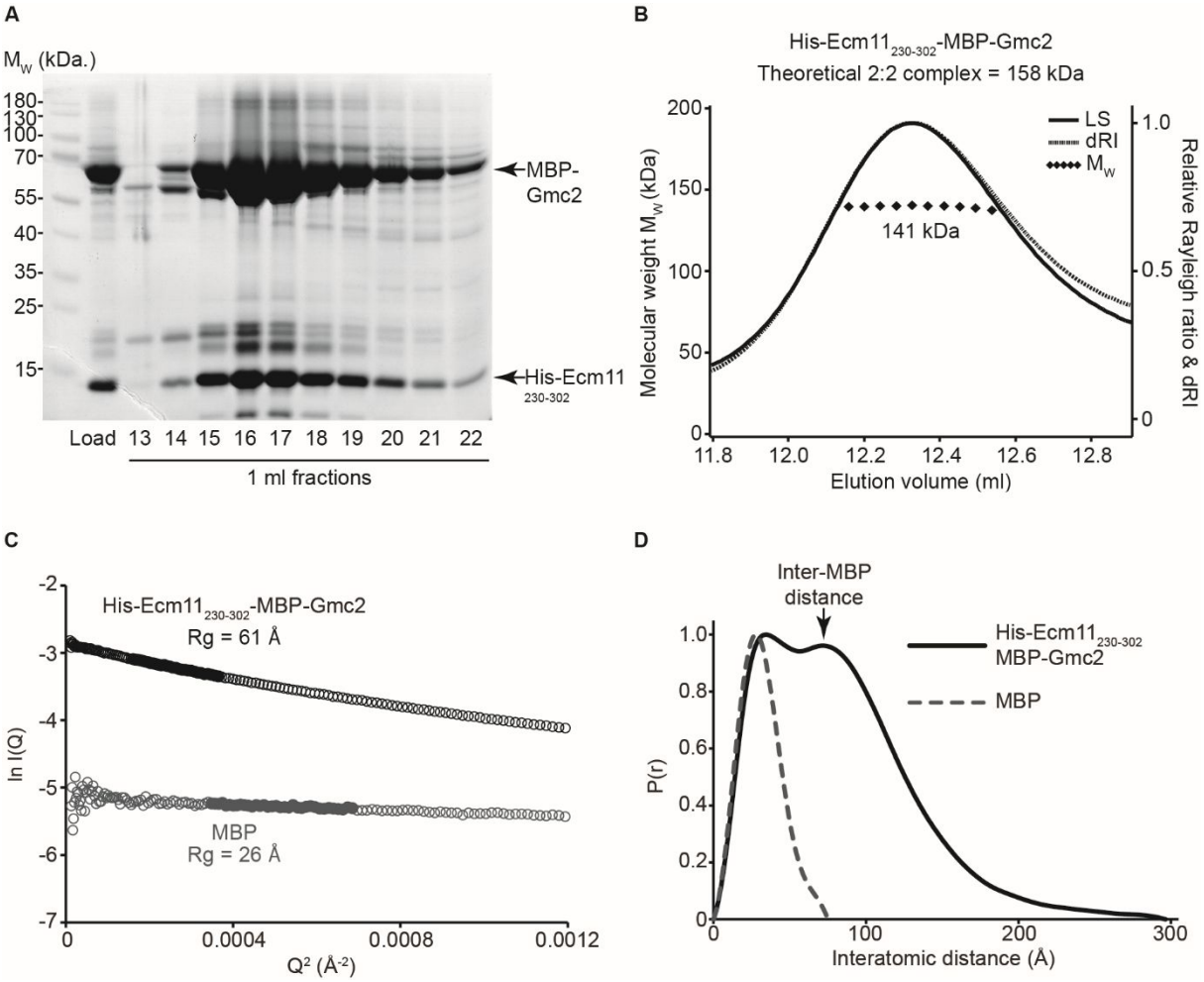


Figure 5.30. Purification and structural analysis of His-Ecm11$_{230-302}$-MBP-Gmc2. (A) SDS-PAGE showing the purification of the fusion complex by anion exchange chromatography, following Ni-NTA and amylose affinity chromatography. Material was prepared by Mumen Chowhury. (B) SEC-MALS analysis showing a single 141 kDa peak. A theoretical 2:2 complex has an estimated 158 kDa molecular weight suggesting a 1:2 or 2:2 complex formation. (C) Guinier fit to determine the radius of gyration (Rg) as 61 Å (black), with the MBP control as 26 Å (grey). Clear circles are the data, solid circles represent the regions used for the Guinier fit. (D) Paired real-space distribution plot. His-Ecm11$_{220-302}$-MBP-Gmc2 has a maximum dimension of 296 Å, with an inter-MBP peak at 71 Å as indicated (black line). The MBP control has a maximum dimension of 75 Å (grey dashed line).

Following the purification of the His- Ecm11$_{linker}$-MBP-Gmc2 complex, the fusion proteins were incubated with TEV protease and the cleaved Ecm11$_{linker}$-Gmc2 complex was purified by anion exchange chromatography (Figure 5.31 A). The His-tag was successfully cleaved from the Ecm11$_{linker}$ fragment, indicating that the insertion of the "3TGS" linker has successfully allowed the accessibility of the TEV recognition site to the TEV protease in order for tag removal.

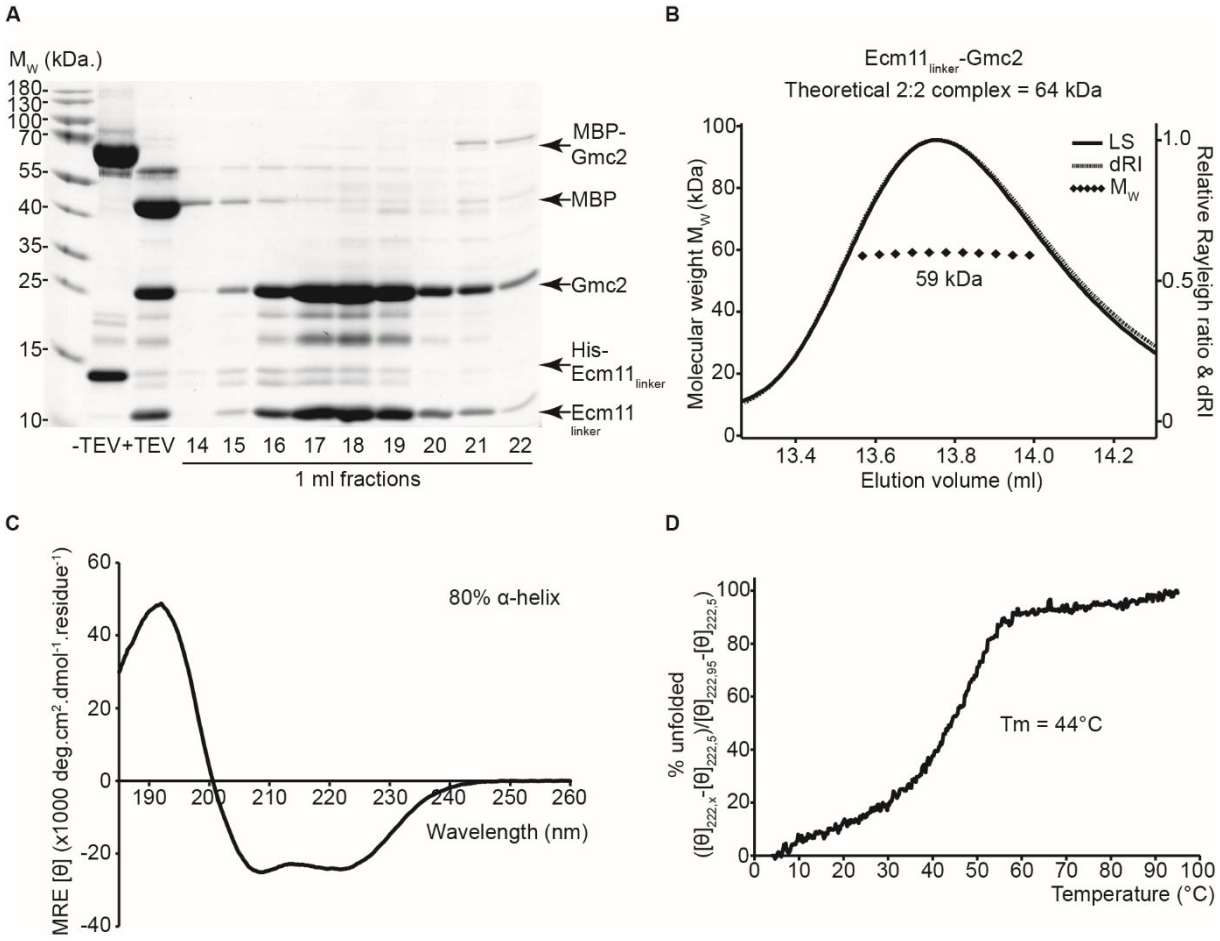


Figure 5.31. Purification and structural analysis of Ecm11$_{linker}$-Gmc2. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. Material was prepared by Mumen Chowhury under my supervision. (B) SEC-MALS analysis showing a single 59 kDa peak. A theoretical 2:2 complex has an estimated 64 kDa molecular weight indicating the formation of a 2:2 complex. (C) CD scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 80% alpha-helix, 6% β-sheet, 6% turns and 8% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 44°C, taken as the point at which the sample is 50% unfolded.

The insertion of the "3TGS" linker enabled the purification of the cleaved Ecm11$_{230-302}$-Gmc2 complex for analysis. A CD wavelength scan was recorded to determine the secondary structure of Ecm11$_{linker}$-Gmc2, where a strong helical signal was detected, with an 80% helical content estimated by deconvolution (Figure 5.31 C). CD thermal denaturation shows a melting curve that exhibits cooperative unfolding. Interestingly, when compared to previous Ecm11-Gmc2 constructs, a single unfolding event is observed (Figure 5.31 D), differing from the two-

stage unfolding previously identified. The melting temperature was estimated to be 44°C, which is a higher temperature than the 36°C melting temperature determined for the Ecm11$_{206-302}$-Gmc2 complex. The increased thermal stability of Ecm11$_{linker}$-Gmc2, and the single cooperative unfolding event, suggests that the structural core of the Ecm11-Gmc2 complex has been isolated, and that the Ecm11 amino acids 206-229 (and the rest of the Ecm11 N-terminus) may partake in additional helical interactions that may account for the early unfolding step.

The molecular mass of the Ecm11$_{linker}$-Gmc2 species was calculated to be 59 kDa by SEC-MALS analysis (Figure 5.31 B). A theoretical 2:2 complex is estimated to be 64 kDa, suggesting that Ecm11$_{linker}$-Gmc2 still exists as a 2:2 complex and further implies that the N-terminal Ecm11 region is not required for oligomerisation.

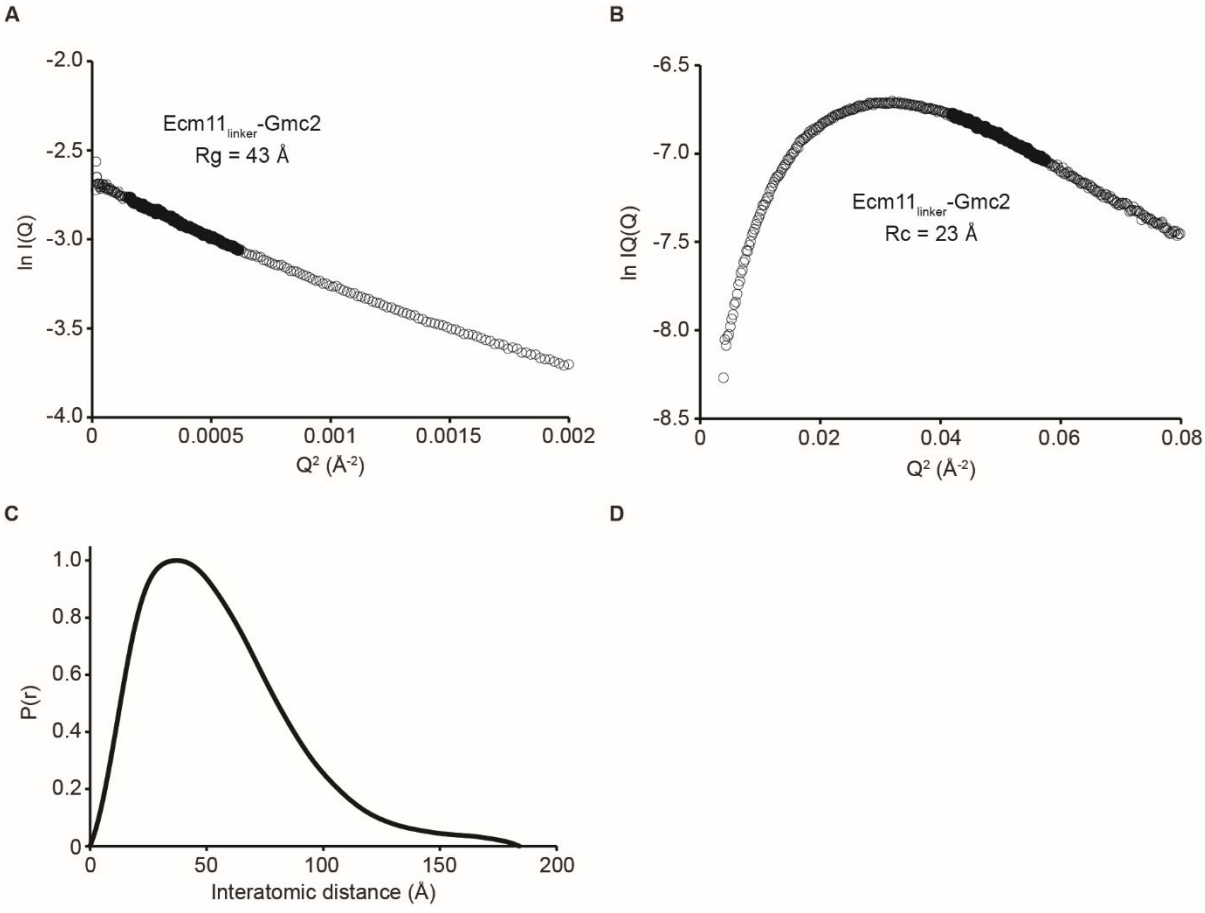


Figure 5.32. SEC-SAXS analysis of Ecm11$_{linker}$-Gmc2. (A) Guinier fit to determine the radius of gyration (Rg) as 43 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 23 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. Ecm11$_{linker}$-Gmc2 has a maximum dimension of 184 Å.

SEC-SAXS was used to study the low resolution solution structure of Ecm11$_{linker}$-Gmc2. Guinier analysis determines the Rg as 43 Å (Figure 5.32 A), indicating that a relatively compact structure is formed. The Rc was determined as 23 Å, suggesting a bulkier structure rather than a simple four helical bundle (Figure 5.32 B). The distance distribution curve signifies a predominantly globular structure, with the tailing of the peak representing an element of elongation for the Ecm11$_{linker}$-Gmc2 complex (Figure 5.32 C). The maximum dimension was estimated to be 184 Å by real space analysis (Figure 5.32 C). The maximum dimension for Ecm11$_{206-302}$-Gmc2$_{37-188}$ was determined as 275 Å, meaning that the alterations in construct boundaries has caused almost a 100 Å difference in complex length, with Ecm11$_{linker}$-Gmc2 forming the more compact and globular structure.

The His-Ecm11$_{230-302}$-Gmc2 complex was also purified for analysis and structural studies. Following co-purification of the fusion complex, TEV protease incubation was performed to remove the MBP tag from Gmc2, but maintaining the His-tag on Ecm11$_{230-302}$ (by TEV recognition site removal) to prevent a mixture of species being formed after the TEV reaction. Following TEV protease incubation, anion exchange chromatography was performed to purify the His-Ecm11$_{230-302}$-Gmc2 complex (Figure 5.33 A).

CD wavelength scans, and subsequent deconvolution, estimates the helical content of His-Ecm11$_{230-302}$-Gmc2 to be 78%, corresponding to 228 helical residues (Figure 5.33 C). The His-Ecm11$_{230-302}$-Gmc2 melting curve, produced by CD thermal denaturation, displays cooperative unfolding with an estimated melting temperature of 43°C (Figure 5.33 D). As observed with Ecm11$_{linker}$-Gmc2, a single unfolding event is recorded.

SEC-MALS analysis of His-Ecm11$_{230-302}$-Gmc2 calculates a 67 kDa species, which corresponds to a 2:2 complex (Figure 5.33 B). SEC-SAXS experiments determine the Rg to be 46 Å (Figure 5.34 A) and the Rc to be 24 Å (Figure 5.34 B), suggesting a relatively compact structure. The real space analysis indicates a predominantly globular shape, with the tailing of the peak signifying some elongation (Figure 5.34 C). The maximum dimension is estimated as 198 Å, forming a more compact structure than Ecm11$_{206-302}$-Gmc2$_{37-188}$. *Ab initio* modelling was performed to produce the low resolution molecular envelope for His-Ecm11$_{230-302}$-Gmc2 (Figure 5.34 D). This demonstrates the globular core of the complex with some elongation likely provided by the N-termini of Ecm11-Gmc2.

To uncover the molecular details that govern the Ecm11-Gmc2 interaction, crystallographic studies were initiated to obtain a high resolution model. One condition from commercial screening yielded non-birefringent micro-crystals that were too small for harvesting (Figure 5.35). Crystal optimisation was carried out with varying concentrations of protein and screening of the well solution components. Unfortunately, the crystals were not

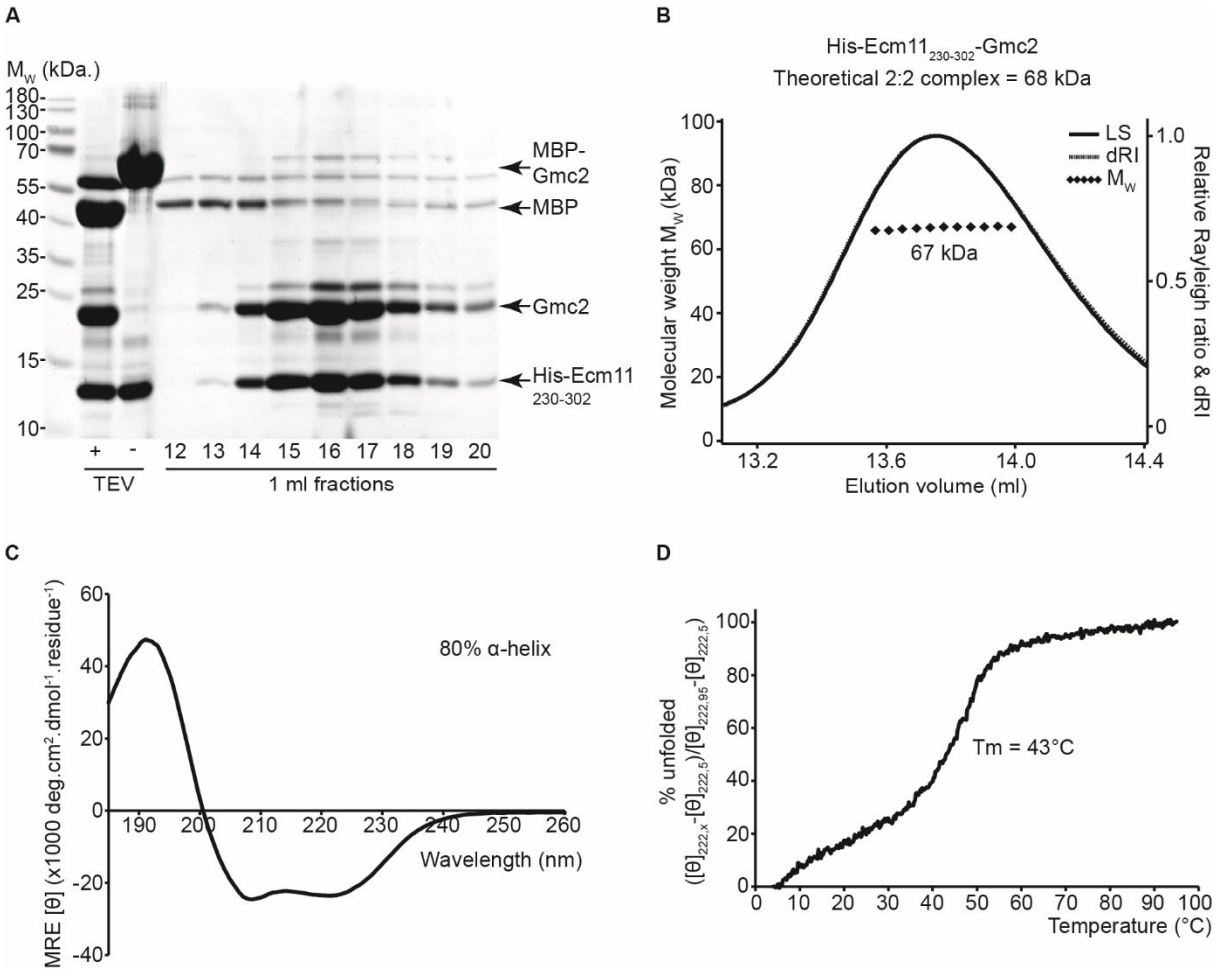


Figure 5.33. Purification and structural analysis of His-Ecm11$_{230-302}$-Gmc2. (A) SDS-PAGE showing the purification of the cleaved complex by anion exchange chromatography, following tag removal by TEV protease incubation. Material was prepared by Mumen Chowhury under my supervision. (B) SEC-MALS analysis showing a single 67 kDa peak. A theoretical 2:2 complex has an estimated 68 kDa molecular weight indicating a 2:2 complex formation. (C) CD wavelength scan between 260-185 nm, plotted as MRE ([θ]) (x1000 deg.cm^2.dmol^{-1}.residue^{-1}). Deconvolution estimates 80% alpha-helix, 7% β-sheet, 4% turns and 9% unordered. (D) CD thermal denaturation between 5°C and 95°C, plotted as % unfolded. Tm estimated as 43°C, taken as the point at which the sample is 50% unfolded.

reproduced upon optimisation and no other conditions yielded crystal growth from further commercial screening. The lack of robust crystal formation may imply that the complex is unsuitable for crystallisation. Although shorter Gmc2 fragments may produce a more suitable Ecm11-Gmc2 structural unit, the full Gmc2 sequence has stabilised the 2:2 complex formation. Further optimisation may be necessary for construct boundaries to allow for successful crystallisation to determine the high resolution structure.

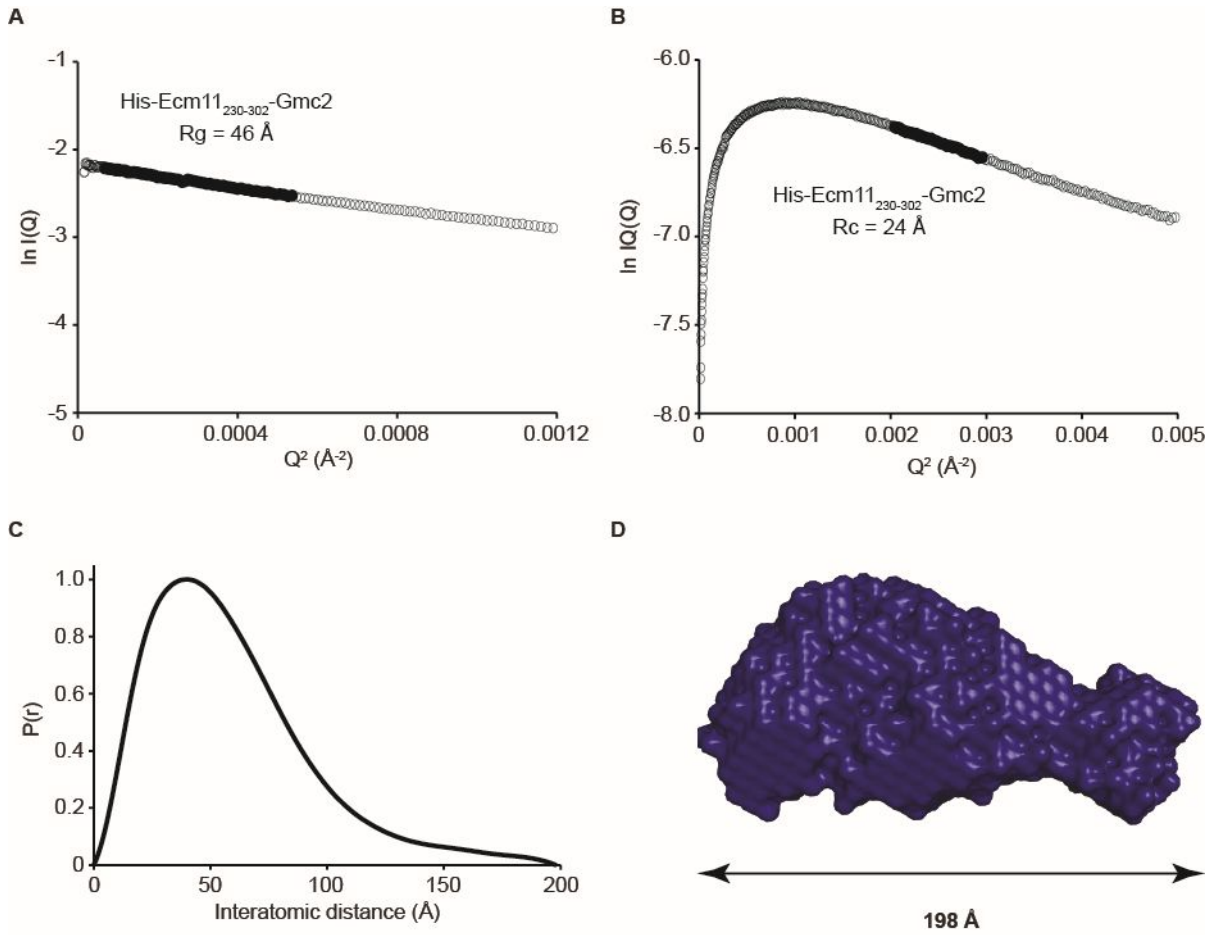


Figure 5.34. SEC-SAXS analysis of His-Ecm11$_{230-302}$-Gmc2. (A) Guinier fit to determine the radius of gyration (Rg) as 46 Å. Clear circles are the data, solid circles represent the region used for the Guinier fit. (B) Guinier analysis to determine the radius of gyration of the cross-section (Rc) as 24 Å. Clear circles are the data, solid circles represent the region used for the fit. (C) Paired real-space distribution plot. His-Ecm11$_{230-302}$-Gmc2 has a maximum dimension of 198 Å. (D) *Ab initio* DAMMIF model presented as the molecular envelope.

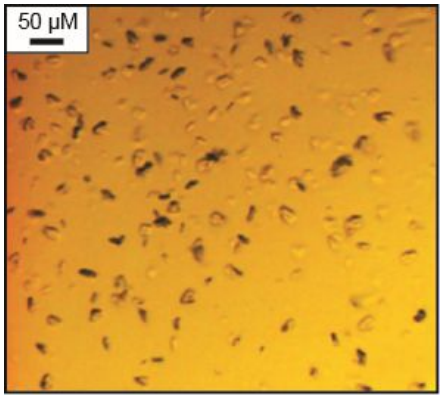


Figure 5.35. Crystallisation of His-Ecm11$_{230-302}$-Gmc2. Micro-crystals formed in 0.2 M sodium citrate, 0.1 M sodium cacodylate pH 6.5, 30% 2-propanol (v/v).

5.2.9 The role of Smt3 SUMOylation in central region assembly

No direct interaction has been detected between the central region components Ecm11-Gmc2 and the Zip1 N-terminus and therefore the molecular mechanism of the yeast SC remains unclear. However, a major role for SUMO, Smt3 in yeast, has been implicated in the assembly of the central region for full SC elongation. It has been shown that Ecm11 is SUMOylated *in vivo* in a Gmc2-dependant manner and that these events are required for Zip1 assembly in the SC (Leung *et al*., 2015). Two SUMOylation sites have been identified in Ecm11, K5 and K101 and therefore it is hypothesised that Smt3 may interact with Ecm11 through the N-terminus and that Zip1 may also interact with Smt3 due to its switch from poly-complex formation to SC incorporation upon Ecm11 SUMOylation (Leung *et al*., 2015).

To investigate whether Smt3 directly interacts with the central region proteins, a yeast 2-hybrid screen was carried out with Ecm11, Gmc2 and the Zip1 N-terminal 1-348 fragment. Consistent with the *in vivo* data, Smt3 was found to interact with Ecm11, with no interaction observed between Smt3 and Gmc2 (Figure 5.36). Further to this, the interaction between Ecm11 and Smt3 is abolished when the Ecm11 is truncated to amino acids 206-302, likely due to the removal of the K5 and K101 candidate SUMO sites.

As hypothesised by the recruitment of Zip1 into the SC upon Ecm11 SUMOylation (Leung *et al*., 2015), an interaction was identified between Smt3 and the Zip1 N-terminus. It has also been shown that the Zip1 N-terminus is required for the Ecm11 SUMOylation, so the identified interaction between Zip1 and Smt3 could suggest that Zip1 recruits Smt3 to Ecm11.

5.3 Discussion

5.3.1 The Zip1 N-terminus forms alpha-helical structures that bind DNA

Zip1 is a long helical molecule that assembles at chromosomal axes through its C-terminus and spans the distance between the lateral and central element to form head-to-head associations in the central element. The N-terminus of Zip1 is therefore thought to interact with central element components, and may also form self-associations (Liu *et al*., 1996). Zip1 also undergoes poly-complex formation, whereby Zip1 forms head-to-head and back-to-back interactions. Zip1 is encouraged to incorporate into the SC, in favour of the poly-complexes, upon Ecm11-Gmc2 SIC recruitment and subsequent Ecm11 SUMOylation (Leung *et al*., 2015).

Purification of Zip1 N-terminal fragments, and subsequent biophysical analysis, demonstrates the alpha helical nature of Zip1, with the very N-terminus being largely unstructured. Through molecular weight analysis, these Zip1 fragments were predominantly monomeric, with some propensity for dimerisation, driven by the 252-336 residues. Due to the experimental set-up, SEC-MALS analysis, and therefore oligomer determination, was

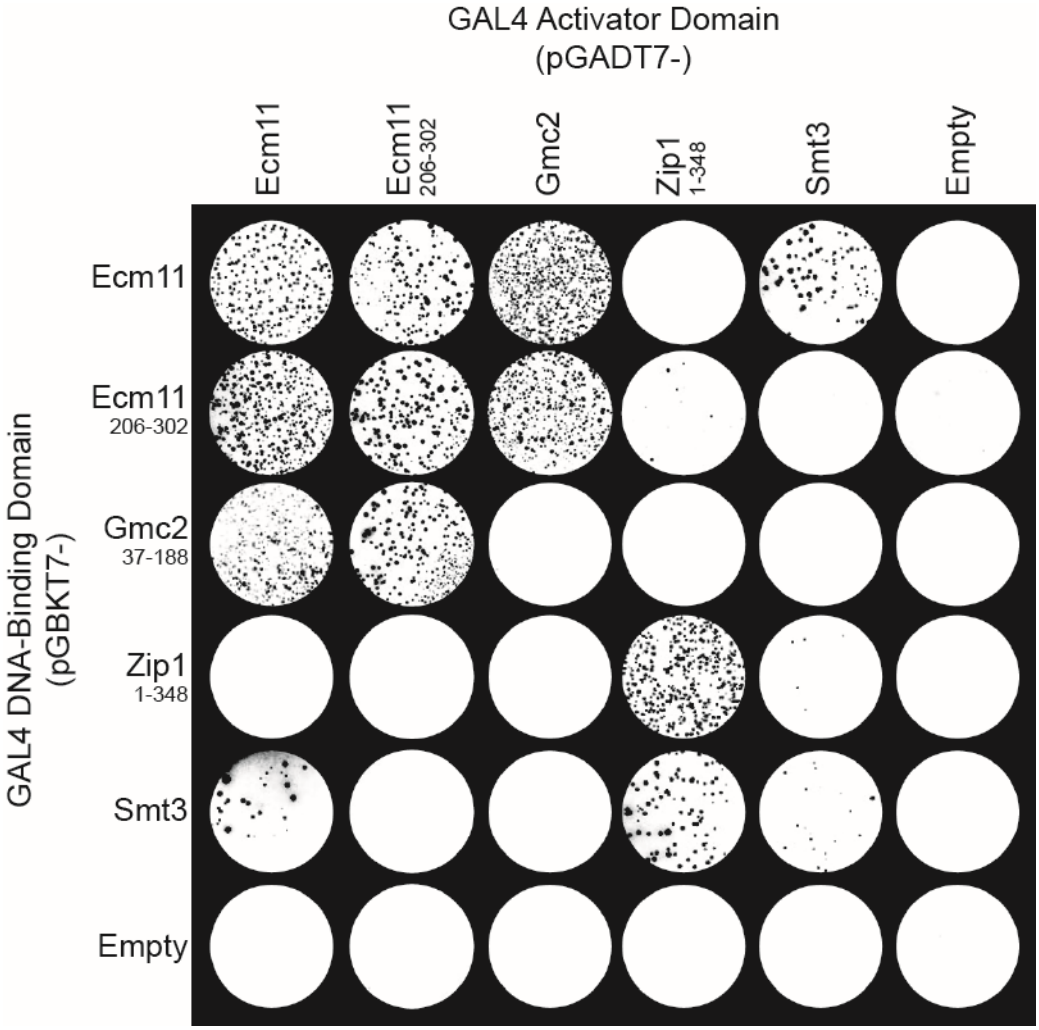


Figure 5.36. Yeast two-hybrid analysis of Smt3 and yeast SC proteins. Y187 (pGBKT7-bait) strains were mated with Y2H Gold (pGADT7-target) strains and plated on SD/-Leu/-Trp/-Ade/-His/x-α-gal plates for colony selection, subject to reporter gene activation (ADE1, HIS3 and MEL1) upon positive interactions. Experiment was performed by Chandni Ravindan.

performed at room temperature. From the CD thermal denaturation studies, it was observed that the Zip1 N-terminal fragments had estimated melting temperatures of 16°C to 33°C, with many of the constructs exhibiting a two-stage unfolding. This suggests that these fragments may be monomeric at room temperature, with the more stable constructs demonstrating a greater propensity for dimer formation (Figure 5.37 A). Therefore, to accurately determine the oligomeric status of the Zip1 N-terminus, the SEC-MALS analysis would preferably be repeated at 4°C. The $Zip1_{175-251}$ fragment is monomeric with no dimer formation and $Zip1_{1-251}$ exhibits a single unfolding event with a low helical content when compared to $Zip1_{1-336}$, 21% and 53% (Table 5.2). This suggests that the 175-251 region is a partially unstructured monomer and that the 252-336 region is highly helical and responsible for oligomerisation.

Analysis of the full length protein would be required to understand the complete structural architecture of Zip1. However, although the full length sequence (and other longer constructs) were successfully expressed in bacteria, they were heavily degraded and in some cases substantially insoluble. To improve this, the construct boundaries and affinity tag placement could be optimised. The expression temperature could also be varied, along with the lysis buffer composition. Another solution could be to use an eukaryotic expression system, such as yeast or insect cells, which would provide post-translational modifications that may be required for protein folding and stability. In the absence of the full length Zip1 analysis, the overall structure has not been determined. However, a model can be proposed through combining the N-terminal characterisation and sequence analysis, where the helical dimer continues and potentially forms coiled-coil interactions (Figure 5.37 B). The C-terminus, approximately amino acids 749-875, is predicted to be relatively unstructured and may be responsible for recruitment to the chromosome axis. Therefore, the N- and C-termini may function to interact with the central element and chromosome axis respectively, with the long central region forming dimeric coiled-coils to form the elongated structural bridge.

The Zip1 N-terminus was found to bind DNA. As this region is located in the central region of the SC, the function of this Zip1-DNA interaction is unclear. It could be postulated that the DNA binding may occur at DSB sites, where DNA would pass over the central region. The Zip1 DNA binding may function to stabilise Holliday junction structures at the DSB sites to aid with break repair and crossover formation (Figure 5.37 C). To test this, the EMSA's could be performed with Holliday junction DNA as the substrate. These experiments could also be performed with corresponding regions of SYCP1 to investigate whether this property is conserved in mammals to provide further insight into SYCP1 function. From the studies of the human SYCP1 transverse filaments, it is proposed that the C-termini are recruited to the chromosomal axes through SYCP1-DNA interactions and that this allows for alignment and head-to-head interactions of the N-termini to provide the preliminary lattice (Dunce *et al.*, 2018). Therefore, Zip1 may undergo a similar mechanism of assembly through DNA interactions at the C-terminus.

5.3.2 Ecm11-Gmc2 form a stable 2:2 complex

Ecm11-Gmc2 have been characterised as the yeast SC central element proteins. Their co-localisation has been established within the SC central region and have been found to interact by Y2H and immunoprecipitation (Humphryes *et al.*, 2013). Further Y2H studies have identified that the Ecm11-Gmc2 interaction occurs via the C-terminal portions of both proteins, corresponding to amino acids 206-302 and 96-188, respectively. A self-interaction between

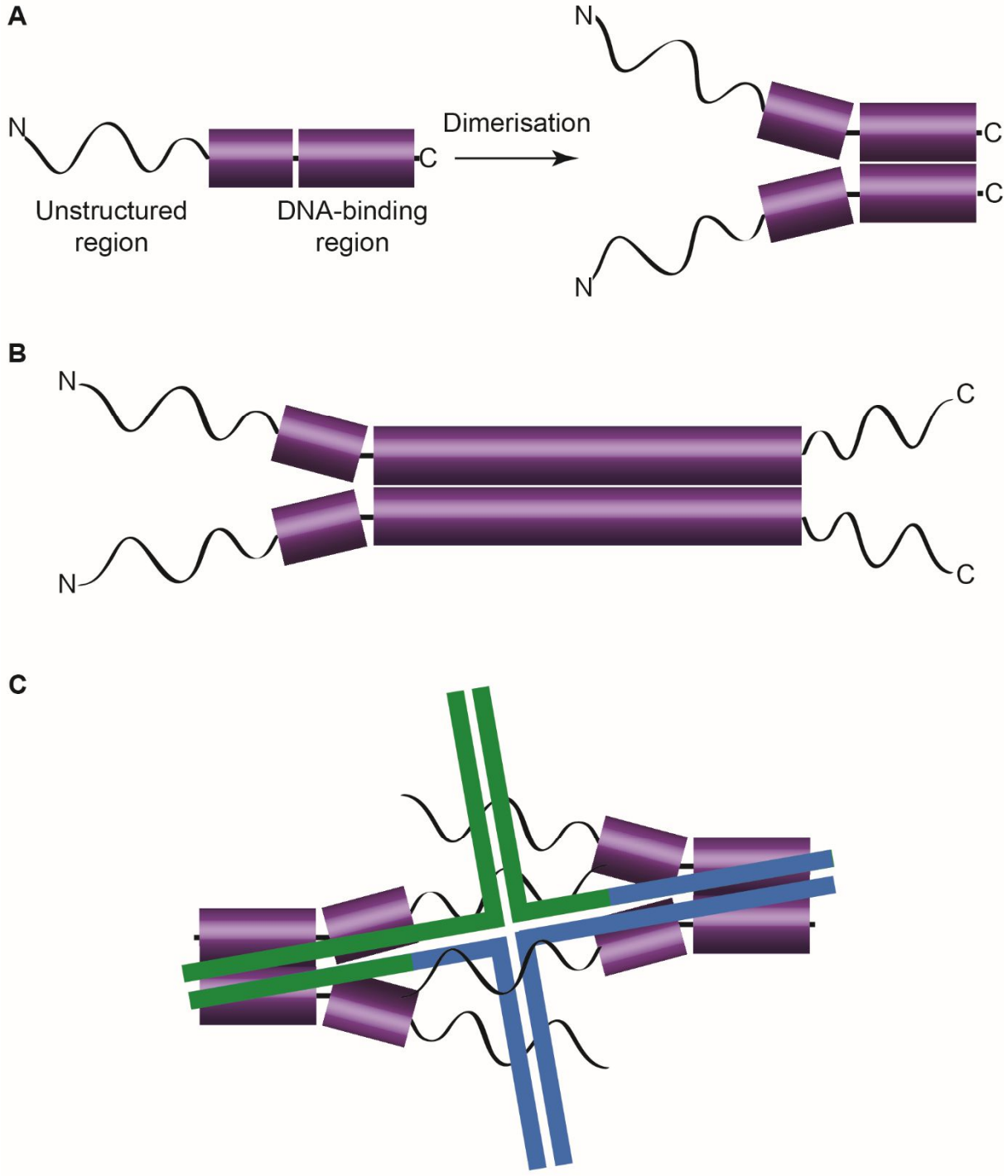


Figure 5.37. Schematic models for Zip1 structure in the SC. (A) The N-terminus of Zip1 (amino acids 1-174) is largely unstructured. The 175-251 region is partially helical and monomeric with the 252-336 region able to bind DNA and undergo dimerisation. (B) Schematic of the full length Zip1, proposing the unstructured N- and C-termini with the central helical region potentially forming dimeric coiled-coil interactions. (C) Proposed model for Zip1 binding to DNA Holliday junction structures. Dimeric N-terminal regions, stemming from opposing lateral elements, may align in the central region to coordinate Holliday junction binding.

Ecm11 was also identified through the 206-302 region, but no self-interaction was detected for Gmc2.

Through biophysical characterisation of a series of constructs, Ecm11 and Gmc2 are found to form a 2:2 complex that likely forms in a parallel arrangement. The removal of the N-terminal Gmc2 region (amino acids 1-36) encouraged aggregation of Ecm11-Gmc2 and also exhibited mild dissociation of Ecm11 from the complex. This implies that these residues are required for stabilisation of the complex. Upon SAXS analysis of the full length Gmc2 with the C-terminal Ecm11 region (amino acids 230-302), the complex was found to be a primarily globular structure. The maximum distance within the complex is 198 Å, suggesting that the Gmc2 may fold back on itself, as its estimated length as an elongated helix is 280 Å (Figure 5.38 A). Therefore, the Gmc2 1-36 residues may take part in the folding back of the Gmc2 chain and consequently its removal contributes to the destabilisation of the complex. The N-terminal region of Ecm11 is predicted to be largely unstructured and therefore may protrude from the complex to assist with assembly in the central element. This is supported by the location of the detected Ecm11 SUMO sites, residues K5 and K101 (Voelkel-Meiman *et al*., 2013), in the unstructured region, allowing these residues to be accessible for SUMOylation (Figure 5.38 B).

5.3.3 Smt3 SUMOylation drives yeast SC central region assembly

No direct interaction has been detected between Zip1 and Ecm11 or Gmc2, although an interaction may only be detected between Zip1 and the pre-formed Ecm11-Gmc2 complex. However, it has been proposed that SUMOylation by Smt3 in yeast is responsible for propagating the assembly of the SC (Leung *et al*., 2015). Therefore, the Zip1 N-terminus may be incorporated into the central element by recursive SUMOylation events. It has been demonstrated *in vivo* that the presence of Zip1 and Gmc2 is required for Ecm11 SUMOylation (Humphryes *et al*., 2013), with Gmc2 likely needed for the formation of the Ecm11-Gmc2 complex. As a direct interaction has been identified between the Zip1 N-terminus and Smt3 through Y2H screening, it could be postulated that Zip1 recruits Smt3 to Ecm11 in the central element (Figure 5.39). A direct interaction between Smt3 and Ecm11 has also been identified, with the interaction abolished by removal of the N-terminus, and therefore the removal of the K5 and K101 SUMO sites.

Ecm11 and Gmc2 are recruited to the synapsis initiation complex (SIC) where they likely form the characterised 2:2 complex. Zip1 is then required for the recruitment of Smt3, with the N-terminal region sufficient for this (Humphryes *et al*., 2013), where Smt3 SUMOylates the Ecm11 N-terminal tail. The Ecm11 SUMOylation recruits further Zip1

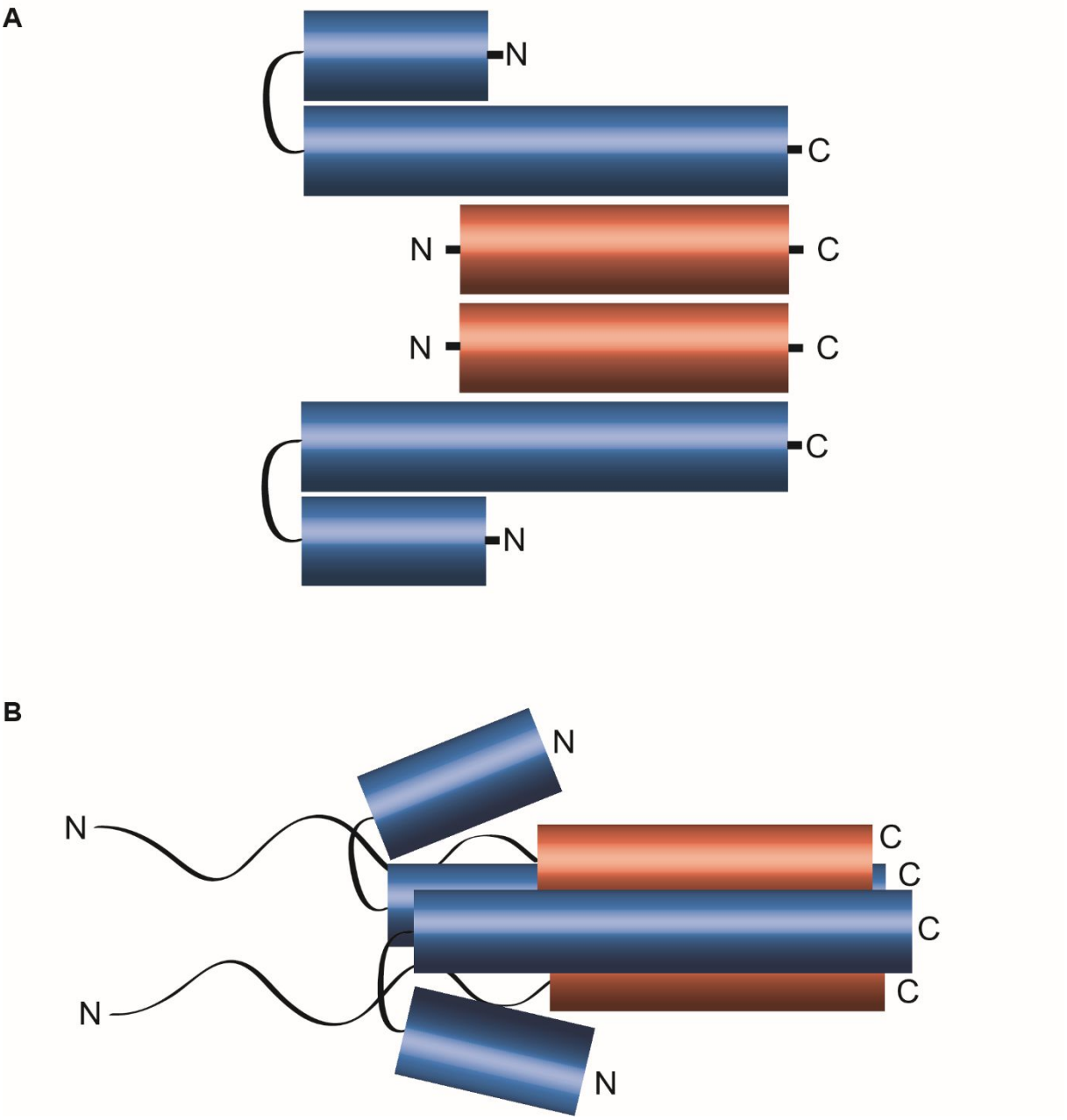


Figure 5.38. Schematic models for the Ecm11-Gmc2 complex. Ecm11 in orange and Gmc2 in blue (A) 2D schematic of the Ecm11$_{230-302}$-Gmc2 complex. The Ecm11-Gmc2 interaction occurs through the C-termini in a parallel nature, with the Gmc2 N-terminus folding back. (B) 3D schematic of the Ecm11-Gmc2 complex. The structured Ecm11 and Gmc2 chains may form a helical bundle, with the Gmc2 N-terminus stabilising the structure. The Ecm11 N-terminal region is largely unstructured and may protrude from the structural core of the complex.

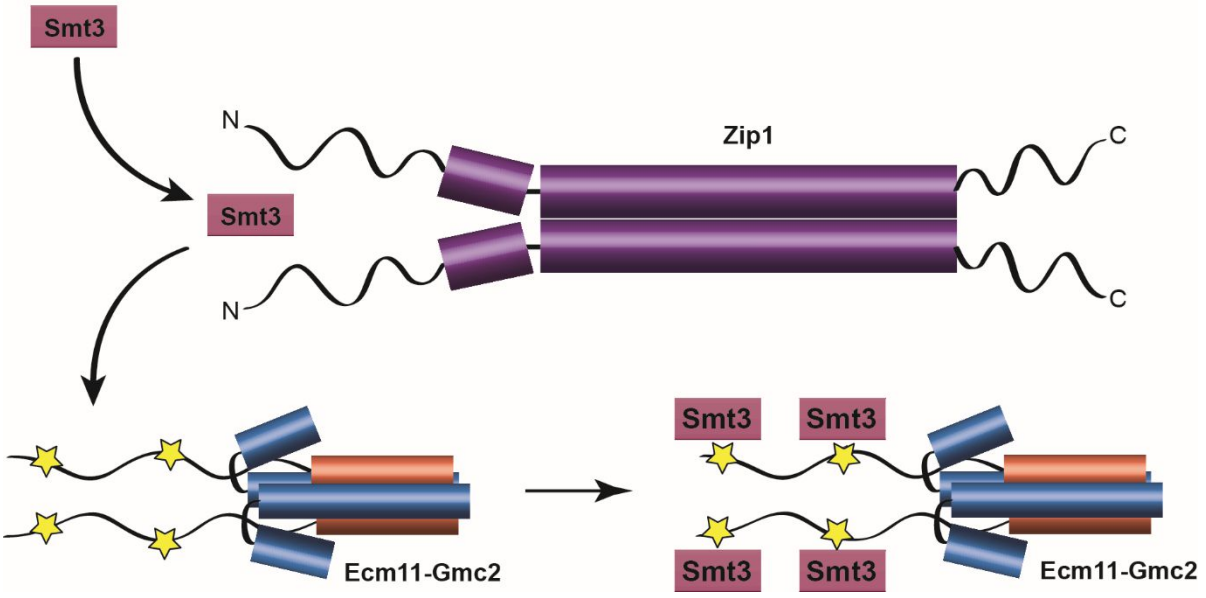


Figure 5.39. Schematic for the recruitment of Smt3 to Ecm11. Ecm11 in orange, Gmc2 in blue and yellow stars highlighting K5 and K101 SUMO sites in Ecm11. Smt3 directly interacts with the Zip1 N-terminus which then recruits Smt3 to the Ecm11-Gmc2 complex where it SUMOylates the Ecm11 N-terminal K5 and K101 residues.

molecules, and therefore more Smt3. The SUMOylation of Ecm11 also encourages the assembly of Zip1 into the SC in preference to the poly-complexes. This positive feedback loop allows for the recursive assembly of the Ecm11-Gmc2 complex with opposing Zip1 transverse filaments. This process is driven by the Smt3 SUMOylation events in order for the elongation of the central element to form the mature and functional SC structure (Figure 5.40).

It is proposed that there are two stages of SUMOylation that occur during yeast meiosis in order to assemble the fully functional SC. The first stage involves the initial recruitment of Zip1 to the SIC sites in order to form initial sites of synapsis. The second stage then involves the extensive Zip1 assembly to polymerise throughout the central element (Voelkel-Meiman *et al.*, 2013). Other SUMO events occur during yeast meiosis such as the SUMOylation of Rad52 and Sgs1, which have roles in DSB repair. It has been suggested that SUMOylation of DNA repair proteins has caused re-localisation of damaged DNA to sites of DNA repair (Nottke *et al.*, 2017). Furthermore, the topoisomerase Top2 in yeast has been shown to be SUMOylated during meiosis. The SUMOylated Top2 protein is thought to interact with Zip1 at centromeres

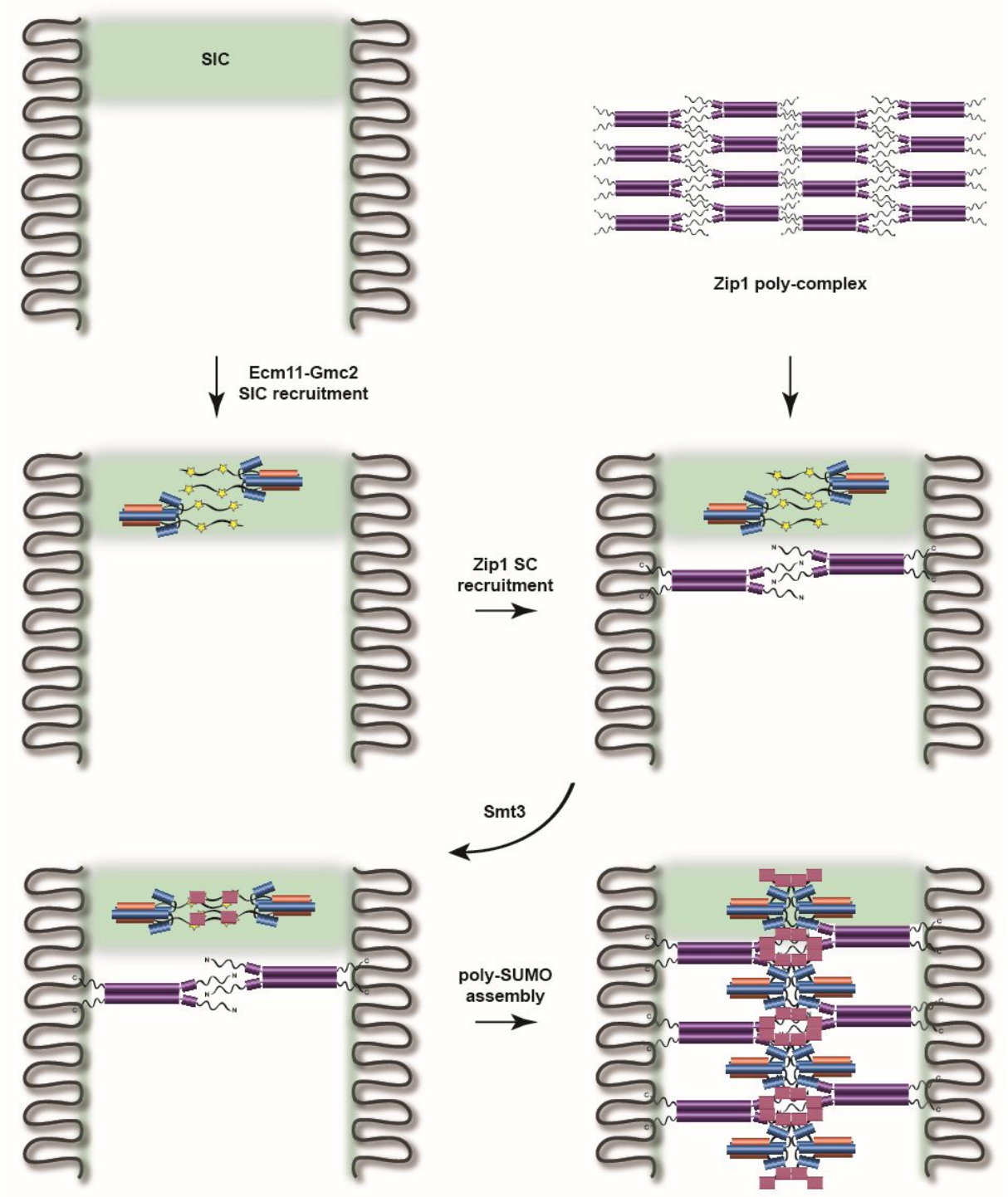


Figure 5.40. Schematic for central element assembly through poly-SUMOylation. SIC (synapsis initiation complex) in green, Zip1 in purple, Smt3 in pink, Ecm11 in orange, Gmc2 in blue and yellow stars highlighting K5 and K101 SUMO sites in Ecm11. Ecm11-Gmc2 is recruited to the SIC which encourages Zip1 SC recruitment from poly-complexes. Zip1 recruits Smt3 that SUMOylates the Ecm11 N-terminus. The SUMOylation of Ecm11 recruits more Zip1 molecules to form a positive feedback loop that causes elongation of the SC through recursive assembly by poly-SUMOylation.

(Cheng *et al*., 2006), thus suggesting that Top2 may have a role in chromosome synapsis. A range of proteins have been identified to be SUMOylated during yeast meiosis by Smt3 but the exact molecular role that SUMOylation plays in choreographing meiotic events remains unclear. However, the range of meiotic proteins that appear to be regulated by Smt3 modifications highlights the importance of SUMOylation in meiotic cell division. It is likely that the targeted Smt3 modifications act as signalling pathways in order to recruit necessary components. Also, the reversible nature of SUMOylation by the ULP enzymes provides a potential mechanism for SC disassembly in yeast (Gareau and Lima, 2010). The extensive SUMOylation of SC components by Smt3, which drives SC polymerisation (Voelkel-Meiman *et al*., 2013), may be removed by the ULP enzymes and thereby causing disassociation of the SC components in a timely manner during meiotic division.

Chapter Six.

Conclusions

6.1 SYCE3 is incorporated into the SYCP1 lattice for SC progression

The mammalian SC transverse filament protein undergoes recursive self-assembly at both its N-and C-termini, which allows for the formation of a preliminary SC lattice. However, this rudimentary lattice is not sufficient for the synapsis of homologous chromosomes and therefore, requires stabilisation by the central element. The central element components act in concert to form the protein-dense midline structure that can assemble along the entire length of homologues to achieve SC maturation upon regular incorporation of the transverse filaments.

Through phenotyping during mammalian prophase I, the overall pattern of central element recruitment, and the respective roles of each component, were eluded to by studying the impact on homologue synapsis upon the deletion of each protein. These studies showed that the lateral element proteins were first recruited to the axial element structures, which then form the lateral elements upon SC assembly. The transverse filaments are then recruited to the chromosome axis and undergo self-assembly to form initial alignments through head-to-head interactions in the midline. The central element proteins then have two sub-groups. The initiation complex is formed by SYCE1, SYCE3 and SIX6OS1 and is required for the initiation of synapsis, with deletion of any of these three proteins resulting in a complete failure of homologue synapsis. The SYCE2-TEX12 complex forms the elongation complex, which is required for the extension of the central element for full chromosome synapsis. The absence of either SYCE2 or TEX12 results in a failure of complete SC formation, but short stretches of synapsis do occur, due to the ability of the initiation complex to form regions of synapsis initiation. These studies indicate that the initial stabilisation of the transverse filaments upon central element assembly is performed by SYCE3.

In this thesis, the N-terminus of SYCP1, which resides in the central SC region, has been shown to directly interact with the central element component SYCE3 by Y2H interaction studies and *in vitro* biophysical analysis, consistent with previous reports. The respective SYCP1 region exists as a homo-tetramer and the free SYCE3 forms a compact dimer that undergoes higher-order assembly. Upon binding of SYCE3 to the SYCP1 N-terminal fragment, both proteins undergo remodelling in order to form a 2:1 complex. This suggests that the stabilisation of the SYCP1 preliminary lattice is achieved by disrupting its N-terminal self-assembly for SYCE3 incorporation. The hetero-trimeric complex has been shown here to form an elongated complex, with the likely interpretation that all three chains are in an extended conformation and in a parallel arrangement.

In a previous report, the SYCP1 101-111 residues were shown to form head-to-head interactions to form an elongated tetramer between two opposing parallel dimers. In this study, these head-to-head interactions are maintained when SYCE3 is bound to SYCP1. This has been

demonstrated by SEC-MALS analysis showing the formation of a 2:1 species and also a 4:2 species. SEC-SAXS analysis and examination of the solution structures for both species shows that the Rc for both species are similar but the length of the 4:2 species is much longer and therefore, the 4:2 complex is likely formed by the SYCP1 head-to-head interaction. This was further examined by mutation or removal of the SYCP1 101-111 amino acid residues. The analysis of both the truncated and mutated forms resulted in the formation of a single 2:1 species when co-purified with SYCE3. This provides further confidence that the 4:2 species is indeed formed by a head-to-head interaction between opposing SYCP1-SYCE3 complexes as the SYCP1 101-111 region is shown to form an antiparallel four helical association between opposing SYCP1 dimers in the SYCP1 N-terminal X-ray crystal structures. These findings suggests that although SYCE3 binding disrupts the existing SYCP1 lattice, it does not abolish the general mechanisms of SYCP1 assembly. Therefore, SYCE3 may act to guide and regulate the incorporation of the transverse filaments upon central element assembly, but SYCE3 binding does not prevent the underlying SYCP1 tethering.

SYCE3 has previously been show to form a compact dimeric four-helical bundle in the X-ray crystal structure, with solution data in this study confirming the presence of the same dimeric conformation. However, SYCE3 also undergoes higher-order assemblies to form other distinct species than the SYCE3 dimer. Through dissection of the potential mechanisms governing these assemblies, it has been revealed that opening of the SYCE3 loop is required to form these larger assemblies. This has been identified by mutation of the SYCE3 loop to impose a closed loop structure and in a separate mutant, to force an open structure. The analysis of both of these mutants shows that the forced closed loop structure can only form the compact dimer form, as observed in the X-ray crystal structure. The forced open loop structure was observed to form large molecular weight species and does not form the compact dimeric bundle. Further mutational studies of SYCE3 highlights two sets of amino acid residues that also appear to be implicated in the higher-order assemblies. The N-terminal Y11 and L15 resides stabilise the elongated tetrameric form and mutation of these residues abolishes SYCE3 assemblies. The W41 and Y44 residues seemingly enhance the higher-order assemblies as mutation of these residues also prevents assembly formation. However, these W41 and Y44 mutations do not abolish the formation of the extended SYCE3 tetramer and so W41 and Y44 are proposed to establish staggered lateral interactions between the SYCE3 dimers and tetramers in order to form the large assemblies. Therefore, a model is proposed whereby the larger species are formed by opening of the SYCE3 loops and capping by closed loop SYCE3 chains in order to produce a tetrameric species where each half contains the antiparallel helical core that is observed in the crystal structure. These tetramer species are stabilised by amino acid residues

Y11 and L15 and the tetramer species can then produce larger assemblies by the formation of staggered lateral interactions governed by W41 and Y44.

The non-assembling SYCE3 mutants that were used to decipher the mechanisms of SYCE3 assembly were all found to maintain binding to SYCP1, other than the imposed closed dimer form of SYCE3, thus suggesting that SYCE3 binds to SYCP1 in the extended form. The biological relevance of these assemblies was investigated by mixing free SYCE3 with SYCP1-SYCE3 complexes, which elucidated the capacity of the free SYCE3 material to tether the 2:1 SYCP1-SYCE3 complexes. This suggests that initial SYCP1 disruption is achieved by SYCE3 binding to form the hetero-trimer complex. Further transverse filament stabilisation is then performed through the self-assembly of SYCE3 between SYCP1-SYCE3 lateral units in order to form the initial parallel central element cables. It could be postulated that the open chain SYCE3 in complex with SYCP1 provides an interaction site for free SYCE3 assemblies to bind and that neighbouring SYCP1-SYCE3 units cap the SYCE3 assemblies.

SYCE3 has also been shown to interact with the SYCE1 and SYCE2 central element proteins by Y2H studies and biochemical *in vitro* interaction studies. Further to this, SYCE1 interacts with SIX6OS1 and SYCE2 forms a complex with TEX12. Therefore, SYCE3 is proposed to act as the molecular adaptor within the central element to bridge the various components in order to achieve SC maturation. SYCE3, SYCE1 and SIX6OS1 are thought to act as the central element initiation complex to prime the preliminary SYCP1 lattice and provide vertical support prior to full SC maturation. The SYCE2-TEX12 complex is then proposed to form the central element elongation complex that undergoes fibrous assemblies to extend the central element, and therefore the full SC, along the entire length of chromosomes. As SYCE3 has been shown to interact with the SYCE2 and to the SYCE2-TEX12 complex, it is thought that SYCE3 may provide interaction sites in the central element to guide SYCE2-TEX12 fibrous assembly.

6.2 SYCE2-TEX12 form fibrous assemblies to achieve SC elongation

SYCE2 and TEX12 were proposed to be responsible for the elongation of the SC due to their respective phenotypes upon their deletion, where short regions of synapsis occur by assembly of the other central element proteins and initial stabilisation of SYCP1. However, these initial synapsed stretches do not propagate and therefore full synapsis is not achieved. A previous study showed that SYCE2 and TEX12 form an equimolar complex that undergoes higher-order assembly, with these assemblies thought to drive SC extension by connecting the central element structure.

In this thesis, the mechanisms of the SYCE2-TEX12 assemblies were investigated, with the additional finding that TEX12 forms a stable unit in the absence of SYCE2. The study of TEX12 in isolation, and when in complex with SYCE2, have elucidated mechanistic insights into the molecular basis of the SYCE2-TEX12 assemblies.

The C-terminus of TEX12 was found to be crucial for the formation of the previously reported 4:4 complex and subsequent filament formation. Removal and mutation of TEX12 amino acid residues 114-123 did not prevent an interaction between SYCE2 and TEX12, but rather restricted the oligomeric state to a non-assembling 2:2 complex in solution. This 2:" complex was characterised by solution scattering experiments and revealed the formation of an elongated complex with a similar length to the 4:4 complex but a cross-sectional radius half of the 4:4 complex. This suggests that the TEX12 C-terminus mediates 4:4 complex formation, and subsequent higher-order assembly, by the interaction between two stable elongated 2:2 units. The corresponding truncation for TEX12 in isolation caused a switch from a compact dimeric state to an extended structure, with an indication of a larger oligomer in solution. Further investigation into the mechanisms underlying this observation revealed two modalities of the TEX12 C-terminus that contribute to the higher-order assembly.

Coiled-coil heptads at the TEX12 C-termini were found to drive the formation of the 4:4 complex and subsequent filament formation, and were also necessary to maintain the compact dimer state of the TEX12 alpha-helical core. This was identified by mutating four hydrophobic residues in the TEX12 C-terminus that are predicted to fall in the "*a*" and "*d*" positions in the coiled coil heptad. The mutation of these residues caused TEX12 to switch from a compact dimeric form to an extended stable tetramer and caused SYCE2-TEX12 to form a single 2:2 species, in both the context of the full length and core helical complexes.

Another group of TEX12 C-terminal residues, which are predicted to form a surface exposed interface away from the hydrophobic coiled-coil face, also had implications for TEX12 structure and a milder effect on SYCE2-TEX12 assembly. Mutations of these residues caused TEX12 to form an extended dimer in the X-ray crystal structure with protrusion of the C-termini. The exposed C-termini then undergo coiled-coil interactions to form a "dimer of dimers" extended structure and also a tetramer interface. These coiled-coil interactions are driven by the four hydrophobic residues that were mutated to block SYCE2-TEX12 4:4 complex formation. The mutation of the surface exposed residues did not prevent SYCE2-TEX12 4:4 complex formation and did not prevent higher-order assembly of the full length complex. However, the extent of higher-order assembly in the core helical complex was compromised upon mutation of these residues, thereby suggesting that the surface exposed patch has a contributory role for higher order assembly.

Therefore, although the molecular details of the SYCE2-TEX12 interaction and the various oligomeric states remain unclear, the TEX12 C-terminal coiled-coil is vital for the 4:4 complex formation and the higher-order structures, with the surface exposed residues likely contributing to transient interactions between SYCE2-TEX12 units to aid filament formation. Crystals of the SYCE2-TEX12 complex were found to produce an X-ray fibre diffraction pattern that is consistent with the k-m-e-f class of coiled-coil proteins. This finding supports the idea that the filamentous structures detected both in solution and by TEM experiments are indeed fibrous coiled-coil assemblies that are likely biologically relevant for the role of SYCE2-TEX12 as the central element elongation complex.

From combining the structural and biophysical characterisation of the various complexes studied in this thesis, along with the reported *in vivo* data, a model is proposed for the molecular basis of the mammalian central element assembly in order to achieve SC maturation.

In this model SYCE3 acts as the molecular adaptor that recruits each central element protein, whether directly or indirectly, with the initial function of SYCE3 to disrupt the preliminary SYCP1 lattice. The formation of the SYCP1-SYCE3 complex within the parallel cables in the central region allows for the stabilisation and regular incorporation of the transverse filaments into the central element. Further SYCE3 molecules are then able to self-assemble between the lateral SYCP1-SYCE3 units to stabilise the SYCP1 incorporation to form the bilayered junction. SYCE3 then recruits SYCE1, which in turn binds to SIX6OS1, to the central element. The proposed SYCE1 function is to provide the vertical support for SYCP1 and the central element, which is likely aided by the interaction with SIX6OS1. The iterative recruitment of SYCE3, SYCE1 and SIX6OS1 allows for the stabilisation of the SYCP1 transverse filaments and for short range formation of vertical and longitudinal central element support and synapsis initiation.

The SYCE2-TEX12 complex is then incorporated into the central element through periodical interactions with SYCE3. SYCE3 may provide the recognition platform to allow the guided fibrous assembly of SYCE2-TEX12 along the central element in order to drive the extension of the SC by connecting the short regions of synapsis initiation through longitudinal and transverse support.

6.3 Ecm11 and Gmc2 form an equimolar complex in the yeast central element

The yeast SC protein components differ to the mammalian SC, but the two SC structures form the same overall architecture and both achieve homologue synapsis for the repair of DSBs and the formation of crossovers for meiotic progression. Therefore, the underlying mechanisms

are likely comparable. Many meiotic events have been characterised in yeast to elucidate the mechanisms that occur in mammals. Therefore, the *in vitro* characterisation of the yeast central element, with the amenability to perform the corresponding experiments *in vivo*, could be used to infer mechanisms of the mammalian SC functions.

In this thesis, the N-terminal region of the Zip1 transverse filament protein was investigated for comparison to the properties of the mammalian SYCP1 transverse filament and to understand the incorporation into the central element. This was studied by expression, purification and biophysical analysis of a variety of Zip1 N-terminal constructs. This analysis identified a putative dimerisation site within the Zip1 N-terminal region, with a basic patch that was found to bind to double stranded DNA. However, the low thermal stability and minimal oligomerisation suggested that the Zip1 N-terminus was not in its native state during analysis. This could also suggest that the Zip1 N-terminus requires a binding partner for stabilisation, potentially with a central element component as identified for SYCP1.

The Ecm11-Gmc2 complex has been characterised *in vivo* to form the yeast central element. However, these proteins had not previously been purified for *in vitro* characterisation and isolation of the Ecm11-Gmc2 complex. Ecm11 was found to be unstable when expressed without Gmc2 but Gmc2 was stable in the absence of Ecm11. The full length Gmc2 protein was predominantly monomeric in solution but some dimeric species were observed. Thermal denaturation revealed a non-cooperative melting curve indicative of gradual unwinding of an alpha helix. Therefore, it is likely that Ecm11 and Gm2 require each other for stability.

By co-expression of a series of Ecm11 and Gmc2 constructs, the Ecm11-Gmc2 complex was isolated and found to form a stable unit, which was characterised by a range of biophysical techniques. An alpha helical core region at the Ecm11 C-terminus was found to form a stable 2:2 complex with Gmc2, with the inferred organisation of the chains to be in a parallel arrangement according to solution scattering experiments.

The N-terminus and central region of Ecm11 is predicted to be largely unstructured, and contains the two lysine residues that are SUMOylated by Smt3 to induce central element assembly. The SUMO modification of the Ecm11 unstructured region was reported to occur in a Gmc2-dependant manner. Based on the structural studies of Ecm11-Gmc2, it is possible that Gmc2 is required to stabilise Ecm11 through the interaction with the structured Ecm11 C-terminal helical region, with the 2:2 Ecm11-Gmc2 complex allowing for the protrusion of the Ecm11 N-terminal region and subsequent access for SUMOylation.

These 2:2 units may then form the building blocks of the central element assembly, with the incorporation of the transverse filaments achieved by the previously reported polySUMO feedback loop that causes the iterative recruitment of Ecm11-Gmc2 and further transverse

filaments. This suggests that, unlike the mammalian SYCE2-TEX12 complex that self-assembles to achieve SC elongation, the Ecm11-Gmc2 units are tethered and assembled as a result of SUMOylation.

6.4 Overall conclusions

In this thesis, three key meiotic assembly mechanisms have been investigated. Preliminary analysis of the yeast SC central region demonstrates that Ecm11-Gmc2 form a stable alpha helical 2:2 complex with the unstructured Ecm11 allowing for SUMOylation and subsequent incorporation of the transverse filaments. The study of the sub-stages of mammalian central element progression has built upon existing data to provide molecular insights into the mechanistic events of central element assembly. SYCE3 has been shown to act as the molecular adaptor of the SC central region by providing the sole link between the transverse filaments and the central element. Further to its bridging function with the transverse filaments, SYCE3 acts as molecular glue to recruit further central element proteins in a two-step process. Firstly, the SYCE1-SIX6OS1 complex is recruited to the SYCE3-recapitulated SYCP1 lattice in order to form the initiation complex, by providing short regions of three dimensional central element structure in order to initiate synapsis. SYCE3 then provides a recognition platform for SYCE2-TEX12 fibrous assembly to mediate SC maturation for the completion of successful and faithful homologous chromosome synapsis that is vital for meiotic progression and consequent fertility.

Publications in preparation

Salmon, L. J., Dunce, J. D., Davies, O. D., SYCE2-TEX12 fibrous assemblies mediate long-range meiotic chromosome synapsis.

Dunne, O. M., Salmon, L. J., Dunce, J. D., Davies, O. D., Structural maturation of SYCP1-mediated meiotic chromosome synapsis through conformational remodelling by molecular adapter SYCE3.

Salmon, L.J., Pastok, M.W., Sandhu, S., Hunter, J. E., Wilson, C.L., Perkins, N. D., Jennings, C., Hunter, N., Davies, O.R., McClurg, U. L., Meiotic TEX12 is a novel centrosomal protein controlling fidelity of chromosome division in cancer cells.

Appendix I- *E. coli* codon optimised sequences used in this thesis (Chapter 2, Section 2.2.1)

SYCP1

```
ATGGAAAAACAGAAACCGTTTGCGCTGTTTGTGCCGCCGCGCAGCAGCAGCAGCCAGGTG
 M  E  K  Q  K  P  F  A  L  F  V  P  P  R  S  S  S  S  Q  V
AGCGCGGTGAAACCGCAGACCCTGGGCGGCGATAGCACCTTTTTTAAAAGCTTTAACAAA
 S  A  V  K  P  Q  T  L  G  G  D  S  T  F  F  K  S  F  N  K
TGCACCGAAGATGATTTTGAATTTCCGTTTGCGAAAACCAACCTGAGCAAAAACGGCGAA
 C  T  E  D  D  F  E  F  P  F  A  K  T  N  L  S  K  N  G  E
AACATTGATAGCGATCCGGCGCTGCAGAAAGTGAACTTTCTGCCGGTGCTGGAACAGGTG
 N  I  D  S  D  P  A  L  Q  K  V  N  F  L  P  V  L  E  Q  V
GGCAACAGCGATTGCCATTATCAGGAAGGCCTGAAAGATAGCGATCTGGAAAACAGCGAA
 G  N  S  D  C  H  Y  Q  E  G  L  K  D  S  D  L  E  N  S  E
GGCCTGAGCCGCGTGTATAGCAAACTGTATAAAGAAGCGGAAAAAATTAAAAAATGGAAA
 G  L  S  R  V  Y  S  K  L  Y  K  E  A  E  K  I  K  K  W  K
GTGAGCACCGAAGCGGAACTGCGCCAGAAAGAAAGCAAACTGCAGGAAAACCGCAAAATT
 V  S  T  E  A  E  L  R  Q  K  E  S  K  L  Q  E  N  R  K  I
ATTGAAGCGCAGCGCAAAGCGATTCAGGAACTGCAGTTTGGCAACGAAAAAGTGAGCCTG
 I  E  A  Q  R  K  A  I  Q  E  L  Q  F  G  N  E  K  V  S  L
AAACTGGAAGAAGGCATTCAGGAAAACAAAGATCTGATTAAAGAAAACAACGCGACCCGC
 K  L  E  E  G  I  Q  E  N  K  D  L  I  K  E  N  N  A  T  R
CATCTGTGCAACCTGCTGAAAGAAACCTGCGCGCGCAGCGCGGAAAAAACCAAAAAATAT
 H  L  C  N  L  L  K  E  T  C  A  R  S  A  E  K  T  K  K  Y
GAATATGAACGCGAAGAAACCCGCCAGGTGTATATGGATCTGAACAACAACATTGAAAAA
 E  Y  E  R  E  E  T  R  Q  V  Y  M  D  L  N  N  N  I  E  K
ATGATTACCGCGTTTGAAGAACTGCGCGTGCAGGCGGAAAACAGCCGCCTGGAAATGCAT
 M  I  T  A  F  E  E  L  R  V  Q  A  E  N  S  R  L  E  M  H
TTTAAACTGAAAGAAGATTATGAAAAAATTCAGCATCTGGAACAGGAATATAAAAAAGAA
 F  K  L  K  E  D  Y  E  K  I  Q  H  L  E  Q  E  Y  K  K  E
ATTAACGATAAAGAAAAACAGGTGAGCCTGCTGCTGATTCAGATTACCGAAAAAGAAAAC
 I  N  D  K  E  K  Q  V  S  L  L  L  I  Q  I  T  E  K  E  N
AAAATGAAAGATCTGACCTTTCTGCTGGAAGAAAGCCGCGATAAAGTGAACCAGCTGGAA
 K  M  K  D  L  T  F  L  L  E  E  S  R  D  K  V  N  Q  L  E
GAAAAAACCAAACTGCAGAGCGAAAACCTGAAACAGAGCATTGAAAAACAGCATCATCTG
 E  K  T  K  L  Q  S  E  N  L  K  Q  S  I  E  K  Q  H  H  L
ACCAAAGAACTGGAAGATATTAAAGTGAGCCTGCAGCGCAGCGTGAGCACCCAGAAAGCG
 T  K  E  L  E  D  I  K  V  S  L  Q  R  S  V  S  T  Q  K  A
CTGGAAGAAGATCTGCAGATTGCGACCAAAACCATTTGCCAGCTGACCGAAGAAAAAGAA
 L  E  E  D  L  Q  I  A  T  K  T  I  C  Q  L  T  E  E  K  E
ACCCAGATGGAAGAAAGCAACAAAGCGCGCGCGGCGCATAGCTTTGTGGTGACCGAATTT
 T  Q  M  E  E  S  N  K  A  R  A  A  H  S  F  V  V  T  E  F
GAAACCACCGTGTGCAGCCTGGAAGAACTGCTGCGCACCGAACAGCAGCGCCTGGAAAAA
 E  T  T  V  C  S  L  E  E  L  L  R  T  E  Q  Q  R  L  E  K
AACGAAGATCAGCTGAAAATTCTGACCATGGAACTGCAGAAAAAAAGCAGCGAACTGGAA
 N  E  D  Q  L  K  I  L  T  M  E  L  Q  K  K  S  S  E  L  E
GAAATGACCAAACTGACCAACAACAAAGAAGTGGAACTGGAAGAACTGAAAAAAGTGCTG
 E  M  T  K  L  T  N  N  K  E  V  E  L  E  E  L  K  K  V  L
GGCGAAAAAGAAACCCTGCTGTATGAAAACAAACAGTTTGAAAAAATTGCGGAAGAACTG
 G  E  K  E  T  L  L  Y  E  N  K  Q  F  E  K  I  A  E  E  L
AAAGGCACCGAACAGGAACTGATTGGCCTGCTGCAGGCGCGCGAAAAAGAAGTGCATGAT
 K  G  T  E  Q  E  L  I  G  L  L  Q  A  R  E  K  E  V  H  D
CTGGAAATTCAGCTGACCGCGATTACCACCAGCGAACAGTATTATAGCAAAGAAGTGAAA
 L  E  I  Q  L  T  A  I  T  T  S  E  Q  Y  Y  S  K  E  V  K
```

Appendix I continued

```
GATCTGAAAACCGAACTGGAAAACGAAAAACTGAAAAACACCGAACTGACCAGCCATTGC
 D  L  K  T  E  L  E  N  E  K  L  K  N  T  E  L  T  S  H  C
AACAAACTGAGCCTGGAAAACAAAGAACTGACCCAGGAAACCAGCGATATGACCCTGGAA
 N  K  L  S  L  E  N  K  E  L  T  Q  E  T  S  D  M  T  L  E
CTGAAAAACCAGCAGGAAGATATTAACAACAACAAAAAACAGGAAGAACGCATGCTGAAA
 L  K  N  Q  Q  E  D  I  N  N  N  K  K  Q  E  E  R  M  L  K
CAGATTGAAAACCTGCAGGAAACCGAAACCCAGCTGCGCAACGAACTGGAATATGTGCGC
 Q  I  E  N  L  Q  E  T  E  T  Q  L  R  N  E  L  E  Y  V  R
GAAGAACTGAAACAGAAACGCGATGAAGTGAAATGCAAACTGGATAAAAGCGAAGAAAAC
 E  E  L  K  Q  K  R  D  E  V  K  C  K  L  D  K  S  E  E  N
TGCAACAACCTGCGCAAACAGGTGGAAAACAAAAACAAATATATTGAAGAACTGCAGCAG
 C  N  N  L  R  K  Q  V  E  N  K  N  K  Y  I  E  E  L  Q  Q
GAAAACAAAGCGCTGAAAAAAAAAGGCACCGCGGAAAGCAAACAGCTGAACGTGTATGAA
 E  N  K  A  L  K  K  K  G  T  A  E  S  K  Q  L  N  V  Y  E
ATTAAAGTGAACAAACTGGAACTGGAACTGGAAAGCGCGAAACAGAAATTTGGCGAAATT
 I  K  V  N  K  L  E  L  E  L  E  S  A  K  Q  K  F  G  E  I
ACCGATACCTATCAGAAAGAAATTGAAGATAAAAAAATTAGCGAAGAAAACCTGCTGGAA
 T  D  T  Y  Q  K  E  I  E  D  K  K  I  S  E  E  N  L  L  E
GAAGTGGAAAAAGCGAAAGTGATTGCGGATGAAGCGGTGAAACTGCAGAAAGAAATTGAT
 E  V  E  K  A  K  V  I  A  D  E  A  V  K  L  Q  K  E  I  D
AAACGCTGCCAGCATAAAATTGCGGAAATGGTGGCGCTGATGGAAAAACATAAACATCAG
 K  R  C  Q  H  K  I  A  E  M  V  A  L  M  E  K  H  K  H  Q
TATGATAAAATTATTGAAGAACGCGATAGCGAACTGGGCCTGTATAAAAGCAAAGAACAG
 Y  D  K  I  I  E  E  R  D  S  E  L  G  L  Y  K  S  K  E  Q
GAACAGAGCAGCCTGCGCGCGAGCCTGGAAATTGAACTGAGCAACCTGAAAGCGGAACTG
 E  Q  S  S  L  R  A  S  L  E  I  E  L  S  N  L  K  A  E  L
CTGAGCGTGAAAAAACAGCTGGAAATTGAACGCGAAGAAAAAGAAAAACTGAAACGCGAA
 L  S  V  K  K  Q  L  E  I  E  R  E  E  K  E  K  L  K  R  E
GCGAAAGAAAACACCGCGACCCTGAAAGAAAAAAAAGATAAAAAAACCCAGACCTTTCTG
 A  K  E  N  T  A  T  L  K  E  K  K  D  K  K  T  Q  T  F  L
CTGGAAACCCCGGAAATTTATTGGAAACTGGATAGCAAAGCGGTGCCGAGCCAGACCGTG
 L  E  T  P  E  I  Y  W  K  L  D  S  K  A  V  P  S  Q  T  V
AGCCGCAACTTTACCAGCGTGGATCATGGCATTAGCAAAGATAAACGCGATTATCTGTGG
 S  R  N  F  T  S  V  D  H  G  I  S  K  D  K  R  D  Y  L  W
ACATCTGCCAAAAATACTTTATCTACACCATTGCCAAAGGCATATACAGTGAAGACACCA
 T  S  A  K  N  T  L  S  T  P  L  P  K  A  Y  T  V  K  T  P
ACAAAACCAAAACTACAGCAAAGAGAAAACTTGAATATACCCATTGAAGAAAGTAAAAAA
 T  K  P  K  L  Q  Q  R  E  N  L  N  I  P  I  E  E  S  K  K
AAGAGAAAAATGGCCTTTGAATTTGATATTAATTCAGATAGTTCAGAAACTACTGATCTT
 K  R  K  M  A  F  E  F  D  I  N  S  D  S  S  E  T  T  D  L
TTGAGCATGGTTTCAGAAGAAGAGACATTGAAAACACTGTATAGGAACAATAATCCACCA
 L  S  M  V  S  E  E  E  T  L  K  T  L  Y  R  N  N  N  P  P
GCTTCTCATCTTTGTGTCAAAACACCAAAAAAGGCCCCTTCATCTCTAACAACCCCTGGA
 A  S  H  L  C  V  K  T  P  K  K  A  P  S  S  L  T  T  P  G
TCTACACTGAAGTTTGGAGCTATAAGAAAAATGCGGGAGGACCGTTGGGCTGTAATTGCT
 S  T  L  K  F  G  A  I  R  K  M  R  E  D  R  W  A  V  I  A
AAAATGGATAGAAAAAAAAAACTAAAAGAAGCTGAAAAGTTATTTGTTTAA
 K  M  D  R  K  K  K  L  K  E  A  E  K  L  F  V  -
```

Appendix I continued

SYCE3

```
atggatgatgcggatccggaagaacgcaactatgataacatgctgaaaatgctgagcgat
 M  D  D  A  D  P  E  E  R  N  Y  D  N  M  L  K  M  L  S  D
ctgaacaaagatctggaaaaactgctggaagaaatggaaaaaattagcgtgcaggcgacc
 L  N  K  D  L  E  K  L  L  E  E  M  E  K  I  S  V  Q  A  T
tggatggcgtatgatatggtggtgatgcgcaccaacccgaccctggcggaaagcatgcgc
 W  M  A  Y  D  M  V  V  M  R  T  N  P  T  L  A  E  S  M  R
cgcctggaagatgcgtttgtgaactgcaaagaagaaatggaaaaaaactggcaggaactg
 R  L  E  D  A  F  V  N  C  K  E  E  M  E  K  N  W  Q  E  L
ctgcatgaaaccaaacagcgcctg
 L  H  E  T  K  Q  R  L
```

SYCE2

```
atggagcgacagggagtggacgtgccccatgtgaaatgcaaagaccaggaaccgcagccc
 M  E  R  Q  G  V  D  V  P  H  V  K  C  K  D  Q  E  P  Q  P
ttgggggagagcaaggagcatccgcggtgggaagagaactgcgaggaggaagctggtgga
 L  G  E  S  K  E  H  P  R  W  E  E  N  C  E  E  E  A  G  G
gggccagctagtgccagttgccagctgacggtcctggaagggaagtcgggactctacttc
 G  P  A  S  A  S  C  Q  L  T  V  L  E  G  K  S  G  L  Y  F
tcctctctggactcaagcattgacatcctgcagaagagagcccaggagctgatcgaaaac
 S  S  L  D  S  S  I  D  I  L  Q  K  R  A  Q  E  L  I  E  N
atcaacaagagccggcaaaaggaccatgcactcatgaccaacttcaggaacagcctgaag
 I  N  K  S  R  Q  K  D  H  A  L  M  T  N  F  R  N  S  L  K
accaaggtttcggatctgacagagaaattagaggagaggatctatcagatttataatgac
 T  K  V  S  D  L  T  E  K  L  E  E  R  I  Y  Q  I  Y  N  D
cacaacaagatcatccaggaaaagctccaagagttcacccagaaaatggcaaagatcagc
 H  N  K  I  I  Q  E  K  L  Q  E  F  T  Q  K  M  A  K  I  S
catttggagacagagctcaaacaagtctgccacagcgtggagactgtgtacaaagacctg
 H  L  E  T  E  L  K  Q  V  C  H  S  V  E  T  V  Y  K  D  L
tgtctccagcctgagcagagcctaagactcagatgggggccagaccactctaggggaaag
 C  L  Q  P  E  Q  S  L  R  L  R  W  G  P  D  H  S  R  G  K
tccccaccacgtcccggcaactcacagcccccagacgtgttcgtttcttctgtggctgaa
 S  P  P  R  P  G  N  S  Q  P  P  D  V  F  V  S  S  V  A  E
actacttctcaggccactgcttcagaagtacagaccaacagagatggtgaatgc
 T  T  S  Q  A  T  A  S  E  V  Q  T  N  R  D  G  E  C
```

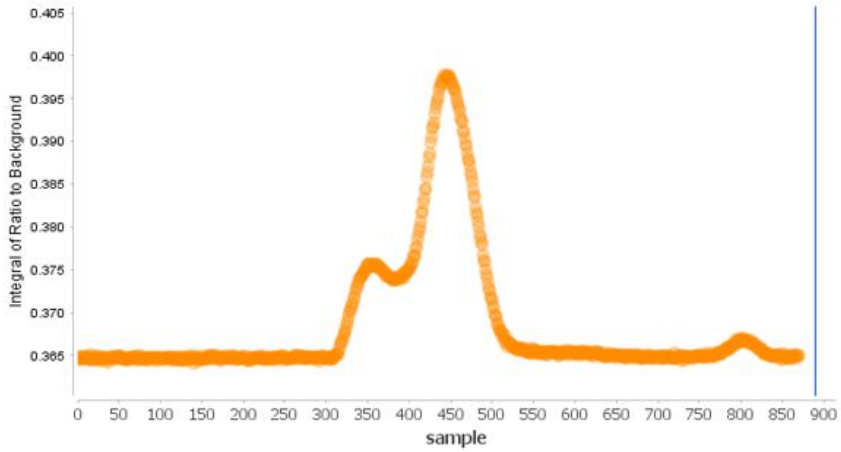
TEX12

```
atgatggcaaatcaccttgtaaagcctgataataggaattgcaagaggccaagagaattg
 M  M  A  N  H  L  V  K  P  D  N  R  N  C  K  R  P  R  E  L
gagtctccagtgccagatagtccacagctgtcctctcttggaaaatcagattcatctttc
 E  S  P  V  P  D  S  P  Q  L  S  S  L  G  K  S  D  S  S  F
tctgaaatttccggactattttataaagatgaagccttggagaaagatttaaatgatgtg
 S  E  I  S  G  L  F  Y  K  D  E  A  L  E  K  D  L  N  D  V
agcaaggaaattaatctaatgttgtctacctatgcaaagcttttaagtgagagagcagca
 S  K  E  I  N  L  M  L  S  T  Y  A  K  L  L  S  E  R  A  A
gtagatgcatcttacattgatgagatagatgaactcttcaaagaagccaatgctattgaa
 V  D  A  S  Y  I  D  E  I  D  E  L  F  K  E  A  N  A  I  E
aactttctaatacaaaaaagagagttcctgcgacagaggtttacagtgattgcaaacaca
 N  F  L  I  Q  K  R  E  F  L  R  Q  R  F  T  V  I  A  N  T
ttacacaga
 L  H  R
```

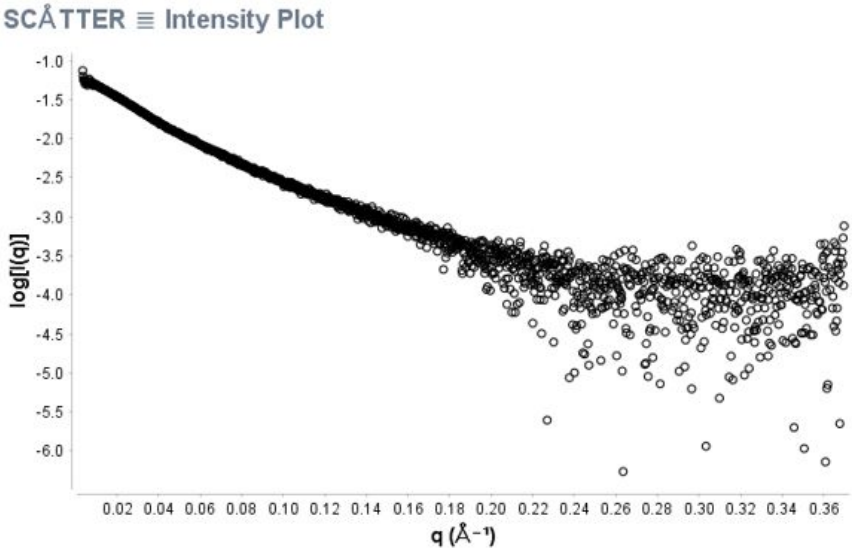
Appendix II- Size exclusion chromatography small-angle X-ray scattering (SEC-SAXS) processing and presentation (Chapter 2, Section 2.6.5)

SEC-SAXS data collection was performed at the B21 beamline, Diamond Light Source, and data was processed using ScÅtter 3.0 software and programs from the ATSAS suite. Data collection and initial processing was performed using the ScÅtter 3.0 software.

The scattering signal across the elution profile was inspected for an appropriate baseline region for selection to be used for buffer signal subtraction. The sample peaks were then selected and the corresponding buffer signal was subtracted to allow analysis of the sample peak. The subtracted scattering data was generated for each frame within the selected sample peak. The log intensity plot against scattering angle was manually inspected for each frame in order to select the appropriate frames within the peak to average, ensuring that frames with poor signal are omitted. The selected peaks were then averaged to produce the averaged scattering curve (see example below).

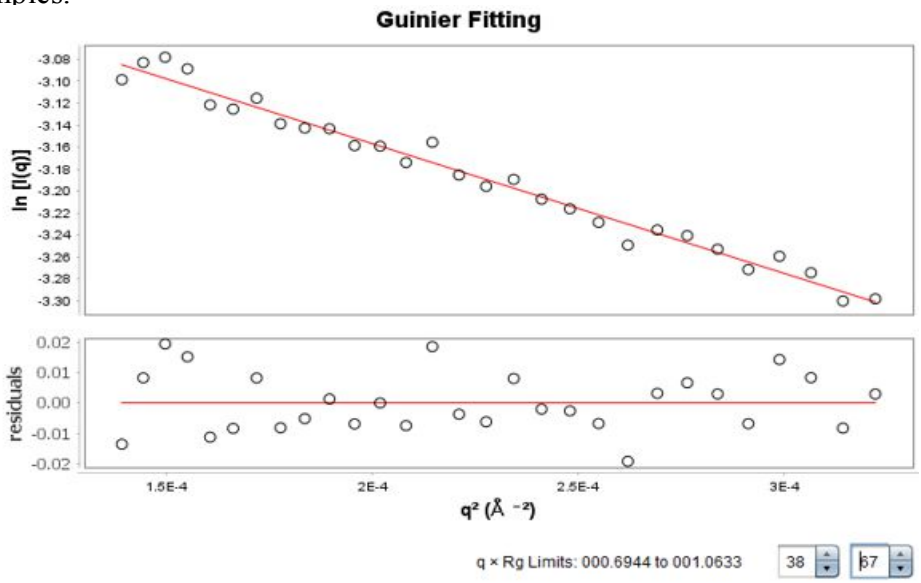


ScÅtter graph of scattering signal across the SEC elution profile

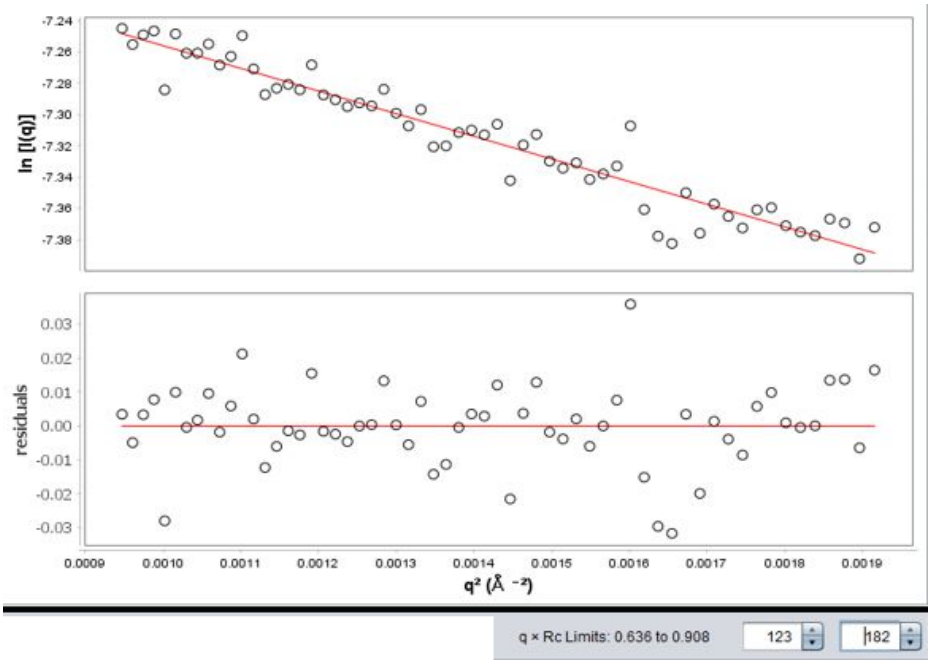


ScÅtter graph of intensity plot after buffer subtraction and peak frame averaging

The functions in the ScÅtter 3.0 software were then used to analyse the data. Guinier analysis was performed by selecting the first linear section within the low Q region of the ln I(Q) v. Q^2 ($Å^{-2}$) plot, where Q is the scattering angle and I is the scattering intensity. The data points were truncated to ensure the selected region had q x Rg limits of 1.3 or less and this region was then used to determine the radius of gyration (Rg), calculated by the software. The cross-sectional radius of gyration (Rc) was determined in a similar manner by selecting the first linear section within the low Q region of the ln IQ(Q) v. Q^2 ($Å^{-2}$) plot. The data points were truncated to ensure the selected region had q x Rg limits of 1.3 or less and this region was then used to determine the Rc. In some cases, Rc values could not be determined due to the globular nature of samples.

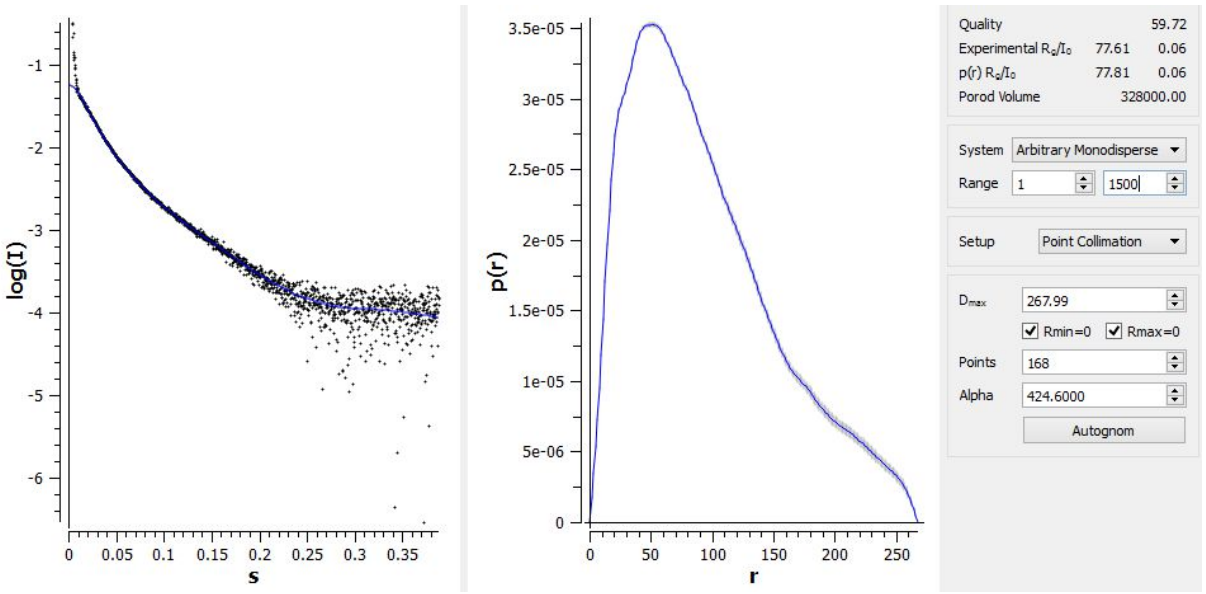


ScÅtter graph of Guinier peak analysis to pick linear region for Rg determination

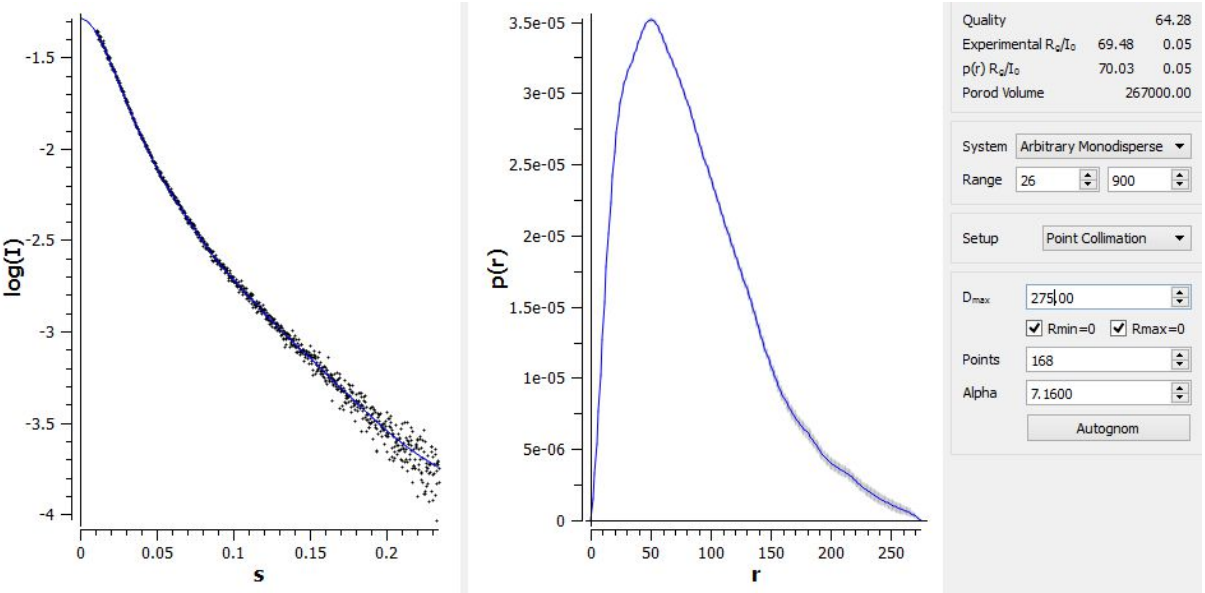


ScÅtter graph of Guinier peak analysis to pick linear region for Rc determination

The PRIMUS programme from the ATSAS suite was used to perform distance distribution analysis using the averaged data file produced in the ScÅtter 3.0 software. The paired-distance distribution function, P(r), plotted against interatomic distances observed within the sample was generated. Data was truncated accordingly in order to remove noisy data from the high Q region and the low Q region if necessary. The Dmax value was adjusted until a satisfactory curve was produced, aided by monitoring the quality value produced.

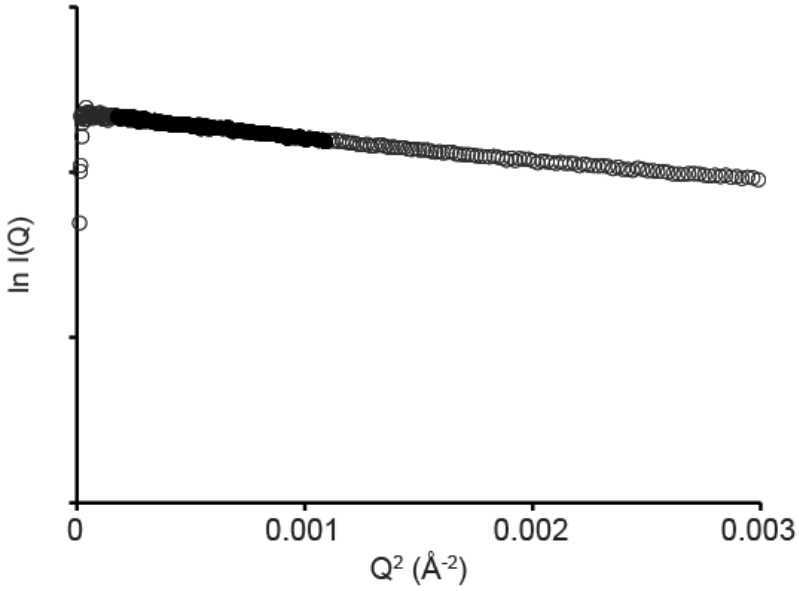


PRIMUS distance distribution before data truncation and Dmax adjustment

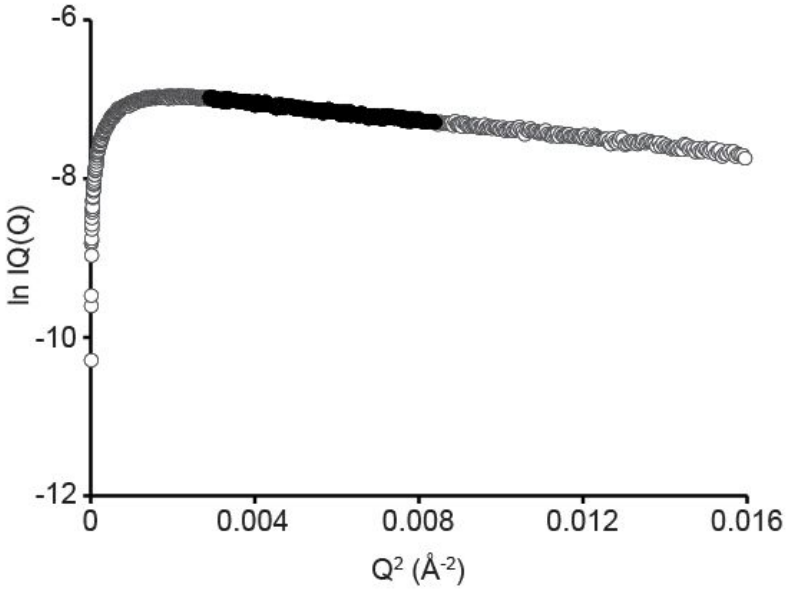


PRIMUS distance distribution analysis after data truncation and Dmax adjustment

In order to produce the graphs included in this thesis, the data file generated from ScÅtter 3.0 and the distance distribution output file from PRIMUS were opened in Microsoft Excel. The file generated from ScÅtter 3.0 supplied the Q and I(Q) data from the averaged scattering curve and was used to generate the Rg and Rc plots. For the Rg plots, the Q^2 and natural log of the I(Q) were calculated, respectively. The Q^2 was plotted against the ln I(Q) and the linear region used to generate the Rg value (determined in the ScÅtter 3.0 software as described above) was highlighted. For the Rc plots, Q^2 was plotted against the natural log of the I x Q and the linear region used to generate the Rc value (determined in the ScÅtter 3.0 software as described above) was highlighted.

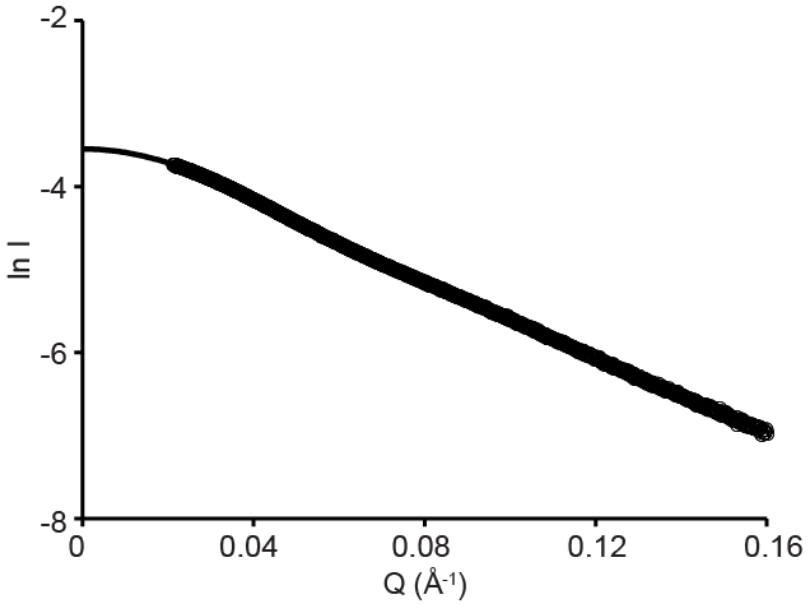


Example of Rg plot after manual data processing of the ScÅtter averaged data

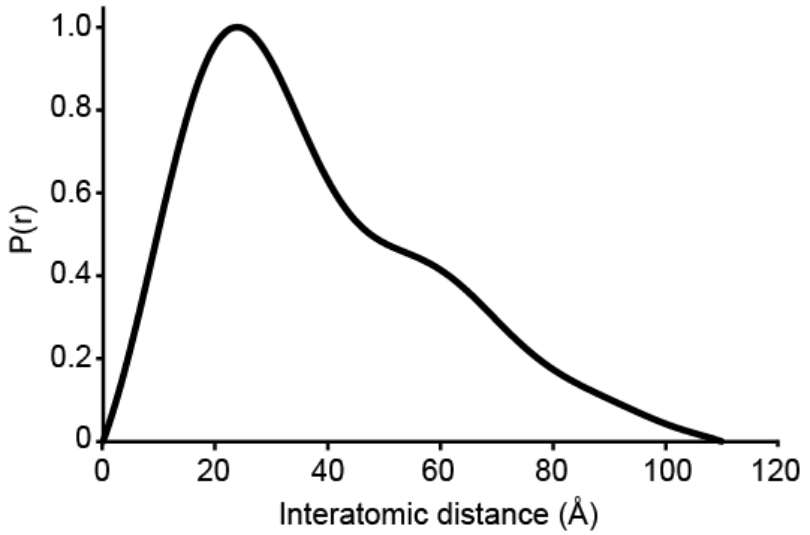


Example of Rc plot after manual data processing of the ScÅtter averaged data

The output file from the PRIMUS P(r) processing was used to generate the scattering curves and distance distribution plots. The file supplied the experimental and fitted data for the scattering intensity, I, against Q. The natural log of both the truncated experimental data and the fitted scattering intensity was calculated and plotted on the same graph against Q. The PRIMUS output file also supplied the P(r) and distance distribution data generated in order to plot the interatomic distances against the P(r). In order for the distance distribution plots to be comparable between samples, the P(r) values were normalised so that the highest point of the curve had a P(r) value of 1 and the other points were adjusted accordingly.

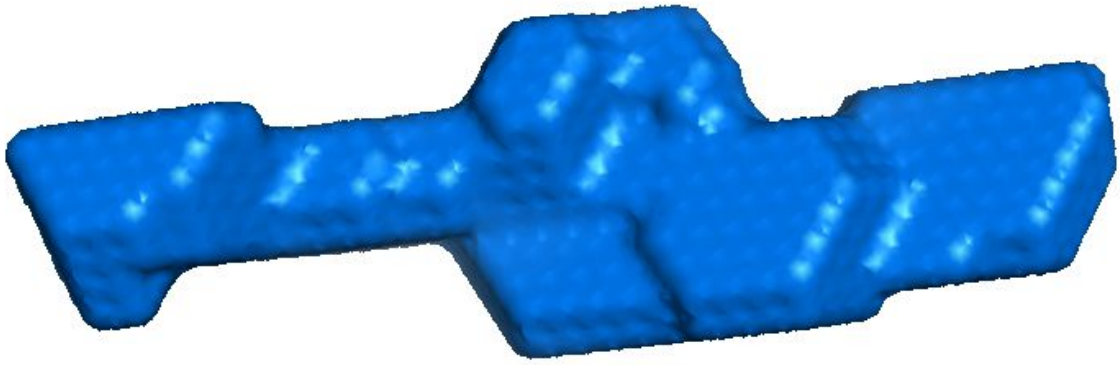


Example of scattering curve after manual data processing of the PRIMUS output

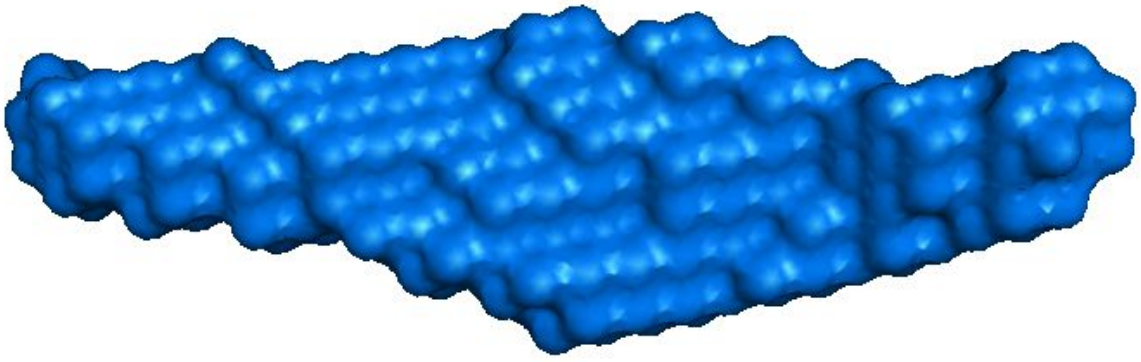


Example of P(r) curve after manual data processing of the PRIMUS output

After generation of the distance distribution plots in PRIMUS, the output file was used for *ab initio* modelling for those samples where possible using programs within the ATSAS suite. DAMMIF was used to perform initial shape determination using a single phase dummy atom model and 10-20 runs were performed for each sample. DAMSEL was then used to compare all the generated models to find the most probable model and to identify any outliers. DAMAVER then used the DAMSEL output to average the suitable models followed by DAMFILT to filter the averaged model. The final DAMFILT models were presented with surface representation.

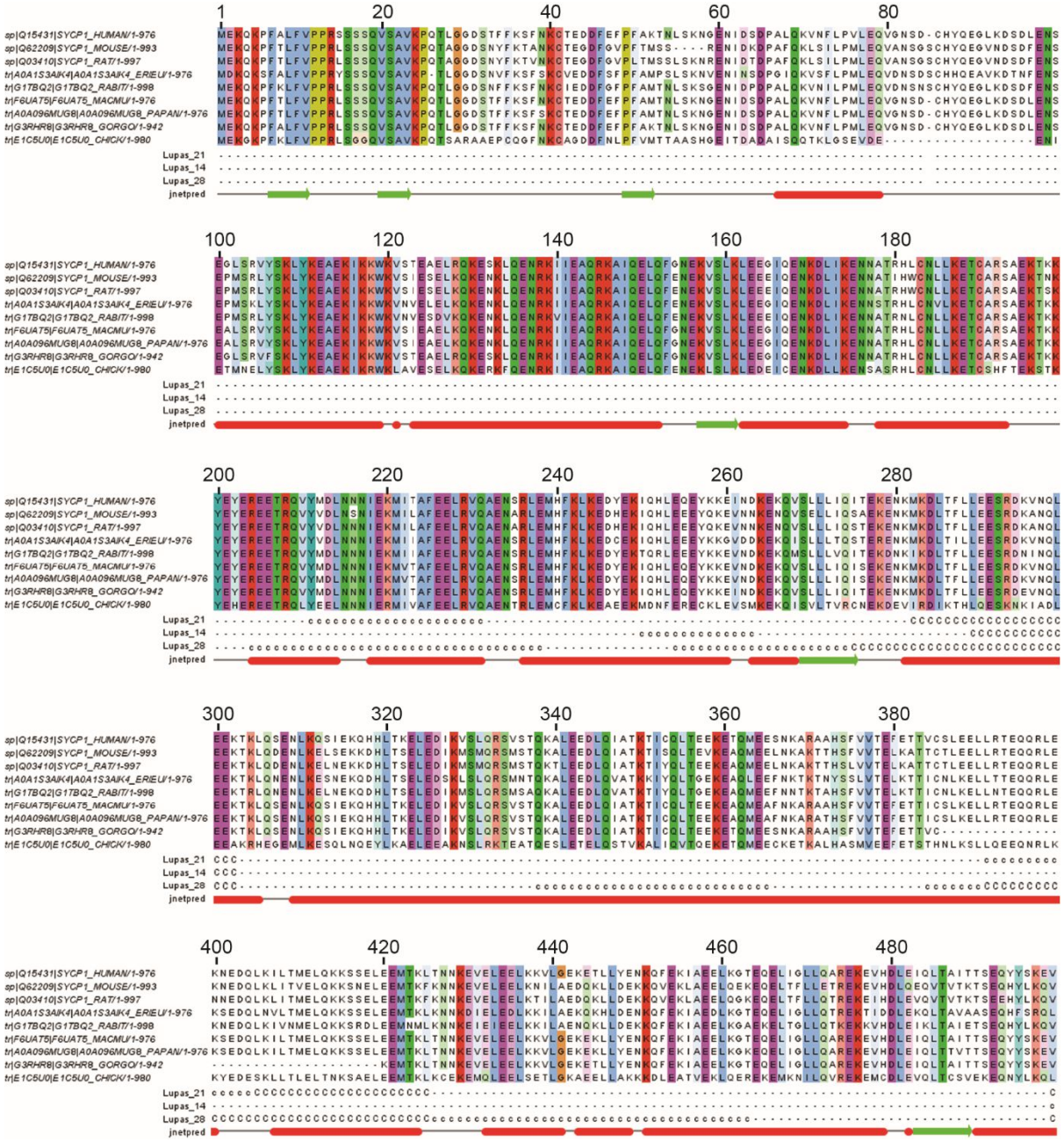


Surface representation of an initial single DAMMIF modelling run



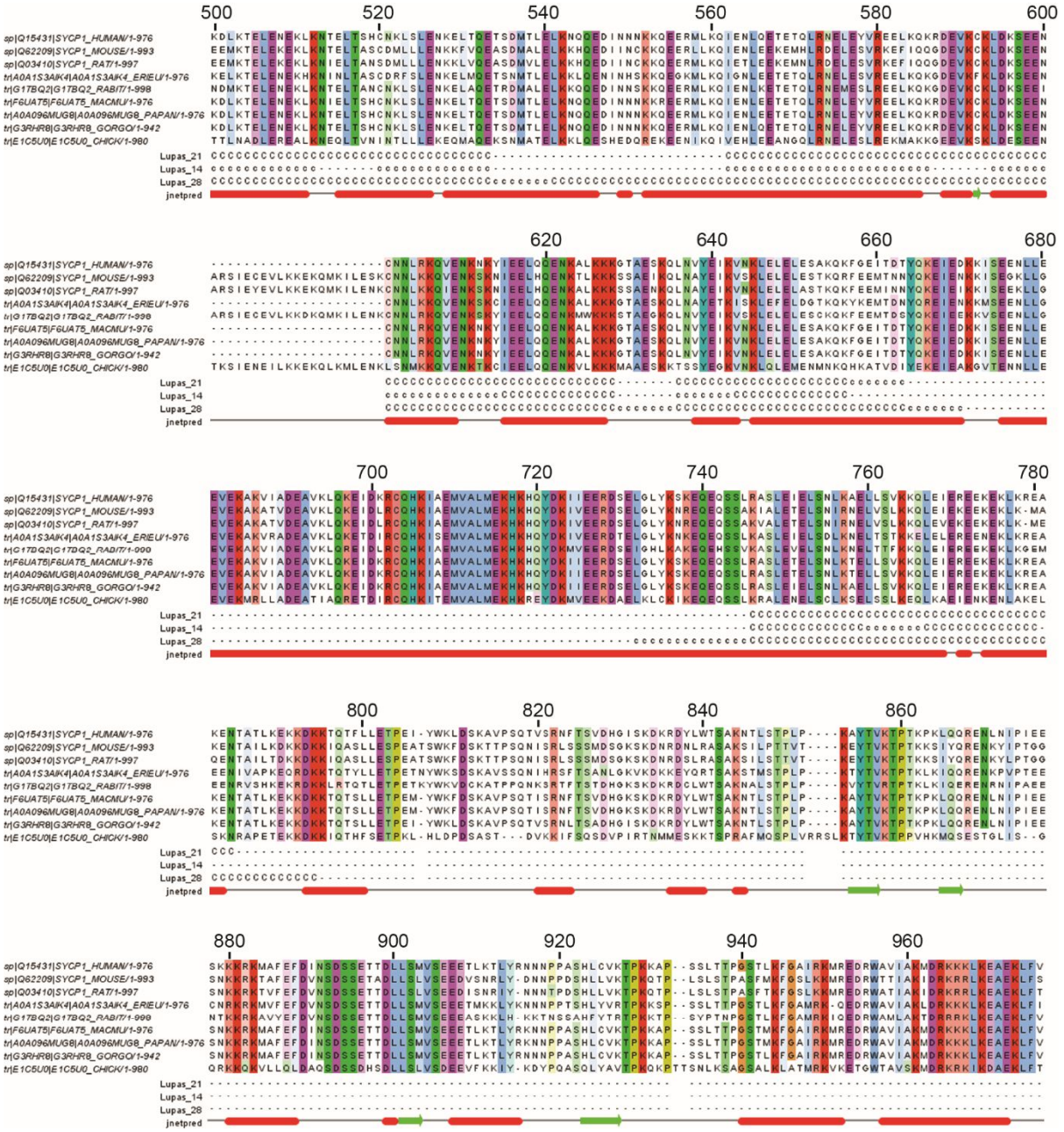
Surface representation of the final DAMFILT model from the same sample as above

Appendix III



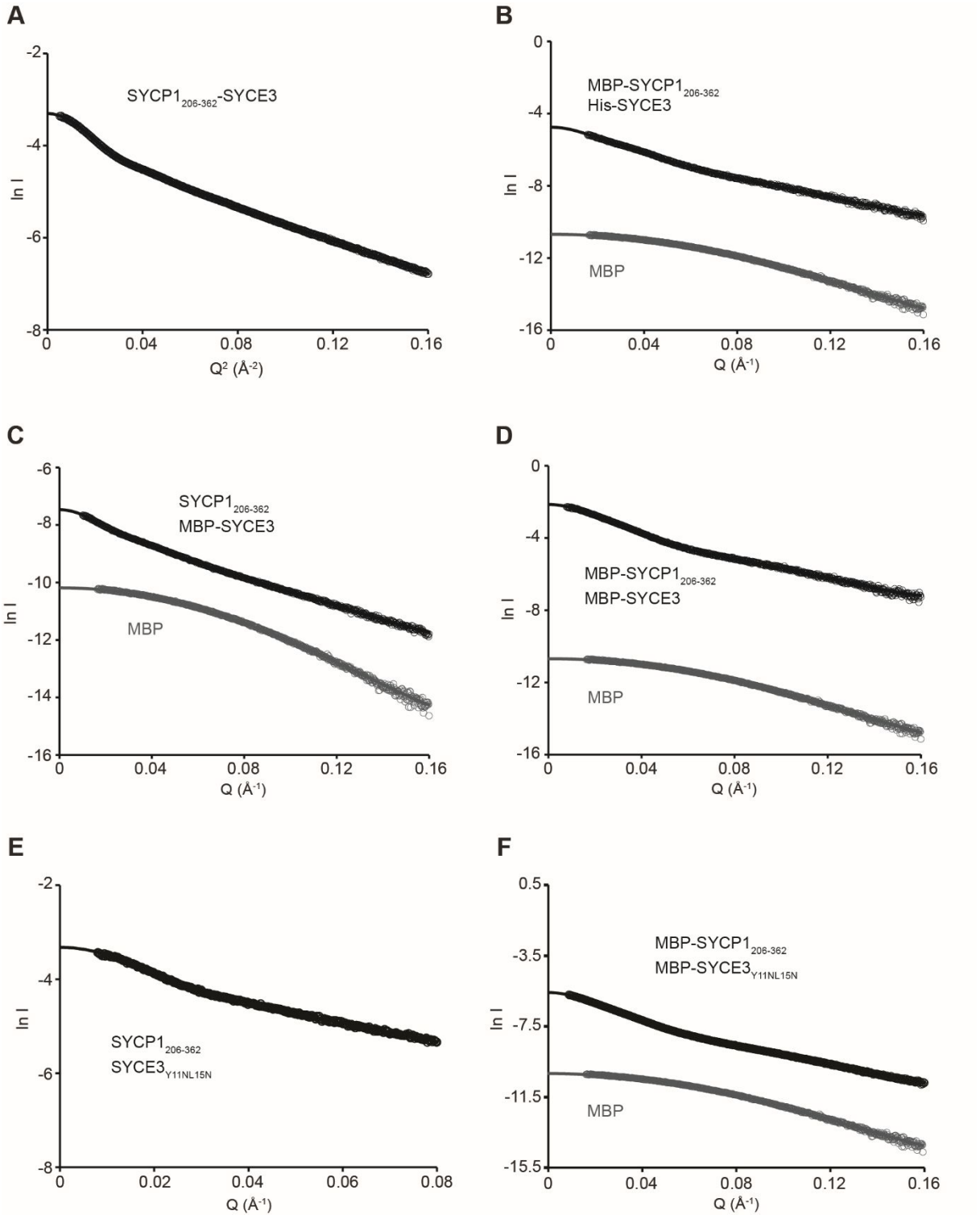
SYCP1 sequence analysis, corresponding to Chapter three, Figure 3.3, page 67. Multiple sequence alignments were generated by MUSCLE and secondary structure prediction performed by JNET. Analysis visualised in Jalview, amino acids coloured by conservation, with labelled amino acid positions corresponding to the human SYCP1 sequence. Alpha helix (red rods), beta sheets (green arrow), unstructured (grey line) presented below with the COILS Lupas coiled-coil prediction.

Appendix III continued



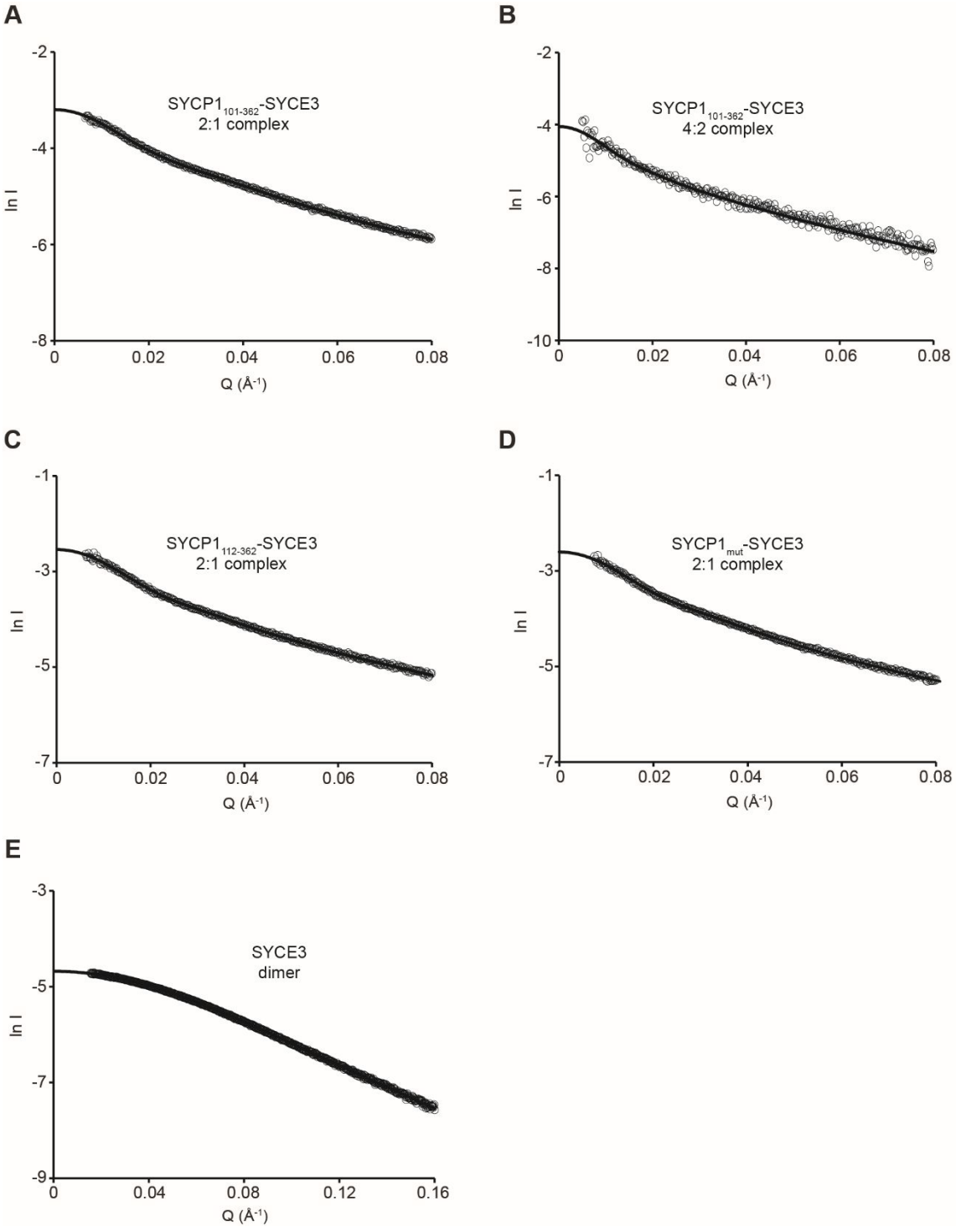
SYCP1 sequence analysis, corresponding to Chapter three, Figure 3.3, page 67. Multiple sequence alignments were generated by MUSCLE and secondary structure prediction performed by JNET. Analysis visualised in Jalview, amino acids coloured by conservation, with labelled amino acid positions corresponding to the human SYCP1 sequence. Alpha helix (red rods), beta sheets (green arrow), unstructured (grey line) presented below with the COILS Lupas coiled-coil prediction.

Appendix IV



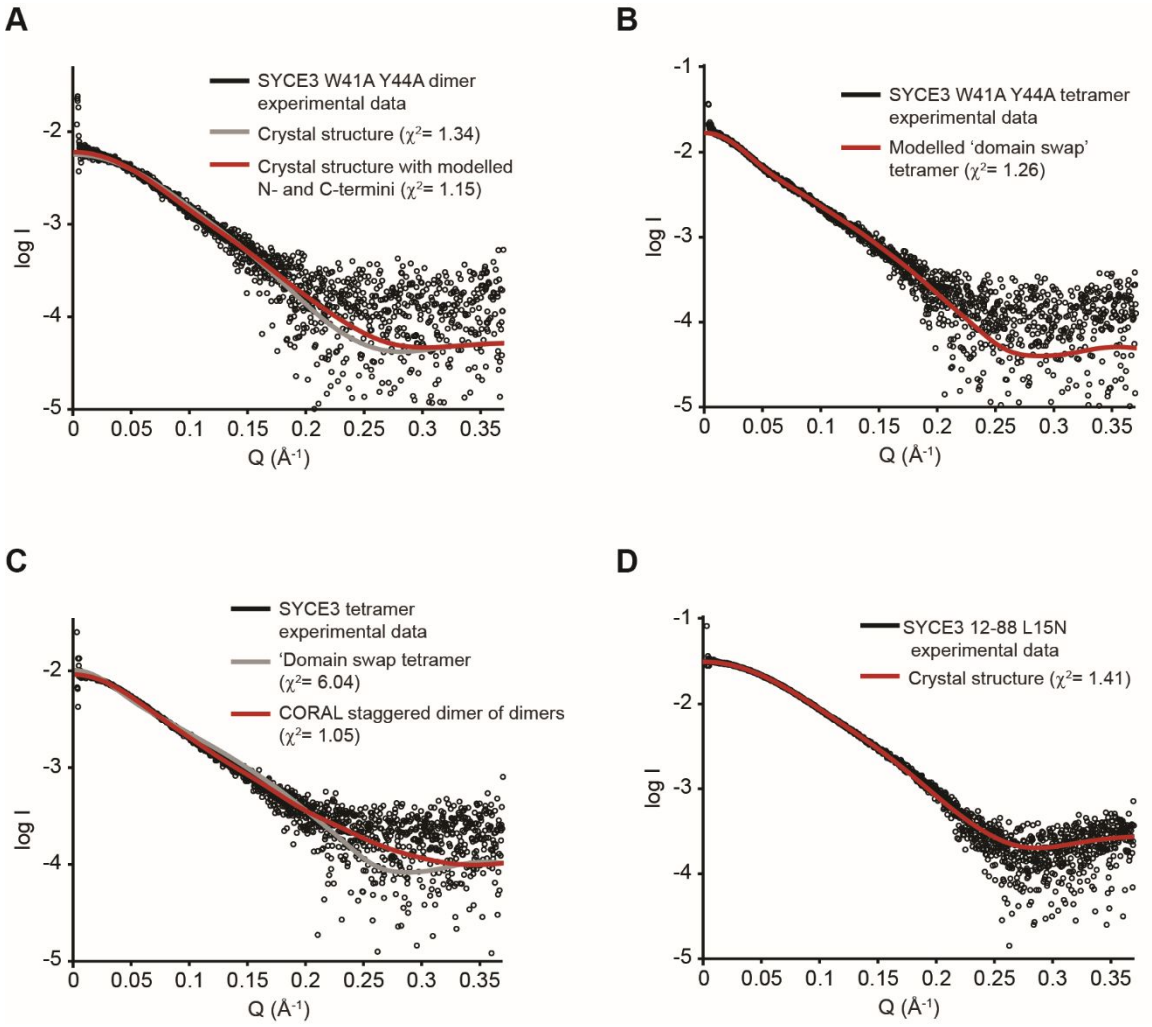
SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter three. (A) SYCP1$_{206-362}$-SYCE3. Corresponding to Figure 3.10, page 78. (B) MBP-SYCP1$_{206-362}$-His-SYCE3. Corresponding to Figure 3.11, page 79. Free MBP plotted in grey. (C) SYCP1$_{206-362}$-MBP-SYCE3. Corresponding to Figure 3.12, page 80. Free MBP plotted in grey. (D) MBP-SYCP1$_{206-362}$-MBP-SYCE3. Corresponding to Figure 3.13, page 81. Free MBP plotted in grey. (E) SYCP1$_{206-362}$-SYCE3$_{Y11NL15N}$. Corresponding to Figure 3.30, page 108. (F) MBP-SYCP1$_{206-362}$-MBP-SYCE3. Corresponding to Figure 3.29, page 107. Free MBP plotted in grey.

Appendix V



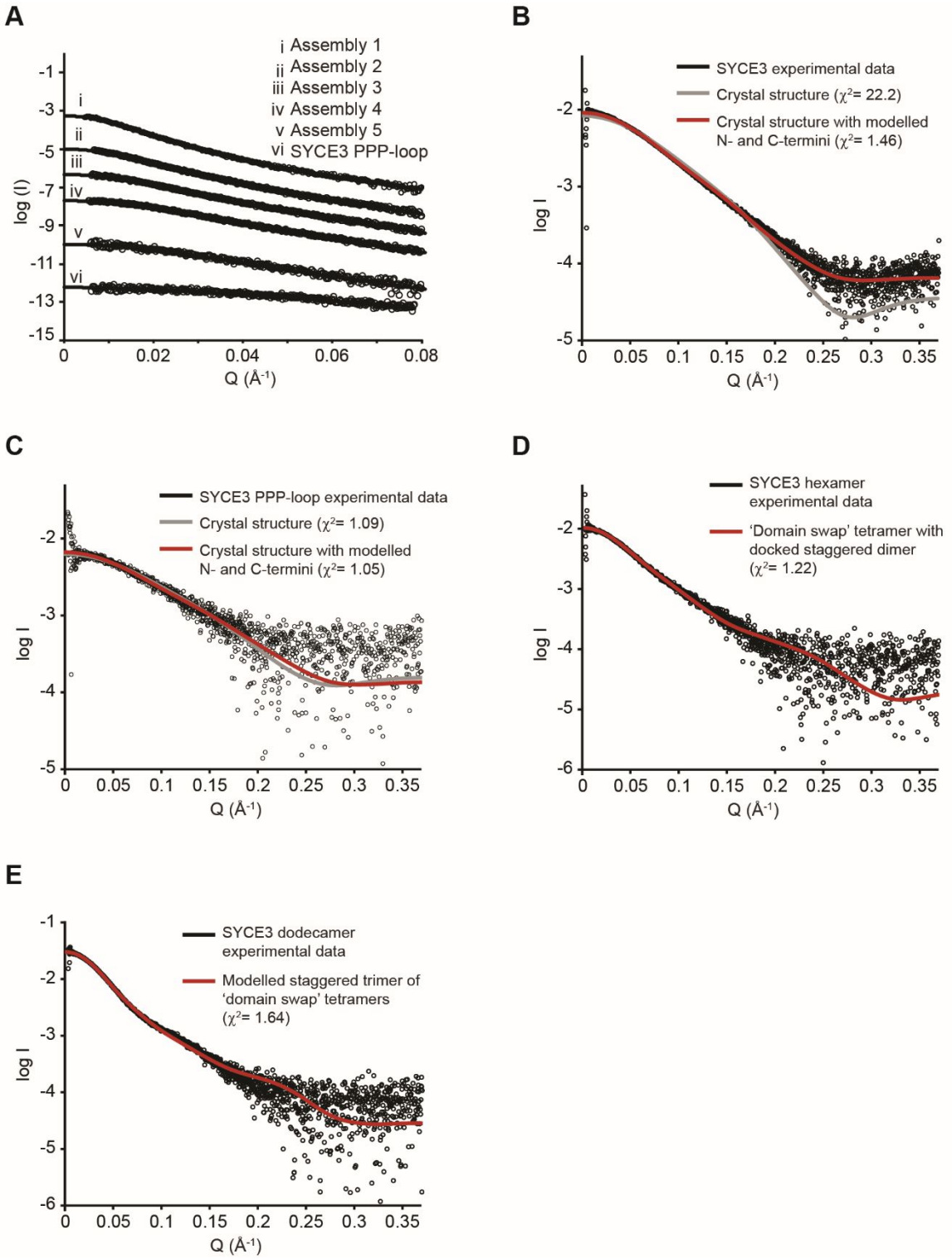
SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter three. (A) SYCP1$_{101-362}$-SYCE3 2:1 complex. Corresponding to Figure 3.16, page 86. (B) SYCP1$_{101-362}$-SYCE3 4:2 complex. Corresponding to Figure 3.17, page 87. (C) SYCP1$_{112-362}$-SYCE3. Corresponding to Figure 3.18, page 88. (D) SYCP1$_{mut}$-SYCE3. Corresponding to Figure 3.19, page 90. (E) SYCE3 dimer. Corresponding to Figure 3.21, page 93.

Appendix VI



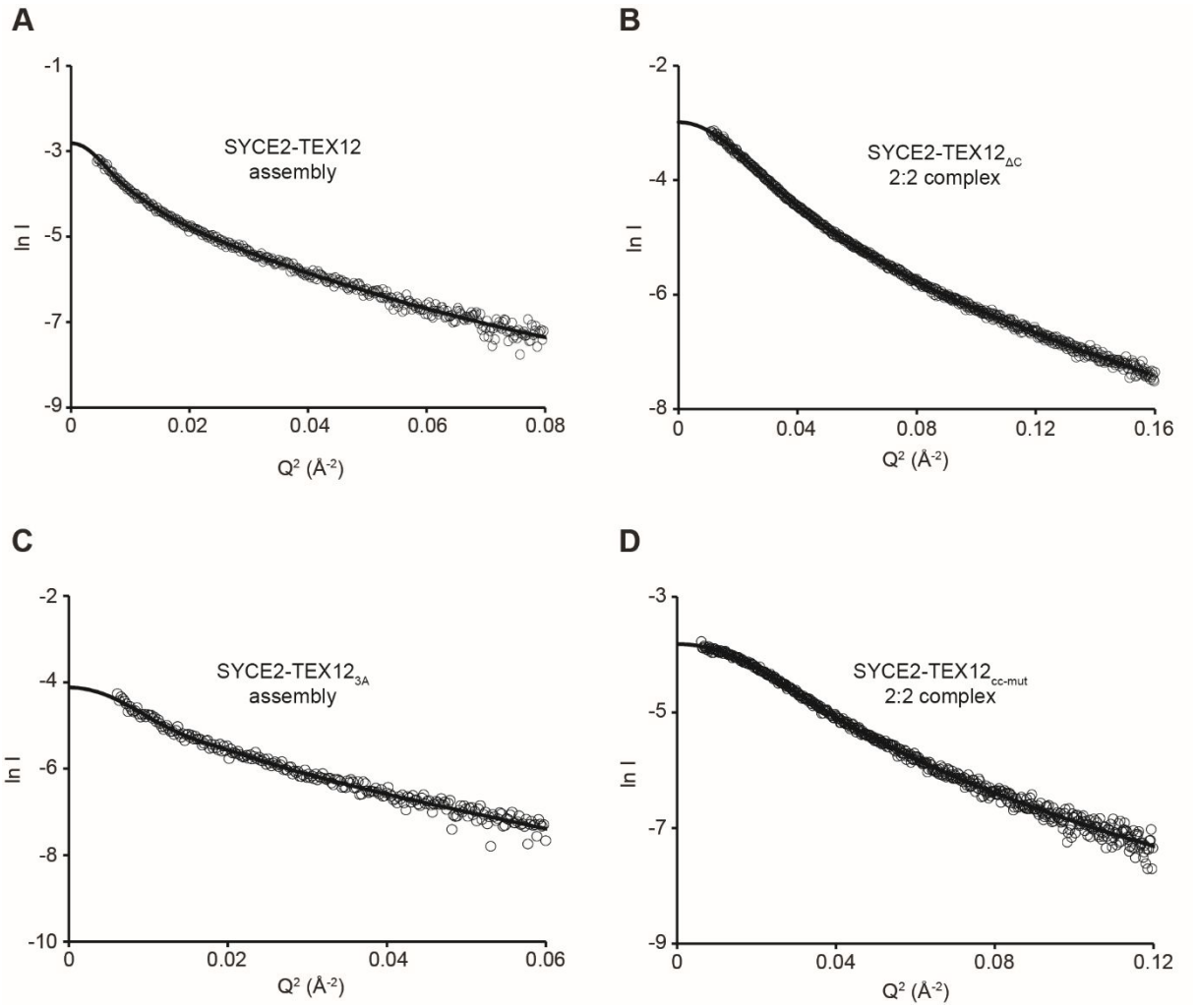
SEC-SAXS scattering curves plotted with the modelled structure fits, corresponding to Chapter three. (A) SYCE3 W41A Y44A dimer with the original and refined docked structures. Corresponding to Figure 3.26, page 101. (B) SYCE3 W41A Y44A tetramer with the modelled tetramer docking. Corresponding to Figure 3.26, page 101. (C) SYCE3 tetramer fitted with the docked domain swap tetramer and the staggered dimer of dimers. Corresponding to Figure 3.26, page 101. (D) SYCE3 12-88 L15N dimer with the fit from the docked crystal structure. Corresponding to Figure 3.24, page 98.

Appendix VII



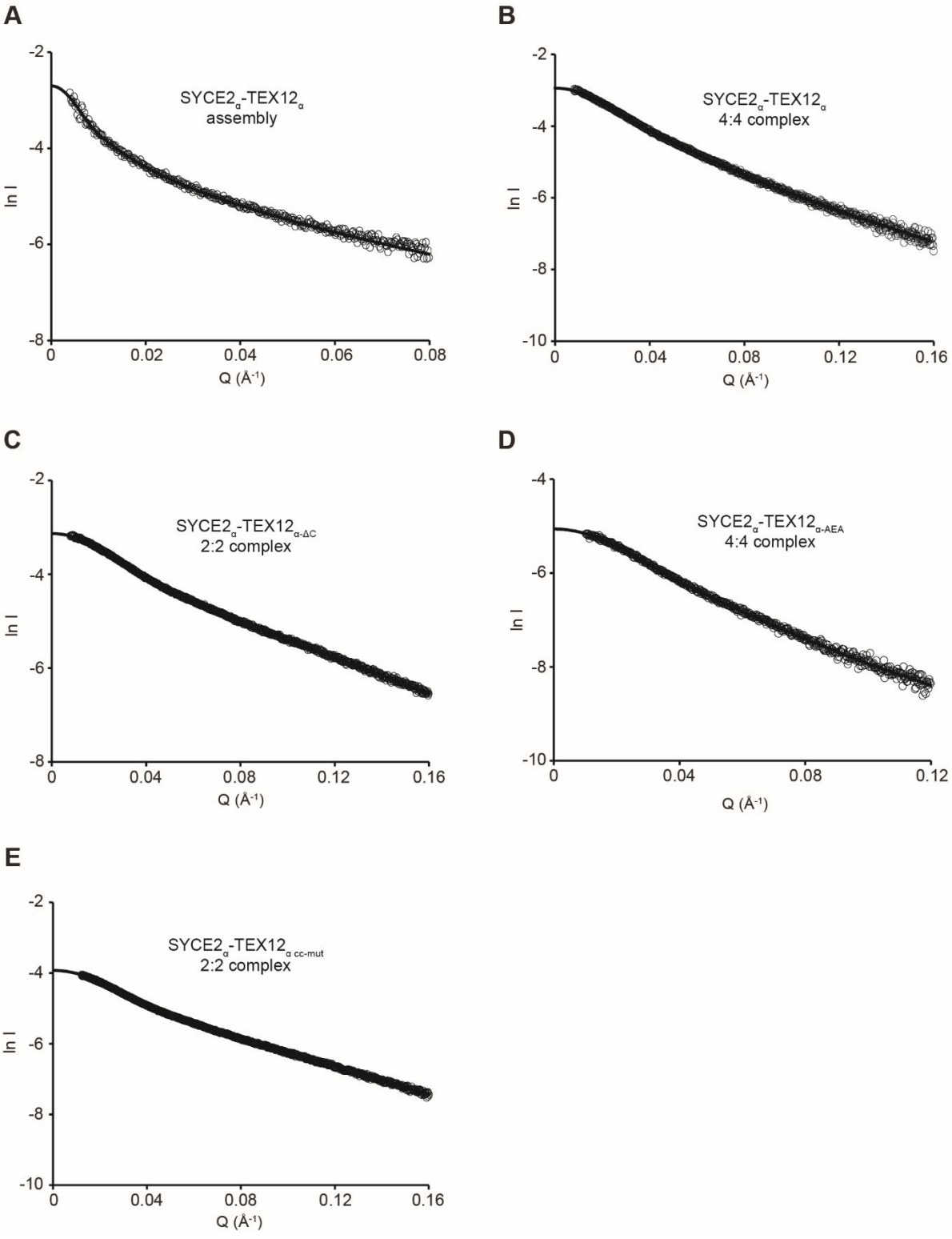
SEC-SAXS scattering curves plotted with the modelled structure fits, corresponding to Chapter three. (A) Overlaid SYCE3 scattering curves for the P53Q assembly peaks and the PPP-loop dimer. Corresponding to Figure 3.25, page 100. (B) SYCE3 dimer with the original and refined docked structures. Corresponding to Figure 3.21, page 93. (C) SYCE3 PPP-loop dimer with the original and refined docked structures. Corresponding to Figure 3.25, page 100. (D) SYCE3 hexamer with the fit from the docked domain swap tetramer with a staggered dimer model. Corresponding to Figure 3.26, page 101. (E) SYCE3 dodecamer with the fit from the docked staggered trimer of domain swap tetramers model. Corresponding to Figure 3.26, page 101.

Appendix VIII



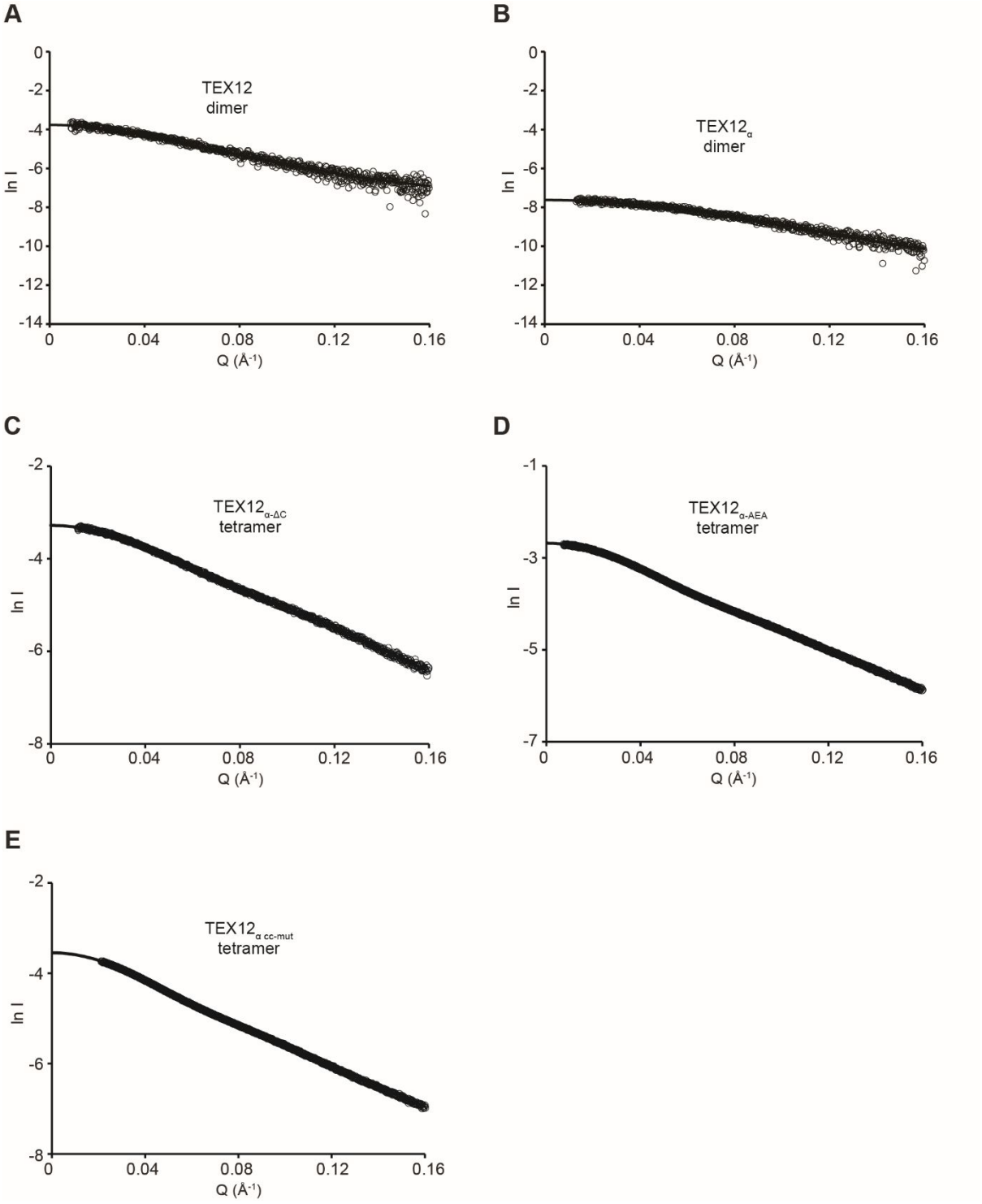
SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter four. (A) SYCE2-TEX12 assembly peak. Corresponding to Figure 4.6, page 132. (B) SYCE2-TEX12$_{\Delta C}$ 2:2 complex. Corresponding to Figure 4.8, page 136. (C) SYCE2-TEX12$_{3A.}$ Corresponding to Figure 4.10, page 139. (D) SYCE2-TEX12$_{cc\text{-}mut.}$ Corresponding to Figure 4.11, page 140.

Appendix IX



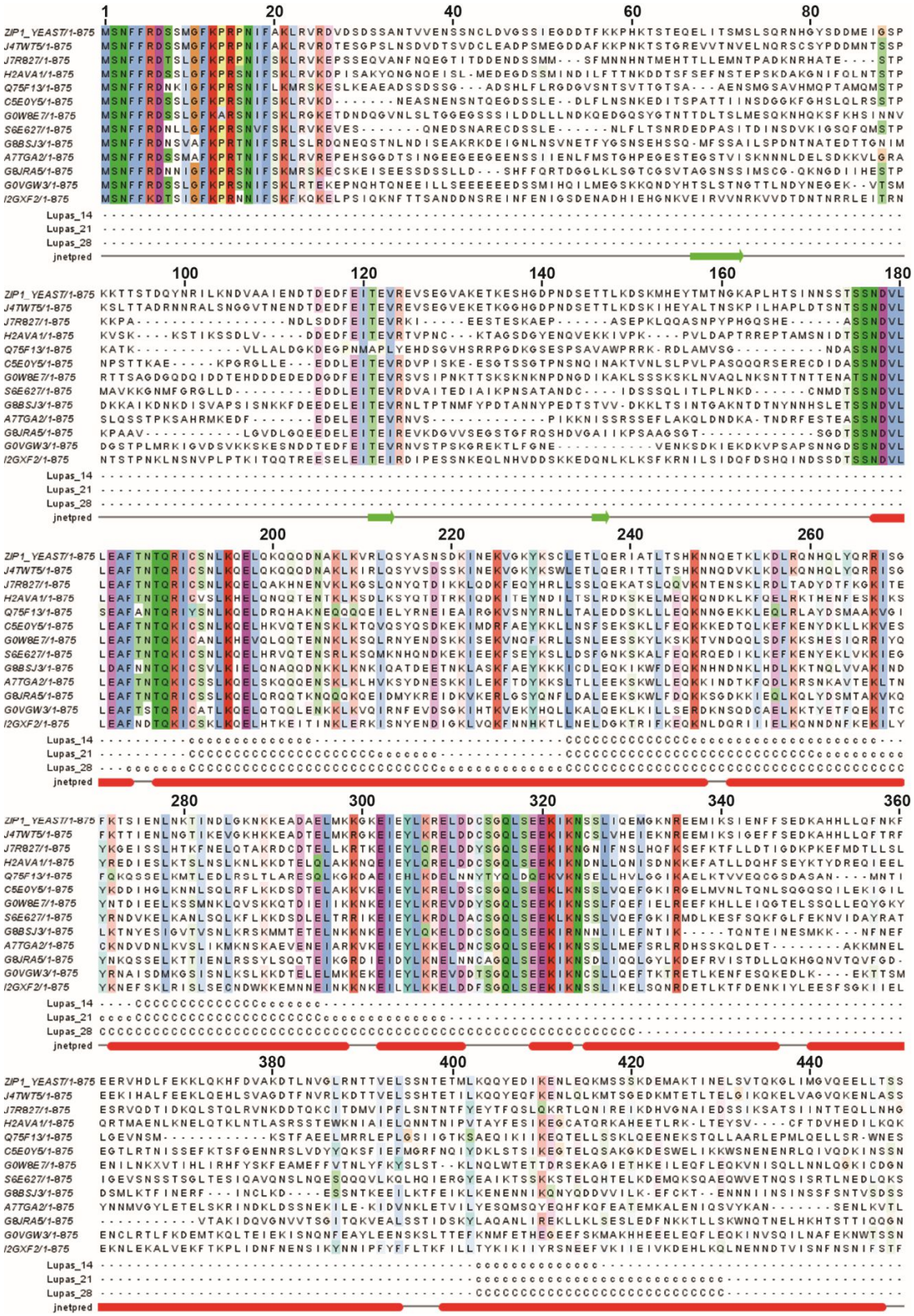
SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter four. (A) SYCE2$_{\alpha}$-TEX12 $_{\alpha}$ assembly peak. Corresponding to Figure 4.13, page 145. (B) SYCE2$_{\alpha}$-TEX12$_{\alpha}$ 4:4 complex. Corresponding to Figure 4.14, page 146. (C) SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-}\Delta C.}$ Corresponding to Figure 4.17, page 150. (D) SYCE2$_{\alpha}$-TEX12$_{\alpha\text{-AEA.}}$ Corresponding to Figure 4.22, page 157. (E) SYCE2$_{\alpha}$-TEX12$_{\alpha\text{ cc-mut.}}$ Corresponding to Figure 4.20, page 154.

Appendix X



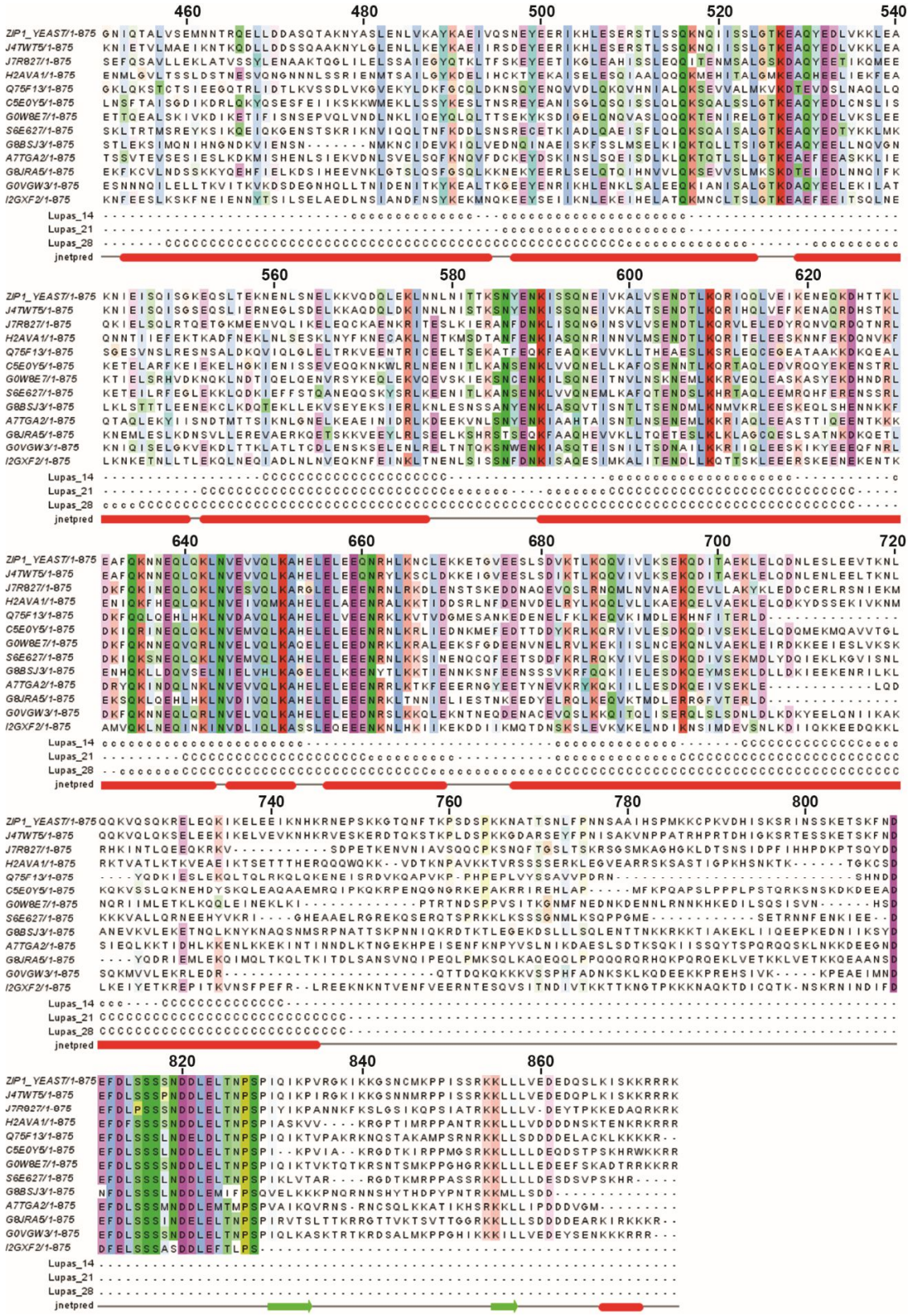
SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter four. (A) TEX12. Corresponding to Figure 4.24, page 160. (B) TEX12$_{\alpha}$. Corresponding to Figure 4.26, page 163. (C) TEX12$_{\alpha\text{-}\Delta C}$. Corresponding to Figure 4.28, page 165. (D) TEX12$_{\alpha\text{-}AEA}$. Corresponding to Figure 4.31, page 169. (E) TEX12$_{\alpha\ cc\text{-}mut}$. Corresponding to Figure 4.29, page 167.

Appendix XI



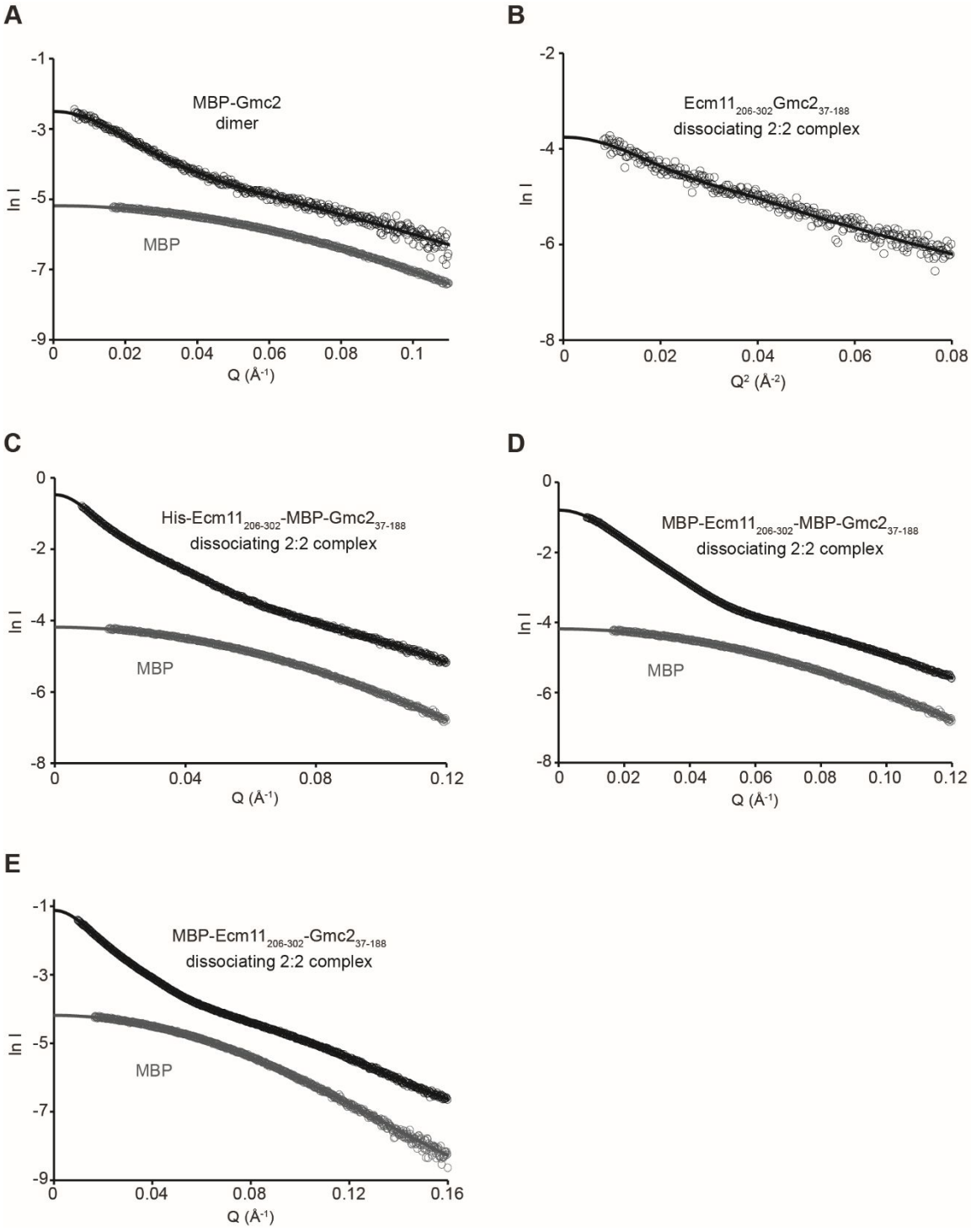
Zip1 sequence analysis, corresponding to Chapter five Figure 5.3, page 195. Multiple sequence alignments were generated by MUSCLE and secondary structure prediction performed by JNET. Analysis visualised in Jalview, amino acids coloured by conservation, with labelled amino acid positions corresponding to the human Zip1 sequence. Alpha helix (red rods), beta sheets (green arrow), unstructured (grey line) presented below with the COILS Lupas coiled-coil prediction.

Appendix XI continued



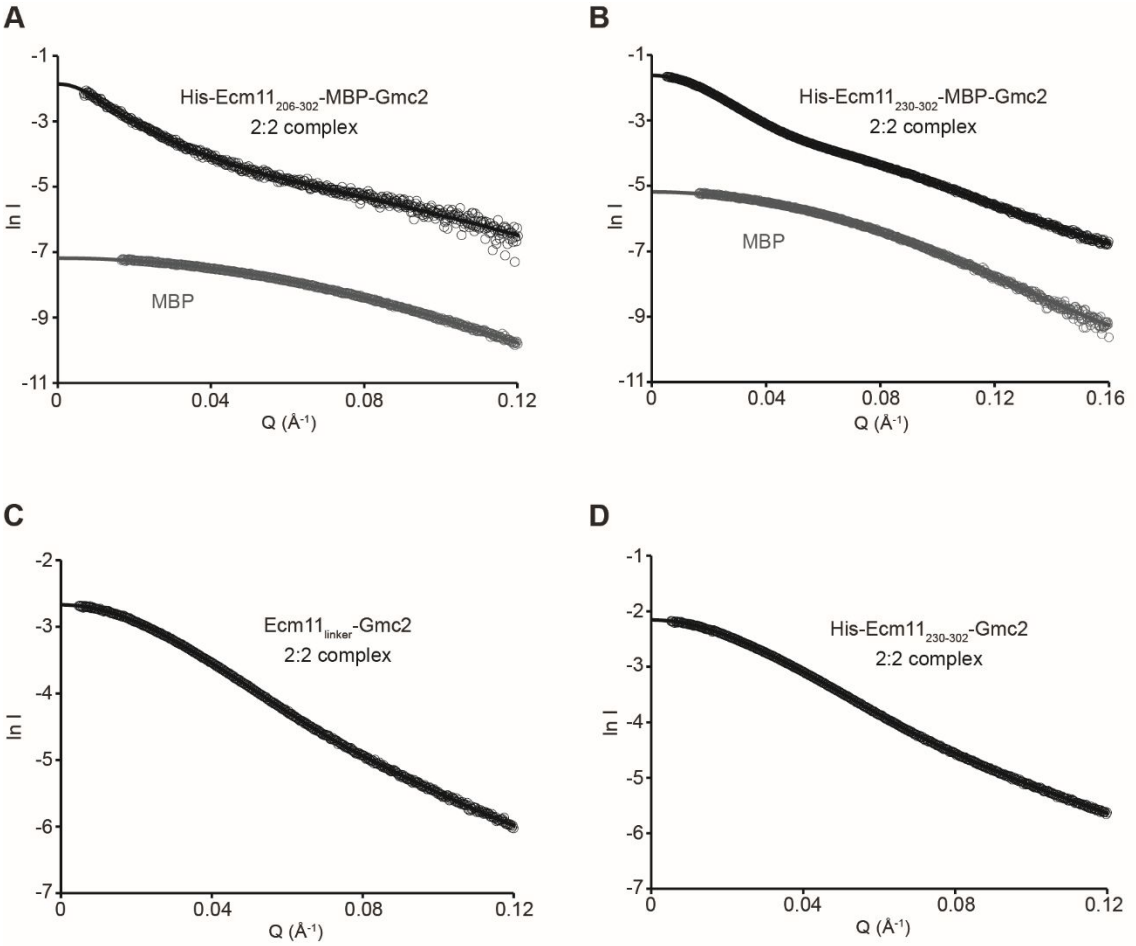
Zip1 sequence analysis, corresponding to Chapter five Figure 5.3, page 195. Multiple sequence alignments were generated by MUSCLE and secondary structure prediction performed by JNET. Analysis visualised in Jalview, amino acids coloured by conservation, with labelled amino acid positions corresponding to the human Zip1 sequence. Alpha helix (red rods), beta sheets (green arrow), unstructured (grey line) presented below with the COILS Lupas coiled-coil prediction.

Appendix XII



SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter five. (A) MBP-Gmc2. Corresponding to Figure 5.18, page 224. Free MBP plotted in grey. (B) Ecm11$_{206-302}$-Gmc2$_{37-188}$. Corresponding to Figure 5.22, page 230. (C) His-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$. Corresponding to Figure 5.23, page 231. Free MBP plotted in grey. (D) MBP-Ecm11$_{206-302}$-MBP-Gmc2$_{37-188}$. Corresponding to Figure 5.24, page 233. Free MBP plotted in grey. (E) MBP-Ecm11$_{206-302}$-Gmc2$_{37-188}$. Corresponding to Figure 5.25, page 234. Free MBP plotted in grey.

Appendix XIII



SEC-SAXS scattering curves plotted with the P(r) fits, corresponding to Chapter five. (A) His-Ecm11$_{206-302}$-MBP-Gmc2. Corresponding to Figure 5.26, page 236. Free MBP plotted in grey. (B) His-Ecm11$_{230-302}$-MBP-Gmc2. Corresponding to Figure 5.30, page 241. Free MBP plotted in grey. (C) Ecm11$_{linker}$-Gmc2. Corresponding to Figure 5.32, page 243. (D) His-Ecm11$_{230-302}$-Gmc2. Corresponding to Figure 5.34, page 246.

References

ADAMS, P. D., AFONINE, P. V., BUNKOCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr,* 66**,** 213-21.

AGARWAL, S. & ROEDER, G. S. 2000. Zip3 provides a link between recombination enzymes and synaptonemal complex proteins. *Cell,* 102**,** 245-55.

ALBERTS, J., LEWIS , RAFF, ROBERTS., AND WALTER 2002. An Overview of the Cell Cycle. Available. *Molecular Biology of the Cell*.

ASTBURY, W. T., STREET, A. 1932. The X-ray studies of the structure of hair, wool and related fibres. *Philosophical transactions of the Royal Society of London,* 230**,** 75-101.

BAIER, A., ALSHEIMER, M. & BENAVENTE, R. 2007. Synaptonemal complex protein SYCP3: Conserved polymerization properties among vertebrates. *Biochim Biophys Acta,* 1774**,** 595-602.

BAUDAT, F., IMAI, Y. & DE MASSY, B. 2013. Meiotic recombination in mammals: localization and regulation. *Nat Rev Genet,* 14**,** 794-806.

BENSON, F. E., STASIAK, A. & WEST, S. C. 1994. Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA. *Embo j,* 13**,** 5764-71.

BIBBY, J., KEEGAN, R. M., MAYANS, O., WINN, M. D. & RIGDEN, D. J. 2012. AMPLE: a cluster-and-truncate approach to solve the crystal structures of small proteins using rapidly computed ab initio models. *Acta Crystallogr D Biol Crystallogr,* 68**,** 1622-31.

BISHOP, D. K. 1994. RecA homologs Dmc1 and Rad51 interact to form multiple nuclear complexes prior to meiotic chromosome synapsis. *Cell,* 79**,** 1081-92.

BISHOP, D. K., PARK, D., XU, L. & KLECKNER, N. 1992. DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell,* 69**,** 439-56.

BLACKBURN, E. H. 2000. Telomere states and cell fates. *Nature,* 408**,** 53-6.

BLACKBURN, E. H., EPEL, E. S. & LIN, J. 2015. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science,* 350**,** 1193-8.

BLACKWELL, L. J., BJORNSON, K. P., ALLEN, D. J. & MODRICH, P. 2001. Distinct MutS DNA-binding modes that are differentially modulated by ATP binding and hydrolysis. *J Biol Chem,* 276**,** 34339-47.

BOLCUN-FILAS, E., COSTA, Y., SPEED, R., TAGGART, M., BENAVENTE, R., DE ROOIJ, D. G. & COOKE, H. J. 2007. SYCE2 is required for synaptonemal complex assembly, double strand break repair, and homologous recombination. *J Cell Biol,* 176**,** 741-7.

BOLCUN-FILAS, E., HALL, E., SPEED, R., TAGGART, M., GREY, C., DE MASSY, B., BENAVENTE, R. & COOKE, H. J. 2009. Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair. *PLoS Genet,* 5**,** e1000393.

BROOKER, A. S. & BERKOWITZ, K. M. 2014. The roles of cohesins in mitosis, meiosis, and human health and disease. *Methods Mol Biol,* 1170**,** 229-66.

CHEN, X., SUHANDYNATA, R. T., SANDHU, R., ROCKMILL, B., MOHIBULLAH, N., NIU, H., LIANG, J., LO, H. C., MILLER, D. E., ZHOU, H., BORNER, G. V. & HOLLINGSWORTH, N. M. 2015. Phosphorylation of the Synaptonemal Complex Protein Zip1 Regulates the Crossover/Noncrossover Decision during Yeast Meiosis. *PLoS Biol,* 13**,** e1002329.

CHENG, C. H., LO, Y. H., LIANG, S. S., TI, S. C., LIN, F. M., YEH, C. H., HUANG, H. Y. & WANG, T. F. 2006. SUMO modifications control assembly of synaptonemal complex and polycomplex in meiosis of Saccharomyces cerevisiae. *Genes Dev,* 20**,** 2067-81.

CHIANG, T., DUNCAN, F. E., SCHINDLER, K., SCHULTZ, R. M. & LAMPSON, M. A. 2010. Evidence that weakened centromere cohesion is a leading cause of age-related aneuploidy in oocytes. *Curr Biol,* 20**,** 1522-8.

CLARKE, L. & CARBON, J. 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature,* 287**,** 504-9.

CLOUD, V., CHAN, Y. L., GRUBB, J., BUDKE, B. & BISHOP, D. K. 2012. Rad51 is an accessory factor for Dmc1-mediated joint molecule formation during meiosis. *Science,* 337**,** 1222-5.

COSTA, Y., SPEED, R., OLLINGER, R., ALSHEIMER, M., SEMPLE, C. A., GAUTIER, P., MARATOU, K., NOVAK, I., HOOG, C., BENAVENTE, R. & COOKE, H. J. 2005. Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. *J Cell Sci,* 118**,** 2755-62.

CRICK, F. H. 1952. Is alpha-keratin a coiled coil? *Nature,* 170**,** 882-3.

CRICKARD, J. B., KANIECKI, K., KWON, Y., SUNG, P. & GREENE, E. C. 2018. Spontaneous self-segregation of Rad51 and Dmc1 DNA recombinases within mixed recombinase filaments. *J Biol Chem,* 293**,** 4191-4200.

DAVIES, O. R., MAMAN, J. D. & PELLEGRINI, L. 2012. Structural analysis of the human SYCE2–TEX12 complex provides molecular insights into synaptonemal complex assembly. *Open Biology,* 2**,** 120099.

DAVYDENKO, O., SCHULTZ, R. M. & LAMPSON, M. A. 2013. Increased CDK1 activity determines the timing of kinetochore-microtubule attachments in meiosis I. *J Cell Biol,* 202**,** 221-9.

DE LOS SANTOS, T. & HOLLINGSWORTH, N. M. 1999. Red1p, a MEK1-dependent phosphoprotein that physically interacts with Hop1p during meiosis in yeast. *J Biol Chem,* 274**,** 1783-90.

DE VRIES, F. A., DE BOER, E., VAN DEN BOSCH, M., BAARENDS, W. M., OOMS, M., YUAN, L., LIU, J. G., VAN ZEELAND, A. A., HEYTING, C. & PASTINK, A. 2005. Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev,* 19**,** 1376-89.

DE VRIES, L., BEHAR, D. M., SMIRIN-YOSEF, P., LAGOVSKY, I., TZUR, S. & BASEL-VANAGAITE, L. 2014. Exome sequencing reveals SYCE1 mutation associated with autosomal recessive primary ovarian insufficiency. *J Clin Endocrinol Metab,* 99**,** E2129-32.

DONG, H. & ROEDER, G. S. 2000. Organization of the yeast Zip1 protein within the central region of the synaptonemal complex. *J Cell Biol,* 148**,** 417-26.

DUNCE, J. M., DUNNE, O. M., RATCLIFF, M., MILLÁN, C., MADGWICK, S., USÓN, I. & DAVIES, O. R. 2018. Structural basis of meiotic chromosome synapsis through SYCP1 self-assembly. *Nature Structural & Molecular Biology,* 25**,** 557-569.

EIJPE, M., HEYTING, C., GROSS, B. & JESSBERGER, R. 2000. Association of mammalian SMC1 and SMC3 proteins with meiotic chromosomes and synaptonemal complexes. *J Cell Sci,* 113 (Pt 4)**,** 673-82.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

FENG, J., FU, S., CAO, X., WU, H., LU, J., ZENG, M., LIU, L., YANG, X. & SHEN, Y. 2017. Synaptonemal complex protein 2 (SYCP2) mediates the association of the centromere with the synaptonemal complex. *Protein Cell,* 8**,** 538-543.

FRANKE, D. & SVERGUN, D. I. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr,* 42**,** 342-346.

FRASER, R. D., MACRAE, T. P. & MILLER, A. 1964. THE COILED-COIL MODEL OF ALPHA-KERATIN STRUCTURE. *J Mol Biol,* 10**,** 147-56.

FRASER, R. D. B. & MACRAE, T. P. 1961. The molecular configuration of α-keratin. *Journal of Molecular Biology,* 3**,** 640-IN22.

FRAUNE, J., BROCHIER-ARMANET, C., ALSHEIMER, M. & BENAVENTE, R. 2013. Phylogenies of central element proteins reveal the dynamic evolutionary history of the mammalian synaptonemal complex: ancient and recent components. *Genetics,* 195**,** 781-93.

FRAUNE, J., BROCHIER-ARMANET, C., ALSHEIMER, M., VOLFF, J. N., SCHUCKER, K. & BENAVENTE, R. 2016. Evolutionary history of the mammalian synaptonemal complex. *Chromosoma,* 125**,** 355-60.

FRIEDMAN, D. B., HOLLINGSWORTH, N. M. & BYERS, B. 1994. Insertional mutations in the yeast HOP1 gene: evidence for multimeric assembly in meiosis. *Genetics,* 136**,** 449-464.

GAO, J. & COLAIACOVO, M. P. 2018. Zipping and Unzipping: Protein Modifications Regulating Synaptonemal Complex Dynamics. *Trends Genet,* 34**,** 232-245.

GAREAU, J. R. & LIMA, C. D. 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol,* 11**,** 861-71.

GEISINGER, A. & BENAVENTE, R. 2016. Mutations in Genes Coding for Synaptonemal Complex Proteins and Their Impact on Human Fertility. *Cytogenet Genome Res,* 150**,** 77-85.

GILBERT, S. F. 2000a. Meiosis. *Developmental Biology.* Sunderland (MA): Sinauer Associates.

GILBERT, S. F. 2000b. Oogenesis. *Developmental Biology.* Sinauer Associates.

GILBERT, S. F. 2000a. Spermatogenesis. *Developmental Biology.* Sinauer Associates.

GOLDSTEIN, P. 1987. Multiple synaptonemal complexes (polycomplexes): origin, structure and function. *Cell Biol Int Rep,* 11**,** 759-96.

GÓMEZ-H, L., FELIPE-MEDINA, N., SÁNCHEZ-MARTÍN, M., DAVIES, O. R., RAMOS, I., GARCÍA-TUÑÓN, I., DE ROOIJ, D. G., DERELI, I., TÓTH, A., BARBERO, J. L., BENAVENTE, R., LLANO, E. & PENDAS, A. M. 2016. C14ORF39/SIX6OS1 is a constituent of the synaptonemal complex and is essential for mouse fertility. *Nature Communications,* 7**,** 13298.

GOMEZ, H. L., FELIPE-MEDINA, N., SANCHEZ-MARTIN, M., DAVIES, O. R., RAMOS, I., GARCIA-TUNON, I., DE ROOIJ, D. G., DERELI, I., TOTH, A., BARBERO, J. L., BENAVENTE, R., LLANO, E. & PENDAS, A. M. 2016.

C14ORF39/SIX6OS1 is a constituent of the synaptonemal complex and is essential for mouse fertility. *Nat Commun,* 7**,** 13298.

GORBSKY, G. J., SAMMAK, P. J. & BORISY, G. G. 1987. Chromosomes move poleward in anaphase along stationary microtubules that coordinately disassemble from their kinetochore ends. *J Cell Biol,* 104**,** 9-18.

GRANDIN, N. & CHARBONNEAU, M. 2008. Protection against chromosome degradation at the telomeres. *Biochimie,* 90**,** 41-59.

GRAY, J. J., MOUGHON, S., WANG, C., SCHUELER-FURMAN, O., KUHLMAN, B., ROHL, C. A. & BAKER, D. 2003. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol,* 331**,** 281-99.

HAERING, C. H., LOWE, J., HOCHWAGEN, A., NASMYTH, K. 2002. Molecular architecture of SMC proteins and the yeast cohesin complex. *Molecular Cell,* 9**,** 773-788.

HAMER, G., GELL, K., KOUZNETSOVA, A., NOVAK, I., BENAVENTE, R. & HOOG, C. 2006. Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex. *J Cell Sci,* 119**,** 4025-32.

HAMER, G., WANG, H., BOLCUN-FILAS, E., COOKE, H. J., BENAVENTE, R. & HOOG, C. 2008. Progression of meiotic recombination requires structural maturation of the central element of the synaptonemal complex. *J Cell Sci,* 121**,** 2445-51.

HANDEL, M. A. & SCHIMENTI, J. C. 2010. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nat Rev Genet,* 11**,** 124-36.

HARVEY, S. H., KRIEN, M. J. & O'CONNELL, M. J. 2002. Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. *Genome Biol,* 3**,** Reviews3003.

HERNANDEZ-HERNANDEZ, A., MASICH, S., FUKUDA, T., KOUZNETSOVA, A., SANDIN, S., DANEHOLT, B. & HOOG, C. 2016. The central element of the synaptonemal complex in mice is organized as a bilayered junction structure. 129**,** 2239-49.

HEYTING, C., J. DETTMERS, R., J. J. DIETRICH, A., REDEKER, W. & C. G. VINK, A. 1988. *Two major components of synaptonemal complexes are specific for meiotic prophase nuclei*.

HOLLINGSWORTH, N. M. & BYERS, B. 1989. HOP1: a yeast meiotic pairing gene. *Genetics,* 121**,** 445-62.

HOLLINGSWORTH, N. M. & PONTE, L. 1997a. Genetic Interactions Between HOP1, RED1 and MEK1 Suggest That MEK1 Regulates Assembly of Axial Element Components During Meiosis in the Yeast Saccharomyces cerevisiae. *Genetics,* 147**,** 33-42.

HOLLINGSWORTH, N. M. & PONTE, L. 1997b. Genetic interactions between HOP1, RED1 and MEK1 suggest that MEK1 regulates assembly of axial element components during meiosis in the yeast Saccharomyces cerevisiae. *Genetics,* 147**,** 33-42.

HUMPHRYES, N., LEUNG, W. K., ARGUNHAN, B., TERENTYEV, Y., DVORACKOVA, M. & TSUBOUCHI, H. 2013. The Ecm11-Gmc2 complex promotes synaptonemal complex formation through assembly of transverse filaments in budding yeast. *PLoS Genet,* 9**,** e1003194.

JEONG, J. Y., YIM, H. S., RYU, J. Y., LEE, H. S., LEE, J. H., SEEN, D. S. & KANG, S. G. 2012. One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. *Appl Environ Microbiol,* 78**,** 5440-3.

JORDAN, P. W., KARPPINEN, J. & HANDEL, M. A. 2012. Polo-like kinase is required for synaptonemal complex disassembly and phosphorylation in mouse spermatocytes. *J Cell Sci,* 125**,** 5061-72.

KABSCH, W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr,* 66**,** 125-32.

KEENEY, S. 2008. Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome Dyn Stab,* 2**,** 81-123.

KEENEY, S., GIROUX, C. & KLECKNER, N. 1997. Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. . *Cell* 3.

KELLY, S. M., JESS, T. J. & PRICE, N. C. 2005. How to study proteins by circular dichroism. *Biochim Biophys Acta,* 1751**,** 119-39.

KLEIGER, G., SAHA, A., LEWIS, S., KUHLMAN, B. & DESHAIES, R. J. 2009. Rapid E2-E3 assembly and disassembly enable processive ubiquitylation of cullin-RING ubiquitin ligase substrates. *Cell,* 139**,** 957-68.

KLEIN, F., MAHR, P., GALOVA, M., BUONOMO, S. B., MICHAELIS, C., NAIRZ, K. & NASMYTH, K. 1999. A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. *Cell,* 98**,** 91-103.

KOBAYASHI, W., HOSOYA, N., MACHIDA, S., MIYAGAWA, K. & KURUMIZAKA, H. 2017. SYCP3 regulates strand invasion activities of RAD51 and DMC1. 22**,** 799-809.

KOLAS, N. K., YUAN, L., HOOG, C., HENG, H. H., MARCON, E. & MOENS, P. B. 2004. Male mouse meiotic chromosome cores deficient in structural proteins SYCP3 and

SYCP2 align by homology but fail to synapse and have possible impaired specificity of chromatin loop attachment. *Cytogenet Genome Res,* 105**,** 182-8.

KONAREV, P. V., VOLKOV, V. V., SOKOLOVA, A. V., KOCH, M. H. J. & SVERGUN, D. I. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography,* 36**,** 1277-1282.

KORNBERG, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science,* 184**,** 868-71.

KOUZNETSOVA, A., BENAVENTE, R., PASTINK, A. & HOOG, C. 2011. Meiosis in mice without a synaptonemal complex. *PLoS One,* 6**,** e28255.

KOUZNETSOVA, A., NOVAK, I., JESSBERGER, R. & HOOG, C. 2005. SYCP2 and SYCP3 are required for cohesin core integrity at diplotene but not for centromere cohesion at the first meiotic division. *J Cell Sci,* 118**,** 2271-8.

KOZIN, M. B. & SVERGUN, D. I. 2001. Automated matching of high- and low-resolution structural models. *Journal of Applied Crystallography,* 34**,** 33-41.

KSCHONSAK, M., MERKEL, F., BISHT, S., METZ, J., RYBIN, V., HASSLER, M. & HAERING, C. H. 2017. Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes. *Cell,* 171**,** 588-600.e24.

KSHIRSAGAR, R., GHODKE, I. & MUNIYAPPA, K. 2017. Saccharomyces cerevisiae Red1 protein exhibits nonhomologous DNA end-joining activity and potentiates Hop1-promoted pairing of double-stranded DNA. *J Biol Chem,* 292**,** 13853-13866.

LEE, J. 2013. Roles of cohesin and condensin in chromosome dynamics during mammalian meiosis. *J Reprod Dev,* 59**,** 431-6.

LEUNG, W. K., HUMPHRYES, N., AFSHAR, N., ARGUNHAN, B., TERENTYEV, Y., TSUBOUCHI, T. & TSUBOUCHI, H. 2015. The synaptonemal complex is assembled by a polySUMOylation-driven feedback mechanism in yeast. *J Cell Biol,* 211**,** 785-93.

LIU, J. G., YUAN, L., BRUNDELL, E., BJORKROTH, B., DANEHOLT, B. & HOOG, C. 1996. Localization of the N-terminus of SCP1 to the central element of the synaptonemal complex and evidence for direct interactions between the N-termini of SCP1 molecules organized head-to-head. *Exp Cell Res,* 226**,** 11-9.

LU, J., GU, Y., FENG, J., ZHOU, W., YANG, X. & SHEN, Y. 2014. Structural insight into the central element assembly of the synaptonemal complex. *Sci Rep,* 4**,** 7059.

LUGER, K., MADER, A. W., RICHMOND, R. K., SARGENT, D. F. & RICHMOND, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389**,** 251-60.

LYNN, A., SOUCEK, R. & BORNER, G. V. 2007. ZMM proteins during meiosis: crossover artists at work. *Chromosome Res,* 15**,** 591-605.

MALUMBRES, M. 2014. Cyclin-dependent kinases. *Genome Biol,* 15**,** 122.

MAOR-SAGIE, E., CINNAMON, Y., YAACOV, B., SHAAG, A., GOLDSMIDT, H., ZENVIRT, S., LAUFER, N., RICHLER, C. & FRUMKIN, A. 2015. Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia. *J Assist Reprod Genet,* 32**,** 887-91.

MARTINI, E., DIAZ, R. L., HUNTER, N. & KEENEY, S. 2006. Crossover homeostasis in yeast meiosis. *Cell,* 126**,** 285-95.

MCLACHLAN, A. D. S., M. 1975. Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *Journal of Molecular Biology,* 98**,** 293-304.

MCMAHILL, M. S., SHAM, C. W. & BISHOP, D. K. 2007. Synthesis-dependent strand annealing in meiosis. *PLoS Biol,* 5**,** e299.

MESELSON, M. S. & RADDING, C. M. 1975. A general model for genetic recombination. *Proceedings of the National Academy of Sciences of the United States of America,* 72**,** 358-361.

MEUWISSEN, R. L., OFFENBERG, H. H., DIETRICH, A. J., RIESEWIJK, A., VAN IERSEL, M. & HEYTING, C. 1992. A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *Embo j,* 11**,** 5091-100.

MILLER, M. P., UNAL, E., BRAR, G. A. & AMON, A. 2012. Meiosis I chromosome segregation is established through regulation of microtubule-kinetochore interactions. *Elife,* 1**,** e00117.

MIYAMOTO, T., HASUIKE, S., YOGEV, L., MADURO, M. R., ISHIKAWA, M., WESTPHAL, H. & LAMB, D. J. 2003. Azoospermia in patients heterozygous for a mutation in SYCP3. *Lancet,* 362**,** 1714-9.

MOSES, M. J. 1956. Chromosomal structures in crayfish spermatocytes. *J Biophys Biochem Cytol,* 2**,** 215-8.

NASMYTH, K. & HAERING, C. H. 2009. Cohesin: its roles and mechanisms. *Annu Rev Genet,* 43**,** 525-58.

NOTTKE, A. C., KIM, H. M. & COLAIACOVO, M. P. 2017. Wrestling with Chromosomes: The Roles of SUMO During Meiosis. *Adv Exp Med Biol,* 963**,** 185-196.

OFFENBERG, H. H., SCHALK, J. A., MEUWISSEN, R. L., VAN AALDEREN, M., KESTER, H. A., DIETRICH, A. J. & HEYTING, C. 1998. SCP2: a major protein component of the axial elements of synaptonemal complexes of the rat. *Nucleic Acids Res,* 26**,** 2572-9.

OLLINGER, R., ALSHEIMER, M. & BENAVENTE, R. 2005. Mammalian protein SCP1 forms synaptonemal complex-like structures in the absence of meiotic chromosomes. *Mol Biol Cell,* 16**,** 212-7.

ORTEGA, S., PRIETO, I., ODAJIMA, J., MARTIN, A., DUBUS, P., SOTILLO, R., BARBERO, J. L., MALUMBRES, M. & BARBACID, M. 2003. Cyclin-dependent kinase 2 is essential for meiosis but not for mitotic cell division in mice. *Nat Genet,* 35**,** 25-31.

PALM, W. & DE LANGE, T. 2008. How shelterin protects mammalian telomeres. *Annu Rev Genet,* 42**,** 301-34.

PARRY, D. A. D., FRASER, R.D.B., SQUIRE, J. M. 2008. Fifty years of coiled-coils and α-helical bundles: A close relationship between sequence and structure. *Journal of Structural Biology,* 163**,** 258-269.

PERANEN, J., RIKKONEN, M., HYVONEN, M. & KAARIAINEN, L. 1996. T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. *Anal Biochem,* 236**,** 371-3.

PETRONCZKI, M., SIOMOS, M. F. & NASMYTH, K. 2003. Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell,* 112**,** 423-40.

ROCKMILL, B., FUNG, J. C., BRANDA, S. S. & ROEDER, G. S. 2003. The Sgs1 helicase regulates chromosome synapsis and meiotic crossing over. *Curr Biol,* 13**,** 1954-62.

ROCKMILL, B. & ROEDER, G. S. 1988. RED1: a yeast gene required for the segregation of chromosomes during the reductional division of meiosis. *Proc Natl Acad Sci U S A,* 85**,** 6057-61.

SALI, A. & BLUNDELL, T. L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol,* 234**,** 779-815.

SCHMEKEL, K., MEUWISSEN, R. L., DIETRICH, A. J., VINK, A. C., VAN MARLE, J., VAN VEEN, H. & HEYTING, C. 1996. Organization of SCP1 protein molecules within synaptonemal complexes of the rat. *Exp Cell Res,* 226**,** 20-30.

SCHRAMM, S., FRAUNE, J., NAUMANN, R., HERNANDEZ-HERNANDEZ, A., HOOG, C., COOKE, H. J., ALSHEIMER, M. & BENAVENTE, R. 2011. A novel mouse synaptonemal complex protein is essential for loading of central element proteins, recombination, and fertility. *PLoS Genet,* 7**,** e1002088.

SCHÜCKER, K., HOLM, T., FRANKE, C., SAUER, M. & BENAVENTE, R. 2015. Elucidation of synaptonemal complex organization by super-resolution imaging with isotropic resolution. *Proceedings of the National Academy of Sciences,* 112**,** 2029-2033.

SCHWACHA, A. & KLECKNER, N. 1995. Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell,* 83**,** 783-91.

SEO, E. K., CHOI, J. Y., JEONG, J. H., KIM, Y. G. & PARK, H. H. 2016. Crystal Structure of C-Terminal Coiled-Coil Domain of SYCP1 Reveals Non-Canonical Anti-Parallel Dimeric Structure of Transverse Filament at the Synaptonemal Complex. *PLoS One,* 11**,** e0161379.

SHELDRICK, G. M. 2008. A short history of SHELX. *Acta Crystallogr A,* 64**,** 112-22.

SHINOHARA, A., OGAWA, H. & OGAWA, T. 1992. Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell,* 69**,** 457-70.

SLATER, M. L., SHARROW, S. O. & GART, J. J. 1977. Cell cycle of Saccharomycescerevisiae in populations growing at different rates. *Proc Natl Acad Sci U S A,* 74**,** 3850-4.

SMITH, A. V. & ROEDER, G. S. 1997. The yeast Red1 protein localizes to the cores of meiotic chromosomes. *J Cell Biol,* 136**,** 957-67.

SNOWDEN, T., ACHARYA, S., BUTZ, C., BERARDINI, M. & FISHEL, R. 2004. hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Mol Cell,* 15**,** 437-51.

SOLARI, A. J. & MOSES, M. J. 1973. The structure of the central region in the synaptonemal complexes of hamster and cricket spermatocytes. *J Cell Biol,* 56**,** 145-52.

SREERAMA, N. & WOODY, R. W. 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem,* 287**,** 252-60.

SVERGUN D.I., B. C. K. M. H. J. 1995. CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst.,* 28**,** 768-773.

SYM, M., ENGEBRECHT, J. A. & ROEDER, G. S. 1993. ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell,* 72**,** 365-78.

SYM, M. & ROEDER, G. S. 1994. Crossover interference is abolished in the absence of a synaptonemal complex protein. *Cell,* 79**,** 283-92.

SYM, M. & ROEDER, G. S. 1995. Zip1-induced changes in synaptonemal complex structure and polycomplex assembly. *J Cell Biol,* 128**,** 455-66.

SYRJANEN, J. L., HELLER, I., CANDELLI, A. & DAVIES, O. R. 2017. Single-molecule observation of DNA compaction by meiotic protein SYCP3. 6.

SYRJANEN, J. L., PELLEGRINI, L. & DAVIES, O. R. 2014. A molecular model for the role of SYCP3 in meiotic chromosome organisation. 3.

TAKAGI, J. & SHIMAMOTO, Y. 2018. Mechanical heterogeneity and roles of parallel microtubule arrays in governing meiotic spindle length. *bioRxiv*, 385633.

THOMAS, J. O. & KORNBERG, R. D. 1975. An octamer of histones in chromatin and free in solution. *Proc Natl Acad Sci U S A,* 72**,** 2626-30.

TOTH, A. & JESSBERGER, R. 2016. Oogenesis: Ageing Oocyte Chromosomes Rely on Amazing Protein Stability. *Curr Biol,* 26**,** R329-31.

TRIPET, B., VALE, R. D., HODGES, R. S. 1997. Demonstration of coiled-coil interactions within the kinesin neck region using synthetic peptides. *Journal of Biological chemistry,* 272**,** 8946-8956.

TRUEBESTEIN, L. & LEONARD, T. A. 2016. Coiled-coils: The long and short of it. *Bioessays,* 38**,** 903-16.

TUNG, K. S. & ROEDER, G. S. 1998. Meiotic chromosome morphology and behavior in zip1 mutants of Saccharomyces cerevisiae. *Genetics,* 149**,** 817-32.

TURNER, J. M. 2007. Meiosis 2007--where have we got to and where are we going? *Chromosome Res,* 15**,** 517-21.

VIVEIROS, M. M., O'BRIEN, M., WIGGLESWORTH, K. & EPPIG, J. J. 2003. Characterization of Protein Kinase C-δ in Mouse Oocytes Throughout Meiotic Maturation and Following Egg Activation1. *Biology of Reproduction,* 69**,** 1494-1499.

VOELKEL-MEIMAN, K., TAYLOR, L. F., MUKHERJEE, P., HUMPHRYES, N., TSUBOUCHI, H. & MACQUEEN, A. J. 2013. SUMO localizes to the central element of synaptonemal complex and is required for the full synapsis of meiotic chromosomes in budding yeast. *PLoS Genet,* 9**,** e1003837.

VON WETTSTEIN, D., RASMUSSEN, S. W. & HOLM, P. B. 1984. The synaptonemal complex in genetic segregation. *Annu Rev Genet,* 18**,** 331-413.

VONRHEIN, C., FLENSBURG, C., KELLER, P., SHARFF, A., SMART, O., PACIOREK, W., WOMACK, T. & BRICOGNE, G. 2011. Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D,* 67**,** 293-302.

WESTERGAARD, M. & VON WETTSTEIN, D. 1972. The synaptinemal complex. *Annu Rev Genet,* 6**,** 71-110.

WILLIAMS, R. C. & SPENGLER, S. J. 1986. Fibers of RecA protein and complexes of RecA protein and single-stranded phi X174 DNA as visualized by negative-stain electron microscopy. *J Mol Biol,* 187**,** 109-18.

WINKEL, K., ALSHEIMER, M., OLLINGER, R. & BENAVENTE, R. 2009. Protein SYCP2 provides a link between transverse filaments and lateral elements of mammalian synaptonemal complexes. *Chromosoma,* 118**,** 259-67.

WU, L. & HICKSON, I. D. 2003. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature,* 426**,** 870-4.

XU, D. & ZHANG, Y. 2012. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins,* 80**,** 1715-35.

YANG, F., DE LA FUENTE, R., LEU, N. A., BAUMANN, C., MCLAUGHLIN, K. J. & WANG, P. J. 2006. Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis. *J Cell Biol,* 173**,** 497-507.

YUAN, L., LIU, J. G., ZHAO, J., BRUNDELL, E., DANEHOLT, B. & HOOG, C. 2000. The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. *Mol Cell,* 5**,** 73-83.

YUAN, L., PELTTARI, J., BRUNDELL, E., BJORKROTH, B., ZHAO, J., LIU, J. G., BRISMAR, H., DANEHOLT, B. & HOOG, C. 1998. The synaptonemal complex protein SCP3 can form multistranded, cross-striated fibers in vivo. *J Cell Biol,* 142**,** 331-9.

ZAVEC, A. B., COMINO, A., LENASSI, M. & KOMEL, R. 2008. Ecm11 protein of yeast Saccharomyces cerevisiae is regulated by sumoylation during meiosis. *FEMS Yeast Res,* 8**,** 64-70.

ZAVEC, A. B., LESNIK, U., KOMEL, R. & COMINO, A. 2004. The Saccharomyces cerevisiae gene ECM11 is a positive effector of meiosis. *FEMS Microbiol Lett,* 241**,** 193-9.

ZICKLER, D. & KLECKNER, N. 1999. Meiotic chromosomes: integrating structure and function. *Annu Rev Genet,* 33**,** 603-754.

ZICKLER, D. & KLECKNER, N. 2015. Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harb Perspect Biol,* 7.