Structural and Biochemical Studies of the *S. cerevisiae* Condensin Complex

---

Author
Christopher Cawood

Supervisor
Luis Aragon

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Cell Cycle Group

MRC London Institute of Medical Sciences
Imperial College London

April 2020

Abstract

The Structural Maintenance of Chromosomes (SMC) complex family forms an essential part of the genome's organisational machinery in all domains of life. In eukaryotes condensin is one such complex that mediates the packaging of DNA for segregation during cell division. This pentameric ATPase is comprised of two elongated Smc proteins (Smc2 and Smc4), a kleisin type protein (Brn1) and two HEAT repeat proteins (Ycg1 and Ycs4).

Despite the important role of condensin in such fundamental cell processes its overall structure, and consequently its molecular mechanism, are currently unknown. This study describes the processes leading to production of the first atomic resolution model of the *Saccharomyces cerevisiae* condensin complex in the ATP free apo state by cryo electronmicroscopy. This model demonstrates that several features such as a folded conformation in the Smc protein arms are common to multiple members of the SMC complex family. Furthermore, previously described structural features such as the location of the Ycs4 subunit in contact with the Smc protein head domains, and a discontinuity in the arms of the Smc proteins called the joint, are also described as part of the whole complex. Finally, the position of the Ycg1 subunit is also described as sitting flexibly under the Smc head domains, anchored by the Brn1 subunit to the rest of the complex.

Also described in this study is preliminary work to further characterise the biochemical properties of the condensin complex in the context of chromatin using the Lumix C-trap optical tweezer system. Condensins' ability to compact DNA in real time and to translocate on a fixed DNA substrate in the C-trap system are demonstrated, as is ongoing work to chromatinise the same DNA substrate *in situ*. This chromatinisation involves the purification of recombinant *S. cerevisiae* histone octamers and histone chaperone Nap1 which have been successfully applied to DNA within the C-trap to create a chromatin substrate suitable for future experimentation.

Decleration of Originality

All work contained within this thesis is my own except where indicated. Any non-original works are appropriately referenced and credited where appropriate.

Copyright Decleration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I would like to thank Professor Luis Aragon for his guidance the opportunity to work on this project. I would also like to thank all members of the Cell Cycle Group past and present but especially Dr Pilar Gutierrez-Escribano and Dr Jonay Garcia-Luis for their support and input throughout. Additionally, I would like to thank Dr Alejandra F. Cid for taking the time to train me in the earliest parts of the project, and Joseph Davy for his assistance throughout the final year.

I would also like to thank members the Lowe laboratory (LMB, Cambridge) for their work in cryo EM, especially Dr Byung-Gil Lee whose collaboration and technical knowledge was essential to this project. I would also like to acknowledge members of the Single Molecule Imaging Group (Imperial College London) for their collaborative effort on the single molecule assays.

Finally, I would like to thank all my friends and family who have been with me for everything outside the lab, with special thanks to my Mum and Faye.

Contents

Abstract 2

Decleration of Originality 3

Copyright Decleration 3

Acknowledgements 4

List of Figures and Tables 8

List of Abbreviations 10

Chapter 1 – Introduction 11

1.1. The SMC Complexes 11
1.1.1. Subunits of the SMC complexes 12

1.2. Smc subunits of the SMC complexes 15
1.2.1 The ATPase domain 15
1.2.2 Coiled-coils 19
1.2.3 Hinge domain 22

1.3. The Kleisin Subunit 24

1.4. KITEs and HAWKs 27
1.4.1. KITES 28
1.4.1 i Bacterial KITEs 28
1.4.1 ii Eukaryotic KITEs 29
1.4.2 HAWKs 30
1.4.2 i HAWKs in Condensin 31
1.4.2 ii HAWKs in cohesin 33
1.4.2 iii Regulation of Cohesin Through HAWKs 34

1.5. Current Models of SMC Complex Function 34
1.5.1 DNA Looping – Passive vs Active Loop Formation 35
1.5.2 Molecular Mechanisms Underpinning the SMC Complex Motor 35
1.5.2 i Evidence for an ATP Driven Mechanism 35
1.5.2 ii DNA Binding Sites and Conformational Changes 36
1.5.3 Applying the SMC Pump Mechanics to Asymmetric or Symmetric Loops 39

1.6. Condensin chromatin dynamics 40
1.6.1 Organisation of the genome 41
1.6.2 Cell cycle stages 44
1.6.3 Condensin in Mitosis 46
1.6.3. i. The rDNA and the Centromere in Mitosis 48
1.6.3 ii. Mitotic Post-Translational Modification 50
1.6.3. iii Condensin-protein Interactions in Mitosis 51
1.6.4. Condensin in Meiosis 52
1.6.5. Condensin in Interphase 53
1.6.5. i Condensin and Interphase TADs 54
1.6.5. ii Condensin and Chromatin Territories 55
1.6.5 iii Condensin and Further Transcriptional Regulation 55

1.8. Condensin in disease 56

1.9 Project Aims 57

Chapter 2 – Purification of the *S. cerevisiae* Condensin Complex 59

2.1 Purification and Cryo EM analysis of the Condensin Complex 59

2.2 Purification and Cryo EM Analysis of Condensin Subcomplexes 63

Chapter 3 – Purification of Ycs4 69

3.1 Purification of Ycs4 for Crystallography – Preliminary GST Tagging and Optimisation 69

3.2 Purification of Ycs4 with Histidine Tagging 77

3.3 Purification of C. albicans Ycs4 85

3.4 Purification of S. cerevisiae Ycs4 in Insect Cells 86

Chapter 4 – Co-purification of Ycs4 and Condensin Subunit Fragments 89

4.1 Purification of a Brn1 270 – 423 Fragment 89

4.2 Purification of the Smc4 Head 95

4.3. Final Structure of the S. cerevisiae Condensin Complex 101

Chapter 5 – Characterisation of Condensin 106

5.1. Condensin DNA binding 106

5.2. Ycs4 DNA Binding 107

5.3. Single Molecule Assays with Condensin 109

5.3.1 Condensation Assays 111

5.3.2 Translocation assays 113

Chapter 6 – Chromatinisation of a Single Molecule Substrate for Condensin Assays . 115

6.1 Histone Purification Optimisation 115

6.2 Chaperone Mediated in situ Chromatin Assembly 119

Chapter 7 – Discussion 126

7.1 Analysis of Key Condensin Structural Features 126

7.1.1. The Condensin Elbow 126

7.1.2. Future Analysis of the Condensin Elbow 127

7.1.3. The Condensin Joint 128

7.1.4. Future Analysis of the Condensin Joint 128

7.1.5 Analysis of the HEAT Repeat Subunits Ycs4 and Ycg1 in the Condensin Structure 129

7.1.5. i. Ycg1 129

7.1.5 ii Future Analysis of Ycg1 130

7.1.5 iii Ycs4 131

7.1.5 iv Future Analysis of Ycs4 132

7.2 Technical Improvements to Purification Techniques 133

7.3 Preliminary Single Molecule Analysis of Condensin 134

7.4 Future Single Molecule Analysis of Condensin on a Chromatinised Substrate 135

7.4 Concluding remarks 135

Chapter 8 - Methods 137

8.1 Purification of condensin pentamers, tetramers and trimers 137

8.2 Protein Expression Assays 138

8.3 Protein Solubility Assays *138*
8.4 Purification of S. cerevisiae Ycs4-6xHis from E. coli *139*
8.5 Purification of C. albicans 6xHis-TEV-Ycs4 from E. coli *140*
8.6 Purification of S. cerevisiae Ycs4-6xHis from Insect Cells *141*
8.7 Purification of S. cerevisiae Ycs4-6xHis and Brn1 270-423 from E. coli *141*
8.8 Electrophoretic Mobility Shift Assays *142*
8.10 Cross-linking Mass Spectrometry *146*
8.11 Western Blotting *148*
8.12 Condensin Force-clamp and Force-extension Optical Tweezer Experiments *148*
8.13 Condensin Translocation Kymographs *149*
8.14 S. cerevisiae Histone Octamer purification *149*
8.15 Individual Histone Purification *150*
8.16 Nap1 purification *151*
8.17 ISWIa complex purification *152*
8.18 In vitro chromatin assembly *152*
8.19 Micrococcal nuclease digestion assay *153*
8.20 Histone labelling *153*
8.21 Protein quantification *154*
8.22 Tables of Strains, Constructs, Antibodies and Rights and Permissions *154*

Bibliography 158

List of Figures and Tables

Figure 1.1: *Cartoon representation of eukaryotic and bacterial SMC complexes*…… 14
Figure 1.2: *Smc protein heads*……..... 16
Figure 1.3: *Summary of the knobs-in-holes model of coiled-coil assembly in Smc proteins*.. 19
Figure 1.4: *Cartoon representation of the potential conformations adopted by Smc dimer*... 21
Figure 1.5: *Kleisin interaction with Smc heads* …….. 25
Figure 1.6: *Cartoon representation of the kleisin compartment* …… 26
Figure 1.7: *KITE-Kleisin interaction* ……. 28
Figure 1.8: *HAWK-Kleisin interaction*…… 31
Figure 1.9: *Models of loop extrusion*…….. 39
Figure 1.10: *Cartoon summary of the major structures found in the interphase genome*….. 43
Figure 1.11: *The cell cycle*…….. 45
Figure 2.1: *Purification of the condensin pentamer*…… 60
Figure 2.2: *Widefield image of WT S. cerevisiae condensin*……... 61
Figure 2.3: *Cryo EM of the condensin pentamer*…… 62
Figure 2.4: *Purification of condensin trimer and tetramers*……... 64
Figure 2.5: *Western blot analysis of whole cell extracts taken from cells expressing condensin sub complexes*…….. 65
Figure 2.6: *Widefield image of WT S. cerevisiae condensin*……... 66
Figure 2.7: *Cryo EM of the condensin ΔYcg1 tetramer*…….. 67
Figure 2.8: *Low resolution 3D structure of the condensin complex*……... 68
Figure 3.1: *GST-Ycs4 expression screen*……. 70
Figure 3.2: *Further GST-Ycs4 expression screening*…….. 71
Figure 3.3: *Purification of S. cerevisiae GST-Ycs4*……. 73
Figure 3.4: *Purification of C. albicans GST-Ycs4*……... 75
Figure 3.5: *GST-Ycs4 solubility and expression screen*…….. 77
Figure 3.6: *Purification of Ycs4-6xHis*…… 79
Figure 3.7: *Solubility screening of Ycs4-6xHis purification*……... 81
Figure 3.8: *Summary of S. cerevisiae Ycs4-6xHis expression and purification*…….. 83
Figure 3.9: *Summary of final S. cerevisiae Ycs4-6xHis purification*…….. 84
Figure 3.10: *Summary of final C. albicans Ycs4-6xHis purification*……... 86

Figure 3.11: *Summary of S. cerevisiae Ycs4-6xHis expressed in Sf9 insect cells*…….. 88
Figure 4.1: *Design of Brn1 fragments for purification*……....90
Figure 4.2: *Brn1 270-423 fragment expression screening*……...92
Figure 4.3: *Purification of Brn1 270-423*…… 93
Figure 4.4: *Co-purification of Brn1 270-423 and Ycs4-6xHis*…… 95
Figure 4.5: *Illustration of CLMS cross-links between Ycs4 and Smc4 in the ΔYcg1 tetramer*……... 96
Figure 4.6: *Selection of the Smc4 head domain for purification*……. 97
Figure 4.7: *Smc4 head expression screen*……... 98
Figure 4.8: *Further expression and purification of the Smc4 head*……. 99
Figure 4.9: *Further expression screening of Smc4 head*……... 100
Figure 4.10: *Comparison of S. cerevisiae and C. thermophilum Ycs4*……..… 102
Figure 4.11: *Final structure of the condensin complex*……. 105
Figure 5.1: *Electrophoretic mobility shift assay of pentameric and ΔYcg1 tetramer complexes*…….. 107
Figure 5.2: *Electrophoretic mobility shift assays of condensin and Ycs4 homologs*…… 108
Figure 5.3: *Cartoon summary of the Lumix C Trap setup*……. 110
Figure 5.4: *Single molecule condensation assays*……. 112
Figure 5.5: *Condensin translocation kymographs*…… 114
Figure 6.1: *Histone purification*…… 118
Figure 6.2: *Purification of components for in vitro assembly of chromatin*……. 121
Figure 6.3: *Chromatin test assembly*……. 122
Figure 6.4: *Assembly of chromatin in situ*……. 123
Figure 6.5: *Fluorescent labelling of histones*…… 125

Table 1: *Summary of the major subunits of SMC complexes*…….. 12
Table 2: *List of yeast strains*……..……..……. 154
Table 3: *List of constructs*……. 156
Table 4: *List of antibodies*……. 156
Table 5: Rights and Permissions……... 157

List of Abbreviations

Abbreviation	Meaning
ABC	ATP Binding Cassette
Bs	*Basillus subtillis*
CLMS	Cross Linking Mass Spectometry
Cryo EM	Cryo Electron Microscopy
Ct	*Chaetomium thermophillum*
CV	Column Volume
GST	Glutathione S-Transferase
HA	Haemaglutinin
HAWK	HEAT proteins Associated With Kleisins
HEAT	Hunutintin, Elongation factor 3, protein phosphatase 2A, Tor1
His	Histidine
HJURP	Holiday Junction Recognition Protein
ITC	Isothermal Titration Calorimetry
KITE	Kleisin-Interacting tandem Winged helix Element
MAGE	Melanoma Antigen Gene
OD	Optical Density
RING	Really Interesting New Gene
Sc	*Saccharomyces cerevisiae*
SEC	Size Exclusion Chromatography
SMC	Structural Maintenance of Chromosomes
SUMO	Small Ubiquitin-like Modifier
TAD	Topologically Associated Domain
TFIIIC	Transcription Factor III C
TEV	Tobaco Etch Virus
WHD	Winged Helix Domain
WT	Wild Type

Chapter 1 – Introduction

1.1. The SMC Complexes

The Structural Maintenance of Chromosomes (SMC) family of protein complexes form a key part of the genome's regulatory and structural machinery. These large assemblies are found in related forms throughout all three domains of life, contributing essentially to such fundamental processes as the packaging and topological arrangement of DNA throughout the cell cycle (Kimura and Hirano 1997), sister chromatid cohesion (Losada et al. 1998), DNA damage repair (De Piccoli et al. 2006, Heale et al. 2006, Potts et al. 2006, Bauerschmidt et al. 2010) and transcriptional regulation (Dowen et al. 2013). As a consequence of their critical role in basic cell processes, SMC complexes are remarkable in both their ubiquity and diversity. Mutations within SMC complexes have been linked to human diseases, such as microcephaly (Martin et al. 2016), T-cell lymphoma (Woodward et al. 2016) and Cornelia de Lange syndrome (Mehta et al. 2013). Therefore, the broad reach of the SMC complexes have made them important candidates for detailed studies.

In eukaryotes there are 3 related SMC complexes, cohesin, condensin and SMC5/6, each fulfilling distinct roles within the cell. The first, cohesin, is loaded during DNA replication, entrapping newly formed sister chromatids from S phase to anaphase. This complex was discovered simultaneously by (Michaelis et al., 1997) and (Guacci et al., 1997) via screens for mutants displaying sister chromatid cohesion defects. This cohesion function is essential for proper alignment of chromosomes along the metaphase plate such that they can be equally segregated into newly formed daughter cells. Cohesin has also been shown to have important functions in interphase, where it regulates the formation of topologically associated domains (TADs). These domains arrange the regulatory elements of the genome in space, for example bringing distal enhancer and promoter elements into proximity, thus allowing fine control of gene expression.

Condensin, another SMC complex and the focus of my study, is the key mediator of DNA condensation during mitosis and was originally identified and characterised in *Xenopus laevis* egg extracts by (Hirano et al., 1997). During cell division, the genome undergoes a

significant structural transition from that of loosely organised interphase chromatin to the much more compacted form present on mitotic chromosomes with tight loops organised along an axial core. This transition is essential for proper genome segregation during cell division, facilitates the removal of catenation and provides mechanical strength to chromosomes to allow them to resist the pulling stress of the mitotic spindle.

The third eukaryotic SMC complex is SMC5/6, and it is mainly involved in the process of homologous recombination repair in response to DNA damage, as well as being implicated in telomere maintenance and replication fork stability. The whole complex was first identified in *Schizosaccharomyces pombe* through tagging and purification of the Smc6 subunit (Fousteri et al., 2000).

Bacteria also contain SMC complexes: MukBEF (found in enterobacteria and species of gamma-proteobacteria) (Niki et al., 1991; Yamanaka et al., 1996), MksBEF (a recently identified family in *Pseudomonas aeruginosa*) (Petrushenko et al., 2011), and Smc-ScpAB (found in most sequenced species) (Mascarenhas et al., 2002). Bacterial SMC complexes are broadly involved in the proper segregation of DNA during cell division, which often leads to them being considered as bacterial condensins. MukBEF and Smc-ScpAB are typically mutually exclusive, though MksBEF can be found alongside other bacterial condensins.

While the roles of the individual complexes are diverse, they each possess core structural motifs within conserved subunits that have been preserved throughout evolution. In turn each of these elements will be discussed hereafter, summarising our current understanding of SMC structure and mechanics.

1.1.1. Subunits of the SMC complexes

The SMC complexes are comprised of multiple proteins whose interactions generate dynamyc structures producing a diverse range of functions. Table 1 summarises the known SMC families and their subunits, as well as some of the known roles that each complex plays.

Domain	**Eukaryote**					**Bacteria**		
SMC Family	**Condensin**			**Cohesin**	**SMC5/6**	**Smc-ScpAB**	**MukBEF**	**MksBEF**
	I	**II**	**DC**					
Function	**Chromatin organisation during cell division**	**Chromatin organisation during cell division, interphase chromatin regulation**	**Dosage compensation of C. elegans X chromosome**	**Sister chromatid cohesion, interphase genome organisation into topological domains**	**Double strand break repair, resolution of recombination structures, replication fork stability**	**Genome compaction in cell division and segregation**	**Genome compaction in cell division and segregation**	**Genome compaction in cell division and segregation**
SMC proteins	Smc2/Smc4	Smc2/Smc4	Smc2/DPY 27 (Smc4)	Smc1/Smc3	Smc5/Smc6	Smc	MukB	MksB
Kleisin	Brn1 (CAP H)	- (CAP H2)	- (DPY 26)	Scc1 (Rad21), Rec8*	Nse4	ScpA	MukE	MksE
HAWK	Ycs4 (CAP D2) / Ycg1 (CAP G)	- (CAP D3) / - (CAP G2)	- (DPY 28) / - (CAP G)	Scc3 (SA1, SA2) / Pds5, Scc2	- / -	-	-	-
KITE	-	-	-	-	Nse1 / Nse3	ScpB	MukF	MksF
Other	-	-	-	-	Nse2 / Nse5 / Nse6	-	-	-

Table 1. Summary of the major subunits of SMC complexes. Brackets indicate subunit homologs in higher eukaryotes. (*) indicates a meiosis specific variant of the cohesin kleisin subunit Rec8.

At the core of every SMC complex is a common trimeric protein assembly, consisting of two Smc proteins bridged by a kleisin type protein (Note: For the purpose of clarity, the SMC complexes shall hereafter be distinguished by use of capital letters from the individual Smc proteins. Smc proteins will be referred to using the yeast nomenclature).

In eukaryotes, six Smc proteins (Smc1 – Smc6) pair to form three discrete heterodimers which define the three known SMC complex types; cohesin (Smc1 and Smc3), condensin (Smc2 and Smc4) and the SMC5/6 complex (Smc5 and Smc6). Bacteria differ from eukaryotes in that they usually have a single expressed Smc protein that forms homodimers. In Smc-ScpAB the protein forming these homodimers is simply called Smc. In MukBEF the Smc subunit is MukB and in MksBEF it is MksB. In addition to the core Smc-kleisin trimer, additional protein subunits are also present on SMC complexes. These can be divided into two types: KITEs (Kleisin-Interacting tandem Winged helix Element) and HAWKs (HEAT proteins Associated With Kleisins). KITEs are associated with the SMC5/6 complex and the bacterial SMC complexes while HAWKs are larger proteins associated with the cohesin and condensin complexes.

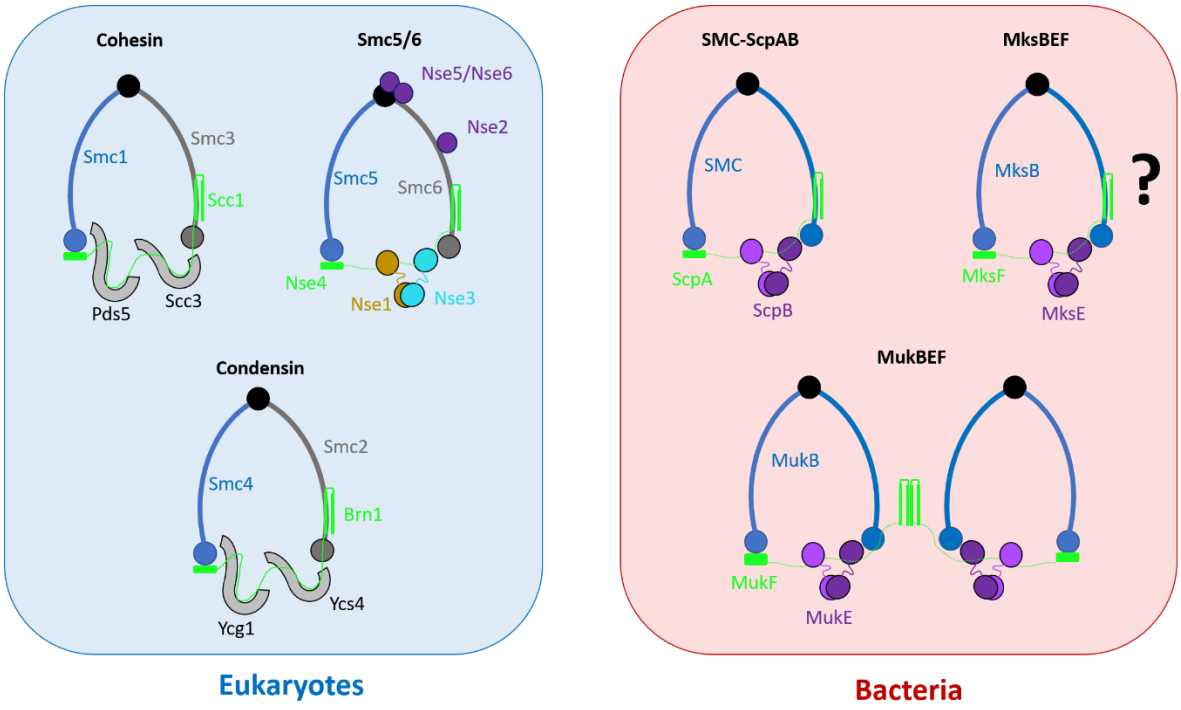


Figure 1.1. Cartoon representation of eukaryotic and bacteria SMC complexes with their respective subunits. Note that the higher order structure of the bacterial MksBEF is currently unknown. It is possible that due to its structural similarities to the related complex MukBEF it exists as a dimer. Eukaryotic complexes use the nomenclature for the subunits found in yeast.

1.2. Smc subunits of the SMC complexes

Structurally the Smc proteins are large proteins with conserved ATPase domains. Another characteristic feature is the presence of two coiled coil domains separated by a hinge region. The long (up to 50 nm), antiparallel stretches of alpha helices in the coils are separated by a globular hinge domain, which orients the coils such that the Smc's N and C terminal domains terminate together. These juxtaposed N and C termini form a bi-lobed, globular "head" domain, which cooperates with the head domain of the dimerised second Smc protein to form an "ATP Binding Cassette" (ABC) type ATPase (Lowe et al., 2001, Wilkens, 2015).

1.2.1 The ATPase domain

The two opposing heads are structurally similar, even in heterodimers, and oppose each other in an anti-parallel fashion. This orientation creates 2 ATP binding pockets sandwiched between key conserved motifs of the first lobe of one Smc (known as RecA-like or lobe I) and the second lobe of the other Smc (known as helical or lobe II).

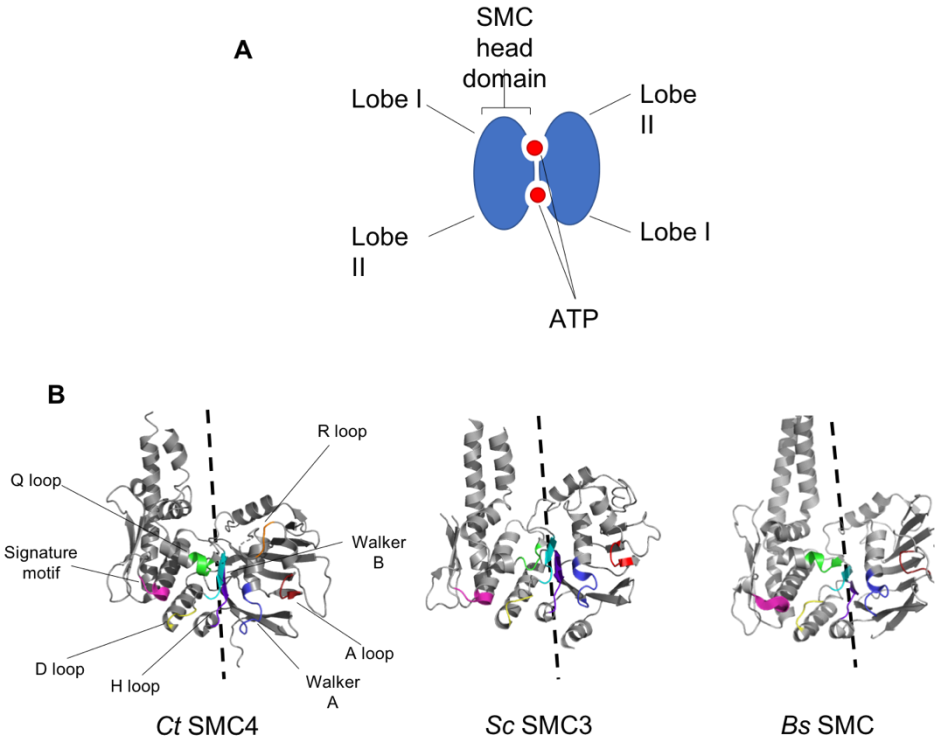


Figure 1.2. Smc protein heads (A) Cartoon representation of engaged Smc ATPase head domains binding two ATP molecules (red). Lobe I of the left most ATPase head forms a binding pocket with Lobe II of the right most ATPase head and viceversa. (B) Comparison between published structures of the Smc ATPase heads of Chaetomium thermophillum Smc4 (PDB 6qj2, Hassler et al., 2019), Sacharmoyces cerevisiae Smc3 (PDB 4ux3, Gligoris et al., 2014) and Bacilus subtilis Smc (PDB 5h66, Kamada et al., 2017). Key motifs are indicated by colour: A loop (red), Walker A (blue), R loop (orange), Q loop (green), Signature (magenta), Walker B (cyan), D loop (yellow) and H loop (purple). Dotted lines separate the helical lobe (lobe II, left) from the Rec-A-like lobe (lobe I, right).

The first lobe of the head domain is structurally similar to that of Rec-A (Rec-A-like) (Story et al., 1992, Ye et al., 2004), and can be found in many proteins containing ABC type ATPases such as the DNA repair protein Rad50 (Hopfner et al., 2000), helicases (Subramanya et al., 1996) and transmembrane transport proteins (Chen et al., 2003). Within this lobe can be found the Walker A motif (also known as the P loop), with the general consensus sequence GXXXXGK[T/S], where X is any amino acid (Saraste et al., 1990; Walker et al., 1982). The motif is essential for the coordination of ATP through an electrostatic interaction between the conserved lysine and the beta/gamma phosphates of ATP. Mutations in this region are poorly tolerated in Smc proteins, resulting in loss of function (Hudson et al., 2008, Kinoshita et al., 2015). In addition, a conserved aromatic

residue upstream of the Walker A motif, found in a sequence known as the A loop, provides a second ATP interaction site within the first lobe (Ambudkar et al., 2005). Here, the side chain of a residue such as tyrosine, phenylalanine or tryptophan is thought to form a π-π interaction with the adenine ring of ATP, stacking the aromatic rings parallel (Ambudkar et al., 2005).

Within the N terminal portion of lobe I can also be found the R and Q loops. The R loop is named because of the presence of a strongly conserved arginine residue that has been implicated in the up-regulation of ATPase activity in response to DNA binding (Lammens et al., 2004). The Q loop, as with the R loop, is named because of a single highly conserved glutamine residue. This eponymous glutamine maintains an interaction with the gamma phosphate of the ATP via a coordinated Mg^{2+} ion which contributes to nucleotide binding (Hudson et al., 2008, Hopfner et al., 2000). It has long been suggested that it is this motif that acts as a "lever" in members of the ABC family through which ATP binding and hydrolysis is coupled to conformational changes that drive molecular function (Hopfner and Tainer 2003). Recently, in a *C. thermophilum* Smc4 head construct, this motif was shown to form part of a gating mechanism in which repositioning of the Q loop to coordinate Mg^{2+} upon ATP binding at the P loop, resulted in a conformational change within the Smc head that forces dissociation of one of condensin HEAT repeat subunits (discussed in section 1.4.2). This uncovered the Smc4 head, and allowing its dimerization with Smc2 (Hassler et al., 2019). Mutation of this same glutamine in Smc2 *in vivo* prevented loading of condensin onto chromosomes (Hudson et al., 2008).

Within the C terminal portion of lobe I, the Walker B motif can be found. This motif has the consensus sequence hhhhDE, where h is any hydrophobic residue (Orelle et al., 2003; Walker et al., 1982). Molecular dynamic simulations have suggested several mechanisms through which ATP hydrolysis might take place, with the requirement in all cases that a proximal residue within the catalytic site can act as a proton acceptor and catalytic base (Zhou et al., 2013, Hsu et al., 2016, Prieß et al., 2018) One proposed key residue is a glutamate within the Walker B region (Orelle et al., 2003), though the H loop (another C terminal motif with conserved histidine) has also been implicated in the proton acceptor role in other ABC family members (Zhou et al., 2013).

As with the allosteric repositioning of the Q loop, the ATP hydrolysis cycle is likely translated into mechanical work through conformational changes in the Smc head domains.

Not only does the binding of an ATP to one catalytic site induces conformational shifts, but also binding of ATP provides an intermediate through which opposing Smc heads within the same dimer can interact, in addition to their constitutive hinge dimerization. Head engagement is a requirement for ATP catalysis.

Two further conserved motifs on the opposing lobe II bind ATP *in trans*, sandwiching the two ATP's at both binding sites. One of these motifs is the signature motif or C loop. This is described as a gamma phosphate sensor containing the sequence LSGGQ, in which the conserved serine and second glycine reach across to the ATP bound by the Walker A domain of the opposing Smc head, and form hydrogen bonds with the gamma phosphates oxygen atoms (Hopfner et al., 2003, Arumugam et al., 2006). The hydrolysis of ATP to ADP, releases the gamma phosphate to which the signature motif is bound, suggesting a possible means by which the two head domains become unbound. In addition, the D loop is another motif that acts in *trans*, helping to present the attacking nucleophilic water in the opposing lobe I (Hopfner et al., 2000). Its proximity to the Walker B motif makes the D loop a strong candidate for a mediator of allosteric control between catalytic sites (Hopfner et al., 2000).

The exact relationship between the two ATPase sites, and whether there are significant functional differences has recently begun to be explored. Evidence in Rad50 (Hopfner et al., 2000), cohesin (Arumugam et al., 2003) and condensin (Hudson et al., 2008) suggest that ATP catalysis is cooperative and proper Smc function is dependent upon binding at both sites, but the results of this binding in terms of mechanical output are largely unknown.

A very recent study in condensin has highlighted striking differences in the contribution of neighbouring ATPase sites to condensin function (Elbatsh et al 2019). The study showed that one site constituting lobe I of Smc4 and lobe II of Smc2 (known as AS1), mediates hyper-activation of DNA condensation activity, while the second site (AS2) acts to attenuate it (Elbatsh et al 2019). It is interesting to note that lobe I of AS1 has been implicated in the Q loop gating mechanism that precedes head dimerization, described above (Hassler et al., 2019). As a result, ASI appears to favour the formation of loops, possibly by involvement in the specific conformational changes required in the earliest stages of loop extrusion (discussed in section 1.5). These differences may represent a fundamental part of the SMC complexes underlying mechanism.

1.2.2 Coiled-coils

The coiled-coils of Smc proteins are made up of anti-parallel strands of alpha helix that twist around each other to form a near continuous extended structure, linking the head and hinge domains (Niki et al., 1991, Strunikov et al., 1993). The physical properties of this domain are subject to ongoing investigation, particularly with regards to their flexibility. Current evidence suggests they could form flexible rope-like structures (Eeftens et al., 2016) or more rigid rod-like structures (Deibold-Durand et al., 2017). The presence of a coiled-coil domain is highly characteristic of Smc proteins, and thus key to their function, yet unified models linking the structure of the coils to the various genomic functions of SMC complexes are still only speculative.

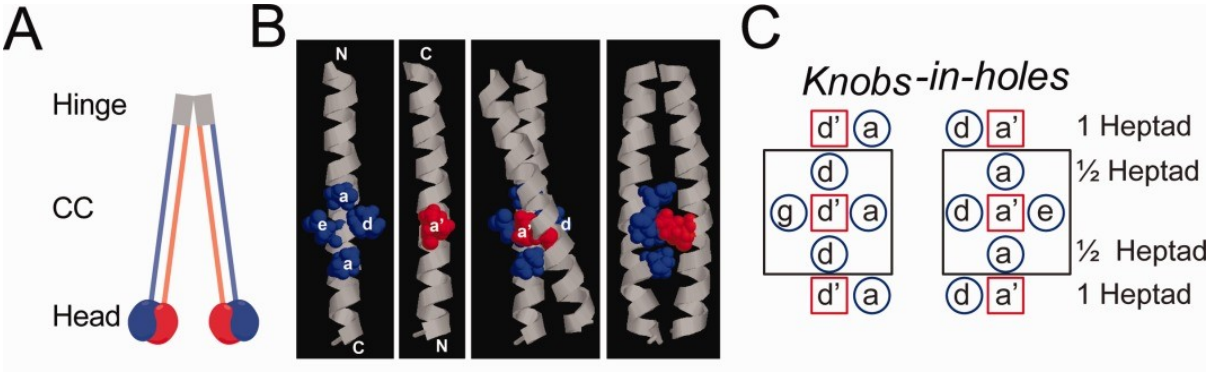


Figure 1.3. Summary of the knobs-in-holes model of coiled-coil assembly in Smc proteins. (A) Cartoon schematic showing dimerised Smc proteins with N terminal residues in blue and C terminal residues in red. (B) Structure of the heptad repeat shown on an antiparallel coiled-coil (PDB 1hf9, Gordon-Smith et al., 2001). The 7 residues of the heptad repeat are labelled a-g with residues b, c, f and g being omitted. Residues a, d and e of the N terminal strand form a hole into which the a residue of the C terminal (a') can fit. (C) Schematic of two types of knobs-in-holes packing in antiparallel coiled-coils showing how d' of the C terminal strand fits into a pocket formed by residues a, g and d across two heptads and how a' of the C terminal strand fits in a pocket of residues a, d and e on the C terminal strand. Figure adapted from Waldman et al., 2015 with permission from the copyright holder Wiley Periodicals Inc.

The coiled-coils within one Smc protein form through an interaction known as knobs-in-holes, in which a repeated heptad motif of amino acids presents a stripe of hydrophobic residues along one side of a single strand of alpha helices. These hydrophobic residues are arranged such that they form binding pockets, or holes, into which hydrophobic residues from

a heptad on the opposing strand can fit (Figure 1.3 A-C). The orientation of these reciprocal binding pockets curves around the surface of the alpha helical strand creating the characteristic twisted appearance of the coiled-coils (Waldmann et al., 2015, Lupas and Gruber, 2005).

The structural relationship between the coiled-coils of dimerised Smc proteins is still under investigation, though cross-linking mass spectrometry (CLMS) has shown multiple inter coiled-coil contacts between both condensin Smc2/Smc4 dimers and cohesin Smc1/Smc3 dimers. These contacts are dynamic, which suggests that opening and closing of the inter-arm space between Smc proteins by juxtaposing the coiled-coils can take place (Barysz et al., 2015, Kulemzina et al. 2016). Furthermore, multiple conformations have been observed in condensin Smc2/Smc4 dimers including rod-like structures with tightly juxtaposed coiled coils (Anderson et al., 2002), as well as V shaped dimers that interact solely at their hinge domains (Eeftens et al., 2016). Most interestingly, the Smc proteins also appeared to bend around their mid-point when analysed with atomic force microscopy (AFM), demonstrating hinge domains can bind one or both of the Smc head domains in a P or B like arrangement (Eeftens et al., 2016). These experiments suggested a significant amount of flexibility in the coiled-coils which is interesting because it may be assumed that a degree of rigidity would be required for the transfer of mechanical force through the condensin complex (Eeftens et al., 2016)

Structural and sequential homology are at their most conserved in the head and hinge domains of Smc proteins, with the most interspecies variation seen in the coiled-coils (Griese et al., 2010, Bürmann et al., 2017), however, the coiled-coils show a striking correlation in length between species. This clustering of arm lengths is periodic, with peaks at steps of roughly 100 amino acids and this high degree of structural conservation would seem to indicate a need to link the hinge and head domains in a very specific orientation to conserve complex function. Introduction of flexible peptide inserts into the coiled-coil of *B. subtilis* Smc homodimers has deleterious effects on bacterial SMC complex function that are independent of ATPase activity or subunit assembly, meaning a degree of rigidity is required in the bacterial SMC complexes (Burmann et al., 2017). This is possibly to allow conduction of conformational change as part of an ATP-dependant mechanism. The flexibility observed in isolated Smc dimers considering this evidence may point to a need for other SMC complex subunits to stabilise the structure. In eukaryotes, a similar disruption to the condensin Smc2

coiled-coil by insertion of 5 Myc repeats (the smc2-8 allele in yeast) results growth arrest (Freeman et al., 2000).

In addition to the above, recent analysis of bacterial MukBEF and *S. cerevisiae* cohesin by cryo electron microscopy (Cryo EM) have revealed an intriguing new structural feature to the coiled-coils that presents itself in the context of the whole complex (Burmann et al., 2019). Around the mid-point of the coil-coils in these complexes a discontinuity called the "elbow" can be found which allows the SMC complex to adopt a folded conformation that brings the hinge and head domain into proximity. The coiled-coils are juxtaposed in this configuration (figure 1.4).

The conservation of this "elbow" fold between complexes in such distantly related species would suggest a role in a shared mechanism. Exactly what this mechanism may be is unclear (discussed in section 1.5), but what seems apparent is that rigidity in the coils is needed for switching between multiple conformations in an ATP-dependant manner, and that the coils are doing more than simply entrapping DNA.

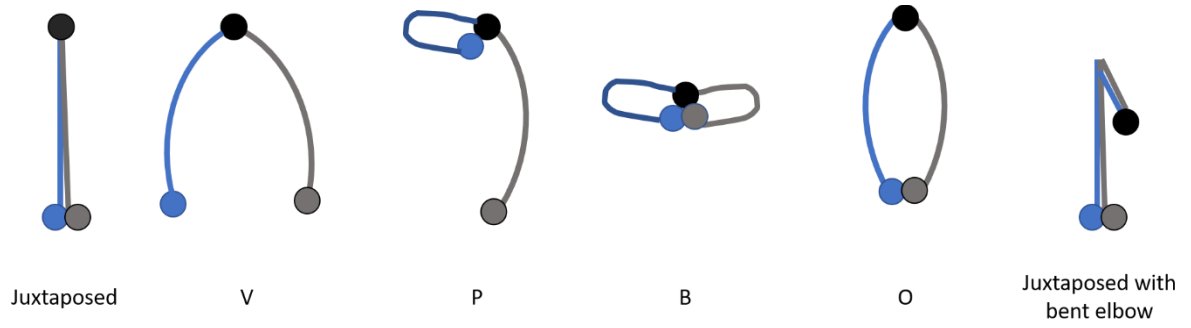


Figure 1.4. Cartoon representation of the potential conformations adopted by Smc dimers. Individual Smc coiled-coils and heads are represented in blue or grey, with their dimerised hinge domains depicted in black.

A proposed consequence of the coiled-coils' unique architecture is that they are able to form the walls of a large DNA binding pocket between the hinge and head domains (Ulhmann, 2016). Evidence of an ATP-dependant salt resistant mode of binding in purified condensin complexes strongly supports the idea of topological entrapment (Terakawa et al., 2017) and similarly, cohesin has been shown to bind circular DNA minichromosomes which can be released upon opening of either the DNA ring or proteolytic opening of the coiled-coils using an artificially introduced Tobacco Etch Virus (TEV) site (Ivanov et al., 2005, Gruber et al.,

2003). A similar observation is made with proteolytic cleavage of *S. cerevisiae* Smc4 coiled-coils resulting in strongly reduced binding to chromosomes (Cuylen et al., 2011).

Part of the topological gating mechanism may involve shifting the coiled-coils from a rod-like "closed" conformation to a more open, or O-like conformation, to generate space for DNA, possibly in response to an ATP-dependant twist of the head domains relative to the hinge (Nunez et al., 2019).

Crystal structures of both the Smc-ScpAB and condensin hinge and proximal coiled-coil domains suggest this transient opening may be coupled to the binding of DNA at the hinge (Soh et al., 2015). In the Smc-ScpAB complex, binding of DNA at the hinge is possible when the lumen of the coiled-coils is in an open conformation. Switching of DNA between the hinge site and a second Smc binding site formed by the engaged head domains occurs within the open lumen (Nunez et al., 2019). These binding sites were not detected in cohesin but this may be due to their transient nature (Chapard et al., 2019).

1.2.3 Hinge domain

The hinge domain can be found at the midpoint of the Smc protein primary structure and forms the main point of contact between dimers. This interface is structurally well conserved across SMC complexes and in different species. Crystal structures for murine condensin and cohesin hinge heterodimers (Kurze et al., 2011, Griesse et al 2010), in addition to homodimers of Smc proteins from *P. furiosus* (Griese and Hopfner, 2011) and *T. maritima* (Haering et al., 2002), have revealed a shallow U-shaped globular structure that, when dimerised, forms a toroid with its binding partner. This leaves a ~10 A charged hole in the middle of the hinge domain. This hole is positively charged due to the presence of several arginine and lysine residues along its surface (Kurze et al., 2011). The smaller hinge domain of *B. subtilis* MukB shares little sequence homology with other Smc members, but does produce a similar toroidal shape upon dimerization, albeit with an occluded central pore and little by way of charged residues (Li et al., 2010).

Much like the Smc head domains, the hinge can be divided into two subdomains (I and II) that interact with the second Smc protein to produce a pseudo 2-fold symmetrical structure.

The dimer is formed at two separate interfaces which are characterised in eukaryotes by the presence of a GX_6GX_3GG motif, and in bacteria by a GX_5GGX_3GG motif found in the last 2 β-strands of subdomain II of the hinge (Alt et al., 2017). This motif is found in most eukaryotic and bacterial structures solved to date, but not in the *S. pombe* SMC5/6 hinge structure where these glycines are either partially (Smc6) or wholly (Smc5) missing. This marked divergence in hinge architecture may be significant for the unusual functions of SMC5/6 in homolgous DNA repair, but why this might be is still unclear (Alt et al., 2017).

The hinge domain takes on special importance when found in species with multiple SMC complexes formed from Smc heterodimers. It appears that despite strong conservation, Smc proteins can distinguish their correct heteromeric binding partners based on the inclusion of several divergent amino acids in otherwise conserved motifs at the binding interface (Griesse et al., 2010). This makes the hinge domain an essential part of complex formation.

In addition to being the primary interface of the Smc proteins, several studies have also implicated the hinge domain in the process of entrapping DNA inside the coiled-coil pocket of cohesin. Closure of the *S. cerevisiae* cohesin heads by cross linking of the kleisin sub-unit Scc1 would seem to indicate the complex can still topologically load DNA, presumably through a gate at the hinge (Gruber et al., 2006). DNA found in the lumen of the Smc-kleisin tripartite ring requires that the hinge be sealed to prevent topological unloading of DNA under denaturing conditions (Haering et al., 2008). The process of topologically entrapping DNA through loading at the hinge apparently involves the charged central pore, with select mutations of key lysines resulting in a complex than can bind DNA and hydrolyse ATP but not topologically entrap DNA (Srininvasan et al., 2018). It is currently unclear whether this mechanism applies to all SMC complexes but the high degree of conservation in the central pore would suggest it does.

Several studies have reported a propensity to bind DNA at the hinge domain in multiple SMC complexes in addition to cohesin. In condensin this DNA binding is single stranded (Griesse et al., 2010) and induces a conformational change that partly leads to opening of the hinge toroid to allow ssDNA access, but not all the way to the coiled-coil lumen (Uchiyama et al., 2015). ssDNA binding is also preferred in the archaeal Smc of *Pyrococcus furiosus,* which shares a high degree of structural homology with its bacterial and eukaryotic homologs (Griesse and Hopfner, 2011). The same is also true for the SMC5/6 hinge which, despite its

structural divergence, displays the same ssDNA preference. Additionally, several studies in bacteria have also implicated the DNA binding ability of the hinge domain in stimulating ATPase activity (Hirano and Hirano, 2006)

1.3. The Kleisin Subunit

The kleisin subunit of SMC complexes is unique amongst the SMC subunits in that this class of proteins can be found interacting with all SMC proteins, including HAWKs and many of the KITES. This makes them central to the overall architecture of the complex, acting as a 3rd member of the core tripartite ring that characterises SMC complexes. Kleisins are largely disordered proteins, thought to bridge the opposing Smc heads, thus closing the lumen created by the coiled coils, even in the absence of ATP mediated dimerization at the Smc heads.

The eukaryotic kleisins can be subdivided based on sequence homology and their associated Smc proteins. Alpha kleisin assembles into cohesin, beta kleisin into condensin II and gamma kleisin into condensin I, though not all types of kleisins are present in all eukaryotes (e.g. beta kleisin is absent from yeast). Bacterial kleisins are regarded separately and assemble exclusively into bacterial SMC complexes (Schllieffer et al., 2003). ScpA is found in Smc-ScpAB, MukF in MukBEF and MksE in MksBEF.

Klesins N and C terminal domains show the most sequence homology, with the central linking domains between species being very diverse. Typically, the disordered linker region is shortest in the bacterial kleisins (~250 total residues) with gamma kleisins being the longest, having up to 1000 residues in total (Schllieffer et al., 2003). Alpha and beta kleisins display an intermediate length between 550 and 700 amino acids (Schllieffer et al., 2003).

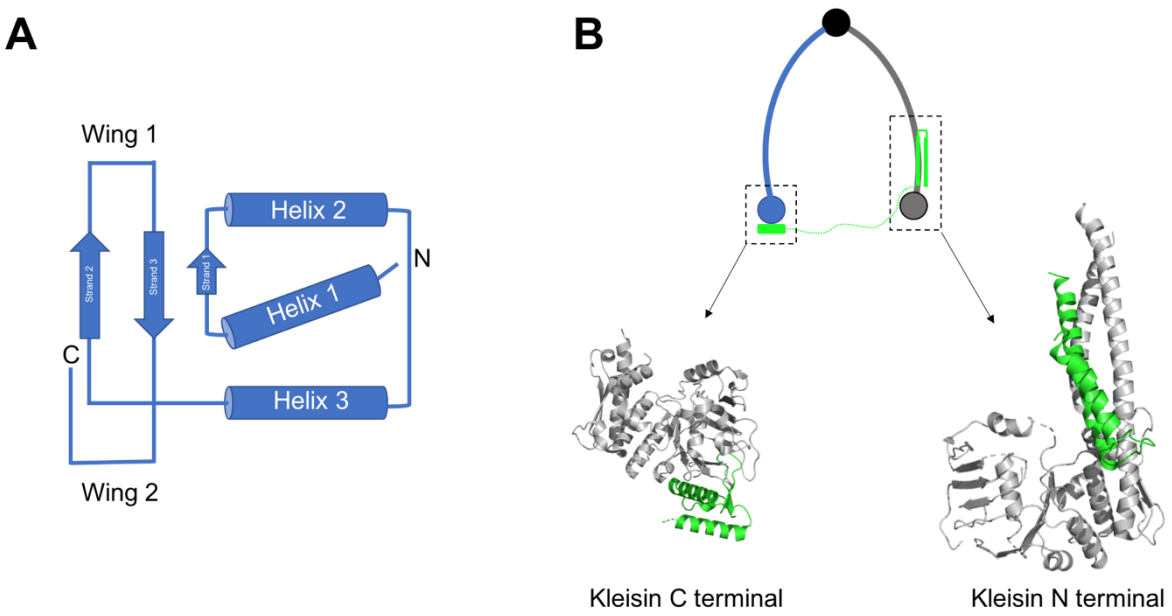


Figure 1.5. Kleisin interaction with Smc heads. (A) Cartoon schematic of a generic winged helix domain (Adapted from Gajiwala et al., 2000). (B) Structure of the Smc-ScpA N and C terminal interfaces. ScpA residue are shown in green, with the winged-helix C terminal domain bound to the underside cap of the Smc head depicted on the left (PDB 4i99, Shin et al., 2013). The N terminal ScpA bound to the head proximal Smc coiled-coils is depicted on the right (PDB 3zgx, Burmann et al., 2013). Both Smc heads and the head proximal coiled-coils are shown in grey (Burmann et al., 2013).

Principally, kleisins anchor themselves to the Smc proteins through two main asymmetrical interfaces. Firstly, the C terminal of kleisins folds in to a "winged helix" domain (WHD) (Haering et al., 2004) which is a type of helix-turn-helix consisting of 3 alpha helical segments and 2 or 3 beta strands linked by large loops that form the "wings" (Gajiwala et al., 2000) (figure 1.5 A). In cohesin, this domain interacts with the C terminal beta strands of Smc1 through hydrophobic residues to form a "cap" on the head (Haering et al., 2004). A similar arrangement is seen with the Smc-kleisin C terminal interfaces in *P. furiosus* (Smc-ScpA) (Burhman et al., 2013) and *E. coli* MukB-MukF (Woo et al., 2009). Generally, Smc proteins binding this C terminal WHD can be referred to as a κSmc and it is thought this mechanism is conserved throughout the SMC complexes. Binding of the kleisin has also been shown to stimulate the ATPase activity of the head domains (Arumugum et al., 2006, Zawadzaka et al 2018).

The N-terminal domains of kleisin subunits associate themselves with the head-proximal coiled-coils of Smc proteins (figure 1.5 B). The Smc-ScpA interphase for example, is formed

from an N terminal alpha helix combining with the coiled-coils of the Smc to form a 3-helix bundle (Burhman et al., 2013). A similar structure is seen in the *S. cerevisiae* cohesin complex, though in this instance it is a 4 helical bundle with 2 N terminal Scc1 (kleisin) alpha helices forming the interface with Smc1 (Gligoris et al., 2014). In both cases the Smc protein bound by the kleisin N terminal helices is referred to as νSmc. As with the C terminal Smc-kleisin architecture, this mode of binding is largely preserved in structures observed thus far, with the exception of the bacterial MukBEF complex, in which an N terminal helical bundle not only interacts with the head proximal coiled coil of MukB (νSmc), but also mediates dimerization with a second complex (Zawadzka et al., 2018). The heterodimeric nature of eukaryotic Smc protein interactions ensures that the central tripartite ring is asymmetrical in nature. A similar asymmetric interaction is seen in the bacterial Smc-ScpAB despite homodimerization of the Smc protein. Interestingly this asymmetry is apparently essential for the function of the complexes (Burhmann et al 2013).

There is also strong evidence that the kleisin subunit forms an alternate topological binding pocket between itself and the engaged heads of the Smc proteins in both the Smc-ScpAB complex and cohesin (Valasquez-Nunez et al., 2019, Chapard et al., 2019) (figure 6). This kleisin pocket is potentially opened to the coiled-coil lumen upon disengagement of the Smc heads following ATP hydrolysis, and has been shown to continue entrapping DNA even when the coiled-coils are in their juxtaposed position (Chapard et al., 2019).

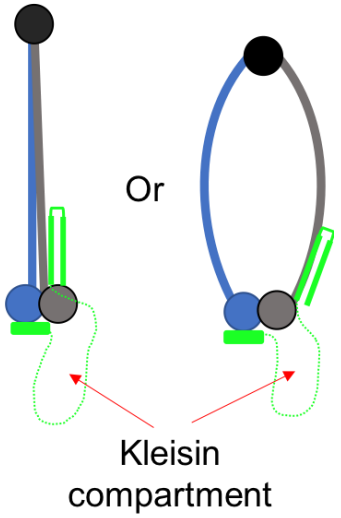


Figure 1.6. Cartoon representation of the kleisin compartment formed by juxtaposition of the Smc protein coiled-coils or engagement of the Smc protein heads. Kleisin binding subunits not shown.

Another DNA binding domain formed by the kleisin in cooperation with one of the HEAT repeat subunits has also recently been described. Structures of the *C. thermophilum* Ycg1 subunit, co-crystalised with a central fragment of the kleisin Brn1 revealed DNA bound along the inner surface of the Ycg1 solenoid with Brn1 topologically encircling it (Kschonsak et al., 2017). These interactions are potentially conserved in condensin I (Hara et al., 2019) and cohesin (Li et al., 2018) (further discussed in section 4.2).

In addition to the afformentioned structural and mechanistic contributions to SMC complex function, the cohesin kleisin subunit also possess novel functionality, essential to its specific role in sister chromatid cohesion. Dynamic opening of the kleisin-Smc interface is an important feature of cohesin in both interphase and mitosis/meiosis, with stabilisation of this action forming a key part of its cell cycle regulation (Rowland et al., 2009, Unal et al., 2008). Similarly, proteolytic cleavage of the cohesin kleisin is another means of regulation crucial to its unique role in the cell cycle (Uhlmann et al., 2000). Such is the importance of the kleisin as a target for cohesin regulation that specific variants exist for mitosis and meiosis (Parisi et al., 1999).

1.4. KITEs and HAWKs

In addition to the core Smc-kleisin trimer, additional non-Smc subunits have been described that associate with SMC complexes. These were recently recategorized based on their evolutionary divergence into two groups: KITEs (Palecek and Gruber, 2015) and HAWKs (Wells et al., 2017).

KITE proteins are found in bacteria and the eukaryotic SMC5/6 complex and are thought to be the more ancient of the two groups. At some point these proteins were displaced in the last common eukaryotic ancestor by the HEAT repeat proteins, giving rise to the modern eukaryotic condensin and cohesin complexes (Wells et al., 2017). It is not known whether the SMC5/6 complex gained its HAWKs and subsequently lost them or if it diverged from condensin and cohesin prior to this exchange (Wells et al., 2017).

1.4.1. KITES

Structurally, KITEs are composed of tandem winged helix domains, similar to those found in the C terminal of the kleisin subunit, that can come together in a head-to-head arrangement to form dimers. Interactions between two N terminal WHDs are mediated by hydrophobic residues within the first alpha helix of one WHD, and two loops between the last alpha helix/beta sheet and the last "wing" of the other WHD (Paleck and Gruber., 2015, Burhman et al., 2013). Such interactions are further reinforced through WHD mediated binding with the kleisin subunit. Here, the C terminal WHD aligns its first alpha helix with another in the central region of the kleisin, with additional non-helical bonds also contacting the dimerised N terminal WHDs (Rajeskar et al., 2019, Burhman et al 2013, Doyle et al., 2010, Newman et al., 2016).

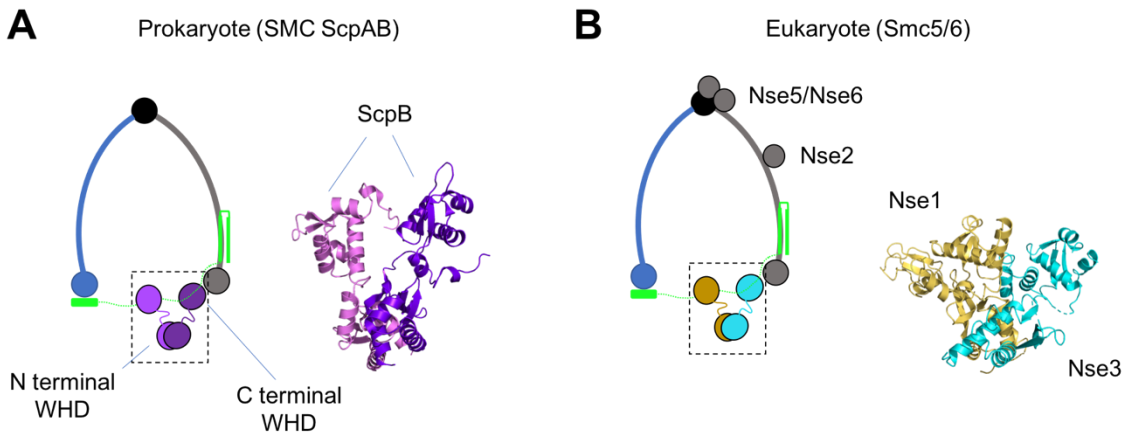


Figure 1.7. KITE-Kleisin interaction (A) Cartoon schematics of bacteral SMC complexes with KITE subunits. A bacterial homodimer of the KITE ScpB from Geobacillus stearothermophilus (PDB 2z99, Kamada et al., 2013) is depicted in purple. (B) Cartoon schematics of eukaryotic SMC complexes with KITE subunits. The eukaryotic heteromeric dimer of Nse1 (yellow) and Nse3, also called MAGE1 (Melanoma Antigen Gene 1) (cyan) is depicted on the right (PDB 5wy5, Doyle et al., 2010).

1.4.1 i Bacterial KITEs

It is still unclear exactly what the role of the KITE proteins is, and if their functions are conserved between species. Both bacterial ScpB and MukE have been shown to be important for the proper assembly of their respective complexes and deletion of these subunits leads to

segregation defects in dividing cells, similar to deletions of the Smc proteins themselves (Soppa et al., 2002, Kamada et al., 2013 ref).

ScpB appears to be important for inducing conformational changes that allow ScpA to upregulate the ATPase activity of Smc in *B. subtilis* (Kamada et al., 2013). It is thought that conformational changes are brought about by the release of a portion of the kleisin ScpA by one ScpB upon ATP binding to the Smc heads. This in turn allows binding of a formerly concealed ScpA helix to the Smc head whereupon the interaction facilitates a higher turnover of ATP (Kamada et al., 2013). This mechanism may be conserved in the MukBEF complex (Woo et al., 2009, Kamada et al., 2013).

The bacterial MukBEF complex has recently been shown to have some additional, novel structural features mediated by its MukE subunit, that do not appear to be common to other KITE containing complexes. Not only was MukE structurally important for the interaction of MukF and MukB but also it has been shown to have a homodimeric interface in addition to that found between MukE's of the same complex (Rajeskar et al., 2019). This new interface appears to be mediating inter-complex interactions. It is unclear why inter-MukE interactions are important, but mutations in key interfaces lead to temperature sensitivity without affecting MukE binding to MukF and MukB (Rajeskar et al., 2019). A possible explanation is that the binding of MukE prevents MukF from binding the same MukB at the head and neck interfaces thus ensuring formation of a complete tripartite Smc-kleisin core. It is also possible that the presence of MukE compacts the complex such that the MukB heads more readily engage (Rajeskar et al., 2019).

1.4.1 ii Eukaryotic KITEs

Like the bacterial KITEs, the eukaryotic SMC5/6 KITES themselves also appear to be essential to the complex (Pebernard et al., 2004). As with the bacterial complexes, two KITEs (Nse1 and Nse3) bind the kleisin subunit (Nse4) via their tandem C terminal WHD's, though in this case the KITES form heterodimers at their N terminal WHDs (Guerineau et al., 2012, Hudson et al., 2011). Unlike its bacterial cousins though, the Nse1 subunit, in addition to its tandem WHD, contains a RING (Really Interesting New Gene) domain which bestows ubiquitin ligase activity. Mutation of this domain leads to a DNA damage sensitive phenotype

in yeast (Pebernard et al., 2008), which suggests a novel gain of function in addition to that shared with its bacterial homologs. Likewise, Nse3 is novel in that it seems to play an important role in binding DNA via the Nse1-Nse3-Nse4 sub complex (Zabrady et al., 2016). The mechanical significance of this binding is unknown but in fission yeast its disruption is lethal (Zabrady et al., 2016).

Despite a lack of structural homology between the KITEs and HAWKs, the observation that Nse1-Nse3-Nse4 sub-complex can mediate DNA binding in a similar way to the Ycg1-Ycs4-Brn1 subcomplex in condensin, suggests perhaps a common need for multiple DNA binding sites within the SMC complexes being met by unrelated subunits (Piazza et al., 2014, Kschonsak et al., 2017). Note, while SMC5/6 does also contain 3 additional peripheral subunits [Nse5, Nse6 and the SUMO (Small Ubiquitin-like Modifier) ligase Nse2], their role appears at this time to be unique to the SMC5/6 complex and thus will not be discussed further.

1.4.2 HAWKs

The HAWKs (HEAT Proteins Associated with Kleisins), encompass a family of proteins found in eukaryotes (Neuwald et al., 2000). These proteins, as their name suggests, contain HEAT (huntingtin, elongation factor 3, protein phosphatase 2A, Tor1 kinase) repeats which are comprised of two amphipathic alpha helices joined by a short loop (Grinthal et al., 2010). Stacking of these motifs in the HAWKs produces a solenoid type structure that bends into a “harp” or U shape (Kschonsak et al., 2017, Kikuchi et al., 2016, Lee et al., 2016, Hara et al., 2014). This structure has been demonstrated to lend a significant amount of elasticity to the HEAT repeat proteins, allowing them to act as molecular springs, though what role this flexibility might play is unknown (Grinthal et al., 2010, Kappel et al., 2010, Manalastas-Cantos et al., 2019).

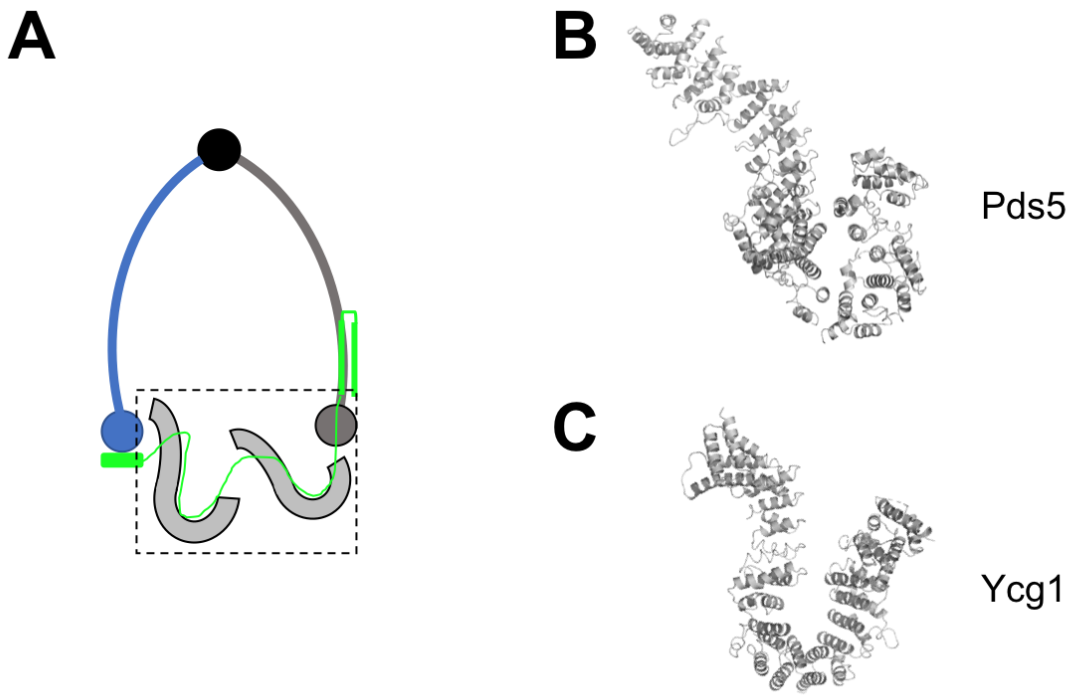


Figure 1.8. HAWK-kleisin interaction. (A) Cartoon schematic of the generic interaction of HAWKs with the kleisin subunit of a eukaryotic SMC complex. (B) Structure of the S. cerevisiae cohesin subunit Pds5 (PDB 5f0n, Lee et al., 2016) (C) Structure of the S. cerevisiae condensin subunit Ycg1 (PDB 50qq, Kschonsak et al., 2017). Note that these two proteins do not appear together in a single complex.

1.4.2 i HAWKs in Condensin

The yeast condensin complex contains two types of HAWK; Ycg1 and Ycs4. In higher eukaryotes two homologs of these proteins exist (CAP G/CAP G2 and CAP D2/CAP D3 respectively), which in addition to either beta or gamma kleisins define the condensin I and II complexes. As with the KITEs, the HAWKs association with their respective complexes is mediated through interaction with the kleisin subunit. Unlike the KITEs however, there is very little evidence of direct interaction between different HAWK subunits (Piazza et al., 2014; Kimura et al., 2000).

Once again, as with the KITEs, the role of the HAWKs is poorly understood, but current data points towards a role in non-topological DNA binding, possibly as a first step in complex loading. *In vitro* studies of the non-Smc sub-complex, consisting of Ycs4, Ycg1 and Brn1 in *S. cerevisiae* showed that it binds DNA alone, with the presence of both HEAT repeats greatly increasing the affinity over single HAWK sub-complexes (Piazza et al., 2014).

Interestingly, analysis of the *X. laevis* non-Smc sub-complex shows a slightly different picture, in that binding of DNA by the non-Smc sub-complex is extremely weak and easily disrupted by low ionic strength buffers (100 mM KCl) (Kimura et al., 2000).

The importance of these HEAT-kleisin-DNA interactions is demonstrated by *in vivo* mutation analysis. For example, mutations in either of condensins' HEAT repeat subunit strongly reduces the amount of condensin associated with chromosomes (Lavoie et al., 2002; Bhalla et al., 2002; Savvidou et al., 2005). A recent study in *X. laevis* used a mutant condensin complex and wild type (WT) depleted cell extracts to dissect the individual contribution of the HAWKs to chromosome condensation, showing that loss of CAP G (Ycg1) in condensin I results in axial loading of condensin and a highly compacted central axis. Loss of CAP D2 however, allowed condensin to bind but not to form axes which suggests that the mechanistic contributions of the CAP G subunit is driving axis formation while CAP D2 opposes it, likely striking a balance in the wild type (Kinoshita et al., 2015).

A recent structure of the *Chaetomium thermophilum* Ycg1 has uncovered perhaps the most convincing evidence of a key mechanistic role for the non-Smc, non-kleisin subunits to date. Here it was shown that in conjunction with Brn1 (kleisin), Ycg1 forms an essential binding pocket for dsDNA, in which positively charged residues form a groove that can contact the phosphate backbone in a sequence independent manner. Brn1 stabilises this arrangement by lining the inner surface of Ycg1, closing the biding pocket and encircling DNA via a "buckle" and "latch" interaction between its alpha1 and alpha4 helices (Kschonsak et al., 2017). A similar structure is also seen in the human condensin I wherein the homologs of Ycg1 and Brn1 (hCAP G and hCAP H respectively) form an analogous DNA binding site (Hara et al., 2019).

The second condensin HAWK Ycs4 has also recently been described in detail, being implicated in a regulatory role within the ATPase cycle. Not only is Ycs4 bound to the kleisin subunit, but it also concurrently interacts with the head domain of Smc4 via a motif known as the W loop (Hassler et al., 2019). Release of Ycs4 from the W loop upon ATP binding at the Smc4 head allows access to the Smc2 head by Smc4 and subsequent binding of a second ATP. This conformational shift also corresponds to the release of the N terminal portion of Brn1, implicating Ycs4 in the mechanism for opening the tripartite ring of condensin (Hassler et al., 2019). In earlier Cross-linking mass spectrometry (CLMS) experiments involving the

chicken DT40 condensin complex, no evidence of Smc-Ycs4 interaction was found, which may suggest that either the W loop release mechanism is not wholly conserved, or the interaction is transient in different species (Barysz et al., 2015).

1.4.2 ii HAWKs in cohesin

The cohesin HEAT repeat protein Scc3 also provides a positively charged groove for DNA interaction. In this case it is not known if the kleisin Scc1 contributes a stabilising loop as is seen with Ycg1-Brn1 (Li et al., 2018), though it has been demonstrated that disruption of the key residues mediating interaction between Scc1 and Scc3 prevent proper cohesin loading on DNA (Hu et al., 2011). It is unclear what this apparently conserved binding domain is required for, though ablative mutation of key DNA binding residues in Scc3 lead to loss of function (Li et al 2018), as is the case with Ycg1 in condensin (Kschonsak et al, 2017). Interestingly, loss of the DNA binding domain in condensin is much less well tolerated than in cohesin, suggesting that perhaps part of its function is in the loading of the complex onto DNA. This function may be partially compensated for by cohesin loader complexes (Ciosk et al., 2000), thus explaining the discrepancy (Kschonsak et al., 2017, Li et al., 2018).

Like condensin, cohesin also possesses another stably associated HAWK called Pds5. Like Scc3, Pds5 binds the cohesin complex through association with the kleisin Scc1, which crystal structures show lines the inner concave surface of Pds5 (Lee et al., 2016). Mutation of key residues mediating the Pds5-Scc1 interaction lead to lethality *in vivo*, with the association of Pds5 being strongly implicated in regulation of cohesin stability on DNA during the cell cycle (Chan et al., 2013). There is as yet no evidence showing that Pds5 is capable of forming a DNA binding domain like Scc3 or Ycg1.

An additional HAWK can also be found associated with cohesin in a more dynamic mode. The Scc2 subunit of the cohesin loader complex (Scc2-Scc4) also has an extensive C terminal HEAT repeat domain that mediates interaction with the cohesin kleisin Scc1 (Kikuchi et al., 2016). Crucially this Scc1 interacting domain overlaps with the biding domain of Pds5. This means, in contrast to the apparently constitutive association of the condensin HAWKs, Pds5 and the Scc2-Scc4 loader associate with Scc1 reciprocally.

1.4.2 iii Regulation of Cohesin Through HAWKs

In addition to its role in DNA binding the Scc3 HAWK also forms a key platform for the cell cycle specific regulation of cohesin. In particular, the cohesion regulators Wapl and SgoI have been shown to antagonistically compete for binding to the Scc3 human homolog SA2 within cohesins located on the mitotic centromere.

Not only does SgoI stabilise cohesin DNA binding by preventing the Wapl C-terminal from binding Scc3, but it also recruits another regulator, PP2A which in turn stabilises the binding of sororin to the HAWK Pds5 (Shintomi et al., 2009). Sororin antagonises the N terminal binding of Wapl (Hara et al., 2014).

Though sororin is notably missing in yeast species, regulation of cohesin stability on DNA through the Pds5-Wapl interaction is an important and conserved HAWK-based regulator. The Wpl1–Pds5 complex is thought to promote DNA release through an exit gate at the Smc3–Scc1 interface prior to S phase. Not only does this complex compete for the binding of Scc1 at the Scc2-Scc4 loader binding site, but it has also been suggested that binding of Wapl -Pds5 holds the DNA gate open by stiffening Scc1 between the two Smc head domains (Makrantoni et al., 2018). Paradoxically however, Pds5 association of cohesin apparently promotes acetylation of Smc3 by Eco1 which itself is important for stabilising cohesin on DNA, in apparent contrast to the role that Pds5 has in Wapl mediated dissociation (Chan et al., 2013). Pds5 then, provides a platform from which loading and unloading of the complex is controlled in fine detail.

1.5. Current Models of SMC Complex Function

While the individual contributions of the SMC complexes to the structure and regulation of the genome differs widely between different complexes, the structural similarities present across the family point towards a potentially common mode of action. The formation of loops within the genome is the primary consequence of this action and is thought to underpin a wide range of SMC phenomena (condensation, cohesion, bacterial genome segregation etc.). The importance of the SMC complexes contribution to genome looping and organisation is demonstrated by *in vivo* studies (Rao et al., 2017, Wutz et al., 2017, Gibcus et al., 2018,

Wang et al., 2017), but how the SMC complexes carry out their functions and the specific protein-DNA and protein-protein interactions that facilitate them are not fully understood.

1.5.1 DNA Looping – Passive vs Active Loop Formation

Many models of DNA loop formation and enlargement have been proposed previously, from the simple act of stochastic cross-linking stabilising interactions between distal DNA elements (Cheng et al., 2015), to active formation of DNA loops by the SMC complexes themselves (Nasmyth, 2001). Current theoretical and experiment evidence seems to point towards the SMC complexes mediating a form of loop capture and progressive enlargement (Fudenberg et al., 2016, Goloborodko et al., 2016), but the mechanism underpinning this mode of "loop extrusion" is unclear.

Several models of passive extrusion have been put forward previously. Based on evidence of cohesin accumulation at convergent transcription sites (Lengronne et a., 2004, Ocampo-Hafalla et al., 2016) and its ability to be translocated by bacterial polymerases *in vitro* (Davidson et al., 2016) it has been suggested that the driving force behind loop extrusion could be external to the SMC complexes which simply would be acting to hold DNA while the transcription machinery pushes the complexes towards the base of loops. This model is not supported by apparent formation of cohesin dependant genomic loops where transcription is repressed (Hug et al., 2017, Flyamer et al., 2017), nor does it track with condensin based looping in mitosis in which transcription is also down regulated. It is therefore most likely that the loop extrusion phenomenon requires an intrinsic motor or DNA pumping mechanism within the structure of the SMC complexes.

1.5.2 Molecular Mechanisms Underpinning the SMC Complex Motor

1.5.2 i Evidence for an ATP Driven Mechanism

Evidence for the active formation of loops by purified SMC complexes has recently been demonstrated *in vitro*, with direct observation of condensin and cohesin dependant looped

structures forming on single molecule DNA substrates (Ganji et al., 2018, Davidson et al., 2019, Kim et al., 2019). These key studies would seem to eliminate the need for additional factors to drive the formation and enlargement of DNA loops, meaning the SMC complexes must be capable of active loop extrusion in an ATP-dependant fashion.

This intrinsic ATP dependence is supported by i*n vivo* evidence from the bacteria *B. subtilus* which demonstrates that ATP activity is required for the Smc-ScpAB mediated juxtaposition of chromosome arms (Wang et al., 2018, Wang et al., 2017). Time-resolved HiC experiments measuring the rate of this juxtaposition clearly show that reduction of ATPase activity of Smc-ScpAB correlates with a reduction of juxtaposition rate (Wang et al., 2018).

Furthermore, mutation of ATPase domains in bacterial Smc-ScpAB leads to accumulation of the complex at its loading site, without translocation along juxtaposing chromosome arms (Minnen et al., 2016). This evidence suggests that the mechanism of translocation (i.e. loop extrusion) requires ATP but loading of the SMC complex on to DNA does not. Similarly, ATPase mutants of condensin and cohesin are capable of loading DNA but not extruding loops (Thadani et al., 2018, Eeftens et al., 2017).

Despite the apparent need for ATP hydrolysis in loop extrusion, the structure of the SMC complexes does not immediately suggest how this hydrolysis could be converted into active restructuring of DNA. The most likely explanation is ATP hydrolysis powers a form of conformational change within the SMC complex to drive pumping of DNA (Nasmyth, 2001, Ganji et al., 2018, Davidson et al., 2019).

1.5.2 ii DNA Binding Sites and Conformational Changes

At a molecular level there are several proposed mechanisms for the SMC complex motor. ATP induced movement within the complex is often named as the “power-stroke”, but the precise conformation change or changes required to accomplish this power-stroke are still speculative.

It is possible that ATP induced conformational changes are translated through the coiled-coils. It has recently been demonstrated that engagement of the two ATPase sites in the yeast

condensin rotates them 15 – 25 degrees, consequently rotating the head-proximal coiled coils (Hassler et al., 2019). Coupling of this rotation to bending of the coiled coil elbow from an extended conformation to a folded one could be providing the power-stroke that feeds DNA through the lumen of the coiled-coils in a levering motion (Burman et al., 2019) (figure 1.9 A), or through a second compartment formed between the kleisin and engaged Smc heads (Vazquez-Nunez et al., 2019, Chapard et al., 2019). Binding of DNA to the hinge domain could provide a point of contact for the lever. In opposition to this model is the observation that mutations of key residues in the charge central pore of the hinge region do not preclude translocation of condensin, which likely eliminates the possibility that this domain is a key binding site attaching DNA to the moving lever (Vazquez-Nunez et al., 2019).

Anchoring of DNA at a second binding site within the SMC complex could be providing a stable platform from which the base of a loop is held. Candidates for these anchor points include the previously described binding domains formed by the HEAT repeat subunit Ycg1 with the kleisin Brn1 (Kschonsak et al., 2017), though the entire non-Smc sub-complex working cooperatively is also a candidate (Piazza et al., 2014). SMC5/6 and its Nse1/Nse4/Nse3 sub-complex could be fulfilling a similar role with Nse3 known to mediate DNA binding (Zabrady et al., 2016), as could the Scc3 binding domain in cohesin (Li et al., 2018). It is however unclear if an equivalent domain is formed in the bacterial SMC complexes (Fennel-Fezzie et al., 2005).

Another use of the coiled-coil bend might be to force the DNA through the lumen between the coiled-coils by effectively zipping up the space (figure 9 B). If a loop of DNA is in the lumen when the coils juxtapose and the hinge bends towards the heads, the loop would be forced through the shrinking lumen (Burmann et al., 2019). This model would also require the anchoring of DNA at the kleisin-HEAT domain but not at the hinge, which may be more compatible with available data showing that condensin translocation is still possible in hinge mutants (Vazquez-Nunez et al., 2019). Movement of DNA into the recently described kleisin pocket as the ATPase heads disengage following hydrolysis would allow the complex to reset while still topologically embracing the DNA. This motion could then be followed by re-engagement of the heads, opening of the coiled-coil lumen and influx of a fresh length of DNA for looping (figure 1.9 B).

Concomitant with the engagement of the Smc heads is the dissociation of the N terminal region of the kleisin subunit as seen in both condensin (Hassler et al., 2019) and cohesin (Muir et al., 2020). While this may seem to preclude the kleisin compartment as a reservoir of looped DNA following the power-stroke, it is possible that engagement of another part of the kleisin to the Smc heads, either directly or through one of the HEAT repeat subunits would close the compartment upon ATP induced head dimerization (Hassler et al., 2019).

Alternatively, the spring-like structure of the HEAT repeats has also been suggested as part of an “inch worm” type movement in which splitting of the Smc dimer into an open V shape would stretch the HEAT repeats along the DNA, followed by loop formation as the head domains re-engaged, reeling the DNA in with them (Nicols and Corces, 2018) (figure 1.9 C). This model considers the elastic properties of the HEAT repeat subunits (Yoshimura and Hirnao, 2016), though whether this is possible in KITE containing proteins is still unclear, particularly given the shorter length of the bacterial kleisins (Schliffen et al., 2003). While atomic force microscopy (AFM) experiments on Smc dimers suggest they could adopt an open V conformation, CLMS and structural studies seem to indicate that when heads are disengaged the coiled-coils are juxtaposed, not walking (Burmann et al., 2019). Furthermore, binding of the kleisin Brn1 to Ycg1 has been demonstrated to increase the rigidity of the HEAT repeat protein (Manalastas-Cantos et al., 2019).

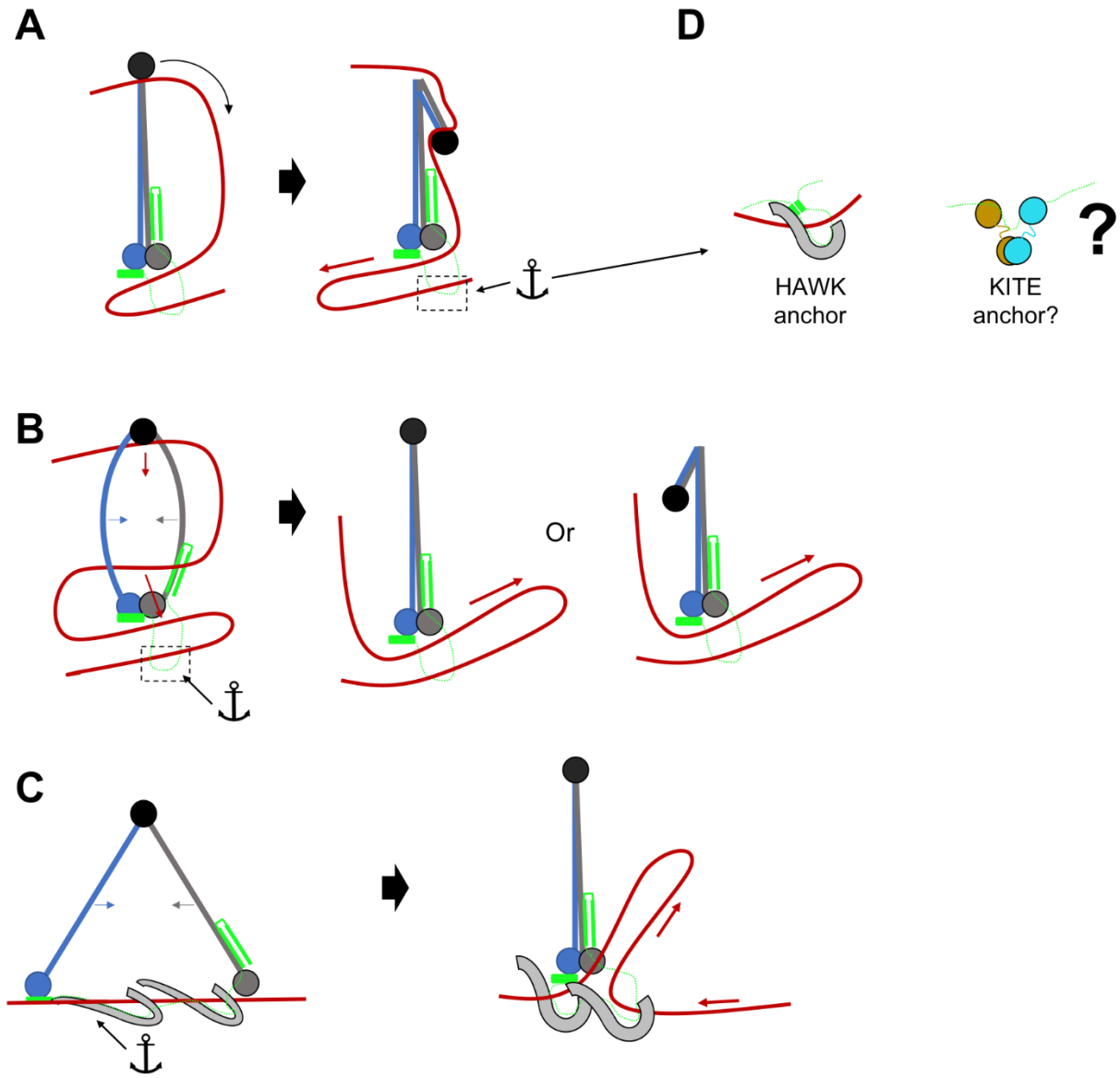


Figure 1.9. Models of loop extrusion. (A) Model of extrusion in which bending of the coiled-coil hinge provides the power stroke, feeding DNA through the kleisin pocket. (B) Model of extrusion in which juxtaposition of the coiled coils forces extrusion of DNA through the kleisin pocket. (C) Model of extrusion in which separation of the Smc ATPase head domains along DNA is followed by juxtaposition of the ATPase head domains with DNA being reeled in by the juxtaposition. (D) Cartoon schematic of the putative kleisin-HEAT (HAWK) or kleisin-KITE anchor sites. SMC complex subunits are depicted as in previous figures, with Smc protins in blue or grey, kleisins in green and HAWKs in light grey. KITEs are shown in orange and blue. DNA is shown in red and putative DNA anchor sites are highlighted by anchor symbols.

1.5.3 Applying the SMC Pump Mechanics to Asymmetric or Symmetric Loops

Asymmetric loop extrusion is well supported in condensin by both its unidirectional translocation on immobilised DNA substrates (Terakawa et al., 2017) and direct observations of fluorescently labelled DNA structures *in vitro (*Ganji et al., 2018). This presents a problem however, in that asymmetric loop extrusion is in theory insufficient to mediate the degree of compaction seen *in vivo.* This is because extruding in opposite directions will leave un-looped gaps between them (Banigan and Mirny, 2019). A possible explanation for this discrepancy may have been found in the observation that condensin complexes can traverse one another while extruding asymmetrically, forming "Z-loops" that would allow loop extrusion to progress into areas of DNA between adjacent complexes extruding in opposite directions (Kim et al., 2020).

However, recent observations of loop extrusion in the human cohesin complex differ somewhat from those of condensin in that loops appear to be extruded symmetrically. It is theoretically possible that an asymmetrical mechanism could produce symmetrical loops through cooperative action of two complexes, but strikingly both condensin and cohesin appear capable of looping DNA alone (Ganji et al., 2018, Davidson et al., 2019). Alternatively, a single asymmetrically extruding complex could produce a symmetrical loop by switching between the two DNA strands of the loop, which is supported in theory by molecular dynamic simulation data (Hassler et al., 2018, Banigan et al., 2019). Semi-diffusive extrusion, i.e. a model in which the fixed anchor point is allowed to slide somewhat, thus allowing the non-actively extruded strand of the loop to be incorporated by passive diffusion, has also been suggested. The seat-belt mechanisms observed in the Ycg1-Brn1 DNA binding pocket might suggest a mechanism through which this could be controlled, but this model is poorly supported by simulations (Banigan et al., 2019).

In bacteria, juxtaposition of chromosome arms *in vivo* by SMC complexes also appears to be symmetrical, beginning at *parS* sites (Wang et al., 2017; Tran et al., 2017). While the overall formation of loops is symmetrical, again it is possible that this is mediated by complexes acting cooperatively using an asymmetric mechanism.

1.6. Condensin chromatin dynamics

From the detailed study of the whole SMC family a picture is beginning to emerge of their shared molecular mechanisms. Despite the apparent structural similarity between different

SMC complexes, the regulation and function of each complex differs greatly. The following sections will therefore focus mainly on the dynamic regulation of condensin for simplicity as it is our main object of study.

1.6.1 Organisation of the genome

In eukaryotes there are several layered modes of organising genomes. These organisational features are not static, but instead change throughout the cell cycle to meet different structural and regulatory needs. Further convoluting the picture is the fact that the genome is never homogenous, and the accessibility and concentration of local genomic factors changes throughout time and space. It is in this complex environment that the condensin complex carries out its function, and with the changing of its molecular environment, so too does condensin's activity and specificity behaviour change.

Typically, each chromosome within the nucleus contains a centromere upon which the kinetochore assembles for mitotic spindle attachment during cell division. Interactions between sister chromatids prior to separation at anaphase are also maintained here and these regions are defined by sequence specificity (point centromere) (Henikoff and Henikoff, 2012) or within a region of DNA by epigenetic marks (regional centromere) (Pluta et al., 1995). The exact position of the centromere can vary between different chromosomes and in some species may be found in multiple places along the chromosome arms, forming what is known as holocentric chromosomes, notably in *C. elegans* (Melters et al., 2012). The regions of DNA constituting the rest of the chromosome in species containing monocentric chromosomes can be referred to as p (shorter) or q (longer) arms. DNA immediately adjacent to the centromere, typically consisting of repetitive, transcriptionally inactive DNA (regional centromeres only) is known as the pericentromeric DNA. In addition to the centromeres, chromosomes also contain telomeres at the ends of their arms that are comprised of short G-rich repeats (Meyne et al., 1989). These regions cap the ends of the chromosomes, providing a disposable buffer of untranscribed DNA that is progressively lost during incomplete DNA replication (Witzany, 2008).

Relative to the size of the cell, the DNA is very large. For example, the 46 chromosomes of the human genome consisting of some 6.4 billion base pairs, laid end to end would measure 2

meters in length. This material must fit into a nucleus measuring up to 10 μm in diameter, necessitating substantial packaging and organisation. The first level of this organisation is called the "10 nm fibre" and consists of arrays of nucleosomes comprising ~146 bp of DNA wrapped 1.6 times around octamers of histones (White et al., 2001). Histones are small ~10-16 kDa proteins. There are 4 core histone members; H2A, H2B, H3 and H4. H2A and H2B form dimers, as do H3 and H4, which then assemble into the higher order globular octamer (Luger et al., 1999) (figure 1.10 A).

In addition to the extensive packaging effects of wrapping DNA into nucleosomes, the disordered N and C terminal tails of these histones provide an important target for epigenetic regulators to carry out chemical modifications that are then recognised by other proteins interacting with the genome (figure 1.10 B). These signals contribute to the formation of further ordered domains known as compartments. These compartments consist of euchromatin (actively transcribed regions) and heterochromatin (non-transcribed regions, both constitutive and facultative) (figure 1.10 C). More specialised histone variants such as CenH3 or H2A.Z can also be found organising DNA in the centromere and in embryonic cells respectively. A final histone type, H1, is found to be associated with heterochromatin, and its effect is to promote further levels of compaction.

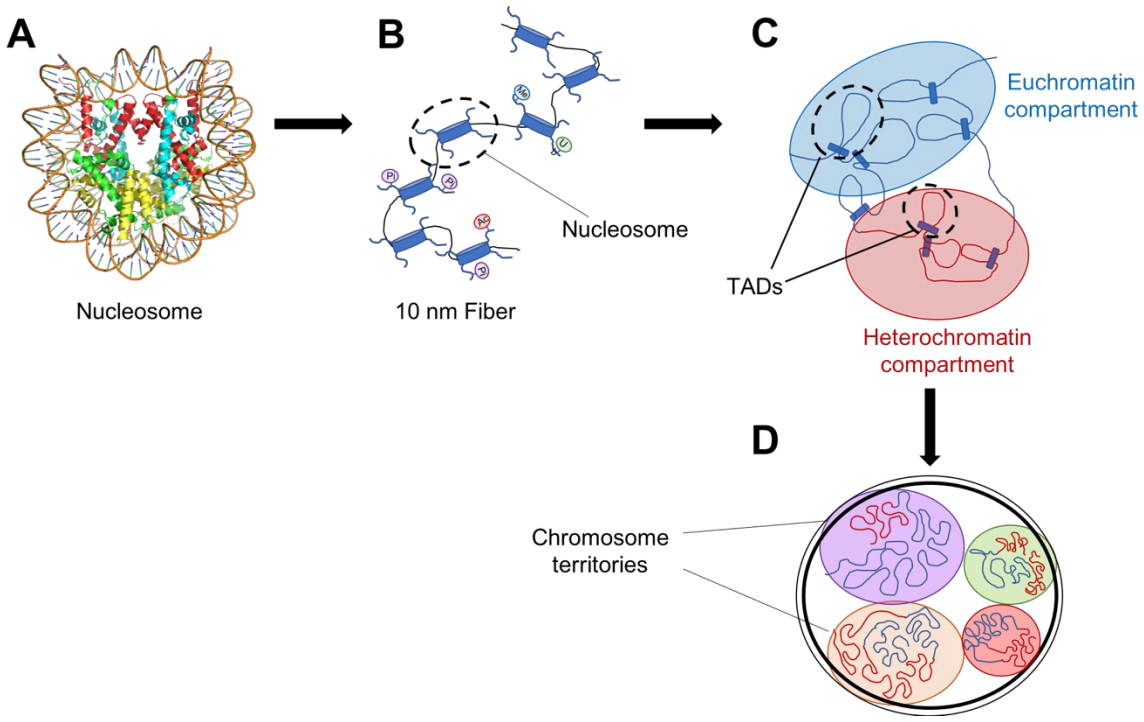


Figure 1.10. Cartoon summary of the major structures found in the interphase genome. (A) Crystal structure of the S. cerevisiae nucleosome (PDB 1id3, White etal., 2001). Histone H2A is shown in green, H2B in yellow H3 in red and H4 in cyan. Histone tails are not shown. (B) Cartoon schematic of the 10 nm fibre. Example, post-translational modifications are depicted on the tails of the histones. Pi is phosphorylation, Me is methylation, Ub is ubiquitination and Ac is acetylation. (C) Cartoon representation of chromatin compartments and the topologically associated domains within them. Euchromatin is depicted in blue and heterochromatin in red. Cohesin is depicted at the base of chromatin loops as a blue box. (D) Cartoon representation of chromatin territories within a nucleus. Individual chromosome territories are marked by coloured circles, with euchromatin shown in blue and heterochromatin shown in red.

Within the sequence of the DNA there are a multitude of regulatory units that modulate transcription of the genome (promoters, transcription start sites, terminators, enhancer, repressors, etc). Sequence specificity plays a role in defining these regions, as does the presence of epigenetic markers such as histone modifications and the 3-dimensional organisation of the DNA within the nucleus. In addition to the presence of the transcription machinery throughout the genome, there is also the transient action of the DNA replication machinery, DNA damage repair mechanisms and local folding of regions into topologically associated domains (TADs) (Szabo et al., 2019) (figure 1.10 C). Cohesin drives the folding of these TADs, which helps to guide distant regulatory elements to their target genes during interphase. During cell division, these topologically associated domains are lost in favour of large-scale compaction of the genome by condensin.

Finally, on top of all the aforementioned processes that shape the genome, interphase chromatin is further ordered by segregation into chromosome territories in which the DNA of a single chromosome preferentially interacts with itself (Cremer and Cremer, 2010) (figure 1.10 D). Within these territories, heterochromatin is usually oriented towards the periphery of the nucleus (Holla et al., 2020).

1.6.2 Cell cycle stages

The cell cycle can be divided into several morphologically distinct stages that are triggered through a cascade of protein-protein and protein-DNA interactions. Non-dividing, quiescent cells are said to be in G0 phase, while those that are preparing to undergo mitosis can be in periods of cell growth (G1 and G2 phases) before and after the DNA replication stage (S phase). Together these phases make up a period known as interphase. Mitosis can be divided further into 5 stages, beginning with prophase in which the DNA begins to condense. This is followed by prometaphase wherein the nuclear envelope breaks down, metaphase where condensed chromatids align along the metaphase plate, anaphase where sister chromatids are separated by the pulling of the mitotic spindles and telophase where the nuclear envelope begins to reform and DNA starts to uncondensed. Finally, cytokinesis completes the cell cycle wherein fission of the cytoplasm gives rise to the individualised daughter cells (Figure 1.11 A).

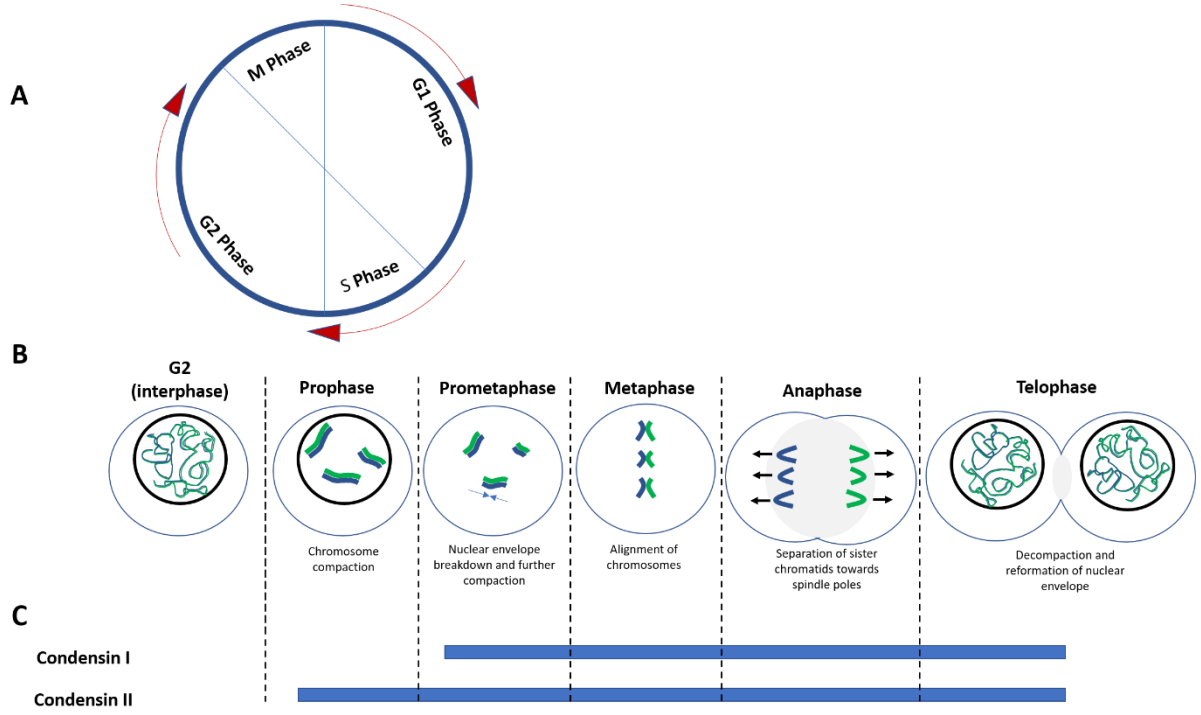


Figure 1.11. The cell cycle. (A) Summary of the phases of the cell cycle. (B) Summary of the cell cycle stages of M phase (mitosis) from G2 to telophase showing the morphological changesof chromosomes associated with each stage. Sister chromosomes are shown in blue and green, the nuclear envelope in black and cell membrane in dark blue. The preceding stages of G1 and interphase are not shown. (C) General summary of the chromatin association of the condensin I and II complexes in higher eukaryotes throughout mitosis (Batty and Gerlich, 2019).

The specific pool of regulatory proteins active during the cell cycle differs somewhat between species, but in eukaryotes at least, the major regulators belong to two families: the cyclin dependant kinases (Cdk's) and the cyclins. Generally, the cell cycle is a rolling cascade of phosphorylation and protein activation in which heterodimers of cell cycle stage specific Cdk-cyclins accumulate and regulate both downstream stage-specific regulators and subsequent Cdk-cyclin pairs.

In budding yeast, stage specificity is conferred through a change in the cyclin component, with a single Cdk (Cdc28) being present throughout the cell cycle. Early in the cycle the cyclins Cln1-3 associate with Cdc28 and are required to drive the cell cycle through G1, followed in S phase by replacement with early-expressed B-type cyclins Clb5 and Clb6 (Bloom and Cross, 2007). Further association of Cdc28 with the B-type cyclins Clb1-4 mediates the cells passage through mitosis.

In higher eukaryotes, multiple Cdk's are used throughout the cell cycle in a stage specific manner, in addition to many cyclin partners. The number and type of Cdks is species specific, for example, humans require 4 Cdks during the cell cycle (Cdk1, 2, 4 and 6), which themselves associate in an overlapping fashion with 4 cyclins (Cyclin D, E, A and B).

1.6.3 Condensin in Mitosis

The mechanical understanding of condensin gives rise to several potential models of its action, but the larger question is how does this molecular mechanical action manifests within whole genome architecture, and what do the resulting phenotypical and morphological observations tell us about condensin's action.

Early experiments in *X. laevis* cell extracts demonstrated the importance of condensin in chromosome morphology. It was shown that condensin-containing cell extracts reproduce morphologically the formation of mitotic chromatids *in vitro* (Hirnao et al., 1997). Further work has since shown that, while there are several additional factors required for mitotic chromosome condensation (Shintomi et al., 2015), even in the absence of histones, condensin alone can mediated DNA compaction (Shintomi et al., 2017).

We now know that in mitosis the genome undergoes a radical restructuring and even before the discovery of condensin it was proposed that this restructuring involved the formation of tight loops on chromatin (Earnshaw and Laemmli, 1984, Marsden and Laemmli, 1979). Local and long-range regulatory interactions in the interphase genome are lost entirely in favour of local looped structures throughout each chromosome (Naumova et al., 2013). The formation of these loops proceeds rapidly through prophase, with interphase structures being entirely absent by late prophase (Gibcus et al., 2018).

In yeast, there is a single condensin complex present throughout the cell cycle which must perform all the complex specific functions, however, in higher eukaryotes there is a division of labour between two complexes: condensin I and condensin II. Typically, condensin II is constitutively present in the nucleus throughout all stages of the cell cycle, while condensin I will only enter the nucleus upon NEB breakdown in prometaphase (Hirota et al., 2004).

In species expressing only one variant of condensin, such as the yeasts, the interphase access of condensin varies between different species. In *S. cerevisiae* the condensin complex is nuclear throughout the entire cell cycle (Freeman et al., 2000), with only small increases in enrichment at selected chromosomal locations such as the centromere and telomers during mitosis. The amount of available condensin appears to be limited by the expression levels of the Ycg1 subunit (Doughty et al., 2016). In fission yeast, condensin is largely excluded from the nucleus, much like condensin I in higher eukaryotes, until the onset of mitosis, though a small population does persist through interphase (Nakazawa et al., 2015).

In higher eukaryotes condensin II begins to compact bulk DNA during prophase, prior to NEB, and forms an initial axial core with large loops (Hiroto et al., 2004). In human cells these loops are around 450 kb, while they are ~400 kb in chicken DT40 cells (Gibcus et al., 2018). Upon entry of condensin I into the nucleus the large loops are further compacted to around 90 kb in human cells or 80 kb in DT40 cells (Gibscus et al., 2018), with condensin I localising around the central condensin II axes (Walther et al., 2018) (by contrast, the onset of mitosis in fission yeast appears to result in the formation of ~35 kb loops within the genome (Schalbetter et al., 2017)).

The loading of condensin I onto chromosomes during mitosis takes place in two phases; the first occurs in promethaphse in conjunction with the nuclear envelope breakdown, and the second in anaphase (Walther et al., 2018). While the condensin I complex is more abundant during mitosis (though the exact ratio varies between species) (Shintomi et al., 2011), it has been shown to be more dynamic than condensin II, with shorter residence times and higher turnover (Gerlich et al., 2006). Loss of either of the condensin complexes leads to morphological changes in the chromosomes and segregation defects. Condensin II depleted cells exhibit longer, less rigid chromosomes lacking the usual defined axis (Ono et al., 2003; Abe et al., 2011), whereas loss of condensin I disrupts lateral compaction leading to wider fuzzier chromatids (Baur et al., 2012; Green et al., 2012). The morphological changes observed upon disruption of either complex point to distinct roles in the organisation of mitotic chromatin, with condensin II being required for a broad, base compaction that establishes the core axis, which is then built on by the formation of compound loops within loops by condensin I.

Recent HiC analysis coupled to molecular dynamic simulation has further refined the differential effects of the two condensin complexes, suggesting that while the initial condensin II mediated large loops are random throughout prophase, in prometaphase the loop axis arranges itself into a stepped helix that progressively "tightens" as cells progress into anaphase (Gibcus et al., 2018).

1.6.3. i. The rDNA and the Centromere in Mitosis

In eukaryotes the structural RNA component in ribosomes (rRNAs) is transcribed from multicopy genes that are spread over the genome, or clustered in certain regions of the genome (rDNA). In *S. cerevisiae* rDNA copies are located in the right arm of Chromosome XII, and include 150-200 tandemly repeated copies of a 9.1 kb rDNA unit (Petes, 1979) that is highly transcribed (Warner, 1999). Due to its highly repetitive nature and high rate of transcription, the rDNA is considered a difficult region to properly segregate, mostly due to the formation of inter-repeat linkages (D'amours et al., 2004).

During anaphase, the mitotic exit phosphatase Cdc14 plays a central role in regulating condensin binding to the rDNA (D'Ambrosio et al., 2008, de los Santos-Velazquez., 2017). Cdc14 is a cell cycle phosphatase that is essential for the transition from mitosis to G1 in *S. cerevisiae* (Stegmeier and Amon, 2004). Yeast Cdc14 mainly dephosphorylates Cdk targets and it is inhibited during most of the cell cycle, only becoming active in early anaphase. At this cell cycle stage, the APC^{Cdc20} complex is activated, causing the activation of the FEAR network (Cdc14 Early Anaphase Release) which in turn produces the inhibition of RNA pol I transcription, which is responsible for the transcription of rRNA genes. Downregulation of rDNA transcription by Cdc14 targets the condensin complex to this locus (Clemente-Blanco et al., 2009) allowing its axial compaction and correct segregation during anaphase. Additionally, condensin is implicated in bringing the rDNA into proximity with the centromere during mitosis which is thought to aid in either separation of the sister chromatids in this region and/or protection of the DNA between the centromere and the rDNA loop from shearing damage during anaphase (Paul et al., 2018).

In budding yeast, the centromere comprises an extended spur of DNA that projects perpendicular to the chromatid (Yeh et al., 2008). At the apex of this projection a centromere

specific variant of the nucleosome binds an exposed section of DNA creating a double looped "basket" that mediates contacts between the chromosome and the kinetochore/mitotic spindle (Bloom et al., 2017, Lawrimore et al., 2016). Condensin and cohesin complexes are localised in pericentromeric regions (Megee et al., 1999, D'Ambrosio et al., 2008) where they are thought to perform a crucial role in organising the DNA into a looped form that orients the centromere projection perpendicular to the chromatid axis. This "bottle brush" structure has been proposed to act as a molecular shock absorber or spring that localises the pulling forces of the mitotic spindle to an area stiffened by organised loops (Lawrimore and Bloom., 2019). These loops have been proposed to be organised by a basal core of condensin that localises to the area most closely aligning with the axis bisecting the spindle pole bodies, while cohesin is more radial, possibly mediating and stabilising interloop interactions (Stephens et al., 2013). Force generated by the stalling loop extrusion activity of condensin in the pericentromere during biorientation has been demonstrated, by molecular dynamic simulation, to be important in stabilising the kinetochore-microtubule interaction, and allowing spontaneous changes in microtubule length to be smoothed out (Lawrimore et al., 2018, Lawrimore et al., 2019, Akiyoshi et al., 2010).

In species containing regional centromeres the situation is slightly different. In this case the centromere is less uniform, though the enrichment of condensin and cohesin at the pericentromere is still a key feature. Repetitive DNA sequences are marked by the presence of the CENP-A histone H3 variant which requires both HJURP (Holiday Junction Recognition Protein) and condensin II for loading and maintenance during mitosis and interphase (Samoshkin et al., 2009, Bernard et al., 2011, Barnhart-Dailey et al., 2017). During mitosis condensin II but not condensin I co-localises with CENP-A at the centromere (Shintomi and Hirnao, 2011, Ono et al., 2004). Like in point centromeres, association of condensin is important for the stiffening of centromeric DNA in order to resist mitotic spindle forces (Ribeiro et al., 2009). Unlike the single Cse4 (yeast centromeric histone variant) mediated kinetochore attachment found in *S. cerevisiae*, chromosomes with regional centromeres can exhibit several kinetochore sites within the same centromere, each of which must be orientated by structuring of the surrounding chromatin such that they all can bind to the mitotic spindle (Lawrimore and Bloom, 2019).

1.6.3 ii. Mitotic Post-Translational Modification

Condensin activity is heavily regulated by extensive post-translational phosphorylation, which tunes its activity throughout the cell cycle. In yeast, multiple kinases including Cdk1, Cdc5 and Ipl1 phosphorylate condensin. Cdk1 phosphorylation occurs at a cluster of consensus sites located on the N terminal region of Smc4, and these phosphorylations appear to control the rate of complex turnover during mitosis (Robellet et al., 2015, Thadani et al., 2018). Priming by Cdk1 is also required for the subsequent phosphorylation by Cdc5, which hyperactivates the complexes DNA supercoiling activity (St Pierre et al., 2009). Hyperphosphorylation appears to lead to defects in chromatin condensation despite mediating longer residence times on chromatin (Thadani et al., 2018), but conversely abrogation of Cdc5 phosphorylation leads to cell death in yeast (St-Pierre et al., 2009). This suggests that the phosphorylation state of condensin needs to be kept finely balanced to maintain viability.

Furthermore, in addition to regulation by Cdk1 and Cdc5, in budding yeast Ipl1 kinase has also been implicated in the phosphorylation of condensin subunits (Lavoie et al., 2004). Similarly, in fission yeast phosphorylation by the Ipl1 homolog aurora B at condensin's Kleisin promotes recruitment of the complex to mitotic chromosomes (Nakazawa et al., 2011) and facilitates telomere separation in anaphase by promoting condensin loading (Reyes et al., 2015).

In higher eukaryotes, condensin has been shown to require Cdk1 phosphorylation for *in vitro* compaction (Kimura et al., 1998, 2001). Cdk1 targets condensin II at a conserved site on the CAP D3 subunit (Abe et al., 2011). This is required not only for the proper formation of mitotic chromosomes in prophase but also for the association of and hyperphosphorylation by Polo-like kinase (Plk1, Cdc5 in *S. cerevisiae*) (Abe et al., 2011).

Later studies have also shown that in higher eukaryotes Plk1 phosphorylates CAP-H2 which stabilises condensin II on mitotic chromosomes, opposing Cdc20 mediated degradation (Kagami et al., 2017). Further degradation of CAP-H2 in *D.melanogaster* appears to be mediated by ubiquitin ligase activity but this system is not entirely conserved across species (Buster et al., 2013, Kagami et al., 2017). In higher eukaryotes, Aurora B is also important in the regulation of condensin I, aiding recruitment to the mitotic chromatin (Lipp et al., 2007, Takemoto et al., 2007).

Finally, the Msp1 mitotic kinase which is involved in the spindle assembly checkpoint, segregation and cytokinesis (Liu and Winey, 2012) has also been shown to govern the loading of condensin II in HeLa cells by targeting the CAP-H2 subunit (Kagami et al., 2014).

1.6.3. iii Condensin-protein Interactions in Mitosis

Condensin is one of many proteins or protein complexes enriched in mitotic chromatin. An extensive network of factors are involved in the compaction and segregation process, many of which interact directly with condensin, or indirectly in related processes. Three key non-SMC regulators of the process are discussed below.

Histones are of course ever present in chromatin but their role in condensin-mediated compaction is still unclear. Direct interactions between condensin and histone H2A and H2B tails suggests a role (Wilkins et al., 2014, Tada et al., 2011), although the requirement for the FACT complex (Shintomi et al., 2015) and the correlation of condensin with nucleosome depleted regions (Toselli-Mollereau et al., 2016) would seem to paint histones as more of a barrier than a helper for condensin functions on chromatin. Deacetylation of histone H4 tails has been suggested to mediate an interaction between individual nucleosomes that may contribute to the condensation but the significance of this to condensin function is unclear (Wilkins et al., 2014).

Topoisomerase II is known to colocalise with condensin II at the chromatin axis and is known to be lost upon disruption of condensin binding (Leonard et al., 2015). Either through direct or indirect interactions with condensin, the presence of topoisomerase II is important for the proper segregation of sister chromatids (Charbin et al., 2014). It has been suggested that the positive supercoiling activity of condensin creates a substrate upon which topoisomerase II can preferentially decatenate sisters (Baxter et al., 2011, Baxter and Aragon, 2012), a function which is essential for proper segregation.

Another protein highly enriched on mitotic chromatin is Kif4A, a kinesin motor protein found in the chromosome axis in mitosis. Kif4A has been shown to interact with condensin I via direct binding to the CAP-G subunit through its C terminal region. The role of Kif4A is

thought to be assisting in the alignment of condensin to the chromatin axis (Samejima et al., 2012, Takahashi et al., 2016), though not in the centromere (Stumpff et al., 2012). Loss of Kif4 leads to shorter wider chromosomes similar to loss of condensin I, as well as impacting on the rigidity of mitotic chromosomes (Takahashi et al., 2016).

1.6.4. Condensin in Meiosis

As with mitosis, proper segregation of chromosomes into gametes during the meiotic division requires condensin (Lee, 2017). Briefly, meiosis consists of the usual 3 stages of interphase which are followed by 2 consecutive cell divisions (meiosis I and II) without intervening DNA replication that separate both maternal and paternal copies of each chromosome, followed by separation of replicated sister chromatids into daughter cells.

During prophase I, in conjunction with chromatin condensation, a copy of each of the paternal and maternal homologous chromosomes engage in "synapsis" mediated by the synaptonemal complex. This process aligns homologous chromosomes by sequence and binds them together. Subsequently, a process of recombination creates further linkages called chiasmata in which homologous chromosomes are physically linked by DNA crossovers between them. These paired homologs remain cohesed to their replicated sister chromatids similar to mitosis. The associated homologs are known as bivalents and these linkages allow alignment in the metaphase I plate using the spindle microtubules which attach to the pair of sister chromatids that form each homolog.

At anaphase I chiasmata and synapsis are resolved, and co-oriented sister chromatids are pulled towards the same spindle pole. Telophase I and cytokinesis then resolve meiosis I, giving rise to 2 daughters (Bolcun-Filas and Handel, 2018).

Meiosis II further divides the daughter cells' genetic content. Condensation is usually maintained while sister chromatids are bioriented in metaphase II and again separated, this time to opposite spindle poles in anaphase II. A final telophase II and cytokinesis separate one of each sister chromatid into 4 daughter cells with haploid genomes (Bolcun-Filas and Handel, 2018).

Condensin is essential for the prophase condensation of meiotic chromatin in higher and lower eukaryotes (Yu and Koshland, 2003, Hagstrom et al., 2002). In mouse oocytes, condensin II is essential for the maintenance of compaction in bivalents, with targeted proteolytic inactivation leading to the unravelling of the chromosomes (Houlard et al., 2015). Additionally, condensin is required to resolve non-cohesin interlinks between sister chromatids and homologs before metaphase I (Chan et al., 2004).

A further specialised role is seen for condensin at the rDNA array in *S. cerevisiae* where it aids in double strand break suppression and repair, downregulating Spo11 mediated breaks and ensuring that any breaks made are resolved before anaphase (Li et al., 2014).

Condensin also appears to be important for the recruitment of synaptonemal complex components (Yu and Koshland, 2003, Hernandez et al., 2018) implicating it in the establishment of meiosis specific interchromatin interactions. Another SMC complex, cohesin is heavily involved in the establishment of these interactions, and it appears that condensin is important for cohesins' stability on chromatin (Hernandez et al., 2018).

Separase mediated loss of cohesin plays a role in both anaphase I and II. Loss of cohesin from the chromosome arms during anaphase I is required for homolog segregation (Kudo et al., 2006), but simultaneously, a pool of centromeric cohesin must remain stably bound until cleavage in anaphase II to prevent early segregation of sister chromatids (Tachibana-Konwalski et al., 2010). Condensin is implicated in both the maintenance of centromeric cohesin, which requires the suppression of Wapll-mediated dissociation (Hernandez et al., 2018), and also in the selective loss of cohesin from chromosome arms in anaphase I (Yu et al., 2005), thus ensuring the correct segregation of meiotic chromatin.

1.6.5. Condensin in Interphase

While the principle role of condensin is generally thought to be during cell division, there is a small pool of evidence suggesting that it is also involved in the regulation of gene expression and genome architecture during interphase. In those species containing two condensin complexes, it is condensin II that seems primarily responsible for this function, being present in the nucleus throughout the cell cycle. Condensin I plays a role more restricted to mitotic

states, entering the nucleus at prometaphase when the nuclear envelope breaks down and being rapidly removed over the course of G1 following the completion of mitosis (Zhang et al., 2016). A small population of condensin I has been reported to persist in the nucleus following mitosis (Li et al., 2015; Kim et al., 2013), but the role of this population is still under investigation.

1.6.5. i Condensin and Interphase TADs

The use of high-resolution techniques such as HiC and genome paints has revealed that the interphase genome is made up of a dynamic hierarchy of 3D interactions that bring regulatory elements into proximity. Local, 1-2 Mb loops of DNA form that are known as topologically associated domains in which the DNA contained within the looped section forms a self-interacting and insulated unit (Dixon et al., 2012). Establishment of these domains through G1 following mitosis is heavily associated with cohesin and the structural protein CTCF (Schwarzer et al., 2017, Li et al., 2020), however, condensin II has also been shown to be enriched at the borders of these TADs, with enrichment correlating with a greater degree of insultation from adjacent TADs (van Bortle et al., 2014). TAD border proximity is correlated with a higher degree of transcription in genes, particularly histone genes in mammalian cells and tRNA genes; genes which condensin II is thought to cluster together in the interphase nucleus (Yuen et al., 2017, Haeusler et al., 2008). Transcription factor III C (TFIIIC) and its association with RNA polymerase III are thought to play an important part in localising condensin II to these dense clustered regions, especially given that TFIIIC and condensin II localise to enhancer and promoter elements within the TADs, offering a hint at why this arrangement might result in higher rates of transcription (Yuen et al., 2017).

Additional stress-induced mechanisms are at play in flies where the loss of condensin II subunit CAP-H2 at TAD boundaries corresponds to an increase in long-range interactions within the genome (Li et al., 2015). This may be part of a mechanism in which TAD boundaries are made more flexible to allow a transcriptional response to stress, though its larger significance is unclear.

Interestingly, a very recent study has contradicted the previous assertion that condensin II is playing a role in TAD formation and gene regulation in interphase (Abdenur et al., 2019).

This study in mouse hepatocytes found no change in genome organisation by HiC, unlike deletion of the cohesin regulator Nipbl which led to large scale changes, nor were drastic changes in gene expression observed. Further studies in yeast support the idea that condensin is not regulating gene expression levels, but instead global changes in interphase transcription are a result of segregation defects in prior cell cycles (Paul et al., 2018, Hocquet et al., 2018).

1.6.5. ii Condensin and Chromatin Territories

The local folding of TADs is found within another layer of organisation known as chromatin compartments, in which homologous interactions between active or inactive chromatin form related domains. In addition to the TADs and compartments, individual chromosomes exhibit preferences in the space they occupy within the nucleus (Rosin et al., 2018, Zhang et al., 2019, Iwasaka et al., 2016). These chromosome territories limit the amount of inter-mixing between chromosomes to a greater or lesser extent depending on the cell type and cell cycle stage. Work in *D. melanogaster* shows that local condensation of the genome into chromosome territories is undone upon loss of condensin II subunit CAPH2, with consequences for gene expression in interphase (Rosin et al., 2018). Similarly, in fission yeast, chromosome territory formation is also driven by condensin (Iwasaka et al., 2016, Aono et al., 2002).

Additonally, in flies, the chromosome territories mechanisms is at odds with a unique architectural feature known as transvection in which homologous chromosomes regulate each other *in trans*. This situation is antagonised by condensin II, likely through its control of chromosome territories (Hartl et al., 2008). While the aforementioned studies in yeast contradict the idea of condensin regulating gene expression levels due to both TADs and loss of chromosome territories, in both cases the genome volume is altered significantly (Rosin et al., 2018, Hocquet et al., 2018).

1.6.5 iii Condensin and Further Transcriptional Regulation

Condensin localisation correlates with areas depleted of nucleosomes, with one model of condensin localisation suggesting that nucleosomes constitute a barrier to condensin loading and that acetylation dependant remodelling of histones creates nucleosome depleted regions of DNA onto which condensin can load (Toselli-Mollereaux et al., 2016). These regions

correspond with transcription start sites (Sadeh & Allis, 2011) and are linked to transcription mediated remodelling by several factors including the RSC (Remodelling the Structure of Chromatin) (Toselli-Mollereaux et al., 2016) and FACT complexes (Facilitates Chromatin Transcription) (Dewari & Bhargava, 2014).

In addition to its association with and clustering of tRNA and histone genes, condensin also has a complex association with the rDNA array, with several condensin subunits having been reported to localise to the rDNA array in interphase (Cabello et al., 2001). *In S. cerevisiae*, silencing of rDNA transcription is mediated by condensin in a Sir2p-dependant manner (Machin et al., 2004). In humans, binding to the rDNA array by condensin negatively regulates transcription of rRNA by antagonising binding of CTCF, which acts in an RNA pol I dependant manner (Huang et al., 2013).

C. elegans expresses an additional condensin complex, named the dosage compensation complex, that is defined by a unique Smc4 homolog DPY27. In interphase this complex is involved in regulation of gene expression on the X chromosome, providing dosage compensation in XX hermaphrodites (Crane et al., 2015). In this role the dosage compensation complex prevents the association of RNA pol II at its promoter, reducing the overall gene expression from the X chromosome and compacting this chromosome up to 40% (Albrittron et al., 2018). It is interesting to note how the widely reported association of the dosage compensation complex with this type of promoters seems to have been co-opted to provide additional functionality.

1.8. Condensin in disease

A growing body of evidence is beginning to implicate the condensin complex and its subunits in several pathologies. Misregulation or mutation of all subunits of condensin I and II in humans has been associated with cancer pathology in a range of tissue and tumour types. For example, mutation or misregulation of the condensin Smc proteins is found in types of lymphoma (Ham et al., 2007), neuroblastoma (Murakami-Tonami et al., 2014), colorectal (Je et al., 2014), breast, ovarian (Kar et al., 2016), lung and prostate cancers amongst others (reviewed in Wang et al., 2018). Similarly, all the non-Smc subunits of both condensin I and II are implicated in cancer such as upregulation of NCAPH and NCAPG being correlated

with melanoma (Ryu et al., 2007). All the subunits of condensin II (NCAPG2, D3 and H2) have also been found to be misregulated in malignancies (Deutschman et al., 2018, Zhan et al., 2017, Wang et al., 2018). The final subunit to be implicated in cancer pathology, hCAP D2, was most recently found to be positively associated with increased metastasis and poor prognosis in triple negative breast cancers (Zhang et al., 2020).

Analysis by siRNA of individual condensin genes has shown diminished condensin expression is associated with increased genome instability phenotypes (Baergen et al., 2019). Genomic instability manifesting as a gain or loss of genomic material is strongly associated with cancer phenotypes such tumour heterogeneity (Jefford et al., 2006), metastasis (Gao et al., 2016), cellular transformation (Nowak et al., 2002), and drug resistance (Lee et al., 2011, Duesberg et al., 2000), suggesting that loss of condensin function could play an important role in the pathological progression of cancer by allowing further damage to occur within the genome (Baergen et al., 2019). The important role of condensin in maintaining genome stability could in future make it an important prognostic factor in cancer diagnosis, and also as a potential therapeutic target.

In addition to its role in cancer progression, condensin mutations are also associated with congenital defects, including microencephaly, which likely occurs as a consequence of impaired chromosome segregation (Perche et a., 2013, Martin et al., 2016). Additional studies have also pointed to apparent neurological, ocular, urogenital, skeletal, and digit defects in patients with CAP-G2 mutations (Khan et al., 2019). The emerging clinical significance of condensinopathies and related SMC mutant pathologies such as cornelia de lange syndrome, marks the condensin complex as an important object of clinical study. Understanding both the underlying mechanisms, key to the fundamental process of cellular division, and the clinical repercussions of the disruption of such mechanisms in humans, will no doubt be important to our wider understanding of the genome and its maintenance.

1.9 Project Aims

Since the discovery of the pentameric condensin complex in 1997, much work has been done to uncover the underlying molecular mechanisms that facilitate its aforementioned roles in mitosis, meiosis and interphase. This accumulating evidence has provided tantalising

glimpses into the mechanisms but a unified mode of action that reconciles these observations is not yet complete. A key step on the path to this unified model of condensin function would be to produce a detailed atomic-resolution structure of the entire complex. Such a structure would provide a solid basis for predictive experimentation and a core around which previous data can be reinterpreted and applied.

In addition to the atomic resolution model, it is also important to consider condensin in its proper context, to best understand the molecular mechanisms at play *in vivo*. Recent single molecule studies have focused on assaying condensin *in vitro* on naked DNA, however this work largely does not account for the fact that condensin *in vivo* will be required to carry out its molecular functions on chromatinised templates, which also include DNA binding proteins and other obstacles. These gaps in the literature give rise to two main project aims:

1) To produce an atomic-resolution structure of the full condensin complex using cryo electron microscopy (cryo EM).

2) To use purified histones and optical tweezers technology to assay the ability of condensin to compact, translocate and form loops on chromatinised DNA.

Chapter 2 – Purification of the *S. cerevisiae* Condensin Complex

2.1 Purification and Cryo EM analysis of the Condensin Complex

The purification of the *Saccharomyces cerevisiae* condensin complex was first described in (St Pierre et al., 2009) and has since been adapted for use in many studies (Terakawa et al., 2017, Ganji et al., 2019). Adapting this reliable, high yield method was the first step towards addressing the aims of this project and figure 2.1 summarises the purification process used. Genes encoding the 5 subunits of condensin (Smc2, Smc4, Brn1, Ycg1 and Ycs4) were inserted into 2 plasmids containing markers for growth on -trp and -leu selective media and the plasmids were transformrd in *S. cerevisiae* W303-1a cells. Two of the condensin subunits, Smc4 and Brn1 were modified to contain affinity tags; specifically, 3xStrepII repeats on the C terminal of Smc4 with a short linker and 12x histidine (His) repeats on the C terminal of Brn1. Brn1 also contained a 3x haemagglutinin (HA) affinity tags between the C terminal and the 12 x His tag that was not used for purification.

Expression was induced by addition of 2% galactose at an optical density (OD) of 1. This triggers expression of the 5 subunits under GAL 1/10 or GAL7 promoters. Following lysis, a three-step purification of condensin from the whole cell extract results in a largely pure, stoichiometric complex as shown by SDS PAGE with Coomassie staining in figure 2.1A. This process involves sequential affinity pulldowns exploiting the 12x His and the 3x StrepII tags as described previously (St Pierre et al., 2009, Terakawa et al., 2017), followed by size exclusion chromatography (SEC).

SEC, also known as gel filtration (GF), separates protein on the basis of size by passing a protein sample over a porous resin (stationary phase). Large proteins or aggregates are excluded from the pores of the resin by their size and therefore elute from the column first as they run entirely between gaps in the packed resin bed. The volume of buffer (mobile phase) required to elute these molecules is called the void volume. Smaller molecules which have full or partial access to the resin pores are retained on the column. The smaller a molecule is

the more access it has to the resin, thus proportionately increasing its retention time and elution volume. Monomeric condensin is separated from higher molecular weight aggregates by SEC and the final SEC eluate was concentrated to around 1.8 – 2 mg/mL. It is this complex that is used in all further experiments. While there is some limited evidence that the higher molecular weight aggregates may have a functional role, analysis of these complexes was beyond the scope of this study (Keenholtz et al., 2017).

Figure 2.1B shows the elution profile of condensin following the final SEC step. The elution of protein from the SEC column is measured by the degree of absorbance of the aromatic side chains of amino acids within the protein at wavelength of 280 nm (ultraviolet). This is plotted against the volume of buffer flowed over the column following injection of the concentrated protein sample.

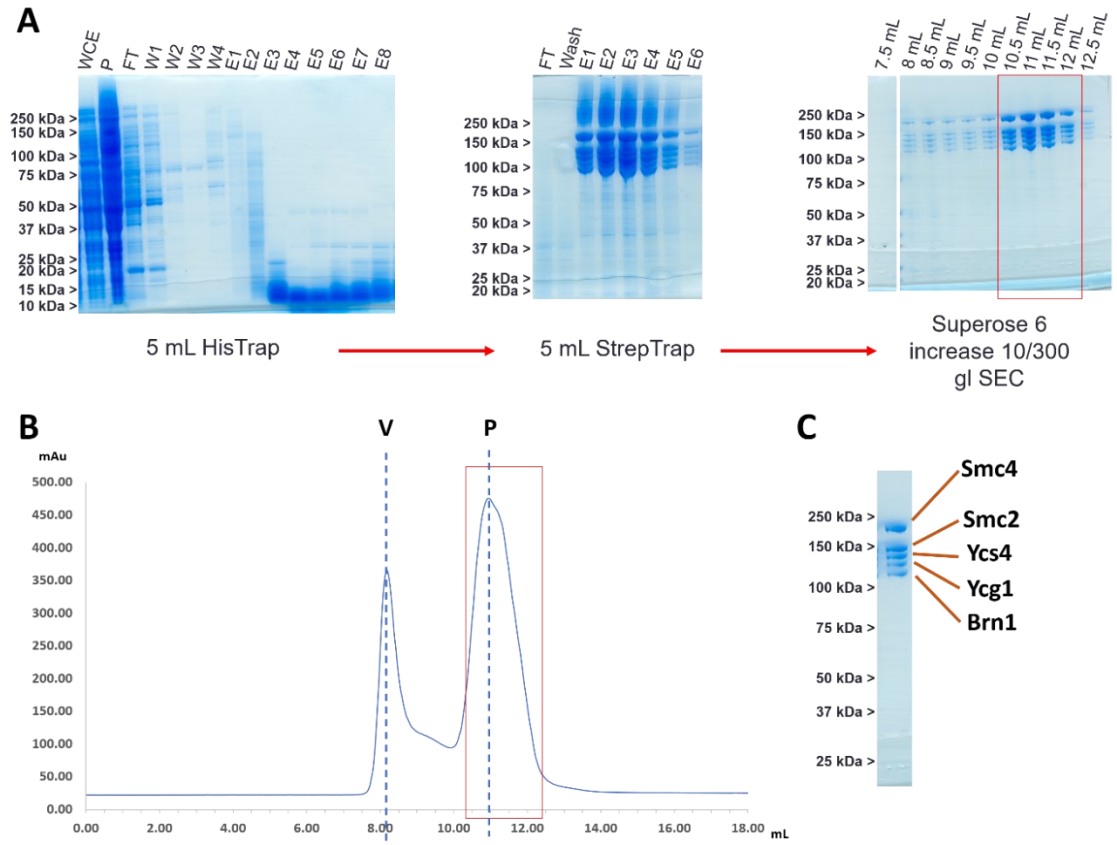


Figure 2.1. Purification of the condensin pentamer. (A) Summary of condensin purification workflow and elution fractions as assayed by coomasie stained SDS PAGE gel. Whole Cell Extract (WCE), Pellet (P), Flow-through (FT), Wash (W), and Elution fractions (E) are indicated. Red box indicates the 0.5 mL elution fractions that were pooled and concentrated to ~2 mg/mL for further experiments. (B) Chromatogram of the gel filtration step. Red box indicates the peak fractions that were pooled, as indicated in (A). V indicates the void volume and P indicates the elution volume of the pentameric condensin complex. (C) Coomassie stained SDS PAGE gel of the final condensin indicating the positions of the individual subunits.

With the purified WT condensin pentamer in hand, the process of analysis by cryo EM was started.

Imaging of the purified pentamer was done in collaboration with the Lowe laboratory (LMB, Cambridge) and preliminary electron micrographs revealed a largely homogenous sample (figure 2.2) which was then subjected to further refinement.

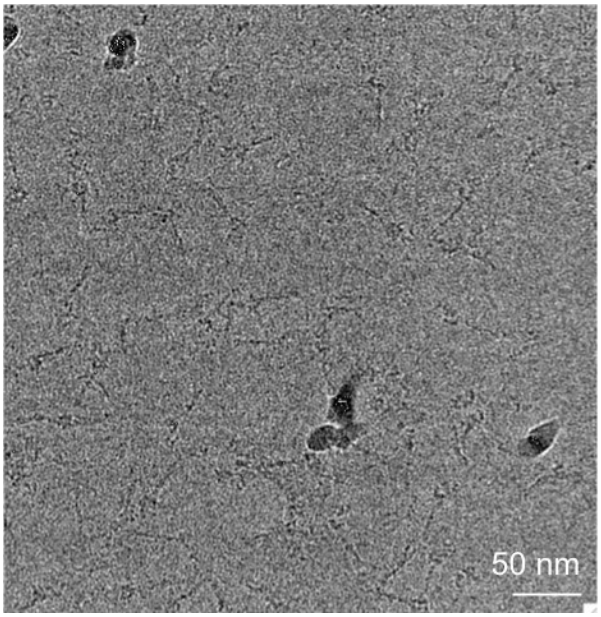


Figure 2.2. Widefield image of WT S. cerevisiae condensin. Produced in collaboration with Dr Byung-gil Lee.

2D class averages are shown in figure 2.3 which revealed an elongated structure consistent with the known properties of the coiled-coil, with a globular domain at one end and a further poorly defined floppy mass beneath this globular domain. Identities were tentatively assigned to these low-resolution structures, with the assumption that the elongated mass most likely corresponded to the coiled-coils and the lower globular domain and associated floppy mass constituted the Smc2/Smc4 heads and HEAT repeat subunits (figure 2.3 C). Previously published structures of the HEAT repeats correlated well with the sub set of images in which the lower mass resolved into a U-shaped mass (Kschonsak et al., 2017).

It was also clear, however, that the complex displayed a preferred orientation. Figure 2.3 B demonstrates this phenomenon in the form of a heat map, with the relative number of images taken from a given coordinate about a 3D sphere surrounding the complex being shown by a

colour gradient. Clustering of signal intensity around a limited number of coordinates indicates that the complex predominantly lies flat on the cryo EM grid in a limited number of orientations. In order to produce a full 3D rendering of the complex at atomic resolution, images must be captured in a full 360-degree sphere, with a sufficient number of images for each orientation available to average them into a high-resolution image.

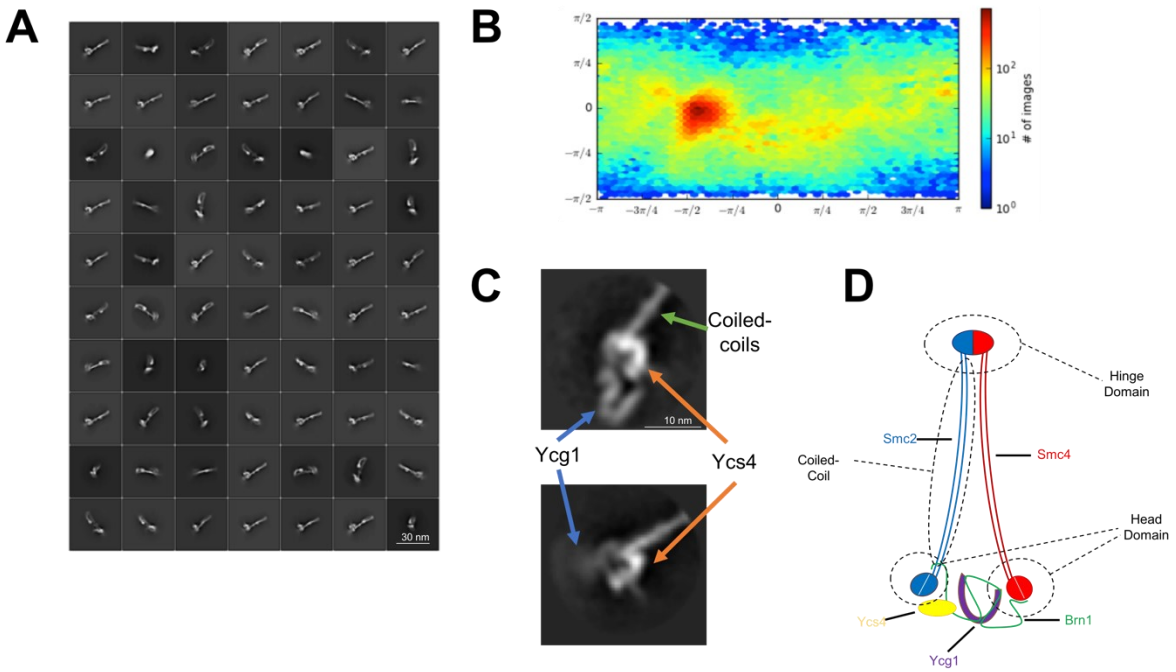


Figure 2.3. Cryo EM of the condensin pentamer. (A) 2D class average electron micrographs of the purified condensin complex. (B) HEAT map showing the relative distribution of condensin complex orientations. (C) Close up view of condensin class averages indicating key putative domains (D) Cartoon schematic of the condensin pentamer based on known structural data. Produced in collaboration with Dr Byung-gil Lee.

To rectify the preferred orientation problem, we next attempted to produce sub-assemblies of the condensin complex missing the HEAT repeat subunits. We reasoned that removal of these subunits could free the complex from the grid without disrupting the core architecture of the Smc-kleisin trimer. It was hoped that by freeing the complex from the grid this might then allow it to be frozen and imaged from a greater range of angles.

2.*2* Purification and Cryo EM Analysis of Condensin Subcomplexes

Purification of a trimer of core Smc-kleisin protein and tetramers missing either Ycs4 (ΔYcs4) or Ycg1 (ΔYcg1), was attempted. As with the pentamer, these complexes were purified in a three-step process using His and StrepII tag affinity purification followed by SEC. Figure 2.4 shows a summary of the final purified material in this process for each of the attempted complexes. While it was possible to eliminate the Ycg1 subunit and maintain a stable, stoichiometric tetramer (figure 2.4 C), removal of both HEAT repeat proteins or Ycs4 resulted in disassembly of the complex to a stoichiometric dimer of the Smc sub-units with sub-stoichiometric Brn1 only appearing in the void volume (figures 3A-3B). Analysis by SDS PAGE demonstrates the apparent loss of Ycg1 signal as detectable by Coomassie in the case of the ΔYcs4 tetramer accompanied by a significant shift in the major peak to ~12.2 mL.

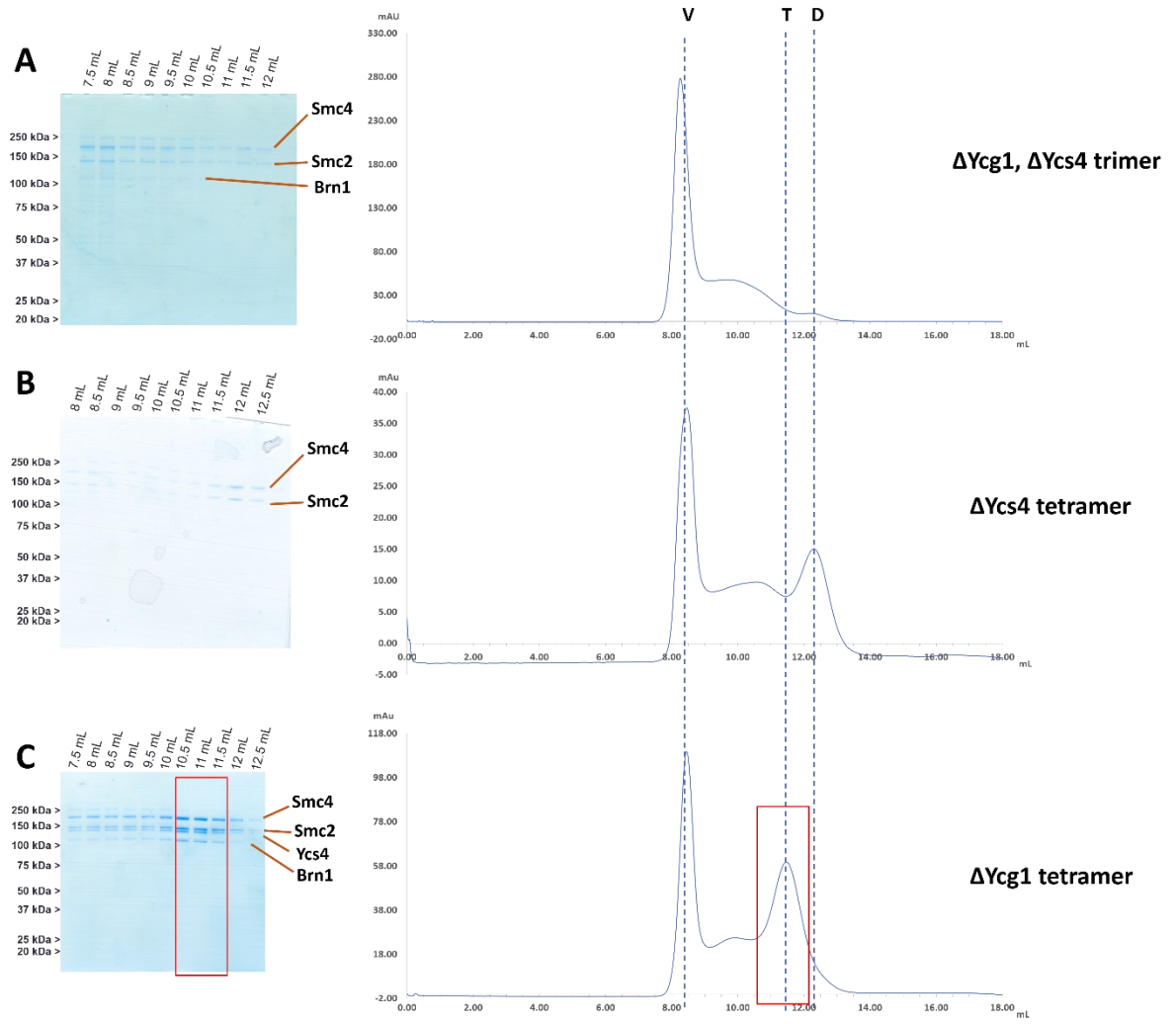


Figure 2.4. Coomasie stained SDS PAGE gels of peak elution fractions for condensin sub complexes. V, indicates void volume, T, the elution volume of the condensin tetramer and D the elution volume of a putative Smc2/Smc4 dimer. (A) SDS PAGE and chromatogram for the attempted condensin trimer purification containing Smc2, Smc4 and Brn1 only. Void volume and dimer peak elutions are indicated. (B) SDS PAGE and chromatogram for the ΔYcs4 tetramer sub-complex. Void volume and dimer peaks are again indicated. (C) SDS PAGE and chromatogram for the ΔYcg1 tetramer sub-complex. Red boxes on both the SDS PAGE gel and chromatogram indicate the peak fractions that were pooled and concentrated for further experimentation.

To confirm that the loss of the non-Smc subunits was due to complex instability and not under-expression of Brn1 during induction, the relative amounts of expressed Brn1 were analysed by western blot. Staining with antibodies against the C terminal 3 x HA tag revealed that Brn1 was expressed in all three strains carrying sub-complexes. Though there was some apparent degradation of Brn1 in the trimer strain, it was still present and both the ΔYcs4 tetramer and ΔYcg1 tetramer showed equivalent expression. Therefore, loss of the non-Smc subunits in the final SEC elution was not due to under-expression.

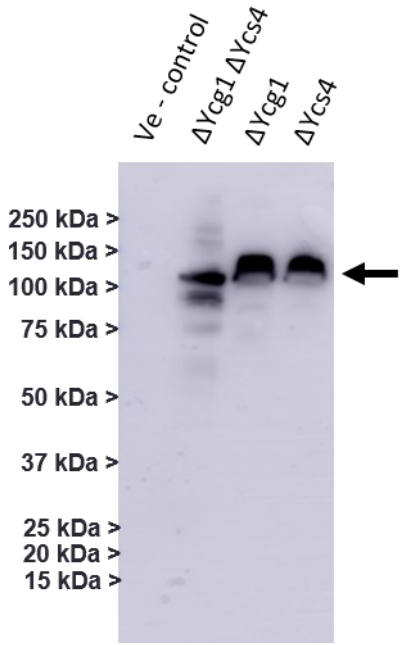


Figure 2.5, Western blot analysis of whole cell extracts taken from cells expressing condensin sub-complexes. 30 ug of extracted protein was loaded for control cells with no expression vector and cells expressing ΔYcs4 tetramer, ΔYcs4 tetramer and the ΔYcs4 ΔYcg1 trimer. Brn1 was detected by anti-HA antibody, indicated by a black arrow.

It was reported by our collaborators that on the cryo EM grid inclusion of the detergent CHAPS, while detrimental to the association of the HEAT repeats, was helpful for stabilisation of the core trimer (data not shown). Inclusion of CHAPS in the purification process was also attempted to stabilise the Ycg1 tetramer or the core trimer, though this did not improve the purification.

As with the pentameric condensin complex, the ΔYcg1 tetramer was subject to cryo EM analysis, revealing a homogenous sample similar to the pentamer samples (figure 2.6).

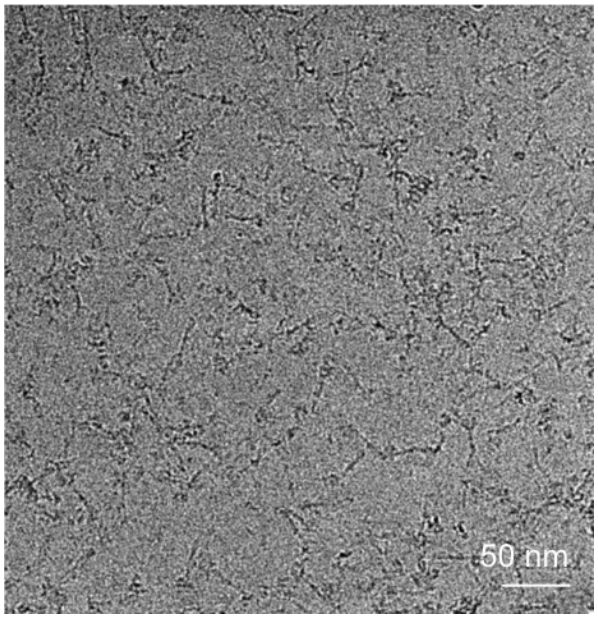


Figure 2.6. Widefield image of WT S. cerevisiae condensin. Produced in collaboration with Dr Byung-gil Lee.

2D class averages of the ΔYcg1 tetramer are shown in figure 2.7 A, which lead to two important conclusions when compared to the pentameric complex. Firstly, the ΔYcg1 tetramer appears to show improved orientation (figure 2.7 B) suggesting that removal of Ycg1 has indeed somewhat freed the complex. Comparison of pentameric 2D electromicrographs to the tetrameric complex also allowed tentative identification of Ycg1 and possibly Ycs4 signals, on the basis of elimination (figure 2.7 C). The largely diffuse and poorly averaged mass below the main body of the complex is lost when imaging the ΔYcg1 tetramer (figure 2.7 C). This loss suggests that the density correlates with the Ycg1 sub-unit, which shows poor resolution suggesting a degree of floppiness in the pentamer images.

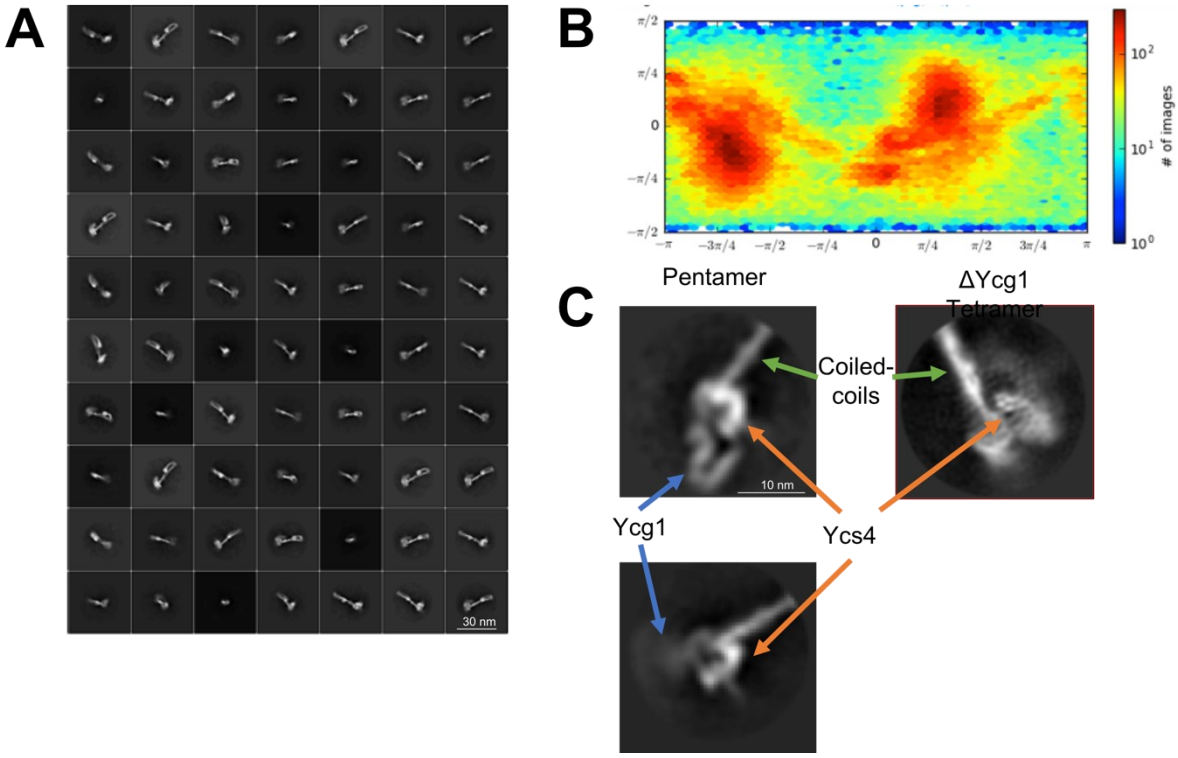


Figure 2.7. Cryo EM of the condensin ΔYcg1 tetramer. (A) 2D electron micrographs of ΔYcg1 condensin tetramer. (B) HEAT map showing the relative distribution of ΔYcg1 tetramer complex orientations. (C) Comparison of 2D electron micrographs between the condensin pentamer and the ΔYcg1 tetramer complex, indicating putative subdomains. Produced in collaboration with Dr Byung-gil Lee.

3D averaging of both the pentameric and ΔYcg1 tetrameric complexes led to a maximum resolution of 4.5 Å at the head/kleisin/HEAT domain, 6.5 Å at the hinge/coiled-coils, and 8 Å overall (figure 2.8) but this resolution was ultimately insufficient to confidently assign individual residues within the complex. It would therefore be necessary to fit existing crystal structures of subdomains of condensin, or fit highly conserved domains crystalised in related family members, to produce a pseudo-atomic model. A number of these domains had been solved previously, including the Smc1 and Smc3 heads and their respective interactions with the kleisin subunit in *S. cerevisiae* (Haering et al., 2004, Gligoris et al., 2014), the *S. cerevisiae* condensin Ycg1-Brn1 sub-complex (Kschonsak et al., 2017), the mouse condensin hinge domain (Griesse et al., 2010) and portions of the coiled-coils and hinge of *P. furiosus* condensin (Soh et al., 2015). Crucially though, at this time there was no suitable structure for the Ycs4 subunit, making it necessary to purify and crystallise the protein *de novo*.

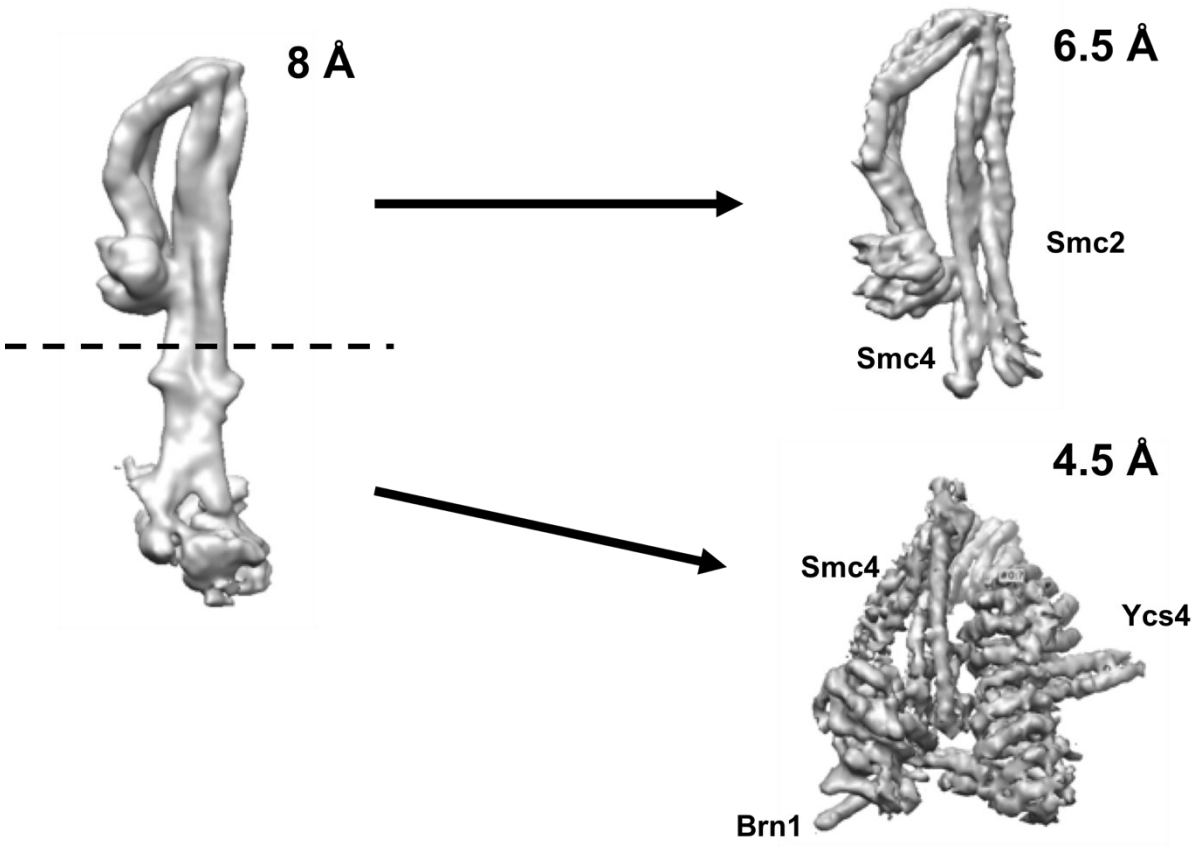


Figure 2.8. Low resolution 3D structure of the condensin complex and associated resolutions of individual head and coiled-coil domains prior to crystal structure fitting. Produced in collaboration with Dr Byung-gil Lee.

Chapter 3 – Purification of Ycs4

3.1 Purification of Ycs4 for Crystallography – Preliminary GST Tagging and Optimisation

As there were no previously described structures of Ycs4 it was decided the best approach initially would be to attempt to purify several homologs, in order to increase the chances of one or more producing crystals. Ycs4 genes from *S. cerevisiae, Candida albicans, Picchia pastoris and Homo sapiens* were selected for this purpose and these genes were cloned into pGEX-6p-1 expression vectors. This introduced an N-terminal glutathione S-transferase (GST) tag to Ycs4 in all cases, and a preScision protease site in an intervening linker region that allowed removal of the GST tag following affinity purification with glutathione agarose beads.

Expression vectors were transformed into BL21 (DE3) strain *E. coli*. At first these were induced in a single condition to gauge the efficiency of expression of each gene. Cells were grown to OD 0.6, then induced with 0.5 mM IPTG for 4 hours of overnigh. (figure 3.1). 1 mL samples of cell culture were taken and pelleted, followed by lysis in buffer containing 1% SDS at 95^{o}C, then equal amounts of sample were then loaded onto SDS PAGE gels and band intensities at the expected molecular weights were compared.

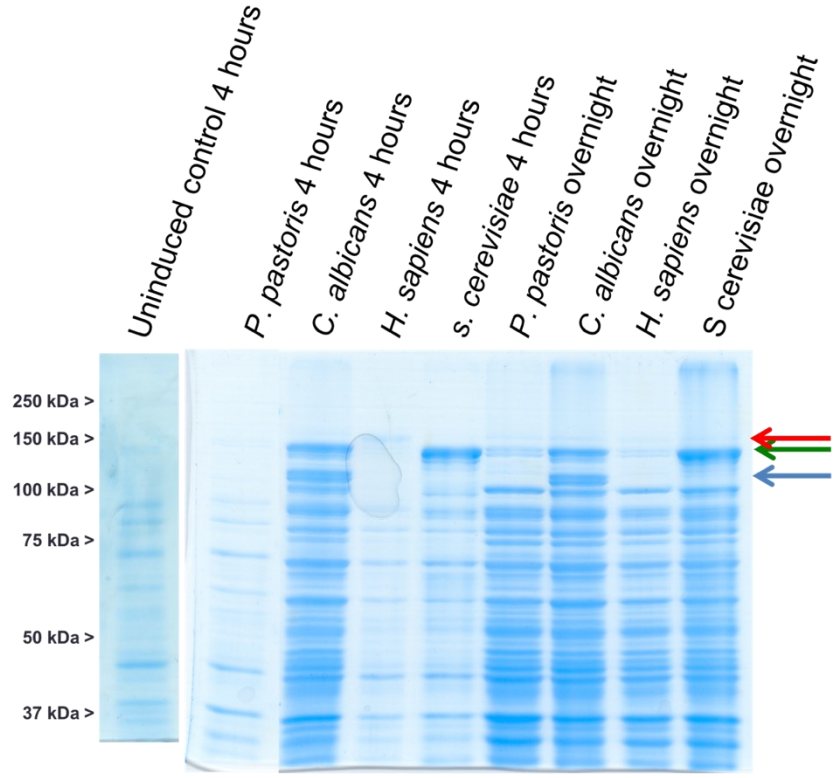


Figure 3.1. Coomasie stained SDS PAGE gels of BL21 (DE3) E. coli whole cell extracts following transformation with expression vectors containing N terminal GST tagged Ycs4 homologs. E. coli were induced at OD 0.6 with 0.5 mM IPTG for 4 hours of overnight. Green arrows indicate the expected molecular weight of S. cerevisiae and C. albicans GST-Ycs4, red arrows indicate H. sapiens GST-NCAP D2.

S. cerevisiae and *C. albicans* GST-Ycs4 were carried over for further expression optimisation based on an initially favourable induction. Optimal induction time, media type, IPTG concentration and expression temperature were screened (figure 3.2). and assayed by SDS PAGE of whole cell extracts as before (figure 3.1). From these experiments it appeared that 4 hour inductions at 37°C produced the best expression (figure 3.2). It was also noticed that an additional lower molecular weight band appeared during expression of the *C. elegans* Ycs4, possibly as a truncated version of the protein.

Induction of *P. pastoris* Ycs4 yielded little or no protein at the expected molecular weight and *H. sapiens* appeared to also express poorly, therefore further optimisation of these proteins was discontinued.

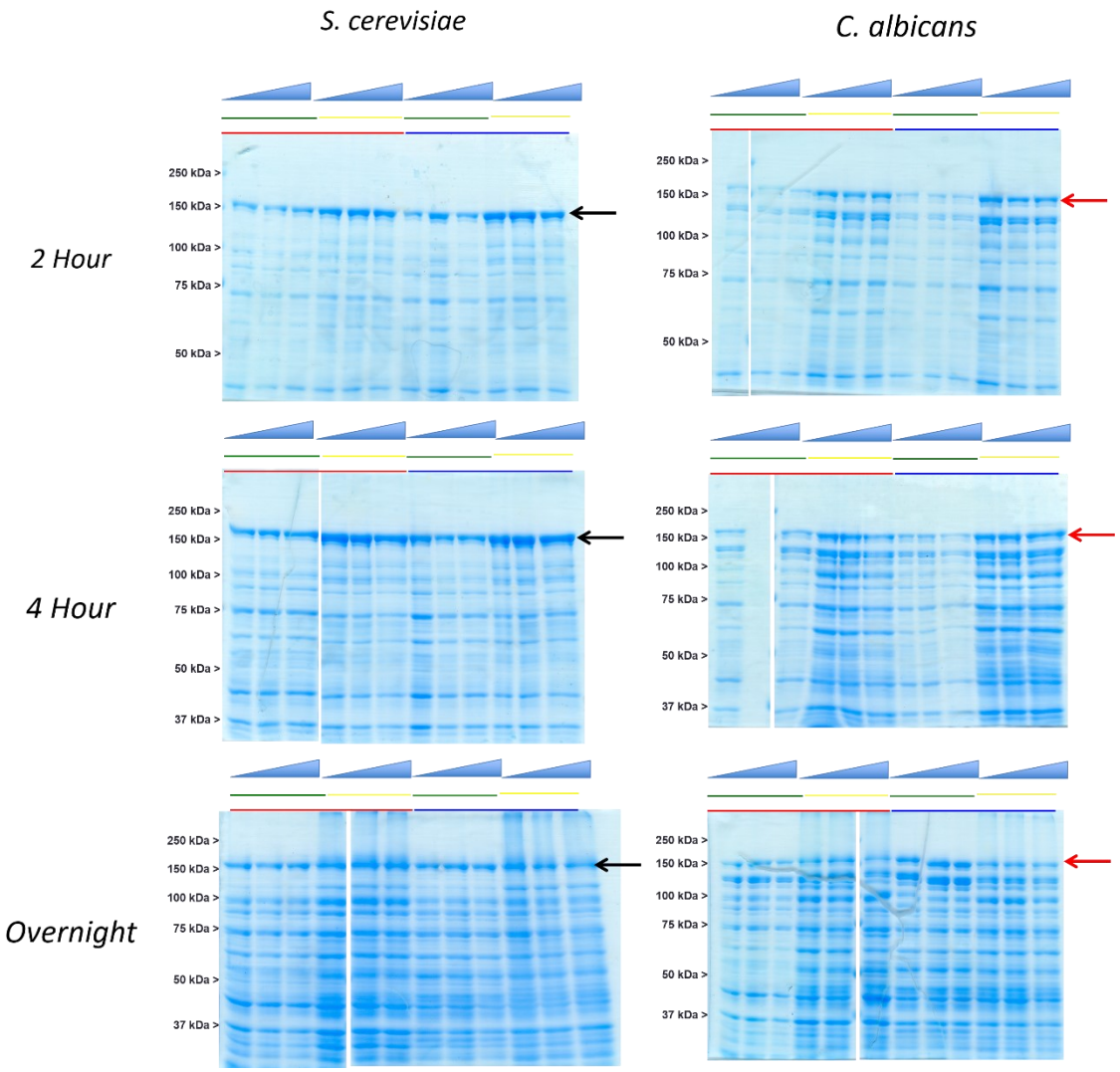


Figure 3.2. Further expression screening of S. cerevisiae (left gels) and C. albicans (right gels) GST-Ycs4 by SDS PAGE. Coloured bands above gels correspond to expression conditions. Red bands indicate extracts from cells grown in LB media and blue bands cells grown in 2 x TY media. Green bands indicate cells induced at 20°C and yellow bands at 37°C. Blue triangles indicate increasing concentrations of IPTG from 0.2 mM to 0.5 mM to 1 mM.

Test induction conditions were scaled up to 6 L cultures and a 3-step process of purification was devised. First, cells were lysed by incubation at 4°C with 0.5 mg/mL lysozyme for an hour, followed by sonication. Cleared lysates were treated with benzonase nuclease to remove DNA contamination and reduce sample viscosity, then subjected to affinity purifications over glutathione agarose beads. Proteins were eluted using preScission protease to remove the 26 kDa GST tag. This was done to reduce any interference of the tag with the crystallisation process (figure 3.3A).

The theoretical isoelectric point (pI) of 5.43 was determined from the protein sequence in ProtParam and then used to select an appropriate ion exchange resin for the second purification step. Buffer pH was set to 8 to ensure Ycs4 was appropriately charged for interaction with a Mono Q 5/50 anion exchange resin. This resin binds protein in buffers with pH above their isoelectric point, and the small size of the beads comprising the resin allows for higher resolution separation of elution peaks compared to related resins such as sepharose or SOURCE. This high-resolution separation is appropriate for intermediate purifications where large amounts of contaminants have already been removed. Ycs4 was eluted with a gradient up to 1.6 M NaCl, which allows lower affinity contaminants to be removed first before eluting the target protein (figure 3.3 B-C).

The final purification step involved SEC over a superdex 200 10/300 gl column. In addition to removing contaminants this step was selected because it also allows for aggregated or misfolded target proteins to be separated from globular monomeric protein (figure 3.3 D-E).

Insoluble material from the lysis step was collected by centrifugation and resuspended in 8M urea to the same volume as the supernatant in order to check the solubility of the expressed protein. Lysis was deemed to be almost total visually, by 100x oil immersion light microscopy. Comparison of the supernatant to the resuspended pellet (figure 3.3 A) showed that while the absolute amount of protein produced was significant, the majority was appearing in the insoluble fraction and was thus unusable. Continuation of the purification process through to SEC with the remaining soluble fraction also showed that most of the material eluted in the void volume (Figure 3.3 D), suggesting it was either unfolded or solubly aggregated.

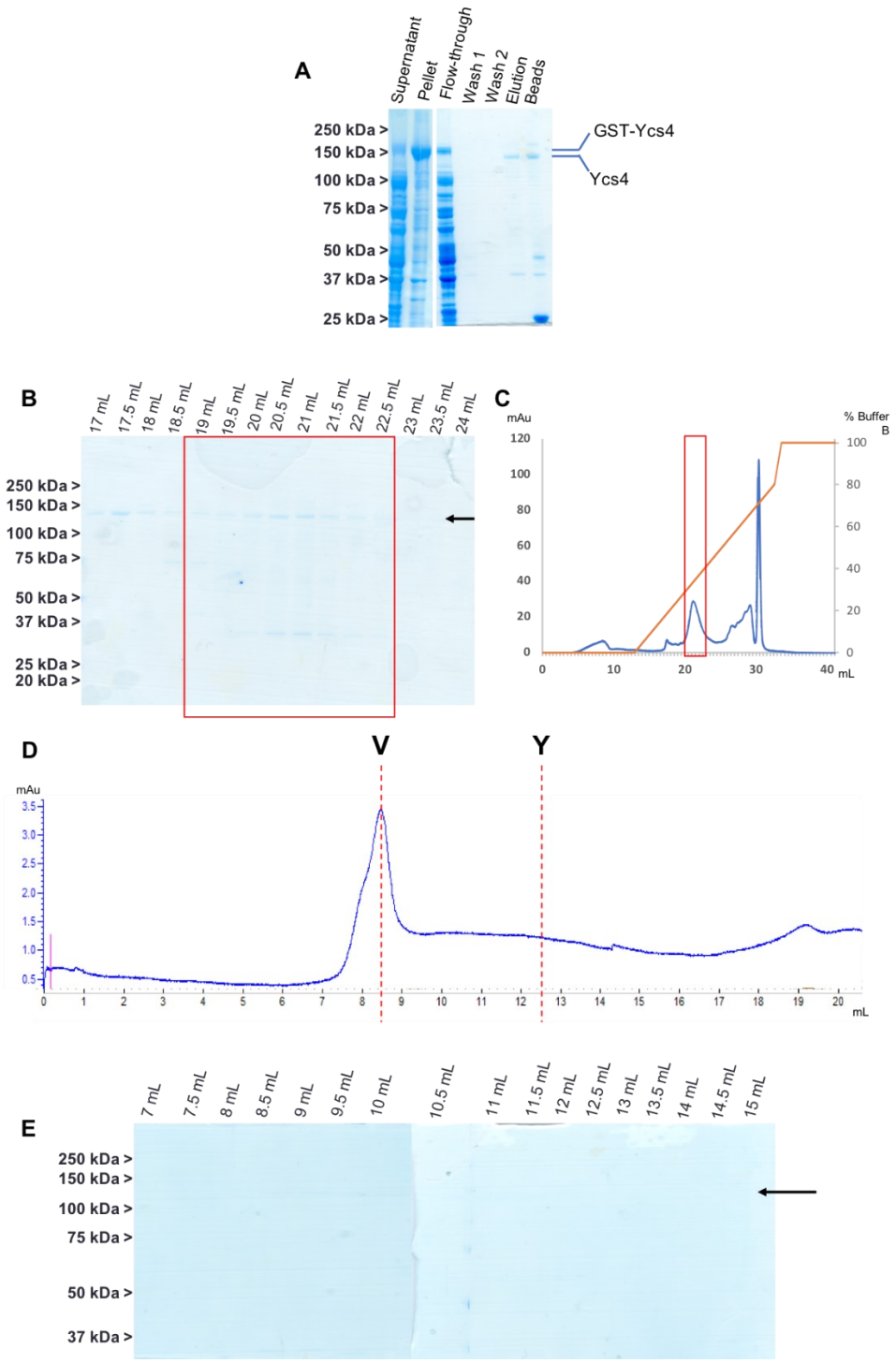


Figure 3.3. Purification of S. cerevisiae GST-Ycs4. (A) Coomasie stained SDS PAGE of glutathione bead pulldown. The elution fraction was subsequently loaded on to a 1 mL MonoQ column for further purification. (B) SDS PAGE of 0.5 mL elution fractions from MonoQ anion exchange column. Red box indicates the fractions pooled for subsequent gel filtration. (C) Chromatogram of MonoQ elution. Red box indicates the fractions pooled for subsequent gel filtration. (D) Elution profile of GST-Ycs4 over superdex 200 10/300 gl. V indicates the column void volume, P indicates the approximate elution volume of GST-Ycs4. (E) SDS PAGE of 0.5 mL SEC fractions.

The same purification strategy was used for the *C. albicans* Ycs4 homolog. As before the theoretical pI (5.05) was calculated using ProtParam, justifying use of the same anion exchange strategy. Based on the elution profile of the *S. cerevisiae* Ycs4, the final gradient was lowered to 1M NaCl over the same length to improve separation of peaks (figure 3.3 C).

As with the *S. cerevisiae* homolog it appeared that a large portion of the expressed protein was insoluble following lysis (figure 3.3 A), and of the soluble material carried through purification process the majority ran abnormally during SEC (Figure 3.4 D-E). Furthermore, although there did appear to be a very small amount of *C. albicans* Ycs4 around the predicted elution volume, as seen by SDS PAGE analysis of peak fractions following SEC, the amount was undetectable by Bradford reagent (figure 3.3 D-E). It would theoretically be possible to scale up purification with a larger starting cell culture to increase yield, however given that the starting culture was already 6 L and a full crystal screen requires 1.5 – 3 mg of pure protein at a concentration of 5 – 10 mg/mL, the amount of starting culture required would likely be impractical

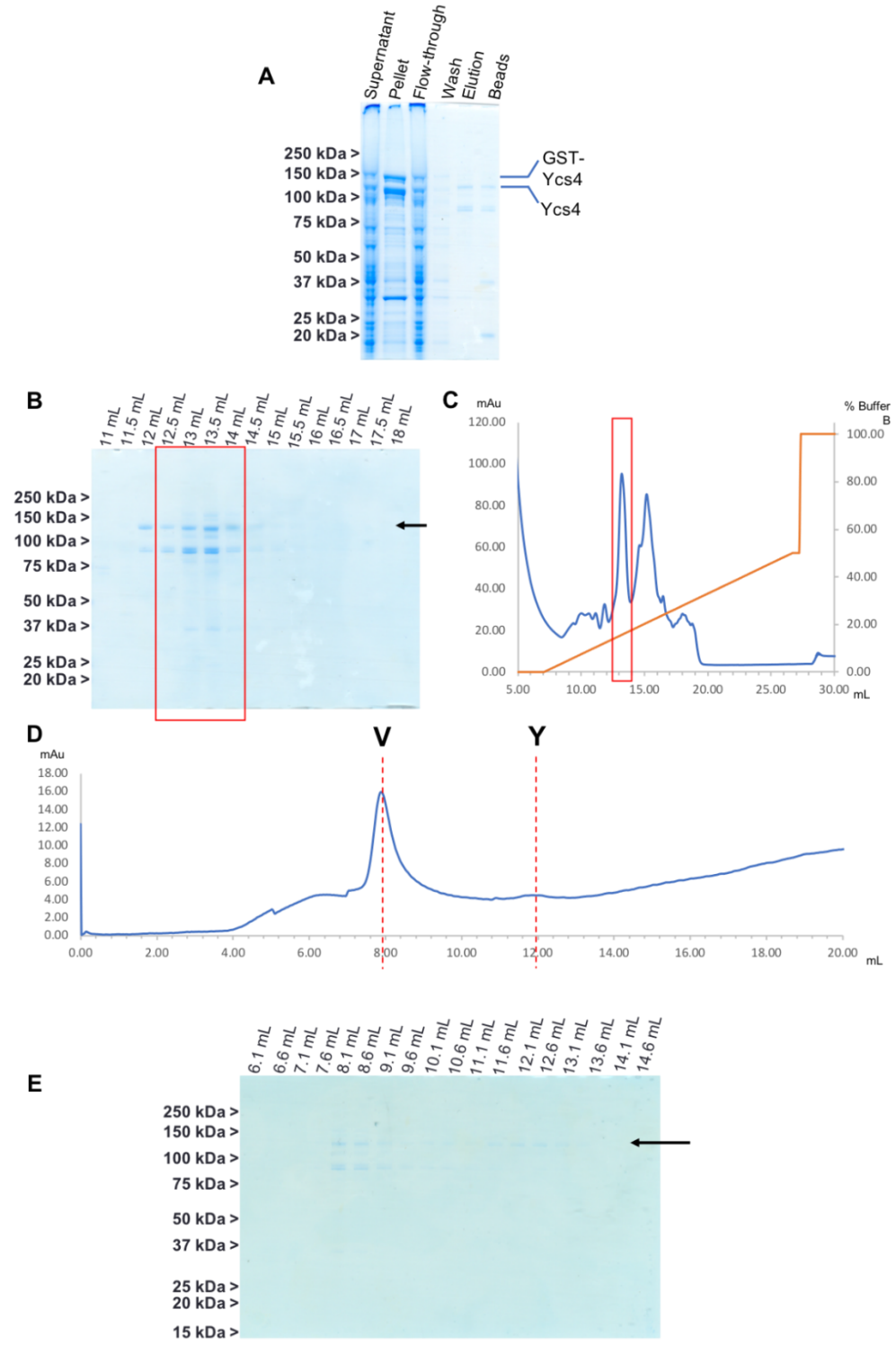


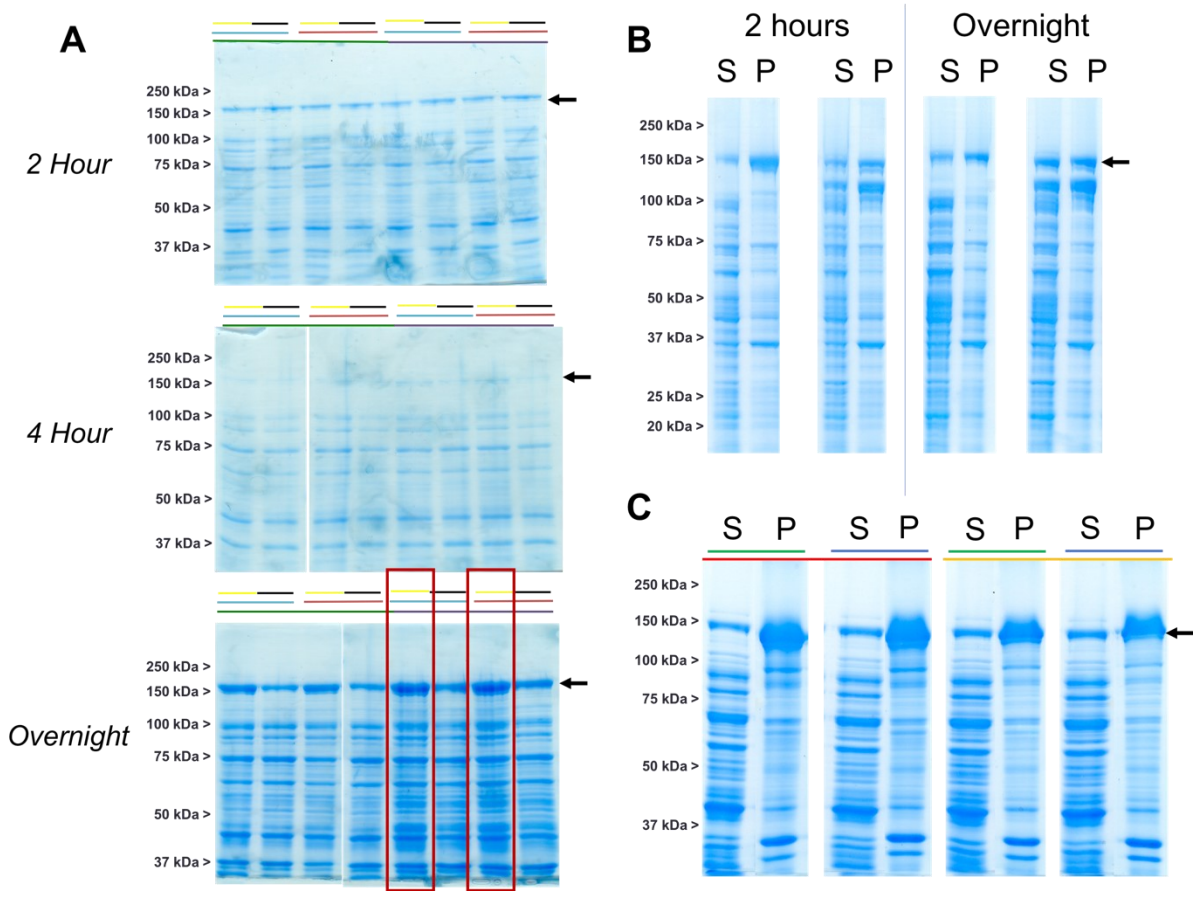
Figure 3.4. Purification of C. albicans GST-Ycs4. (A) Coomasie stained SDS PAGE of glutathione bead pulldown. Elution fraction was subsequently loaded on to a 1 mL MonoQ column for further purification. (B) SDS PAGE of 0.5 mL elution fractions from MonoQ anion exchange column. Red box indicates the fractions pooled for subsequent gel filtration. (C) Chromatogram of MonoQ elution. Red box indicates the fractions pooled for subsequent gel filtration. (D) Elution profile of GST-Ycs4 over superdex 200 10/300 gl. V indicates the column void volume, P indicates the approximate elution volume of GST-Ycs4. (E) SDS PAGE of 0.5 mL SEC fractions.

Given the apparent instability of GST-Ycs4 under the previously tested conditions it was decided that attempting to increasing its solubility during expression would be the next most efficient route for optimisation. It was reasoned that it would be best to increase the chances of recovering more usable protein per litre of starting culture, than to simply scale up. Selected conditions are outlined in figure 3.5.

The *S. cerevisiae* homolog expression vectors were transformed into BL21 (DE3) pLysS or BL21 (DE3) codon + (RIL) strains and expressed at 20^{o}C in either LB or 2xTY media, with variable IPTG concentrations. The pLysS strain was selected to reduce the rate of protein expression and allow proper folding, while the codon + strain was selected in an attempt to increase the overall concentration of expressed protein and thus the concentration of soluble protein per L of culture. Reduction of the expression temperature to 20^{o}C was also done to aid protein stability.

Overnight expression in 2x TY media with 0.5 mM IPTG in the pLysS strain produced the most protein as judged by SDS PAGE of whole cell extracts, and so these conditions were subjected to solubility assays. Cells induced for 2 hours in the same conditions were also carried forward as it was reasoned that the expressed proteins would have had less opportunity to aggregate after a short induction (figures 3.5 A).

Soluble and insoluble fractions were separated by centrifugation following lysis by sonication, with the insoluble material being resuspended in an equal volume of 1x Laemmli buffer. Sampling of these fractions by SDS PAGE for both *S. cerevisiae* and *C. albicans* Ycs4 expression still showed a large amount of protein appearing in the insoluble fractions (figure 3.5 B). Further attempts were made to improve protein solubility of the *S. cerevisiae* protein by including 1% glucose in the growth media to slow expression overnight and allow folding, or by increasing the concentration of NaCl in the lysis buffer to 500 mM to break apart aggregates (figure 3.5 C). As before these conditions had little effect on protein solubility. Insoluble pellets were also resistant to high concentrations of detergent, with white pellets remaining insoluble in 10 – 20% SDS.


A
2 Hour
4 Hour
Overnight
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
B
2 hours
Overnight
S P S P S P S P
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
25 kDa >
20 kDa >
C
S P S P S P S P
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
25 kDa >
20 kDa >

expressed in *E. coli* due to a lack of appropriate chaperones to aid its folding. Secondly, the large (26 kDa) GST tag itself may be responsible for the instability given that *E. coli* is known to have difficultly expressing and folding large recombinant proteins (Baneyx et al., 2004).

In order to determine if the inclusion of the GST tag was affecting the solubility of Ycs4, the codon optimised *S. cerevisiae* Ycs4 sequence was inserted into a pET 30a (+) vector on our behalf by Genescript and transformed into BL21(DE3) Star strain *E. coli.* This insertion replaced the N terminal GST tag and preScission site with a C terminal 6xHis tag. This 6x His tag is much smaller than GST, having a mass of only 0.82 kDa.

The same insert was also subcloned into the viral vector pFastBAC to allow expression in Sf9 insect cells with the same C terminal 6 x His tag. Eukaryotic insect cell expression systems are advantageous in the expression of protein that otherwise mis-fold in bacteria because they express a range of chaperones that can be co-opted to maintain recombinant protein stability.

For the material expressed in *E. coli* an initial expression and pulldown step using cells induced with 0.5 mM IPTG at 37^{o}C for 4 hours was performed. A three-step purification process was used again as before, replacing the glutathione agarose beads of the first purification step with Nickel-NTA beads and eluting with 500 mM imidazole (figure 3.6 A). Initial material from the first elution was produced by Genescript and thawed for further purification by MonoQ anion exchange (figure 3.6 B-C) and size exclusion chromatography. As with the GST tagged protein the majority of detectable material appeared in the void volume following a final SEC step, suggesting protein aggregation or misfolding (Figure 3.6 D-E). Repeats of this purification using a gradient elution from 0 – 500 mM imidazole over a 5 mL HisTrap column in the first step produced similar results.

The same purification process was applied to Ycs4-6xHis expressed in insect cells, again with identical results. The majority of material detected in the final SEC step also appeared in the void volume, suggesting soluble aggregation or misfolding (figure 3.6 F-G).

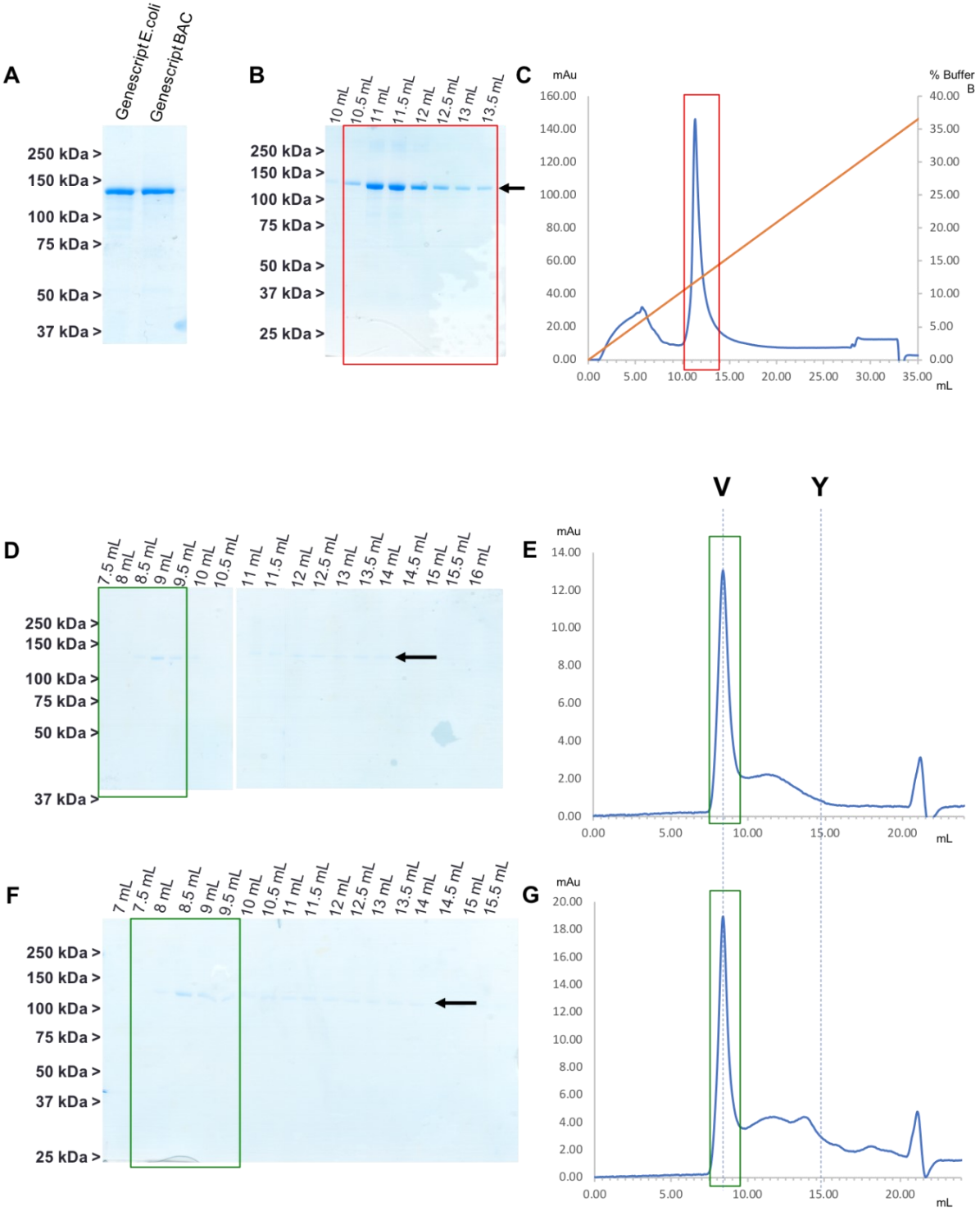

A
Genescript E.coli
Genescript BAC
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
B
10 mL
10.5 mL
11 mL
11.5 mL
12 mL
12.5 mL
13 mL
13.5 mL
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
25 kDa >
C
mAu
% Buffer B
mL
D
7.5 mL
8 mL
8.5 mL
9 mL
9.5 mL
10 mL
10.5 mL
11 mL
11.5 mL
12 mL
12.5 mL
13 mL
13.5 mL
14 mL
14.5 mL
15 mL
15.5 mL
16 mL
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
E
V
Y
mAu
mL
F
7 mL
7.5 mL
8 mL
8.5 mL
9 mL
9.5 mL
10 mL
10.5 mL
11 mL
11.5 mL
12 mL
12.5 mL
13 mL
13.5 mL
14 mL
14.5 mL
15 mL
15.5 mL
250 kDa >
150 kDa >
100 kDa >
75 kDa >
50 kDa >
37 kDa >
25 kDa >
G
mAu
mL

Figure 3.6. Purification of Ycs4-6xHis. (A) SDS PAGE of Nickel column purified material produced by Genescript from E. coli and Sf9 cells (BAC). (B) SDS PAGE of peak fractions following MonoQ 5/50 anion exchange with material produced in E. coli, following 6xHis affinity purification. Red box corresponds to fractions taken from the elution volumes indicated by the red box in chromatogram C. (C) Chromatogram of Ycs4-6xHis, expressed in E. coli, eluting from MonoQ 5/50. Both the absorbance (blue line) and the percentage of buffer B in the elution gradient (orange line) are indicated. Red box indicates the fractionated volume sampled in B and the fractions carried forward for further purification. (D) SDS PAGE of fractions following SEC of Ycs4-6xHis following anion exchange. Arrows indicate the expected molecular weight of Ycs4-6xHis. Green box corresponds to the fractions covering the most abundant peak of protein measured by absorbance in E. (E) Size exclusion chromatogram of Ycs4-6xHis following nickel affinity and anion exchange purification. V indicates the void volume of the superpose 6 increase column used for separation, Y indicates the expected elution volume of Ycs4-6xHis. (F) SDS PAGE of fractions following size exclusion chromatography of Ycs4-6xHis, produced in Sf9 cell transfected with baculovirus (Ycs4-6xHis-BAC). As with E. coli derived Ycs4-6xHis, protein was produced by Genescript and purified by Nickel affinity chroamography followed by anion exchange and pooling of peak fractions (not shown). Arrows indicate the expected molecular weight of Ycs4-6xHis. Green box corresponds to the fractions covering the most abundant peak of protein measured by absorbance in G. (G) Size exclusion chromatogram of Ycs4-6xHis-BAC following nickel affinity and anion exchange purification.

Interestingly, despite failure to elute at the expected volume following SEC, the change to a C terminal 6xHis tag appeared to shift a large portion of expressed protein in the *E. coli* strain to the soluble fraction (figure 3.7). Further analysis of samples taken at various stages in the purification process though, seemed to indicate that despite this initially favourable improvement in solubility, Ycs4 was prone to precipitation over time. Samples taken following imidazole elution from the HisTrap column, and following buffer exchange to remove imidazole for Mono Q anion exchange both showed evidence of insoluble material appearing after overnight incubation at 4ºC (figure 3.7 A-B).

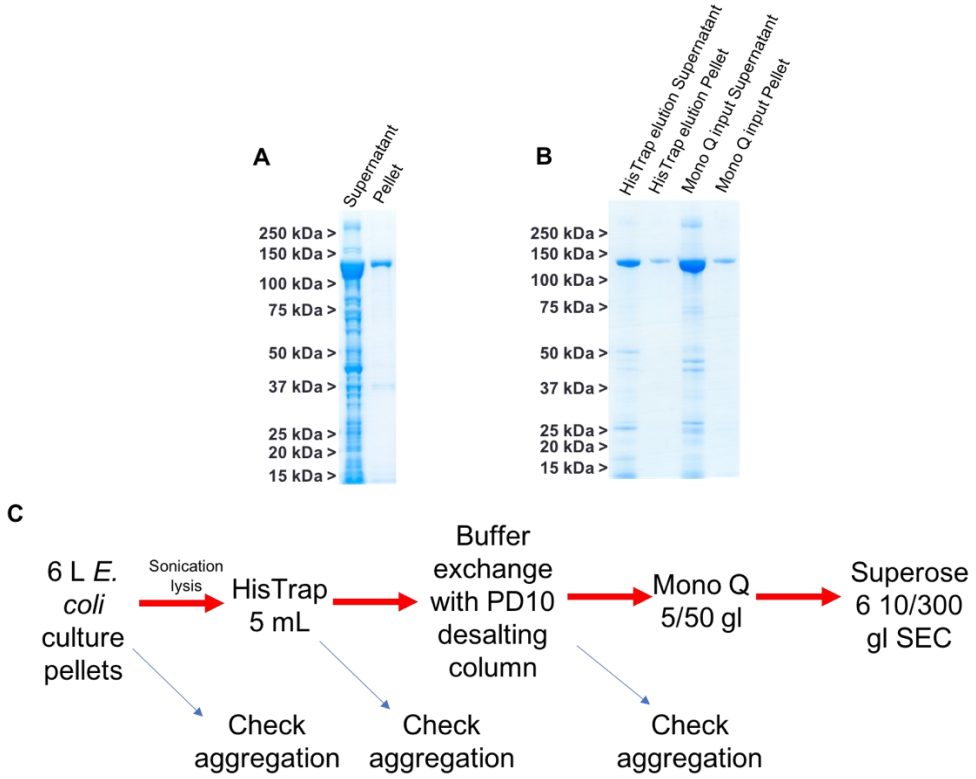


Figure 3.7. Solubility screening of Ycs4-6xHis purification. (A) SDS PAGE analysis of WCE extracts of BL21 (DE3) Star strain E.coli expressing Ycs4-6xHis. Insoluble material was collected by centrifugation and pellets resuspended in an equal volume of 1x Laemmli buffer. (B) SDS PAGE analysis of 500 uL sample taken following elution of Ycs4-6xHis from the HisTrap column with 500 mM imidazole and following buffer exchange with a PD10 desalting column to remove imidazole. Samples were incubated overnight at 4°C, then insoluble material was collected by centrifugation and resuspended in an equal volume of 1x Laemmli buffer. (C) Summary of the initial work flow for the purification of Ycs4-6xHis indicating the points at which protein solubility was checked by centrifugation and SDS PAGE.

We next suspected that a further refinement of the first purification step may remove contaminants and, possibly, early aggregates in such a way as to help maintain the stability of the soluble protein throughout the purification. To this end, the elution step of the Nickel affinity purification, done over a 5 mL HisTrap column, was separated into steps of 20 mM, 50 mM, 100 mM, 200 mM, 300 mM and 500 mM imidazole. Each elution step was maintained at a fixed concentation of imidazole until the UV absorbance had returned to baseline in order to minimise the overlap of different populations of Ycs4 and contaminants in the eluent (figure 3.8 B).

The peak of Ycs4 was seen eluting at a concentration of 100 mM (E1) and 200 mM (E2) imidazole by SDS PAGE (Figure 3.8 A-B). In order to analyse the solubility of Ycs4, 500 uL of both E1 and E2 were individually subjected to SEC over a superpose 6 increase 10/300 gl column without further purification. To best compare the relative amounts of Ycs4 appearing at the expected molecular weight, sample loops were overfilled to produce a homogenous injection.

Comparison of the elution profile of E1 (figure 3.8 C) and E2 (figure 3.8 D) following SEC, clearly showed that the majority of Ycs4 appeared at the expected molecular weight iin E1, with little by way of aggregated or misfolded protein in the void volume. Ycs4 in E2 however largely eluted in the void volume (figure 3.8 F), with a much smaller proportion of material at the expected molecular weight for globular Ycs4. We speculated that this is because soluble aggregates can bind the HisTrap column with multiple tags, requiring higher concentrations of imidazole to remove. Without this initial step these aggregates may act as nucleation points for further aggregation of the remaining monomers prior to the final gel filtration.

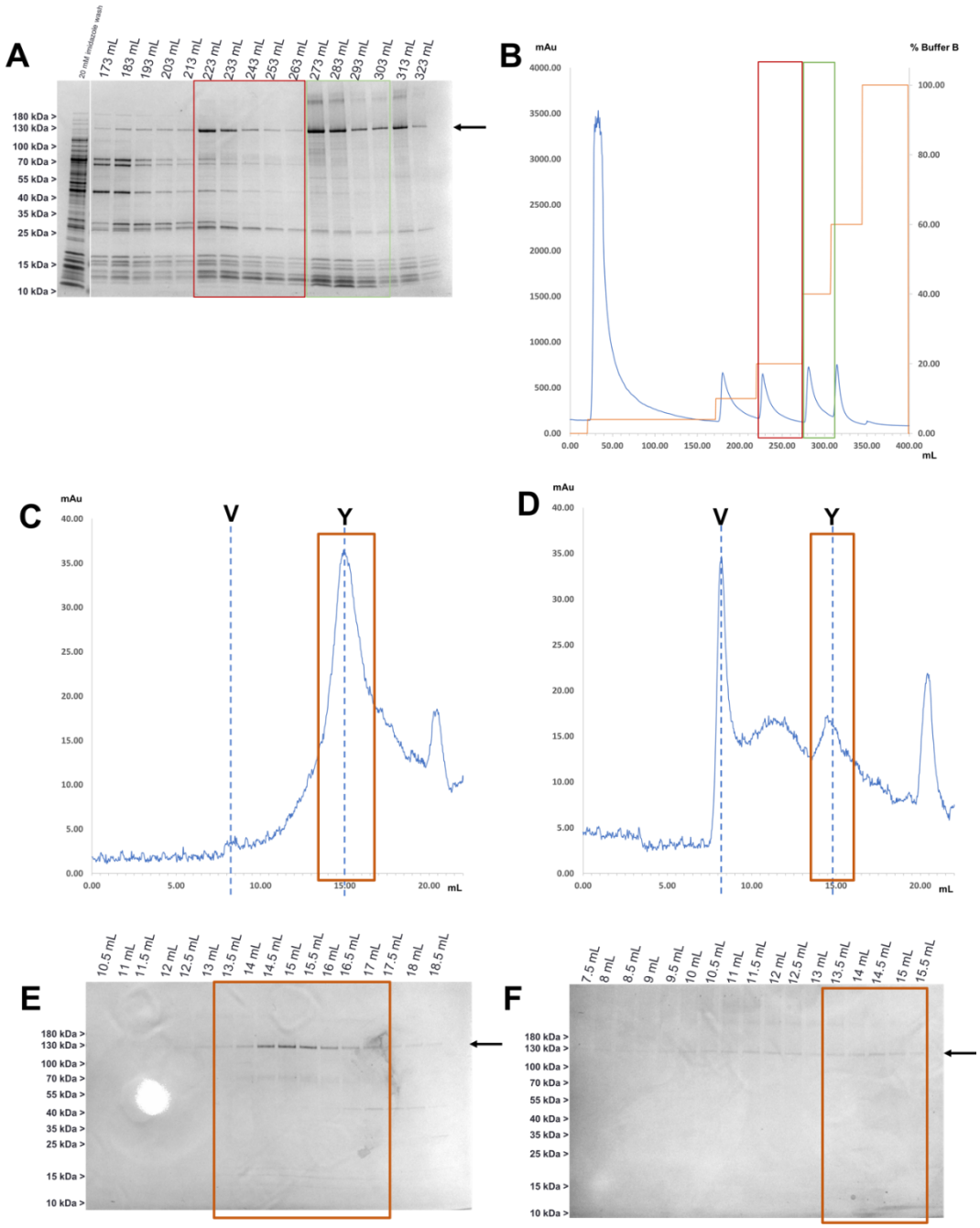


Figure 3.8. Summary of S. cerevisiae Ycs4-6xHis expression and purification. (A) SDS PAGE of peak fractions from 5 mL nickel affinity purification of Ycs4-6xHis expressed in Bl21 (DE3) Star (12L culture). Red box indicates peak fractions eluted at 100 mM imidazole and green box indicates peak fractions eluted at 200 mM imidazole. Black arrow indicates the expected molecular weight of Ycs4-6xHis. (B) Elution profile of the Ycs4-6xHis nickel affinity purification. Red box corresponds to those fractions indicated in A that elute at 100 mM imidazole and green to those that elute at 200 mM imidazole. (C) Elution profile following SEC of pooled fractions eluted at 100 mM imidazole during the nickel affinity purification. SEC was performed over a Superose 6 increase 10/300 gl column. (D) Elution profile following SEC of pooled fractions eluted at 200 mM imidazole during nickel affinity purification. Again, SEC was performed over the same Superose 6 increase 10/300 gl column. (F) SDS PAGE of 0.5 mL fractions eluted during SEC of the 100 mM imidazole elution pool. (G) SDS PAGE of 0.5 mL fractions eluted by SEC of the 200 mM imidazole elution pool.

With an improved first purification step in hand we returned to the previously established 3 step purification process of HisTrap, anion exchange and SEC, this time adding an additional intermediate step following anion exchange over a HiTrap heparin column. This was done to further improve the purity of the final protein.

Figure 3.9 A details the now 4-step purification process. The final successful production of soluble, pure *S. cerevisiae* Ycs4 from 17 L of cultured *E. coli* is shown in figure 3.9 B with the accompanying SEC elution profile demonstrating that this protein was indeed soluble (figure 3.9 C)

Approximately 400 uL of pure protein at a concentration of ~10 mg/mL was produced and subsequently applied to a crystal screen. Unfortunately, in this case no suitable conditions were found for crystal growth, however, with an optimised Ycs4 purification process available we now focused on further adapting it to produce variants of Ycs4 that may themselves crystallise.

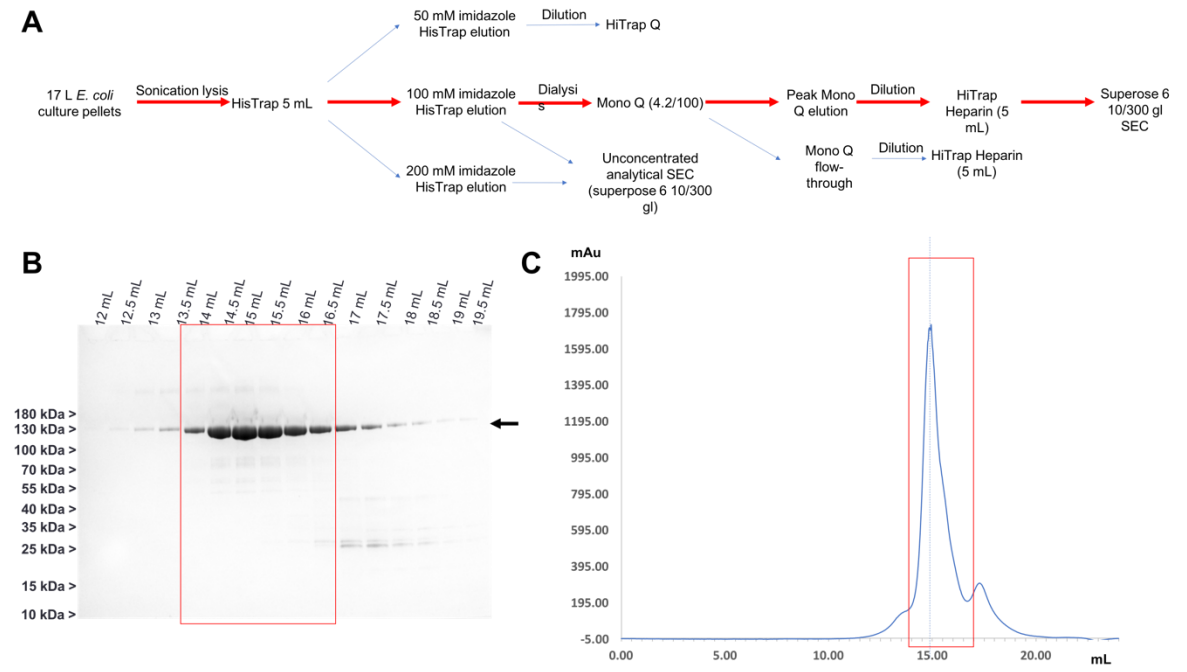


Figure 3.9. Summary of final S. cerevisiae Ycs4-6xHis purification. (A) Summary workflow of the final purification process for S. cerevisiae Ycs4-6xHis. Red arrows indicate the final purification pathway that produced th final material for crystal screening. (B) SDS PAGE of S. crevisiae Ycs4-6xHis eluted in 0.5 mL SEC fractions from superpose 6 increase 10/300 gl. Red box indicates the final fractions pooled and concentrated for further experiments and the black arrow indicates the expected molecular weight of Ycs4-6xHis. (C) Elution profile of S. cerevisiae Ycs4 6x-His. Red box indicates pooled fractions.

3.3 Purification of *C. albicans* Ycs4

Using the now optimised purification process we revisited the *C. albicans* Ycs4, subcloning the codon optimised insert into a pPRO-EX-HTB vector. This process removed the previously attempted N terminal GST tag and preScission protease site and replaced them with an N terminal 6xHis tag with an adjacent TEV cleavage site to allow removal. We suspected that removing the 6xHis tag during purification may increase our chances of successful crystal formation. The same tagging scheme was also applied to the codon optimised inserts for *P. pastoris, H. sapiens* and *S. cerevisiae* Ycs4 which were then test expressed under a single set of conditions from BL21 (DE3) Star strain *E. coli* to gauge their potential (figure 3.10 A). As with the GST tagged version, *P. pastoris* Ycs4 showed little to no expression, though the human homolog did show minor improvement.

As with *S. cerevisiae* Ycs4, a 4-step process of purification was carried out for the *C. albicans* homolog (Figure 3.10 B). An additional overnight digestion step with TEV protease was included between the HiTrap heparin and SEC steps, with a short incubation with loose nickel affinity resin prior to concentration for SEC to remove uncleaved protein. Additionally, the SEC resin used in this purification was changed from superose 6 to superdex 200 increase. This was done to improve the resolution of peaks in the 10000 Da to 600000 Da range.

The final purified *C. albicans* Ycs4 fractions following SEC are shown in figure 3.10 C-D. Sufficient material was produced for crystal screening (~7.5 mg/mL in 300 uL), though with a comparatively lower purity than the *S. cerevisiae* Ycs4. As before purified material was subjected to crystal screening though unfortunately once again no crystals could be produced.

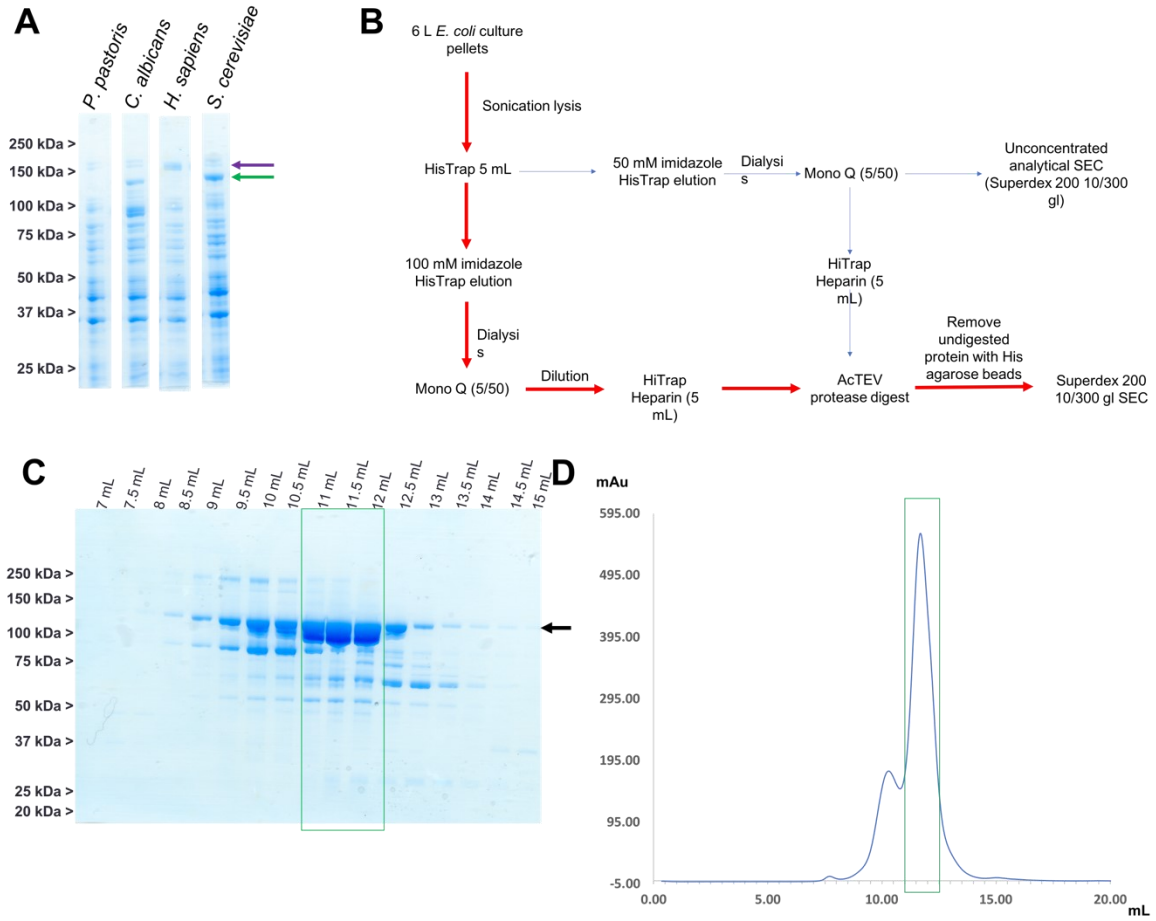


Figure 3.10. Summary of final C. albicans Ycs4-6xHis purification (A) Whole cell extracts of E. coli expressing 6xHis-TEV-Ycs4 from P. pastoris, C. albicans, H. sapiens and S. cerevisiae. Cells were induced with 0.5 mM at 37°C IPTG overnight and 1 mL aliquots lysed in 1% SDS at 95°C. Purple arrows indicate the expected molecular weight of H. sapiens Ycs4, green arrows indicate the approximate molecular weight of all other homologs. (B) Workflow of C. albicans Ycs4 purification. Red arrows indicate the final purification pathway that produced the final material for crystal screening. (C) SDS PAGE of final SEC fractions. Green box indicates those fractions pooled and concentrated for further experimentation. Black arrow indicates the expected molecular weight of Ycs4. (D) Elution profile of C. albicans Ycs4 following SEC. Green box iindicates the same frctions indicated in C.

3.4 Purification of *S. cerevisiae* Ycs4 in Insect Cells

It being the case that two homologs of Ycs4 expressed in *E. coli* had failed to crystallise once pure, further expression and purification of the potentially promising *H. sapiens* and *S. cerevisiae* Ycs4 homologs with N terminal 6xHis was halted. It was reasoned that we could return to these should other approaches also fail. We further speculated that perhaps

misfolding or lack of post translational modifications were preventing crystal formation. Therefore, the 6xHis tagged version of the *S. cerevisiae* Ycs4 was again expressed in 12 L of Sf9 insect cells and subject to the optimised purification process, with minor modifications. These are summarised in (figure 3.11 A).

Following elution from the nickel affinity column at 100 mM imidazole, pooled material was desalted over a PD10 spin column to remove imidazole and loaded onto a HiTrap heparin column. The anion exchange step used previously was forgone at this time in an attempt to streamline the purification process. Elution from this column was separated into 300 mM NaCl pools and 400 mM NaCl. The 300 mM NaCl elution pool was subject to SEC over a superdex 200 10/300 gl column, with peak fractions being recovered and pooled (figure 13B-C). As the purity of the eluted material was somewhat less than that seen previously with *E. coli* derived material, SEC was followed by the missing anion exchange step (figure 13D-E). This final material was collected and concentrated to 6.9 mg/mL in 300uL for crystal screening as with previous purifications. Unfortunately, there was once again no crystal formation when this materia was screened.

The 400 mM NaCl heparin pool appeared to contain several contaminant bands that the 300 mM NaCl pool did not. Therefore, several additional conditions were tried in an attempt salvage this material. Gel filtrations were attempted with either 600 mM NaCl or 150 mM NaCl and 0.1% CHAPS to break away any interacting protein being pulled down with Ycs4. Elution from the gel filtration column remained dirty compared to that of the 300 mM pool and MonoQ anion exchange failed to improve it. This material was therefore discarded.

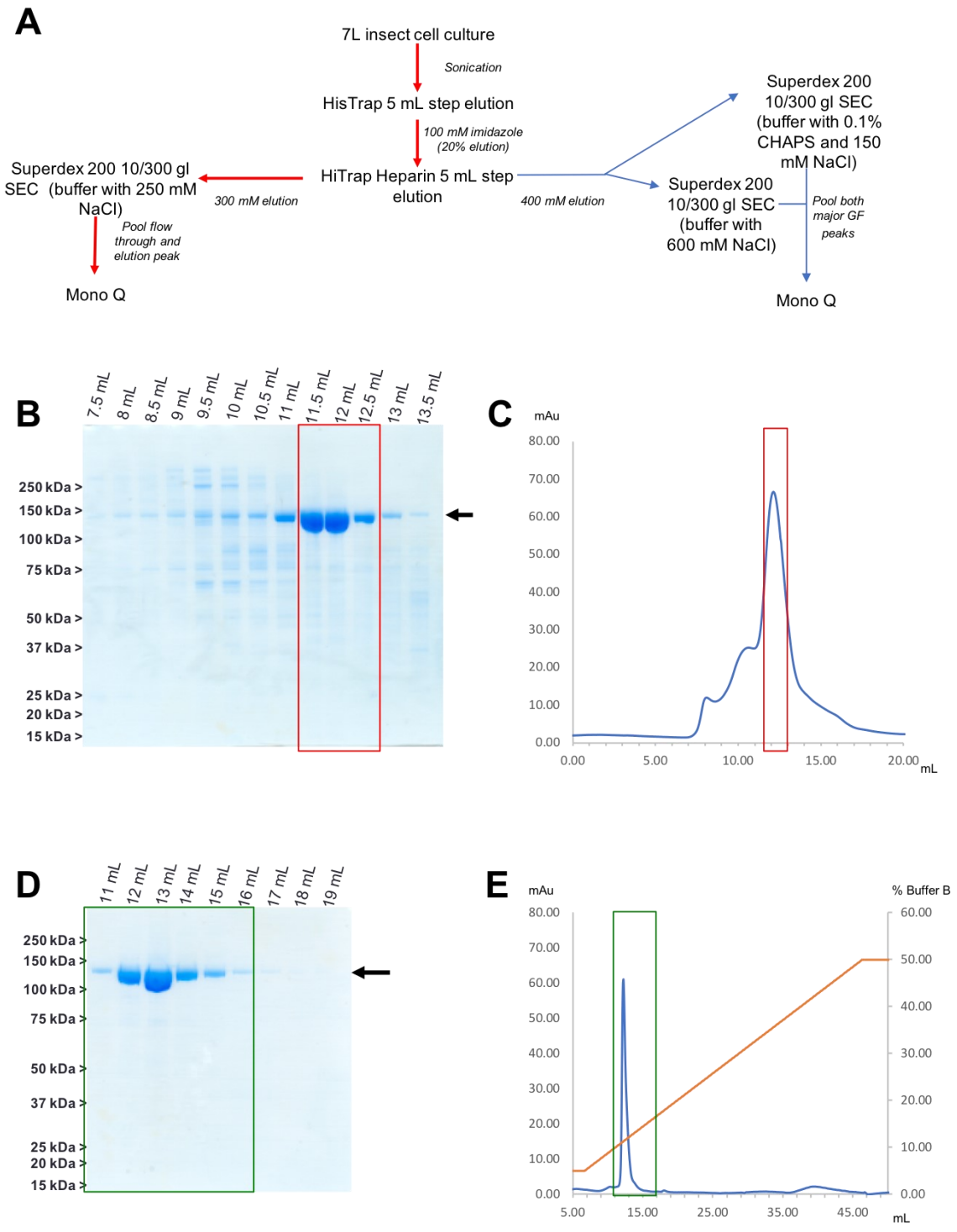


Figure 3.11. Summary of S. cerevisiae Ycs4-6xHis expressed in Sf9 insect cells (A) Summary of the optimisation process undertaken during purification, detailing the purification methods applied to the insect cell lysate. Red arrows indicate the final purification workflow. (B) SDS PAGE showing peak fractions eluted from superdex 200 10/300 gl. Red box indicates peak fractions collected for further purification. (B) Elution profile of Ycs4-6xHis from superdex 200 10/300 gl. Red box indicates peak fractions collected and pooled for further purification. (D) SDS PAGE of peak elution fractions from MonoQ 5/50. Green box indicates fractions collected and pooled for further experimentation. (E) Elution profile of Ycs4-6xHis from MonoQ 5/50. Green box indicates peak fractions collected and concentrated for further experiments.

Chapter 4 – Co-purification of Ycs4 and Condensin Subunit Fragments

4.1 Purification of a Brn1 270 – 423 Fragment

The unsuccessful crystalisation of the individually expressed Ycs4 next prompted an attempt to co-crystallise Ycs4 with known fragments of binding partners in the *S. cerevisiae* condensin complex; specifically, an N terminal fragment of the kleisin Brn1, and the Smc4 head domain. Ycs4's association with Brn1 is well established in the literature (Piazza et al., 2014) and the 4.5 Å structure of the head domain (figure 2.8) seemed to support the association of Ycs4 and Smc4 interacting. It was speculated that by including a binding partner the Ycs4 structure might stabilised, allowing crystal formation.

The fragment of Brn1 to be co-expressed was chosen by comparison to cross linking mass spectrometry data gained from the previously purified ΔYcg1 tetrameric condensin complex (figure 4.1 A and B). Selection of this fragment was further narrowed down by sequence alignments with an N terminal fragment of Brn1 from *C. thermophillum* which had been previous co-purified with Ycs4 (Kschonsak et al., 2017). This alignment was followed by sequence alignments with a further 8 yeast homologs of Brn1 to identify areas of sequence conservation (figure 4.1 C).

The CLMS data appeared to show 3 patches within Brn1 that interacted with Ycs4 (figure 4.1 B). A conservative approach to fragment generation wa taken in order to eliminate as many unstructured regions as possible from the final purification. This led to elimination of patch 1 from the final fragment, as there was very little conservation of this region between Brn1 homologs in addition to no obvious density around bound parts of Ycs4 from the unfinished cryo EM structure. Patch 2 showed a good degree of conservation among Brn1 homologs in addition to strong density on the inner surface of Ycs4. Patch 3 was much the same, also appearing adjacent to an area of high conservation and showing appreciable density. Based on this analysis, Brn1 amino acids 270 -423 were selected to be co-expressed with Ycs4, which encompassed all of patch 2 with an extended N terminus to include the surrounding conserved regions, and the most conserved amino acids from patch 3 (figure 4.1 C).

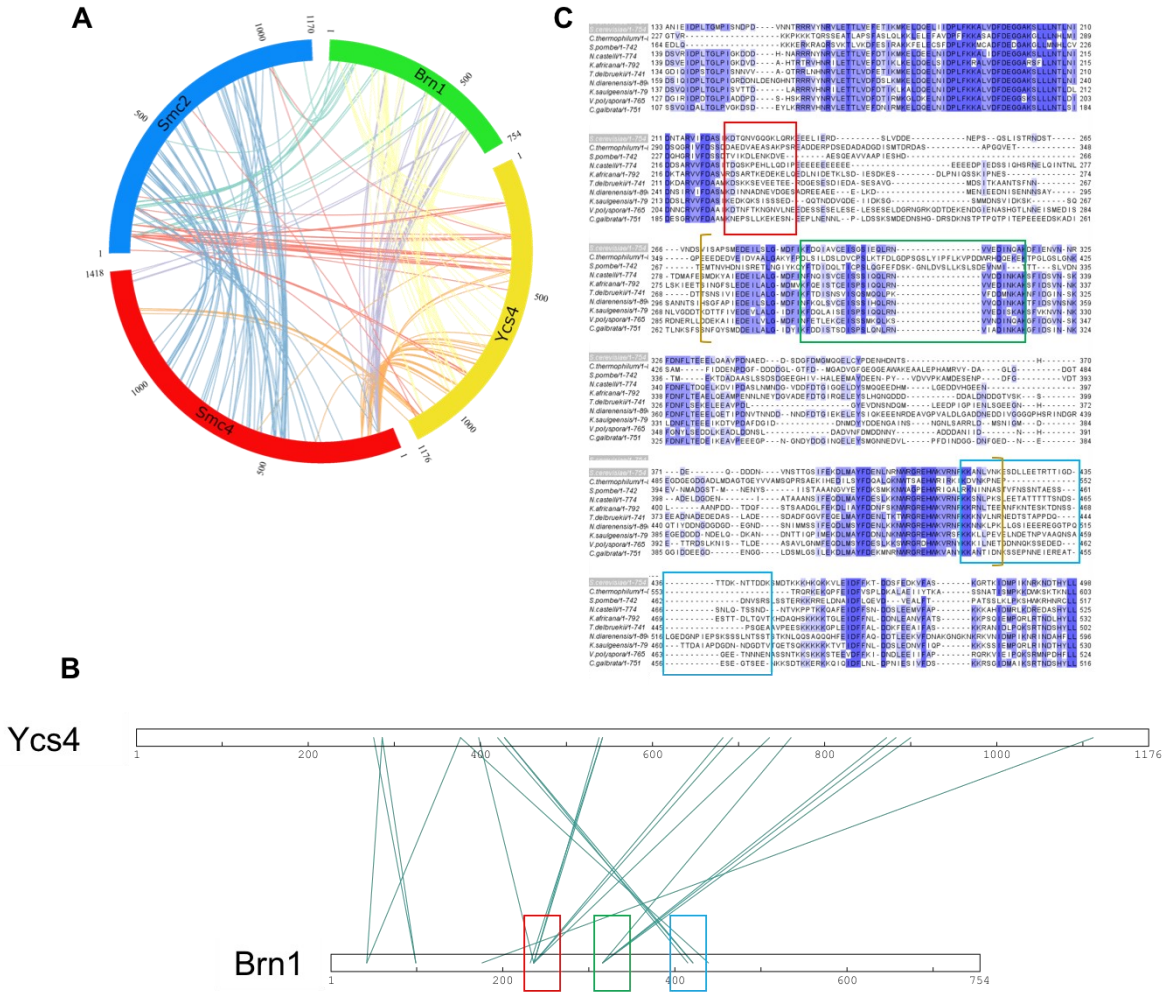


Figure 4.1. Design of Brn1 fragments for purification. (A) Cross-linking mass spectrometry (CLMS) summary showing interactions between subunits of the condensin ΔYcg1 tetramer. (B) Close view of the Ycsg1 and Brn1 subunits and their CLMS interactions with a match score of 13 or higher within the *ΔYcg1 tetramer. Three patches of clustered interaction within Brn1 are indicated by coloured boxes. Red is patch 1, green patch 2 and blue patch 3. (C) Sequence alignments of Brn1 homologs (S. cerevisiae, Chaetomium thermophillum, Schitzosacharamyces pombe, Naumovozyma castellii, Kazachstania africana, Torulaspora delbrueckii, Naumovozyma dairenensis, Kazachstania saulgeensis, Candida glabrata and Vanderwaltozyma polyspora). Alignments correspond to the region identified as interacting with Ycs4 by CLMS in S. cerevisiae. Red box indicates patch 1, green patch 2 and blue patch 3. Orange brackets indicate the S. cerevisiae area selected for C-terminal StrepII tagging and co-expression with Ycs4-6xHis. Produced in collaboration with Dr Byung-gil Lee, with CLMS by Dr. Francis O'Reilly (Rappsilber lab).*

To begin, individual expression of the Brn1 fragment with a C terminal StrepII tag, was attempted. It was speculated that there would be a greater chance of crystal formation if the ratio of Brn1 to Ycs4 could be titrated in the crystal screen. Therefore, expressing and purifying each protein individually would be preferential to co-expression.

The Brn1 270-423 fragment sequence was inserted into a pET-20 (b) vector with a C terminal StrepII tag and transformed into BL21 (DE3) Star and BL21 (DE3) pLysS *E. coli* strains. As with the Ycs4 homologs, expression conditions were varied to optimise protein yield. For both transformed strains, IPTG concentration, media type and induction time were screened (figure 4.2 A).

SDS PAGE analysis of whole cell extracts showed little evidence of robust induction under any condition. However, anti-strep western blotting of BL21 (DE3) Star whole cell extracts following induction with 0.5 mM IPTG for 2 hours, 4 hours or overnight showed protein was detectable, with the best induction occurring overnight (figure 4.2 B).

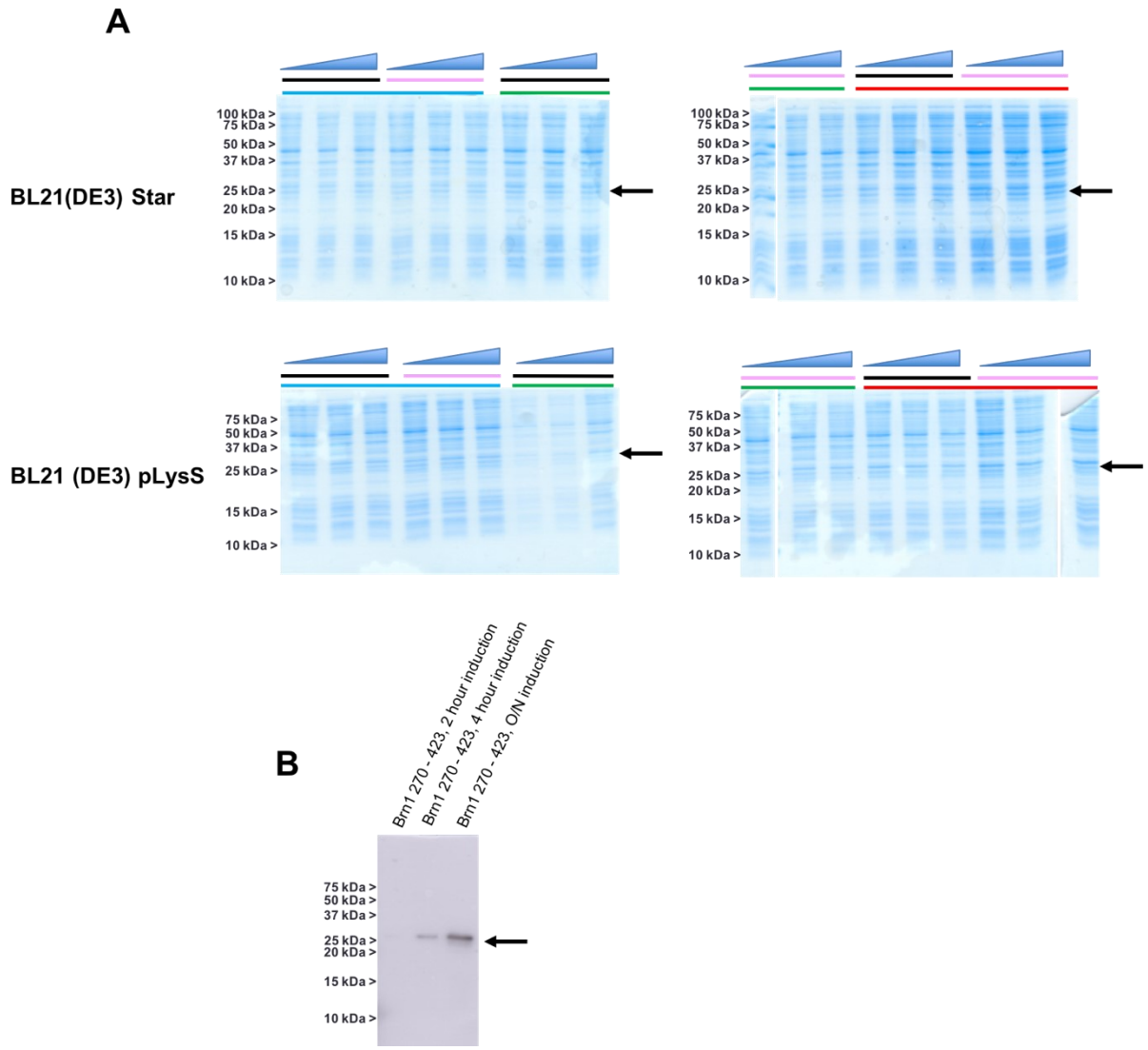


Figure 4.2. Brn1 270-423 fragment expression screening. (A) Test expression screen for S. cerevisiae Brn1 270-423 fragment in Bl21(DE3) Star and BL21 (DE3) pLysS. Black arrows indicate the expected molecular weight of the fragment. Blue triangles indicate increasing IPTG concentrations from 0.2 mM to 05 mM to 1 mM. Bands above lanes indicate conditions used. Black bands indicate samples grown in LB media, pink bands indicate samples grown in Terrific broth. Blue bands indicate cells induced for 2 hours, green bands indicate cells induced for 4 hours and red bands indicate cells induced overnight. (B) Anti-strepII western blot of whole cell extracts from BL21 (DE3) Star cells transformed with Brn1 270-423 fragment vector. Cells were induced with 0.5 mM IPTG in Terrific broth for the indicated time. Black arrow indicates the Brn1 270-423 fragment.

To test the stability of the protein that was induced under these conditions, cell cultures were scaled up to 5L and a 3-step purification process was devised. A preliminary affinity purification step using a StrepTrap 5 mL column was followed by anion exchange over a HiTrap Q column and SEC with a superdex 200 10/300 gl column. The purification

workflow is detailed in figure 4.3 A. Anion exchange was again selected on the basis of the theoretical isoelectric point. This was calculated form the fragment sequence in ProtParam to be 4.22. The HiTrap Q resin was favoured over Mono Q in this instance simply to allow a faster runtime at the expense of resolution. In addition to the standard pH 8 Tris based buffer used for initial purification thus far, a second parallel purification was attempted using a pH 6 MES buffer on the basis of the particularly low pI. It was hoped that bringing the buffer pH closer to that of the pI would help make the anion exchange step more selective, reducing or reversing the charge of some contaminant protein to prevent binding.

Initial pulldown experiments appeared to give unfavourable results, with the majority of the material appearing in the void volume with heavy contamination, though a small population of the Brn1 270- 423 fragment did appear at the approximately correct elution volume (figure 4.3 B-C). Reducing the buffer pH did not appear to have any major effect on either the purity or stability of the Brn1 270-423 fragment.

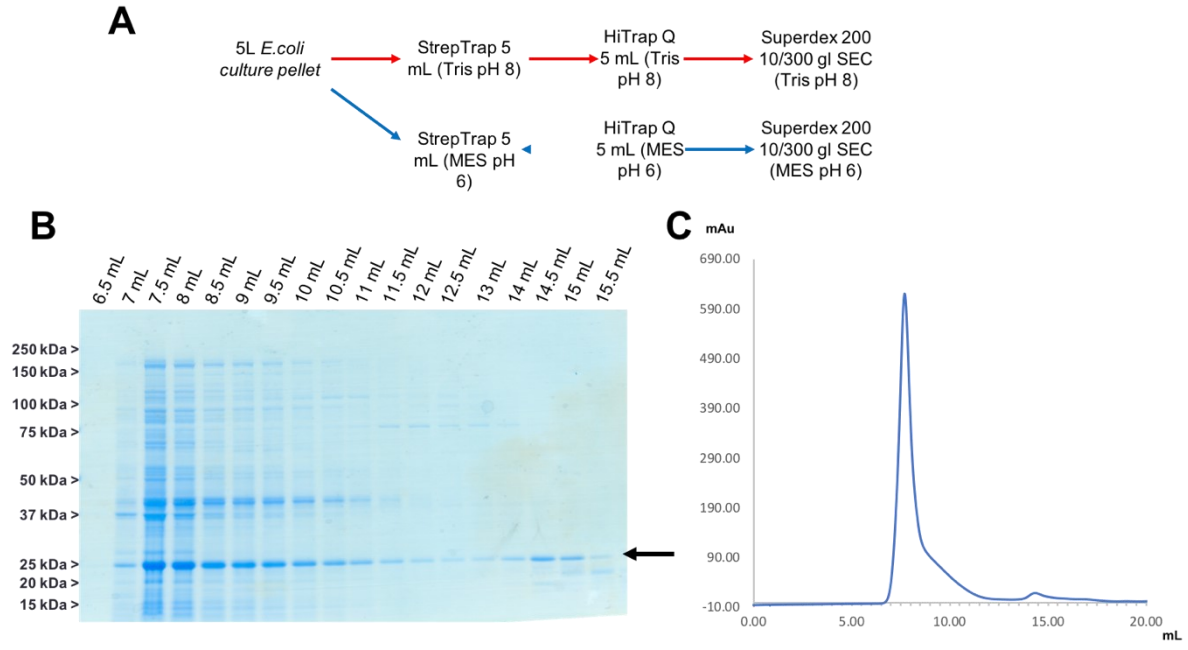


Figure 4.3. Purification of Brn1 270-423. (A) Summary of the workflow during the optimisation of the purification of the Brn1 270-423 fragment alone. Red arrows indicate the path taken to produce final material shown in B and C. (B) SDS PAGE of peak fractions following SEC over a superdex 200 10/300 gl column. Black arrow indicates the expected molecular weight of Brn1 270-423. (C) Chromatogram of Brn1 270-423 SEC corresponding to B.

The poor yield during the purification of Brn1 alone at this juncture prompted us to consider the alternate strategy of co-purifying the fragment with Ycs4. The pET30 (a) *S. cerevisiae* Ycs4-6xhis vector was co-transformed with the Brn1 270-423 fragment vector into BL21 (DE3) Star cells and a 15 L culture was grown and induced overnight with 0.5 mM IPTG. The purification workflow for the co-expressed complex is detailed in figure 4.4 A. The StrepTrap affinity step in this case was followed by a HisTrap affinity step which takes advantage of the 6xHis affinity tag in the co-expressed Ycs4. These two affinity steps were followed by SEC as before. By sequentially pulling down two affinity tags we aimed to not only eliminate contaminant protein but to also ensure recovery of mainly heterodimers as opposed to a mixed population of heterodimers and Brn1 monomers.

The final purified material following SEC is shown in figure 4.4 B with accompanying chromatogram in figure 4.4 C. Analysis of the SEC elution profile appears to show a large peak of material eluting around 10.5 mL (peak 1) with a second smaller shoulder eluting later around 11.2 mL (peak 2). Peak 1 and 2 were separated and fractions pooled, as far as resolution would allow, then submitted for crystal screening as before. The exact stoichiometry of the two peaks is unclear, but comparison of fractions by SDS PAGE seemed to indicate equivalent amounts of both proteins in each pool. It was also unclear if the ratio of Ycs4 to Brn1 270-423 was 1:1 or if there was a higher ratio of Ycs4. The Coomassie stain intensity for Ycs4 was much greater, but concurrently the molecular weight of Ycs4 was also much greater, meaning more mass per molecule to stain. It is possible that these two populations represent alternate conformations of the heterodimer or higher order assemblies of multiple subunits.

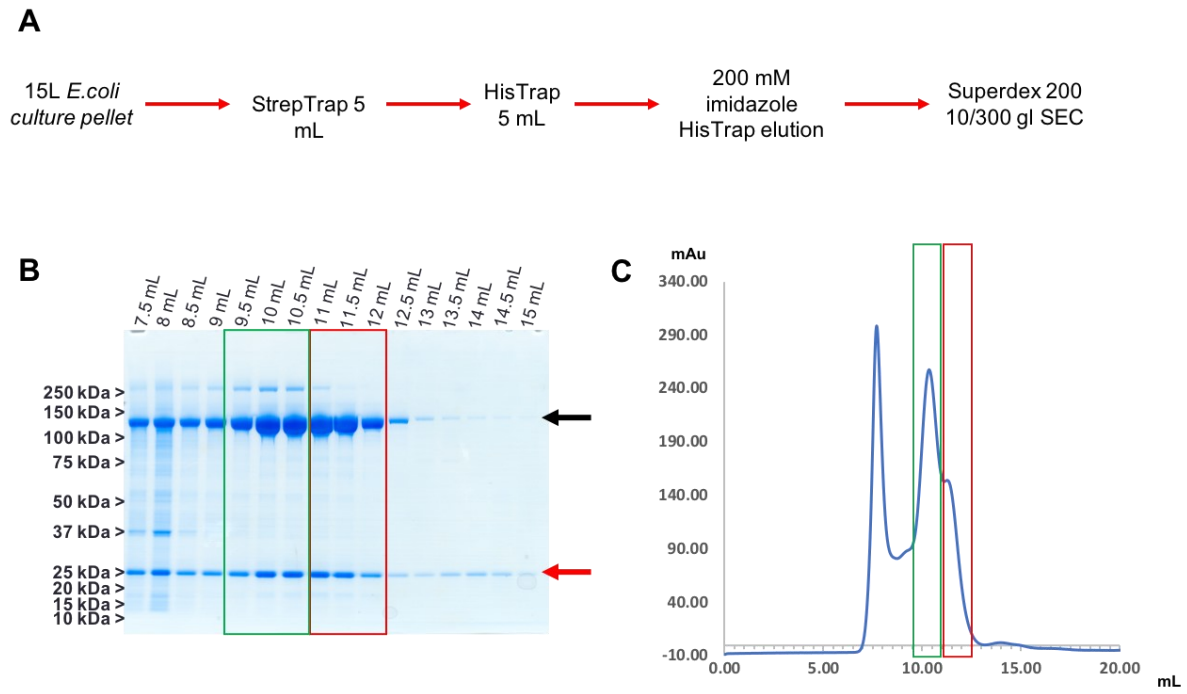


Figure 4.4. Co-purification of Brn1 270-423 and Ycs4-6xHis. (A) Diagram showing the workflow for co-purification of S. cerevisiae Ycs4 and Brn1 270 – 423 Strep. (B) SDS PAGE of final SEC fractions. Black arrow indicates the expected molecular weight of Ycs4-6xHis. Red arrow indicates the expected molecular weight of the Brn1 270-423-Strep fragment. Red box indicates those fractions pooled corresponding to the highest molecular weight peak eluting after the void volume (peak 1). Green box indicates those fractions pooled corresponding to the lowest molecular weight peak eluting after the void volume (peak 2). (C) SEC chromatogram of co-expressed Ycs4-6xhis and Brn1 270-423-Strep over superdex 200 10/300 gl. Red box indicates peak 1 and its corresponding pooled fractions. Green box indicates peak 2 and its corresponding pooled fractions.

Unfortunately, as with all previously purified Ycs4 variants, the Brn1 fragment heterodimer also failed to produce crystals.

4.2 Purification of the Smc4 Head

Both structural evidence (figure 2.8) and CLMS (figure 4.5) supported the notion that the head domain of Smc4 bound to Ycs4. This has since been further substantiated in the literature (Hassler et al., 2019). Therefore, as with the Brn1 270-423 fragment, co-crystallisation of Ycs4 with a fragment of Smc4 head domain was attempted next.

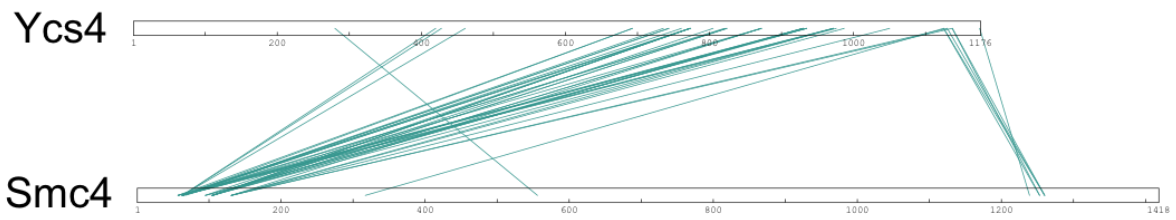


Figure 4.5. Illustration of CLMS cross-links between Ycs4 and Smc4 in the ΔYcg1 tetramer.

Selection of the Smc4 head domain to be expressed was done on the basis of a PSIPRED secondary structure prediction of the N and C terminal domains (figure 4.6 B), in conjunction with a partially complete low resolution cryo EM structure with homologous published crystal structures fitted, to give an estimate of where secondary structure domains might lie (figure 4.6 A).

Figure 4.6 A indicates the approximate cut-off point within the head proximal coiled-coils of both Smc4 N and C terminals that we aimed to produce. The precise residues selected are indicated in relation to the secondary structure prediction in figure 4.6 B, with the corresponding DNA sequences being inserted into a pET20 (b) vector for expression. The N and C terminal halves of the Smc4 head domain were fused by inclusion of a flexible 5 amino acid linker (GSGSG) between the truncated coiled-coils to allow expression as a single polypeptide.

In addition to the core RecA and helical lobes of the Smc4 head domain there is also a large apparently unstructured region at the N-terminus according to secondary structure prediction. A large unstructured domain like this may present a barrier to crystallisation once purified, therefore it was decided to produce two variants of the Smc4 head domain, one with and one without this N terminal region. A C terminal StrepII tag was added to both constructs for affinity purification.

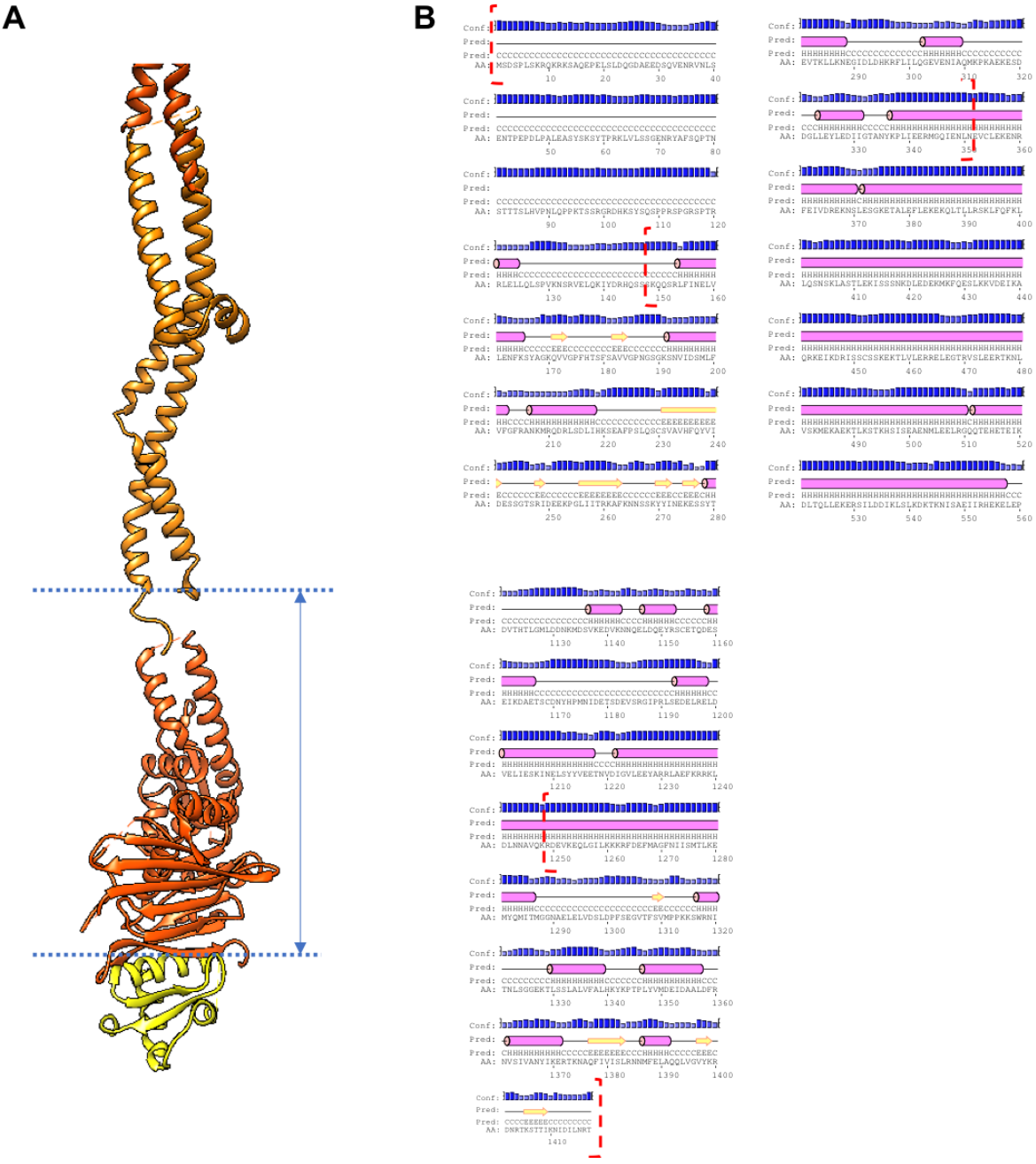


Figure 4.6. Selection of the Smc4 head domain for purification. (A) Incomplete structure of the Smc4 head and head-proximal coiled coil domain fitted with previously solved B. subtilis coiled-coil structures (orange) and S. cerevisiae cohesin head (red). Blue lines indicate the approximate region selected for truncation. (B) Secondary structure prediction of the N and C terminal Smc4 regions encompassing the two parts of the head domain. Red dashed brackets indicate the regions selected for the two constructs. Both a full-length N terminal construct (smc4 hd + N term), and a truncated N terminal construct missing the first 147 amino acids (Smc4 hd) were selected. Both constructs expressed the full C terminal part of the Smc4 head which was linked to the N terminal (full length or truncated) by a 5 amino acid GSGSG linker. Produced in collaboration with Dr Byung-gil Lee.

As with the Brn1 270-423 fragment vectors, the Smc4 head with a truncated N terminal (Smc4 hd) and Smc4 head with full length N terminal (Smc4 hd + N term) vectors were transformed into BL21 (DE3) star and Bl21 (DE3) pLysS *E. coli* strains and test induced in a range of condition (figure 4.7).

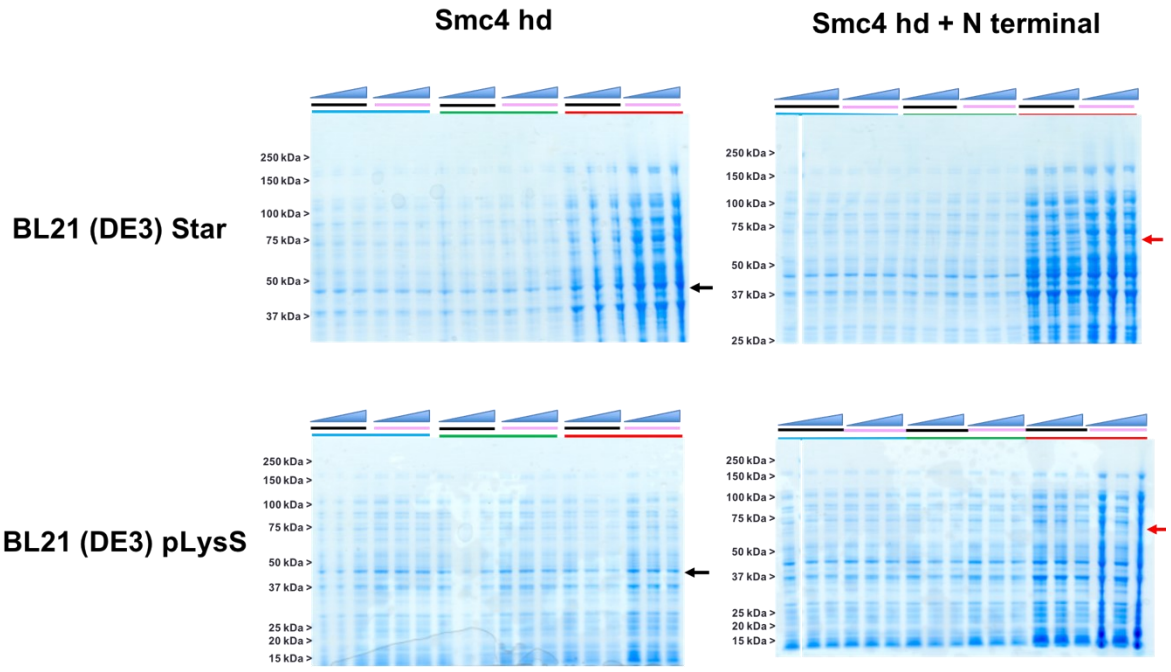


Figure 4.7. Test expression for the Smc4 head and Smc4 head + N terminus constructs in BL21 (DE3) Star and BL21 (DE3) pLysS strains. Coloured bands indicate expression conditions used. Blue bands indicate induction for 2 hours, green bands induction for 4 hours and red bands induction overnight. Black bands indicate cells grown in LB media and pink bands indicate cells grown in Terrific broth. Blue traingles indicate increasing concentration of IPTG used for induction from 0.2 mM to 0.5 mM to 1 mM. Black arrows indicate the expected molecular weight of the Smc4 head only and red arrows indicate the expected molecular weight of the Smc4 head with full N terminal.

SDS PAGE analysis did not appear to show obvious evidence of induction, however, much like the Brn1 fragment, anti-Strep western blotting of whole cell extracts demonstrated the presence of induced protein at the expected molecular weights (figure 4.8 A). Unlike Brn1 however, preliminary purification steps taken with material produced in the conditions assayed by western blot did not recover sufficient material to either justify further scaling up under the same conditions, or eliminate the Smc4 head constructs on the basis of protein

instability (figure 4.8 B). SDS PAGE analysis indicated that very little material could be pulled down by Streptactin beads, though further western blot analysis of small-scale affinity purifications with Streptactin resin did suggest both the Smc4 hd and Smc4 hd + N term constructs produced potentially purifiable protein (figure 4.8 C).

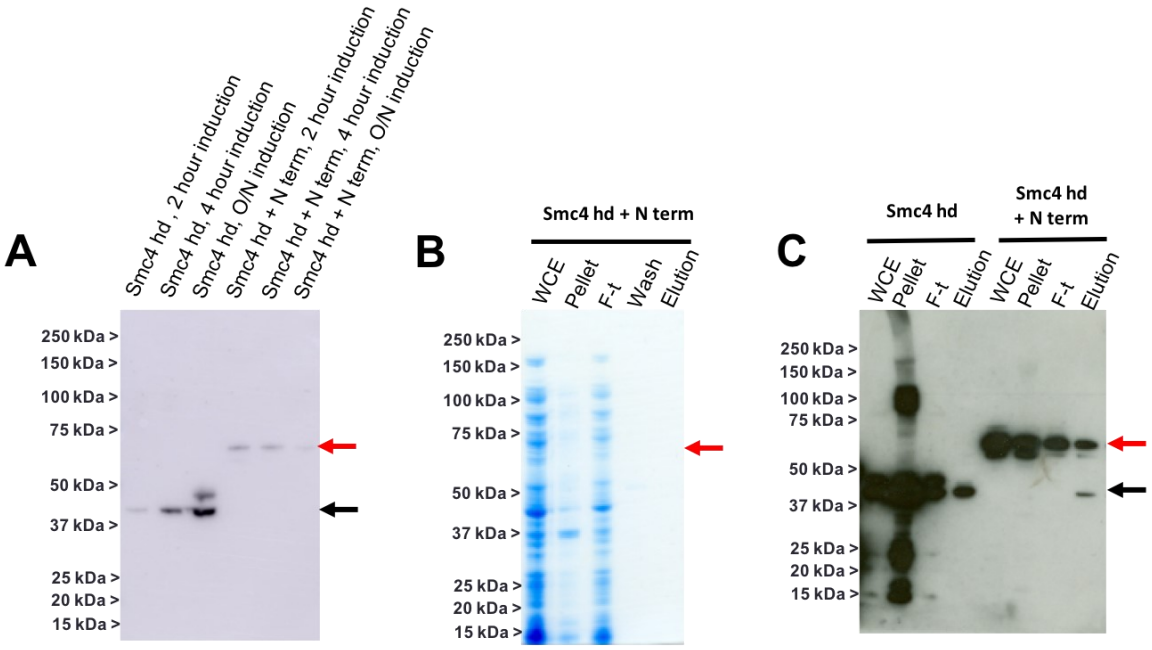


Figure 4.8. Further expression and purification of the Smc4 head. (A) Anti-strepII western blot of whole cell extracts from BL21 (DE3) Star cells transformed with the Smc4 hd and Smc4 hd + N term constructs. Cells were induced with 0.5 mM IPTG in Terrific broth for the indicated time. Black arrow indicates the expected olecular weight of the Smc4 head only and the red arrow indicates the expected molecular weight of the Smc4 head with full N terminal. (B) SDS PAGE analysis of a test pulldown of the Smc4 head with full N terminal with streptactin beads. Red arrow indicates the expected molecular weight of the Smc4 head with full N terminal. (C) Anti-StrepII western blot analysis of Smc4 head and Smc4 head with full N terminal test pulldowns with Streptactin beads.

Based on the observation that protein could be purified, but not in sufficient quantities for crystal screening, we decided to return to optimising the expression steps in order to increase the amount of starting material for purification. For this purpose, 3 further *E. coli* host strains were selected and transformed with both the Smc4 hd and Smc4 hd + N term constructs (figure 4.9).

The BL21 (DE3) Rosetta Gammi B strain aids in folding of expressed proteins by providing an oxidative environment conducive to the formation of stabilising disulphide bridges between cysteines. Rosetta strains have additional copies of key tRNA genes to eliminate

codon bias. Similarly, the Bl21 (DE3) codon + (RIL) strains also have additional tRNA genes, specifically for arginine, isoleucine and leucine. The BL21 (DE3) strain was selected to minimise the burden on the host strains transcriptional machinery, it being the case that there are no additional factors being produced in this strain to aid expression.

Of the strains selected only one showed a marked improvement in the expression level of the Smc4 hd protein. The BL21 (DE3) codon + (RIL) strain produced an observable band at the expected molecular weight indicated in figure 4.9 C.

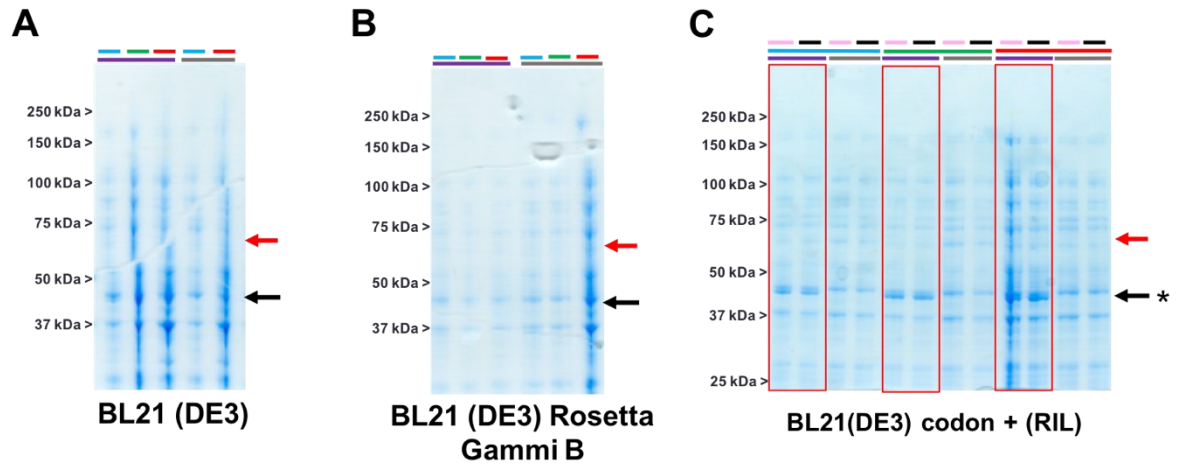


Figure 4.9. Further SDS PAGE analysis of whole cell extracts from BL21(DE3) cells expressing the Smc4 head (purple band) and the Smc4 head with full N terminus (grey band). Blue bands indicate cells induced for 2 hours, green for 4 hours and red overnight. Pink bands indicate cells grown in LB media and black cells grown in Terrific broth. Black arrows indicate the expected molecular weight of the Smc4 head only and red arrows indicate the expected molecular weight of the Smc4 with full N terminal. (A) Whole cell extracts from BL21 (DE3) strains expressing Smc4 heads. (B) Whole cell extracts from BL21 (DE3) rosetta gammi B strains expressing Smc4 heads. (C) Whole cell extracts from BL21 (DE3) codon + (RIL) strains expressing Smc4 heads. Asterisked black arrow indicate the expected molecular weight of the Smc4 head only and the appearance of a positive band by SDS PAGE. Red boxes indicate those conditions in which the positive band appears.

At this time, continuation of the Smc4 head optimisation and purification process would have proceeded as done previously with Ycs4 and the Brn1 270-423 fragment; building on the finding that the Smc4 hd was more robustly expressed with an ultimate goal to either purify the head alone or co-purify with Ycs4 as with the Brn1 270-423 fragment.

However, significant changes in the literature at this time presented an opportunity to complete the condensin structure without the need to further crystalise Ycs4. The crystal

structure of a related homolog of Ycs4 from the themophillic yeast *Chaetomium thermophillum* was published along with fragments of the Brn1 homolog and Smc head domains (Hassler et al., 2019). Thus, the ongoing purification project was discontinued and attention was shifted to analysis of the now complete condensin structure and furthering in vitro analysis of the complex.

4.3. Final Structure of the S. cerevisiae Condensin Complex

Fortunately, due to the high degree of structural conservation, it was possible to fit the published *C. thermophillum* Ycs4 structure to the 4.5A head structure of the originally purified condensin pentamer and ΔYcg1 tetramer determined from cryo EM. Figure 4.10 shows a comparison of the final *S. cerevisiae* Ycs4 structure as determined from the cryo EM structure and the C. *thermophillum* structure from which it was derived (PDB: 6qj3).

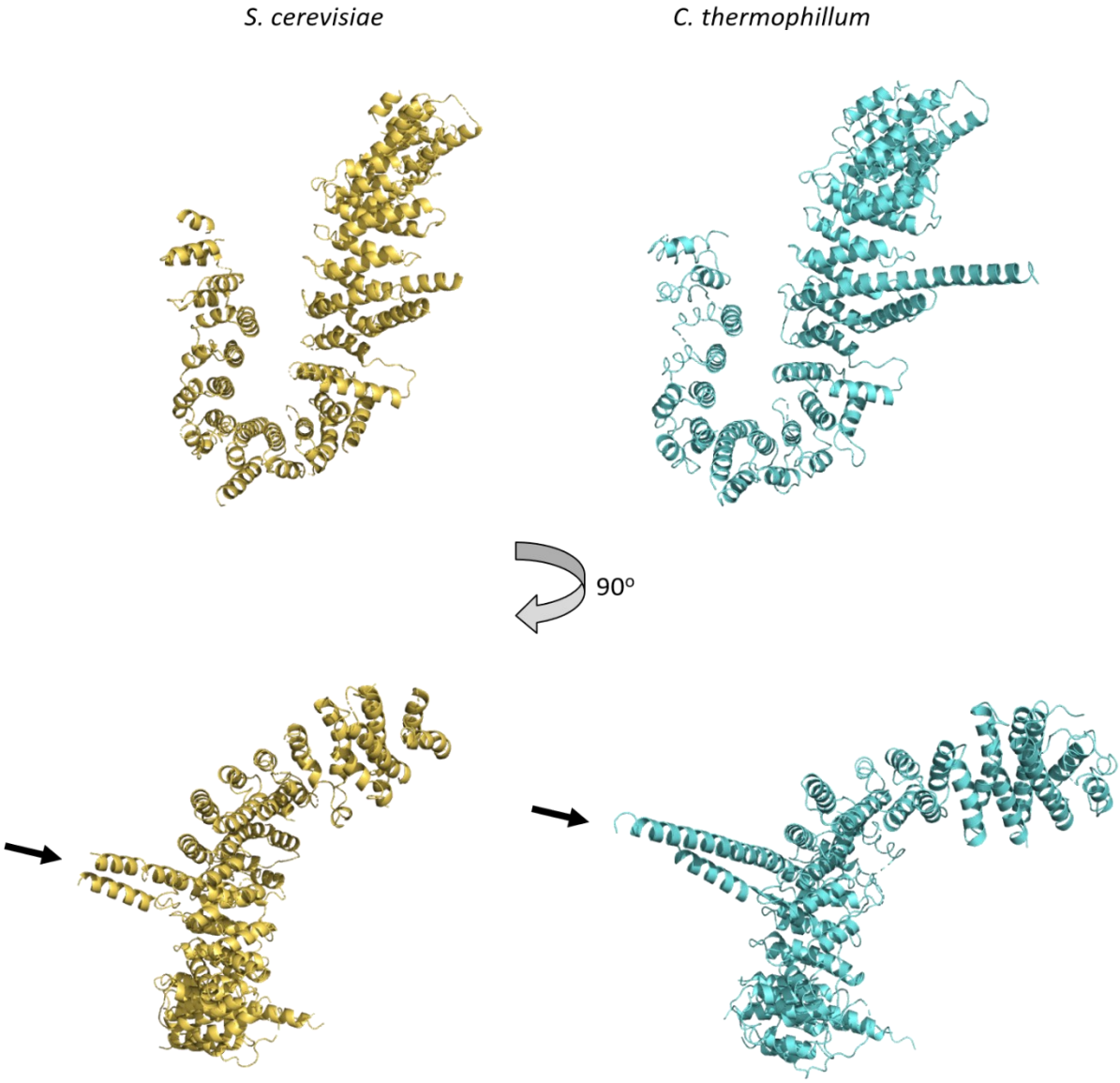


Figure 4.10. Comparison of the final S. cerevisiae Ycs4 structure determined from fitting the C. thermophillum structure to the cryo EM density of the pentameric condensin and ΔYcg1 tetramer, with the C. thermophillum crystal structure from which it was derived (PDB 6qj3, Hassler et al., 2019). Arrows indicate an apparently conserved structural motif comprising an extended "proboscis"-like helix perpendicular to the core solenoidal harp comprising the majority of Ycs4.

With the fitting of the Ycs4 crystal structure to the previously determined density maps of the condensin head domains, our collaborators in the Lowe laboratory were able to generate a final structure for the entire condensin complex (figure 4.11). Further imaging and refinement of head and coiled-coil domains was carried out by the Lowe laboratory during the Ycs4 purification process which increased their resolution to 4.2 and 5.5 Å respectively, with the whole complex resolving at 7 Å. Further domains were fitted with high resolution crystal

structures available in the literature (Soh et al., 2015, Hassler et al., 2019), with further comparison to CLMS data allowing refinement of the final structure, shown in its entirety in figure 4.11 C. With the structure of the condensin Smc2 and Smc4 heads also being solved along with Ycs4, these too could be used to finalise the structure in leu of the original *S. cerevisiae* cohesin head structure found in the literature. Mapping of crosslinks (generated by the Rappsilber lab) onto our final structure and simultaneous structures generated by the Hearing Laboratory reveals only 8% of significant BS3 and 3% of sulfo-SDA cross-links are overlength compared to the cryo EM structure (Lee et al., unpublished). This suggests a fairly homogenous structure in agreement with the CLMS.

Analysis of the final condensin structure reveals several notable features. In the nucleotide free "apo" state, the Smc2 and Smc4 coiled-coils appear to be juxtaposed through the majority of their length. This juxtaposition is confirmed by CLMS of both the pentameric and tetrameric complexes. Furthermore, the coiled-coils also exhibit a hitherto unobserved discontinuity in their length. This bend or "elbow" appears around 2/3^{rds} of the way up the coiled-coil domain, bending through nearly 180^{o} to bring the Smc2/Smc4 hinge domain towards the Smc2/Smc4 head domain and Ycs4. Proximity of the Smc4 hinge to the Smc2 coiled-coil as a result of this bend is confirmed by CLMS data, indicating that this bending at the elbow is present in solution as well as on cryo EM grids. The distance between the closest hinge and coiled-coil residues is ~4-5 Å. This closely proximity strongly suggests an interaction, though further work will be required to confirm it.

Previous work by (Burman et al., 2019) predicted the presence of an elbow in the condensin complex based on similar observation in the *S. cerevisiae* cohesin Smc proteins and *E. coli* MukB. This prediction was based on conservation of a discontinuity in the predicted coiled-coils around residues 360-420 and 760-830 in Smc2. Similar predictions were made between approximate residues 510-570 and 960-1030 in Smc4. This prediction is in good agreement with our actual observed elbow discontinuities between Smc2 residues 384-399 and 785-792, and Smc4 residues 538-544 and 968-973. While a discontinuity may be conserved there is very little sequence conservation between any of the structurally determined elbows (including the condensin elbow presented here) or those further predicted by Burmann et al., 2017 (*S. cerevisiae* Smc5, Smc6, *B. subtilis* Smc and *Mycoplasma alligatoris* Smc)

In addition to the elbow, the coiled-coils of both Smc subunits exhibit another less dramatic discontinuity about 1/3rd of the way between the head domains and the elbow, labelled here as the "joint". Again, as with the elbow, this joint domain is conserved in both MukBEF and cohesion, with predictions by (Burmann et al., 2019) largely agreeing with structural observations. As with the elbow the condensin joint is not sequentially conserved between Smc proteins (Smc2 residues 205-255 and 930-990, Smc4 residues 370-420 and 1140-1190).

The Ycs4 position is in good agreement with previously described structures of the head domains only (Hassler et al., 2019), binding through the W-loop to the Smc4 head. Ycg1 on the other hand appears to sit flexibly below the main body of the complex in the apo state, anchored by a flexible length of Brn1 which keeps it in contact with the complex through binding of the C-terminal to the cap region of the Smc4 head and the N terminal to head proximal coiled coils of Smc2. As expected, the Smc2 and Smc4 heads, while in proximity, are unengaged in the absence of ATP.

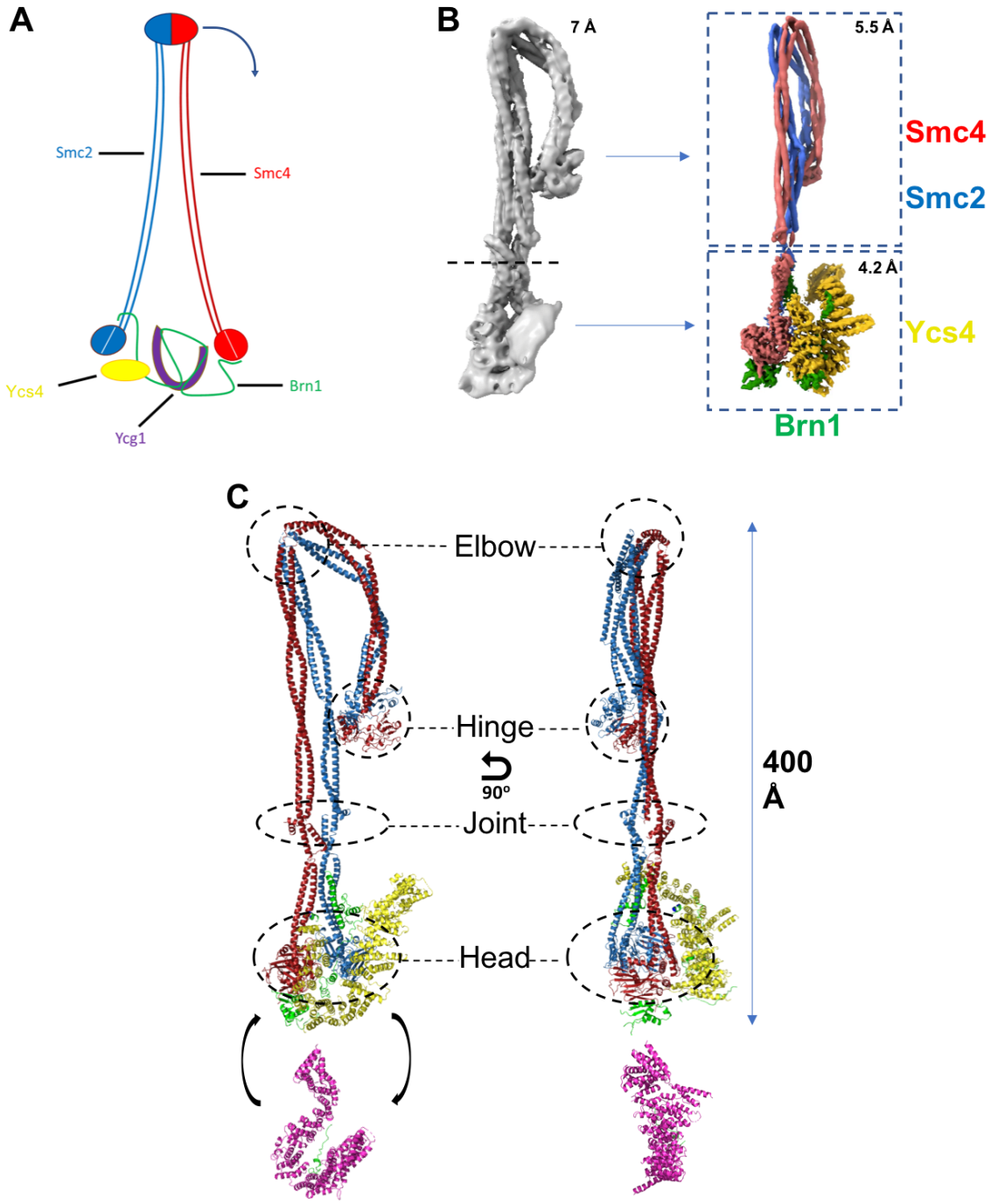


Figure 4.11. Final structure of the condensin complex. (A) Cartoon representation of the condensin complex based on previously published structural data in the literature. (B) 7A model of the ΔYcg1 condensin tetramer prior to crystal structure fitting and highest resolution images of the coiled-coil-hinge domain and the head-Ycs4 domain. (C) Cartoon ribbon diagram of the entire condensin complex following crystal structure fitting. Ycg1 is shown as sitting below the head domain of the complex, attached via unstructured parts of Brn1 (not shown). Produced in collaboration with Dr Byung-gil Lee.

Chapter 5 – Characterisation of Condensin

5.1. Condensin DNA binding

It has previously been shown that Ycg1 and its homologs in *C. thermophillum*, *S. cerevisiae* and human condensin complexes contain a DNA binding pocket that is essential for the proper function of the condensin complex in vivo (Kschonsak et al., 2017, Hara et al., 2019). We speculated that by removing the Ycg1 subunit, the affinity of the condensin complex for DNA would be reduced.

To test this hypothesis, we performed Electrophoretic Mobility Shift Assays (EMSA) with both the condensin WT pentamer and the ΔYcg1 tetramer using 35 bp oligomers of DNA in the presence and absence of ATP to test the complexes binding affinity.

Loss of the Ycg1 subunit caused a significant decrease in the ability of condensin to bind DNA. This loss is not affected by addition of ATP, nor is binding of the pentamer. This suggests that the binding does not represent a topological interaction powered by the ATP hydrolysis cycle. Note however that ATP was not included in the running buffer for the 2 mM ATP EMSA raising the possibility that complexes may lose any ATP dependant binding as they move into the resolving gel.

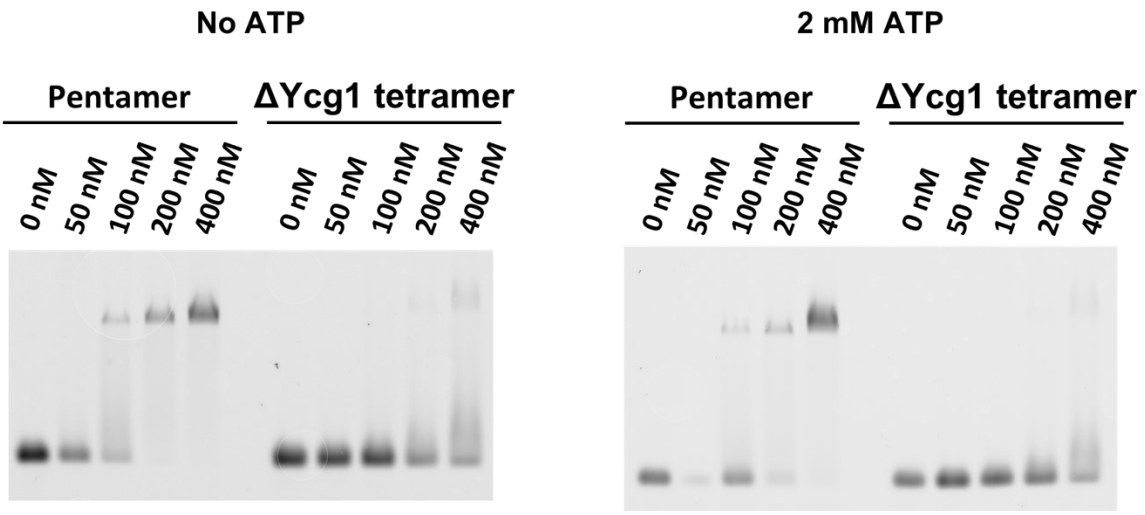


Figure 5.1. Electrophoretic mobility shift assay of pentameric and ΔYcg1 tetramer complexes with 50 nM of fluorescently labelled 35 bp DNA. Complexes were incubated for 10 minutes in EMSA buffer with or without 2 mM ATP before being resolved on a 1 % seakem agarose gel.

5.2. Ycs4 DNA Binding

Given the structural similarities between Ycs4 and Ycg1, we speculated that perhaps Ycs4 could also act as another DNA binding site. Analysis of the Ycs4 position in the final crystal structure (credit to our collaborators in the Lowe laboratory) suggested a potential route for this binding in that a DNA molecule running along the concave surface of Ycs4 parallel to the conserved "proboscis" motif (figure 4.10) could fit through a DNA sized space found between the head proximal coiled coils of Smc2 and Smc4 in the ATP free apo state. These two pieces of information prompted us to assay our purified Ycs4 homologs and fragmentary sub-complexes using electrophoretic mobility shift assays (EMSA) to gauge whether they possessed DNA binding capacity and if so, with what affinity.

EMSA analysis (figure 5.2) revealed that *S. cerevisiae* Ycs4 alone weakly binds DNA at sufficiently high concentrations, as does the *C. albicans* homolog, although the latter to a much lesser degree. However, comparison of the complex alone to the Ycs4-Brn1 270-423 fragment heterodimer DNA mobility shift suggested that even this modest binding capacity is completely abrogated in the preence of the Brn1 fragment. This is true for either of the two populations found during purification by SEC.

Given the relatively small amount of condensin needed to induce a complete mobility shift of DNA compared to Ycs4 it is likely that the observed shift is due to non-specific binding by excess protein or contaminants. It may be the case that the association of the Brn1 fragment stabilises a flexible structure which may be capable of binding non-specifically. The notion of Brn1 stabilising HEAT repeat proteins has been previously demonstrated in Ycg1, whereupon it is seen to stiffen the structure (Manalastas-Cantos et al., 2019).

While we cannot conclusively rule out a role for Ycs4 in binding DNA, these observations suggest that under our tested conditions it does not bind DNA while associated with its binding partner Brn1, as it would be *in vivo*. It maybe that a much greater length of Brn1 is needed to form a closed DNA binding structure as is seen with Ycg1, though this remains to be determined experimentally.

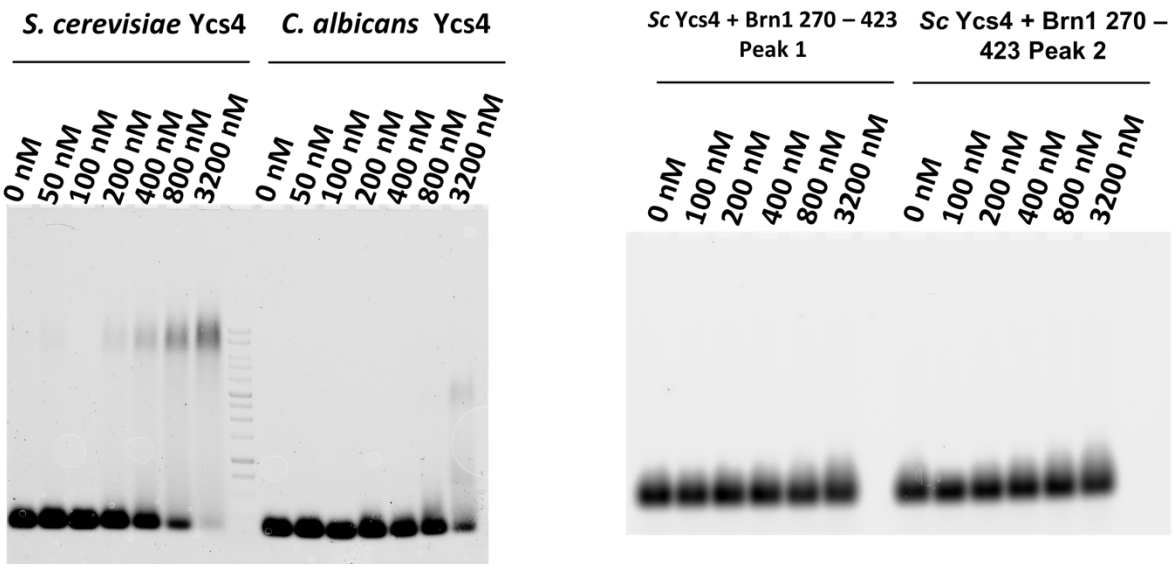


Figure 5.2. Electrophoretic mobility shift assays of condensin and Ycs4 homologs. Increasing concentrations of purified proteins were incubated with 50 nM of fluorescently labelled 35 bp DNA for 10 minutes at room temperature then imaged with fluorescent excitation at 473 nm after resolution on 1% seakam agarose gel.

5.3. Single Molecule Assays with Condensin

A number of recent publications have demonstrated the characteristic behaviour of condensin in single molecule systems. These publications used a range of techniques including magnetic tweezers, DNA curtains and fluorescence imaging of tethered DNA substrates, to demonstrate the ability of condensin to compact DNA substrates and form looped structures *in vitro* (Ganji et al., 2018, Terakawa et al., 2017, Eeftens et al., 2018, Kim et al., 2019).

Using these experiments as a guide we aimed to further characterise condensin using optical tweezers. In collaboration with the Single Molecule Imaging Group (Imperial College London) we sought to apply our purified recombinant condensin complex to the Lumix C-Trap model of optical tweezers. Initially we sought to confirm the activity of our complex and, utilising the unique features of the C trap, develop new assays for further characterisation of condensin mechanisms.

The optical tweezer system has several advantages over other single molecule approaches, which make it ideal for biophysical analysis of condensin. The set-up is summarised in figure 5.3. In brief, the system uses a pair of infrared lasers aimed parallel into a transparent flow cell, which can be used to trap microscopic polymer beads at their focal point. These beads, when coated with suitable ligands such as streptavidin or anti-digoxigenin, can be used to capture tagged DNA. The first advantage of the optical tweezer system is that not only can this individual piece of captured DNA be imaged through use of a confocal microscope above the flow cell, but also by moving the trapping lasers, forces can be applied to and measured on the same piece of DNA.

DNA is readily tagged with intercalating dye such as sytox which can be used to assay protein induced structural changes on the DNA. Simultaneously fluorescently tagged proteins can also be imaged directly allowing observations of protein DNA interactions and behaviour. Addition of these proteins can be finely controlled by the optical tweezers' second great advantage: the flow cell. In this system a series of channels are created parallel to one another within the flow cell by laminar flow, allowing up to 5 different conditions to be present in the same experiment. Often channels 1 and 2 are occupied by a solution of beads and DNA respectively. Movement from one condition to the other is as simple as moving the

trapping lasers, allowing for instant changes in buffer conditions, protein composition and co-factors.

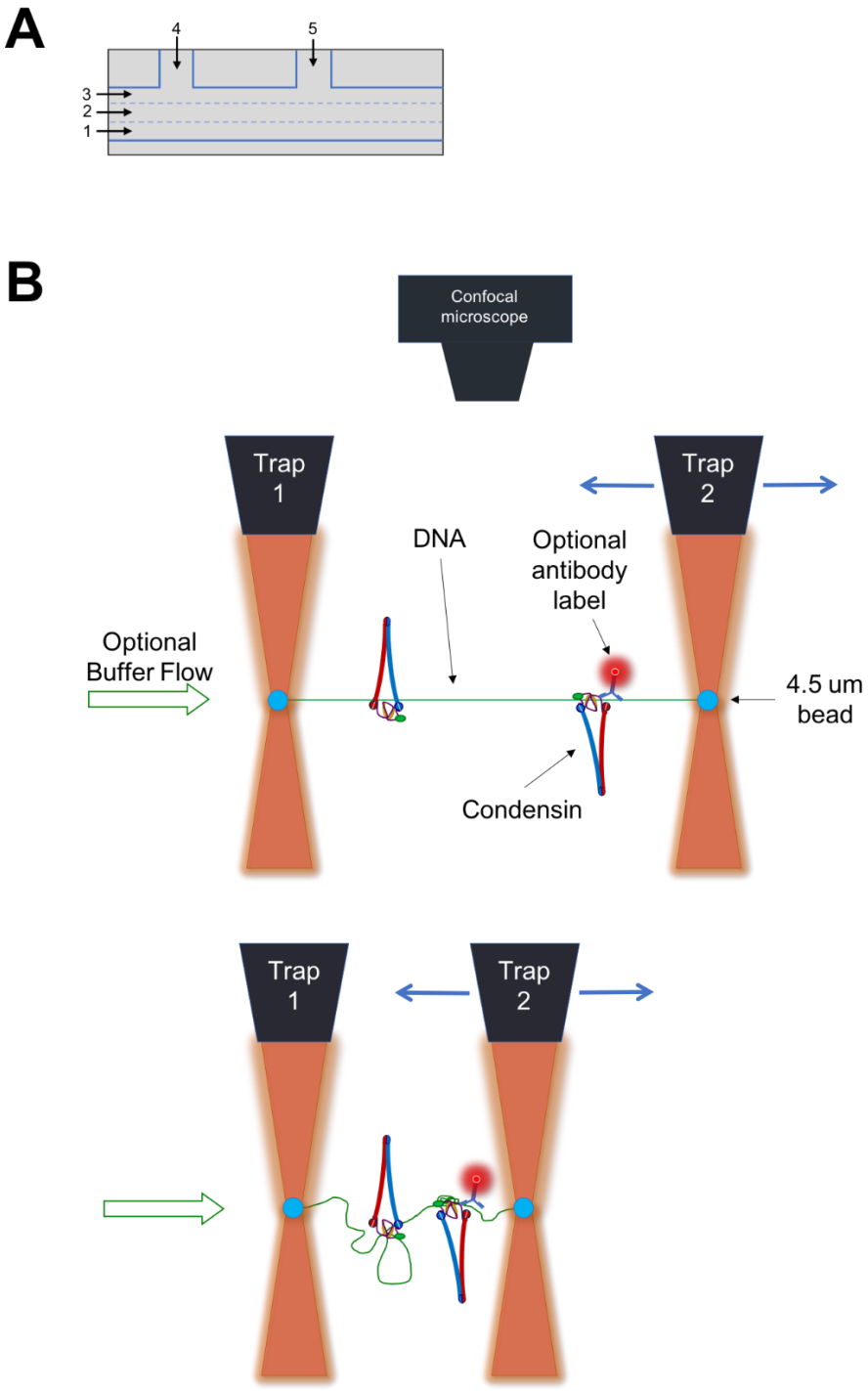


Figure 5.3. Cartoon summary of the Lumix C-Trap setup. (A) Top down schematic of the flow cell showing the arrangement of the 5 availablew flow channels. (B) Side-on, cross-sectional view of the flow cell depicting captured DNA at contour length (top) and relaxed (bottom). 4.5 um beads are shown in blue anchoring DNA in green between the focal point of the two trapping lasers.

5.3.1 Condensation Assays

It has previously been reported that purified condensin complexes are capable of compacting DNA against forces of < 2 pN *in vitro* (Eeftens et al., 2017). This compaction is ATP dependant and salt sensitive, with compaction events being disrupted in NaCl concentrations above 500 mM.

We reasoned that reproduction of these observations in an alternate system would be a valid demonstration of the activity of our purified complex, but also by finding conditions in which DNA compaction could be reliably observed with the C-trap we might form the basis for future experiments in which the effect on condensation of additional co-factors and alternative substrates could be examined

By fixing one of the trapping lasers in place and applying a constant stretching force of 1 pN to the DNA we were able to produce an experimental set up similar to that of magnetic tweezers. In the presence of condensin a progressive shortening of DNA was observed as a decrease in the distance between the trapped beads. Figure 5.4 A shows that condensin alone, at a concentration of 5 nM with 2 mM ATP, causes a decrease in bead distance when bead 2 is trapped at 1 pN constant force. Compaction of DNA is apparently a dynamic process as we do see reversals of the compaction activity, though with a clear trend towards DNA shortening. Previous work has suggested a concentration dependant compaction rate of upto 200 nm/s (~588 bp/s). Rates observed in the optical tweezer experiments (figure 5.4) of between 126.4 and 151.8 bp/s, while not identical, are close given differences in experimental parameters.

Reversal of this DNA compaction by re-extension of the DNA to its contour length of 14 μm manifests as a series of spikes in the force-extension curve profile following ATP dependant compaction (figure 5.4 B, orange arrows). The spikes likely correspond to the physical breakage of inter-strand DNA bridges mediated by condensin, a phenomenon that would be consistent with loop formation as demonstrated by (Ganji et al., 2019).

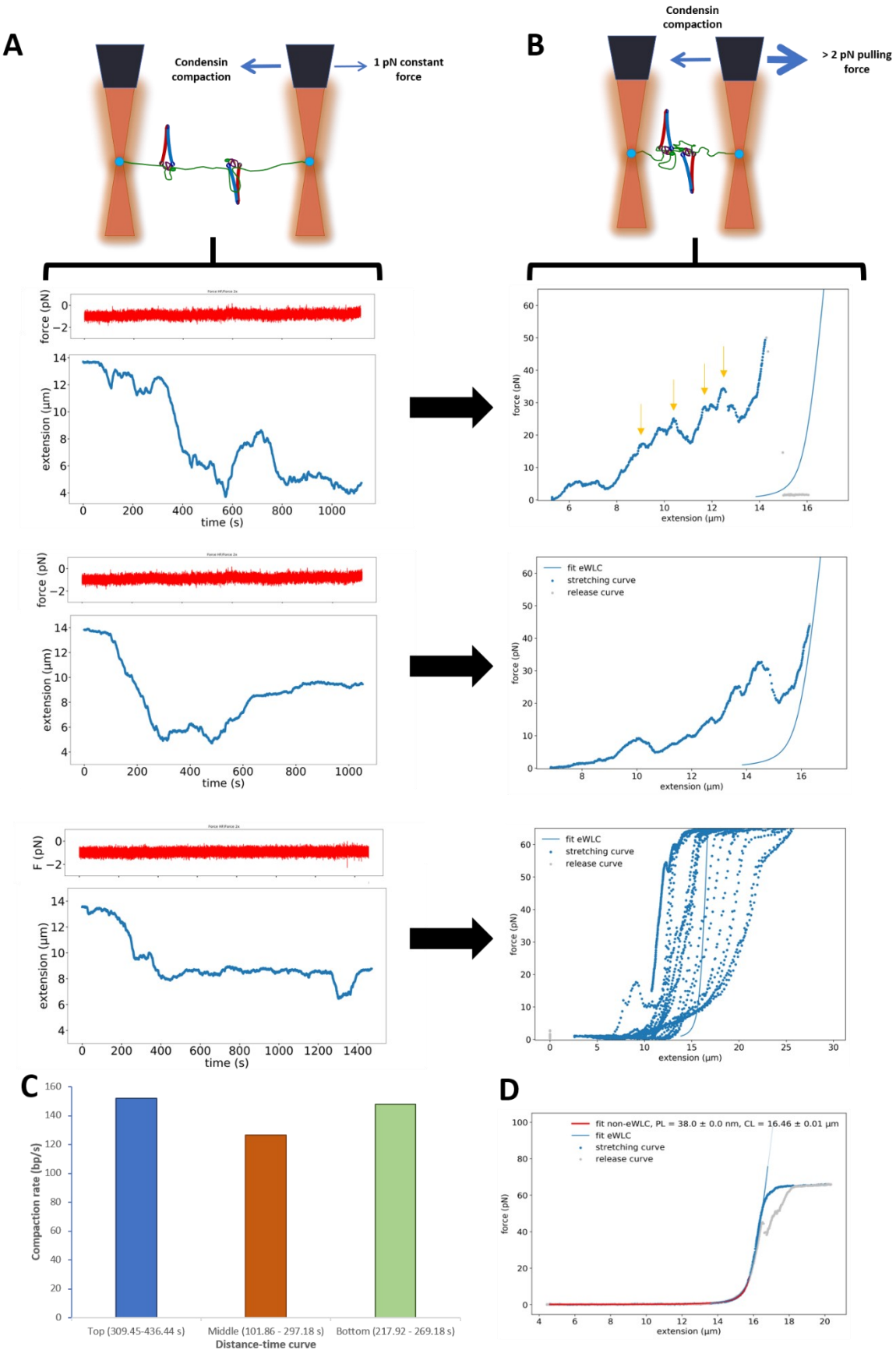

A
Condensin compaction
1 pN constant force
B
Condensin compaction
> 2 pN pulling force
force (pN)
extension (µm)
time (s)
fit eWLC
stretching curve
release curve
F (pN)
C
Compaction rate (bp/s)
Top (309.45-436.44 s)
Middle (101.86 - 297.18 s)
Bottom (217.92 - 269.18 s)
Distance-time curve
D
fit non-eWLC, PL = 38.0 ± 0.0 nm, CL = 16.46 ± 0.01 µm
fit eWLC
stretching curve
release curve

Figure 5.4. Single molecule condensation assays. (A) Distance-time curve for λ DNA in the presence of 5 nM condensin and 2 mM ATP. DNA was incubated with condensin in Tweezer buffer with a constant stretching force of 1 pN being applied and the distance between both force-trapped beads recorded over time. (B) Reversal of compaction events for the 3 DNA molecules shown in A. Following condensation, force-trapped beads were moved to stretch DNA back to its contour length of 14 μm. The force required to extend the DNA as measured by the distance between beads was recorded. Examples of "spikes" in the force extension curve are highlighted by orange arrows. Note the third molecule did not stretch back to its original length on the first pull and was stretch multiple times. Reference force extensuion curves are shown as solid blue lines following the worm-like chain model of naked DNA (WLC). (C) Estimated rates of compaction for the distance-time curves in figure A. Rates were estimated by selecting areas of linear compaction and calculating the change in bead distance in base pairs (bp) per second. (D) Control compaction reversal with condensin but not ATP. Produced in collaboration with Dr Artur Kaczmarczyk.

5.3.2 Translocation assays

It has also been demonstrated previously that condensin, when applied to double tethered DNA can act as a molecular motor in an ATP dependant manner (Terakawa et al., 2017). It is unclear if this is a phenomenon that occurs *in vivo*, though it has been observed that *S. cerevisiae* condensin will relocate from the centromere to the chromosome arms during anaphase (Leonard et al., 2015). Alternatively, this translocation is thought to be a by-product of the condensin motor failing to real in tethered DNA. As a result, the anchor slips and condensin appears to translocate. In either case the translocation activity and its rate are closely linked to the complex activity and thus potentially serve as a reliable assay for activity.

By fixing the trapping lasers in place we were able to create a DNA substrate very similar to the double tethered DNA found in the DNA curtains set up used previously (Terakawa et al., 2017). 1D imaging of the DNA substrate in the presence of ATP and condensin tagged with anti-HA antibodies conjugated to Q-dot allowed us to track the movement of individual complexes over time along the DNA. As predicted the tagged complex is capable of translocating along the DNA in a unidirectional manner (figure 5.5). Similar to the aforementioned compaction assays, not only does this assay provide validation of our complexes activity, but it also has potential as a base assay for future studies. Furthermore, preliminary analysis of the rate of translocation in the first condensin and ATP kymograph suggests the complex in this case is moving at a rate of ~136 bp/s, which would be roughly

consistent with previous observations (Terakawa et al., 2017). Further data will be needed to refine this observation.

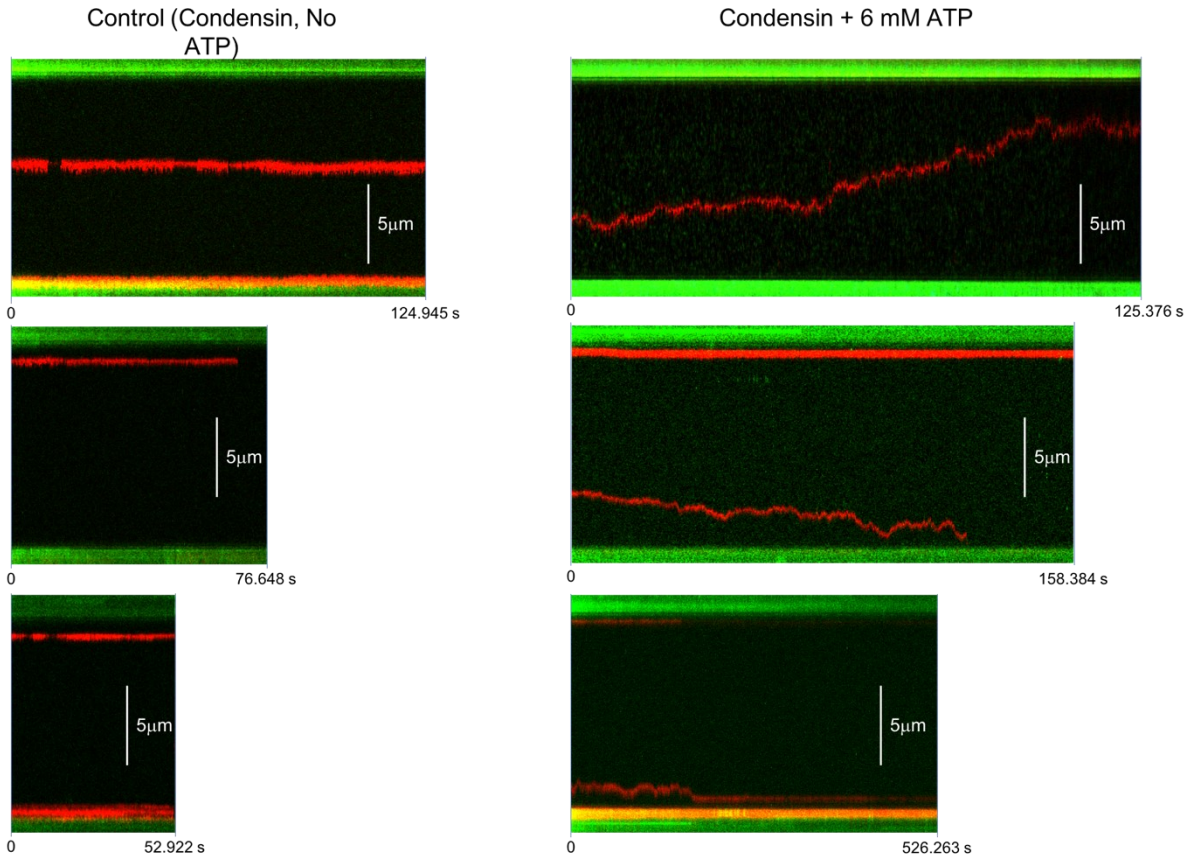


Figure 5.5. Condensin translocation kymographs. 5 nM of condensin labelled with anti-HA-Qdot 705 conjugates was incubated with DNA with minimal flow for 30 seconds to allow loading on to DNA. DNA was then moved to a buffer-only channel containing no ATP for controls or 6 mM ATP. DNA was held at contour length of around 14 um and imaged continuously with a 1D scan. Fluorophores were excited with 488 nm laser. Control N = 3, 6 mM ATP N = 5.

Chapter 6 – Chromatinisation of a Single Molecule Substrate for Condensin Assays

Having reproduced and adapted a small selection of assays using the Lumix C-trap system we next aimed to build on them further. The ultimate aim was to produce a chromatinised substrate within the flow cell, and to characterise condensin behaviour on this substrate. At time of writing this work is still ongoing, but what follows is an up to date summary of the current achievements of this project.

6.1 Histone Purification Optimisation

Chromatinisation of DNA *in vitro* can be undertaken in several ways, and multiple methods have been described previously. Each presents their own advantages, which prompted us at the outset to seek for an optimum workflow that best suited our goal of introducing chromatin to the C-trap system.

While methods do exist for the formation of nucleosome arrays using undefined cell extracts (Bennink et al., 2001, Lusser and Tadonaga, 2004), early on in our optimisation process it was decided that a more defined substrate was needed to avoid any potential artefacts produced by contaminating proteins. Therefore, the first step in assembling a nucleosome array on the DNA template was to purify recombinant histones *in vitro*. The *S. cerevisiae* histones were chosen for this purpose in keeping with our *S. cerevisiae* condensin complex.

We employed 3 different methods, eventually settling on one based on its balance of efficiency and yield. Figure 6.1 A gives a summary of the three methods used. Method 1 was first described in 1999 by Luger et al., and involves the purification of individual histones from inclusion bodies under denaturing conditions. Inclusion bodies are formed during expression of recombinant histones in *E. coli,* and soluble material is removed by repeated washes with buffer containing 1% triton x100. These inclusion bodies are then solubilised with DMSO and 6M Guanidine HCl buffer, followed by SEC in 8M urea buffer and cation

exchange. In order to assemble complete nucleosomes each of the four individual histones H2A, H2B, H3, and H4 must be produced separately. Figure 6.1 B shows the final purification of each of these proteins by method 1. The yield for this method was considered good for all histones giving up to 20 mg, except for H4 which often produced only around 1 mg final protein per 6 L of culture in our hands. While this method is robust, its main drawback was how time consuming it is in comparison to later methods. The production of inclusion bodies to begin each purification takes several hours of repeated washes and sonication to remove contaminants, followed by a 3-day purification under denaturing conditions in urea buffer. Finally, after the histones themselves are produced there is a further day of work to assemble the lyophilised monomers into octamers for nucleosome assembly. The paucity of histone H4 also made repeat purification necessary to produce sufficient amounts for octamer assembly.

Like method 1, method 2 also involves the purification of each histone individually. Unlike method 1 though, this method forgoes the time-consuming process of producing inclusion bodies from bacterial pellets in favour of acid precipitation of contaminants, followed by recovery of the histones by acetone precipitation. This method was attempted principally to overcome our problems with histone H4 yield, which was comparatively low, though it proved to be no better in this regard.

While time consuming, the main advantage of individual purifications is that histones can be swapped out in favour of specific variants for example, Cse4 (yeast centromere specific H3). It is also possible for specific post-translational modifications to be introduced as synthetic peptide tails by conjugation with truncated histones (Bartke et al., 2010). These and other potential embellishments to the chromatin substrate could in theory pave the way to a range of experiments that examined the interplay between condensin and context-specific chromatin modifications and states. However, after careful consideration it was decided that simply building and thoroughly probing condensins' interaction with a chromatinised DNA substrate with the C-trap would be more feasible in the time available. Therefore, it was reasoned that a more recently described method (method 3) of simultaneous expression of all 4 histones in the same cell would be more expedient (Kingston et al., 2011, Kurat et al., 2015).

Figure 6.1 C and D detail the purification of histone octamers by co-expression. All 4 *S. cerevisiae* histones are present on two plasmids in BL21 (DE3) codon + RIL cells and form soluble complexes than can be purified over a heparin affinity column. Lower molecular weight combinations of histone tetramers, dimers and monomers are then separated away by SEC, with only the octamers being recovered. This method not only streamlines the production of the histone themselves but also negates the need for a separate octamer assembly step in the process. While it is not clear if PTMs could be introduced to the purified octamer, the rapidity and ease of this method compensates for its reduced versatility.

This same method was also used unmodified to produce octamers containing mutant H2A with reactive cysteine residues in either the N or C terminal tail. These residues were later used as conjugation sites for Alexa Flour 532-malaemide labelling reactions (figure 6.2 A - B).

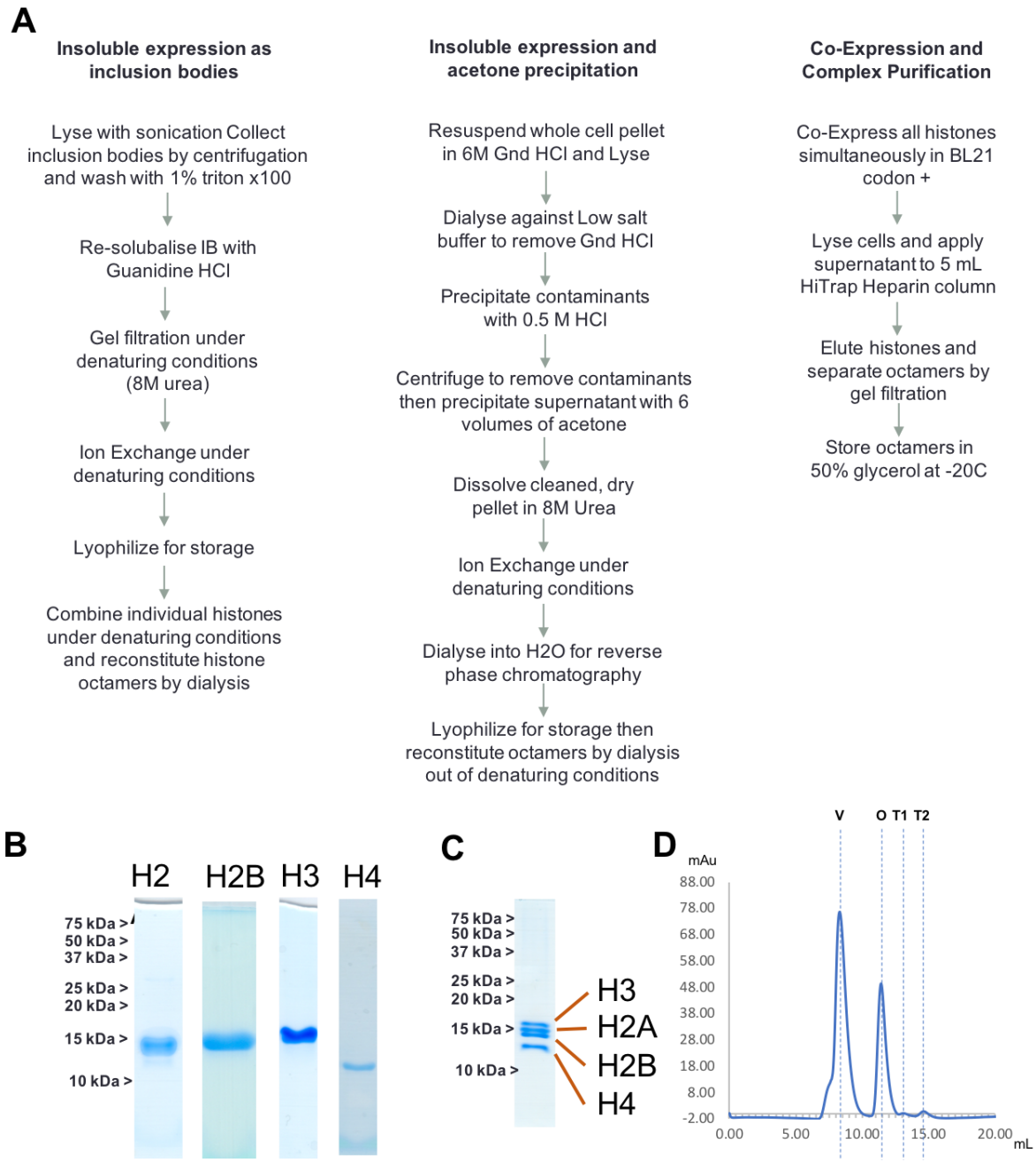


Figure 6.1. Histone purification. (A) Summary workflows for the 3 attempted methods of histone purification. (B) Comparison of individually purified histones H2A, H2B, H3 and H4 by extraction from insoluble inclusion bodies under denaturing conditions. (C) Peak fraction from the final SEC purification step in S. cerevisiae histone octamer purification. Histones produced here were taken from the co-expression method. (D) Chromatogram of the final SEC step in S. cerevisiae histone octamer purification shown in C. V indicates void volume, O elution volume of histone octamers, T1 indicates the elution volume for histone H3 and H4 tetramers and T2 indicates the putative elution volume for histone H2A and H2B tetramers or dimers.

6.2 Chaperone Mediated in situ Chromatin Assembly

As with the purification of histones, the assembly of nucleosome arrays can be achieved in multiple ways. The simplest of these methods utilises the spontaneous loading of histones on to DNA by removal of salt with dialysis. Here, preassembled nucleosome octamers are mixed with DNA in a high salt buffer (2M KCl or NaCl) which is then slowly exchanged for a low salt buffer (250 mM KCl or NaCl) overnight.

This method can be used to assemble highly defined and regular chromatin arrays with known DNA templates such as the "widom 601" DNA, which has been used extensively throughout the literature to study chromatin dynamics (Luger et al., 1999). This specific sequence of around 187 bp DNA is popular in that it strongly binds and positions histones in a fixed position allowing formation of defined arrays. Formation of homogenous nucleosome arrays using this method is certainly robust and well described, however, for our purposes it presents several disadvantages. For the 601 sequence, the thermal energy required to induce sliding on 601 DNA is reported to be above 55ºC (Perales et al., 2011, Flaus, 2012). If the dynamic repositioning of histones were to be important for condensin activity in some way, this strong positioning might introduce artefacts.

A number of other more labile positioning sequences are available such as the 5s rDNA sequence (Simpson and Stafford 1983), but these sequences also run into an additional drawback in that the production of defined arrays often includes many areas of repetitive DNA which are susceptible to recombination in host strains. Production of a homogenous solution of very large arrays of many kilobase, particularly those made exclusively of the positioning sequence with short linkers, is therefore non-trivial. Finally, the speed of buffer exchange is very important for proper assembly of nucleosomes by salt dialysis. If the salt concentration is reduced too quickly there is a strong possibility that nucleosomes and DNA will precipitate (Lee and Narliker, 2001), with even a small amount of precipitation capable of disrupting laminar flow within the C-trap.

A number of drawbacks resulting from the use of repetitive array DNA for defined positioning can be overcome by simply using random sequence of DNA, but salt dialysis deposition of nucleosomes on these templates is itself random, thus creating areas of naked

DNA which condensin may interact with differently to fully chromatinised regions (Lusser and Kadonaga, 2004).

We also considered production of pre-assembled nucleosome arrays which could be caught by beads as naked DNA would. This method has been effective with magnetic tweezers (Kaczmarczyk et al., 2017) but ultimately, we reasoned that for the C-trap this would not be feasible due to the ping-pong method of catching DNA with beads. To tether DNA on the C-trap, two untethered beads are moved back and forth with the trapping lasers relative to one another in the presence of tagged DNA. When DNA is caught between both beads it is detected by an increase in the force needed to move the beads apart. This movement of the beads would likely result in partial disassembly of nucleosome arrays upon catching, therefore, a method of assembling nucleosomes *in situ* is required. Preferably, this method would produce a roughly physiological array on the existing λ DNA that was already available for the C-trap.

Chaperone and histone remodeller-mediated chromatin assembly has been previously described in the literature (Kurat et al., 2017, Vary et al., 2004), and while this method requires the production of more recombinant proteins than the salt dialysis method, once the initial investment of protein purification has been achieved, assembly reaction times are faster than salt dialysis, and the substrate produced can have different sequences. We reasoned that including the components for a chaperone-mediate nucleosome assembly reaction in one of the 5 flow cell channels would allow us to chromatinise DNA after it had been caught *in situ*.

We began by purifying recombinant, histone-remodeller Nap1, which is required first to load the purified histone octamers onto DNA. The purification of this chaperone was adapted from (Vary et al., 2004 and Kurat et al., 2016) with minor changes and is detailed in figure 6.2 A-B. An N terminally GST tagged version of the *S. cerevisiae* Nap1 was expressed in *E. coli* and purified in a two-step process of glutathione agarose affinity and MonoQ anion exchange with a gradient elution. GST was removed by cleavage at a preScission protease site which allowed elution of Nap1 from glutathione agarose resin.

The chromatin remodeller ISWIa, which is a complex of two proteins ISW1 and IOC3, can also be included in assembly reactions with ATP to space the formed arrays evenly throughout the DNA length. This evenly spaced array is closer to that found *in vivo* and does not require specific positioning sequences. Purification of this complex was also adapted from (Kurat et al., 2016) and is detailed in figure 6.2 C-D. The endogenous IOC3 subunit of the complex is tagged with a short 3xFLAG peptide in *S. cerevisiae* cells and purified by FLAG affinity. IOC3 is copurified with the second subunit ISW1 and the whole complex is further purified by MonoQ anion exchange with a 20 column volume gradient elution up to 1 M NaCl.

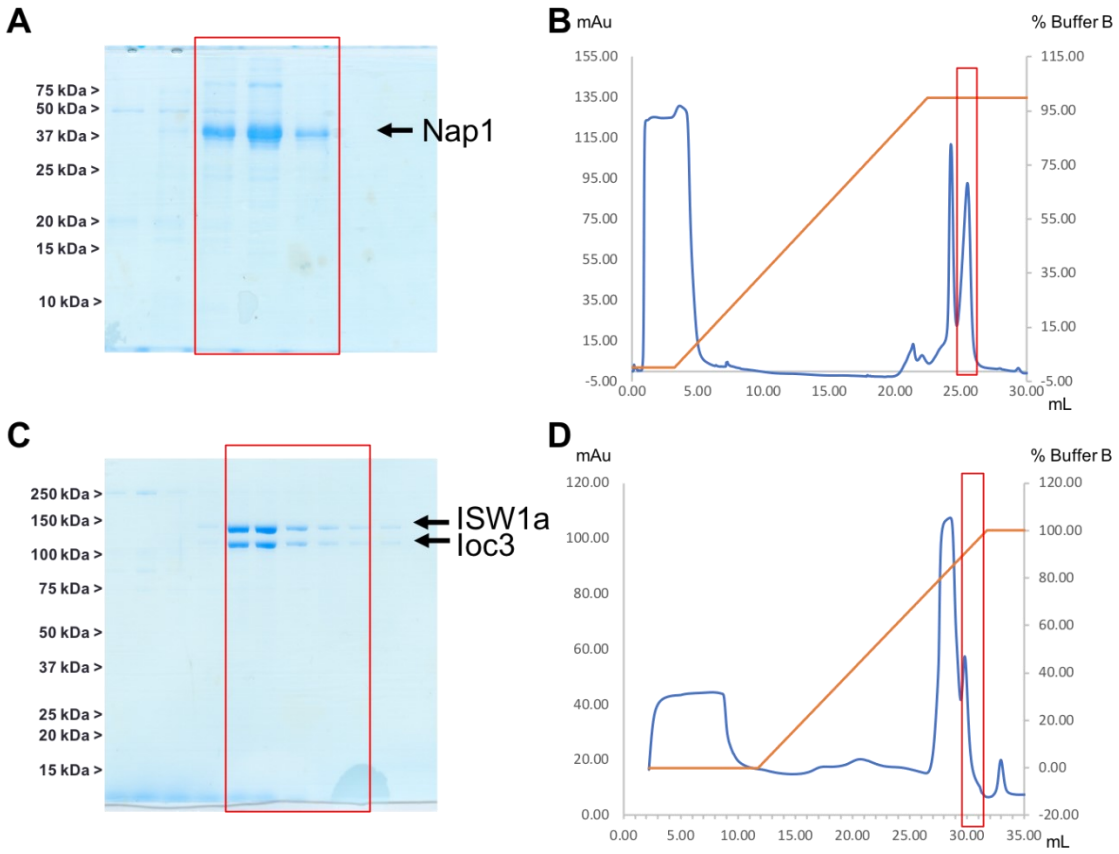


Figure 6.2. Purification of components for in vitro assembly of chromatin. (A) SDS PAGE of the final anion exchange step for purification of the histone chaperone Nap1. Red box indicates the fractions pooled and concentrated for chromatin assembly experiments. Red arrow indicates the expected molecular weight of Nap1. (B) Elution profile of Nap1 from HiTrap Q 5 mL anion exchange column. Red box indicates the fractions pooled and concentrated for further chromatin assembly experiments. (C) SDS APGE of the final elution from MonoQ 5/50 gl column of the histone remodeller ISWIa complex proteins. IOC3 and ISW1 are indicated by blue and green arrows respectively. Red box indicates fractions that were pooled and concentrated for further chromatin assembly experiments. (D) Elution profile of ISWIa from MonoQ 5/50 gl anion exchange column. Red box indicates the fractions pooled and concentrated for further chromatin assembly experiments.

Test assemblies of nucleosome arrays are shown in figure 6.3. Here nucleosome arrays were formed on linearized pRS426 plasmids in 30 uL reactions with ATP and an ATP regeneration system (creatin phosphokinase and creatine phosphate) for 4 hours. Once arrays were formed, 5-minute digests with MNase were done to produce a "ladder" of mono, di, tri etc. nucleotide DNA which was protected from digestion by the assembled nucleosomes. This ladder was visualised by agarose gel electrophoresis. Even laddering suggests even spacing of nucleosomes (figure 6.3 B).

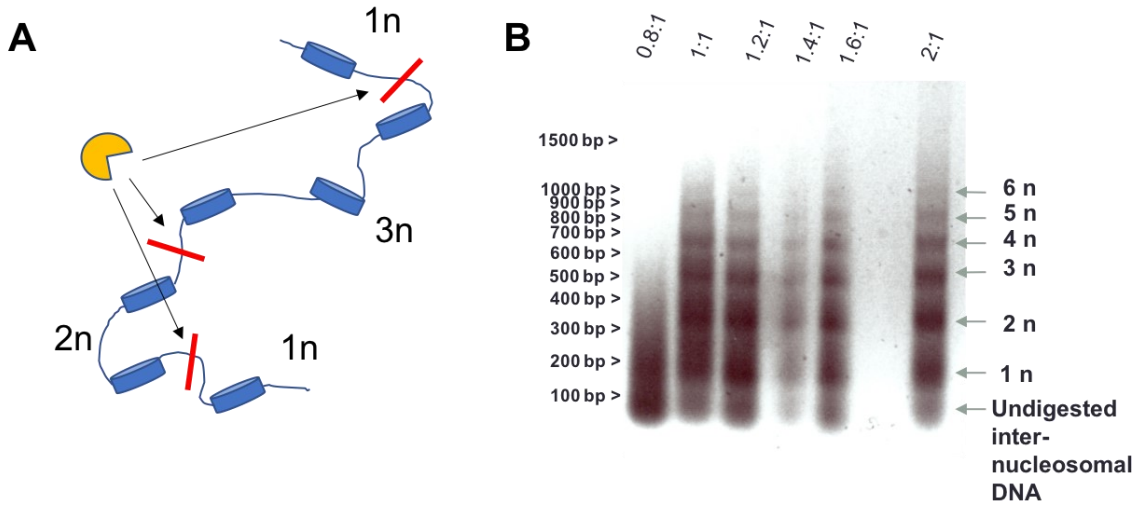


Figure 6.3. Chromatin test assembly. (A) Cartoon representation of a typical Mnase digest experiment. MNase is depicted in yellow and chroamtinised DNA in blue. Cylinders represent nucleosomes and red lines indicate potential digestion sites between nucleosomes. (B) Agarose gel of DNA fragments formed by the digestion of chromatin for 5 minutes with 50 units Mnase in 30 uL. Ratios of DNA to histone octamer by weight are indicated. Arrows indicate bands corresponding to mono, di, tri etc. nucleosome DNA that has been protected from digestion by the bound histones.

With the confirmation that the assembly machinery is active, we next attempted to adapt this system for use on the C-trap, assembling chromatin *in situ*. In this case the nature of the flow cell precludes MNase digestion as means of assaying chromatin assembly, but spontaneous compaction of DNA in the presence of histone octamers (figure 6.4 A) and chaperones, and

force extension curve "tears" suggests wrapping of DNA (figure 6.4 B). Thus far, Nap1 only has been used as a loader to refine conditions for *in situ* chromatinization. This DNA wrapping also appears to be possible without the inclusion of Nap1 (figure 6.4 C), suggesting spontaneous loading of histones, though whether these are correctly wrapped without the chaperone is unclear and requires further investigation.

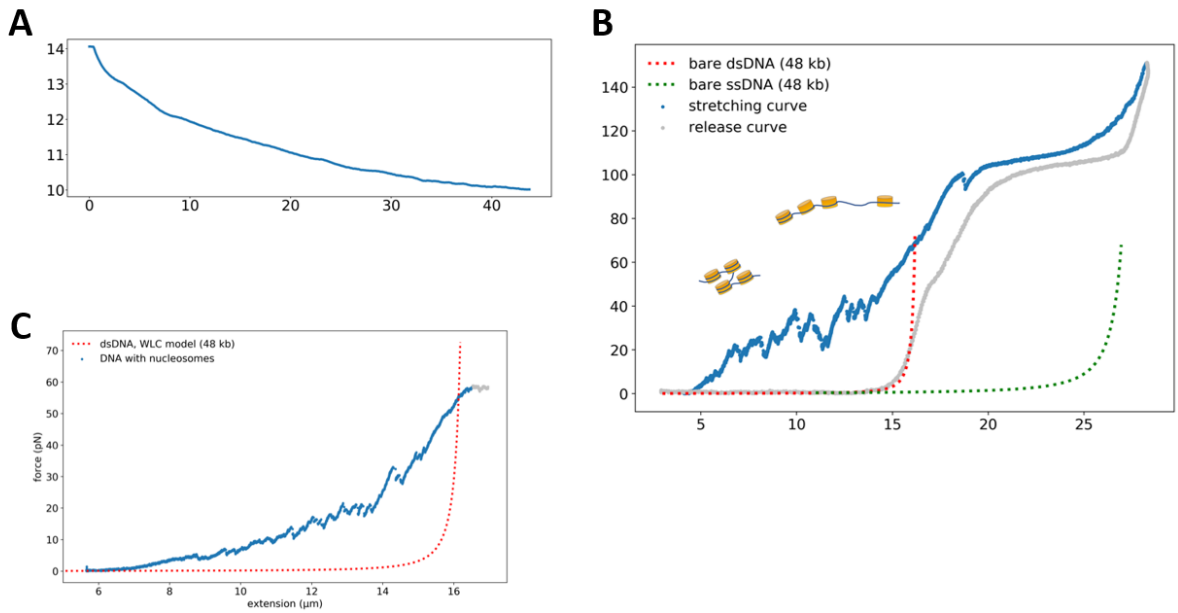


Figure 6.4. Assembly of chromatin in situ. (A) Distance-time curve of C-trap biotinylated beads, holding λ phage DNA, held at 1 pN in the presence of purified histone octamer and Nap1. (B) Force extension curve of DNA following incubation with purified histone octamer and Nap1. Blue line indicates stretching curve following nucleosome assembly. Grey line indicates the return curve of stretched DNA to the un-extended bead distance of 5um. Red and green dotted lines indicate reference force-extension curves for dsDNA and ssDNA respectively. (C) Force-extension curve for DNA following incubation with purified histones only. Produced in collaboration with Dr Artur Kaczmarczyk.

Additionally, direct imaging of chromatin assembly was possible with the inclusion of fluorescently labelled H2A octamers (figure 6.5). We were able to introduce an Alexa Fluor 532 (AF532) tag by purifying histone octamers containing a histone H2A variant with a single cysteine substitution in its N terminal tail. AF532-maleimide conjugates were then reacted with the side chains of this cysteine residue to covalently link the tag. Octamers were subjected to SEC to remove any disassembled tetramers produced during the labelling reaction (figure 6.5 A) and fluorescent labelling of the re-recovered octamers was tested by

running protein on an SDS PAGE gel and scanning for H2A fluorescence with excitation at 488 nm (figure 6.4 B).

Labelled octamers were subsequently used for *in situ* chromatin assembly within the C-trap flow cell and imaged directly by confocal microscopy following incubation with Nap1. The appearance of intermittent spots of fluorescent signal indicates that the labelled histones are indeed assembling on to the DNA substrate (figure 6.5 C).

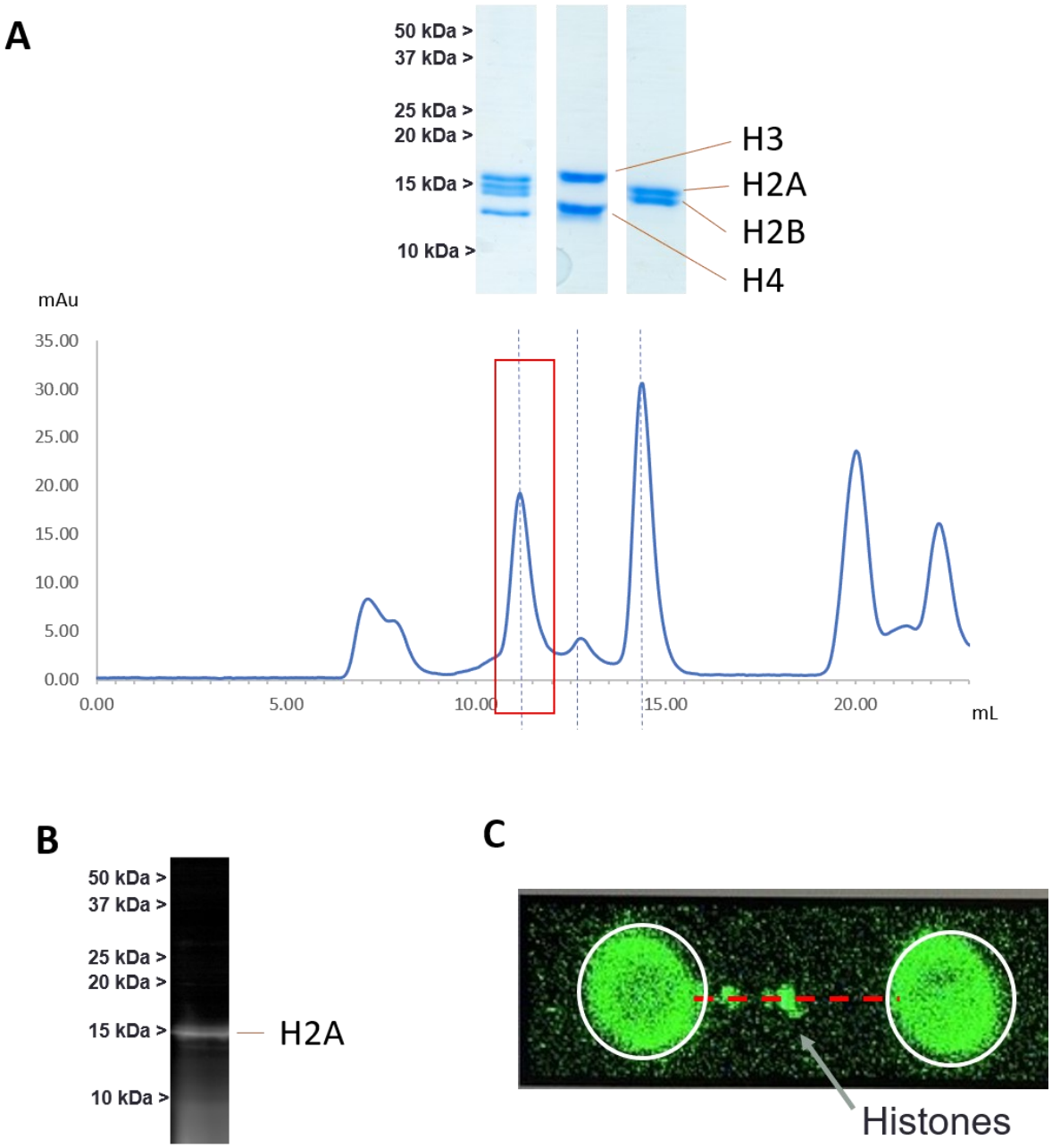


Figure 6.5. Fluorescent labelling of histones. (A) Elution profile of purified S. cerevisiae histone octamers from superdex 200 increase 10/300 gl column following a 2 hour labelling reaction with malaemide-Alexa Flour 532. Red box indicates the octamer peak following the reaction which was pooled and concentrated for single molecule applications. (B) SDS PAGE of the labelled octamer imaged under 532 nm laser to confirm that H2A had bound successfully to the label. (C) Flourescent imaging of captured DNA on the C-trap following incubation with labelled histone octamer and Nap1. Produced in collaboration with Dr Artur Kaczmarczyk.

While these results are promising, much work will need to be done to properly characterise the chromatin assembled here, as well as the effect of this chromatin on condensin activities.

Chapter 7 – Discussion

7.1 Analysis of Key Condensin Structural Features

7.1.1. The Condensin Elbow

Solving the structure of the *S. cerevisiae* condensin complex is an important step in our understanding of the SMC complex function at large. Analysis of the structure and comparison to the previously described structures reveals striking similarities within the family. Notably, the presence of the elbow raises interesting questions about the molecular mechanism underpinning condensin, especially given the presence of similar elbow domains in the bacterial MukBEF complex and *S. cerevisiae* cohesin (Burmann et al., 2019).

It seems likely then given the high degree of conservation that bending of the coiled-coils is an important and conserved feature of the SMC complex family in general. This will no doubt be key to future models of loop extrusion and in fact our findings are already consistent with several previously described models (Hassler et al., 2018, Burmann et al., 2018).

It remains possible that the bending of the coiled-coil is the power stroke that feeds DNA through the complex (figure 1.9), though recent evidence suggests that DNA binding to the hinge domain is expendable for bacterial SMC function. If this is also the case for condensin then it is unlikely that this feature provides the power stroke (Valasquez Nunez et al., 2019, Burmann et al., 2017). More likely, it is possible that the pumping action of the condensin complex is achieved by closing of the coiled-coil lumen (figure 1.9). Why the bend would be necessary for this as opposed to simply juxtaposing the coiled-coils is unclear but depending on the intermediate conformations that the coils might take when switching between open and folded-juxtaposed states it is possible that an asymmetric bend is important for the directionality of the loop extrusion. Alternatively, folding may be necessary to relieve tension in the coiled-coils, or for some other as yet undiscovered purpose.

Topological entrapment of the DNA in the cohesin coiled-coil lumen alone has recently been shown to not be simultaneously possible with entrapment of DNA in the kleisin pocket

(chapard et al., 2019). However, if a loop were to be passed into the lumen it would not be detected by their assay, which would be consistent with a closing lumen model of loop extrusion (Chapard et al., 2019).

7.1.2. Future Analysis of the Condensin Elbow

To probe the role of the elbow further, future work will need to focus on how its movement may be coupled to ATP hydrolysis and how transfer of conformational change is mediated by the coiled-coils. One approach may be targeting of the mid-coiled-coil residues themselves with proline substitutions to fix the elbow conformation. Purification of these mutant complexes and biochemical/single molecule assessment using the assays described in chapters 5 and 6 would be an appropriate approach, as would simultaneous introduction of the same mutants into *S. cerevisiae* hosts. The great advantage of having solved the condensin structure for a well characterised model organism such as *S. cerevisiae* is that predictions made from the finished structure can be readily tested for their phenotypic consequences.

Targeting of the elbow residues themselves, however, presents the problem of separating effects related to loss of bending from effects of disruption to the periodicity of the coiled coils and subsequent coupling of head domains to hinge (Burmann et al., 2017). Another approach could be to target the charged residues on the Smc4 hinge mediating interaction with the coiled-coil in the folded state. This could perhaps be done by introducing charge reversal mutants to eliminate the interaction, or cross linkable cysteine residues to seal the elbow in the bent conformation. Again, these mutants should also be introduced *in vivo* to check how essential the elbow is to proper condensin complex function.

Targeting of the hinge is based on the assumption that any interaction with the coiled-coil is either essential to the folded conformation, or stabilising the conformation enough that its loss would manifest as a significant change in the complexes ability to extrude loops. Interestingly, replacement of the hinge domain does not seem to impair complex function in bacterial MukBEF, which would suggest that in these complexes, at least, an interaction with the coiled-coil is not essential, if present at all (Burmann et al., 2017). It is possible that this interaction is needed to stabilise the bent elbow conformation in condensin or other higher eukaryotic complexes but not in MukBEF. This may have consequences for

the interpretation of any assays that target the hinge-coiled-coil interaction specifically. Ideally, both the elbow and the hinge should be targeted.

7.1.3. The Condensin Joint

Another region that should also be further inspected is the condensin joint.
This short discontinuity in the coiled-coils could potentially be providing a pivot point to allow the coiled-coils to twist in response to ATP dependant head engagement. Previous work on the *P. furiosus* Smc protein has shown this domain to be conserved and is potentially important for orienting the Smc heads relative to each other to prevent steric clashes (Diebold-Durand et al., 2017). Our observation is of a break in the continuous left-handed spiral of the coils in which several helices connected by short loops on the Smc4 C terminal strand (approximate residues 1110-1215) wrap around the N terminal strand. A similar structure is seen with the Smc2 C terminal coils (approximate residues 937-983). It is possible that this arrangement lends some flexibility to this region, though atomic force microscopy studies of Smc2 and Smc4 dimers didn't appear to show a double-jointed appearance when bent, assuming the elbow would account for one bend and the joint for the other (Eeftens et al., 2016). It could be that the joint allows for mediation of twisting without disruption of the periodic relationship between the head and the hinge domains.

7.1.4. Future Analysis of the Condensin Joint

As with the elbow, fixing the conformation of the joint by replacing key residues with more rigid prolines might be an effective approach to investigate its functionality. Another approach may be the creation of complexes entirely missing the joint region. If this region does provide a pivot point for the opening of the coiled-coils it might be expected that ATP binding to the head domains would not result in dimerization and subsequent ATP hydrolysis. Alternatively, it may be that ATP can bind and dimerise the heads, but the coiled-coil lumen is prevented from opening, or the torsional strain induced by dimerization breaks the hinge Smc2-Smc4 interface. Analysis of the conformational change induced by introduction of ATP to purified complexes missing their joint could be done by CLMS with the activity of these complexes being assayed by single molecule and *in vivo* approaches as described for the elbow.

Again, as with the elbow the difficultly with targeting a domain within the coiled-coil is that it may disrupt the link between heads and hinge. When removing the joint, it may be prudent to design several constructs with varying lengths of linkers between the N and C terminal portions of the two coiled-coil strands to give the best chance of maintaining their periodicity. These linkers could be flexible to minimise the chances of misaligning the heads and hinge, though this may affect the coiled-coils ability to transfer torsional stress. Another approach may be to base the linker region on known coiled-coil heptad repeats, introducing multiples of repeats until a stable complex can be formed.

7.1.5 Analysis of the HEAT Repeat Subunits Ycs4 and Ycg1 in the Condensin Structure

7.1.5. i. Ycg1

The finding that the Ycg1 subunit sits in a flexible position below the disengaged heads in the ATP free apo state, anchored via its interaction with Brn1, is also a new observation that remains to be explored. The exact purpose of this flexible positioning is unclear but it would mean that the previously described DNA binding site formed by Ycg1 and a belt-like loop of Brn1 is also mobile. Flexible movement of this bound DNA relative to the rest of the complex could be part of the condensin loop extrusion mechanism, or perhaps flexibility in the Ycg1 binding pocket aids in loading of the complex on to DNA. Analysis of our ΔYcg1 tetramer by EMSA seems to suggest that even in the absence of ATP Ycg1 is required for robust non-topological binding of DNA by the complex, which may support a role in loading. This observation has also been recorded in the other studies (Kschonsak et al., 2017).

In parallel with our experiments additional observations, currently unpublished, were made by collaborators (Lee et al., unpublished), which are currently being interpreted in conjunction with our own results for a future publication. In this study it was found by the Haering laboratory (European Molecular Biology Laboratory, Heidelberg) that upon engagement of Smc heads by ATP not only does the N terminal of Brn1 dissociate from the head proximal coiled-coils of Smc2 along with Ycs4 dissociation from the Smc4 head (Hassler et al., 2019), but also the Ycg1 subunit engages the Smc2 head directly (Lee et al., unpublished).

It is unclear why the engagement of Ycg1 with Smc2 might be important, but it may be speculated that disengagement of the N terminal of Brn1 would initially seem to open the kleisin pocket, potentially releasing any topologically entrapped DNA. The same is true for cohesin (Muir et al., 2020). Binding of Ycg1 to the Smc2 head therefore may maintain the integrity of the kleisin pocket, but why this would be necessary and why the N terminal of the kleisin may need to disengage at all is not known.

One hypothesis could be that the binding of Ycg1 to the Smc2 head serves to shrink the lumen of the kleisin pocket, and that by shrinking the size of this pocket DNA loops or topologically entrapped strands are prevented from sliding backwards while a new segment of DNA is captured within the lumen of the coiled-coils.

Another possibility may be that the flexible attachment of Ycg1 to the complex aids in its capture of a DNA strand in the initial loading stage of the loop extrusion process, but later binding of Ycg1 to the Smc2 head provides a more stable anchor point from which the complex can gain leverage when extruding loops. Moving of the anchor point closer to the coiled coil lumen to aid in segment capture when the lumen opens (if this is indeed how the complex pumps DNA) may also be a reasonable explanation.

7.1.5 ii Future Analysis of Ycg1

The importance of Ycg1 switching from its flexible positioning below the Smc heads, described in figure 4.11, to its more rigid association with the Smc2 head observed by the Haering lab will need to be explored further in the future. Initially, it should be determined how essential the interaction with Smc2 is by mutations of the key interacting residues, with analysis of the resulting phenotypes and complex activity. Further work should then aim to address the hypotheses put forward here.

If shrinking of the kleisin lumen is important in stabilising extruded loops while new DNA is being loaded, then increasing the size of the lumen by increasing the length of Brn1 between its Ycg1 binding site and its C terminal winged helix domain with insertion of flexible peptides should affect the complexes ability to extrude loops.

Comparative measurements of compaction rates between WT and mutants assembled with a longer Brn1 could be analysed in single molecule approaches such as optical tweezers or previously established methods with higher throughput, such as magnetic tweezers. Alternatively, the force against which compaction can occur could be measured. If extruded loop stability is affected by changing the kleisin lumen size, then it may be expected that condensin with a longer Brn1 would not be able to maintain loops against the same force as a WT condensin, thus no compaction would be observed at lower force.

To assay the role of Ycg1's apparent flexibility in the ATP free apo state, an opposite approach could be taken by shortening the Brn1 to reduce its mobility relative to the rest of the complex. A similar effect might be achieved without the risk of preventing Ycg1's ability to associate with the Smc2 head by producing a fusion of Smc2 and Ycg1, bridged by a short flexible linker, though the length of this single construct may make expression difficult.

Measurement of DNA loading and unloading dynamics of complexes with fixed Ycg1 may be best done with biochemical techniques such as isothermal titration calorimetry (ITC), though fluorescently labelled complexes could also be directly observed loading DNA in single molecule experiments with optical tweezes or DNA curtains (Terakawa et al., 2017).

7.1.5 iii Ycs4

The conservation of several features of Ycs4 between homologs raises several interesting questions about its function. First, it appears that the position of Ycs4 would make it an ideal candidate for the anchoring of DNA in a similar way to Ycg1 or Scc3 described previously (Hassler et al., 2019, Lee et al, 2017). Binding of DNA to the head proximal coiled-coils is an important feature for some SMC complexes (Valasquez-Nunez et al., 2019), which in the ATP-free apo state might present a situation in which the DNA bound at the heads can align with the outer surface of the Ycs4 subunit below its N terminal or along the concave inner surface of the Ycs4 solenoid, as is the case with Ycg1. However, our analysis by EMSA of purified Ycs4 suggests that this is not the case, though longer Brn1 fragments would need to be tested to determine if a stabilising kleisin loop is required as with Ycg1 (Kschonsak et al., 2017). If Ycs4 is not acting as another binding site for DNA, then its principle function may

be to act as a regulatory subunit at the Smc heads, with its detachment upon ATP binding allowing further conformational changes of the whole complex (Hassler et al., 2019).

Another interesting feature of the Ycs4 subunit is the conserved "proboscis" seen in both our condensin structure and the previously published *C. thermophilum* structure (Hassler et al., 2019) (figure 4.10). This structure is made of two extended alpha helices that project from the main body of the Ycs4 solenoid. This region covers residues from approximately L441 to V528, though residues 457-517 are unresolved in our structure. The fact that this motif is conserved makes it an interesting target for further investigation. Furthermore, a previous study looking at the phosphorylation state of Ycs4 identified 5 residues within the unresolved proboscis region that are phosphorylated in anaphase (S464, T471, S495, S499 and S503) (St Pierre et al., 2009). It is likely then that the projection is acting as a docking site for regulators such as Cdc5, but the down-stream consequences of PTMs at the proboscis are not known.

Finally, from our observations during the purification of the condensin complex and its subassembly into tetramers/trimer we can draw inferences that may guide future work. The observation that loss of the Ycs4 subunit appeared to destabilise the complex during purification could potentially suggest some unexplored structural need for Ycs4, or a potentially a role for it in the assembly of the non-Smc protein subunits. Loss of Brn1 in both the condensin trimer and ΔYcs4 tetramer purifications is unlikely to be due to under expression, as the same plasmid was used for both the wild-type pentameric complex and all sub-complexes. Detection of Brn1 by western blotting shows that it is equally expressed in both the ΔYcg1 and ΔYcs4 tetramers, though there is some evidence of degradation of Brn1 when expressed with no HEAT repeats subunits. It may then be possible that binding of either Ycg1 or Ycs4 helps stabilise overexpressed Brn1 but only Ycs4 is necessary for its further proper assembly into the condensin complex. Future work will be required to explore this further.

7.1.5 iv Future Analysis of Ycs4

The purification of Ycs4 and its homologs may present future opportunities now that the process has been optimised. Work could focus on more in depth characterisation of the subunit and its binding partners, for example, small angle X-ray scattering (SAXS)

approaches were recently used to quantify the flexibility of the Ycg1 subunit with and without binding to Brn1 fragments (Manalastas-Cantos et al., 2019). The apparent flexibility of the HEAT repeat structure has been suggested for the inch-worm type model of condensin translocation (figure 1.9 C), wherein separation of the Smc protein arms and stretching of the Brn1 subunit provide the loop extrusion motor. Demonstrating that Ycs4's flexibility is significantly reduced upon binding of Brn1 as is the case with Ycg1 would be important to further eliminate this model.

To further assess the effect of the aforementioned phosphorylation around the proboscis region, purification and biochemical/biophysical analysis of phosphomimetic mutants of Ycs4 may be a good approach. Assessment of the effects of PTMs on the flexibility of Ycs4 and its ability to bind other condensin subunits such as Brn1 and the Smc4 head by ITC would be an appropriate measure of how PTMs affect subunit interaction and stability. Introduction of the same mutations into the whole condensin complex would also be appropriate, both *in vivo* and for purified complexes.

An important follow up to our observation of Ycs4 affecting purified complex stability could be the inclusion of a small tag on the Ycg1 subunit to allow western blotting in extracts from cells expressing the ΔYcs4 tetramer. This could be used to confirm the presence of the Ycg1 subunit and further demonstrate loss of Ycs4 is in fact resulting in poor assembly, despite sufficiently available Ycg1.

7.2 Technical Improvements to Purification Techniques

In order to further characterise the condensin complex and its subunits and build on our current work, much can be done to streamline our purification processes. Future work building on the three-step purification process should focus on generating variants of the complex with an aim to improve the quality of grids for cryo EM. While the elimination of the Ycg1 subunit does indeed make a significant difference to the orientation of the complex when frozen, imaging still required the use of a tilting stage to capture enough imaging angles. This was done over multiple imaging sessions. If steps to improve the complex orientation can be taken in the purification process itself this could reduce the time needed to resolve the structure of any future complexes harbouring experimental modifications.

Changing or removing tags by inclusion of TEV protease sites in linker regions would be a logical next step. Inclusion of a TEV protease incubation step prior to SEC, possibly with an affinity step to remove uncleaved complexes and reduce heterogeneity would be recommended here. Furthermore, elimination of the 3 x HA Brn1 tag entirely may be another approach for cryo EM only, as it is not needed for purification. Additional removal of one of either the 12 x His tag or the 3 x StrepII tag may also be possible if one of the affinity steps could be replaced by a non-affinity such as heparin or ion exchange.

If complex purification and imaging could be usefully streamlined, the next steps for analysis should involve imaging of condensin with DNA at different stages of the ATP hydrolysis cycle. This can be done by inclusion of non-hydrolysable ATP analogues to stably lock the complex in a heads-engaged conformation with DNA in addition to further imaging of complexes in the apo state.

7.3 Preliminary Single Molecule Analysis of Condensin

The most significant finding of the single molecule analysis of condensin is that previously described phenomena can be corroborated in an unrelated optical tweezer system. Our observation that condensin compacts DNA against force of 1 pN is an important confirmation of the more in-depth studies undertaken on magnetic tweezers (Eeftens et al. 2017) (figure 5.4). The same can also be said for our limited reproduction of the condensin translocation phenomenon on a fixed DNA substrate (Terakawa et al., 2017) (figure 5.5). This assay may give us a future means to measure both complex activity and strength of complex binding by comparing mutant and WT translocation rates.

In addition, our observation of force-extension rips is another parameter that was not previously analysed and may be adapted in the future as an additional assay for condensin function (figure 5.4). For example, we may use this assay to analyse the strength of condensin loop formation, though future work on this assay will need to focus on generating repeats for statistical analysis. It is likely that different interactions between separate parts of the condensin complex and DNA can be broken at different forces. If breaking of these interactions can be mapped to conserved tears in the force-extension curve, then we will have a potent tool for mutant complex analysis by way of the shift in the tears.

7.4 Future Single Molecule Analysis of Condensin on a Chromatinised Substrate

Our ongoing work attempting to chromatinise a substrate for use in the C-trap still requires further optimisation, but successful production of chromatinised DNA *in vitro* with purified recombinant protein is a major step towards realising this aim. While we can see apparent histone induced compaction of DNA and the localisation of fluorescent signals when incubating λ DNA with tagged histones and Nap1, it remains to be seen if this corresponds with the assembly of octamers wrapped with a full physiological 1.6 turns of DNA or if nonspecific binding or partial assembly of tetramers/dimers is responsible for our observations. Further refinement of assembly conditions is required, with titration of Nap1 and histone ratios into the flow cell being a good next step. Systematic adjustment of buffer conditions to vary salt and pH may also help in stabilising histone interactions with DNA.

Future work should involve the introduction of the ISW1 complex to the assembly reaction with fluorescently labelled histones followed by analysis of the spacing of fluorescent signals after incubation.

Once a stable chromatinised substrate has been established there are a myriad of potential experiments that could be performed. Real time compaction of chromatinised DNA in the presence of condensin with comparison to unchromatinised DNA would be one such experiment. Average rate of compaction in both experimental conditions would be the first test, but coupling of this compaction with confocal microscopy of fluorescently labelled condensin and nucleosomes would also be informative. This would potentially allow us to assess how nucleosomes affect condensin stability in the face of obstacles and if DNA can be compacted by condensin past nucleosomes without additional factors to mediate passage. Further analysis of condensins ability to translocate over chroamtinsised DNA could be done using the kymograph translocation assays based on the previous observations (Terakawa et al., 2017).

7.4 Concluding remarks

Building on a wealth of previous work we have here described the first complete structure of the *S. cerevisiae* condensin complex. It is our hope that not only will this new information be of immediate use in modelling the molecular mechanisms of condensin, but also that it will

be used to guide future studies. As more structural studies begin to uncover what is most conserved amongst the SMC complex family, we believe that the structure of condensin will take on even greater significance as it is used to help answer questions about a range of incredibly important protein complexes found in all domains of life.

Furthermore, it is inevitable that the biochemical study of condensin will move more towards increasingly complex experiments that aim to better understand its interaction with chromatin. While our work in this direction is preliminary it provides a good starting point for future work to be built on. In conjunction with the complexes structure we hope this too will lead to a more comprehensive understanding of DNA compaction in cell division and its consequences for all eukaryotic life.

Chapter 8 - Methods

8.1 Purification of condensin pentamers, tetramers and trimers

Purification procedure was based on previously described methods (Terakawa et al. 2017), (St-Pierre et al. 2009). Pentameric condensin was produced from W303-1a S. cerevisiae stains transformed with plasmids GC1267 [URA3 leu2-d pGAL7-SMC4-StrepII3 pGAL10-SMC2 pGAL1-BRN1-HA3-His12] and GC1268 [TRP1 leu2-d pGAL10-YCS4 pGAL1-YCG1] (Hearing lab). Tetrameric condensin, containing Smc2, Smc4, Brn1 and Ycs4, was produced in strains harbouring plasmids GC1267 and GC1259 [TRP1 pGAL10-YCS4]. Tetrameric condensin containing Smc2, Smc4, Brn1 and Ycg1 was produced in strains harbouring plasmids GC1267 and GC1258. Trimeric condensin containing Smc2, Smc4 and Brn1 was produced from strains containing GC1267 only.

Cells were inoculated at an OD of 0.2 and grown to OD 1 at 30 ^{o}C in minimal media with 2 % raffinose, before induction overnight with 2% galactose (24 ^{o}C). Pelleted cells were then resuspended in an equal volume of His Buffer (50 mM Tris pH 7.5, 200 mM NaCl, 5% (v/v) glycerol, 5 mM BME, 20 mM imidazole and EDTA free protease inhibitor cocktail (Roche)) and frozen in liquid nitrogen. Frozen cells were lysed by freezer milling and melting, before incubation with benzonase for 1 hour. Lysates were centrifuged at 45,000 g for 60 minutes, followed by filtration through 0.45 µm and 0.22 µm syringe filters (Sartorius). The cleared lysates were then applied to 5 mL HisTrap affinity column (GE) and washed with 3 column volumes of His Buffer and 4 column volumes each of Wash Buffer 1 (His Buffer with 500 mM NaCl), Wash Buffer 2 (His Buffer with 1 mM ATP, 10 mM KCl and 1 mM $MgCl_2$) and Wash Buffer 3 (His Buffer with 40 mM imidazole). Protein was eluted in His Buffer with 200 mM imidazole.

Peak fractions were pooled and exchanged into Strep Buffer (50 mM Tris pH 7.5, 200 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 1mM EDTA) with either overnight dialysis or by desalting over 5 mL HiTrap desalting column with sephadex G-25 (GE). Exchanged material was centrifuged at 20000 g for 15 minutes then loaded on to a 5 mL StrepTrap column (GE).

Column was washed with 5 column volumes of Strep Buffer followed by elution of bound proteins with Strep buffer + 5 mM desthiobiotin.

Peak fractions were again pooled and concentrated for injection onto superose 6 increase 10/300 GL (GE) equilibrated in GF buffer (Strep Buffer + 1 mM MgCl2). Peak fractions were concentrated to around 2 mg/mL by centrifugation in 100 kDa MWCO centricons (GE) where necessary.

8.2 Protein Expression Assays

Expression constructs were transformed into their respective host strains according to manufacturers instructions. Cells were inoculated from glycerol stocks into 10 mL overnight pre-cultures in LB media, containing selective antibiotics ampicillin, kanamycin, chloramphenicol or spectinomycin as necessary. Cells were grown at 37ºC then the following morning, 10 mL test cultures were inoculated from precultures with a 1:100 dilution in LB, 2xTY or Terrific broth as indicated. Cells were grown to OD 0.6 before induction with IPTG (0.2 mM, 0.5 mM or 1 mM) and samples taken at 2 hours, 4 hours or overnight post induction. 1 mL samples were pelleted by centrifugation at 4000g for 2 minutes in a benchtop centrifuge before resuspending in 100 uL of breaking buffer (20 mM Tris pH 7, 150 mM NaCl, 1 % SDS) to lyse. 5 uL samples were taken from lysed pellets and mixed 1:1 with 1 x Laemmli buffer and run on 4-12 % Bis-Tris polyacrylamide gels (Invitrogen) or 17.5% polyacrylamide gels. Gels were then stained with InstantBlue Coomasie stain (Expedeon).

8.3 Protein Solubility Assays

Overnight cultures were inoculated from glycerol stocks and grown overnight at 37ºC. 2 x 1 L cultures were then inoculated from overnight cultures into 2xTY, with one of the 1 L also containing 1% glucose. Cells were grown to OD 0.6 before induction with 0.2 mM IPTG for 4 hours followed by pelleting of 2 x 500 mL of each media condition at 4000g for 20 minutes.

A single pellet grown in each media condition was then resuspended in Solubility buffer A (SBA) (20 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT, PIC) or Solubility buffer B (SBB) (20 mM Tris pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM DTT, PIC). Lysozyme was added to 1 mg/mL to each resuspended pellet with benzoate nuclease and incubated for an hour at 4ºC. Cells were then sonicated at 5 minutes, 5 seconds on 5 seconds off at 18% output with microprobe (Branson digital sonicator) to lyse. Lysis was checked by light microscope.

1 mL samples of lysed material were then taken and centrifuged for 45 minutes (21000g, 4ºC) to separate soluble and insoluble fractions. Supernatant was removed and insoluble material was resuspended in an equal volume of 2 x concentrated Laemmle buffer. Equal amounts of each supernatant and pellet were sampled and loaded onto SDS PAGE gels (4-12%) and stained with InstantBlue coomasie stain.

8.4 Purification of S. cerevisiae Ycs4-6xHis from E. coli

Expression constructs were transformed into BL21 (DE3) Star according to manufacturer's instructions. Overnight cultures were inoculated from glycerol stocks and used to inoculate 17 L of 2xTY media with ampicillin. Cells were grown to OD 0.6 and induced with 0.5 mM IPTG overnight at 15ºC. Cells were pelleted by centrifugation at 6000g for 20 minutes then resuspended in 300 mL of Ycs4 buffer A (YBA) (50 mM Tris pH 8, 250 mM NaCl, 5 mM BMET and 10% glycerol and PIC). DNase and RNase were added to resuspended cells which were then lysed by sonication for 3 x 3 minutes, 2/4 seconds on/off with 3/8" probe probe at 70% output (Branson digital sonicator).

Lysate was cleared by ultracentrifugation at 100000g for 1 hour then loaded onto a 5 mL HisTrap column. Protein was eluted with a step gradient between YBA and Ycs4 Buffer B (50 mM Tris pH 8, 500 mM imidazole, 50 mM NaCl, 5 mM BMET and 10% glycerol) at 4%, 10%, 20%, 40%, 60% and 100%. Peak fractions eluted with 20% (100 mM imidazole) and 40% (200 mM imidazole) YBB were pooled, with 500 uL samples being subject to SEC on a superpose 6 increase 10/300 gl column. The remaining 20% YBB pool was applied to a MonoQ 4.6/100 PE anion exchange column following 2 x 1-hour dialysis into 4 L of Mono Buffer A (MBA) (50 mM Tris pH 8, 50 mM NaCl 10% glycerol and 5 mM BME). Column

was washed with 4 column volumes (CV) of MBA and eluted with a 40 CV gradient up 25% Mono Buffer B (MBB) (50 mM Tris pH 8, 2 M NaCl 10% glycerol and 5 mM BME).

Peak fractions from Mono Q were pooled and diluted down to 50 mM NaCl with MBA containing no NaCl, then loaded onto a HiTrap Heparin 5 mL column (GE). Protein was eluted with a 20 CV gradient up to 50% MBB. Peak fractions were pooled and concentrated for SEC over a superpose 6 10/300 gl column. Peak material was concentrated with 50 kDa MWCO centricons (GE) to 10 mg/mL and frozen in liquid nitrogen for storage.

8.5 Purification of C. albicans 6xHis-TEV-Ycs4 from E. coli

Expression constructs were transformed into BL21 (DE3) Star according to manufacturer's instructions. Overnight cultures were inoculated from glycerol stocks and used to inoculate 6 L 2xTY with ampicillin. Cells were grown to OD 0.6 and induced with 0.5 mM IPTG overnight at 20ºC. Cells were pelleted at 6000g for 20 minutes and stored at -80ºC.

Pellets were thawed and resuspended in YBA with benzonase nuclease, then lysed by the addition of 0.5 mg/mL lysozyme for 1 hour at 4ºC followed by sonication for 2 x 3 minutes 30 seconds, 3 seconds on, 5 off at 40% output (Branson digital sonicator, 3/8" probe). Lysate was cleared by centrifugation at 45000g for 1 hour followed by filtration through 0.2 um syringe filters. Cleared lysate was then loaded onto 5 mL HisTrap HP (GE). Column was washed with 4% YBB and elution was done with a step gradient of 10%, 20%, 40% and 60% YBB.

Fractions eluting at 20% YBB (100 mM imidazole) were pooled and dialysed overnight into MBA. Protein was then loaded on to a MonoQ 5/50 gl column and washed with 4 CV MBA. Elution was performed with a gradient over 20 CV up to 50% MBB. Peak fractions were collected and diluted with MBA to less than 150 mM NaCl and loaded on to a 5 mL HiTrap Heparin. Elution from the heparin column was performed over 15 CV with a gradient up to 50% MBB.

Following Heparin elution, eluent was mixed with AcTEV proteasae overnight at 4ºC to remove His tags. Digested material was mixed with 1 mL Ni NTA agarose beads and flow-

through collected and concentrated for SEC over a superdex 200 10/300 gl column in (50 mM Tris pH 8, 250 mM NaCl, 5% glycerol, 5mM BMET). Peak fractions were concentrated and frozen in liquid nitrogen for storage at -80ºC.

8.6 Purification of S. cerevisiae Ycs4-6xHis from Insect Cells

Sf9 cell cultures were transfected with viral vectors containing (pFastBAC-Ycs4-6xHis) and grown to an appropriate density in SF900II SFM (1X) at 27ºC before pelleting by centrifugation. Cell pellets were stored at -80ºC. Pellets were thawed and resuspended in YBA with benzonase nuclease followed by sonication to lyse for 4 minutes, 4 on/ 4 off, 40% output (Branson digital sonicator, 3/8" probe). Lysis was checked by light microscopy at 100 x magnification. Lysate was cleared by centrifugation at 45000g for 45 mins and filtered through 0.2 um syringe filters (sartorius).

Cleared lysate was loaded on to a 5 mL HisTrap HP column and washed with 4 CV of YBA followed by step gradient elution with YBB at 4%, 10%, 20%, 30%, 40%, 60% and 100%. 20% (100 mM imidazole) YBB elution pool was dialysed overnight into 2 x 2L MBA and loaded onto a 5 mL HiTrap Heparin column which was also eluted with a step gradient at 5%, 10%, 15%, 20%, 25%, 30%, 40% and 50% MBB. Material eluting at 15% YBB (300mM NaCl) was pooled and concentrated in a 50 kDa MWCO centricon (GE) and loaded onto a superdex 200 10/300 gl for SEC. Peak fractions were again collected and diluted to <150 mM NaCl with MBA then run over a MonoQ 5/50 gl column with a 20 CV gradient elution with MBB up to 50%.

8.7 Purification of S. cerevisiae Ycs4-6xHis and Brn1 270-423 from E. coli

Expression constructs GC1425 and GC1524 were transformed into BL21 (DE3) codon + (RIL) cells according to manufacturer's instructions. Overnight cultures were inoculated from glycerol stocks and used to inoculate 15 L of 2xTY containing chloramphenicol, ampicillin and kanamycin. Cells were grown to OD 0.6 and induced with 0.5 mM IPTG overnight at 15ºC. Cells were pelleted at 6000g for 20 minutes and stored at -80ºC.

Pellets were thawed and resuspended in YBA (150 mM NaCl) with 0.5 mg/mL lysozyme and benzonase nuclease for an hour. Cells were then sonicated at 40% output for 6 x 2 minutes, 4 seconds on/ 4 seconds off. Lysate was then cleared by centrifugcation at 45000g for 1 hour followed by filtration through 0.2 um syring filters. Cleared lysate was then loaded on to a 5 mL StrepTrap column and washed with 8 CV of YBA (150 mM NaCl), followed by elution with YBA (150 mM NaCl) + 5 mM desthiobiotin.

The following day peak eluted fractions were loaded onto a 5 mL HisTrap column and washed with 4 CV YBA (150 mM NaCl). Step gradient elution was performed with YBB at 4%, 10%, 20%, 40%, 60% 100%. Eluted fractions at 20% (100 mM imidazole) YBB were pooled and concentrated for SEC over a superdex 200 10/300 gl column. Two sets of peak fractions were pooled independently and concentrated with 50 kDa MWCO centricon, followed by freezing in liquid nitrogen and storage at -80^{o}C

8.8 Electrophoretic Mobility Shift Assays

Flourescin labelled 35 bp complimentary oligos were mixed at equimolar concentrations of 10 uM in annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA) and annealed with a temperature gradient from 95^{o}C to 4^{o}C, with temperature decreasing at at 0.1 ^{o}C/s.

Purified recombinant proteins/complexes were mixed at concentrations from 0 nM to 3200 nM with 50 nM of fluorescin labelled dsDNA in 40 mM TRIS–HCl pH 7.5, 125 mM NaCl, 5 mM MgCl2, 10 % glycerol, 1 mM DTT. Reactions were incubated for 10 minutes at room temperature before running on a 0.8% agarose or 1% seakem agarose gel for 1 hour at 110 V in 1x TAE buffer. DNA mobility shift was imaged by fluorescence excitation at 473 nm, 700 V, 25 um pixel size with a fugifilm FLA-5100 reader.

8.9 Cryo EM analysis of the Condensin Complex

Methods detailed in this section were executed by Dr Byung-Gil Lee of the Lowe laboratory on our behalf. The following sections (8.9.1 - 8.9.4) were produced by him and have been reproduced with minimal changes.

8.9.1 Grid preparation

Prior to applying samples onto EM grids, aliquots of purified condensin tetramer or pentamer samples were thawed and injected onto a Superose 6 Increase 3.2/300 column (GE Healthcare) in nucleotide-free buffer (25 mM Tris-HCl, 125 mM NaCl, 1 mM TCEP, pH 7.5). Peak fractions were immediately used for cryo-EM grid preparations, using 3-3.5 µl of sample at a concentration of 0.15 - 0.25 mg ml-1 that was applied to freshly glow-discharged Quantifoil Cu/Rh 2/2 holey carbon 200 mesh grids (Quantifoil). The grids were blotted for 1.5 – 2 s at 4 ˚C with humidity at 100 % and were flesh frozen using an FEI Vitrobot Mark IV (Thermo Fisher) and a liquid-ethane cryostat set to −180°C (Russo et al., 2016).

8.9.2 Data Collection

Images were recorded on Titan Krios electron microscope (FEI) equipped with a K2 or K3 summit direct electron detector (Gatan), mounted behind a GIF Quantum energy filter. Images were collected automatically using software EPU (Thermo Fisher) or SerialEM (Schorb et al., 2019). A total of 13,628 micrographs of condensin tetramer grids were collected with the K2 camera in counting mode through 8 separate data collection sessions with total doses of 44 – 55 electrons per Å2 during exposure times of 8 – 12 s, dose-fractionated into 40 – 55 movie frames, at defocus ranges of 1.6 – 3.6 µm. The magnifications used were 105,000x or 130,000x, resulting in a physical pixel sizes of 1.16, 1.15, 1.047, 1.055, 1.065 Å per pixel. 3 of the eight tetramer datasets with K2 camera (dataset 3, 5, 6, Table S1) were collected at tilts of 40, 30 and 35˚, respectively. 10,050 micrographs of condensin tetramer grids were collected using the K3 camera (dataset 9, Table S1) in super-resolution mode (at a pixel size of 1.085 Å per pixel) with a total dose of 40 electrons per Å2 during total exposure time of 2.25 s, dose-fractionated into 40 movie frames, at a defocus range of 1.8 – 3.3 µm. A total of 10,000 micrographs of condensin pentamer grids were collected with the K2 camera and Volta phase plate (VPP) (Danev et al., 2014) in counting mode (at a pixel size of 1.065 Å per pixel) with a total dose of 45 electrons per Å2 during total exposure time of 9 s, dose-fractionated into 35-42 movie frames. A defocus range of 0.5 – 0.9 µm was used for VPP data collection.

8.9.3 Image processing

Image processing was done in RELION 3.0 (Scheres, 2012). Movies were aligned using 5 x 5 patches using MotionCor2 with dose-weighting (Zheng et al., 2017). CTF parameter were estimated with Gctf (Zhang, 2016). For tilt datasets, local CTF estimation was performed using Gctf after initial processing. All refinements were performed using independent data half-sets (gold-standard refinement) and the resolutions were determined based on the Fourier shell correlation (FSC) = 0.143 criterion (Rosenthal and Henderson, 2003).
Tetramer dataset 2 (750 images) was used for initial model calculation. Particles were picked manually from a few micrographs, and the resulting particles were binned by 1.74 and extracted with a 3602-pixel box size (2 Å/pixel) to generate initial reference-free 2D class averages. Good 2D class averages were used as a template for reference-based auto-picking. The resulting ~ 83 k particles (again from tetramer dataset 2) were subjected to multiple rounds of 2D classifications, and a final ~ 19 k particles were used to calculate the initial 3D model in RELION (see Lee et al., unpuiblished, Fig. S1B).

Particle picking for each dataset was performed separately using the same strategy, starting with auto-picking using the 2D classes from dataset 2; the resulting particles were then downscaled into 4 Å/pixel (1802-pixel box size) for initial processing. 3D classification into 1 class against the 40 Å low-pass filtered initial model was performed with an offset search range of 20 pixel, a limit resolution E-step of 10 Å and an angular sample of 7.5˚. Subsequently, reference-free 2D classification without image alignment was performed on the aligned particles from the 3D classification. After selecting good classes, duplicated particles within 100 Å distance were removed because picking used a very low threshold. For tilted datasets, local CTF estimation was performed with the refined particle coordinates using Gctf followed by re-extraction of particles. The resulting particles from each dataset were then merged for further processing (~ 3M particle images of tetramer and ~ 1.5M of pentamer). The angular assignments resulting from these initial steps were also used for focused 3D classifications for the condensin arm and head, after re-centering particles.

The set of particle images obtained was subjected to further 3D classification into 7 classes with no image alignment. The best class with clear density of Smc2, Smc4, Brn1-C, Ycs4 consisting of 205,786 particles was re-extracted with a 3602-pixel box (2 Å/pixel), and used for a final round of 3D refinement with a soft-edged mask and applying solvent-flattened

FCSs. Post-processing with RELION yielded a final overall map of apo condensin at 7.0 Å resolution.

For focused processing of the condensin 'head' part, containing Smc2-4 ATPases, Brn1 and Ycs4, the particle images from the initial processing were re-centered on the heads and re-extracted with a box size of 320 pixels (1.2 Å/pixel). An initial model for this part of condensin was produced by removing the density of the coiled-coil arms above the joint using UCSF Chimera (Pettersen et al., 2004). The re-centered particle images of the tetramer heads (~3M) were subjected to several rounds of 3D classification. The best class consisting of 403,128 particles was 3D refined to 4.6 Å resolution, followed by CTF refinement and Bayesian polishing in RELION, resulting in a final resolution of 4.3 Å. The pentamer dataset (~ 1.5 M particle images) was processed with the same strategy and resulted in 5 Å resolution reconstruction with 172,888 particles. In the pentamer reconstruction, density for Ycg1 was not clearly visible, presumably because of its mobility with respect to the rest of the condensin complex. The final refined map of the pentamer sample was near-identical to the one obtained for the tetramer. We hence merged both datasets at the level of refined particle images and the merged datasets yielded a final 4.17 Å resolution map.

For the processing of the condensin 'arms', containing the coiled coils and hinge domains, the particles of the tetramer sample from the initial processing were re-centered and re-extracted with a box size of 3002 (1.187 Å/pixel). Again, an initial model was produced using UCSF chimera. After multiple rounds of 3D classifications and 2D classifications with no image alignments, 636,446 particle images were selected and 3D auto-refined, and resulted in the final map at 5.5 Å resolution (Fig. S3B).

8.9.4 Model building

To generate a pseudo-atomic model of the Smc4 ATPase, Brn1 C-terminal domain and Ycs4, previous crystal structures of Smc4-Brn1 (PDB code 6QJ2) from C. thermophilum and Ycs4-Brn1 from C. thermophilum (PDB code 6QJ4) (Hassler et al., 2019) were docked into our cryo-EM head map at 4.2 Å resolution using UCSF chimera. Manual modifications were performed in COOT (Emsley and Cowtan, 2004), including budding yeast sequence assignment, and the residues of 151-365 and 1234-1414 for Smc4, 5-1149 for Ycs4, 643-748 for Brn1 were built manually based on the available density. Homology models of Smc2

ATPase domain and Brn1 N-terminal domain were generated using as templates the crystal structure of Smc3-Scc1 N-terminal domain from S. cerevisiae (PDB code 4UX3) (Gligoris et al., 2014) and the NMR structure of the Brn1 N-terminal domain from C. thermophilum (PDB code 6Q6E) (Hassler et al., 2019), and submitting them to the SWISS-MODEL server (Waterhouse et al., 2018), and were rigid-body docked into our map using Chimera. The stretches 2-212 and 986-1167 for Smc2 and 25-108 for Brn1 were built manually according to the map density.

To generate a pseudo-atomic model of the arm, elbow and hinge regions of condensin, the crystal structure of the hinge domain from S. cerevisiae (PDB code 4RSI, Smc2 residues 453-739, Smc4 residues 560-947) (Soh et al., 2015) was fitted into the 5.5 Å map. Furthermore, the coiled coils were extended up to the elbow, manually building 399-789 for Smc2 and 544-969 for Smc4. The coiled coil model of the region between the joint and the elbow (Smc2 residues 212-383, 792-936 and Smc4 residues 399-539, 773-1114) was first built by placing poly-alanine chains into the density, and the sequences were tentatively assigned based on cross-linking/mass-spec information and the maps appearance. Finally, to obtain an overall atomic model of condensin, the above models of the arm and head regions were fitted into the overall map at 7 Å resultion with Chimera, and the missing residues linking the two parts were manually added using COOT.

8.10 Cross-linking Mass Spectrometry

Cross-linking mass spec experiments were performed on our behalf by Dr. Francis O'Reilly of the Rappsiber laboratory. Descriptions of the following methods are taken from (Lee et al., unpublished) and reported with minor modifications.

The crosslinkers BS3 (bis(sulfosuccinimidyl)suberate) and sulfo-SDA (sulfosuccinimidyl 4,4′-azipentanoate) was purchased from Thermo Scientific Pierce (Rockford, IL, USA). The cross-linkers were dissolved in cross-linking buffer (20 mM HEPES, 150 mM NaCl, 1 mM TCEP, and 5 % glycerol, pH 7.7) to 100 mM before use. For cross-linking with BS3, the purified condensin tetramer or pentamer in cross-linking buffer (20 mM HEPES, 150 mM NaCl, 1 mM TCEP, and 5 % glycerol, pH 7.7) were incubated at 0.7 mg/mL with 2 mM BS3 for 2 h on ice and the reactions were quenched with 50 mM (NH4)HCO3 for 45 min at 4 ˚C.

Reaction products were separated on a Criterion TGX 4-15% SDS-PAGE gel (BioRad). The gel band corresponding to the cross-linked complex was excised and digested with trypsin (Pierce, Germany). The resulting tryptic peptides were extracted and desalted using C18 StageTips (Rappsilber et al., 2003). For photo-crosslinking the nucleotide-free apo condensin with sulfo-SDA, 0.85 mg/ml purified complexes were incubated with sulfo-SDA using three different protein to cross-linker molar ratios (1:250, 1:500, 1:1000) for 2 h at 4 ˚C. The samples were then irradiated with UV light at 365 nm for 20 min and quenched with 50 mM (NH4)HCO3. Subsequently, samples were denatured in 8 M urea, 100 mM (NH4)HCO3 modified with iodoacetamide and digested with LysC endoproteinase (Wako) for 4 h. After dilution of the sample to a urea concentration of 1.5 M, trypsin (Promega) was added and the sample digested for 16 h at 30 °C. The resulting tryptic peptides were extracted and desalted using C18 StageTips (Rappsilber et al., 2003).

Eluted peptides were fractionated using a Superdex Peptide 3.2/300 (GE Healthcare) at a flow rate of 10 µL/min using 30 % (v/v) acetonitrile and 0.1 % (v/v) trifluoroacetic acid as mobile phase. 50 µl fractions were collected and dried. Samples for analysis were resuspended in 0.1 % (v/v) formic acid 1.6 % (v/v) acetonitrile. LC-MS/MS analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, Germany) coupled on-line with an Ultimate 3000 RSLCnano system (Dionex, Thermo Fisher Scientific, Germany). The sample was separated and ionized by a 50 cm EASY-Spray column (Thermo Fisher Scientific). Mobile phase A consisted of 0.1 % (v/v) formic acid and mobile phase B of 80 % (v/v) acetonitrile with 0.1 % (v/v) formic acid. Flow-rates were 0.3 µL/min using gradients optimized for each chromatographic fraction from offline fractionation, ranging from 2 % mobile phase B to 45 % mobile phase B over 90 min. MS data were acquired in data-dependent mode using the top-speed setting with a three second cycle time. For every cycle, the full scan mass spectrum was recorded using the the Orbitrap at a resolution of 120,000 in the range of 400 to 1,500 m/z. Ions with a precursor charge state between 3+ and 7+ were isolated and fragmented. Fragmentation by Higher-energy Collisional Dissociation (HCD) employed a decision tree logic with optimized collision energies (Kolbowski et al., 2017). The fragmentation spectra were then recorded in the Orbitrap with a resolution of 50,000. Dynamic exclusion was enabled with single repeat count and 60-second exclusion duration. A recalibration of the precursor m/z was conducted based on high-confidence (<1% false discovery rate (FDR)) linear peptide identifications. The recalibrated peak lists were searched against the sequences and the reversed sequences

(as decoys) of crosslinked peptides using the Xi software suite (v.1.6.745) for identification (Mendes et al., 2019). Final crosslink lists were compiled using the identified candidates filtered to 1% FDR on link level with xiFDR v.1.4.1 (Fischer and Rappsilber, 2017).

8.11 Western Blotting

Proteins were first resolved on either 4-12% gradient (invitrogen) or 15% SDS PAGE gels. Proteins were then transferred to polyvinylidene fluoride (PVDF) membranes (Hybond-P, Amersham Biosciences) using a TE 77 semi-dry transfer system (Amersham Biosciences). PVDF membranes were pre-soaked in methanol and stacked with filter paper soaked in 1x tris-glycine Transfer buffer (National Diagnostics). Samples were run for 1 hour at 60 mA per gel.

Prestaining of membranes was done with Ponceau S stains to check the quality of transfer, followed by washing in TBST (TBS with 0.1% Tween 20) to remove stain and blocking in 5% skimmed milk powder for 30 minutes. Primary antibody in fresh 5% milk TBST solution was added and incubated with the membrane overnight at 4 ºC with agitation. The following morning membranes were washed 3 x for 10 minutes in TBST to remove unbound antibody and incubated with secondary antibody Horseradish peroxidiase (HRP) conjugates for 1 hour at room temperature in 5 % milk TBST solution. Unbound secondary antibodies were removed by a further 3 x 10 minute was in TBST.

Antibody bound membranes were imaged by adding ECL Plus/Prime Western Blotting Detection System (GE Healthcare) followed by exposure to ECL Hyper Film (GE Healthcare). Alternatively, membranes were imaged directly on an Amersham imager 680 system. For lower signal antibodies, high sensitivity Luminata Forte (Millipore) was used.

8.12 Condensin Force-clamp and Force-extension Optical Tweezer Experiments

Optical tweezer experiments were performed on the Lumix C rap system. Initially, the flow cell was passivated with 0.50% pluronic and 2 mg/ml BSA in PBS solution. Streptavidin-coated polyesterene beads (4.42 μm in diameter, Spherotech) and biotin-labeled double-

stranded λ-DNA were then flowed into 2 of the 5 laminar flow channels, with experimental conditions in the remaining 3.

DNA was captured by laser trapped beads and incubated with 5 nM condensin in Tweezer buffer (50 mM tris pH 7.5, 50 mM NaCl, 1 mM DTT, 0.5 mg/mL BSA, 2.5 mM MgCl2), with and without 2 mM ATP. For force clamping experiments, beads were held at 14 um distance (DNA contour length) and a constant stretching force of 1 pN was applied. Bead distance was then recorded over time using BlueLake software. Following force clamping, DNA compaction was reversed by manual moving of beads back to 14 um distance with recording of force-extension measurements.

8.13 Condensin Translocation Kymographs

Condensin translocation assays were also performed on the Lumix C-trap system with passivation and DNA capture as described in 8.12. A fixed DNA substrate was created by setting beads 14 um apart in a channel with 5 nM condensin in Tweezer buffer with or without 6 mM ATP. Condensin was visualised by inclusion of 1 uL/mL Qdot 705 labelled anti-HA monoclonal antibody (Roche). Antibody labelling was carried out according to manufacturer's instruction using the SiteClick 705 antibody labelling kit (Thermo Fisher).

8.14 S. cerevisiae Histone Octamer purification

Plasmids pETduet.H3-H4 and pCDFduet.H2A-H2B (Diffley lab) were transformed into BL21 (DE3) codon + (RIL) *E. coli* following manufacturer’s instructions. 50 mL overnight cultures were inoculated with glycerol stocks and grown overnight at 37ºC. The following morning, 1-2 L of 2xTY media with chloramphenicol, ampicillin and spectinomycin was inoculated from overnight cultures and grown to OD 0.6. Cells were then induced with 0.3 mM IPTG for 3 hours and cells were pelleted by 20-minute centrifugation at 6000g then stored at -80ºC.

Purification began by thawing pellets on ice, followed by resuspension in Histone Buffer A (HBA) (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 0.1 mM EDTA, 10 mM BME, and PIC

(roche)). Cells were lysed by sonication for 3 x 5 minutes, 5 seconds on 5 seconds off in 10 – 15 mL aliquots, 15% output (Branson digital sonicator, microprobe). Lysis was checked by 100x oli immersion light microscopopy. Lysate was then incubated with benzonase nuclease for 40 minutes following lysis to remove DNA contamination.

Lysates were then cleared by centrifugation for 1 hour at 45000g followed by filtration through 0.2 um syringe filters (sartorius). Cleared lysates were loaded onto a 5 mL HiTrap heparin column followed by 4 column volumes of HBA to wash. Elution was performed with a gradient from 0 – 100% Histone Buffer B (HBB) (20 mM Tris-HCl pH 8.0, 2 M NaCl, 0.1 mM EDTA, 10 mM BME). Peak fractions were pooled and concentrated with a 50 kDa MWCO centricon (GE) for SEC over a superdex 200 increase 10/300 gl column equilibrated in HBB. Peak fractions were again pooled and concentrated to ~4 – 6 mg/mL then mixed 1:1 with 100% glycerol and stored at -20^{o}C.

8.15 Individual Histone Purification

pET 28 expression vectors harbouring genes for *S. cerevisiae* histones H2A, H2B and H3 were transformed into BL21 (DE3) codon + (RIL) *E. coli* following manufacturer's instructions. A pET 28 vector containing histone H4 was also transformed into BL21 (DE3) pLysS. *E. coli* were grown to OD 0.6 and induced for 4 hours with 0.2 mM IPTG. Cells were pellet by centrifugation at 6000g for 20 minutes and inclusion bodies formed according to methods in (Luger et al., 1999).

Pellets were resuspended in Histone Wash Buffer (HWB) (50 mM Tris pH 7.5, 100 mM NaCl, 1 mM EDTA) and cells were lysed by sonication for 2 minutes, 3 seconds on/ 3 seconds off, 20% output. Insoluble material was collected by centrifugation at 12000g for 20 minutes. The insoluble pellet was then resuspended with HWB + 1% triton x100 to remove contaminants and sonicated for a further 30 seconds. Again, insoluble material was collected by centrifugation. This process was repeated twice with HWB + 1% titron x100 and twice with HWB only. The final supernatant was discarded, and the pelleted inclusion bodies stored at -80^{o}C.

Histones were extracted from inclusion bodies by gentle mixing with DMSO followed by resuspension in 7M guanidine HCl, 20 mM Tris pH 7.5, 10 mM DTT. Insoluble material was removed by centrifugation twice at 30000 g for 20 minutes, with supernatant being collected and filtered through 0.45 um syringe filters. Filtered material was then separated over a Sephacryl s200 gel filtration column pre-equilibiratied in SAU 200 (20 mM NaAc, 7 M urea, 200 mM NaCl, 5 mM BMET and 1 mM EDTA) with peak fractions appearing after the major void volume peak being collected. These fractions were then applied to a HiPrep 16/60 SP column for ion exchange. Histones H2A, H2B and H3 were eluted over a 5 CV gradient between 200 and 600 mM NaCl, H4 between 200 mM and 1M. Peak fractions were collected and dialysed into 3 changes of H_2O with 1 mM DTT then aliquoted and lyophilised for storage (Luger et al., 1999).

8.16 Nap1 purification

pGEX-6p-1 vectors containing GST- Nap1 were transformed into BL21 (DE3) codon + (RIL) according to manufacturers instructions (CC5, Diffley lab). Cells were grown to OD 0.4 and induced with 0.4 mM IPTG for 4 hours at 20 ºC. Cells were collected by centrifugation at 6000g for 20 minutes and stored at -80 ºC.

Protein was extracted by resuspending pellets in Nap1 Buffer A (NBA) (50 mM KPO4 pH 7.6, 150 mM KOAc, 5 mM MgCl2, 1% triton x100, 1 mM DTT, protease inhibitor cocktail (sigma)) with 1 mg/mL lysozyme and benzonase nuclease for 1 hour, followed by sonication for 2 x 1 minute at 20% output, 5 seconds on/off (Brnason digital sonicator, microprobe). Lysates were cleared by centrifugation at 30000g at 4 ºC for 40 minutes. Supernatants were collected and spun again for 20 minutes. Lysates were then incubated for 2 hours at 4 ºC with glutathione agarose beads before washing beads with 4 CV of NBA. Nap1 was then eluted in 1 CV of NBA with PreScission protease by incubating overnight.

Eluent was then dialysed overnight into 3 x 2 L of Nap1 MonoQ Buffer A (NMBA) (20 mM Tris pH 7.5, 100 mM NaCl, 0.5 mM EDTA, 10% glycerol, 1mM DTT, 0.1 mM PMSF), before loading onto a MonoQ 5/50 anion exchange column and washing with 5 CV of NMBA. Nap1 was then eluted with a 20 CV linear gradient between NMBA and 100 % NMBB with (20 mM Tris pH 7.5, 1 M NaCl, 0.5 mM EDTA, 10% glycerol, 1mM DTT, 0.1

mM PMSF). Peak fractions were collected and concentrated to ~ 3-4 mg/mL in a viva spin centricon, aliquoted and stored at -80C (Kurat et al, 2017. Vary et al., 2004).

8.17 ISWIa complex purification

Yeast strains containing Ioc3-FLAG under endogenous promoters (CCG14634 (yCFK1), Diffley lab) were grown in YPD media overnight to an OD of between 7 and 9. Cells were resuspended in an equal volume of ISWI Buffer A (IBA) (25 mM HEPES pH 7.6, 300 mM KCl, 0.1 mM EDTA, 2 mM MgCl2, 20 % glycerol, 0.02 % NP40 and 1 mM DTT) and frozen in liquid nitrogen. Cells were then lysed by freezer milling.

Lysed material was melted at 4ºC with additional IBA added up to 80 mL with benzonse for 40 minutes. Lysate was cleared by centrifugation at 45000g for 1 hour followed by filtration through 0.45 um then 0.2 um syringe filters. Cleared lysate was then incubated for 3 hours at 4ºC with 2.25 ml pre-equilibrated EZview Red anti-FLAG M2 beads (sigma) followed by 2 x 20 mL washes with IBA and 2 x 20 mL washes with IBB (25 mM HEPES pH 7.6, 100 mM KCl, 0.1 mM EDTA, 2 mM MgCl2, 20 % glycerol, 0.02 % NP40 and 1 mM DTT). Elution was done at 9ºC in a thermoshaker with 4 x 30-minute incubations in IBB + 1 mg/mL of 3 x FLAG peptide.

Pooled elutions were then loaded onto a MonoQ 5/50 column equilibirated in IBB. Column was washed with 4 CV of IBB before eluting with a 20 CV linear gradient from 0-100% IBC (25 mM HEPES pH 7.6, 600 mM KCl, 0.1 mM EDTA, 2 mM MgCl2, 20 % glycerol, 0.02 % NP40 and 1 mM DTT). Peak fraction were pooled and concentrated for storage at -80ºC (Kurat et al., 2017, Vary et al., 2004).

8.18 In vitro chromatin assembly

Chromatinization reactions were constituted by mixing 300 ng of recombinant histone octamer, 1.8 ug of recombinant Nap1 and 10 ng of recombinant ISWIa complex in Assembly Buffer (10 mM Hepes pH 7.6, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM EGTA, 10% glycerol and 0.1 mg/mL BSA). Reactions were mixed by gentle agitation and left on ice for 30 minutes.

ATP was then added to final concentration of 3 mM with 2000 unit/mL of creatine phosphokinase and 40 mM creatine phosphate. 300 ng of linearized plasmid (pRS426) was also added and the reaction was made up to 30 uL with Assembly Buffer. The reaction was then incubated at 30ºC in a thermoshaker for 4 hours (Kurat et al., 2017)

8.19 Micrococcal nuclease digestion assay

Assembled chromatin was assayed by digestion with micrococcal nuclease (MNase) immediately following assembly. $CaCl_2$ was added to reaction mixtures to a final concentration of 5 mM. 50 units of MNase (NEB) were then added and mixed by gentle agitation, followed by a 5-minute incubation of the whole reaction at 37ºC. The MNase digest was quenched by adding an excess of EGTA to final concentration of ~40 mM. Undigested DNA was then released from nucleosomes by adding proteinase K and incubating at 50ºC for 30 minutes. Undigested DNA fragments were recovered by purification by Qiaquick PCR column and run on a 1.3% agarose gel for 50 minutes at 100 V before imaging.

8.20 Histone labelling

Stored histone octamers containing histone H2A with a C terminal cysteine residue were dialysed into Histone Labelling Buffer (HLB) (10 mM Tris pH 7, 2 M NaCl and 1 mM EDTA) to remove 50% glycerol in the storage buffer. Protein concentration was re-quantified by Quick Start Bradford reagent assay and octamers diluted with HLB to ~ 25 uM (2x H2A). 1 mM TCEP was added for 20 minutes to reduce disulphide bonds.

Alexa Fluor 532 maleimide (Thermo Fischer) was dissolved to a stock concentration of 20 mM then added to the histone octamers to give a final dye concentration of ~250 uM (5x molar excess). Labelling reactions were then allowed to proceed at room temperature for 2 hours.

Excess dye and disassembled histones were then removed by SEC over a superdex 200 10/300 gl column, with the main octamer peak being collected, concentrated to ~5 mg/mL and mixed with glycerol to a final concentration of 50%. Octamers were then stored at –20 ^{o}C. Integration of label was assayed by running an SDS PAGE gel of octamers and scanning on a Fugifilm FLA 5100 for fluorescence excitation at the expected molecular weight of H2A.

8.21 Protein quantification

Protein concentrations were quantified by three methods. Colorimetric Quickstart Bradford reagent assays (biorad) were used to compare a 1/100 dilutions of protein solution in 500 uL Bradford reagent with a BSA standard curve between 0 and 2 mg/mL according to manufacturer's instructions. Protein samples were left to develop in reagent for 5 minutes at room temperature before measuring absorbance at 595 nm. Alternatively, Qbit protein assays (Invitrogen) were used for quantification according to manufacturer's instructions.

For individual histones purifications, final material following dialysis into water with 1 mM DTT was quantified by measuring absorbance at 280 nm. Cocnentrations were determined and using the Beer Lamberts equation ($A = \varepsilon \ell c$) with extinction coefficients generated from sequences in ProtParam.

8.22 Tables of Strains, Constructs, Antibodies and Rights and Permissions

Table 2. List of yeast strains

Strain	Genotype	Source
W303-1a	MATa {leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15}	Haering lab (Terakawa et al., 2017)
CCG14634 (yCFK1)	MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100 bar1::Hyg pep4::KanMX IOC3-3xFLAG-NAT	Diffley lab (Kurat et al., 2016)

CCG14540	MATα ade2-1 ura 3-1 his 3-11,15 trp1-1 leu2-3,112 cqn 1-100 bar1::Hyg pep4::KanMx ura::URA-Spt16-GAL10-3xFLAG-TEV-Pob3	Diffley lab (Kurat et al., 2016)
CCG14619	Mata Lys2::pGAL1-GAL4::LYS2 pep4::HIS3 bar1::hisG ade2-1 trp1D2 can1-100 leu 2-3,112 ura3-52 [GAL7-SMC4-strep-SMC2-GAL-10-BRN1-HA-His URA3 2micron plasmid] [GAL1-YCG1 GAL10-YCS4 TRP1 2micron]	Haering lab (terakawa et al., 2017)
CCG14606	Mata Lys2::pGAL1-GAL4::LYS2 pep4::HIS3 bar1::hisG ade2-1 trp1D2 can1-100 leu 2-3,112 ura3-52 [GAL7-SMC4-strep-SMC2-GAL-10-BRN1-HA-His URA3 2micron plasmid]	This study
CCG14658	Mata Lys2::pGAL1-GAL4::LYS2 pep4::HIS3 bar1::hisG ade2-1 trp1D2 can1-100 leu 2-3,112 ura3-52 [GAL7-SMC4-strep-SMC2-GAL-10-BRN1-HA-His URA3 2micron plasmid] [GAL10-YCG1 TRP1 2micron]	This study
CCG14659	Mata Lys2::pGAL1-GAL4::LYS2 pep4::HIS3 bar1::hisG ade2-1 trp1D2 can1-100 leu 2-3,112 ura3-52 [GAL7-SMC4-strep-SMC2-GAL-10-BRN1-HA-His URA3 2micron plasmid] [GAL10-YCS4 TRP1 2micron]	This study

Table 3. List of constructs

Construct	Genotype	Source
GC1267	URA3 leu2-d pGAL7-SMC4-StrepII3 pGAL10-SMC2 pGAL1-BRN1-HA3-His12	Haering lab (Terakawa et al., 2017)
GC1268	TRP1 leu2-d pGAL10-YCS4 pGAL1-YCG1	Haering lab (Terakawa et al., 2017)
GC1259	TRP1 pGAL10-YCS4	This study
GC1258	TRP1 pGAL10-YCG1	This study
GC1425	pET30a (+)	This study
GC1425	pFastBac1-Ycs4-6xHis	This study
GC1376	pGEX-6-p1 Ycs4 (*Picchia pastoris*)	This study
GC1377	pGEX-6-p1 Ycs4 (*Candida albicans*)	This study
GC1378	pGEX-6-p1 CAP D2 (*Homo sapiens*)	This study
GC1379	pGEX-6-p1 Ycs4 (*Saccharomyces cerevisiae*)	This study
CC1	pPRO-Ex-HTB Ycs4 (*Picchia pastoris*)	This study
CC2	pPRO-Ex-HTB Ycs4 (*Candida albicans*)	This study
CC3	pPRO-Ex-HTB CAP D2 (*Homo sapiens*)	This study
CC4	pPRO-Ex-HTB Ycs4 (*Saccharomyces cerevisiae*)	This study
GC1419	pCDF duet-1 H2A-H2B	Diffley lab (Kurat et al., 2017)
GC1406	pET Duet H3-H4	Diffley lab (Kurat et al., 2017)
GC1508	pCDF duet-1 H2Acys H2B	This study
CC5	pGEX-6-p1 Nap1	Diffley lab (Kurat et al., 2017)
GC1522	pET 20b Smc4 hd 148-351-GSGSG-1249-1414-StrepII	This study
GC1523	pET 20b Smc4 1-351-GSGSG-1249-1418-StrepII	This study
GC1524	pET 20b Brn1 270-423-StrepII	This study

Table 4. List of antibodies

Antibody	Source
Anti-His tag mouse monoclonal	Abcam (ab18184)
Anti-StrepII tag rabbit monoclonal	Abcam (ab180957)
Anti-HA tag rat monoclonal	Roche (11867423001)

Table 5. Rights and Permissions

Page No.	Type of work	Name of Work	Source of Work	Copyright Holder	Permission Request Date	Permission Granted?	Permission Note
18	Figure	Figure 1 (A-C), SMC architecture and coil-coil packing interactions	Waldman et al., 2015	Wiley Publications Inc.	17/04/20	Yes	Written permission via RightsLink. Licence number: 4811390155138

Bibliography

Abdennur, N., W. Schwarzer, A. Pekowska, I. A. Shaltiel, W. Huber, C. H. Haering, L. Mirny and F. Spitz (2018). "Condensin II inactivation in interphase does not affect chromatin folding
or gene expression." BioRxiv: 437459.

Abe, S., K. Nagasaka, Y. Hirayama, H. Kozuka-Hata, M. Oyama, Y. Aoyagi, C. Obuse and T. Hirota (2011). "The initial phase of chromosome condensation requires Cdk1-mediated phosphorylation of the CAP-D3 subunit of condensin II." Genes Dev **25**(8): 863-874.

Akiyoshi, B., K. K. Sarangapani, A. F. Powers, C. R. Nelson, S. L. Reichow, H. Arellano-Santoyo, T. Gonen, J. A. Ranish, C. L. Asbury and S. Biggins (2010). "Tension directly stabilizes reconstituted kinetochore-microtubule attachments." Nature **468**(7323): 576-579.

Albritton, S. E. and S. Ercan (2018). "Caenorhabditis elegans dosage compensation: insights into condensin-mediated gene regulation." Trends in Genetics **34**(1): 41-53.

Alt, A., H. Q. Dang, O. S. Wells, L. M. Polo, M. A. Smith, G. A. McGregor, T. Welte, A. R. Lehmann, L. H. Pearl and J. M. Murray (2017). "Specialized interfaces of Smc5/6 control hinge stability and DNA association." Nature communications **8**(1): 1-14.

Ambudkar, S. V., I.-W. Kim, D. Xia and Z. E. Sauna (2006). "The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker A motif in ABC transporters, is critical for ATP binding." FEBS letters **580**(4): 1049-1055.

Aono, N., T. Sutani, T. Tomonaga, S. Mochida and M. Yanagida (2002). "Cnd2 has dual roles in mitotic condensation and interphase." Nature **417**(6885): 197-202.

Arumugam, P., S. Gruber, K. Tanaka, C. H. Haering, K. Mechtler and K. Nasmyth (2003). "ATP hydrolysis is required for cohesin's association with chromosomes." Curr Biol **13**(22): 1941-1953.

Arumugam, P., T. Nishino, C. H. Haering, S. Gruber and K. Nasmyth (2006). "Cohesin's ATPase Activity Is Stimulated by the C-Terminal Winged-Helix Domain of Its Kleisin Subunit." Current Biology **16**(20): 1998-2008.

Baergen, A. K., L. M. Jeusset, Z. Lichtensztejn and K. J. McManus (2019). "Diminished Condensin Gene Expression Drives Chromosome Instability That May Contribute to Colorectal Cancer Pathogenesis." Cancers **11**(8): 1066.

Baneyx, F. and Mujacic, M., (2004). Recombinant protein folding and misfolding in Escherichia coli. Nature Biotechnology, 22(11), pp.1399-1408.

Banigan, E. J. and L. A. Mirny (2019). "Limits of chromosome compaction by loop-extruding motors." Physical Review X **9**(3): 031007.

Banigan, E. J., A. A. van den Berg, H. B. Brandão, J. F. Marko and L. A. Mirny (2019). "Chromosome organization by one-sided and two-sided loop extrusion." bioRxiv: 815340.

Barnhart-Dailey, M. C., P. Trivedi, P. T. Stukenberg and D. R. Foltz (2017). "HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition." Molecular biology of the cell **28**(1): 54-64.

Barysz, H., J. H. Kim, Z. A. Chen, D. F. Hudson, J. Rappsilber, D. L. Gerloff and W. C. Earnshaw (2015). "Three-dimensional topology of the SMC2/SMC4 subcomplex from chicken condensin I revealed by cross-linking and molecular modelling." Open biology **5**(2): 150005.

Batty, P. and D. W. Gerlich (2019). "Mitotic chromosome mechanics: how cells segregate their genome." Trends in cell biology.

Bauer, C. R., T. A. Hartl and G. Bosco (2012). "Condensin II promotes the formation of chromosome territories by inducing axial compaction of polyploid interphase chromosomes." PLoS genetics **8**(8).

Bauerschmidt, C., C. Arrichiello, S. Burdak-Rothkamm, M. Woodcock, M. A. Hill, D. L. Stevens and K. Rothkamm (2010). "Cohesin promotes the repair of ionizing radiation-induced DNA double-strand breaks in replicated chromatin." Nucleic Acids Res **38**(2): 477-487.

Baxter, J. and L. Aragón (2012). "A model for chromosome condensation based on the interplay between condensin and topoisomerase II." Trends in Genetics **28**(3): 110-117.

Baxter, J., N. Sen, V. L. Martínez, M. M. De Carandini, J. Schvartzman, J. Diffley and L. Aragon (2011). "Positive supercoiling of mitotic DNA drives decatenation by topoisomerase II in eukaryotes." Science **331**(6022): 1328-1332.

Bennink, M., Leuba, S., Leno, G., Zlatanova, J., de Grooth, B. and Greve, J., 2001. Nature Structural Biology, 8(7), pp.606-610.

Bernad, R., P. Sánchez, T. Rivera, M. Rodríguez-Corsino, E. Boyarchuk, I. Vassias, D. Ray-Gallet, A. Arnaoutov, M. Dasso and G. Almouzni (2011). "Xenopus HJURP and condensin II are required for CENP-A assembly." Journal of Cell Biology **192**(4): 569-582.

Bhalla, N., S. Biggins and A. W. Murray (2002). "Mutation of YCS4, a budding yeast condensin subunit, affects mitotic and nonmitotic chromosome behavior." Molecular biology of the cell **13**(2): 632-645.

Bloom, J. and F. R. Cross (2007). "Multiple levels of cyclin specificity in cell-cycle control." Nature reviews Molecular cell biology **8**(2): 149-160.

Bloom, K. and V. Costanzo (2017). Centromere structure and function. Centromeres and Kinetochores, Springer**:** 515-539.

Bolcun-Filas, E. and M. A. Handel (2018). "Meiosis: the chromosomal foundation of reproduction." Biology of reproduction **99**(1): 112-126.

Bürmann, F., A. Basfeld, R. Vazquez Nunez, M.-L. Diebold-Durand, L. Wilhelm and S. Gruber (2017). "Tuned SMC Arms Drive Chromosomal Loading of Prokaryotic Condensin." Molecular Cell **65**(5): 861-872.e869.

Bürmann, F., B.-G. Lee, T. Than, L. Sinn, F. J. O'Reilly, S. Yatskevich, J. Rappsilber, B. Hu, K. Nasmyth and J. Löwe (2019). "A folded conformation of MukBEF and cohesin." Nature Structural & Molecular Biology **26**(3): 227-236.

Bürmann, F., H.-C. Shin, J. Basquin, Y.-M. Soh, V. Giménez-Oya, Y.-G. Kim, B.-H. Oh and S. Gruber (2013). "An asymmetric SMC–kleisin bridge in prokaryotic condensin." Nature Structural &Amp; Molecular Biology **20**: 371.

Buster, D. W., S. G. Daniel, H. Q. Nguyen, S. L. Windler, L. C. Skwarek, M. Peterson, M. Roberts, J. H. Meserve, T. Hartl and J. E. Klebba (2013). "SCFSlimb ubiquitin ligase suppresses condensin II–mediated nuclear reorganization by degrading Cap-H2." Journal of Cell Biology **201**(1): 49-63.

Cabello, O. A., E. Eliseeva, W. He, H. Youssoufian, S. E. Plon, B. Brinkley and J. W. Belmont (2001). "Cell cycle-dependent expression and nucleolar localization of hCAP-H." Molecular biology of the cell **12**(11): 3527-3537.

Chan, K.-L., T. Gligoris, W. Upcher, Y. Kato, K. Shirahige, K. Nasmyth and F. Beckouët (2013). "Pds5 promotes and protects cohesin acetylation." Proceedings of the National Academy of Sciences **110**(32): 13020-13025.

Chan, R. C., A. F. Severson and B. J. Meyer (2004). "Condensin restructures chromosomes in preparation for meiotic divisions." The Journal of cell biology **167**(4): 613-625.

Chapard, C., R. Jones, T. van Oepen, J. C. Scheinost and K. Nasmyth (2019). "Sister DNA entrapment between juxtaposed smc heads and kleisin of the cohesin complex." Molecular cell **75**(2): 224-237. e225.

Charbin, A., C. Bouchoux and F. Uhlmann (2014). "Condensin aids sister chromatid decatenation by topoisomerase II." Nucleic acids research **42**(1): 340-348.

Chen, J., G. Lu, J. Lin, A. L. Davidson and F. A. Quiocho (2003). "A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle." Molecular cell **12**(3): 651-661.

Ciosk, R., M. Shirayama, A. Shevchenko, T. Tanaka, A. Toth, A. Shevchenko and K. Nasmyth (2000). "Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins." Molecular cell **5**(2): 243-254.

Clemente-Blanco, A., M. Mayán-Santos, D. A. Schneider, F. Machín, A. Jarmuz, H. Tschochner and L. Aragón (2009). "Cdc14 inhibits transcription by RNA polymerase I during anaphase." Nature **458**(7235): 219-222.

Crane, E., Q. Bian, R. P. McCord, B. R. Lajoie, B. S. Wheeler, E. J. Ralston, S. Uzawa, J. Dekker and B. J. Meyer (2015). "Condensin-driven remodelling of X chromosome topology during dosage compensation." Nature **523**(7559): 240-244.

Cremer, T. and M. Cremer (2010). "Chromosome territories." Cold Spring Harbor perspectives in biology **2**(3): a003889.

Cuylen, S., J. Metz and C. H. Haering (2011). "Condensin structures chromosomal DNA through topological links." Nat Struct Mol Biol **18**(8): 894-901.

D'Amours, D., F. Stegmeier and A. Amon (2004). "Cdc14 and condensin control the dissolution of cohesin-independent chromosome linkages at repeated DNA." Cell **117**(4): 455-469.

D'Ambrosio, C., C. K. Schmidt, Y. Katou, G. Kelly, T. Itoh, K. Shirahige and F. Uhlmann (2008). "Identification of cis-acting sites for condensin loading onto budding yeast chromosomes." Genes & development **22**(16): 2215-2227.

Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J.M., and Baumeister, W. (2014). Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proc. Natl. Acad. Sci. U.S.A. 111, 15635–15640.

Davidson, I. F., B. Bauer, D. Goetz, W. Tang, G. Wutz and J.-M. Peters (2019). "DNA loop extrusion by human cohesin." Science **366**(6471): 1338-1345.

Davidson, I. F., D. Goetz, M. P. Zaczek, M. I. Molodtsov, F. Weissmann, G. Litos, D. A. Cisneros, M. Ocampo-Hafalla, R. Ladurner and F. Uhlmann (2016). "Rapid movement and transcriptional re-localization of human cohesin on DNA." The EMBO journal **35**(24): 2671-2685.

de Los Santos-Velazquez, A. I., I. G. de Oya, J. Manzano-López and F. Monje-Casas (2017). "Late rDNA condensation ensures timely Cdc14 release and coordination of mitotic exit signaling with nucleolar segregation." Current Biology **27**(21): 3248-3263. e3245.

De Piccoli, G., F. Cortes-Ledesma, G. Ira, J. Torres-Rosell, S. Uhle, S. Farmer, J.-Y. Hwang, F. Machin, A. Ceschia, A. McAleenan, V. Cordon-Preciado, A. Clemente-Blanco, F. Vilella-Mitjana, P. Ullal, A. Jarmuz, B. Leitao, D. Bressan, F. Dotiwala, A. Papusha, X. Zhao, K. Myung, J. E. Haber, A. Aguilera and L. Aragón (2006). "Smc5–Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination." Nature Cell Biology **8**: 1032.

Deutschman, E., J. R. Ward, K. T. Ho-A-Lim, T. J. Alban, D. Zhang, B. Willard, M. E. Lemieux, J. D. Lathia and M. S. Longworth (2018). "Comparing and contrasting the effects of Drosophila condensin II subunit dCAP-D3 overexpression and depletion in vivo." Genetics **210**(2): 531-546.

Dewari, P. S. and P. Bhargava (2014). "Genome-wide mapping of yeast histone chaperone anti-silencing function 1 reveals its role in condensin binding with chromatin." PloS one **9**(9).

Dixon, J. R., D. U. Gorkin and B. Ren (2016). "Chromatin domains: the unit of chromosome organization." Molecular cell **62**(5): 668-680.

Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu and B. Ren (2012). "Topological domains in mammalian genomes identified by analysis of chromatin interactions." Nature **485**(7398): 376-380.

Doughty, T. W., H. E. Arsenault and J. A. Benanti (2016). "Levels of Ycg1 limit condensin function during the cell cycle." PLoS genetics **12**(7).

Dowen, Jill M., S. Bilodeau, David A. Orlando, Michael R. Hübner, Brian J. Abraham, David L. Spector and Richard A. Young (2013). "Multiple Structural Maintenance of Chromosome Complexes at Transcriptional Regulatory Elements." Stem Cell Reports **1**(5): 371-378.

Doyle, J. M., J. Gao, J. Wang, M. Yang and P. R. Potts (2010). "MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases." Molecular cell **39**(6): 963-974.

Duesberg, P. and D. Rasnick (2000). "Aneuploidy, the somatic mutation that makes cancer a species of its own." Cell motility and the cytoskeleton **47**(2): 81-107.

Earnshaw, W. C. and U. K. Laemmli (1984). "Silver staining the chromosome scaffold." Chromosoma **89**(3): 186-192.

Eeftens, J. M., S. Bisht, J. Kerssemakers, M. Kschonsak, C. H. Haering and C. Dekker (2017). "Real-time detection of condensin-driven DNA compaction reveals a multistep binding mechanism." EMBO J **36**(23): 3448-3457.

Eeftens, J. M., A. J. Katan, M. Kschonsak, M. Hassler, L. de Wilde, E. M. Dief, C. H. Haering and C. Dekker (2016). "Condensin Smc2-Smc4 Dimers Are Flexible and Dynamic." Cell Rep **14**(8): 1813-1818.

Elbatsh, A. M. O., E. Kim, J. M. Eeftens, J. A. Raaijmakers, R. H. van der Weide, A. García-Nieto, S. Bravo, M. Ganji, J. uit de Bos, H. Teunissen, R. H. Medema, E. de Wit, C. H. Haering, C. Dekker and B. D. Rowland (2019). "Distinct Roles for Condensin's Two ATPase Sites in Chromosome Condensation." Molecular Cell.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Fennell-Fezzie, R., S. D. Gradia, D. Akey and J. M. Berger (2005). "The MukF subunit of Escherichia coli condensin: architecture and functional relationship to kleisins." The EMBO journal **24**(11): 1921-1930.

Fischer, L., and Rappsilber, J. (2017). Quirks of Error Estimation in Cross-Linking/Mass Spectrometry. Anal. Chem. 89, 3829–3833.

Flaus, A., (2011). Principles and practice of nucleosome positioningin vitro. Frontiers in Life Science, 5(1-2), pp.5-27.

Flyamer, I. M., J. Gassler, M. Imakaev, H. B. Brandão, S. V. Ulianov, N. Abdennur, S. V. Razin, L. A. Mirny and K. Tachibana-Konwalski (2017). "Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition." Nature **544**(7648): 110-114.
Freeman, L., L. Aragon-Alcaide and A. Strunnikov (2000). "The condensin complex governs chromosome condensation and mitotic transmission of rDNA." The Journal of cell biology **149**(4): 811-824.

Fousteri, M., 2000. A novel SMC protein complex in Schizosaccharomyces pombe contains the Rad18 DNA repair protein. *The EMBO Journal*, 19(7), pp.1691-1702.

Fudenberg, G., M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur and L. A. Mirny (2016). "Formation of chromosomal domains by loop extrusion." Cell reports **15**(9): 2038-2049.
Gajiwala, K. S. and S. K. Burley (2000). "Winged helix proteins." Current opinion in structural biology **10**(1): 110-116.

Ganji, M., I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering and C. Dekker (2018). "Real-time imaging of DNA loop extrusion by condensin." Science **360**(6384): 102-105.

Gao, C., Y. Su, J. Koeman, E. Haak, K. Dykema, C. Essenberg, E. Hudson, D. Petillo, S. K. Khoo and G. F. V. Woude (2016). "Chromosome instability drives phenotypic switching to metastasis." Proceedings of the National Academy of Sciences **113**(51): 14793-14798.

Gerlich, D., T. Hirota, B. Koch, J.-M. Peters and J. Ellenberg (2006). "Condensin I stabilizes chromosomes mechanically through a dynamic interaction in live cells." Current Biology **16**(4): 333-344.

Gibcus, J. H., K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. T. Kanemaki, L. Xie, J. R. Paulson and W. C. Earnshaw (2018). "A pathway for mitotic chromosome formation." Science **359**(6376): eaao6135.

Gligoris, T. G., J. C. Scheinost, F. Bürmann, N. Petela, K.-L. Chan, P. Uluocak, F. Beckouët, S. Gruber, K. Nasmyth and J. Löwe (2014). "Closing the cohesin ring: Structure and function of its Smc3-kleisin interface." Science **346**(6212): 963-967.

Goloborodko, A., J. F. Marko and L. A. Mirny (2016). "Chromosome compaction by active loop extrusion." Biophysical journal **110**(10): 2162-2168.

Gordon-Smith, D., Carbajo, R., Yang, J., Videler, H., Runswick, M., Walker, J. and Neuhaus, D., 2001. C-Terminal Coiled-Coil Domain from Bovine IF1.

Green, L. C., P. Kalitsis, T. M. Chang, M. Cipetic, J. H. Kim, O. Marshall, L. Turnbull, C. B. Whitchurch, P. Vagnarelli and K. Samejima (2012). "Contrasting roles of condensin I and condensin II in mitotic chromosome formation." Journal of cell science **125**(6): 1591-1604.

Griese, J. J. and K. P. Hopfner (2011). "Structure and DNA-binding activity of the Pyrococcus furiosus SMC protein hinge domain." Proteins: Structure, Function, and Bioinformatics **79**(2): 558-568.

Griese, J. J., G. Witte and K.-P. Hopfner (2010). "Structure and DNA binding activity of the mouse condensin hinge domain highlight common and diverse features of SMC proteins." Nucleic Acids Research **38**(10): 3454-3465.

Grinthal, A., I. Adamovic, B. Weiner, M. Karplus and N. Kleckner (2010). "PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis." Proceedings of the National Academy of Sciences **107**(6): 2467-2472.

Gruber, S., P. Arumugam, Y. Katou, D. Kuglitsch, W. Helmhart, K. Shirahige and K. Nasmyth (2006). "Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge." Cell **127**(3): 523-537.

Gruber, S., C. H. Haering and K. Nasmyth (2003). "Chromosomal cohesin forms a ring." Cell **112**(6): 765-777.

Guacci, V., Koshland, D. and Strunnikov, A., 1997. A Direct Link between Sister Chromatid Cohesion and Chromosome Condensation Revealed through the Analysis of MCD1 in S. cerevisiae. *Cell*, 91(1), pp.47-57.

Guerineau, M., Z. Kriz, L. Kozakova, K. Bednarova, P. Janos and J. Palecek (2012). "Analysis of the Nse3/MAGE-binding domain of the Nse4/EID family proteins." PLoS One **7**(4).

Haering, C. H., A.-M. Farcas, P. Arumugam, J. Metson and K. Nasmyth (2008). "The cohesin ring concatenates sister DNA molecules." Nature **454**(7202): 297-301.

Haering, C. H., J. Lowe, A. Hochwagen and K. Nasmyth (2002). "Molecular architecture of SMC proteins and the yeast cohesin complex." Mol Cell **9**(4): 773-788.

Haering, C. H., D. Schoffnegger, T. Nishino, W. Helmhart, K. Nasmyth and J. Löwe (2004). "Structure and stability of cohesin's Smc1-kleisin interaction." Molecular cell **15**(6): 951-964.

Haeusler, R. A., M. Pratt-Hyatt, P. D. Good, T. A. Gipson and D. R. Engelke (2008). "Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes." Genes & development **22**(16): 2204-2214.

Hagstrom, K. A., V. F. Holmes, N. R. Cozzarelli and B. J. Meyer (2002). "C. elegans condensin promotes mitotic chromosome architecture, centromere organization, and sister chromatid segregation during mitosis and meiosis." Genes & development **16**(6): 729-742.

Ham, M. F., T. Takakuwa, N. Rahadiani, K. Tresnasari, H. Nakajima and K. Aozasa (2007). "Condensin mutations and abnormal chromosomal structures in pyothorax-associated lymphoma." Cancer science **98**(7): 1041-1047.

Hara, K., K. Kinoshita, T. Migita, K. Murakami, K. Shimizu, K. Takeuchi, T. Hirano and H. Hashimoto (2019). "Structural basis of HEAT-kleisin interactions in the human condensin I subcomplex." EMBO reports **20**(5).

Hara, K., G. Zheng, Q. Qu, H. Liu, Z. Ouyang, Z. Chen, D. R. Tomchick and H. Yu (2014). "Structure of cohesin subcomplex pinpoints direct shugoshin-Wapl antagonism in centromeric cohesion." Nature structural & molecular biology **21**(10): 864.

Hartl, T. A., H. F. Smith and G. Bosco (2008). "Chromosome alignment and transvection are antagonized by condensin II." Science **322**(5906): 1384-1387.

Hassler, M., I. A. Shaltiel, M. Kschonsak, B. Simon, F. Merkel, L. Thärichen, H. J. Bailey, J. Macošek, S. Bravo, J. Metz, J. Hennig and C. H. Haering (2019). "Structural Basis of an Asymmetric Condensin ATPase Cycle." Molecular cell **74**(6): 1175-1188.e1179.

Heale, J. T., A. R. Ball, Jr., J. A. Schmiesing, J. S. Kim, X. Kong, S. Zhou, D. F. Hudson, W. C. Earnshaw and K. Yokomori (2006). "Condensin I interacts with the PARP-1-XRCC1 complex and functions in DNA single-strand break repair." Mol Cell **21**(6): 837-848.

Henikoff, S. and J. G. Henikoff (2012). ""“Point” centromeres of Saccharomyces harbor single centromere-specific nucleosomes." Genetics **190**(4): 1575-1577.

Hernandez, M. R., M. B. Davis, J. Jiang, E. A. Brouhard, A. F. Severson and G. Csankovszki (2018). "Condensin I protects meiotic cohesin from WAPL-1 mediated removal." PLoS genetics **14**(5): e1007382.

Hirano, M. and T. Hirano (2006). "Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions." Molecular cell **21**(2): 175-186.

Hirano, T., R. Kobayashi and M. Hirano (1997). "Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a Xenopus homolog of the Drosophila Barren protein." Cell **89**(4): 511-521.

Hirota, T., D. Gerlich, B. Koch, J. Ellenberg and J.-M. Peters (2004). "Distinct functions of condensin I and II in mitotic chromosome assembly." Journal of cell science **117**(26): 6435-6445.

Hocquet, C., X. Robellet, L. Modolo, X.-M. Sun, C. Burny, S. Cuylen-Haering, E. Toselli, S. Clauder-Münster, L. Steinmetz and C. H. Haering (2018). "Condensin controls cellular RNA levels through the accurate segregation of chromosomes instead of directly regulating transcription." Elife **7**: e38517.

Holla, S., J. Dhakshnamoorthy, H. D. Folco, V. Balachandran, H. Xiao, L.-l. Sun, D. Wheeler, M. Zofall and S. I. Grewal (2020). "Positioning Heterochromatin at the Nuclear Periphery Suppresses Histone Turnover to Promote Epigenetic Inheritance." Cell **180**(1): 150-164. e115.
Hopfner, K.-P. and J. A. Tainer (2003). "Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures." Current Opinion in Structural Biology **13**(2): 249-255.

Hopfner, K. P., A. Karcher, D. S. Shin, L. Craig, L. M. Arthur, J. P. Carney and J. A. Tainer (2000). "Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily." Cell **101**(7): 789-800.

Houlard, M., J. Godwin, J. Metson, J. Lee, T. Hirano and K. Nasmyth (2015). "Condensin confers the longitudinal rigidity of chromosomes." Nature cell biology **17**(6): 771-781.

Hsu, W.-L., T. Furuta and M. Sakurai (2016). "ATP hydrolysis mechanism in a maltose transporter explored by QM/MM metadynamics simulation." The Journal of Physical Chemistry B **120**(43): 11102-11112.

Hu, B., T. Itoh, A. Mishra, Y. Katoh, K.-L. Chan, W. Upcher, C. Godlee, M. B. Roig, K. Shirahige and K. Nasmyth (2011). "ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex." Current Biology **21**(1): 12-24.

Huang, K., J. Jia, C. Wu, M. Yao, M. Li, J. Jin, C. Jiang, Y. Cai, D. Pei and G. Pan (2013). "Ribosomal RNA gene transcription mediated by the master genome regulator protein CCCTC-binding factor (CTCF) is negatively regulated by the condensin complex." Journal of Biological Chemistry **288**(36): 26067-26077.

Hudson, D. F., S. Ohta, T. Freisinger, F. Macisaac, L. Sennels, F. Alves, F. Lai, A. Kerr, J. Rappsilber and W. C. Earnshaw (2008). "Molecular and genetic analysis of condensin function in vertebrate cells." Mol Biol Cell **19**(7): 3070-3079.

Hudson, J. J., K. Bednarova, L. Kozakova, C. Liao, M. Guerineau, R. Colnaghi, S. Vidot, J. Marek, S. R. Bathula and A. R. Lehmann (2011). "Interactions between the Nse3 and Nse4 components of the SMC5-6 complex identify evolutionarily conserved interactions between MAGE and EID families." PLoS One **6**(2).

Hug, C. B., A. G. Grimaldi, K. Kruse and J. M. Vaquerizas (2017). "Chromatin architecture emerges during zygotic genome activation independent of transcription." Cell **169**(2): 216-228. e219.

Ivanov, D. and K. Nasmyth (2005). "A topological interaction between cohesin rings and a circular minichromosome." Cell **122**(6): 849-860.

Iwasaki, O., C. J. Corcoran and K.-i. Noma (2016). "Involvement of condensin-directed gene associations in the organization and regulation of chromosome territories during the cell cycle." Nucleic acids research **44**(8): 3618-3628.

Je, E. M., N. J. Yoo and S. H. Lee (2014). "Mutational and expressional analysis of SMC 2 gene in gastric and colorectal cancers with microsatellite instability." Apmis **122**(6): 499-504.

Jefford, C. E. and I. Irminger-Finger (2006). "Mechanisms of chromosome instability in cancers." Critical reviews in oncology/hematology **59**(1): 1-14.

Kagami, Y., K. Nihira, S. Wada, M. Ono, M. Honda and K. Yoshida (2014). "Mps1 phosphorylation of condensin II controls chromosome condensation at the onset of mitosis." Journal of Cell Biology **205**(6): 781-790.

Kagami, Y., M. Ono and K. Yoshida (2017). "Plk1 phosphorylation of CAP-H2 triggers chromosome condensation by condensin II at the early phase of mitosis." Scientific reports **7**(1): 1-12.

Kamada, K., M. Miyata and T. Hirano (2013). "Molecular basis of SMC ATPase activation: role of internal structural changes of the regulatory subcomplex ScpAB." Structure **21**(4): 581-594.

Kappel, C., U. Zachariae, N. Dölker and H. Grubmüller (2010). "An unusual hydrophobic core confers extreme flexibility to HEAT repeat proteins." Biophysical journal **99**(5): 1596-1603.

Kar, S. P., J. Beesley, A. A. Al Olama, K. Michailidou, J. Tyrer, Z. Kote-Jarai, K. Lawrenson, S. Lindstrom, S. J. Ramus and D. J. Thompson (2016). "Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types." Cancer discovery **6**(9): 1052-1067.

Kaczmarczyk, A., Allahverdi, A., Brouwer, T., Nordenskiöld, L., Dekker, N. and van Noort, J., 2017. Single-molecule force spectroscopy on histone H4 tail-cross-linked chromatin reveals fiber folding. Journal of Biological Chemistry, 292(42), pp.17506-17513.

Khan, T. N., K. Khan, A. Sadeghpour, H. Reynolds, Y. Perilla, M. T. McDonald, W. B. Gallentine, S. M. Baig, A. Allori and M. Angrist (2019). "Mutations in NCAPG2 cause a severe neurodevelopmental syndrome that expands the phenotypic Spectrum of Condensinopathies." The American Journal of Human Genetics **104**(1): 94-111.

Kikuchi, S., D. M. Borek, Z. Otwinowski, D. R. Tomchick and H. Yu (2016). "Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy." Proceedings of the National Academy of Sciences **113**(44): 12444-12449.

Kim, E., J. Kerssemakers, I. A. Shaltiel, C. H. Haering and C. Dekker (2020). "DNA-loop extruding condensin complexes can traverse one another." Nature: 1-5.

Kim, J. H., T. Zhang, N. C. Wong, N. Davidson, J. Maksimovic, A. Oshlack, W. C. Earnshaw, P. Kalitsis and D. F. Hudson (2013). "Condensin I associates with structural and gene regulatory regions in vertebrate chromosomes." Nature communications **4**(1): 1-14.

Kim, Y., Z. Shi, H. Zhang, I. J. Finkelstein and H. Yu (2019). "Human cohesin compacts DNA by loop extrusion." Science **366**(6471): 1345-1349.

Kimura, K., O. Cuvier and T. Hirano (2001). "Chromosome condensation by a human condensin complex in Xenopus egg extracts." J Biol Chem **276**(8): 5417-5420.

Kimura, K., M. Hirano, R. Kobayashi and T. Hirano (1998). "Phosphorylation and activation of 13S condensin by Cdc2 in vitro." Science **282**(5388): 487-490.

Kimura, K. and T. Hirano (1997). "ATP-dependent positive supercoiling of DNA by 13S condensin: a biochemical implication for chromosome condensation." Cell **90**(4): 625-634.

Kimura, K. and T. Hirano (2000). "Dual roles of the 11S regulatory subcomplex in condensin functions." Proc Natl Acad Sci U S A **97**(22): 11972-11977.

Kingston, I.J., Yung, J.S., and Singleton, M.R. (2011). Biophysical characterization of

the centromere-specific nucleosome from budding yeast. J Biol Chem 286, 4021-4026.

Kinoshita, K., T. J. Kobayashi and T. Hirano (2015). "Balancing acts of two HEAT subunits of condensin I support dynamic assembly of chromosome axes." Dev Cell **33**(1): 94-106.

Kolbowski, L., Mendes, M.L., and Rappsilber, J. (2017). Optimizing the Parameters Governing the Fragmentation of Cross-Linked Peptides in a Tribrid Mass Spectrometer. Anal. Chem. 89, 5311–5318.

Kschonsak, M., F. Merkel, S. Bisht, J. Metz, V. Rybin, M. Hassler and C. H. Haering (2017). "Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes." Cell **171**(3): 588-600 e524.

Kudo, N. R., K. Wassmann, M. Anger, M. Schuh, K. G. Wirth, H. Xu, W. Helmhart, H. Kudo, M. Mckay and B. Maro (2006). "Resolution of chiasmata in oocytes requires separase-mediated proteolysis." Cell **126**(1): 135-146.

Kulemzina, I., K. Ang, X. Zhao, J.-T. Teh, V. Verma, S. Suranthran, A. P. Chavda, R. G. Huber, B. Eisenhaber and F. Eisenhaber (2016). "A reversible association between Smc coiled coils is regulated by lysine acetylation and is required for cohesin association with the DNA." Molecular cell **63**(6): 1044-1054.

Kurat, C., Yeeles, J., Patel, H., Early, A. and Diffley, J., (2017). Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Molecular Cell, 65(1), pp.117-130.

Kurze, A., K. A. Michie, S. E. Dixon, A. Mishra, T. Itoh, S. Khalid, L. Strmecki, K. Shirahige, C. H. Haering, J. Löwe and K. Nasmyth (2011). "A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion." The EMBO Journal **30**(2): 364-378.

Lammens, A., A. Schele and K.-P. Hopfner (2004). "Structural Biochemistry of ATP-Driven Dimerization and DNA-Stimulated Activation of SMC ATPases." Current Biology **14**(19): 1778-1782.

Lavoie, B. D., E. Hogan and D. Koshland (2002). "In vivo dissection of the chromosome condensation machinery: reversibility of condensation distinguishes contributions of condensin and cohesin." The Journal of cell biology **156**(5): 805-815.

Lavoie, B. D., E. Hogan and D. Koshland (2004). "In vivo requirements for rDNA chromosome condensation reveal two cell-cycle-regulated pathways for mitotic chromosome folding." Genes & development **18**(1): 76-87.

Lawrimore, J. (2018). "Biological Consequences of Chromatin Looping in Pericentric Chromatin."
Lawrimore, J., T. M. Barry, R. M. Barry, A. C. York, B. Friedman, D. M. Cook, K. Akialis, J. Tyler, P. Vasquez and E. Yeh (2017). "Microtubule dynamics drive enhanced chromatin motion and mobilize telomeres in response to DNA damage." Molecular biology of the cell **28**(12): 1701-1711.

Lawrimore, J. and K. Bloom (2019). "The regulation of chromosome segregation via centromere loops." Critical Reviews in Biochemistry and Molecular Biology **54**(4): 352-370.

Lee, A. J., D. Endesfelder, A. J. Rowan, A. Walther, N. J. Birkbak, P. A. Futreal, J. Downward, Z. Szallasi, I. P. Tomlinson and M. Howell (2011). "Chromosomal instability confers intrinsic multidrug resistance." Cancer research **71**(5): 1858-1870.

Lee, B.-G., M. B. Roig, M. Jansma, N. Petela, J. Metson, K. Nasmyth and J. Löwe (2016). "Crystal structure of the cohesin gatekeeper Pds5 and in complex with kleisin Scc1." Cell reports **14**(9): 2108-2115.

Lee, B., Merkel, F., Allegretti, M., Hassler, M., Cawood, C., Lecomte, L., O'Reilly, F., Sinn, L., Gutierrez-Escribano, P., Kschonsak, M., Bravo, S., Nakane, T., Rappsilber, J., Aragon, L., Beck, M., Löwe, J. and Haering, C., 2020. Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. *Unpublished.*

Lee, J. (2017). The regulation and function of cohesin and condensin in mammalian oocytes and spermatocytes. Oocytes, Springer**:** 355-372.

Lee, K. and Narlikar, G., (2001). Assembly of Nucleosomal Templates by Salt Dialysis. Current Protocols in Molecular Biology,.

Lengronne, A., Y. Katou, S. Mori, S. Yokobayashi, G. P. Kelly, T. Itoh, Y. Watanabe, K. Shirahige and F. Uhlmann (2004). "Cohesin relocation from sites of chromosomal loading to places of convergent transcription." Nature **430**(6999): 573-578.

Leonard, J., N. Sen, R. Torres, T. Sutani, A. Jarmuz, K. Shirahige and L. Aragón (2015). "Condensin relocalization from centromeres to chromosome arms promotes Top2 recruitment during anaphase." Cell reports **13**(11): 2336-2344.

Li, P., H. Jin and H.-G. Yu (2014). "Condensin suppresses recombination and regulates double-strand break processing at the repetitive ribosomal DNA array to ensure proper chromosome segregation during meiosis in budding yeast." Molecular biology of the cell **25**(19): 2934-2947.

Li, W., Y. Hu, S. Oh, Q. Ma, D. Merkurjev, X. Song, X. Zhou, Z. Liu, B. Tanasa and X. He (2015). "Condensin I and II complexes license full estrogen receptor α-dependent enhancer activation." Molecular cell **59**(2): 188-202.

Li, Y., J. H. Haarhuis, Á. S. Cacciatore, R. Oldenkamp, M. S. van Ruiten, L. Willems, H. Teunissen, K. W. Muir, E. de Wit and B. D. Rowland (2020). "The structural basis for cohesin–CTCF-anchored loops." Nature: 1-5.

Li, Y., K. W. Muir, M. W. Bowler, J. Metz, C. H. Haering and D. Panne (2018). "Structural basis for Scc3-dependent cohesin recruitment to chromatin." Elife **7**: e38356.

Li, Y., A. J. Schoeffler, J. M. Berger and M. G. Oakley (2010). "The crystal structure of the hinge domain of the Escherichia coli structural maintenance of chromosomes protein MukB." Journal of molecular biology **395**(1): 11-19.

Lipp, J. J., T. Hirota, I. Poser and J.-M. Peters (2007). "Aurora B controls the association of condensin I but not condensin II with mitotic chromosomes." Journal of cell science **120**(7): 1245-1255.

Liu, X. and M. Winey (2012). "The MPS1 family of protein kinases." Annual review of biochemistry **81**: 561-585.

Losada, A., M. Hirano and T. Hirano (1998). "Identification of Xenopus SMC protein complexes required for sister chromatid cohesion." Genes Dev **12**(13): 1986-1997.

Löwe, J., S. C. Cordell and F. van den Ent (2001). "Crystal structure of the SMC head domain: an ABC ATPase with 900 residues antiparallel coiled-coil inserted." Journal of molecular biology **306**(1): 25-35.

Luger, K., T. J. Rechsteiner and T. J. Richmond (1999). Preparation of nucleosome core particle from recombinant histones. Methods in enzymology, Elsevier. **304:** 3-19.

Lupas, A. N. and M. Gruber (2005). The structure of α-helical coiled coils. Advances in protein chemistry, Elsevier. **70:** 37-38.

Lusser, A. and Kadonaga, J., 2004. Strategies for the reconstitution of chromatin. Nature Methods, 1(1), pp.19-26.

Machín, F., K. Paschos, A. Jarmuz, J. Torres-Rosell, C. Pade and L. Aragón (2004). "Condensin regulates rDNA silencing by modulating nucleolar Sir2p." Current biology **14**(2): 125-130.

Makrantoni, V. and A. L. Marston (2018). "Cohesin and chromosome segregation." Current Biology **28**(12): R688-R693.

Manalastas-Cantos, K., M. Kschonsak, C. H. Haering and D. I. Svergun (2019). "Solution structure and flexibility of the condensin HEAT-repeat subunit Ycg1." Journal of Biological Chemistry **294**(37): 13822-13829.

Marsden, M. and U. Laemmli (1979). "Metaphase chromosome structure: evidence for a radial loop model." Cell **17**(4): 849-858.

Martin, C.-A., J. E. Murray, P. Carroll, A. Leitch, K. J. Mackenzie, M. Halachev, A. E. Fetit, C. Keith, L. S. Bicknell and A. Fluteau (2016). "Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis." Genes & development **30**(19): 2158-2172.

Martin, C. A., J. E. Murray, P. Carroll, A. Leitch, K. J. Mackenzie, M. Halachev, A. E. Fetit, C. Keith, L. S. Bicknell, A. Fluteau, P. Gautier, E. A. Hall, S. Joss, G. Soares, J. Silva, M. B. Bober, A. Duker, C. A. Wise, A. J. Quigley, S. R. Phadke, A. J. Wood, P. Vagnarelli and A. P. Jackson (2016). "Mutations in genes encoding condensin complex proteins cause microcephaly through decatenation failure at mitosis." Genes Dev **30**(19): 2158-2172.

Mascarenhas, J., 2002. Cell cycle-dependent localization of two novel prokaryotic chromosome segregation and condensation proteins in Bacillus subtilis that interact with SMC protein. *The EMBO Journal*, 21(12), pp.3108-3118.

Megee, P. C., C. Mistrot, V. Guacci and D. Koshland (1999). "The centromeric sister chromatid cohesion site directs Mcd1p binding to adjacent sequences." Molecular cell **4**(3): 445-450.
Mehta, G. D., R. Kumar, S. Srivastava and S. K. Ghosh (2013). "Cohesin: functions beyond sister chromatid cohesion." FEBS Lett **587**(15): 2299-2312.

Melters, D. P., L. V. Paliulis, I. F. Korf and S. W. Chan (2012). "Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis." Chromosome Research **20**(5): 579-593.

Mendes, M.L., Fischer, L., Chen, Z.A., Barbon, M., O'Reilly, F.J., Giese, S.H., Bohlke-Schneider, M., Belsom, A., Dau, T., Combe, C.W., et al. (2019). An integrated workflow for crosslinking mass spectrometry. Mol. Syst. Biol. 15, e8994.

Meyne, J., R. L. Ratliff and R. K. MoYzIs (1989). "Conservation of the human telomere sequence (TTAGGG) n among vertebrates." Proceedings of the National Academy of Sciences **86**(18): 7049-7053.

Michaelis, C., Ciosk, R. and Nasmyth, K., 1997. Cohesins: Chromosomal Proteins that Prevent Premature Separation of Sister Chromatids. Cell, 91(1), pp.35-45.

Minnen, A., F. Bürmann, L. Wilhelm, A. Anchimiuk, M.-L. Diebold-Durand and S. Gruber (2016). "Control of Smc coiled coil architecture by the ATPase heads facilitates targeting to chromosomal ParB/parS and release onto flanking DNA." Cell reports **14**(8): 2003-2016.
Muir, K. W., Y. Li, F. Weis and D. Panne (2020). "The structure of the cohesin ATPase elucidates the mechanism of SMC–kleisin ring opening." Nature Structural & Molecular Biology: 1-7.

Murakami-Tonami, Y., S. Kishida, I. Takeuchi, Y. Katou, J. M. Maris, H. Ichikawa, Y. Kondo, Y. Sekido, K. Shirahige and H. Murakami (2014). "Inactivation of SMC2 shows a synergistic lethal response in MYCN-amplified neuroblastoma cells." Cell cycle **13**(7): 1115-1131.

Nakazawa, N., K. Sajiki, X. Xu, A. Villar-Briones, O. Arakawa and M. Yanagida (2015). "RNA pol II transcript abundance controls condensin accumulation at mitotically up-regulated and heat-shock-inducible genes in fission yeast." Genes to Cells **20**(6): 481-499.

Nasmyth, K. (2001). "Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis." Annual review of genetics **35**(1): 673-745.

Naumova, N., M. Imakaev, G. Fudenberg, Y. Zhan, B. R. Lajoie, L. A. Mirny and J. Dekker (2013). "Organization of the mitotic chromosome." Science **342**(6161): 948-953.

Neuwald, A. F. and T. Hirano (2000). "HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions." Genome research **10**(10): 1445-1452.

Newman, J. A., C. D. Cooper, A. K. Roos, H. Aitkenhead, U. C. Oppermann, H. J. Cho, R. Osman and O. Gileadi (2016). "Structures of two melanoma-associated antigens suggest allosteric regulation of effector binding." PloS one **11**(2).

Nichols, M. H. and V. G. Corces (2018). "A tethered-inchworm model of SMC DNA translocation." Nature structural & molecular biology **25**(10): 906-910.

Niki, H., A. Jaffe, R. Imamura, T. Ogura and S. Hiraga (1991). "The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli." The EMBO journal **10**(1): 183-193.

Nowak, M. A., N. L. Komarova, A. Sengupta, P. V. Jallepalli, I.-M. Shih, B. Vogelstein and C. Lengauer (2002). "The role of chromosomal instability in tumor initiation." Proceedings of the National Academy of Sciences **99**(25): 16226-16231.

Nunez, R. V., L. B. R. Avila and S. Gruber (2019). "Transient DNA occupancy of the SMC interarm space in prokaryotic condensin." Molecular cell **75**(2): 209-223. e206.

Ocampo-Hafalla, M., S. Muñoz, C. P. Samora and F. Uhlmann (2016). "Evidence for cohesin sliding along budding yeast chromosomes." Open biology **6**(6): 150178.

Ono, T., Y. Fang, D. L. Spector and T. Hirano (2004). "Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells." Mol Biol Cell **15**(7): 3296-3308.

Ono, T., A. Losada, M. Hirano, M. P. Myers, A. F. Neuwald and T. Hirano (2003). "Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells." Cell **115**(1): 109-121.

Orelle, C., O. Dalmas, P. Gros, A. Di Pietro and J.-M. Jault (2003). "The conserved glutamate residue adjacent to the Walker-B motif is the catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA." Journal of Biological Chemistry **278**(47): 47002-47008.

Palecek, J. J. and S. Gruber (2015). "Kite Proteins: a Superfamily of SMC/Kleisin Partners Conserved Across Bacteria, Archaea, and Eukaryotes." Structure **23**(12): 2183-2190.

Parisi, S., McKay, M., Molnar, M., Thompson, M., van der Spek, P., van Drunen-Schoenmaker, E., Kanaar, R., Lehmann, E., Hoeijmakers, J. and Kohli, J., (1999). “Rec8p, a Meiotic Recombination and Sister Chromatid Cohesion Phosphoprotein of the Rad21p Family Conserved from Fission Yeast to Humans”. Molecular and Cellular Biology, 19(5), pp.3515-3528.

Paul, M. R., T. E. Markowitz, A. Hochwagen and S. Ercan (2018). "Condensin depletion causes genome decompaction without altering the level of global gene expression in Saccharomyces cerevisiae." Genetics **210**(1): 331-344.

Pebernard, S., W. H. McDonald, Y. Pavlova, J. R. Yates III and M. N. Boddy (2004). "Nse1, Nse2, and a novel subunit of the Smc5-Smc6 complex, Nse3, play a crucial role in meiosis." Molecular biology of the cell **15**(11): 4866-4876.

Pebernard, S., J. J. P. Perry, J. A. Tainer and M. N. Boddy (2008). "Nse1 RING-like domain supports functions of the Smc5-Smc6 holocomplex in genome stability." Molecular biology of the cell **19**(10): 4099-4109.
Perales, R., Zhang, L. and Bentley, D., (2011). Histone Occupancy In Vivo at the 601 Nucleosome Binding Element Is Determined by Transcriptional History. Molecular and Cellular Biology, 31(16), pp.3485-3496.

Perche, O., A. Menuet, M. Marcos, L. Liu, A. Pâris, K. H. Utami, D. Kervran, V. Cacheux, B. Laudier and S. Briault (2013). "Combined deletion of two Condensin II system genes (NCAPG2 and MCPH1) in a case of severe microcephaly and mental deficiency." European journal of medical genetics **56**(11): 635-641.

Petes, T. D. (1979). "Yeast ribosomal DNA genes are located on chromosome XII." Proceedings of the National Academy of Sciences **76**(1): 410-414.

Petrushenko, Z. M., W. She and V. V. Rybenkov (2011). "A new family of bacterial condensins." Molecular Microbiology **81**(4): 881-896.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605–1612.

Piazza, I., A. Rutkowska, A. Ori, M. Walczak, J. Metz, V. Pelechano, M. Beck and C. H. Haering (2014). "Association of condensin with chromosomes depends on DNA binding by its HEAT-repeat subunits." Nat Struct Mol Biol **21**(6): 560-568.

Pluta, A., A. Mackay, A. Ainsztein, I. Goldberg and W. Earnshaw (1995). "The centromere: hub of chromosomal activities." Science **270**(5242): 1591-1594.

Potts, P. R., M. H. Porteus and H. Yu (2006). "Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks." EMBO J **25**(14): 3377-3388.

Prieß, M., H. Göddeke, G. Groenhof and L. V. Schäfer (2018). "Molecular mechanism of ATP hydrolysis in an ABC transporter." ACS central science **4**(10): 1334-1343.

Rajasekar, K. V., R. Baker, G. L. M. Fisher, J. R. Bolla, J. Mäkelä, M. Tang, K. Zawadzka, O. Koczy, F. Wagner and C. V. Robinson (2019). "Dynamic architecture of the Escherichia coli structural maintenance of chromosomes (SMC) complex, MukBEF." Nucleic acids research **47**(18): 9696-9707.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal. Chem. 75, 663–670.

Rao, S. S., S.-C. Huang, B. G. St Hilaire, J. M. Engreitz, E. M. Perez, K.-R. Kieffer-Kwon, A. L. Sanborn, S. E. Johnstone, G. D. Bascom and I. D. Bochkov (2017). "Cohesin loss eliminates all loop domains." Cell **171**(2): 305-320. e324.

Reyes, C., C. Serrurier, T. Gauthier, Y. Gachet and S. Tournier (2015). "Aurora B prevents chromosome arm separation defects by promoting telomere dispersion and disjunction." Journal of Cell Biology **208**(6): 713-727.

Ribeiro, S. A., J. C. Gatlin, Y. Dong, A. Joglekar, L. Cameron, D. F. Hudson, C. J. Farr, B. F. McEwen, E. D. Salmon and W. C. Earnshaw (2009). "Condensin regulates the stiffness of vertebrate centromeres." Molecular biology of the cell **20**(9): 2371-2380.

Robellet, X., Y. Thattikota, F. Wang, T.-L. Wee, M. Pascariu, S. Shankar, É. Bonneil, C. M. Brown and D. D'Amours (2015). "A high-sensitivity phospho-switch triggered by Cdk1 governs chromosome morphogenesis during cell division." Genes & development **29**(4): 426-439.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745.

Rosin, L. F., S. C. Nguyen and E. F. Joyce (2018). "Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in Drosophila nuclei." PLoS genetics **14**(7).

Rowland, B. D., M. B. Roig, T. Nishino, A. Kurze, P. Uluocak, A. Mishra, F. Beckouët, P. Underwood, J. Metson and R. Imre (2009). "Building sister chromatid cohesion: smc3 acetylation counteracts an antiestablishment activity." Molecular cell **33**(6): 763-774.

Russo, C.J., Scotcher, S., and Kyte, M. (2016). A precision cryostat design for manual and semi-automated cryo-plunge instruments. Rev Sci Instrum 87, 114302.

Ryu, B., D. S. Kim, A. M. DeLuca and R. M. Alani (2007). "Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression." PloS one **2**(7).
Sadeh, R. and C. D. Allis (2011). "Genome-wide "re"-modeling of nucleosome positions." Cell **147**(2): 263-266.

Samejima, K., I. Samejima, P. Vagnarelli, H. Ogawa, G. Vargiu, D. A. Kelly, F. de Lima Alves, A. Kerr, L. C. Green and D. F. Hudson (2012). "Mitotic chromosomes are compacted laterally by KIF4 and condensin and axially by topoisomerase IIα." Journal of Cell Biology **199**(5): 755-770.

Samoshkin, A., A. Arnaoutov, L. E. Jansen, I. Ouspenski, L. Dye, T. Karpova, J. McNally, M. Dasso, D. W. Cleveland and A. Strunnikov (2009). "Human condensin function is essential for centromeric chromatin assembly and proper sister kinetochore orientation." PloS one **4**(8).

Saraste, M., Sibbald, P. and Wittinghofer, A., 1990. The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences*, 15(11), pp.430-434.

Savvidou, E., N. Cobbe, S. Steffensen, S. Cotterill and M. M. Heck (2005). "Drosophila CAP-D2 is required for condensin complex stability and resolution of sister chromatids." Journal of Cell Science **118**(11): 2529-2543.

Schalbetter, S. A., A. Goloborodko, G. Fudenberg, J.-M. Belton, C. Miles, M. Yu, J. Dekker, L. Mirny and J. Baxter (2017). "SMC complexes differentially compact mitotic chromosomes according to genomic context." Nature cell biology **19**(9): 1071-1080.

Schleiffer, A., S. Kaitna, S. Maurer-Stroh, M. Glotzer, K. Nasmyth and F. Eisenhaber (2003). "Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners." Molecular cell **11**(3): 571-575.

Schwarzer, W., N. Abdennur, A. Goloborodko, A. Pekowska, G. Fudenberg, Y. Loe-Mie, N. A. Fonseca, W. Huber, C. H. Haering and L. Mirny (2017). "Two independent modes of chromatin organization revealed by cohesin removal." Nature **551**(7678): 51-56.

Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Shintomi, K. and T. Hirano (2009). "Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl–Pds5 and Sgo1." Genes & development **23**(18): 2224-2236.

Shintomi, K. and T. Hirano (2011). "The relative ratio of condensin I to II determines chromosome shapes." Genes Dev **25**(14): 1464-1469.

Shintomi, K., F. Inoue, H. Watanabe, K. Ohsumi, M. Ohsugi and T. Hirano (2017). "Mitotic chromosome assembly despite nucleosome depletion in Xenopus egg extracts." Science **356**(6344): 1284-1287.

Shintomi, K., T. S. Takahashi and T. Hirano (2015). "Reconstitution of mitotic chromatids with a minimum set of purified factors." Nat Cell Biol **17**(8): 1014-1023.

Simpson, R. and Stafford, D., (1983). Structural features of a phased nucleosome core particle. Proceedings of the National Academy of Sciences, 80(1), pp.51-55.

Soh, Y.-M., F. Bürmann, H.-C. Shin, T. Oda, K. S. Jin, C. P. Toseland, C. Kim, H. Lee, S. J. Kim and M.-S. Kong (2015). "Molecular basis for SMC rod formation and its dissolution upon DNA binding." Molecular cell **57**(2): 290-303.

Soppa, J., K. Kobayashi, M. F. Noirot-Gros, D. Oesterhelt, S. D. Ehrlich, E. Dervyn, N. Ogasawara and S. Moriya (2002). "Discovery of two novel families of proteins that are proposed to interact with prokaryotic SMC proteins, and characterization of the Bacillus subtilis family members ScpA and ScpB." Molecular microbiology **45**(1): 59-71.

Srinivasan, M., J. C. Scheinost, N. J. Petela, T. G. Gligoris, M. Wissler, S. Ogushi, J. E. Collier, M. Voulgaris, A. Kurze and K.-L. Chan (2018). "The cohesin ring uses its hinge to organize DNA using non-topological as well as topological mechanisms." Cell **173**(6): 1508-1519. e1518.

St-Pierre, J., M. Douziech, F. Bazile, M. Pascariu, E. Bonneil, V. Sauve, H. Ratsima and D. D'Amours (2009). "Polo kinase regulates mitotic chromosome condensation by hyperactivation of condensin DNA supercoiling activity." Mol Cell **34**(4): 416-426.

Stegmeier, F. and A. Amon (2004). "Closing mitosis: the functions of the Cdc14 phosphatase and its regulation." Annu. Rev. Genet. **38**: 203-232.

Stephens, A. D., C. W. Quammen, B. Chang, J. Haase, R. M. Taylor and K. Bloom (2013). "The spatial segregation of pericentric cohesin and condensin in the mitotic spindle." Molecular biology of the cell **24**(24): 3909-3919.

Story, R. M., I. T. Weber and T. A. Steitz (1992). "The structure of the E. coli recA protein monomer and polymer." Nature **355**(6358): 318-325.

Strunnikov, A. V., V. L. Larionov and D. Koshland (1993). "SMC1: an essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous protein family." The Journal of cell biology **123**(6): 1635-1648.

Stumpff, J., M. Wagenbach, A. Franck, C. L. Asbury and L. Wordeman (2012). "Kif18A and chromokinesins confine centromere movements via microtubule growth suppression and spatial control of kinetochore tension." Developmental cell **22**(5): 1017-1029.

Subramanya, H. S., L. E. Bird, J. A. Brannigan and D. B. Wigley (1996). "Crystal structure of a DExx box DNA helicase." Nature **384**(6607): 379-383.

Szabo, Q., F. Bantignies and G. Cavalli (2019). "Principles of genome folding into topologically associating domains." Science advances **5**(4): eaaw1668.

Tachibana-Konwalski, K., J. Godwin, L. van der Weyden, L. Champion, N. R. Kudo, D. J. Adams and K. Nasmyth (2010). "Rec8-containing cohesin maintains bivalents without turnover during the growing phase of mouse oocytes." Genes & development **24**(22): 2505-2516.

Tada, K., H. Susumu, T. Sakuno and Y. Watanabe (2011). "Condensin association with histone H2A shapes mitotic chromosomes." Nature **474**(7352): 477-483.

Takahashi, M., T. Wakai and T. Hirota (2016). "Condensin I-mediated mitotic chromosome assembly requires association with chromokinesin KIF4A." Genes & development **30**(17): 1931-1936.

Takemoto, A., A. Murayama, M. Katano, T. Urano, K. Furukawa, S. Yokoyama, J. Yanagisawa, F. Hanaoka and K. Kimura (2007). "Analysis of the role of Aurora B on the chromosomal targeting of condensin I." Nucleic acids research **35**(7): 2403-2412.

Terakawa, T., S. Bisht, J. M. Eeftens, C. Dekker, C. H. Haering and E. C. Greene (2017). "The condensin complex is a mechanochemical motor that translocates along DNA." Science **358**(6363): 672-676.

Thadani, R., J. Kamenz, S. Heeger, S. Muñoz and F. Uhlmann (2018). "Cell-cycle regulation of dynamic chromosome association of the condensin complex." Cell reports **23**(8): 2308-2317.

Toselli-Mollereau, E., X. Robellet, L. Fauque, S. Lemaire, C. Schiklenk, C. Klein, C. Hocquet, P. Legros, L. N'Guyen and L. Mouillard (2016). "Nucleosome eviction in mitosis assists condensin loading and chromosome condensation." The EMBO journal **35**(14): 1565-1581.

Tran, N., Laub, M. and Le, T., 2017. SMC Progressively Aligns Chromosomal Arms in Caulobacter crescentus but Is Antagonized by Convergent Transcription. *Cell Reports*, 20(9), pp.2057-2071.

Uchiyama, S. and K. Fukui (2015). "Condensin in chromatid cohesion and segregation." Cytogenetic and genome research **147**(4): 212-216.

Uhlmann, F. (2016). "SMC complexes: from DNA to chromosomes." Nature reviews Molecular cell biology **17**(7): 399.

Uhlmann, F., D. Wernic, M.-A. Poupart, E. V. Koonin and K. Nasmyth (2000). "Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast." Cell **103**(3): 375-386.

Ünal, E., J. M. Heidinger-Pauli, W. Kim, V. Guacci, I. Onn, S. P. Gygi and D. E. Koshland (2008). "A molecular determinant for the establishment of sister chromatid cohesion." Science **321**(5888): 566-569.

Van Bortle, K., M. H. Nichols, L. Li, C.-T. Ong, N. Takenaka, Z. S. Qin and V. G. Corces (2014). "Insulator function and topological domain border strength scale with architectural protein occupancy." Genome biology **15**(5): R82.

Vary, J., Fazzio, T. and Tsukiyama, T., (2003). Assembly of Yeast Chromatin Using ISWI Complexes. Chromatin and Chromatin Remodeling Enzymes, Part A, pp.88-102.

Waldman, V. M., T. H. Stanage, A. Mims, I. S. Norden and M. G. Oakley (2015). "Structural mapping of the coiled-coil domain of a bacterial condensin and comparative analyses across all domains of life suggest conserved features of SMC proteins." Proteins: Structure, Function, and Bioinformatics **83**(6): 1027-1045.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." The EMBO journal **1**(8): 945-951.

Walther, N., M. J. Hossain, A. Z. Politi, B. Koch, M. Kueblbeck, Ø. Ødegård-Fougner, M. Lampe and J. Ellenberg (2018). "A quantitative map of human Condensins provides new insights into mitotic chromosome architecture." Journal of Cell Biology **217**(7): 2309-2328.

Wang, X., H. B. Brandão, T. B. Le, M. T. Laub and D. Z. Rudner (2017). "Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus." Science **355**(6324): 524-527.

Wang, X., Hughes, A., Brandão, H., Walker, B., Lierz, C., Cochran, J., Oakley, M., Kruse, A. and Rudner, D., 2018. In Vivo Evidence for ATPase-Dependent DNA Translocation by the Bacillus subtilis SMC Condensin Complex. *Molecular Cell*, 71(5), pp.841-847.e5.

Warner, J. R. (1999). "The economics of ribosome biosynthesis in yeast." Trends in biochemical sciences **24**(11): 437-440.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303.

Wells, J. N., T. G. Gligoris, K. A. Nasmyth and J. A. Marsh (2017). "Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins." Current Biology **27**(1): R17-R18.

White, C. L., R. K. Suto and K. Luger (2001). "Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions." The EMBO journal **20**(18): 5207-5218.

Wilkens, S. (2015). "Structure and mechanism of ABC transporters." F1000prime reports **7**: 14-14.

Wilkins, B. J., N. A. Rall, Y. Ostwal, T. Kruitwagen, K. Hiragami-Hamada, M. Winkler, Y. Barral, W. Fischle and H. Neumann (2014). "A cascade of histone modifications induces chromatin condensation in mitosis." Science **343**(6166): 77-80.

Witzany, G. (2008). "The viral origins of telomeres and telomerases and their important role in eukaryogenesis and genome maintenance." Biosemiotics **1**(2): 191-206.

Woo, J.-S., J.-H. Lim, H.-C. Shin, M.-K. Suh, B. Ku, K.-H. Lee, K. Joo, H. Robinson, J. Lee and S.-Y. Park (2009). "Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions." Cell **136**(1): 85-96.

Woodward, J., G. C. Taylor, D. C. Soares, S. Boyle, D. Sie, D. Read, K. Chathoth, M. Vukovic, N. Tarrats, D. Jamieson, K. J. Campbell, K. Blyth, J. C. Acosta, B. Ylstra, M. J. Arends, K. R. Kranc, A. P. Jackson, W. A. Bickmore and A. J. Wood (2016). "Condensin II mutation causes T-cell lymphoma through tissue-specific genome instability." Genes Dev **30**(19): 2173-2186.

Wutz, G., C. Várnai, K. Nagasaka, D. A. Cisneros, R. R. Stocsits, W. Tang, S. Schoenfelder, G. Jessberger, M. Muhar and M. J. Hossain (2017). "Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins." The EMBO journal **36**(24): 3573-3599.

Yamanaka, K., Ogura, T., Niki, H. and Hiraga, S., 1996. Identification of two new genes,mukE andmukF, involved in chromosome partitioning inEscherichia coli. *Molecular and General Genetics MGG*, 250(3), pp.241-251.

Ye, J., A. R. Osborne, M. Groll and T. A. Rapoport (2004). "RecA-like motor ATPases—lessons from structures." Biochimica et Biophysica Acta (BBA) - Bioenergetics **1659**(1): 1-18.

Yeh, E., J. Haase, L. V. Paliulis, A. Joglekar, L. Bond, D. Bouck, E. Salmon and K. S. Bloom (2008). "Pericentric chromatin is organized into an intramolecular loop in mitosis." Current Biology **18**(2): 81-90.

Yoshimura, S. H. and T. Hirano (2016). "HEAT repeats - versatile arrays of amphiphilic helices working in crowded environments?" J Cell Sci **129**(21): 3963-3970.

Yu, H.-G. and D. Koshland (2005). "Chromosome morphogenesis: condensin-dependent cohesin removal during meiosis." Cell **123**(3): 397-407.

Yu, H.-G. and D. E. Koshland (2003). "Meiotic condensin is required for proper chromosome compaction, SC assembly, and resolution of recombination-dependent chromosome linkages." The Journal of cell biology **163**(5): 937-947.

Yuen, K. C., B. D. Slaughter and J. L. Gerton (2017). "Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters." Science advances **3**(6): e1700191.

Zabrady, K., M. Adamus, L. Vondrova, C. Liao, H. Skoupilova, M. Novakova, L. Jurcisinova, A. Alt, A. W. Oliver and A. R. Lehmann (2016). "Chromatin association of the SMC5/6 complex is dependent on binding of its NSE3 subunit to DNA." Nucleic acids research **44**(3): 1064-1079.

Zawadzka, K., P. Zawadzki, R. Baker, K. V. Rajasekar, F. Wagner, D. J. Sherratt and L. K. Arciszewska (2018). "MukB ATPases are regulated independently by the N- and C-terminal domains of MukF kleisin." Elife **7**.

Zhan, P., G. m. Xi, B. Zhang, Y. Wu, H. b. Liu, Y. f. Liu, W. j. Xu, Q. Zhu, F. Cai and Z. j. Zhou (2017). "NCAPG 2 promotes tumour proliferation by regulating G2/M phase and associates with poor prognosis in lung adenocarcinoma." Journal of cellular and molecular medicine **21**(4): 665-676.

Zhang, T., J. R. Paulson, M. Bakhrebah, J. H. Kim, C. Nowell, P. Kalitsis and D. F. Hudson (2016). "Condensin I and II behaviour in interphase nuclei and cells undergoing premature chromosome condensation." Chromosome research **24**(2): 243-269.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. 193, 1–12.

Zhang, Y., F. Liu, C. Zhang, M. Ren, M. Kuang, T. Xiao, X. Di, L. Feng, L. Fu and S. Cheng (2020). "Non-SMC Condensin I Complex Subunit D2 Is a Prognostic Factor in Triple-Negative Breast Cancer for the Ability to Promote Cell Cycle and Enhance Invasion." The American journal of pathology **190**(1): 37-47.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332.

Zhou, Y., P. Ojeda-May and J. Pu (2013). "H-loop histidine catalyzes ATP hydrolysis in the E. coli ABC-transporter HlyB." Physical Chemistry Chemical Physics **15**(38): 15811-15815.