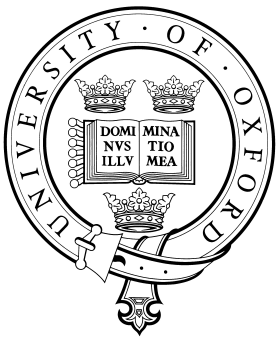
Biochemical studies of the cohesin complex

William R. Upcher

Merton College

Oxford



Dissertation submitted for the degree of Doctor of Philosophy

Department of Biochemistry

University of Oxford

Trinity, October 2012

Biochemical studies of the cohesin complex

William R. Upcher, Merton College, University of Oxford
Dissertation submitted for the degree of Doctor of Philosophy (Biochemistry)
Trinity, October 2012

ABSTRACT

The accurate inheritance of genetic material depends upon the establishment and maintenance of sister chromatid cohesion. Replicated chromosomes are topologically encircled by the large, tripartite protein complex cohesin, allowing bi-orientation in mitosis. To entrap and reversibly dissociate from DNA, the annular complex structure must be disrupted at either the hinge domain between Smc1 and Smc3, or the interfaces created by the kleisin subunit Scc1 bridging the two ABC-like ATPase domains. The aim of this work was to characterise the cohesin complex loading and releasing mechanisms by examining the biochemical requirements for these processes. Although the identity of a chromosomal cofactor could not be assigned, the loading reaction was found to necessitate engagement of ATPase domains in an ATP-dependent manner. Notions of allosteric modulation of ATP binding and NBD engagement by acetylation were discredited. Likewise, a direct and stable physical association of hinge domains with NBDs was shown to be an unlikely conformational intermediate in a reaction thought to promote hinge opening for loading of cohesin onto DNA. The Smc3-Scc1 kleisin interface might be exploited during the opposing process of release of cohesin from DNA. Therefore, a novel protein-protein cross-linking system was adapted for use in *S. cerevisiae*, with a view to (1) confirming the well-founded role of hinge dissociation in topological entrapment, and (2) validating the Smc3-Scc1 interface as the recently conjectured exit gate. Despite promising preliminary kinetics *in vitro*, the SpyTag-SpyCatcher system was considerably less efficient *in vivo*. It was thus deemed unsuitable in its current format for investigating the process of interface dissociation in live cells. Finally, the large, cohesin complex-associated HEAT repeat protein Pds5 has been either speculated or shown to participate in all of the aforementioned processes, potentially modulating the hinge and Smc-kleisin interfaces. Acting as a regulatory node in cohesin function, it was pursued as an informative target for structural studies. Although no diffractive material could be obtained, Pds5 was confirmed to bind a short, N-terminal sequence of Scc1 and was in turn bound at its N-terminus by Wapl. Together, these findings contribute to defining the structures and states of the cohesin complex.

Abbreviations and acronyms

Ω	Aromatic side-chain
ϕ	Hydrophobic side-chain
-ve	Negative
1s	Smc1 with SpyTag insertion at L595
2xTY	Two times tryptone and yeast extract
3s	Smc3 with SpyTag insertion at S606
AAA+	ATPases associated with various cellular activities
ABC	ATP binding cassette
Ac	Acetate
ADE	Adenine
ADP	Adenosine diphosphate
AFM	Atomic force microscopy
APC/C	Anaphase promoting complex/cyclosome
Ame	Associated with microtubules and essential
AMP-PNP	Adenosine 5'-(ß,γ-imino)triphosphate
ARM	Armadillo
Ask	Associated with spindles and kinetochores
ATP	Adenosine triphosphate
ATPγS	Adenosine 5'-*O*-(thiotriphosphate)
B	Bound
B.	*Bacillus (subtilis)*
bp	Base pair
BPA	*p*-benzoyl phenylalanine
BSA	Bovine serum albumin
C	Celsius
CAP	Chromosome associated protein

CEN	Core centromeric sequence
Cep	Centromere protein
Cdc	Cell division cycle
CdLS	Cornelia de Lange Syndrome
CDK	Cyclin dependent kinase
ChIP	Chromatin immunoprecipitation
Chl	Chromosome loss
COMA	Ctf19, Okp1, Mcm21, Ame1
Cse	Chromosome segregation
CTCF	CCCTC-binding factor
Ctf	Chromosome transmission fidelity
Cut	Chromosomes ultimately torn
D.	*Drosophila (melanogaster)*
Dam	Duo1 and Mps1 interacting
Dbf	Dumbell forming
ddH_20	Double-distilled water
DDK	Dbf4- or Drf1-dependent kinase
DMSO	Dimethyl sulfoxide
dNTPs	Deoxynucleoside triphosphates
DNA	Deoxyribose nucleic acid
Drf	Dbf4-related factor
DSB	Double strand break
DTT	Dithiothreitol
E.	*Escherichia (coli)*
ECL	Enhanced chemiluminescence
Eco	Establishment of cohesion
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
FACS	Fluorescence-activated cell sorting

Fig.	Figure
FKBP12	FK506 binding protein 1A, 12kDa
FP	Fluorescence polarisation
FRAP	Fluorescence recovery after photobleaching
Frb	FKBP12-rapamycin binding
FRET	Fluorescence resonance energy transfer
FT	Flow-through
g	Gram
g	Gravitational force
G_1	First gap phase
G_2	Second gap phase
Gal	Galactose
GFP	Green fluorescent protein
H.	*Homo (sapiens)*
HA	Hemagluttinin
Halobac	*Halobacterium sp.*
hd	Head domain
HDAC	Histone deacetylase
HEAT	Huntington, EF3, PP2A, PI3-kinase TOR1
HF	High fidelity
hg	Hinge domain
HIS	Histidine
Hos	Hda one similar
HP	Heterochromatin protein
HPH	Hygromycin-B-phosphotransferase
HR	Homologous recombination
hrs	Hours
HRP	Horse radish peroxidase
Hs	*Homo sapiens*

I	Input
IMAC	Immobilised metal ion affinity chromatography
Iml	Increased minichromosome loss
IP	Immunoprecipitation
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
k	Kilo
KAN	Kanamycin
kb	Kilo bases
K_d	Dissociation constant
kDa	Kilo Dalton
L	Litres
LB	Lysogeny broth
LEU	Leucine
M	Molar
m	Metres
m	Milli
M.	*Mus (musculus)*
M (phase)	Mitotic phase/mitosis
M/s	Million per second
Mad	Mitotic arrest deficient
MAT	Mating type
MBP	Maltose-binding protein
Mcd	Mitotic chromosome determinant
Mcm	Minichromosome maintenance
Mif	Mitotic fidelity of chromosome transmission
mins	Minutes
Mis	Minichromosome instability
Mon	Monomeric SpyCatcher
Mre	Meiotic recombination

MRN	Mre11, Rad50, Nbs1
mRNA	Messenger ribonucleic acid
Mtw	Mis twelve-like
n	Nano
NBD	Nucleotide binding domain
Nbs	Nijmegen Breakage Syndrome
Ndc	Non disjunction of chromosomes
NIPBL	Nipped-B-like
Nkp	Non-essential kinetochore protein
NLS	Nuclear localisation sequence
Nnf	Necessary for nuclear function
NP-40	Nonyl phenoxypolyethoxylethanol-40
Nsl	Nnf1 synthetic lethal
NTP	Nucleotide triphosphate
OD	Optical density
Okp	Outer kinetochore protein
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline + Tween 20
PCR	Polymerase chain reaction
Pds	Precocious dissociation of sisters
PEG	Polyethylene glycol
Pfu	*Pyrococcus furians*
PGK	Phosphoglycerate kinase
PMSF	Phenylmethylsulphonylfluoride
PP2A	Protein phosphatase 2A
Pre-RC	Pre-replication complex
psi	Pounds per square inch

PVDF	Polyvinylidene difluoride
qPCR	Quantitative PCR
Rad	Radiation sensitive
Raff	Raffinose
Rec	Recombination
RFP	Red fluorescent protein
RNA	Ribonucleic acid
rpm	Revolutions per minute
s	Seconds
S.	*Saccharomyces (cerevisiae)*
Sz.	*Schizosaccharomyces (pombe)*
S (phase)	Synthesis phase
SA	Stromal antigen
SAC	Spindle assembly checkpoint
Sbc	Suppressor of RecBC
Sc	*Saccharomyces cerevisiae*
Scc	Sister chromatid cohesion
SCE	Sorbitol, citrate, EDTA
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
Sgo	Shugoshin
SMC	Structural maintenance of chromosomes
Smc	Stability of minichromosomes
SPB	Spindle-pole body
Spc	Spindle pole component
Spr	SMC partner of Rad
Spt	Suppressor of Ty's
Spy	*Streptococcus pyogenes*
Srb	Suppressor of RNA polymerase B

Ssl	Swi6 synthetically lethal
SUMO	Small ubiquitin-like modifier
Swi	Switching deficient
T.	*Thermotoga (maritima)*
Tan	Tandem SpyCatcher
TCA	Trichloroacetic acid
TetO	Tetracycline operator
TetR	Tetracycline repressor protein
TEV	Tobacco etch mosaic virus
T_m	Melting temperature
TMD	Transmembrane domain
TPR	Tetratricopeptide repeat
tRNA	Transfer ribonucleic acid
TRP	Tryptophan
ts	Thermosensitive
µ	Micro
URA	Uracil
UV	Ultraviolet
v/v	Volume/volume
Wapl	Wings-apart like
WCE	Whole cell extract
WT	Wild type
w/v	Mass/volume
X.	*Xenopus (laevis)*
Ybp	Yap1-binding protein
YPD	Yeast extract peptone + dextrose
YPR	Yeast extract peptone + raffinose
YPRG	Yeast extract peptone + raffinose + galactose
Z.	Zygosaccharomyces (rouxii)

Contributions

Parts of this thesis were conducted in collaboration with other researchers. Relevant contributions are as follows:

Several of the DNA constructs used were previously created by others, or were modifications thereof, as listed in *Materials and methods* (Chapter VI).

Crystallisation trials with each protein were performed by Dr Jan Löwe[1].

All SpyCatcher proteins used in *in vitro* experiments were prepared from *E. coli* by Jacob Orry Fierer[2].

Attempts to produce recombinant Smc3hd containing a genetically encoded N-acetyl-lysine were in collaboration with Dr Susan Hancock[1].

Appendix: Figure A-1. Depicts experiments performed by Dr Bin Hu[2].

[1] Medical Research Council Laboratory of Molecular Biology, Cambridge.
[2] Department of Biochemistry, University of Oxford.

Table of contents

CHAPTER I

Introduction

Introduction

THE CELL CYCLE

The unit of life

Over the last several centuries of biological research, our knowledge has grown exponentially. Once popular explanations of life, such as vitalism, have since been replaced by an understanding down to the atomic level. Yet to this day, we lack a concrete definition of life, perhaps in part because it is not so much a tangible substance as it is a descriptive collection of processes. Above all, these processes are self-sustaining in nature - ensuring survival in the face of a changing environment, and continuity via reproduction. When attempting to define life in material terms, one invariably arrives at the structural level of the cell as the most basic unit able to exhibit such properties; it is the building block for the molecular organisation required to sustain these processes.

With the advent of crude microscopes in the 17th century came the first description of cells by Robert Hooke, who famously likened the structures he saw to the cubicles, or cells, where monks slept (Hooke, 1665). Over the next century, cell theory was developed from further observations that led to the central tenets of cell biology, expressed in and summarised by the following quotations from the pioneering researchers of the era:

“The cell is the fundamental element of organisation.” – Henri Dutrochet, 1824

"All living things are composed of cells." – Theodor Schwann, 1839

"Every cell originates from another existing cell like it." – Rudolf Virchow, 1858

This last principle of cell theory, that one cell must arise from another, raises an entire field of questions about the mechanisms required for such a feat. How does a cell divide through binary fission and what are the mechanisms of control?

A cycle of phases

All cells follow a general pattern of activities related to their progression through the cell cycle, culminating in the actual division of the cytoplasm (cytokinesis) to produce two daughter cells, whereupon the process may begin again (Fig. I-1). The progression through the cell cycle is tightly governed at each stage of the sequence through a series of molecular switches that are activated only upon completion of preceding events. The activity of cyclin-dependent kinase (Cdk) is altered at each step via its association with different cyclin proteins, and it is these Cdk-cyclin complexes that act as master initiators of the pathways leading to the major operations required for cell division (Morgan, 1997). Once begun, positive feedback amplifies these pathways and ensures a commitment to follow through, where no intermediate state is possible. With the successful completion of each phase a new cascade of events is set in motion. The combination of these systems ensures that transit through the cell cycle is dedicated, ordered, and unidirectional.

A cell begins its life in G_1 phase (first gap phase) and begins to grow within its environment. When this growth reaches a sufficient level and environmental conditions are suitable, the decision to divide is made and the process of DNA

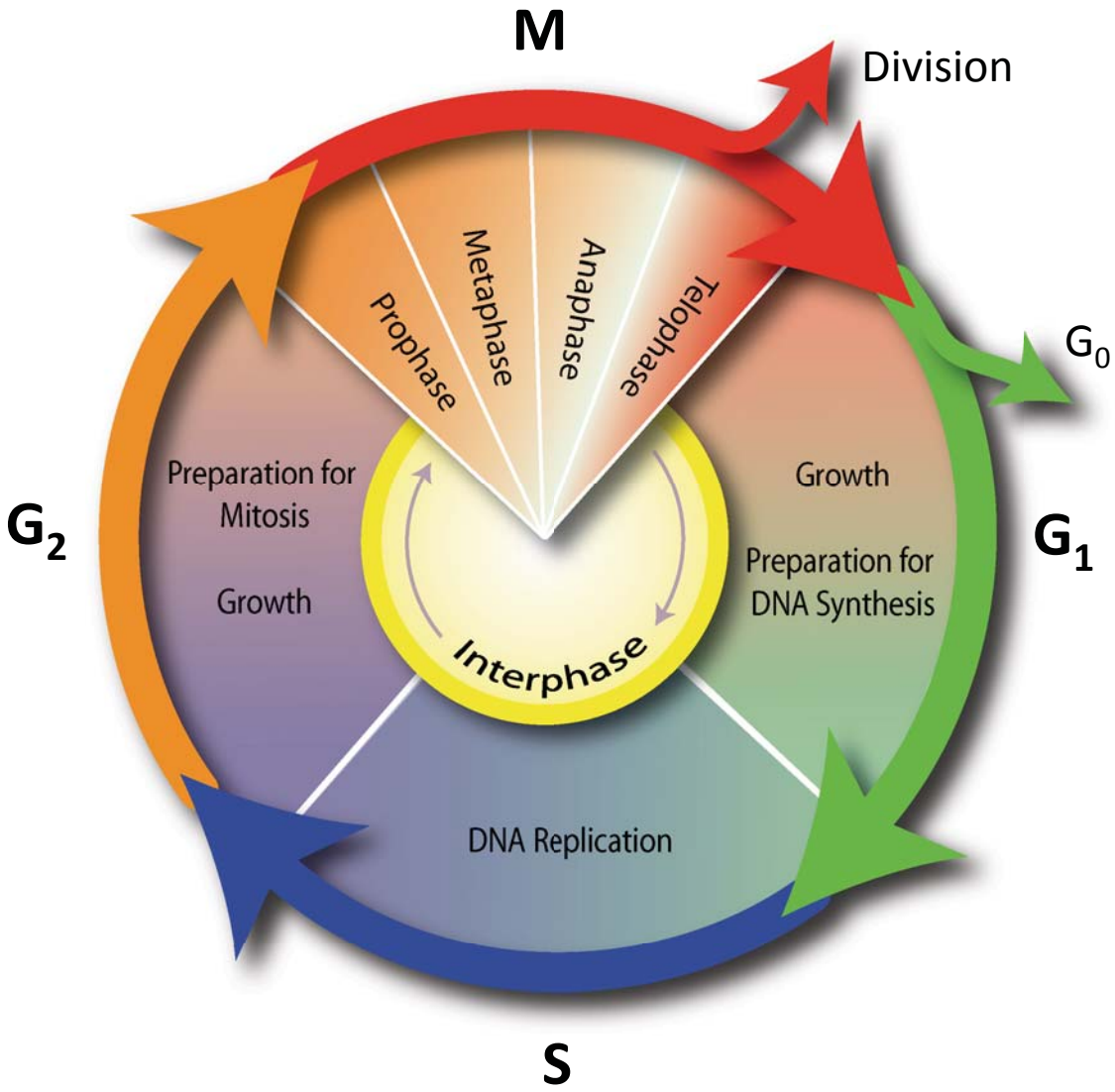


Figure I-1. Overview of a typical eukaryotic cell cycle.
Mitosis (M) is subdivided into the four characteristic stages leading to nuclear and cellular division. A newly formed daughter cell begins the cell cycle in G_1 and depending on conditions and cell type, will either enter a quiescent state (G_0) or commit to undertake division by entering S phase to duplicate the genome. Second gap phase (G_2) precedes mitosis as another period of preparatory growth. G_1 through to G_2 are collectively known as interphase.

replication begins, thus defining the start of S phase (synthesis phase). A checkpoint monitoring the integrity and fidelity of duplicated DNA must be satisfied at the beginning of the second gap phase (G_2) after completion of S phase. During G_2 the cell grows further, synthesising the materials necessary for entering and completing mitosis (M phase).

Early mitosis

Mitosis is the phase at which the commitment to replicate and divide reaches fruition. It is a remarkably complicated period of the cell cycle, consisting itself of multiple distinct stages as outlined in Figure I-2 (adapted from Jackson et al., 2007). The importance and complexity of M phase is highlighted by the fact that all the preceding stages (G_1, S, G_2) are collectively known as interphase, whilst mitosis carries its own designation. Having copied the genetic material in S phase, M phase is the time within which the cell must facilitate the equal distribution of this material between the daughter cells to be produced upon division. Up until M phase, the molecules of DNA that comprise the genome exist without clear structure or separation within the nucleus. Thus, a mechanism for sorting this material first requires its large-scale reorganisation into discrete packets that take the form of condensed chromosomes first observable during prophase. At this point, DNA is visible as sister chromatids: condensed pairs of duplicate DNA held together since their formation in S phase. The physical kinship between duplicate chromatids is known as sister chromatid cohesion. It is conferred partly through the intertwining, or catenation, of DNA strands produced during their synthesis; however, it is the cohesin complex of proteins, discussed later in greater detail,

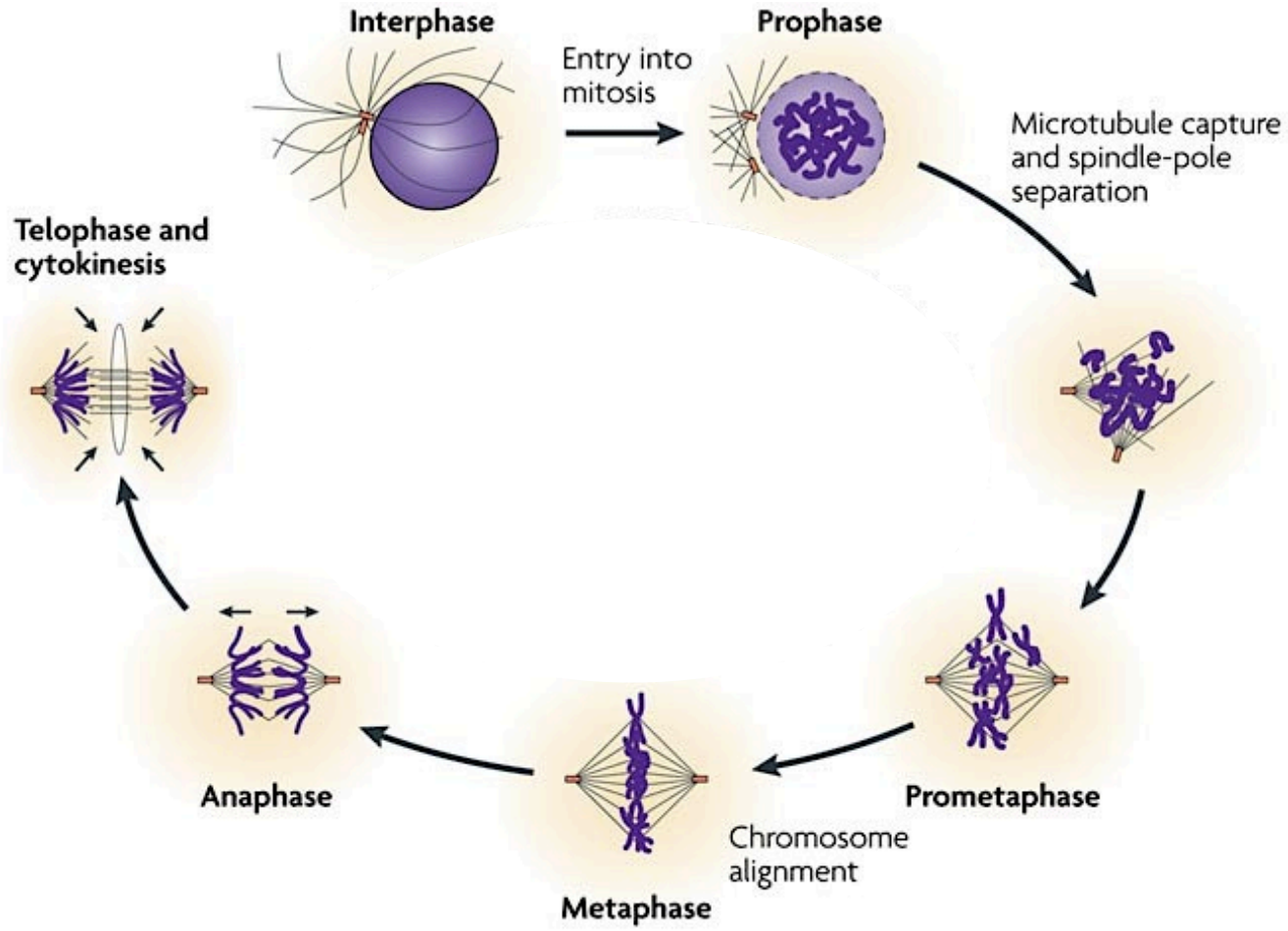


Figure I-2. The events of mitosis.
The stages of mitosis are distinguished by the organisation and manipulation of the genome, beginning with the condensation of DNA into chromosomes in prophase. The breakdown of the nuclear envelope and capture of chromatids by microtubules signifies prometaphase. Sister chromatids congress and align on the metaphase plate at the mid-zone between the bipolar mitotic spindle in metaphase before their separation and segregation to opposite spindle poles in anaphase. Mitosis is complete upon the disassembly of the spindle and the redeposition of the nuclear membrane to form daughter nuclei in telophase. Cytokinesis physically divides the cell in half (adapted by permission from Macmillan Publishers Ltd: *Nature Reviews Cancer*, Jackson et al., copyright 2007).

which provides the most significant contribution to sister chromatid cohesion (Sundin and Varshavsky, 1981; Michaelis et al., 1997; Diaz-Martinez et al., 2008; Farcas et al., 2011). Cohesion is vital to the success of mitosis for a number of reasons. The very fact that chromatids are tethered together allows them to be easily manipulated within the cell as pairs. Following the breakdown of the nuclear envelope in prometaphase (a feature of mitosis not shared by yeast), microtubules emanate from one of two centrosomes or spindle-pole bodies (SPBs) situated on opposite sides of the cell, and are able to attach to chromatids via their kinetochores. Once coupled to microtubules, chromatids can be manoeuvred centrally within the mitotic spindle and are then aligned during metaphase. Once all the chromosomes have congressed at the metaphase plate, there comes a particularly critical point at which the faithful and equal inheritance of genetic information between cells is ensured through bi-orientation, the attachment of sister chromatid kinetochores to microtubules of opposing spindle poles.

Bi-orientation

Incorrect segregation of sister chromatids can produce cells with abnormal numbers of chromosomes, which for single-celled organisms can be lethal. In metazoa, aneuploidy has similarly deleterious effects ranging from oncogenesis in somatic cells, to developmental disorders of offspring if arising in the germ cells of an aberrant meiosis.

It is the phenomenon of bi-orientation that most directly guarantees against these errors of segregation, and central to effective bi-orientation is the

cohesion between sister chromatids. It is only in the case of amphitelic attachment of sister centromeres that the pulling forces from microtubules can generate tension. When both sister chromatids are pulled in the same direction by microtubules from the same pole (syntelic attachment), or when one kinetochore remains unbound (monotelic attachment), there is no force to resist the pull and these weak attachments are abolished (Nicklas and Koch, 1969; Nicklas et al., 1995). The safeguard against improper separation of the genome is the spindle-assembly checkpoint (SAC), which is only satisfied when tension exists between every sister kinetochore pair of the genome (Musacchio and Salmon, 2007). Of course, microtubular pulling forces can only impart tension when sister chromatids remain attached to one another; the function of cohesin is thus central to this entire enterprise.

The mechanisms by which error-correction allows only bipolar attachments remain speculative. Mechanical tension appears to strengthen attachments intrinsically (Akiyoshi et al., 2010). An additional level of active regulation operates on the principle of intra-kinetochore stretching. In this model, Aurora B kinase, occupying the inner centromere, becomes spatially separate from the site of microtubule attachment at the outer kinetochore upon stretching of chromatin under tension (Liu et al., 2009). This prevents Aurora B kinase from phosphorylating and destabilising the proteins involved in securing the microtubule-kinetochore interface, explaining how tension might reinforce the strength of amphitelic attachments (Biggins et al., 1999; Cheeseman et al., 2002). In addition to this error-correction system, unattached kinetochores also catalyse a conformational change within a diffusible protein, Mad2, allowing it

to sequester Cdc20, a factor necessary for the progression to anaphase (Fang et al., 1998a, b). Once all kinetochores are securely attached and under tension, the wait anaphase signal rapidly disappears and mitosis approaches its conclusion (Li and Nicklas, 1995; Xia et al., 2004; Nasmyth, 2005).

Sister chromatid segregation

With all chromatids bi-oriented, the SAC is complied with and anaphase takes place. This is the point at which the coupling between sisters is lost, allowing their segregation towards opposite poles of the spindle. This loss of cohesion between sister chromatids is instigated by the action of the anaphase-promoting complex or cyclosome (APC/C) when bound to Cdc20. Upon activation, APC^{Cdc20} ubiquitinates the securin protein, targeting it for destruction by the proteasome (Cohen-Fix et al., 1996; Shirayama et al., 1999). The function of securin is to bind and inactivate separase, a thiol protease enzyme (Yamamoto et al., 1996; Ciosk et al., 1998; Uhlmann et al., 2000). By destroying securin, the APC/C liberates separase, which in turn can then proteolytically cleave a subunit of the cohesin complex, dissolving cohesion between sister chromatids (Fig. I-3; Irniger et al., 1995; Uhlmann et al., 1999; Hauf et al., 2001). The APC also destroys the mitotic cyclin, cyclin B, thus inactivating Cdk1 in order to prevent both the persistence of arrest signals and the premature progression into another cell cycle following mitosis (Wäsch and Cross, 2002).

With an identical set of chromatids partitioned between the two halves of the cell, the mitotic spindle is dismantled, signifying the end of mitosis, or telophase. Nuclear envelopes reappear around de-condensing chromosomes to

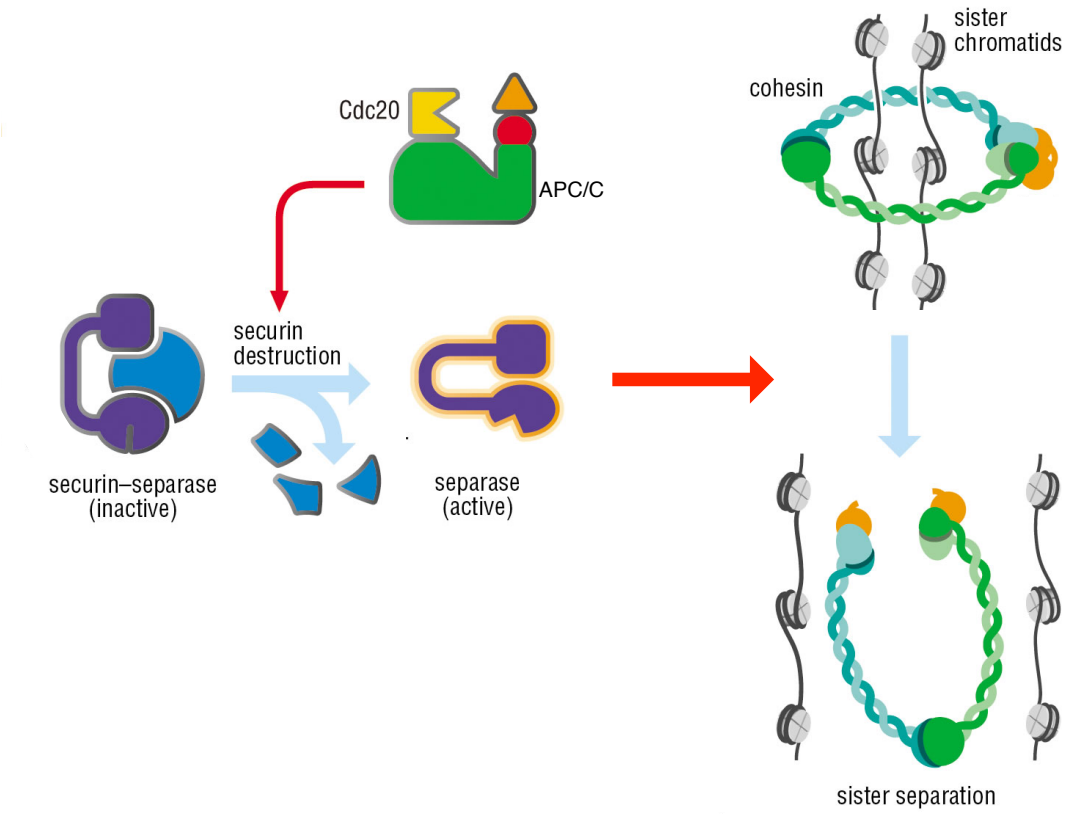


Figure I-3. Separation of sister chromatids at anaphase.
The protease separase is bound to securin and maintained in an inactive form. Upon activation of the APC by Cdc20, securin is ubiquitinated and destroyed, liberating separase and allowing its proteolytic cleavage of cohesin. Loss of cohesion between sister chromatids allows their segregation by the mitotic spindle (adapted by permission from Oxford University Press: *The Cell Cycle: Principles of Control*, Morgan, copyright 2007).

create daughter nuclei, and finally the cell is physically divided in half at the spindle mid-zone through cytokinesis.

CHROMOSOMAL ORGANISATION

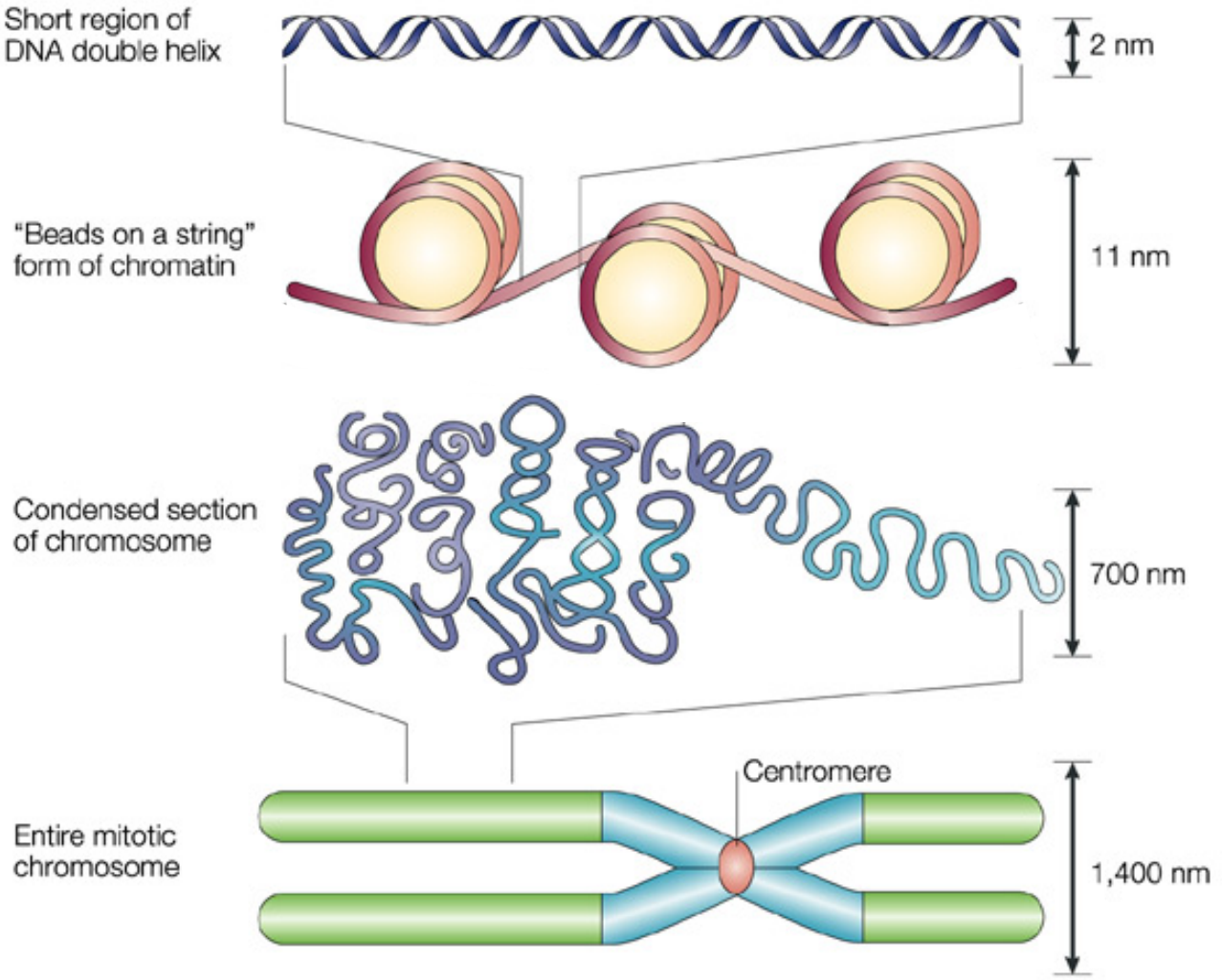
The orders of chromosome architecture

The task of structuring and organising the genetic material of a cell is truly formidable. Vast lengths of DNA must be logically arranged and compacted within the confines of the nucleus. It is therefore unsurprising that there exist a variety of proteins dedicated to this endeavour. Isolated DNA occurs as a linear strand, but in its natural context within the cell, it is tightly wound around a core of histone proteins (Fig. I-4). By providing a structural support for the coiling of DNA, histones confer a significant degree of compaction. Further levels of coiling occur at progressively larger scales to minimise the total length of DNA, which when packaged as such is referred to as chromatin (Felsenfeld and Groudine, 2003).

The packaging of chromatin into dense chromatin fibres is sufficient merely during interphase, at which point the genome appears comparatively diffuse within the nucleus. For the coordinated distribution of two copies of each DNA molecule between daughter cells in mitosis, the seemingly amorphous mass of threads must reconfigure into distinct chromatids. As chromatids, each molecule of DNA becomes microscopically distinguishable, representing a unit of inheritance manageable by the machinery of the mitotic spindle. In fact, in higher eukaryotic cells, the genome is condensed in length almost 10,000-fold at mitosis compared to linear DNA (Cooper, 2000).

To achieve this transformation, a process of extensive chromosomal condensation takes place with the help of the condensin complex (Hirano,

A



B

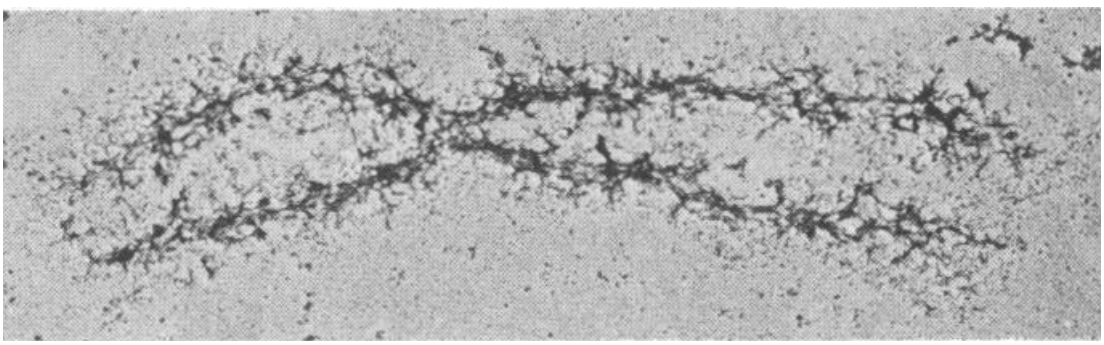


Figure I-4. Compaction of DNA into nucleosomes and metaphase chromosomes.

(A) A linear strand of DNA undergoes the most basic level of packaging via its winding around histone proteins to form a nucleosome. Further levels of coiling allow compaction into metaphase chromosomes (adapted by permission from Macmillan Publishers Ltd: *Nature*, Felsenfeld and Groudine, copyright 2003).

(B) Electron micrograph of the proteinaceous chromosomal scaffold left after digestion of DNA in mitotic chromosomes (Adolphs et al., 1977)

2005). The contribution of condensin to the assembly of mitotic chromosomes differs amongst species. However, cells lacking functional condensin subunits commonly experience defective segregation of sister chromatids at anaphase and a failure to shorten these chromatids during the preparatory phases (Saka et al., 1994; Bhat et al., 1996; Renshaw et al., 2010).

In fact, condensin is closely related to the previously encountered multisubunit protein complex involved in large-scale organisation of chromosomes: cohesin. As discussed, cohesin bears the responsibility of maintaining the propinquity between duplicate molecules of DNA from the moment of their creation in S phase, until separation in anaphase. Both condensin and cohesin belong to a wider family of protein complexes that are instrumental in engineering chromosomal architecture.

The structural maintenance of chromosomes proteins

The importance of the SMC (structural maintenance of chromosomes) family of proteins is underscored by their prevalence throughout all phylogenetic domains of life (Nasmyth and Haering, 2005). Although the discussion has focussed entirely on eukaryotes until now, bacteria and archaea face similar challenges in the segregation of chromosomes (Sherratt, 2003). Despite lacking chromatin and nuclei, prokaryotic cells are able to partition chromosomes in a structured and organised manner. In *Bacillus subtilis*, this has been shown to require the function of the SMC protein, which localises to origins of replication to condense and arrange newly replicated DNA (Britton et al., 1998; Gruber and Errington, 2009; Sullivan et al., 2009). Gram-negative

proteobacteria such as *E. coli* lack the SMC protein. Nonetheless, these species employ the distantly related protein, MukB, as the structural backbone of a complex involved in the same processes; mutants of MukB demonstrate improper positioning of the two arms of the bi-directionally replicating chromosome with the resulting formation of anucleate cells (Niki et al., 1991; Danilova et al., 2007). On account of its activity, bacterial SMC is often known as the prokaryotic condensin, from which the eukaryotic complexes evolved and diversified to produce the protein complexes with a range of specialised activities in chromosome dynamics (Cobbe and Heck, 2004).

In *S. cerevisiae*, there are six SMC proteins dedicated to shaping and segregating chromosomes. Smc1 and Smc3 form the backbone of the cohesin complex, whilst Smc2 and Smc4 constitute the core of condensin. The Smc5 and Smc6 proteins form a relatively unstudied complex that appears to be involved in the repair of DNA double-strand breaks by homologous recombination (Harvey et al., 2004; De Piccoli et al., 2006; Chavez et al., 2011). Another SMC-like family member, Rad50, is also devoted to DNA damage repair pathways in eukaryotes and archaea (Bressan et al., 1999). The SMC-like proteins present amongst a variety of species are displayed in Table I-1.

SMC structure

Evidently, these SMC-like proteins partake in specific and differing functions to maintain collectively the integrity and inheritance of the genome across all living organisms. While differing in sequence identity (MukB in particular is highly diverged), the SMC-like family of proteins share a number of

Species	SMC-like proteins by complex			
	Cohesin	**Condensin**	**Smc 5/6**	**Other**
B. subtilis	BsSMC		SbcC, RecN	
E. coli	MukB		SbcC, RecN	
S. cerevisiae	Smc1 Smc3	Smc2 Smc4	Smc5 Smc6	Rad50
Sz. pombe	Psm1 Psm3	Cut14 Cut3	Spr18 Rad18	rad50
D. melanogaster	DmSMC1 DmSMC3	DmSMC2 DmSMC4	DmSMC5 DmSMC6	rad50
H. sapiens	hSMC1∝ hSMC1ß[§] hSMC3	hCAP-E hCAP-C	hSMC5 hSMC6	hRAD50

TableI-1. Smc-like proteins present in phylogenetically-representative species ranging from bacteria to humans.

§ Vertebrates possess an additional Smc1 isoform (Smc1ß) that acts specifically in meiosis (Revenkova et al., 2001)

defining structural similarities (Cobbe and Heck, 2004; Nasmyth and Haering, 2005). When the *SMC1* gene was first cloned from budding yeast in order to elucidate a mutant phenotype (the nondisjunction and instability of minichromosomes), it was shown to encode a large protein with several predicted domains structurally homologous to gene products in other organisms (Larionov et al., 1985; Strunnikov et al., 1993). These domains consisted of globular regions at the termini and two central stretches thought to form coiled-coil. Crucial for elaborating on these early findings was the fact that all SMC-like proteins were known to act in complexes, and in particular that within these complexes the SMC protein invariably formed a dimer with another SMC. This was first brought to light with the finding that the homologues of condensin SMCs in *Xenopus* (XCAP-C and XCAP-E) associate as a heterodimer (Hirano and Mitchison, 1994). The homologues of the cohesin complex Smc1 and Smc3 proteins heterodimerise similarly in *Xenopus* and *S. cerevisiae* (Losada et al., 1998; Toth et al., 1999; Haering et al., 2002). The bacterial BsSMC and MukB genes are the only SMC protein coding sequences contained within their respective host genomes and were thus anticipated, and later shown, to associate as homodimers instead (Melby et al., 1998). Likewise, the Rad50 DNA repair protein also forms a homodimer (Hopfner et al., 2001). An expansive body of subsequent work has refined our understanding of the unifying structural features of these SMC superfamily members.

Smc1 and Smc3 engage in interactions and adopt conformations that are typical, with a number of exceptions, of the SMC-like proteins at large. Both Smc1 and Smc3 assume rod-shaped tertiary structures with globular ends

(Haering et al., 2002). One end is comprised of a fold at the centre of the primary sequence, forming a 'hinge' domain for heterodimerisation of the two Smc proteins (Haering et al., 2002; Hirano and Hirano, 2002). The other, opposite end is an ABC-like ATPase head formed by the C- and N-termini coming together. The two ends of each protein are separated by 50nm of intramolecular antiparallel coiled-coil.

Each of these domains were identified individually – initially from primary sequence – and for some time it remained uncertain how they were arranged in the context of the Smc1/3 heterodimer. Coiled-coils consist of two amphipathic ∝-helices entwined to bury the hydrophobic residues. In SMC dimers, these helical interactions might equally have been intermolecular or intramolecular, and to complicate further the puzzle of possible configurations, these helices can run either parallel or anti-parallel to one another. The various permutations are summarised in Figure I-5. An experiment that fused fibronectin domains at each terminus of the Smc3 protein and examined the resultant, fully folded structure by EM showed that the two globular regions of each Smc protein are in fact kept together. This provided good evidence that a central fold at the hinge domain gives rise to anti-parallel intramolecular coiled-coil (Haering et al., 2002;). The parallel intermolecular coiled-coil mode of association was also argued against on the basis of an emerging picture of both the ATPase and hinge domain structures in Smcs and similar proteins.

At the N-terminus, Smc proteins have a Walker A motif – a conserved sequence of amino acids concerned with the binding of nucleotide triphosphates (NTPs), in this case, ATP. A related pair of motifs found in the C-

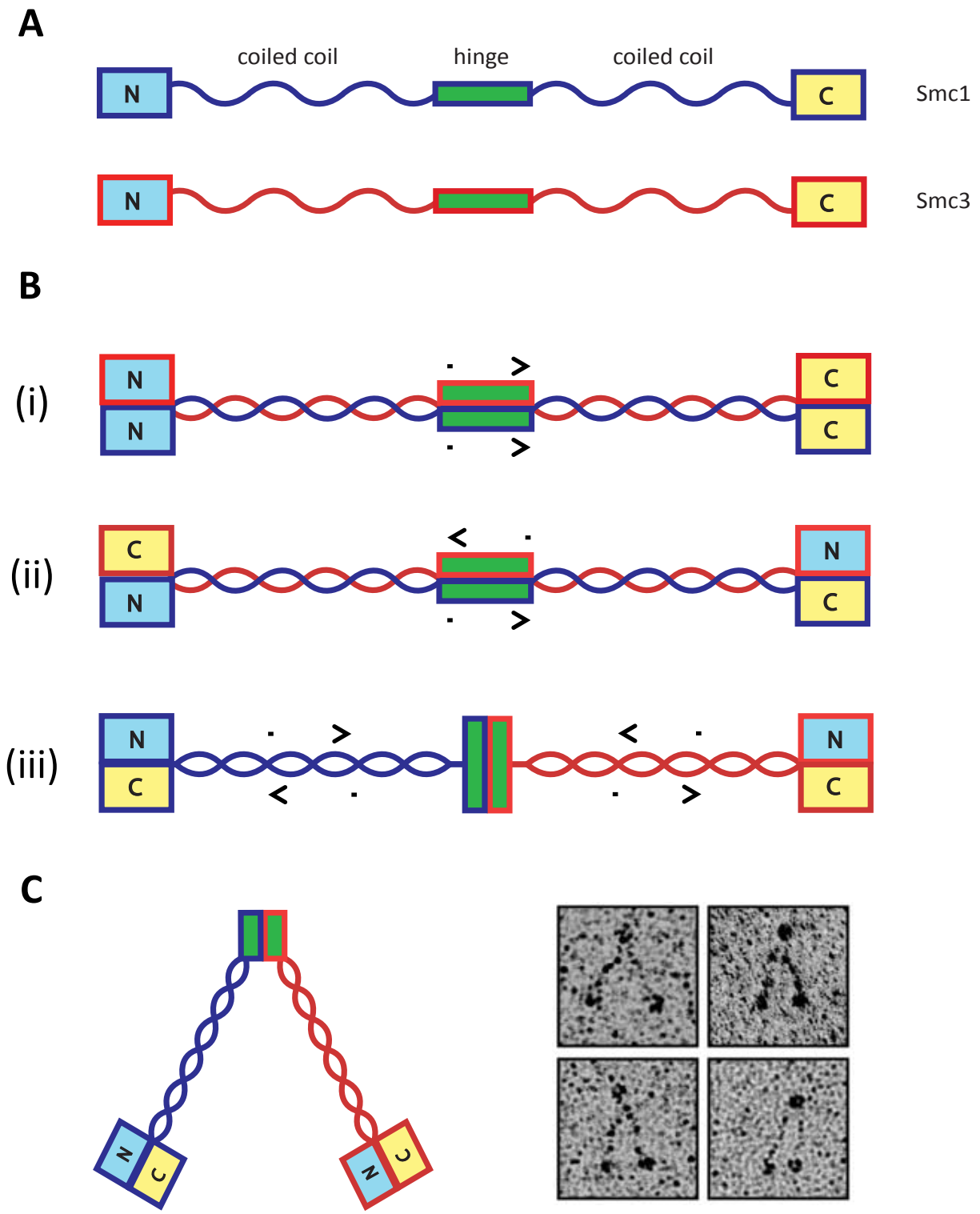


Figure I-5. SMC structure and possible conformations of heterodimers.

(A) The primary sequence of Smc1 and Smc3 predict 5 structural regions: a central hinge domain separated from the globular termini domain by two stretches of sequence predicted to form coiled coil.

(B) Possible arrangements within the heterodimer: (i) Parallel, intermolecular coiled coil alignment, (ii) anti-parallel, intermolecular coiled coil, and (iii) anti-parallel intramolecular coiled coil. Arrows indicate the direction of each sequence (Nasmyth and Haering, 2005)

(C) Smc1/3 heterodimers adopt a V-shaped conformation as visualised by electron microscopy (adapted with permission from Elsevier: *Molecular Cell*, Haering et al., copyright 2002).

terminal region of Smc proteins, the Walker B and signature motifs, together with the Walker A motif make up the necessary parts of a nucleotide-binding domain (NBD), also known as the Smc 'head' domain (Walker et al., 1982; van den Ent et al., 1999). This sort of arrangement was established crystallographically for a number of the SMC proteins and the anti-parallel orientation of the coiled-coil was also confirmed in the process (Hopfner et al., 2000; Löwe et al., 2001). In a class of transmembrane proteins known as ATP-Binding Cassette (ABC) transporters, two of these NBDs are required to dimerise in order to hydrolyse ATP in a concerted mechanism, as the signature motif of one head domain interacts with the nucleotide bound in the pocket created by the Walker A and B motif of the partnered head domain (Fig. I-6; Jones and George, 1999; Hollenstein et al., 2007). The NBDs of SMC-like proteins appear to associate in the same way (Hopfner et al., 2000; Löwe et al., 2001; Haering, 2004; Woo et al., 2009).

The second structural feature of SMC-like proteins is the central domain that acts as the site, or hinge, at which the protein is folded back upon itself. Importantly, it was also shown to be a bone fide dimerisation interface (Haering et al., 2002). Two Smcs folded in this manner associate via this hinge domain and are commonly visualised in a V-shaped conformation by scanning electron micrographs of different isolated SMC heterodimers (Fig. I-5; Melby et al., 1998; Anderson et al., 2002; Haering et al., 2002). While in the case of prokaryotic SMC complexes these interactions are homotypic, the general structure of isolated dimeric hinge domains appears to be nonetheless conserved between *T. maritima* and *M. musculus*. Interestingly, together the dimers produce a torus

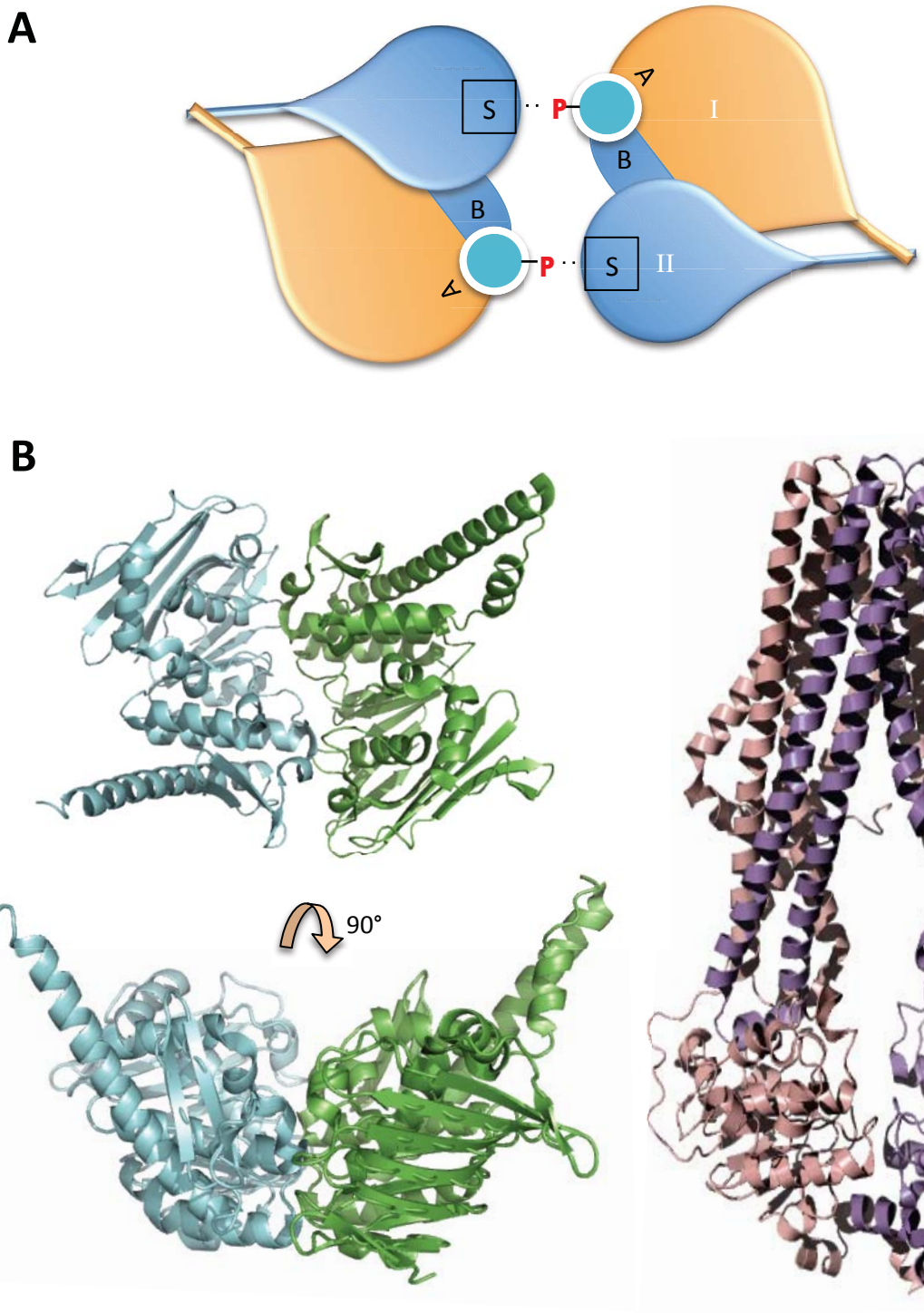


Figure I-6. Structures of ABC-like ATPases.

(A) Schematic of SMC NBDs. Lobe I is comprised of a Walker A motif (A) from N-terminal sequence (orange) and a Walker B motif from C-terminal sequence (blue), required to bind and hydrolyse ATP (turquoise), respectively. Lobe II contains the signature motif (S) that binds the γ-phosphate of ATP in the Walker A/B pocket of an opposing head to form a complete NBD dimer.

(B) Crystal structures of homodimeric Rad50 NBDs (left; Williams et al., 2011) from above and the side, with emerging coiled coils similar to those of a heterodimeric ABC transporter TM287/288 (right; Hohl et al., 2012)

with two distinct pairs of interacting surfaces through the middle of which resides a small, positively charged channel (Fig. I-7; Haering et al., 2002; Mishra et al., 2010; Kurze et al., 2011). This charge distribution is less pronounced in the condensin hinge domains (Smc2 and Smc4 from mouse) and the interacting residues are specific enough to antagonise any potential interactions between the Smcs of different complexes, e.g. the cohesin subunit Smc1 cannot dimerise with the condensin subunit Smc2 (Griese et al., 2010). Amongst the SMC-like proteins, the DNA repair-specific complexes show unique differences in their mode of association. Both Rad50 and bacterial SbcC homodimers lack the toroidal, dual interface structure at the hinge and connect instead by virtue of a zinc-hook (a motif of sequence Cys-X-X-Cys that chelates a Zn^{2+} ion), whilst the RecN protein seems to oligomerise in a rather different process mediated by the coiled-coil (Hopfner et al., 2002; Pellegrino et al., 2012).

The SMC proteins provide the structural backbones to the diversity of multisubunit complexes that curate the genetic material of all cells, and as such they share the aforementioned structural attributes critical to fulfilling their roles. However, these complexes are also highly specialised by function and accordingly, they are regulated in complicated and unique ways. These particularities are bestowed in part by the array of additional proteins that constitute each complex.

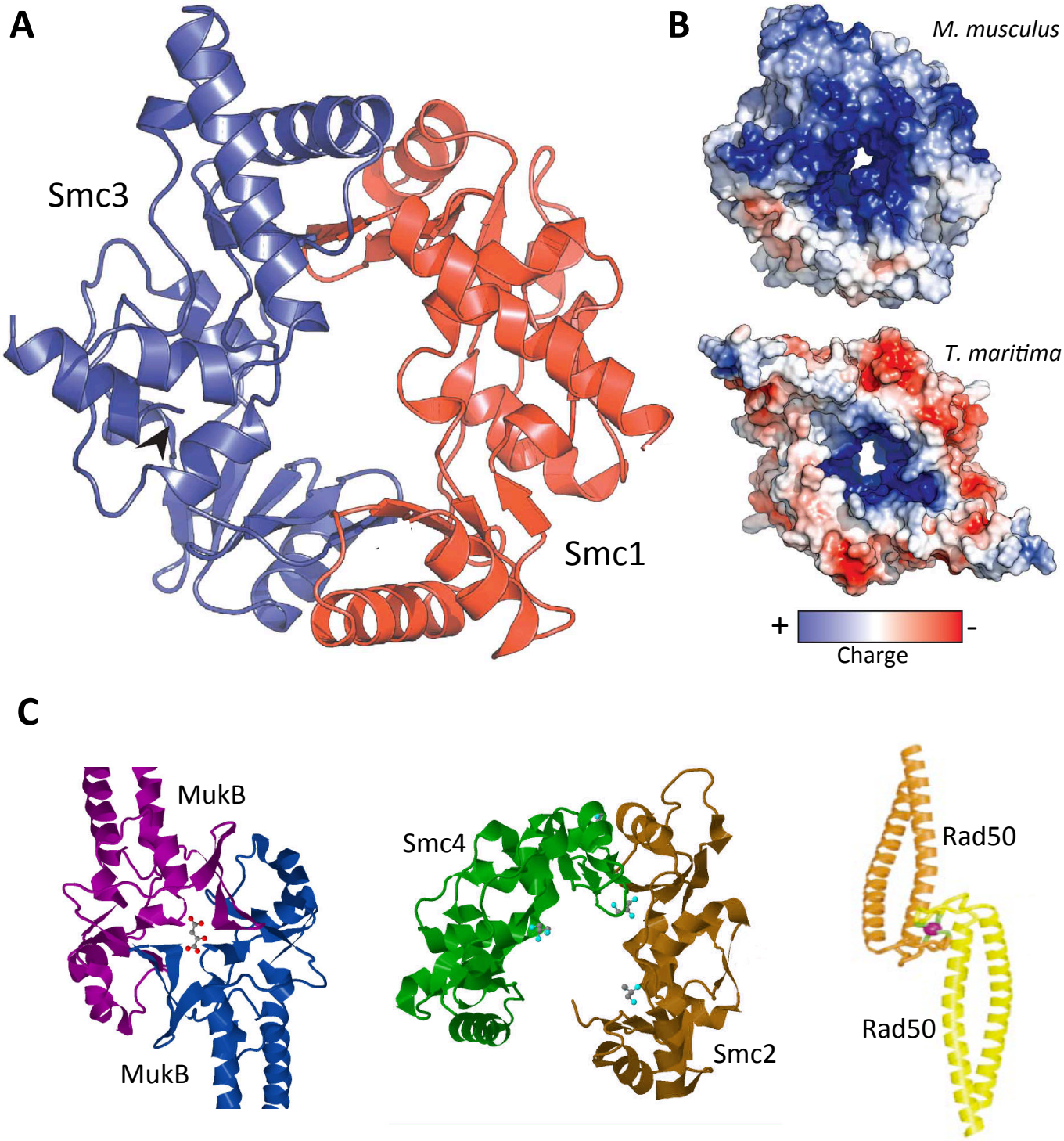


Figure I-7. Structure of SMC and SMC-like protein dimerisation domains.

(A) Crystal structure of the *M. musuculs* Smc1/3 hinge dimer, showing the two distinct interaction surfaces and overall toroidal shape.

(B) Charge-surface mapping reveals positively charged residues within the channel, conserved between *T. maritima* and *M. musculus.*

(C) Hinge structures from other SMC-like proteins, including (from left to right) MukB, Smc2/4 (*M. musculus* condensin crystallised in an open conformation), and the distinct zinc-hook that coordinates a Zn^{2+} for dimerisation of Rad50 molecules. The MukB hinge domain is comparatively small and lacks a central channel (Kurze et al., 2011; Ku et al., 2010; Griese et al., 2010; adapted by permission from Macmillan Publishers Ltd: *Nature,* Hopfner et al., copyright 2002).

THE COHESIN COMPLEX

The topological model

The discussion until now has concerned the commonalities between the SMC complexes, accredited to the conserved Smc proteins that form the architectural skeleton of each complex. In the case of the cohesin complex, which from this point on will be the primary topic of investigation, the long lengths of coiled-coil of the Smc1 and Smc3 proteins are central to enabling sister chromatid cohesion by virtue of their size and 3D shape. To build the cohesin complex, the NBD heads of the Smc1 and Smc3 heterodimers are bridged by the ∝-kleisin, known as Scc1 in budding yeast: the C-terminal winged-helix domain of Scc1 interfaces with the ATPase domain of Smc1, and on the basis of biochemical interactions, it was assumed a similar structure and binding with Smc3 took place at the N-terminus (Gruber et al., 2003; Haering, 2004). Indeed, all true Smc complexes appear to associate with different members of the kleisin family, a group of proteins so named for their part in closing these Smc structures (from kleisimo, Greek for 'closure'; Schleiffer et al., 2003). In the case of cohesin, the binding of Scc1 to connect Smc1/3 completes a tripartite ring structure, and this arrangement provides several clues as to how the complex functions (Gruber et al., 2003).

Historic notions of how cohesin provided cohesion between sister chromatids revolved around the elongated structure physically binding to the DNA of both chromatids and forming a tethering rope between the two. However, after the annular interconnection of subunits was realised, a new hypothesis suggested that the complex might instead entrap chromatid fibres in

a topological manner; the cohesin ring encircles them, and thus keeps them together. Importantly, the approximately 30nm diameter of the ring would be sufficiently large to confine two strands of DNA packaged with histones. This topological embrace model gained strong empirical support when it was shown that enzymatic cleavage of either the cohesin ring, or of its minichromosome substrate, was enough to release the plasmid from its association with cohesin (Ivanov and Nasmyth, 2005). The interpretation of these important experiments was that in the first instance, a proteolytic cut within the cohesin ring allowed escape of the minichromosome. Similarly, linearisation of the minichromosome with a restriction enzyme presumably entailed cohesin sliding off the end of the DNA strand.

A similar line of investigation into how the condensin complex associates with DNA recently evinced that topological entrapment is probably a common mechanistic feature of most SMC complex (Cuylen et al., 2011). The discovery that condensin encircles DNA similarly to cohesin, despite also forming a ring, was not entirely expected given a number of findings that suggested otherwise. For example, the continued association of the complex following cleavage of the Smc2 coiled-coil domain suggested the integrity of the holocomplex was not critical (Hudson et al., 2008). This seemed especially plausible considering that regions of the complex have been shown to bind directly to double stranded and single stranded DNA (Kimura and Hirano, 1997; Bazett-Jones et al., 2002; Chiu et al., 2004; Griese et al., 2010). Indeed, physical and direct protein-DNA interactions are common (e.g. transcription factors, polymerases, histones, etcetera) while topological associations are the exception. Within the Smc

complex superfamily itself, the Rad50-Mre11-Nbs1 complex operates through DNA binding to forge interconnections between the ends of double strand breaks (Fig. I-8; Lamarche et al., 2010; Hohl et al., 2011). This is analogous to how cohesin was originally thought to work: purely physically and without invoking the more complicated spatial relationships now acknowledged. That condensin has the ability both to bind DNA directly and also encircle it is not contradictory. Rather, this may reflect two classes of activity – one concerned with DNA repair (physical) and the other with chromosomal architecture (topological).

Although the results of the experiments with minichromosomes seemed very much indicative of a topological means of association, it was still debatable as to whether simply one cohesin ring performed the job of concatenating DNAs and the possibility of multiple rings acting together could not be dismissed. It was possible that two cohesin rings might act analogously to a pair of handcuffs, with the DNA of different sisters trapped within two physically connected cohesin complexes (Zhang et al., 2008). Alternatively, each of these rings might be topologically interlinked or connected by inter-complex associations that give rise to a larger, conglomerate ring structure (Fig. I-8; Huang et al., 2005). By chemically cross-linking the three interfaces of the cohesin ring[1] and examining the dimerisation between cohesed minichromosomes under denaturing conditions, Haering and colleagues revealed a convincing validation of topological entrapment: cohesion persisted in the absence of quaternary

[1] The Smc1-Smc3 and Smc1-Scc1 interfaces were cross-linked at cysteine substitution sites upon thiol-reactive chemical treatment. The Smc3-Scc1 interface was covalently closed by co-translation as fusion proteins separated by a flexible linker.

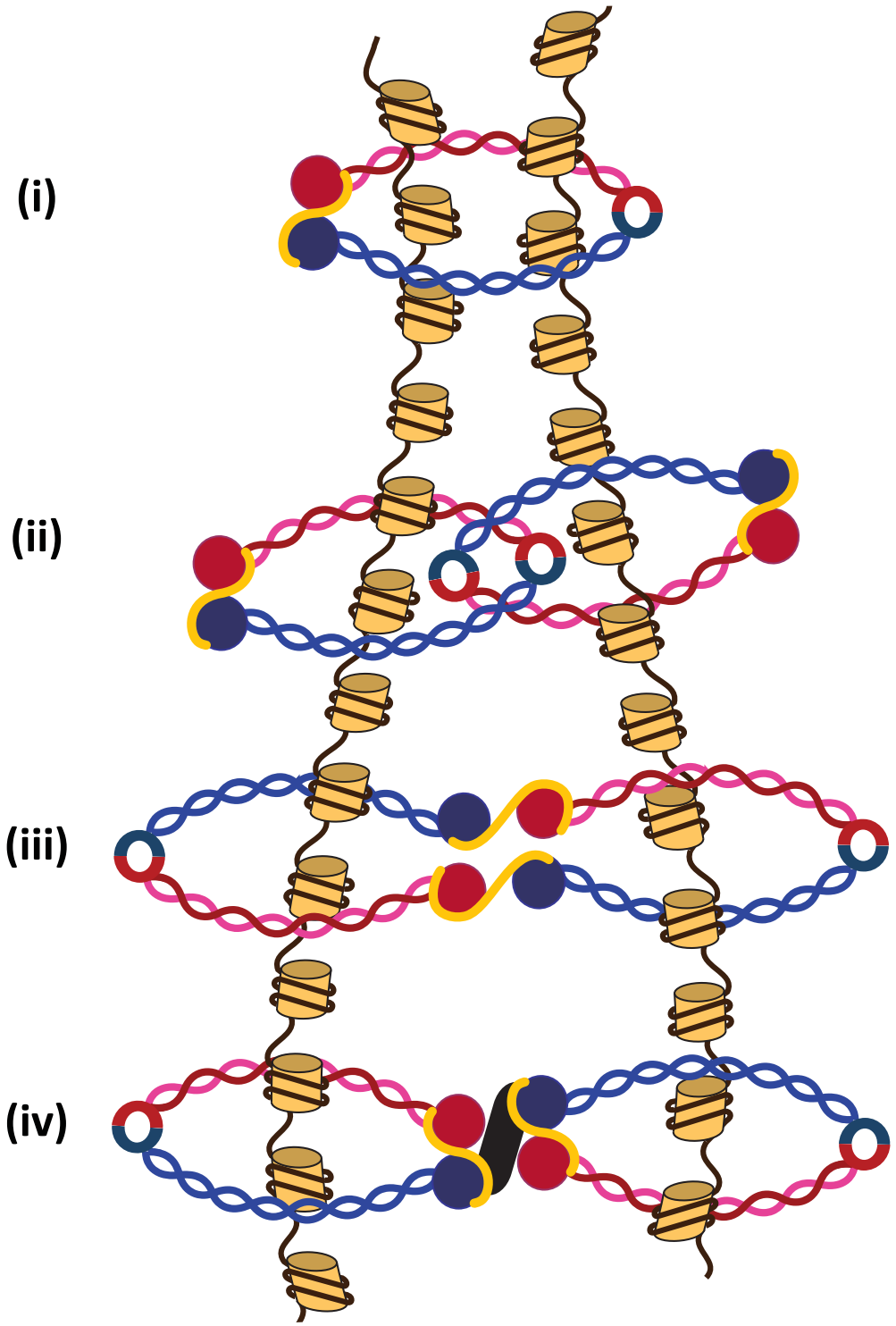


Figure I-8. Models of cohesin-mediated sister chromatid cohesion

i. The single ring model entails entrapment of two chromatin fibres within the one ring of Smc1 (blue), Smc3 (red) and Scc1 (yellow) with an approximately 30nm diameter. This is the simplest and most likely explanation for the provision of cohesion, based on available evidence.

ii. It is formally possible that two rings topologically entrap one another to concatenate chromatids.

iii. Alternatively, cohesin protein-protein interactions might be formed between subunits of multiple Smc complexes to create one larger ring.

iv. The handcuff model relies on a physical interaction to conjoin two cohesin complexes topologically encircling each sister chromatid. The Scc3 subunit (black) might bind in a 1:2 stoichiometry with kleisin subunits.

protein structure (Haering et al., 2008). This went a long way to discounting any models of cohesion necessitating physical interactions between either cohesin and DNA or between two or more cohesin complexes, as all such hypothetical interactions would have been abolished by SDS treatment. Furthermore, Haering and co-workers suggested that individual cohesin rings achieve entrapment of two minichromosomes, and by extension sister chromatids, as opposed to multiple rings. The basis for this analysis was genetic and probabilistic in nature. Minichromosomes extracted from a diploid strain, heterozygous for a TEV protease-cleavable version of Scc1, would show a different percentage loss in cohesion upon protease treatment according to whether this cohesion was reliant on one or two cohesin rings. In summary, the findings were in agreement with the single ring embrace model.

Recall that the timely separation of sister chromatids at anaphase depends on the activity of the separase enzyme. The principal function of separase is to cleave cohesin at two sites within the Scc1 kleisin subunit (Uhlmann et al., 1999). This presents us with a convincing congruence between how cohesion is bestowed and destroyed. The encirclement of sister chromatids by a proteinaceous ring maintains their essential association until anaphase, whereupon simple enzymatic cleavage releases this constraint and catalyses chromosomal segregation.

COHESION CONTROL

Cohesin complex loading

Proteins that operate using topological mechanisms are unusual. Such a mode of function raises an important question, that is, how does the cohesin complex come to find and entrap DNA in the first place? In fact, even the process by which the cohesin complex subunits associate and then localise to DNA remains relatively ill defined. While it was plausible that the complex merely assembled around pre-existent DNA, thereby encircling it, this would itself require intricate control. How is this process directed to occur only at DNA, and indeed only around DNA that has undergone replication to ensure the connection of sister fibres? In actuality, assemblage of the ring components is very unlikely as an explanation for building cohesion for a number of reasons. Foremost, the complex has been shown to exist in a soluble, non-chromatin-associated pool as a complete, trimeric ring (Gruber et al., 2003). This indicates that quaternary structure is established prior to recruitment to DNA. Secondly, in metazoa an intact population of the complex is found to bind to chromatin as early as telophase of mitosis (Shintomi and Hirano, 2009). In budding yeast, cohesin first associates with DNA at the end of G_1 and the start of S phase, correlating with the time at which Scc1 is first expressed in the cell cycle (Michaelis et al., 1997). Expression of Scc1 prior to this point results in its premature cleavage, as separase remains active due to lack of securin resynthesis following anaphase (Uhlmann et al., 1999; Morgan, 1999). In both instances, cohesin associates with DNA prior to the existence of duplicated chromosomes. Clearly, building cohesion must be far more elaborate than

perhaps first suspected. In both instances, cohesin associates with DNA prior to the existence of duplicated chromosomes. Clearly, building cohesion must be far more elaborate than perhaps first suspected.

Regardless of the differences in the cell cycle-dependent timing between yeast and other eukaryotes, there are several known requirements. For a stable, functional association with chromatin, a complex between two proteins, Scc2 and Scc4, is pivotal in acting as a loading factor. Without the Scc2/4 complex, cohesin cannot localise to DNA (Toth et al., 1999; Ciosk et al., 2000). Compared to other players in sister chromatid cohesion, both Scc2 and Scc4 proteins remain structurally and functionally unstudied. They are large proteins; Scc2 in particular ranges in size from approximately 170kDa in *S. cerevisiae* to over 300kDa in humans (Strachan, 2005). It is predicted to be composed of many HEAT (Huntington, Elongation factor 3, Protein Phosphatase 2A, TOR1) repeats (Andrade and Bork, 1995; Neuwald and Hirano, 2000; Panizza et al., 2000). Despite varying in size, it is well conserved and orthologs in *Sz. pombe* and other metazoa have been identified (Furuya et al., 1998; Rollins et al., 1999). Less than half the size of its partner, Scc4 produces abundant repetitive structural motifs of the tetratricopeptide (TPR) kind in prediction algorithms (Blatch and Lässle, 1999; D'Andrea and Regan, 2003). Its sequence is less well conserved, yet like Scc2 it has orthologs in all eukaryotes[2] and is essential for sister chromatid cohesion in each species (Watrin et al., 2006; Seitan et al., 2006). As stated, the

[2] With the exception of obligate, intracellular parasites, such as *Encephalitozoon cuniculi*, a microsporidian lacking several cohesin-related proteins within its highly compacted genome (Katinka et al., 2001; Nasmyth and Schleiffer, 2004). The most likely explanation is not that these proteins are superfluous, but that the host cell provides them instead.

two proteins form a very large complex (Scc2/4[3]) that is crucial for the recruitment of cohesin to DNA. Based on their structure, they are speculated to provide a binding scaffold to mediate multiple protein-protein interactions. On the basis of co-immunoprecipitation (co-IP) experiments, however, Scc2/4 only interacts weakly with cohesin and does not form a stable part of the core complex in either yeast or vertebrates (Toth et al., 1999; Watrin et al., 2006). Rather, it is the transient interaction during the so-called loading reaction in late G_1/S that makes the complex essential, as it is dispensable during later stages of the cell cycle. Inactivation of thermosensitive mutant alleles (*ts*) of either Scc2 or Scc4 in metaphase-arrested cells, for example, does not trigger a premature loss of cohesion and the cohesin ring remains DNA-bound (Ciosk et al., 2000).

The Scc3 protein is a stoichiometric, stably bound member of the holocomplex and as such it is considered a core component. The cohesin complex thus refers to the ring structure formed by Smc1, Smc3, and Scc1, plus Scc3. The latter binds the ring via a tight interaction with Scc1 (Haering et al., 2002; M.B. Roig, unpublished data). Like the Scc2/4 complex, Scc3 is also requisite for cohesin loading (Hu et al., 2011). In contrast, however, Scc3 seems compulsory for the maintenance of cohesion after its mere establishment (Toth et al., 1999). It is therefore unclear what contribution Scc3 might make towards inducing cohesin loading specifically, as opposed to purely maintaining the integrity of the complex at all times.

[3] Also known as kollerin, adherin, Mis4/Ssl3, and in humans, NIPBL/Mau2. For simplicity, the complex and its orthologs will be referred to as Scc2/4.

Loading loci

There are numerous questions posed by the observation that Scc2/4 is imperative for cohesin's association with DNA. Firstly, how is Scc2/4 itself targeted to DNA and is it restricted to specific chromosomal sites? The complex contains no structural domains that would indicate proclivity for particular nucleotide sequence elements or duplex structure. Examining the distribution of Scc2/4 along chromosomes allows some insight and speculation, but actually little is known about the determinants of its positioning. Moreover, there are several differences amongst organisms. Due to its ease of manipulation, the cell-free *Xenopus* egg extract is one of the most commonly used frameworks for researching the cohesin loading requirements. In this system, it has been shown that recruitment of Scc2 and cohesin requires assembly of the pre-replication complex (pre-RC), a large array of proteins that bind to origins of replications in preparation for the initiation of DNA synthesis (Gillespie and Hirano, 2004; Takahashi et al., 2004). In particular, Scc2/4 binds to the Cdc7-Drf1 kinase (Dbf4- or Drf1-dependent kinase, DDK), which phosphorylates the Mcm2-7 helicase for activation (Takahashi et al., 2008; Sheu and Stillman, 2006; Labib, 2010). This interaction may serve as a means of coupling initiation of DNA replication to the presence of cohesin, but the mechanistic details are yet to be explored.

In yeast, the factors targeting Scc2/4 and cohesin to DNA are equally unclear. Fission yeast may employ a similar pathway to *Xenopus*. Early studies identified the requirement for Swi6, the *Sz. pombe* equivalent to heterochromatin protein 1 (HP1), which establishes heterochromatic regions

on pericentromeric DNA where the majority of cohesin resides (Nakayama et al., 2000; Bernard et al., 2001; Nonaka et al., 2002). In keeping with the findings in *Xenopus*, Swi6 binds to DDK and may consequently act as an upstream targeting factor for the replication initiation proteins that recruit Scc2/4 and cohesin (Bailis et al., 2003; Hayashi et al., 2009). In budding yeast, there is no ortholog of Swi6/HP1 and thus the type of chromosomal receptors for Scc2/4 binding must be different. The association of cohesin with DNA is also not reliant on Cdc6 in *S. cerevisiae* (Uhlmann and Nasmyth, 1998). As Cdc6 activity is required for loading of Mcm2-7 to build licensed pre-RCs, downstream factors such as DDK must not be needed for the recruitment, at least initially, of cohesin to DNA in this species (Donovan et al., 1997; Detweiler and Li, 1997).

The most striking criterion for Scc2/4 and cohesin colocalisation identified in budding yeast is the kinetochore at the core centromere. The singular contribution of the centromere in conducting the accumulation of cohesin is exemplified by the fact that minichromosomes can be bi-oriented and stably transmitted. These small, circular plasmids can be engineered to contain a point centromere sequence of ~125bp (CEN) and a minimal length (~ 5kb) of generic nucleotide sequence, which is sufficient for cohesion (Megee and Koshland, 1999). Additionally, the distribution of cohesin on genomic chromosomes was seen to be greatest in pericentromeric regions, i.e. up to 50kb either side of the centromere, as identified by chromatin immunoprecipitation (ChIP; Tanaka et al., 1999; Blat and Kleckner, 1999; Glynn et al., 2004). Again, this was due not simply to the sequence of this pericentromeric region, but specifically the presence of the CEN, which when relocated to ectopic sites

caused an attendant accumulation of cohesin nearby (Megee et al., 1999; Weber et al., 2004). While the CEN is required, it is not sufficient for the localisation of Scc2/4 or cohesin. Rather, it appears that certain components of the kinetochore itself are the factors instrumental for recruitment (Eckert et al., 2007; Fernius and Marston, 2009). Because of the complexity of the kinetochore – composed of at least 70 proteins belonging to numerous subcomplexes – dissecting which factors are implicated in the cohesin loading reaction is far from straight-forward, but nevertheless warrants further research (Westermann et al., 2007).

Although the kinetochore appears most instrumental in loading cohesin, not all cohesin is localised to the surrounding regions. In addition to flanking the centromere, cohesin associates at sites along chromosome arms. In both fission and budding yeast, these sites are typically A-T rich sequences and intergenic regions, particularly in between genes transcribed in a convergent manner (Laloraya et al., 2000; Lengronne et al., 2004; Gullerova and Proudfoot, 2008). Recent work has shown that at the time of the initial loading step, cohesin and Scc2/4 colocalise at both core centromeres and highly transcribed genes (Hu et al., 2011). However, after the establishment of stable cohesion in S phase, Scc2/4 is no longer required for association and the distributions of the two complexes differ, suggesting that following topological chromatin entrapment the ring translocates to intergenic and pericentromeric regions from sites of loading at transcribed genes and the centromere, respectively (Ciosk et al., 2000; Lengronne et al., 2004; Schmidt et al., 2009; Hu et al., 2011; Hu et al., in press).

This sliding across kilobase-distances of chromatin implies, and is consistent with, a topological association. How then, does DNA enter into the ring? A potential explanation is that one of the three protein-protein interfaces may dissociate transiently to allow passage of DNA within. Experiments involving domain fusions and conditional locking of Smc1, Smc3 and Scc1 to one another indicate that of the three interaction surfaces of the trimer, only the hinge appears to require dissociation for cohesin function (Gruber et al., 2006). Congruently, selective closure of the hinge interface prevents the establishment of gainful cohesion. Assuming that the hinge represents the DNA entry gate, it follows that hinge dissociation must be regulated for stable entrapment, i.e. opening must occur to allow DNA inside, and thereon be inhibited to prevent escape.

A candidate for regulating this mechanism is the hydrolysis or binding of ATP by the SMC nucleotide-binding domains. It appears that binding of ATP by Smc1/3 represents an intermediate step in the loading reaction, as Walker A mutants (binding-defective) cannot localise to chromatin. Hydrolysis mutants also show an abnormal association with chromatin, localising with Scc2/4, primarily at kinetochores, but turning over rapidly and being unable to establish cohesion (Arumugam, 2003; Weitzer, 2003; Hu et al., 2011). The ATPase activity therefore seems important for converting a transient loading conformation into one that topologically entraps chromatin. Hydrolysis of ATP could in theory open the hinge, analogously to how ABC transporters accomplish substrate translocation via transmembrane domains.

Establishment of cohesion: temporal control

Uhlmann and Nasmyth first identified the orchestration of cohesion establishment as being linked to S phase (1998). Although able to bind chromosomes when induced after S phase, Scc1 could not build cohesion. In keeping with this, cohesin exists in two populations, defined by their stabilities on chromatin. The time taken for fluorescence recovery after photobleaching (FRAP) of GFP-tagged cohesin subunits, as well as detectable DNA association in chromosome spreads, reveals a shift from a dynamic, unstable state in G_1 to a securely bound state with a long residence time for a subpopulation of cohesin after DNA replication (Gerlich et al., 2006; Bernard et al., 2008). It is this durably bound fraction that provides cohesion, but by what mechanism is it generated?

A clearer picture now exists of the steps involved in the conversion of cohesin from a state of rapid turnover to one conferring true, functional cohesion. In *S. cerevisiae,* this requires the activity of the acetyltransferase protein, Eco1, which acetylates two adjacent lysine residues on the NBD of Smc3[4] (Unal et al., 2008; Ben-Shahar et al., 2008; Zhang et al., 2008; Rowland et al., 2009). The same modification of these highly conserved residues is also critical in humans (Zhang et al., 2008). Acetylation persists from the time of DNA synthesis until the loss of cohesion at anaphase, and is not present before replication takes place (Beckouët et al., 2010). This correlates precisely with the transition of cohesin from short to long residence times on DNA and is thus a mark of stable cohesin, and by extension, cohesion. In keeping with this,

[4] K112, K113 in yeast, and K105, K106 in humans.

mutating the critical lysines to a similarly charged but non-modifiable residue, i.e. arginine, has no effect on the loading of cohesin. However, sister chromatids are prematurely split, indicative of a cohesion defect (Toth et al., 1999; Rowland et al., 2009). On the other hand, substitution with amino acids that partially mimic the acetylated state (glutamine or asparagine) or disrupt the surface in some other way (threonine) can rescue these defects (Unal et al., 2008; Ben-Shahar et al., 2008; Rowland et al., 2009).

Recent evidence has emerged that another post-translational modification, sumoylation, might be equally as significant as acetylation in the establishment phase of cohesion (Almedawar et al., 2012). The small ubiquitin-like modifier (SUMO) of lysine residues had been detected on cohesin subunits previously, both as part of the normal cell cycle as well as the DNA damage response, but the significance was unknown (Stead et al., 2003). Almedawar and colleagues investigated this obscurity, discovering that SUMO, like acetylation, arises in synchrony with DNA duplication and is detectable on Smc1, Smc3, Scc1, Scc3 and Pds5 (2012). Abrogation of cohesin ring sumoylation produces cohesion defects. Tellingly, this is despite normal acetylation. In the situation of DSB repair, discussed below, phosphorylation is known as an upstream, prerequisite modification of Scc1 that serves to recruit Eco1 for subsequent acetylation (Heidinger-Pauli et al., 2009). In other words, phosphorylation and acetylation work in the same pathway in this context. This is in contrast to sumoylation and acetylation, which appear to be entirely separable yet equally vital steps in promoting the functional entrapment of sister chromatids during DNA replication.

There are a number of influences that couple the establishment step to DNA replication. The first is a physical interaction between the processivity clamp, PCNA (proliferating cell nuclear antigen), and Eco1, which hypothetically serves to target the acetyltransferase to active replication forks (Moldovan et al., 2006). Indeed, genetic interactions between Eco1 and numerous replication-associated factors have been documented (Skibbens et al., 1999; Kenna and Skibbens, 2003). A secondary, yet equally important, means of regulation is the cell cycle-dependent degradation of Eco1 itself. Following S phase, Cdk1-dependent phosphorylation of Eco1 leads to ubiquitination and ensuing proteasomal destruction, thereby limiting the ability to build cohesion in G_2/M (Lyons and Morgan, 2011). In fact, when over-expressed during this period, or mutagenised to prevent degradation, Eco1 can generate cohesion anew (Unal et al., 2007).

Although these are artificial situations, it should be noted that cohesion establishment is not strictly limited to the period of active DNA replication, but can also be generated in response to DNA damage. If cells detect DSBs in G2/M, then cohesin localises directly to these sites and is loaded *de novo* at sequences flanking the breaks (Strom et al., 2004; Unal et al., 2004). Remarkably, not only is this newly bound cohesin converted to a cohesive state, but cohesin rings genome-wide become stably bound, even if located on different chromosomes to those damaged (Ström et al., 2007; Unal et al., 2007). While reliant on the acetyltransferase activity of Eco1, it is not the Smc3 head domain modified in this case but residues at the N-terminus of Scc1 instead (Ivanov et al., 2002; Heidinger-Pauli et al., 2009). These precise control mechanisms ensure that

cohesion is established only when appropriate, to allow conjunction of sister DNA molecules upon their synthesis, or to facilitate homologous recombination for DNA repair. As for DNA replication-dependent cohesion, SUMO is emerging here too as a candidate for controlling cohesin dynamics. Cohesin can localise to DSBs, but without sumoylation cannot establish cohesion (Wu et al., 2012; McAleenan et al., 2012).

Establishment of cohesion: mechanism

Exactly how Smc3 acetylation (and in the case of DNA damage, Scc1 acetylation) and cohesin sumoylation afford a functional switch remains enigmatic. Clues were first provided with the identification of suppressor mutations that allow the requirement for Eco1 to be bypassed. The majority of these mutations were pinpointed to three proteins: Scc3, Pds5, and Wapl[5], which all associate with cohesin and form a trimeric subcomplex (Rowland et al., 2009). The implication was that somehow, these cohesin subunits act to prevent the untimely establishment of cohesion, and as such, the purpose of acetylation is to counteract this activity.

Like Scc3, Pds5 is another large, HEAT repeat protein that can bind directly and independently to the cohesin ring (Shintomi and Hirano, 2009). The Wapl subunit, on the other hand, interacts with both Scc3 and Pds5 (and potentially other sites within the ring), but requires Pds5 for this association (Rowland et al., 2009). Loss of either Pds5 or Scc3 function in G_2/M leads to cohesion defects, pointing towards a dual function for these proteins in both

[5] Formally identified as Rad61 but known as Wapl for consistency across species.

preventing the establishment of cohesion prior to S phase, as well as maintaining this same cohesion thereafter (Panizza et al., 2000). Reconciling these opposing actions is the surmised role of Wapl as chief antagonist of cohesin's association with DNA.

The absence of Eco1 could be bypassed by complete deletion of Wapl, implying Wapl is dedicated to inhibiting cohesin complex stability (Rowland et al., 2009; Unal et al., 2008). In fact, previous studies had suggested Wapl acts to remove cohesin from DNA (Kueng et al., 2006; Gandhi et al., 2006; Nishiyama et al., 2010). Recently, Chan et al. made the insightful discovery that fusing Smc3 to Scc1, thereby closing the interface, was able to circumvent the need for Eco1 and acetylation in cohesion establishment and maintenance (Chan et al., 2012). This provided good evidence that acetylation normally serves to close or stabilize the Smc3-Scc1 interface and also that Wapl opposes this. Consistently with this model, expression of Wapl in G_2/M was unable to promote the dissociation of cohesin from DNA in this background (*eco1Δ, GAL1p-WPL1, Smc3-Scc1 fusion*), whereas it could in strains lacking the fusion (*eco1Δ, GAL1p-WPL1*). In higher eukaryotes, an additional protein known as sororin is needed for the maintenance of cohesion[6] (Rankin et al., 2005; Schmitz et al., 2007; Zhang and Pati, 2012). Recruited to the complex at the time of cohesin acetylation, sororin binds Pds5 and competitively displaces Wapl from an interaction with Pds5 but not from the complex entirely (Lafont et al., 2010; Nishiyama et al., 2010; Whelan et al., 2011). Additionally, sororin is no longer required for cohesion maintenance throughout G_2/M when Wapl is depleted.

[6] Orthologous sororin proteins have been identified in most invertebrate phyla, but not in yeast (Nishiyama et al., 2010).

This furthers the case that essentially cohesin exists in equilibrium between recruitment to DNA by Scc2/4 and removal by Wapl, until the action of the latter is neutralised by the establishment step. In yeast, cohesin then stably resides upon DNA until sister chromatid disjunction at anaphase. In metazoa, total loss of cohesion at anaphase is prefaced by the separase-independent and selective removal of cohesin from chromatid arms during the so-called 'prophase pathway'.

The prophase pathway

The prophase pathway is a phenomenon observed during the eponymous stage of mitosis and involves the removal of the majority of cohesin from the arms of paired sister chromatids (Waizenegger et al., 2000). With arms no longer cohesed, sister chromatids are said to be 'resolved' and it is this prophase pathway that gives rise to the characteristic X-shape of paired chromatids in chromosome spreads (Fig. I-9).

Considering that separase cleavage of Scc1 is not required for this process, how cohesin is removed remains an interesting, open question, with the possibility that it requires the disruption of one of the three interfaces of the ring in the same way that occurs during loading. Given that prophase removal of cohesin requires Wapl (Gandhi et al., 2006), it is speculated that the Smc3-Scc1 interface may be responsible for allowing exit of the cohesin ring in the same manner observed with the releasing activity of Wapl against non-acetylated cohesin in yeast. There are a number of differences in these processes, however, that warrant a thorough investigation of this assumption. In the prophase

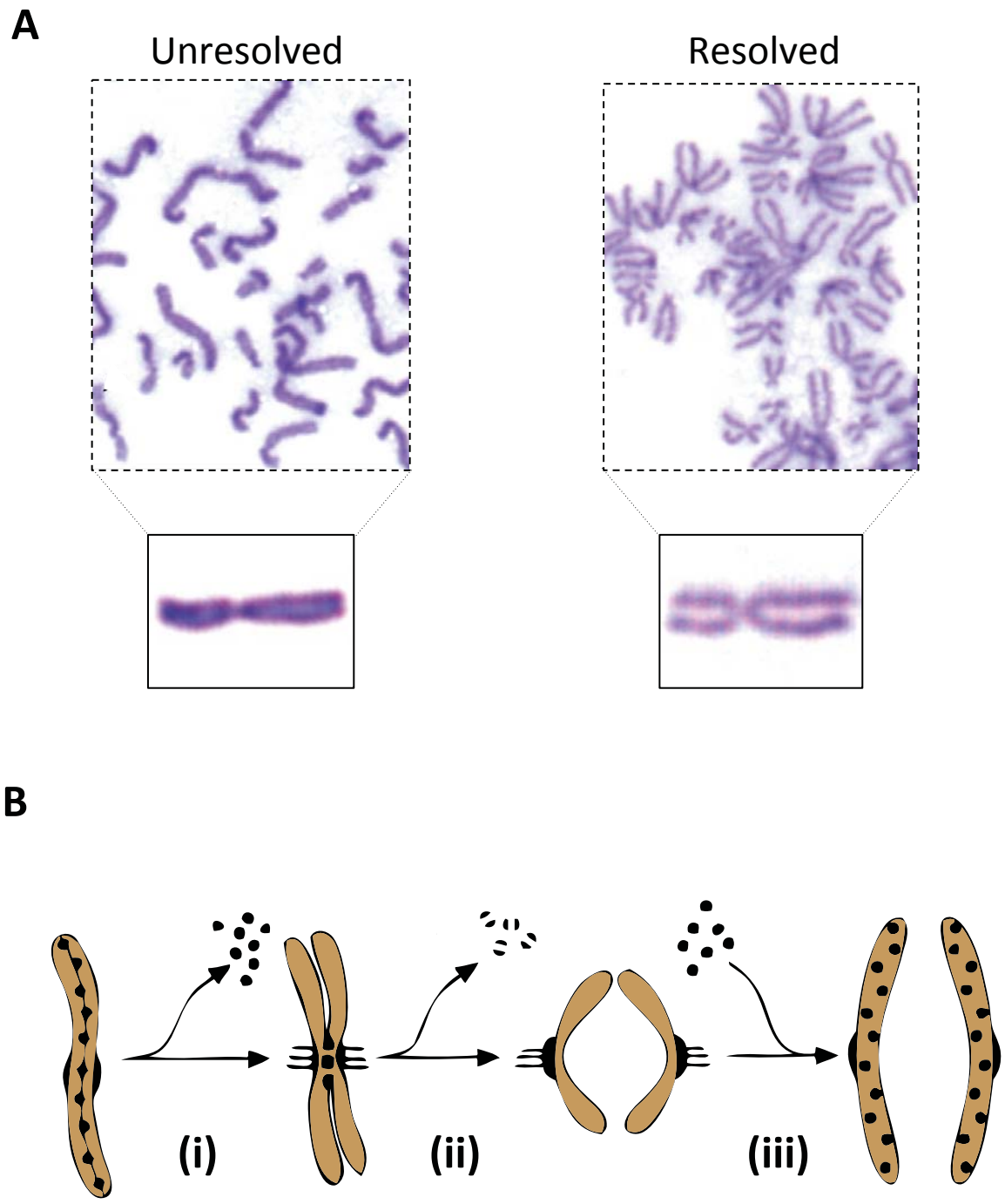


Figure I-9. Sister chromatid resolution and cohesin removal.

(A) HeLa cell chromosome spreads with Giemsa staining of DNA, showing chromatids with cohesed arms (left) and resolved arms during prophase (right; adapted with permission from Elsevier: *Cell,* Kueng et al., copyright 2006; Peters et al., 2008).

(B) The prophase pathway (adapted with permission from Elsevier: *Cell*, Waizenegger et al., copyright 2000). Arm cohesin is removed non-proteolytically (i) and centromeric cohesin is spared this process until separase cleavage of Scc1 and sister chromatid separation (ii). Soluble, intact cohesin released by the prophase pathway reassociates with chromosomes in telophase (iii).

pathway, cohesin is released from chromosome arms irrespective of acetylation status; both acetylated and non-acetylated populations dissociate (Schokel et al., 2011). This suggests that the normal protective function of acetylation is overridden and this may be mediated by the activity of PLK1 and Aurora B in phosphorylating cohesin subunits to make the complex more susceptible to the dissociating activity of Wapl (Sumara et al., 2000; Hauf et al., 2005). In fact, the centromeric cohesin resistant to prophase removal is protected by PP2A, a phosphatase, recruited there by Sgo1 (shugoshin; McGuinness et al., 2005; Xu et al., 2009). While it is feasible that the cohesin holocomplex undergoes a more drastic dissolution, it is perhaps more likely that a single interface is disrupted through which DNA can exit. The complicity of Wapl in the prophase pathway makes the Smc3-Scc1 interface a likely target site.

Unlike in metazoa, in *S. cerevisiae* and *Sz. pombe* cohesion is preserved along chromosome arms until anaphase. As discussed, upon satisfaction of the SAC, separase binds to and cleaves the Scc1 subunit at two sites. Soluble cohesin is spared cleavage, as separase activity is dependent on DNA binding (Sun et al., 2009). With the covalent disintegration of the kleisin, cohesin can no longer resist spindle pulling forces and chromatids separate. It was noted that Smc3 acetylation is lost at the time of cohesin dissociation in anaphase and this has been ascribed to the action of a counterpart to Eco1, namely the deacetylase Hos1 (Beckouët et al., 2010; Borges et al., 2010; Xiong et al., 2010). Again, this is a conserved feature of the cohesin cycle verified by the presence of the vertebrate Smc3 deacetylase, HDAC8, in humans (Deardorff et al., 2012). It is the separase cleavage of Scc1 and the dissociation of the complex from DNA that

coincides with the activity of Hos1. The purpose of deacetylation is conjectured to allow recycling of the Smc subunits - permitting their reassociation with chromatin in the following cell cycle and resetting their state to one competent for replication-coupled cohesion establishment. Evidence suggests that in the absence of deacetylation, the N-terminal separase cleavage fragment of Scc1 is stabilized by virtue of its interaction with Smc3 (F. Beckouët, unpublished data). Deacetylation allows Smc3 to associate with an intact kleisin in the subsequent G_1/S phase transition. This is consistent with the hypothesis that acetylation stabilizes the kleisin-Smc3 interaction (Chan et al., 2012).

Cleavage of Scc1 opens the ring and duly allows segregation of the genome but the function of the prophase pathway is in fact unknown. However, it is hypothesised that the liberation of a large, intact fraction of cohesin provides a ready pool from which to reload the complex in telophase, in order that it can resume its function in the regulation of gene expression during interphase. This notion is also consistent with the absence of the prophase pathway in yeast, where the role of cohesin in regulating gene transcription seems less pronounced.

Cohesin in transcriptional regulation

Cohesin is a regulator of gene expression in metazoa, and to a lesser extent, in yeast too. This activity is best attested to by the fact that the complex continues to play a vital role even in post-mitotic cells, for instance neurons, of metazoa (Pauli et al., 2008). Although this might appear a highly diverged function for this complex, research indicating the way in which it acts to effect

change in expression profiles is consistent with what we know about cohesin thus far. A sophisticated understanding of the elements regulating gene expression has shown that in addition to the mere assembly of arrays of transcription factors at promoters, DNA is subject to rearrangement over large distances to allow for the interaction of enhancers with the transcription initiation complex. In essence, long stretches of DNA can be folded, producing loops, to bring together two sites and their DNA-bound proteins to modulate transcription (Dekker, 2002; Dixon et al., 2012). In addition to holding together DNA of two separate chromosomes, cohesin might equally encourage this looping within chromosomes. As a matter of fact, both inter- and intra-chromosomal regulatory interactions have been reported (Spilianakis et al., 2005; Ling et al., 2007; Splinter et al., 2006). Notably, the mechanism suggested employs topological entrapment. However, this is yet to be convincingly substantiated in this context. As the bulk of interphase cohesin is in dynamic flux between the bound and soluble phase, it is intriguing how these loops might be maintained.

Unsurprisingly, there are a number of other players involved in gene regulation alongside cohesin, one of which is the CTCF (CCCTC binding factor) zinc-finger protein (Parelho et al., 2008; Lee and Iyer, 2012). Characterised as binding to insulators, CTCF helps establish boundaries between transcriptionally active chromatin and silent regions, as well as blocking *cis* interactions between enhancers and cognate genes. This activity is cohesin-dependent (Wendt et al., 2008; Nativio et al., 2009). The mediator complex has also been demonstrated to work in concert with cohesin, binding directly with it

and bringing about looped enhancer-promoter interfaces (Kagey et al., 2010; Ebmeier and Taatjes, 2010). It seems therefore that cohesin can act, depending on its partner, to both augment and repress transcription. Cohesin probably acts uniformly in both circumstances, while the relative spatial arrangements of enhancers and promoters differ (Fig I-10; Cuylen and Haering, 2010).

Cohesin in health and disease

Surveying the myriad processes this extraordinary complex participates in - gene regulation, meiotic recombination, DNA repair, centriole division, chromosomal condensation, and of course sister chromatid cohesion - it is inevitable that a diversity of pathologies would arise from its dysfunction.

Most obviously, when cohesin fails in its canonical role, the incorrect segregation of sister chromatids can produce cells with abnormal karyotypes. For single-celled organisms, this can be lethal. In metazoa, aneuploidy has similarly deleterious effects ranging from neoplasia within somatic tissue, to developmental disorders of offspring if arising in the germ cells of an incorrect meiotic division. In this regard, cohesin might also bear the incredible task of sustaining cohesion within prophase I-arrested homologous chromosomes over periods of several decades in oocytes. Recent work suggests that meiotic cohesin is not continually synthesized and does not turn over from the time it is bestowed at the commencement of meiosis until ovulation and fertilization (Tachibana-Konwalski et al., 2010; Revenkova et al., 2010). Such findings allude strongly to an age-related deterioration in oocyte cohesion, accounting for the striking correlation between increasing maternal age and incidence of

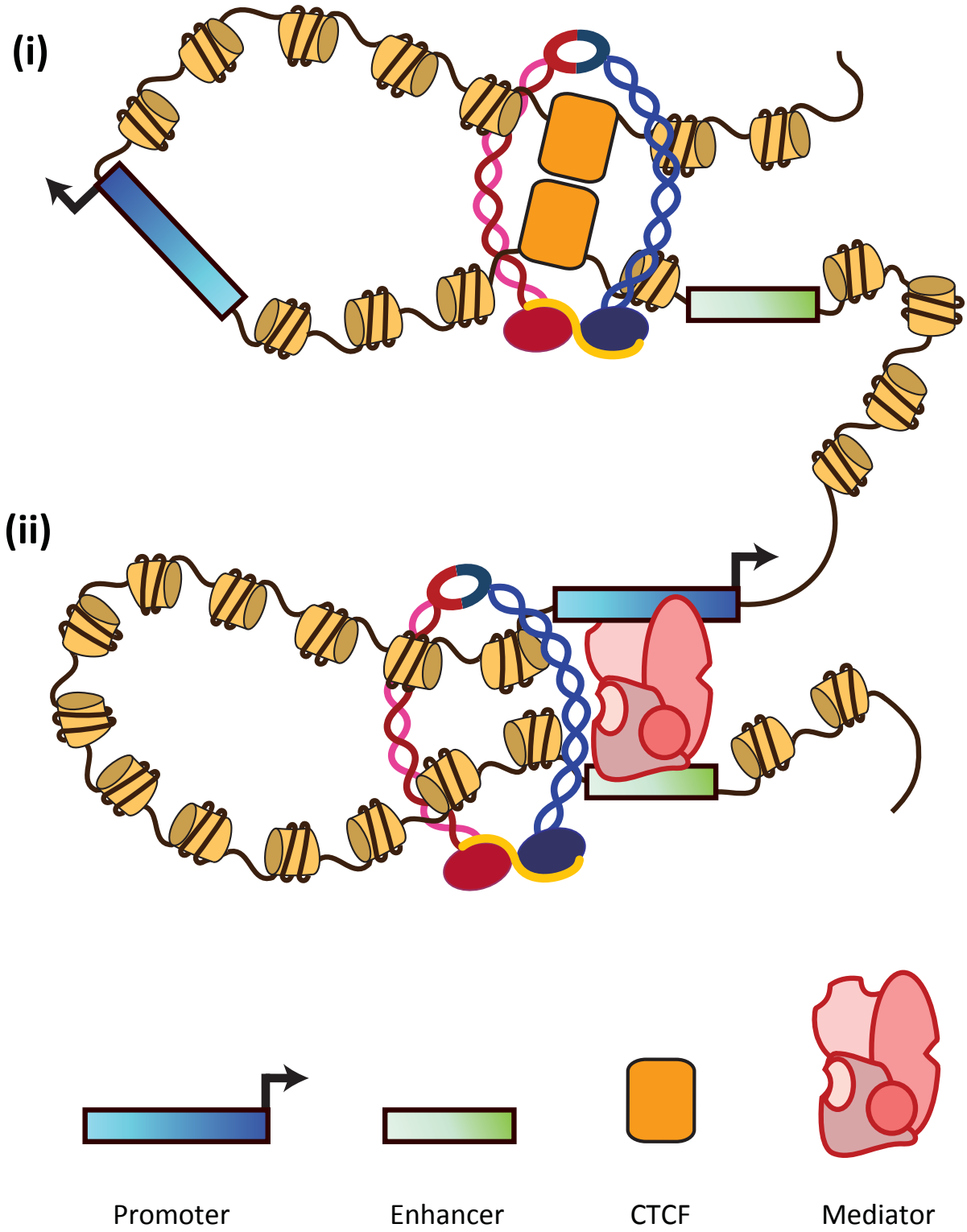


Figure I-10. Mechanisms of transcriptional regulation by cohesin.
Cohesin topologically entraps chromatin to form loops and modulate long-distance interactions between enhancers and promoters. Loops can be coordinated by cohesin in concert with CTCF to insulate gene promoters (i) or with the mediator complex to enhance transcription initiation (ii).

trisomies, e.g. Down syndrome (trisomy 21), in births (Fig. I-11; Newberger, 2000; Jessberger, 2010).

Chromosomal abnormalities and concomitant disease can derive from a time-dependent dysfunction of otherwise satisfactory cohesion, as described above. On the other hand, specific mutations amongst several of the cohesin subunits represent the aetiologies of a specific class of severe developmental disorders (Liu and Krantz, 2008; McNairn and Gerton, 2008). Known together as cohesinopathies, Cornelia de Lange syndrome (CdLS) and Roberts syndrome are two rare diseases, the former occurring approximately every 1 in 10000 live births from spontaneous mutations in cohesin-related genes, predominantly *NIPBL* (Scc2 ortholog), but also *SMC1*, *SMC3* and recently *HDAC8* (Hos1 ortholog; Opitz, 1985; Tonkin et al., 2004; Krantz et al., 2004; Musio et al., 2006; Deardorff et al., 2007; Deardorff et al., 2012). Roberts syndrome is an autosomal recessive mutation in the *ESCO2* gene and its rarity is exceptional (Vega et al., 2005). Both diseases entail a litany of severe physiological abnormalities encompassing mental retardation, growth defects, and limb malformations (Fig. I-11). These conditions are likely caused by dysfunctional activity of cohesin factors in regulating gene expression during development, rather than the possibility that all symptoms represent downstream effects of chromosomal missegregation (Horsfield et al., 2012). Naturally, dissecting these two activities and their relative contributions to cohesinopathogenesis is a challenging enterprise.

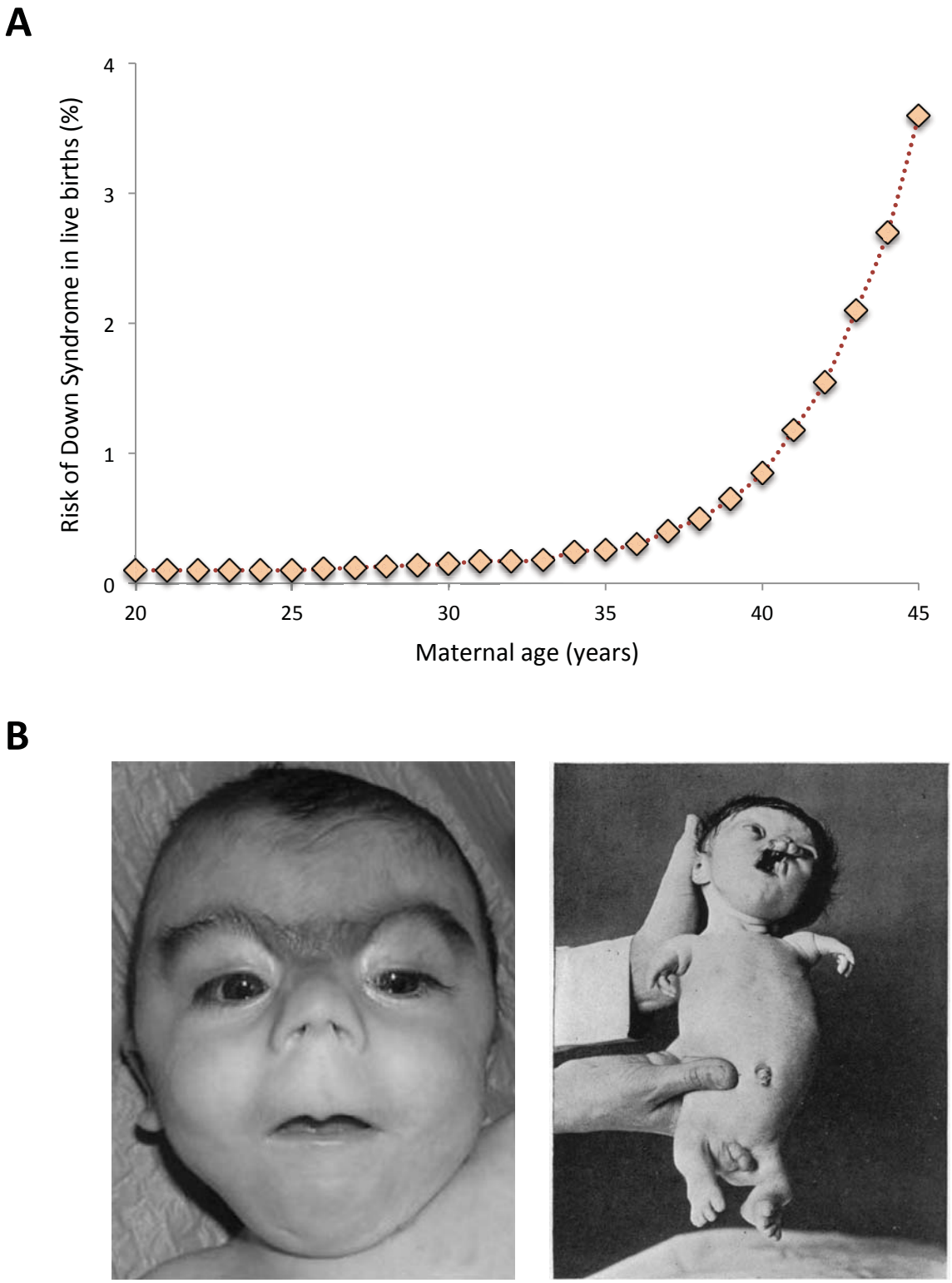


Figure I-11. Clinical implications of cohesin complex dysfunction.

(A) Down syndrome (trisomy 21) is strongly correlated with maternal age. This is speculated to be due to a loss of meiotic cohesion in oocytes with time (data adapted from Newberger, 2000).

(B) The cohesinopathies, Cornelia de Lange syndrome (left panel) and Roberts syndrome (right), are caused by mutations in subunits of the cohesin complex and related regulatory factors. Characteristic facial features and limb malformation are symptomatic of these developmental disorders (reproduced with permission from Wiley-Liss, Inc: *American Journal of Medical Genetics*, Pié et al., copyright 2010; Roberts, 1919).

Thesis objectives

The aims of this thesis were to investigate the mechanisms and nature of how the cohesin complex associates with DNA.

In particular, I sought to identify a potential cofactor for cohesin loading and investigate the engagement of Smc1/3 nucleotide-binding domains as implicated in this process. To address how cohesin is dissociated from DNA in a Wapl-dependent manner, a new technique was explored for conditional cross-linking of an interface of the tripartite complex. Finally, the structure of Pds5 was sought with a view to characterising its mechanism and involvement in the aforementioned processes.

CHAPTER II

Biochemical requirements for cohesin complex loading

Biochemical requirements for cohesin complex loading

INTRODUCTION

Cohesin loading

From the moment its subunits are translated, the cohesin complex begins a remarkable sequence of processes. Subject to exquisite control at each stage, cohesin can support the union of duplicate chromosomes until their punctual disjunction when a cell divides. This point, anaphase, marks the end to another cycle of cohesin's association with DNA. In the daughter cells produced, this cycle starts anew with the directed loading of cohesin onto chromatin. In budding yeast, this occurs towards the end of G1, in the preparatory lead up to replication of the genome.

Reviewing the current body of knowledge, there are a number of factors and steps that have been shown to be vital for loading cohesin onto chromatin. Foremost is the dedicated loading complex, Scc2/4, which may serve as a scaffold upon which cohesin can dock and undergo a necessary conformational change. Without Scc2/4, cohesin cannot localise to DNA whatsoever (Tóth et al., 1999; Ciosk et al., 2000). As there is no evidence that Scc2/4 has intrinsic DNA binding activity, the complex itself requires additional and as yet unknown protein factors for localisation to DNA. In yeast, immunoprecipitation of Scc2 or Scc4 has failed to show any stable binding partners, and therefore the interaction is likely to be weak or transient. Accordingly, such factors may

already be present on DNA. Evaluating the genome-wide loci of Scc2/4 and cohesin at the time of loading can be used to help identify these candidate chromosomal receptors.

The kinetochore is a cohesin loading site

The predominant locus of cohesin loading for budding yeast chromosomes is the centromere. In fact, the accumulation of pericentromeric cohesin is reliant not just on the *CEN* but more specifically protein members of the kinetochore (Weber et al., 2004; Eckert et al., 2007; Fernius and Marston, 2009; Ng et al., 2009). This might be attributable to one or more interactions between Scc2/4 and kinetochore subunits. A fluorescently-labelled Scc2 construct localises to the centromere both independently of cohesin and with high turnover (Hu et al., 2011). The Ctf19 complex subunit Chl4 is required for centromeric enrichment of Scc2/4, yet is dispensable for bi-orientation (B. Hu, unpublished data; Fernius and Marston, 2009). This suggests a greater involvement of this protein in the targeting of Scc2/4 and cohesin to DNA than in the support of a kinetochore structure capable of attaching microtubules. Similarly, Iml3 is another peripheral subunit of the Ctf19 complex that displays a strong mutant phenotype of reduced cohesin association. While these are good candidates, other proteins may well be involved. The kinetochore is a gigantic, intricate structure: greater than 5MDa, consisting of approximately 70 different proteins each organised into at least 7 different subcomplexes, which in turn are present in multiple copies (Fig II-1; Joglekar et al., 2006; Westermann et al., 2007; Santaguida and Musacchio, 2009). Unsurprisingly, we still lack a complete understanding of kinetochore architecture and subcomplex function.

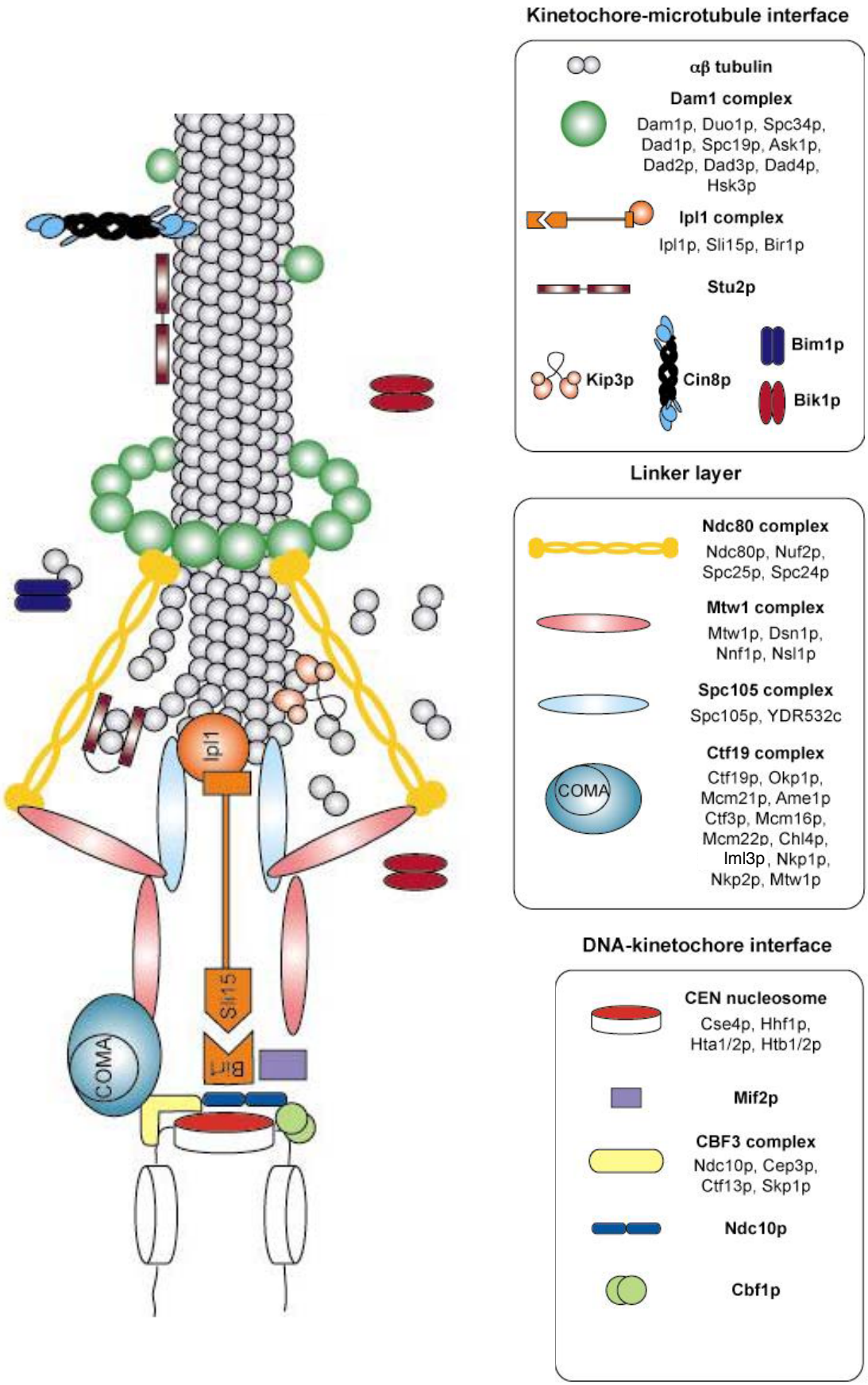


Figure II-1. Structural organisation of the *S. cererevisiae* kinetochore.
The budding yeast kinetochore is comprised of multiple subcomplexes that bind DNA (inner kinetochore) to allow microtubule attachments to impart force. Some subunits are drawn according to structural information (adapted from Westermann et al, 2007).

Cohesin loading is associated with active transcription

In the absence of non-essential kinetochore proteins such as Chl4, centromeric accumulation of cohesin is lost. Nonetheless, sister chromatids can still bi-orient by virtue of cohesin along the arms (Fernius and Marston, 2009). Arm cohesin is thought to require an alternative factor for loading. Contenders for involvement here are proteins involved in transcription, as cohesin in the act of loading via Scc2/4 shows elevated ChIP-SEQ and ChIP-CHIP signals at sites of active transcription (Lengronne et al., 2004; Hu et al., 2011). It was postulated that rather than requiring disparate chromosomal receptors at kinetochores and transcription sites, cohesin and Scc2/4 might interact with a single factor common to both. In vertebrates, it is published that cohesin and Scc2/4 colocalise with the mediator complex, a large collection of proteins that enhance transcription (Kagey et al., 2010; Taatjes, 2010). In yeast, the mediator has also been found at the kinetochore (Esnault et al., 2008). This is of particular interest given the absence of transcription at the centromere, suggesting a non-canonical function. Therefore mediator subunit and transcription-related proteins were tested along with kinetochore proteins for their ability to recruit Scc2/4.

The protein(s) responsible for recruiting Scc2/4 likely interact with cohesin also, as Scc2/4 targeted to a site outside of the kinetochore is insufficient to colocalise cohesin (Hu et al., 2011). Identification of cofactors for cohesin complex loading is an important goal and will be necessary for a complete description of the mechanisms involved.

The majority of studies indicating the importance of kinetochore members have been largely genetic in nature, showing a loss of pericentromeric cohesion in response to inactivation of these various proteins. Dysfunction may arise from downstream effects or interactions with other proteins. The inherent difficulty in interpreting such results calls for a more direct approach to pinpoint the chromosomal components of the loading reaction. Here, I sought to identify a direct physical interaction between the cohesin loading complex and specific members of the kinetochore *in vivo*, using live-cell fluorescent imaging. Theoretically, the assay employed should detect even transient interactions with a high sensitivity. As Scc2/4 shows rapid turnover at the kinetochore by FRAP, conventional biochemical techniques would most likely fail to reveal any associations.

Cohesin Smc ATPase activity

The second principal requirement for loading cohesin onto DNA is the functionality of the ATPase domains. Cohesin complexes defective for binding or hydrolysis of ATP fail to either associate with DNA or localise unusually (Arumugam et al., 2003; Weitzer, 2003). Smc proteins share the NBD structure conserved amongst ABC transporters. In these transmembrane proteins, the NBDs act as motor domains to power substrate transport across bilayers (Jones et al., 2009).

ABC ATPases are identified by six motifs necessary for a normal active site and fully functional hydrolytic activity (Hopfner and Tainer, 2003; Rees et al., 2009). The Walker A motif, also known as the P-loop, has the sequence

GXXGXGK[S/T] and is critical for binding ATP. An aromatic side chain from the N-terminal 'A-loop' assists in coordinating the adenosine ring (Ambudkar et al., 2006). The Walker B motif consists of the sequence ϕϕϕϕDE, where ϕ denotes a hydrophobic residue (Walker et al., 1982; Aasland et al., 2002). In the ABC superfamily of ATPases, the Walker B motif contains the adjacent and conserved glutamate (underlined) identified as the catalytic residue, but in some other classes of ATPases this is absent (Leipe et al., 2003; Orelle et al., 2003). This acidic glutamate is theorised to abstract a proton from water to induce nucleophilic attack on the γ-phosphorous atom, although a comparable contribution to catalysis by the nearby 'switch' loop is conjectured (Geourjon et al., 2001; Zaitseva et al., 2005; Procko et al., 2006). The precise mechanisms for hydrolysis are still ambiguous, but coordination of a Mg^{2+} cation by different residues must also take place for catalysis to occur (Yuan et al., 2001). The above motifs reside collectively within a single NBD, but to constitute an ATPase site, the signature motif (LSGG[Q/E], also known as C-motif) from a partner NBD is required to make contact with the nucleotide γ-phosphate via the serine side-chain (Hopfner et al., 2000; Ye et al., 2004). Thus, NBDs of this sort commonly associate with one another to form active and cooperative domains. The D-loop aspartate, like the signature motif, acts in *trans* across the NBD dimer interface. Its function is not completely understood, but might be the correct positioning of the conserved asparagine in the Walker A motif (see Fig. II-2 for motif conservations), which in turn hydrogen-bonds with the γ-phosphate (De la Rosa and Nelson, 2011). Mutation would thus reduce both ATP binding and hydrolysis. Motifs such as the D-loop and signature motif are means

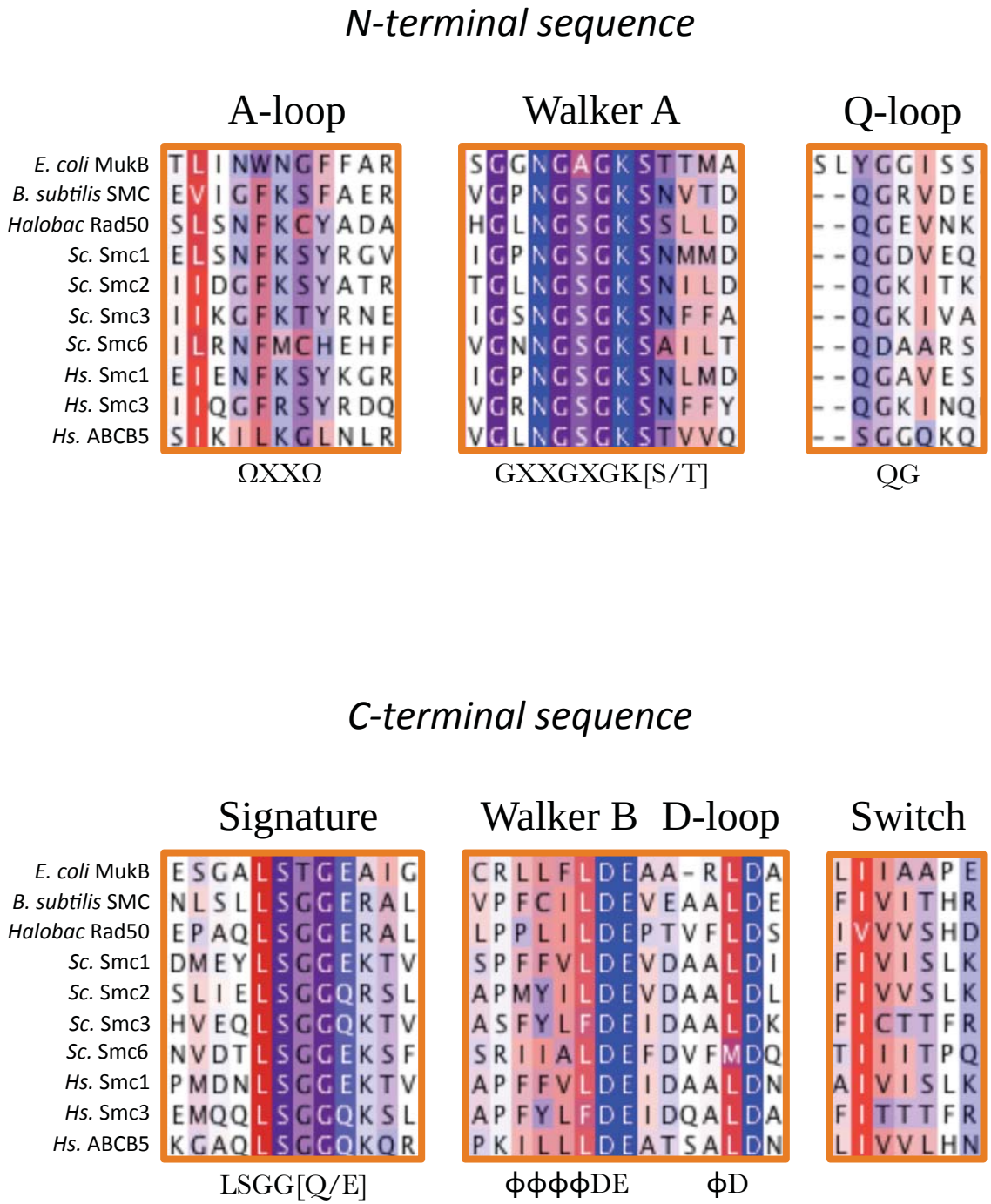


Figure II-2. Sequence alignment of proteins containing ABC ATPase domains.
Across species and complexes – including MukB in *E. coli*, an archaeal Rad50, and an ABC transporter in humans – the critical motifs for binding and hydrolysing ATP are conserved. Multiple sequence alignment was performed using Clustal Omega.
Ω = aromatic residue ϕ = hydrophobic residue.

of communicating across dimer interfaces to produce coordination of hydrolysis between the two active sites (Jones and George, 2011). The Q-loop, on the other hand, might mediate cross-talk between the NBDs and TMDs of ABC transporters; in the case of SMCs, the coiled-coil is perhaps the TMD equivalent (Jones and George, 2002). A high-resolution crystal structure exists of the *S. cerervisiae* Smc1 'head' domain (Smc1hd – the two NBD lobes with abridged coiled-coil and linker), and has been represented to show the orientation of each of the above motifs (Fig II-3).

Walker B mutations, e.g. substitution of glutamate with glutamine, allow the binding of ATP but specifically prevent subsequent hydrolysis. As such, mutations can be used to 'trap' these intermediate states adopted by ATPase proteins, as documented widely in the literature for a variety of these enzymes (Smith et al., 2002; Weibezahn, 2003, Verdon et al., 2003; Hirano and Hirano, 2004). Members of the Nasmyth group exploited these mutations to identify a transitional state in the cohesin loading reaction, whereby hydrolysis-defective complexes can associate with DNA, though not stably (Hu et al., 2011). Whilst wild-type cohesin shows enrichment at the centromere, these mutants accumulate to an even greater degree at this site and fail to associate at other loci. This leads to a model for stepwise cohesin loading, involving interaction of cohesin with its loading factors when the Smc heads are presumably engaged via ATP-dependent association. At this point, cohesin associates unstably with DNA and only becomes more stable after hydrolysis of ATP. Hypothetically, catalysis opens the hinge domains, despite being spatially distant, allowing DNA entry within the ring. Topologically bound (although not cohesive) complexes

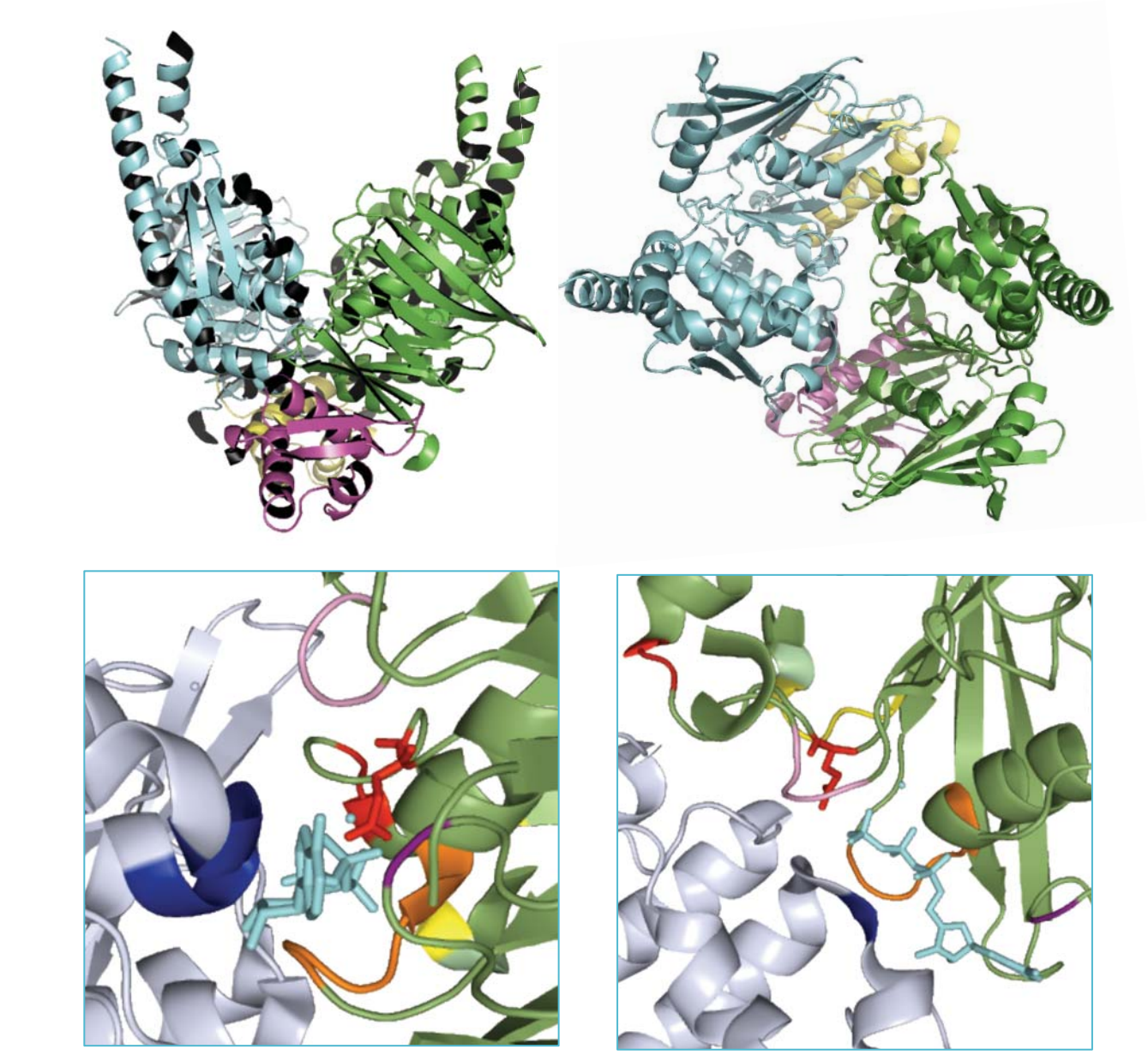


Figure II-3. Crystal structure of the *S. cerevisiae* Smc1 head domain in complex with a C-terminal fragment of Scc1.

The Smc1 NBD homodimerises in the presence of MgATPγS (Haering et al., 2004; 1W1W). The side-on view (left) shows the winged-helix domains of Scc1 bound to Smc1 and the coiled-coils emanating from the NBDs. The top-down view (right) shows the active site and interface formed by two Smc1 heads.

Together, the N- and C-terminal domains of SMC proteins are thought to contribute the Walker A (orange) and B motifs (red, with catalytic glutamate residue highlighted) required for binding and hydrolysis of Mg.ATP (cyan). The ATPase pocket is completed by the ABC signature motif (blue) of a dimerised head domain making contact with the phosphates of the analogue. The D-loop (red, adjacent to the Walker B motif) also seems to form interactions with the partnered head. The switch loop (yellow) helps coordinate the catalysis, while the Q-loop (pink) likely serves to transmit conformational change from hydrolysis to other protein domains. The A-loop (purple) assists in binding the adenosine.

can then translocate to surrounding DNA, potentially sliding along chromatin over large distances driven by RNA polymerase activity.

To date, heterodimeric head engagement has not been confirmed despite the fact that it is thought to occur during the ATP hydrolysis cycle; only homodimers of Smc1 heads have been detected in vitro, and these complexes are able to hydrolyse ATP (Haering et al., 2004; Arumugam et al., 2006). This raises the possibility that ATP hydrolysis is solely carried out by Smc1, with homodimeric NBD engagement occurring between two cohesin complexes. Smc3 may have an accessory or stimulatory role in this process, yet may not actually engage with Smc1.

The ABC ATPases are a large family, most of which are transmembrane transporters of diverse substrates. They are thought to share a common mechanism for effecting transport from the energy of ATP (Linton, 2007; Linton and Higgins, 2007). In this respect, the NBDs act similarly, yet their function depends on the domains with which they are associated. Cohesin's coiled-coil and hinge domains represent a specialised divergence from the transmembrane domains coupled to most ABC transporters. If ATP-driven NBD dimerisation induces an analogous conformational change in these two distinct classes of 'transport' domain, there is a critical difference to consider. Transmembrane domains are by nature short in length and relatively proximal to the NBDs. In contrast, cohesin's hinge is separated from the site of chemomechanical activity by 50nm of coiled-coil. Therefore, transmitting the power from the NBDs is likely to be more complicated. Given the presence of flexible breaks in the lengths of Smc1 and Smc3 coiled-coil, a plausible mechanism for opening the

hinge by the NBDs is the formation of a physical contact between these two domains via folding of the ring structure (Fig. II-4; Gruber et al., 2006). Atomic-force microscopy (AFM) of Smc complexes supports this notion, with some conformations involving apparent head-hinge interactions and significantly bent angles within the coiled-coils (Yoshimura et al., 2002; Sakai et al., 2003). Additionally, overexpression of Smc1hd and Smc1/3 hinge domains revealed a positive co-IP from yeast (Mc Intyre et al., 2007). However, both of these assays suffer from a failure to discriminate against non-specific interactions; overexpression of proteins *in vivo* can produce these and the conditions of AFM are far from physiological. The biochemical relationship between cohesin head and hinge domains is thus revisited.

Given its critical yet unexplored role, it is of great importance to investigate the biochemistry of the cohesin ATPase domains and its relationship to stable loading of the complex onto chromosomes, as well as to identify the requirements for this process in order to be able ultimately to reconstitute the reaction *in vitro*. It was shown herein that specific heterodimerisation does in fact occur between Smc1 and Smc3 and a putative relationship with the hinge domains was addressed biochemically but was found to be lacking. Additional components of the kinetochore potentially involved in such an interaction were screened for yet no candidates were identified. I conclude with a discussion of elucidated cohesin NBD biochemistry in the wider context of ABC transporter function. Finally, potential methodological concerns are considered.

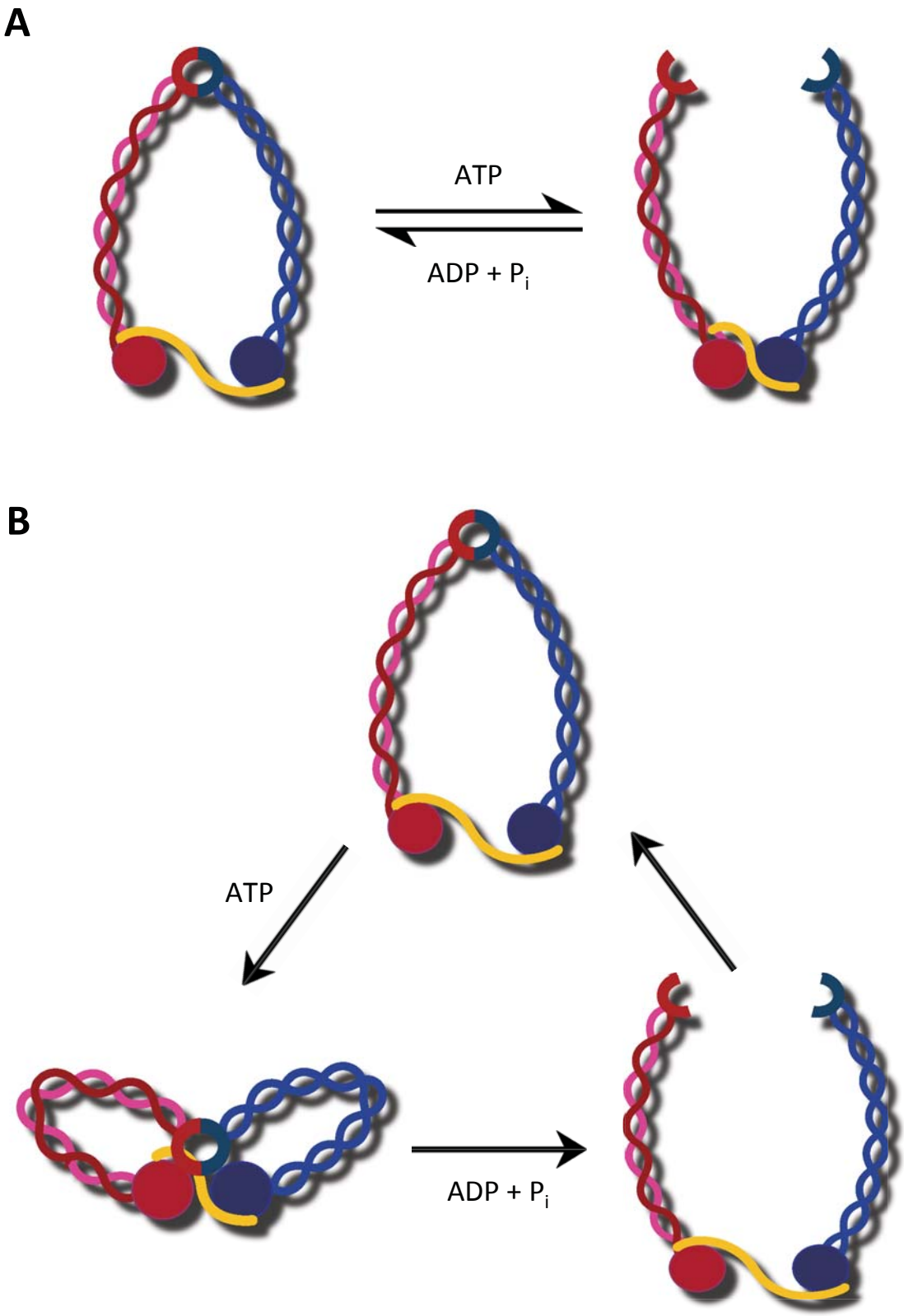


Figure II-4. Potential conformational changes upon ATP binding and hydrolysis
Do the hinge and head domains physically interact as a means of directly transmitting conformational change from ATP hydrolysis to hinge dissociation? (A) Nucleotide binding may dimerise the head domains and open the hinge by conveyance of a conformational change along coiled-coil arms. (B) Alternatively, hinge domains might first interface directly with dimerised NBDs and be opened only upon hydrolysis. Hinge closure is likely to be a passive process.

RESULTS

Heterodimerisation of Smc1 and Smc3 NBDs

To address whether heterodimerisation between the cohesin ATPase cassettes could take place, the NBD domains of Smc1 (co-expressed and purified with a short C-terminal fragment of Scc1 to increase solubility and allow ATP binding) and Smc3 were mutagenised at the Walker B motif to catalytically inactive forms (Smc1hdE1158Q and Smc3hdE1155Q), purified to homogeneity from *E. coli*, and assayed for the ability to form complexes in an ATP-dependent manner via gel-filtration. Gel filtration was chosen as a stringent assay that avoids potential interactions with, and unforeseen steric consequences of, high-affinity binding to matrices.

In the absence of ATP, Smc1hdE1158Q elutes from the gel-filtration column at a retention volume consistent with its size when bound with the C-terminal fragment of Scc1 (~60kDa). The addition of ATP produces an additional peak on the chromatogram, with the elution time corresponding to an approximately 120kDa complex. That this is a homodimer of Smc1hdE1158Q is supported by the earlier elution of protein from the column, as determined by SDS-PAGE (Fig. II-5A). When subjected to the same assay, Smc3hdE1155Q shows no ATP-dependent shift in its retention on the column and elutes primarily as a monomer under both conditions. There is an additional, minor peak at a position in accord with a homodimeric Smc3hdE1155Q complex, the occurrence of which, however, is independent of the presence of ATP (Fig. II-5B). Importantly, engagement appears dependent on the mutation in the

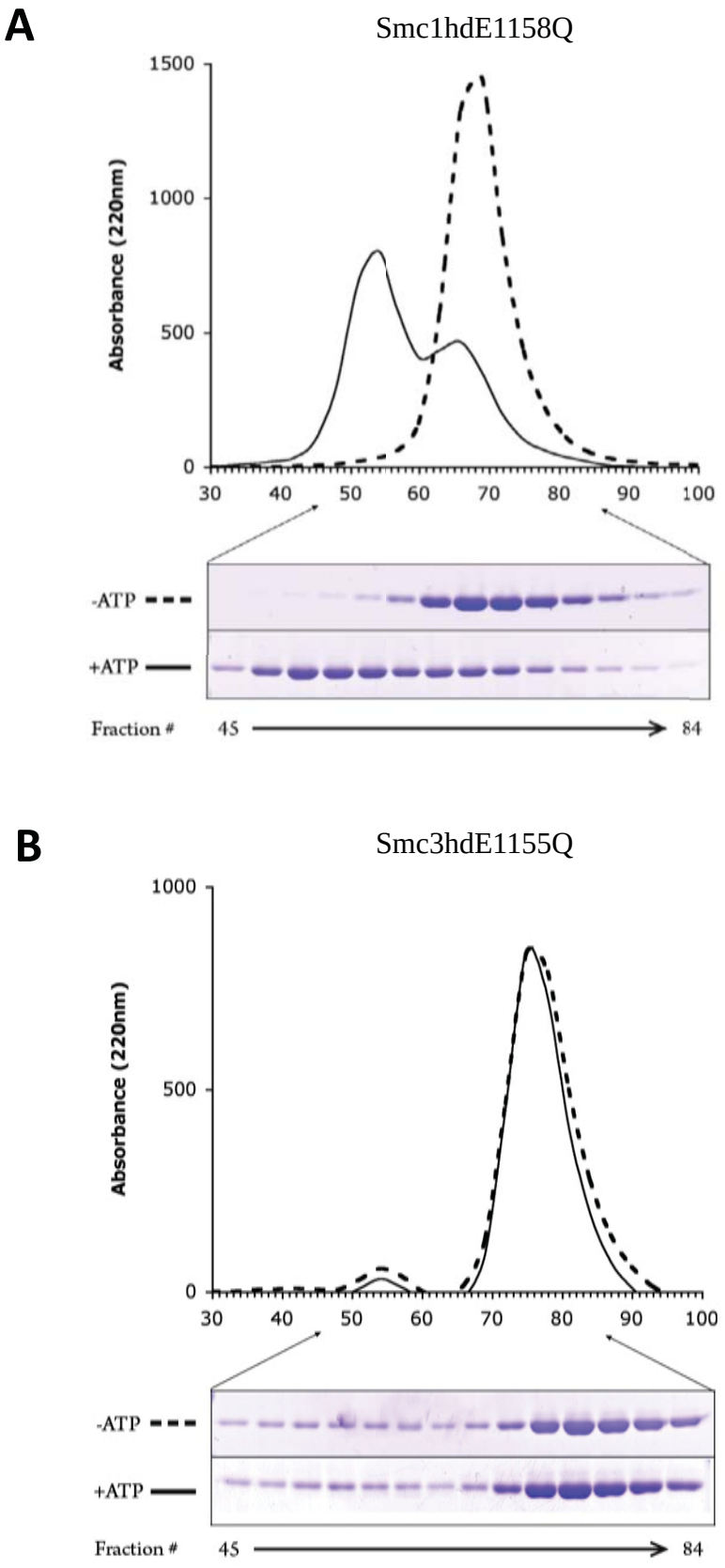


Figure II-5. ATP-dependent homodimerisation of Smc1hdE1158Q

Size-exclusion chromatography of ATP hydrolysis-defective SMC NBDs (3nmol) in the presence or absence of ATP (+/-; 500µM) and visualisation of elution fractions by SDS-PAGE.

(A) Smc1hdE1158Q forms homodimers in an ATP-dependent manner.

(B) Smc3hdE1155Q elutes as a monomer irrespective of the presence of ATP.

Walker B motif that prevents hydrolysis, as isolated wild-type Smc1hd protein is unable to produce a dimeric complex when subjected to the same assay (Fig. II-6).

The notion that only Smc1 participates in ATP binding reactions is supported by these findings. However, Smc1 may show a preference for ATP-dependent heterodimerisation with Smc3, when present, even though Smc3 has no intrinsic ability to homodimerise. Therefore, both Smc1hdE1158Q and Smc3hdE1155Q were mixed together in equimolar amounts in the presence of ATP and their elution profiles from a gel-filtration column measured to assess whether the two proteins can heterodimerise. Both proteins show a change in their elution profiles, indicating the formation of a complex dependent on both the presence of ATP and the Walker B mutation, as wild type protein does not shift in this manner (Fig. II-7A).

Crucially, ATP-dependent complex formation is demonstrably a cooperative process between the Smc1 and Smc3 NBDs, as although Smc1hdE1158Q still displays a shift when wild type Smc3 NBD is present this is presumably due to its ability to homodimerise. The Smc3hdE1155Q protein can only form a high molecular weight complex when Smc1hdE1158Q is present instead of wild type Smc1hd (Fig. II-7B). This suggests that ATP needs to be stably bound and not turning over in the Smc1 active site for the Smc3 NBD to interact via coordination of bound nucleotides.

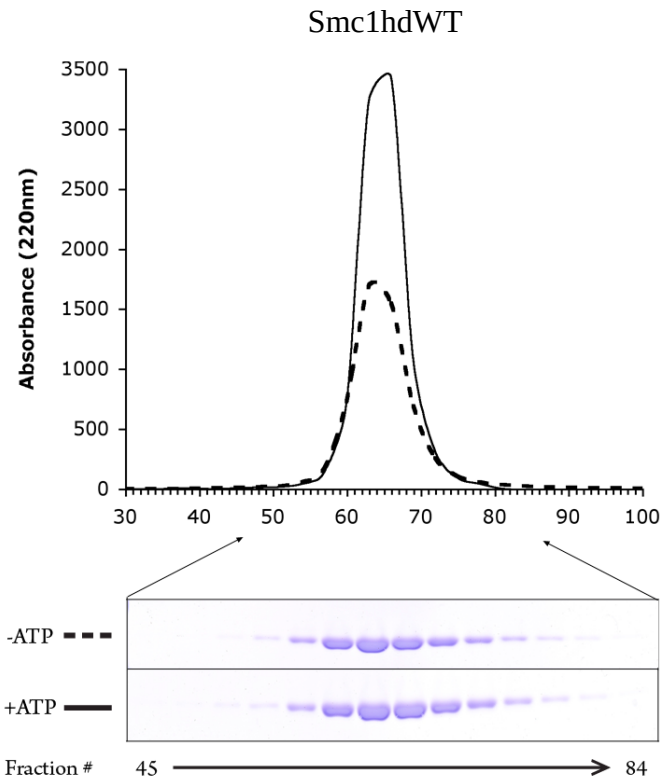


Figure II-6. Wild-type Smc1 NBD cannot stably dimerise
Size-exclusion chromatography of purified wild type Smc1hd (3nmol) +/- ATP (500µM) and visualisation of elution fractions by SDSPAGE. Isolated wild type Smc1 NBD shows no ability to form stable dimers in the presence of ATP.

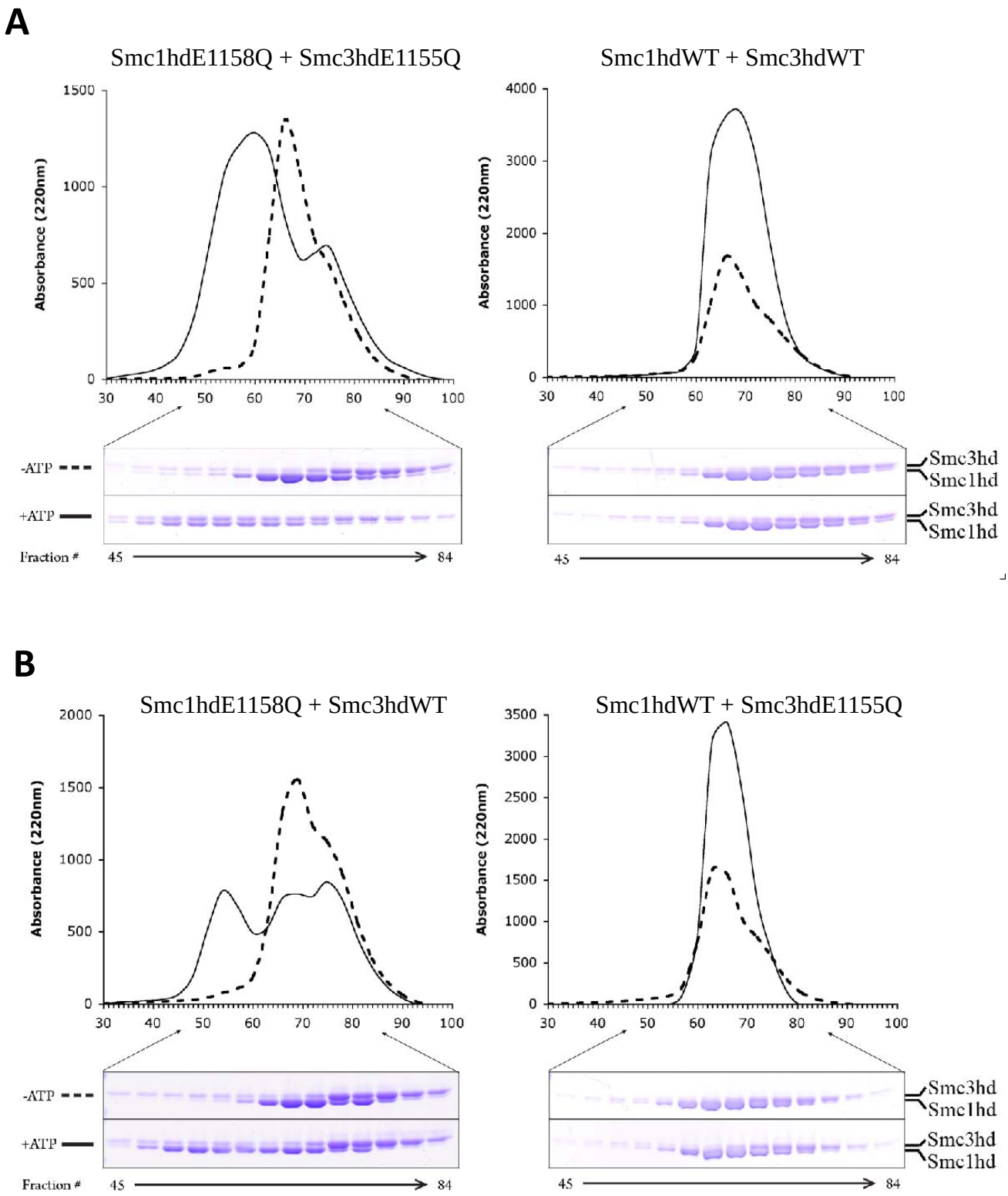


Figure II-7. Heterodimerisation between Smc1 and Smc3 NBDs depends on ATP
Size-exclusion chromatography of Smc NBDs (3nmol) mixed +/- ATP (500µM) and visualisation of elution fractions by SDSPAGE (A) Smc1hdE1158Q and Smc3hdE1155Q shift together in an ATP-dependent manner, whereas wild type Smc1hd and Smc3hd proteins elute as monomers irrespective of the presence of ATP (B) Smc NBDs only co-elute when both are hydrolysis defective, suggesting a cooperative heterodimerisation, as Smc3hdE1155Q does not show a shift when mixed with Smc1hdWT.

Requirements for SMC1/3 NBD dimerisation

Formation of a complex between the two Smc NBDs is seemingly dependent on the presence of both nucleotide and the Walker B motif mutation. These observations are consistent with a model in which ATP bound in the active site of one NBD makes contact with the signature motif of a complimentary NBD. To test this hypothesis, NBDs of Smc1 and Smc3 possessing mutations to both the Walker B and signature motifs were purified from *E. coli* and subjected to the above gel-filtration assay. Mutation of the Smc3hdE1155Q signature motif (Smc3hdE1155Q,S1127R) prevents ATP-dependent heterodimerisation with Smc1hdE1158Q (Fig. II-8A). Mutation of the Smc1hdE1158Q signature motif (Smc1hdE1158Q,S1130R) prevents the formation of heterodimers with Smc3hdE1155Q as well as homodimers (Fig. II-8B). Therefore, the signature motifs of the Smc NBD are sites providing a critical interaction between two head domains. Based on existing evidence from the Smc1hd crystal structure, the signature motif is in contact with the γ-phosphate of the bound ATP. Supporting this notion is the fact that ADP is unable to promote engagement of hydrolysis-defective NBDs (Fig. II-9A). Interestingly, two non-hydrolysable ATP analogues, AMPPNP and ATPγS, were also unable to produce dimers of NBDs (Fig. II-9B), suggesting the chemistry of the γ-phosphate imposes a high degree of steric specificity in supporting an interaction with the signature motif of another NBD.

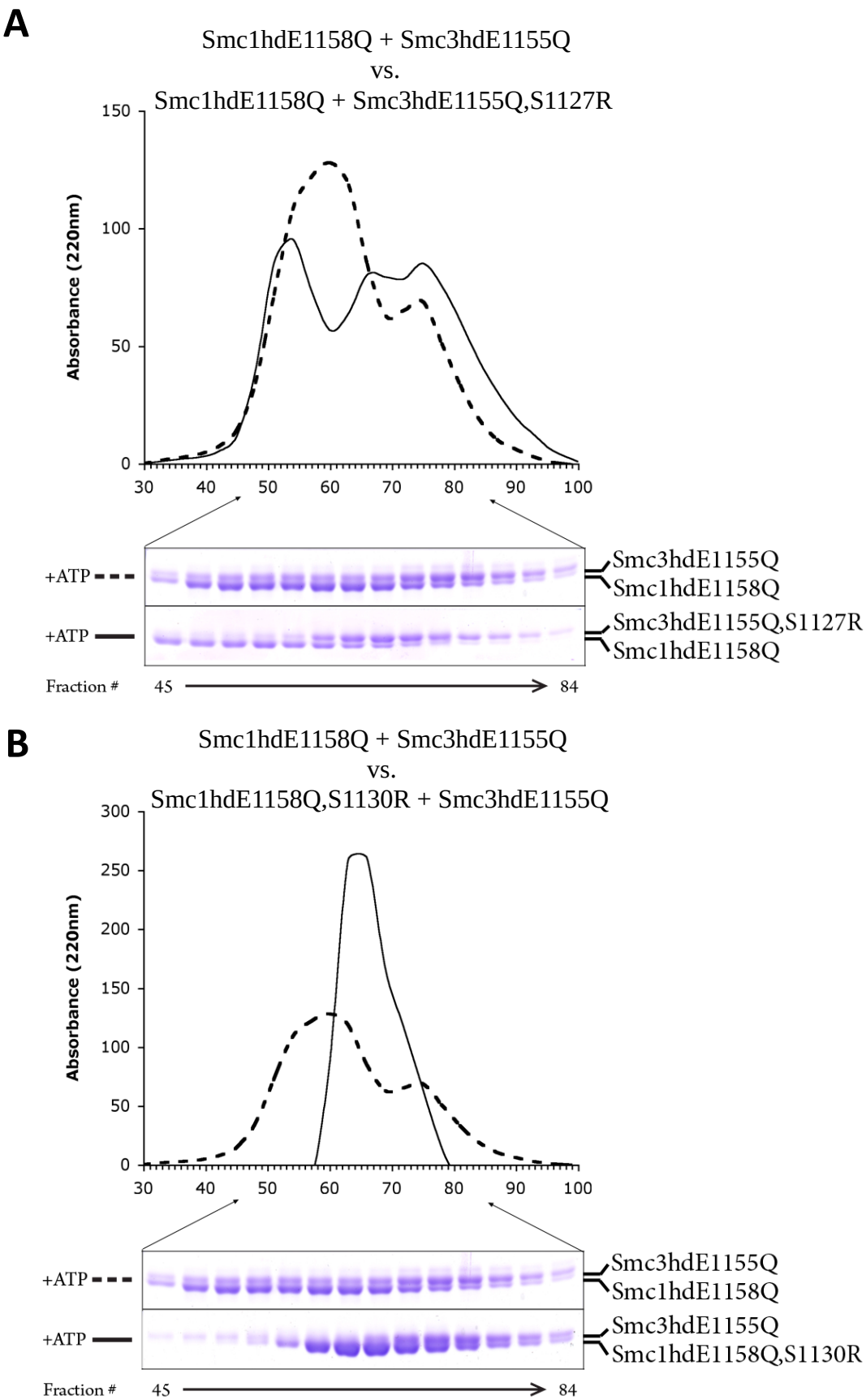


Figure II-8. The signature motif is essential for NBD engagement.
Size-exclusion chromatography of Smc NBDs (3nmol) +/- ATP (500µM) and visualisation of elution fractions by SDSPAGE. (A) Smc1hdE1158Q homodimerizes yet is unable to promote an ATP-dependent shift in Smc3hdE1155Q,S1127R (B) Mutation of the signature motif of Smc1 NBD (Smc1hdE1158Q,S1130R) abrogates both homodimerisation and heterodimerisation with Smc3hdE1155Q.

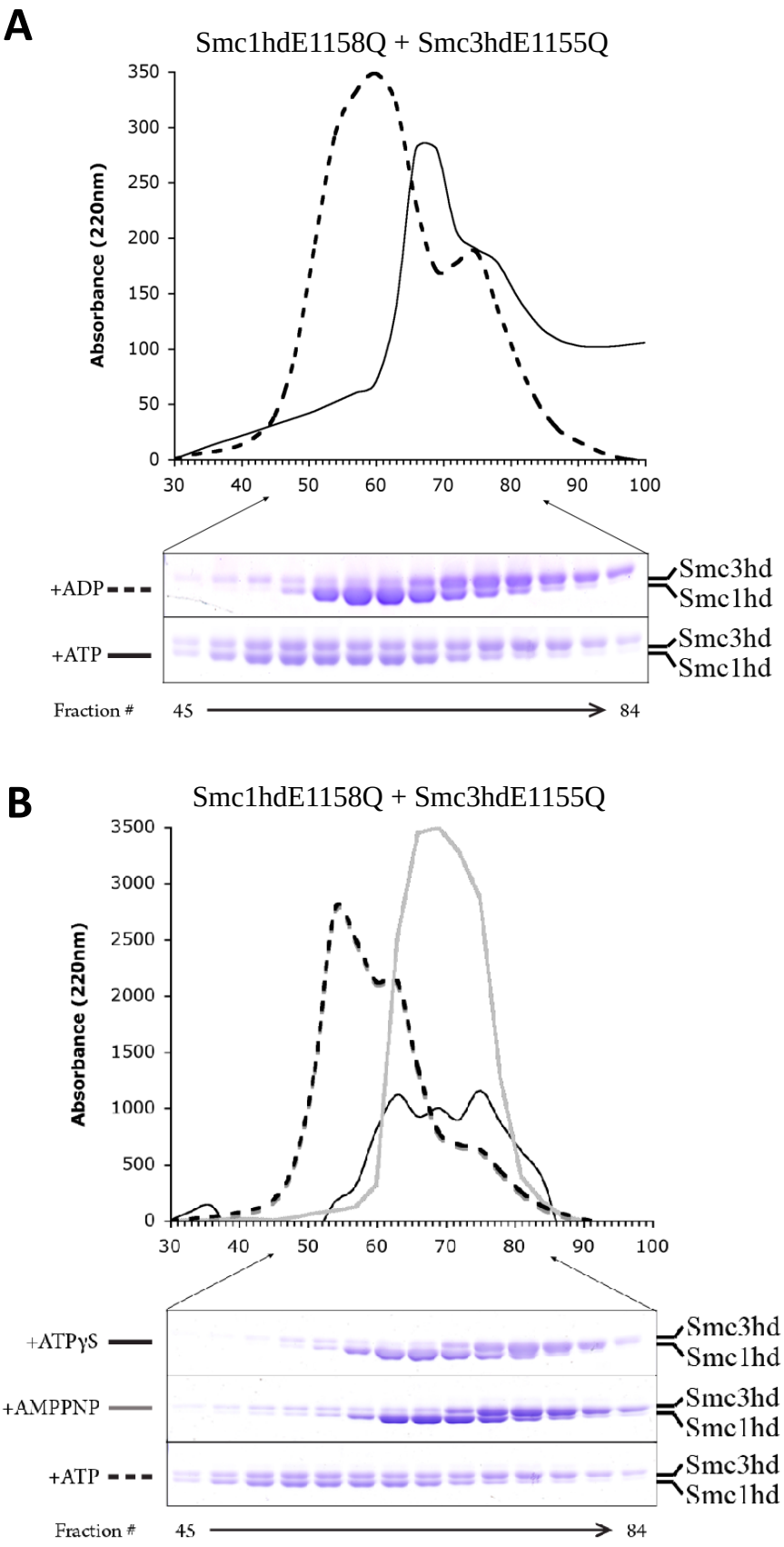


Figure II-9. ADP and ATP analogues cannot support NBD engagement.
Size-exclusion chromatography of ATP hydrolysis-defective SMC NBDs (3nmol) mixed together in the presence of different nucleotides and analogues (500µM). Elution fractions are visualised by SDSPAGE (A) Smc1hdE1158Q and Smc3hdE1155Q heterodimerise in the presence of ATP but not ADP (B) The non-hydrolyzable ATP analogues ATPγS and AMPPNP are unable to support engagement of the heads.

Isolated hinge and head domains do not bind in vitro

Hinge domains of both Smc proteins can be purified to homogeneity from *E. coli*, and exhibit dimerisation *in vitro* when mixed together in equimolar amounts (Fig. II-10A). The binding activity of the hinge complex towards head domains, either in isolation as monomers or as dimers complexed via ATP, was investigated. A shift indicative of a stable interaction between the hinge and NBDs was not observed when the proteins were subjected to gel-filtration, and this was true for both individual wild type NBDs (data not shown) and an ATP-bound NBD heterodimer, when mixed with dimeric hinge protein (Fig. II-10B).

The Smc3 acetylation site does not alter NBD dimerisation

Woo et al. show in the MukB bacterial condensin that ATP-dependent NBD engagement causes dissociation of a subunit domain from the complex (2009). In cohesin, the Scc1 and Smc3 interface is now thought to dissociate and it is believed that the role of Smc3 acetylation in establishment is to prevent further opening (Chan et al., 2012). It was speculated that if this Smc3-Scc1 dissociation occurs analogously to MukB-MukF, then acetylation might serve to prevent this process by inhibiting NBD engagement.

To address this, attempts were initially made to produce a construct of Smc3hd protein with a translation-incorporated, artificial acetyl-lysine by means of an orthogonal tRNA and tRNA synthetase pair (Wang et al., 2006; Neumann et al., 2008). This was approached in collaboration with Dr Susan Hancock at the University of Cambridge in the laboratory of Dr Jason Chin. It was found that use of this system reduced protein expression to an impractical

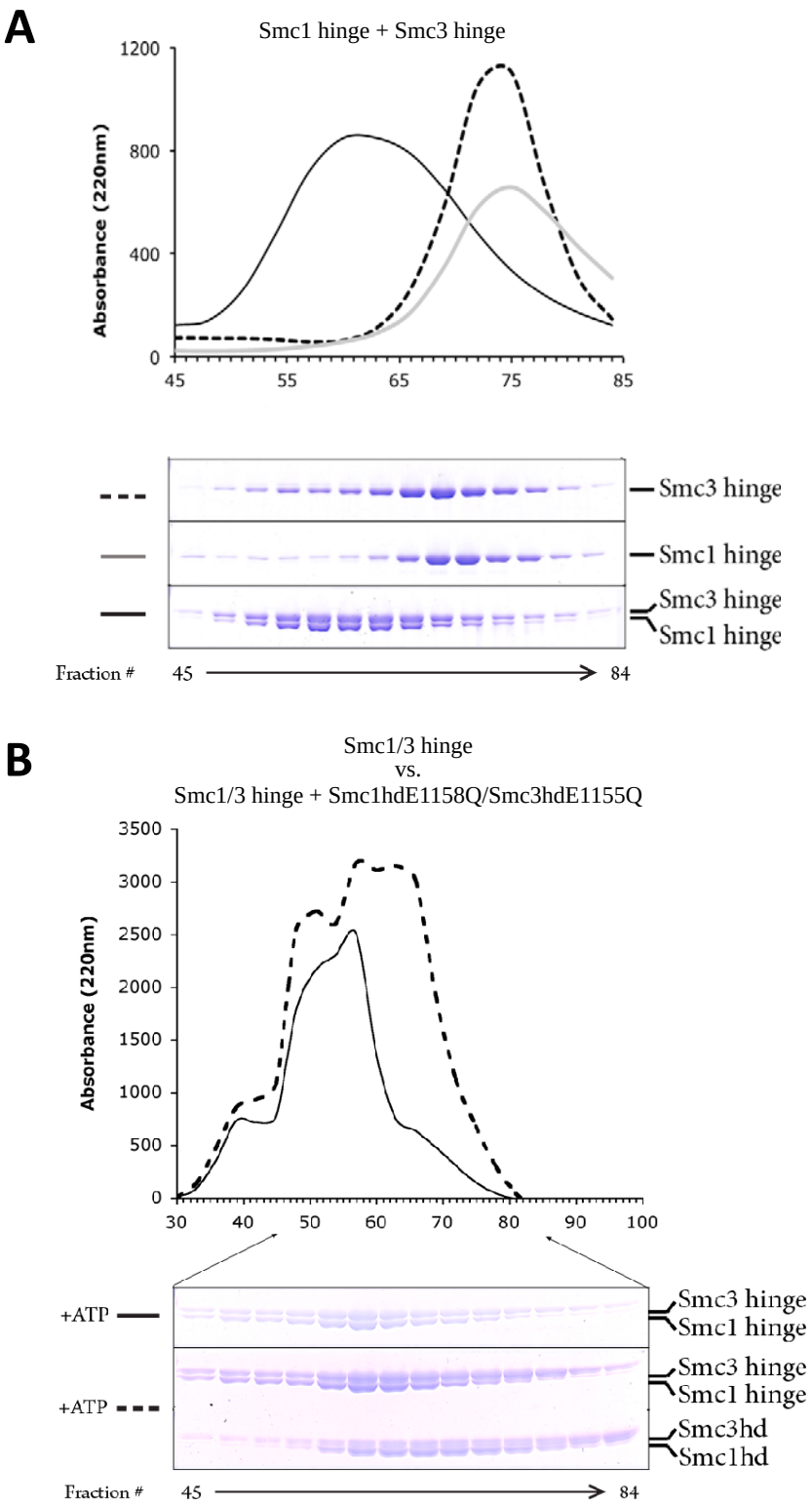


Figure II-10. Dimerised hinge and head domains do not physically interact *in vitro.*
Do the hinge and head domains physically interact as a means of directly transmitting conformational change from ATP hydrolysis to hinge dissociation? (A) Size-exclusion chromatography of purified MBP-Smc1-His_6 and MBP-Smc3-His_6 hinge domains (3nmol) either alone or mixed together, demonstrating heterodimerisation.
(B) Size-exclusion chromatography of the dimerised Smc hinge domains (3nmol) mixed together in the presence or absence of engaged hydrolysis-defective Smc NBDs and visualisation of the elution fractions by SDS-PAGE. ATP (500µM) was present in both reactions. No shift was observed upon combining the two complexes, suggesting no interaction occurs.

level. Thus, the K113Q substitution identified as being sufficient to overcome Eco1 mutation was chosen instead (Unal et al., 2008). The Smc3hdE1155Q,K113Q protein was purified and its ability to heterodimerise with the Smc1hdE1158Q was examined. Gel-filtration analysis showed that in the presence of ATP specifically, this modified head domain could bind durably to its Smc1 NBD partner (Fig II-11).

The stabilised heterodimer complex of Smc1/3hds presented an opportunity to gain structural insight into the Smc3 ATPase domain if crystals were obtainable. Accordingly, Dr Jan Löwe performed numerous high-throughput trials at the University of Cambridge, but no useable crystals matured.

An in vivo assay for detecting kinetochore protein-Scc2 interactions

To mediate hinge opening, cofactors for cohesin loading may be required to elicit conformational change. Thus, using an *in vivo* interaction assay, the requirement for certain kinetochore proteins was investigated. Tandem tet-operator arrays (TetO) bind the tetracycline repressor protein (TetR). The translational fusion of proteins to the TetR domain allows targeting of these constructs to TetO arrays inserted within the budding yeast genome. This provides a system for artificially localising candidate kinetochore proteins to an ectopic locus, i.e. a chromosomal site outside the core kinetochore at the centromere, in this case the *URA3* locus. Fluorescent labelling of Scc2 with GFP allows visualisation of its accumulation at particular foci. Ectopic foci (i.e at the *URA3* locus TetO array) indicative of an interaction with a candidate TetR fusion

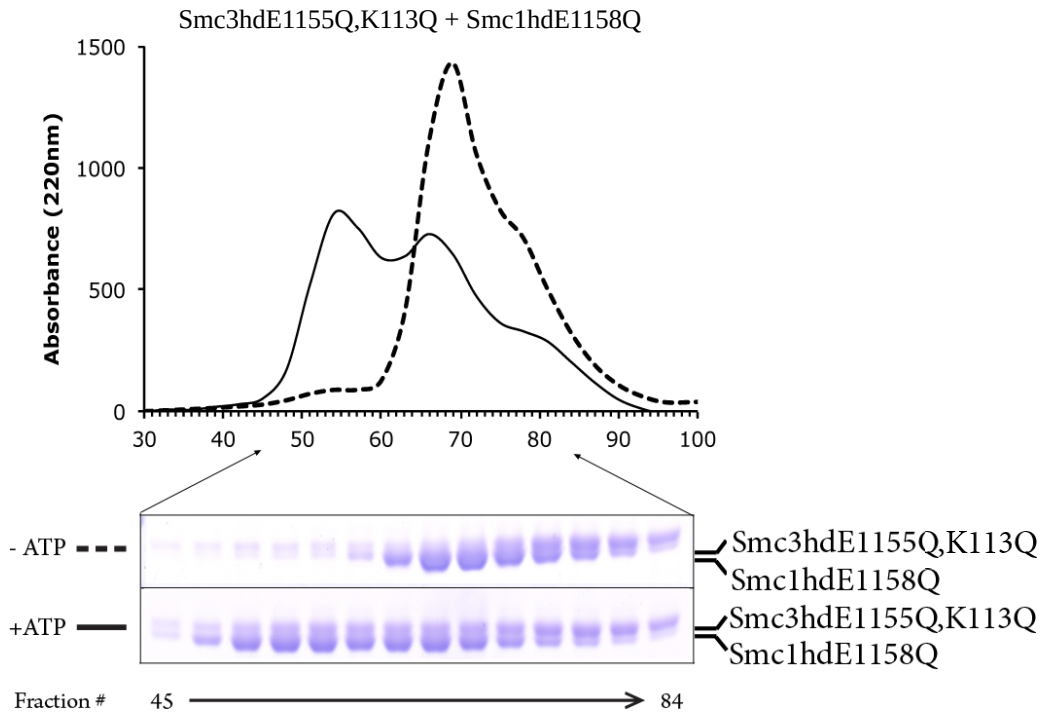


Figure II-11. Acetylation site modification does not disrupt heterodimerisation between ATPase domains.

Size-exclusion chromatography of ATP hydrolysis-defective SMC NBDs (3nmol) +/- ATP (500µM) and visualisation of elution fractions by SDS-PAGE.

Substitution of the principal acetylated lysine of Smc3hd (K113Q) has no effect on the propensity of this ATPase domain to heterodimerise with Smc1hd.

can be distinguished from centromerically-localised Scc2 by marking the centromere with Mtw1-RFP, a fluorescently-tagged, central kinetochore protein.

The ability of this system to detect interactions is validated by the fact that a TetR fusion of Scc2 itself is in turn able to recruit Scc4-GFP to the TetO locus. This manifests as an additional green-fluorescent focus outside of the RFP-marked kinetochores (Fig. II-12). The two centromeric dots represent clusters of 16 chromosomal sister kinetochores separated by amphitelic attachments to microtubules. Scc2/4 by itself is insufficient for loading, as the transition-state complex reproduced by Walker B mutation cannot concentrate at an ectopic site of Scc2/4 enrichment outside of its foci at kinetochores (see Appendix, Fig. A-1).

Kinetochore candidate proteins

A total of 16 kinetochore proteins were C-terminally fused to the TetR domain and were assayed for the capacity to recruit ectopically Scc2-GFP as described above. Haploids harbouring the fusion protein as the only copy of these kinetochore subunits grew healthily for each strain tested, demonstrating that the TetR domain does not interfere with normal function. Diploid cells heterozygous for each marker were used to facilitate imaging.

Of the 16 proteins tested, none showed evidence of ectopic GFP foci. Candidates were chosen on the basis of a combination of pre-existent genetic evidence, *in vivo* data, and hypothetical relationships. Representative images of each TetR fusion strain are shown and the proteins tested are grouped by kinetochore subcomplex (Fig. II-13). Figure II-1 displays the positions of these proteins within the kinetochore superstructure.

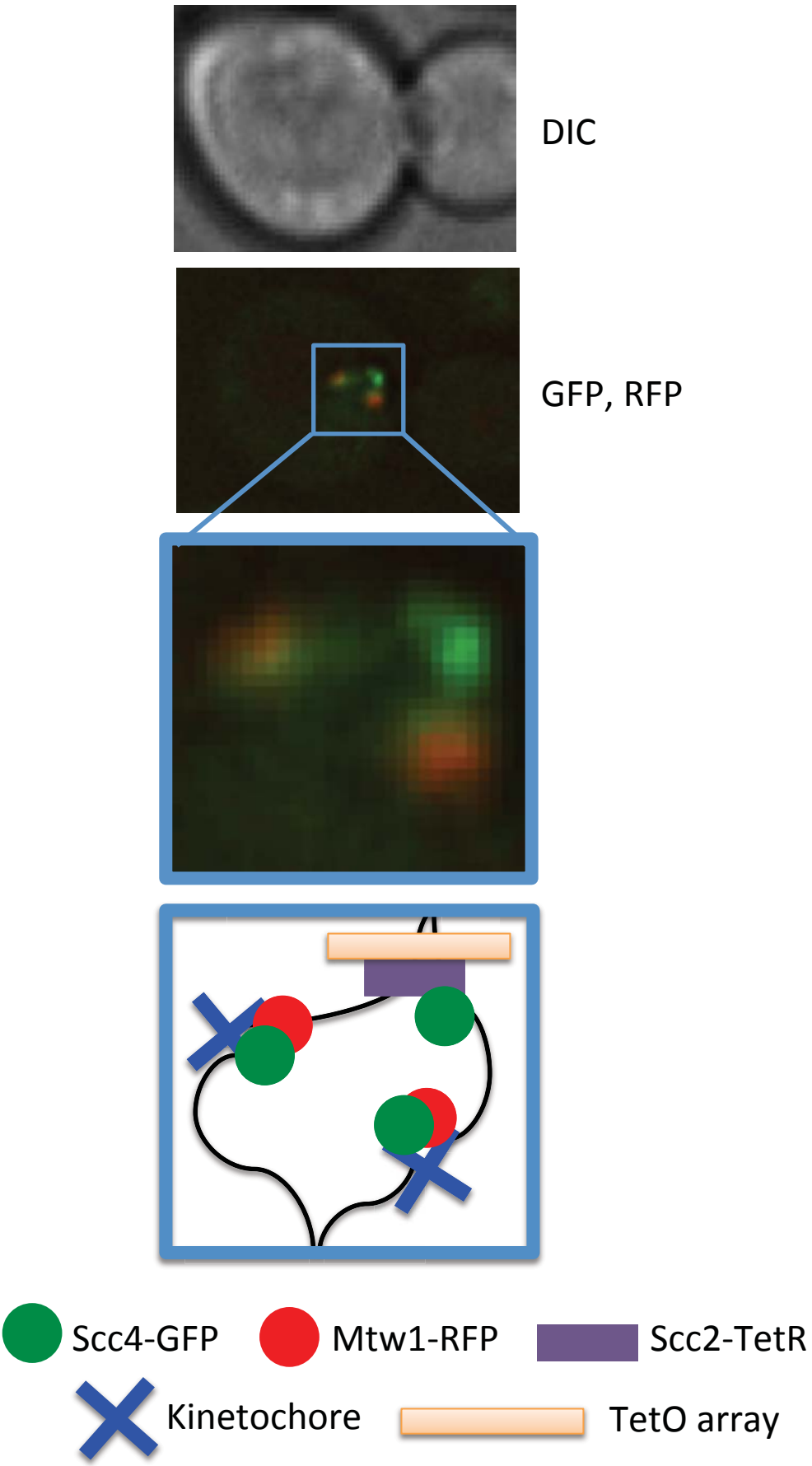


Figure II-12. Scc2-TetR artificially localised to a TetO array can recruit Scc4-GFP.
Live cells were imaged to reveal the separated kinetochores (blue) marked by Mtw1-RFP (red). Scc4-GFP (green) localises centromerically and also to an ectopic site at the TetO array (beige), recruited by the Scc2-TetR fusion (purple) , as represented diagrammatically.

Transcription-related candidate proteins

To test whether the mediator might represent the chromosomal receptor for Scc2/4 and cohesin, two subunits (Cse2 and Srb7) and another Pol II regulating protein (Spt6) were tested via the fluorescence localisation assay. Although there was suggestive evidence that these were good candidate proteins (see *Discussion*), no recruitment of Scc2-GFP was observed (Fig. II-13).

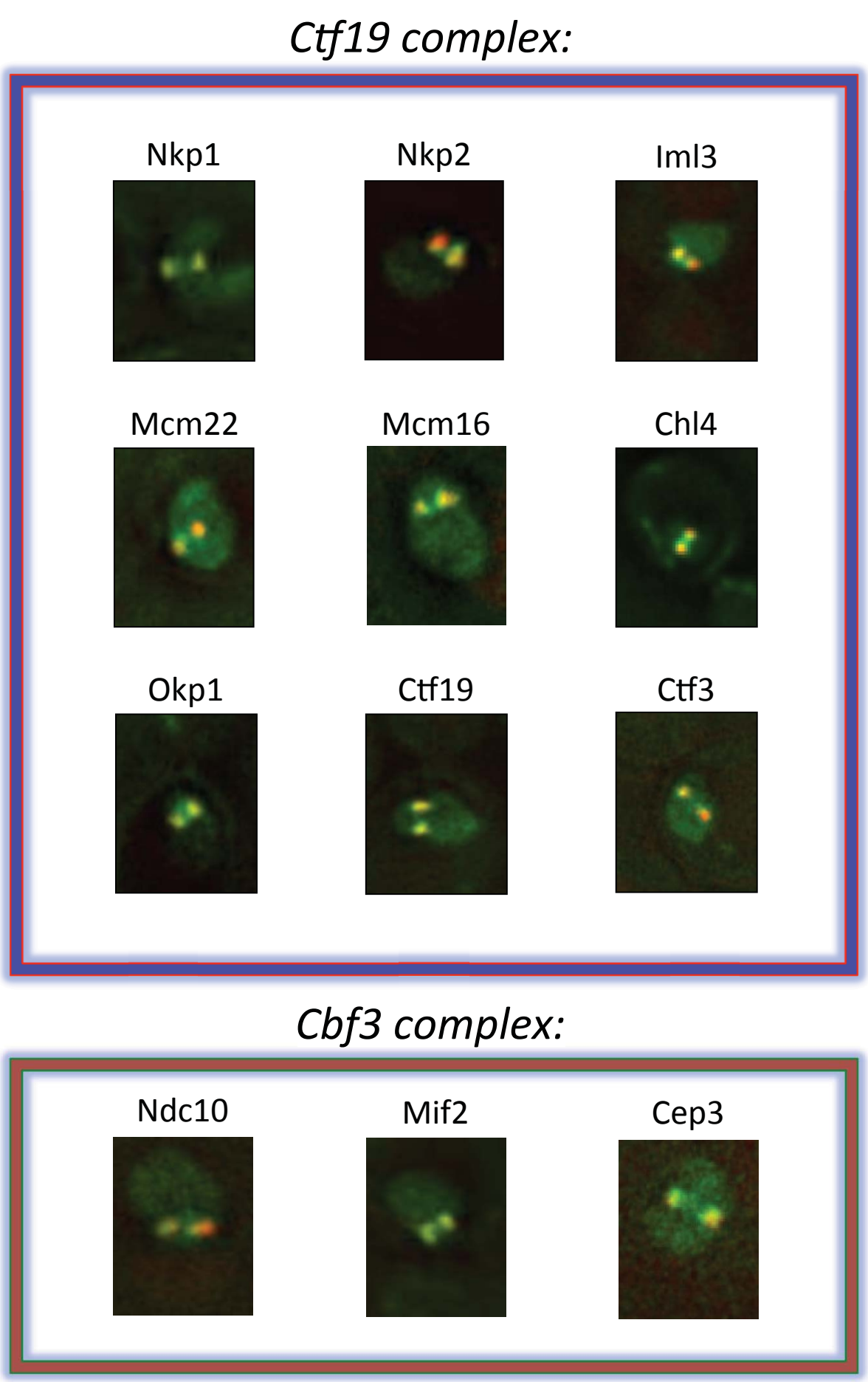


Figure II-13. Analysis of Scc2-GFP ectopic recruitment by candidate-TetR fusion proteins (*continued overleaf).*

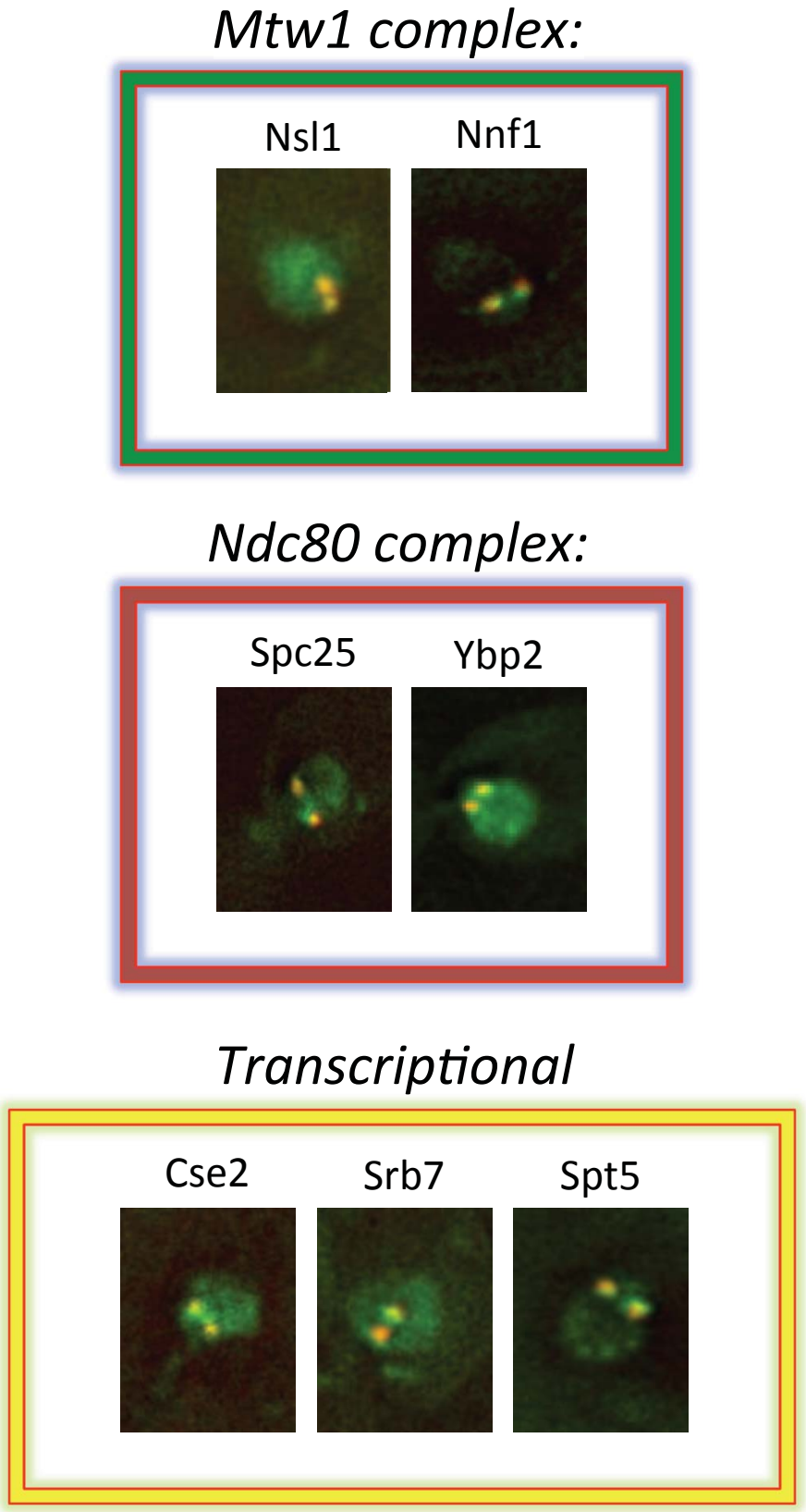


Figure II-13 (continued). Analysis of Scc2-GFP ectopic recruitment by candidate-TetR fusion proteins.
Each image is a merge of GFP and RFP channels and is representative of the foci observed within the population of live, diploid cells examined. Mtw1-RFP was used to mark the centromeric kinetochore. No ectopic foci were observed in any of the fusion strains analysed.

DISCUSSION

ATP-dependent heterodimerisation of Smc nucleotide-binding domains

Although previous studies on the ATPase activity of cohesin suggested NBD engagement occurs between Smc1 and Smc3, this had not been directly demonstrated (Arumugam et al., 2006; Weitzer et al., 2003). Through targeted mutagenesis, purified protein domains could be captured in an intermediate state in the ATPase cycle. Here, the first direct evidence is provided that heterodimerisation does in fact occur between Smc1 and Smc3 heads and that this is dependent upon bound, unhydrolyzed ATP, the γ-phosphate of which is in contact with the signature motif of the apposing head. This configuration is consistent with NBD crystal structures of ABC transporters, SMC family members such as Rad50, bacterial SMC proteins, as well as the yeast Smc1hd homodimer (Hopfner et al., 2000; Smith et al., 2002; Haering et al., 2004; Lammens et al., 2004; Woo et al., 2009; Lammens et al., 2011; Williams et al., 2011).

Differences in structure between the two NBDs may also explain why Smc1hd can homodimerise while Smc3hd cannot. Bacterial SMC complexes are homodimeric and gave rise to heterodimeric complexes by gene duplication in eukaryotes and this has been proposed as an evolutionary explanation for why Smc1 exhibits this behaviour (Haering et al., 2004). It should be noted that although there is a minor amount of Smc3 homodimerisation seen here, it is most likely non-specific as it is independent of ATP. Furthermore, the structure of Smc1 in a state competent in ATP binding appears to be stabilised by the

interaction with the C-terminal winged-helix domain of Scc1 (Arumugam et al., 2003; 2006). Walker A mutants (ATP binding-defective) of Smc1 are unable to bind Scc1, substantiating the notion that the two binding sites affect one another, and this might also be a reason for the differences observed between the two heads.

Whilst it cannot be discounted that Smc1 head homodimerisation occurs between two different cohesin complexes under physiological conditions, there is no evidence to date that this occurs *in vivo.* Also, association of Smc1 and Smc3 via the hinge domain presumably constrains the proximity of heterotypic NBDs, thus favouring intracomplex interactions. Recently published structures of the Rad50-Mre11 complex in both ATPγS-bound and –free states confirm that heads engage within each complex and not between complexes (Lammens et al., 2011).

Structural requirements for dimerisation

For stable binding to occur, it appears both ATP binding sites need to be intact, as mutating the signature motif of either head abolishes heterodimerisation. A decisive role for the signature motif in coordinating ATP is well documented, although this is not without exception (Moncalian et al., 2004). Verdon and colleagues (2003) also report stable heterodimerisation of an ABC ATPase only when a catalytic glutamate is mutated to trap the bound state. However, in this case mutation of a signature motif was tolerable and NBDs heterodimerised just the same. ABC-ATPases frequently harbour a degenerate site within the NBD complex and thus only one active catalytic site

can still permit productive dimer formation. As Smc3hd and Smc1hd each required intact NBDs, neither of the two catalytic domains are likely to be degenerate with respect to ATP binding and dimerisation. Previous studies suggest that full-length Smc3 hydrolyses ATP to the same extent as Smc1, whereas the Smc3hd domain is less active (Arumugam et al., 2006).

Indeed, the necessary integrity of each of the composite sites is highlighted by the fact that the non-hydrolysable analogues ATPγS and AMPPNP were inadequate mimics of regular ATP despite high similarity in their atomic arrangements (Fig. II-14). This also emphasises the need to validate the suitability of these commonly used compounds in experiments of this nature and this is not without precedent; Moody et al. documented the inability of these analogues to substitute for ATP in promoting the dimerisation of an archeael, catalytically-inactive, ATPase cassette (2002). Past experiments with Smc1/3 head domains had also failed to detect heterodimerisation via ATPγS (Haering, pers. comm.). The fact that Smc1hd homodimerised and formed useful crystals in the presence of ATPγS might be attributable to the very high local concentrations of proteins in these conditions, which, when combined with precipitants that act like crowding agents, could allow such associations despite low affinity. In these experiments, Smc1hd homodimerisation by ATPγS was presumably precluded by the stringency of gel-filtration analysis.

Orthovanadate (VO_4^{3-}) serves as a phosphate-group analogue and has been shown to inhibit and trap ATPases in an intermediate state with ADP still bound at the active site following hydrolysis (Urbatsch et al., 1995; Sharma and Davidson, 2000; Chen et al., 2001; Janas et al., 2003). As for the other analogues,

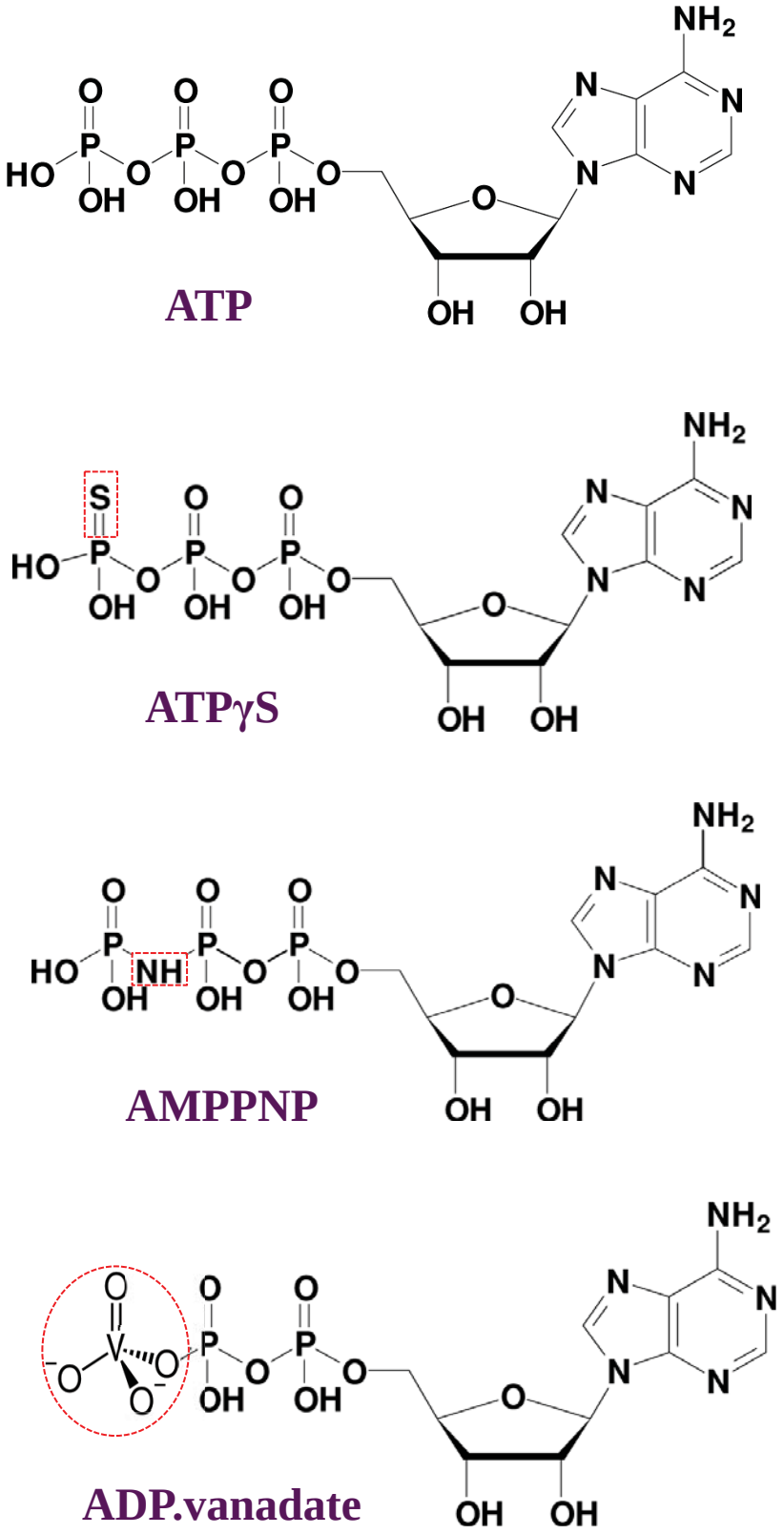


Figure II-14. Chemical structures of nucleotide analogues.
Only ATP could support detectable dimerisation of Smc1 and Smc3 head domains, suggesting even slight changes in the nucleotide structure are intolerable for correct coordination. ATPγS can be slowly hydrolysed, while the ß-γ of AMPPNP cannot be cleaved. Vandate anions adopt the same tetrahedral shape as phosphate and together with ADP can immobilise ATPase active sites.

no such dimerically-trapped state could be detected when vanadate and ATP were mixed with wild-type heads in the assay employed here (data not shown). Beryllium fluoride is an alternative compound isomorphous to phosphate and able to freeze ATPase domains in intermediate conformations, although it was not tested either in crystallisation trials or the above assay (Bigay et al., 1987; Fisher et al., 1995). Likewise, the contributions of the other ABC motifs towards dimerisation were not studied here. Nonetheless, a conformational link between the acetylation modification of Smc3 and NBD function was explored.

As discussed, the close proximity of the Smc3 ATPase site to the lysines acetylated by Eco1 drew speculation of interplay between the two. Additionally, reported dissociation of subunits following NBD engagement is consistent with the known activity of Eco1 in preventing such dissociations (Woo et al., 2009; Chan et al., 2012). If, for example, the ATPase cycle involves opening and closure of the ring at different stages, i.e. engagement and hydrolysis, then inhibition of further cycles could stabilise the ring in a closed conformation. It was shown here that if acetylation does modulate ATPase function, then it is not by inhibiting ATP-dependent dimerisation of NBDs directly. This is in contrast to the findings of Heidinger-Pauli et al., who present genetic evidence that an acetyl-mimicking form of Smc3 can no longer bind ATP (2010b). If that were the case, then NBD dimerisation should be diminished because both active sites require ATP binding, as elucidated herein. However, as determined using direct observation of protein-protein interactions, the Smc3E1155Q,K113Q mutant shows good binding to Smc1hdE1158Q. Thus, acetylation is not likely to exert any potential effect on ATPase activity via a direct modulation of nucleotide

binding as proposed by the authors, although it is feasible that other factors are involved *in vivo*.

Studies of other ABC ATPases show a comparable contribution of ATP-protein and protein-protein contacts to dimer interfaces (Lammens et al., 2004). Nevertheless, high-resolution structural studies of the heterodimer shown here would have been useful, particularly given that the structure of the Smc3 head domain itself is currently unknown and all representations are based on models. Crystallisation trials using ATP-bound Smc NBD Walker B mutant heterodimers failed to grow crystals. This might be due in part to slow turnover of ATP during screens, despite mutation of the glutamate, as residual, low-level hydrolysis is known to occur in other ABC Walker B mutants (Zaitseva et al., 2005). The intricacies of the ATPase domains are interesting not just in regards to how cohesin is regulated, but also in the wider framework of all ABC ATPase proteins and their involvement in myriad diseases, e.g. cystic fibrosis (Efferth, 2001).

The role of ATP in cohesin loading and entrapment

What, then, is the mechanism and function of the cohesin ATPase cycle? Firstly, ATPase experiments have shown a relatively slow rate of hydrolysis, at least *in vitro*, lending support to a 'switch' mechanism of function rather than a continuously required motor activity, for example. As proposed, this switch effects entrapment by opening the cohesin ring, requiring just one cycle of ATP binding and hydrolysis. There are two possible steps for producing work via the NBDs: engagement and hydrolysis. Whether the 'switch' is NBD engagement or

disengagement is still unclear at present. It is seen here that engagement requires ATP bound at both composite sites, as mutation of the signature motif of either head abolished detectable heterodimers. Some models invoke a series of small steps in the ATPase cycle whereby NBDs never dissociate entirely, whilst others entail complete separation (Zoghbi et al., 2012; Aller et al., 2009). Purely on the basis of observations here, these more complicated mechanisms involving multiple conformers produced by hydrolysis first of one ATP then the other seem improbable, since one bound ATP is insufficient to support the dimer. That said, the nuances of NBD binding and hydrolysis cannot be fully addressed here; within the setting of the cohesin holocomplex, with the N-terminus of Scc1 linking NBDs in concert with associated factors, the dynamics of ATP binding and hydrolysis might be very different.

If, as seems likely, DNA enters cohesin via the hinge, how then is this process controlled by the ATPase activity? Although there is pre-existing evidence suggesting a hinge-head interaction does occur, the data remain inconclusive due to the assays employed (Sakai et al., 2003; Mc Intyre et al., 2007). Indeed, when subjected to the more stringent assay used here, such an interaction was not observed. This is perhaps unsurprising, as the cohesin loading reaction is undoubtedly complex and additional factors are likely required. Related protein complexes, such as Rad50, show the potential for large conformational changes upon NBD engagement – undergoing a 60° movement in angle of the coiled-coils that might affect the remote dimerisation domains. Cohesin has flexible regions within the coiled-coil, making a direct correspondence between these two domains unlikely. Also, protease cleavage of

just one strand of the cohesin coiled-coil can be tolerated, further discounting the idea of direct transmittance of conformational change from heads to hinge (Gruber et al., 2003). If the coiled-coil were to conduct force then its rigidity would be paramount and the above finding argues against this model.

The finding that integrity of the hinge is essential for recruitment suggests the hinge is not only closed whilst the NBDs are engaged but also participating in the process more directly, as replacement with generic dimerisation domains (or dimerisation-specific mutations) also abrogates loading (Hu et al., 2011). Therefore, hinge opening may occur in response to ATP hydrolysis, rather than head engagement alone. This is plausible as modelling suggests that the opening of an NBD dimer upon ATP hydrolysis is also a powered event (Newstead et al., 2009). Subsequent hinge closure may simply require passive reassociation.

Kinetochore cofactors in cohesin loading

The apparently important communication between head and hinge domains might be actuated via additional factors. Scc2/4 may recruit cohesin and catalyse the entrapment process, perhaps by stimulating ATP hydrolysis to open the hinge or by inducing a prerequisite conformational change. Nevertheless, Scc2/4 does not act alone during the ATP-bound transitional stage of loading and additional, unknown factors are required for localisation to DNA. Here, attempts to identify these factors using an *in vivo*, fluorescence-based interaction assay failed to yield any positive identifications.

Initially, a potential concern of this assay was that tethering kinetochore subunits to ectopic sites might recruit other kinetochore proteins along with it.

This could foreseeably nucleate the *de novo* assembly of an entire kinetochore capable of attaching to microtubules. This synthetic kinetochore would align with the others by virtue of clustering with the spindle-pole bodies. In essence, this would lead to a merge between the ectopic and centromeric foci and any interactions would be undetectable. There are a number of reasons why this is an unlikely explanation. Foremost, this scenario would produce dicentric chromosomes. Dicentric chromosomes containing two competent kinetochores can form stable, internal amphitelic microtubule attachments. This single chromatid is then torn apart upon attempted segregation at anaphase (Haber et al., 1984). This would lead to a loss of genetic material and growth defects, which were not observed in any of the fusions tested. Because this represents a dominant-negative effect, it could be argued that mutations suppressing this ectopic artificial kinetochore formation might have been inadvertently selected for during isolation of the strain. A mutation causing loss of TetR binding activity would be an instance of such a suppressor. However, it was shown that to create an artificial kinetochore in this manner, the Dam1 and Ask1 proteins are the critical factors that need to be themselves tethered to loci to allow microtubule attachments (Kiermaier et al., 2009; Lacefield et al., 2009). Two research groups concurrently demonstrated that kinetochore proteins could indeed be enlisted to novel loci on minichromosomes with TetR or LacI fusion domains. Even so, only targeting of the direct microtubule binding complex subunits Dam1 or Ask1 were capable of supporting bi-oriented and inheritable minichromosomes (Kiermaier et al., 2009; Lacefield et al., 2009). This indicates that the proteins subjected to the assay presented here were unlikely to build functional ectopic kinetochores.

The most probable explanations for why no protein partners for Scc2 were identified is that the interaction requires: (a) multiple protein-protein contacts with members of one or more subcomplexes, and/or (b) an intrinsic property of complete kinetochore architecture. The kinetochore has been described as assembling in a hierarchical manner (De Wulf, 2003). In other words, the more basal components within the kinetochore are essential for the associations of more peripheral subunits, but not necessarily vice versa. For example, the interdependencies of the kinetochore Ctf19 complex subunits show that the Ctf19 protein is required for recruitment of Chl4, Iml3 and Ctf3, and yet the reciprocal is not true and Ctf19 is still incorporated into the kinetochore in the absence of these same peripheral proteins (Fig. II-15; Pot et al., 2003). However, these binding scaffolds are not entirely unidirectional, as Iml3 appears to reduce the affinity of Chl4 to Ctf19 in immunoprecipitation experiments (Pot et al., 2003). This suggests that at least some multiple protein assemblies are required for robust subunit interactions. This may also explain the fact that ectopic localisation of either Nnf1-TetR or Nsl1-TetR failed to recruit Mtw1-RFP along with it, as only the sister kinetochore foci were observed by microscopy of these strains (Fig. II-13). Thus, multiple interactions between kinetochore subunits appear necessary to stabilise complexes. This is relevant to why Scc2 recruitment was not observed: each candidate protein appears to be localised to the TetO locus independently of other kinetochore proteins, implying Scc2/4 requires interaction with multiple proteins at once for stable binding. Therefore, these interactions would not have been observed outside of the kinetochore.

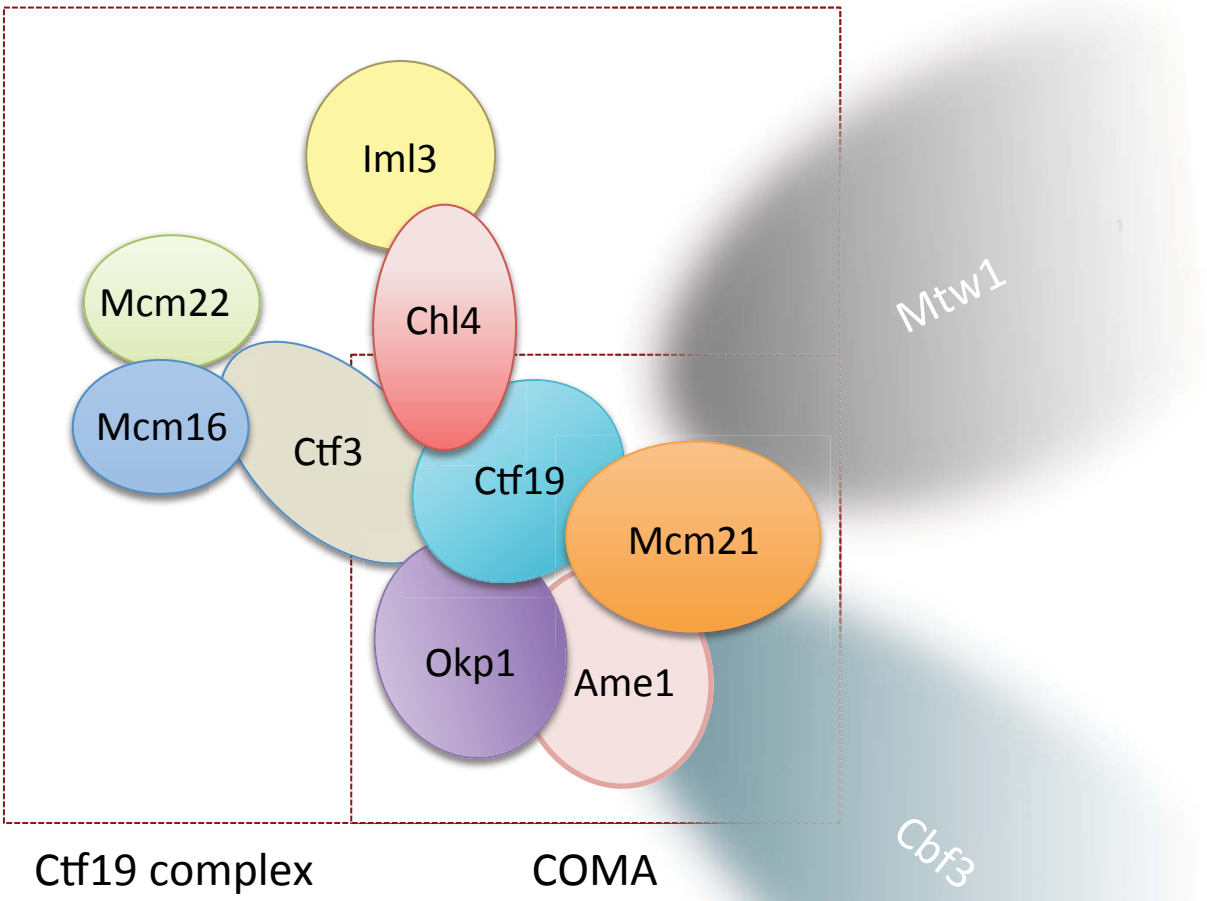


Figure II-15. Inter-relationships of kinetochore proteins of the Ctf19 complex.
The position of each protein corresponds to its dependency upon other subunits for localisation to the kinetochore. Ctf19, Okp1, Mcm21, and Ame1 form the COMA complex, which together with the more peripheral subunits makes up the Ctf19 complex. The COMA complex interacts with Cbf3 and Mtw1 complexes (based on data and representations from Pot et al, 2003; Westermann et al, 2007; and Fernius and Marston, 2009).

Transcriptional cofactors in cohesin loading

The rational for choosing the proteins tested was as follows. Cse2 was first identified as a gene integral for chromosome segregation (Xiao et al., 1993). Subsequent proteomic interaction analyses, such as yeast-two hybrid screens, identified it as associating with both the mediator and elements of the mitotic spindle (Gustafsson et al., 1998; Uetz et al., 2000; Wong et al., 2007). Similarly, Wong et al., found that Srb7 interacts with Scc1 as well as Nkp2, part of the Ctf19 kinetochore complex (2007). Both the mediator and the Ctf19 complex have been implicated in cohesin loading, as discussed. Genetic interactions between genes implicated in transcription, such as mediator subunits, and those conducting mitotic events could merely represent downstream effects of aberrant changes in gene expression. Nevertheless, they may represent the overlap seen between loading of cohesin at arms and kinetochores; a direct physical interaction would indicate a more definite and direct role. No such interaction could be detected here between Scc2 and these proteins. Spt5 is a protein implicated in regulating transcriptional elongation via interactions with Pol II and structuring chromosomes (Andrulis et al., 2000; Bortvin and Winston, 1996). It has also been found to co-IP Smc3 and Smc2, however the sensitivity of mass spectrometry in identifying these interactions lends itself to false positive, non-specific hits (Lindstrom et al., 2003). Here, colocalisation with Scc2-GFP was not seen.

Collectively these results demonstrate that the role of unidentified cofactors in the Scc2/4 cohesin loading reaction is less straightforward than at first anticipated; it may involve the integrity of multiple protein-protein

interactions to build a platform for the binding of Scc2/4 and the cohesin complex. Alternatively, the possibility exists that changes in underlying DNA structure elicited by the kinetochore are what allow cohesin to load to chromatin. Be that as it may, new, peripheral kinetochore proteins are being discovered that affect cohesion. One such protein is Aft1, a transcription factor that associates to the kinetochore via Iml3 and is required for cohesion (Hamza and Baetz, 2011). Together with the characterisation of the Smc1 and Smc3 NBD interactions shown here, the prospective identification of the specific proteins required for cohesin loading will ultimately allow insight into the biochemical mechanisms of this reaction *in vitro*.

CHAPTER III

Applying *in vivo* protein cross-linking to the study of cohesin complex dynamics

Applying *in vivo* protein cross-linking to the study of cohesin complex dynamics

INTRODUCTION

The multi-gated girdle

As a large, ternary ring entrapping sister DNA fibres, the cohesin complex poses an exceptional challenge to description of its mechanism. In one state, the complex exists preassembled and complete in the cytosol (Gruber et al., 2003). In the other, functional state, this same annular structure topologically encircles sister chromatids to provide cohesion (Haering et al., 2008). To transition between these two states requires the passage of 'rope through a ring'. How is such a feat possible? In yeast, the removal of cohesin at anaphase occurs by proteolysis within Scc1, equivalent simply to cutting the ring and allowing escape of chromatids towards opposing spindle poles (Uhlmann et al., 1999). However, a certain population of cohesin cycles between association and dissociation from DNA. Additionally, the initial entrapment of DNA by cohesin cannot be protease-dependent as this is an irreversible operation. The ring must instead open at one of its three protein-protein interfaces, translocate DNA within, and then reclose said interface to prevent escape (Fig. III-1). Such an exercise undoubtedly requires intricate and strict regulation. It must happen at the right place, i.e. upon chromatin, and at the right time, normally preceding DNA replication.

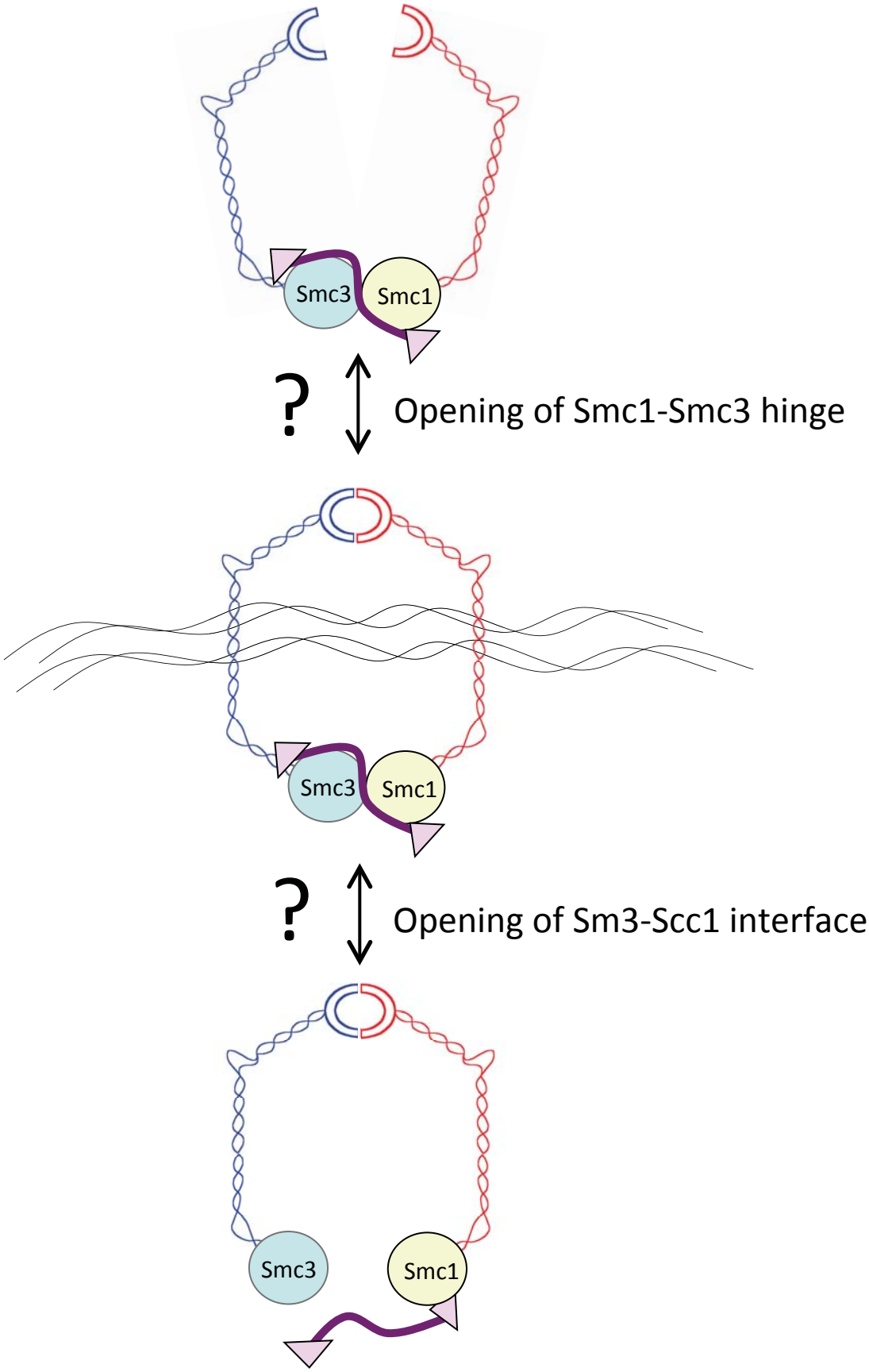


Figure III-I. Might cohesin associate and dissociate from DNA through the disengagement of distinct subunit interfaces?
Available evidence suggests that the cohesin complex might open the Smc1-Smc3 hinge domain to entrap DNA and open the Smc3-Scc1 interface to dissociate from DNA (Gruber et al., 2006; Chan et al., 2012; reviewed in Nasmyth, 2011).

It was seen in the previous chapter that demonstrable engagement of the Smc1 and Smc3 NBDs was a critical requirement for the localisation of the cohesin complex to sites of loading. This intermediary population has a short residence time at the kinetochore foci. For a more stable and broader distribution across the pericentromere, ATP hydrolysis must occur and this is alleged to require hinge dissociation leading to entrapment of DNA. The resultant topological association is stable enough to allow the sliding of cohesin along kilobase distances of chromatin. The combined actions of Scc2/4, an unidentified co-factor, head engagement (Chapter II) and ATP hydrolysis appear imperative for complete loading, yet the actual process is unclear.

Collectively, there is good evidence to indicate that normal hinge domain function is paramount for loading and topological embrace; it is widely considered the entrance gate for DNA (Nasmyth, 2011). The most convincing demonstration of this assertion is that simultaneously locking the other two cohesin ring interfaces permits growth. That the hinge is likely the entrance gate was also addressed more directly by inducing closure of the Smc1-Smc3 interface using rapamycin-lockable protein domains. Locking complexes at this site in G_2/M after loading had already occurred had no detrimental effect on cohesion. In contrast, addition of rapamycin prior to loading inhibited association with chromatin and produced cohesion defects. If cohesin does indeed topologically encircle chromosomes, then the above results cogently point to ring opening at the hinge domain as a prerequisite event.

Disruption of an interface of the cohesin ring is thought to be a mechanism common to both the association and dissociation of cohesin from DNA. The

latter occurs non-proteolytically for unacetylated populations of cohesin; non-acetylated cohesin remains dynamic and turns over throughout G_2/M phase, whereas acetylated cohesin is stable. The details of this process are reviewed in Chapter I. In short, release of unstable cohesin from the DNA is caused by Wapl and the function of acetylation of Smc3 at K112,113 is surmised to counteract its releasing activity. Evidence now indicates that the release from DNA in fact occurs by disruption not of the Smc1-Smc3 hinge domains, but rather the interface between Scc1 and Smc3 (Chan et al., 2012). This process is also thought to be at the heart of the separase-independent removal of cohesin from chromatid arms during the prophase pathway that exists in metazoa.

The prophase pathway

In the prophase pathway cohesin is removed from along the lengths of chromosome arms, leaving a subpopulation intact at the centromere to produce the central constriction. As outlined, this process is distinct from the type of release seen with dynamically-bound, G_1-phase cohesin and there are a number of key differences. For one, it occurs in an orchestrated manner. This is not achieved by the same synchronous release of cohesin witnessed in anaphase either, as separase cleavage of the ring does not occur in prophase (Waizenegger et al., 2000). Secondly, both acetylated and non-acetylated cohesin are removed, as evidenced by the loss of cohesion at arms (sister chromatid resolution).

The non-mitotic removal of cohesin and the prophase pathway are alike, however, in that they both employ Wapl to stimulate release (Gandhi et al.,

2006; Kueng et al., 2006; Chan et al., 2012). Phosphorylation of cohesin subunits, principally SA2 (vertebrate equivalent of Scc3), by PLK1 and Aurora B is required during the prophase pathway (Hauf et al., 2005). It is likely that this modification converts the normally refractory population of acetylated cohesin to one susceptible to Wapl activity. There is very little known about the effects of SA2 phosphorylation. In yeast, Scc3 inactivation in G_2/M leads to prematurely split sister chromatids, signifying an important but largely uncharacterised role in maintaining cohesin's coherence (Tóth et al., 1999). Phosphorylation of SA2 may increase the activity of Wapl in its ability to disrupt Smc3-Scc1 binding, or it may compromise cohesin integrity in some other way that involves dissociation of alternative or multiple interfaces.

Methods for closing an interface

Despite the regulatory complexities of the prophase pathway, the end result must be simple: dissociation of one or more of the interfaces, for instance between Smc3 and the kleisin subunit, to bring about non-proteolytic release of cohesin from chromosome arms. A method for the conditional closure of one or more of the cohesin ring interfaces will allow investigation of the regulatory role of each in determining cohesin-DNA association. Ideally, cohesin would be allowed to load onto DNA normally, i.e. with all three interfaces unaltered such that its mode of association can be expected to be as close to physiologically normal as possible. For example, the Smc3-Scc1 fusion able to prevent Wapl-mediated release of the complex also causes a growth defect in these cells. The amount of chromatin-bound complex is reduced, implying aberrant loading of this fusion construct. Results must be interpreted with this in mind. To avoid

such a complication, closing the target interface only after the preceding steps have occurred is optimal. Achieving this, however, presents a serious technical challenge.

Such a method must meet several criteria: principally, being inducible and rapidly acting. Previous *in vivo* experiments with the cohesin complex have employed a system utilising the dimerisation of the rapamycin binding domains, FRB and FKBP12 (Gruber et al., 2006). There are several potential drawbacks of this system worth mentioning. The FRB and FKBP12 domains are large and thus have the potential to interfere with the natural protein-protein interactions that may occur proximal to the interface in which the domains are inserted. While this may not have been a relevant concern at the hinge interface, if used at unstudied sites it is plausible that phenotypes result from effects downstream of an aberrant loading step, for example, and not from the creation of a physical barrier against the entry or exit of DNA. Small and minimally disruptive insertions/modifications are optimal. Rapamycin treatment also necessitates a mutant background strain that lacks the genes involved in response to this drug. Such mutations can cause defects in normal metabolism, particularly in multicellular organisms, and this is an undesirable complication. Finally, it is also impracticable to observe successful closure directly, as the interaction is lost upon denaturation and hence not seen via SDS-PAGE.

A recently developed technique for covalent bond formation between two protein groups exploits the properties of an extracellular adhesion domain from the pathogenic bacterium *Streptococcus pyogenes* (Zakeri et al., 2012). The CnaB2 domain of the *S. pyogenes* protein FbaB2 contains a spontaneously

formed isopeptide bond, whereby an intramolecular covalent bond is formed between the side chains of aspartate and lysine in the presence of a catalytic glutamate (Fig. III-2; Zakeri et al., 2012).

In its natural context, this isopeptide bond confers remarkable stability to the protein, allowing it to remain folded at pH 2 and up to a temperature of 100°C (Hagan et al., 2010). Zakeri and colleagues split the protein into two parts comprised of a small peptide of 13 amino acids, SpyTag, (containing the aspartate) and the larger structural domain, SpyCatcher (containing the glutamate and reactive lysine). Following rational engineering, the two components can combine to reconstitute the reactive conformation, rapidly producing the covalent, isopeptide bond. It is theorised that this technique can be applied to cross-link two proteins to one another, as occurs at the interfaces of the cohesin ring. By inserting a SpyTag within each of the target proteins, they can be covalently cross-linked to one another upon reaction with two SpyCatcher domains joined by a linker (Tandem SpyCatcher; Fig. III-3).

The SpyCatcher system is attractive for several reasons. The reaction occurs rapidly, at least *in vitro*, and produces an irreversible isopeptide bond that is of high strength due to its covalence. Consequently, cross-linking is preserved and observable following denaturation for SDS-PAGE. The SpyTag unit is small and so insertions within proteins are more likely to be tolerated compared to other larger binding domains, especially before cross-linking. Another advantage of this system is that the cross-linker itself is a protein, rather than a chemical that would typically be more toxic to live cells. Furthermore, being genetically encodable allows SpyCatcher to be placed under

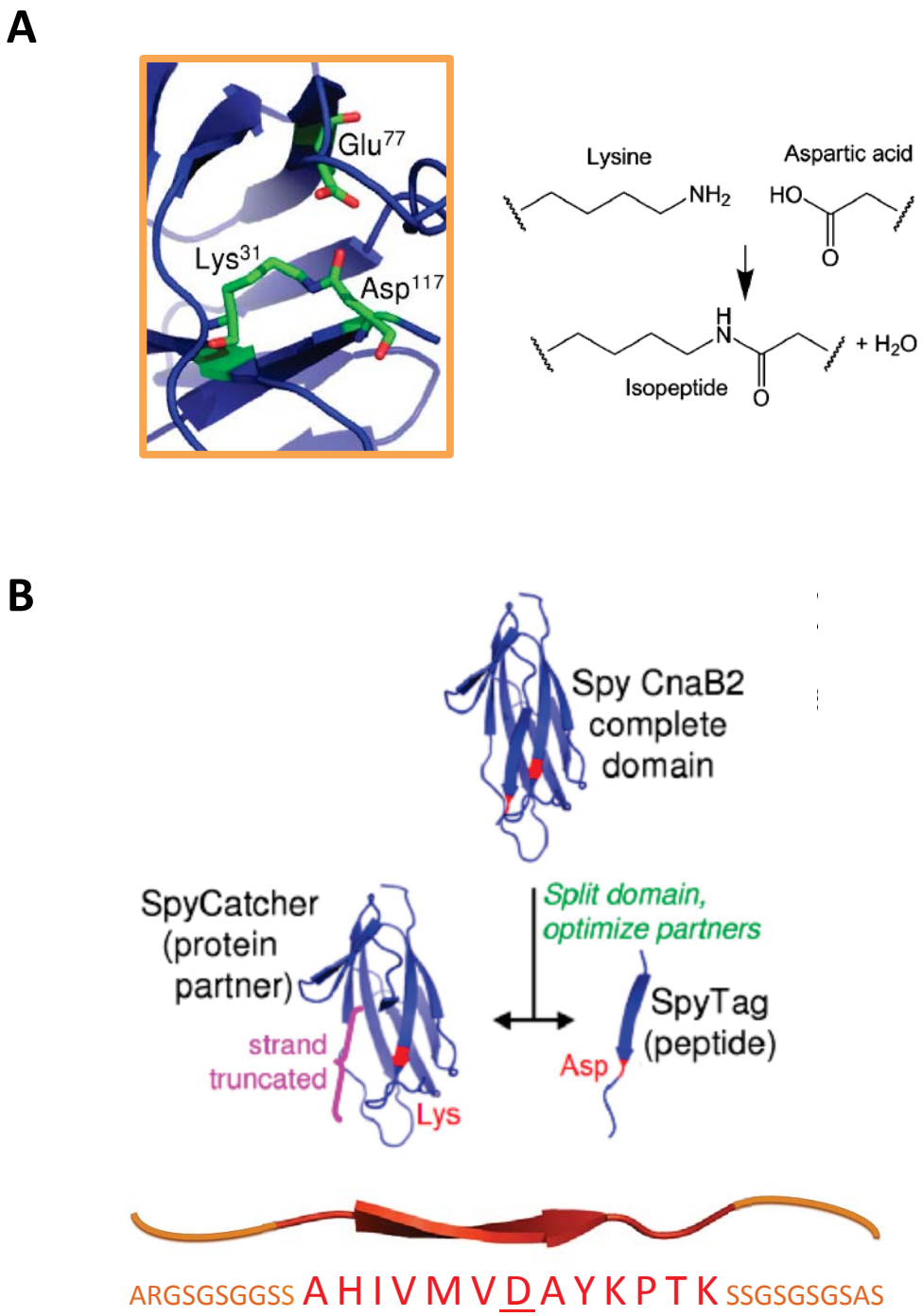


Figure III-2. Construction of the SpyTag-SpyCatcher system by Zakeri et al. (2012).
(A) Isopeptide bond formation between spatially adjacent lysine and aspartate residues within the *S. pyogenes* collagen-adhesin domain (CnaB2) of FbaB. The proximal gluatamate catalyses the reaction.
(B) Howarth and cowokers divided CnaB2 into a short peptide able to reconstitute with the larger domain the catalytic arrangement for covalent bond formation. After rational engineering, the most reactive peptide was 13 amino acids in length.

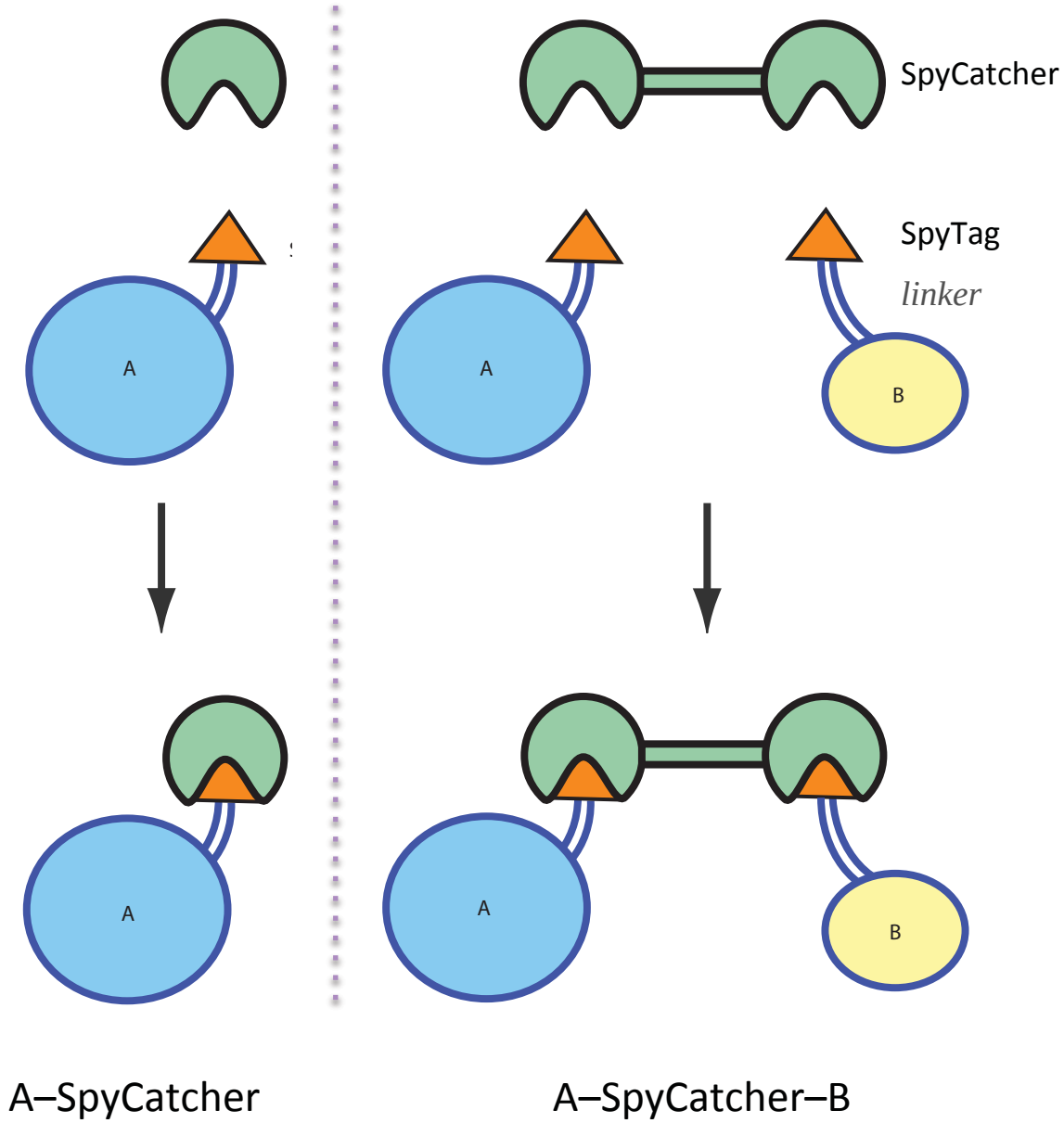


Figure III-3. Schematic of Tandem SpyCatcher system for covalently cross-linking two proteins (A & B).
SpyTag can be inserted internally within a protein domain (A) and can reconstitute with SpyCatcher to form a covalent bond, producing a fusion protein (A-SpyCatcher) visible by denaturing SDS-PAGE (left). Combining this with an insertion of an identical SpyTag sequence within a different protein (B) and expressing two SpyCatcher domains in tandem (Tandem SpyCatcher), should irreversibly conjoin the two if they are naturally proximate.

a promoter of choice, for example *GAL1p*, and its expression induced and controlled as required by the addition of galactose.

Validation of the SpyTag-SpyCatcher system within budding yeast was attempted here to yield a valuable tool for studying the regulation of cohesin's association with DNA. In particular, such an approach would have been useful to effectively address the importance of Smc3-Scc1 interface opening in model metazoan systems such as *Xenopus* and *Drosophila*. Despite showing the prerequisite properties for suitability, the SpyTag-SpyCatcher technique proved inadequate when implemented to close the model Smc1-Smc3 hinge interface.

RESULTS

SpyTag insertion and purification of Smc hinge domains

The Smc1 and Smc3 hinge domains were chosen as a model interface at which to investigate the feasibility of using the SpyTag-SpyCatcher system for protein-protein cross-linking. The structures of the *M. musculus* and *Thermotoga maritima* Smc1 and Smc3 hinge domain heterodimers have been determined previously and they are predicted to be highly similar to the yeast proteins based on sequence homology and modelling (Fig. III-4A). Loops have been identified within which to introduce SpyTags with minimal disruption of tertiary structure, whilst still being spatially proximal. The residues chosen on each of Smc1 (L597) and Smc3 (S606) are the same as those used by Gruber et al. for incorporation of the rapamycin-binding domains as these insertions were viable (2006). To further minimise potential structural disruptions, 10 amino acid linkers of glycine-serine residues were added either side of the inserted SpyTag sequence. Using the known structure of CnaB2 to represent SpyCatcher, each domain appeared free of steric interference using these linker lengths (Oke et al., 2010). The Smc1 hinge (S503-E681) was tagged with MBP to maximise solubility and also to distinguish potential cross-linked species by size. The Smc3 hinge (T495-L705) was FLAG-tagged. The hinges were co-expressed in *E. coli* and affinity purified as a dimeric complex. The stoichiometry of each construct containing the SpyTag insertion was seen to be equivalent to wild type, suggesting that dimerisation was not significantly affected by this modification (Fig. III-4B). Analytical gel filtration of each complex revealed a

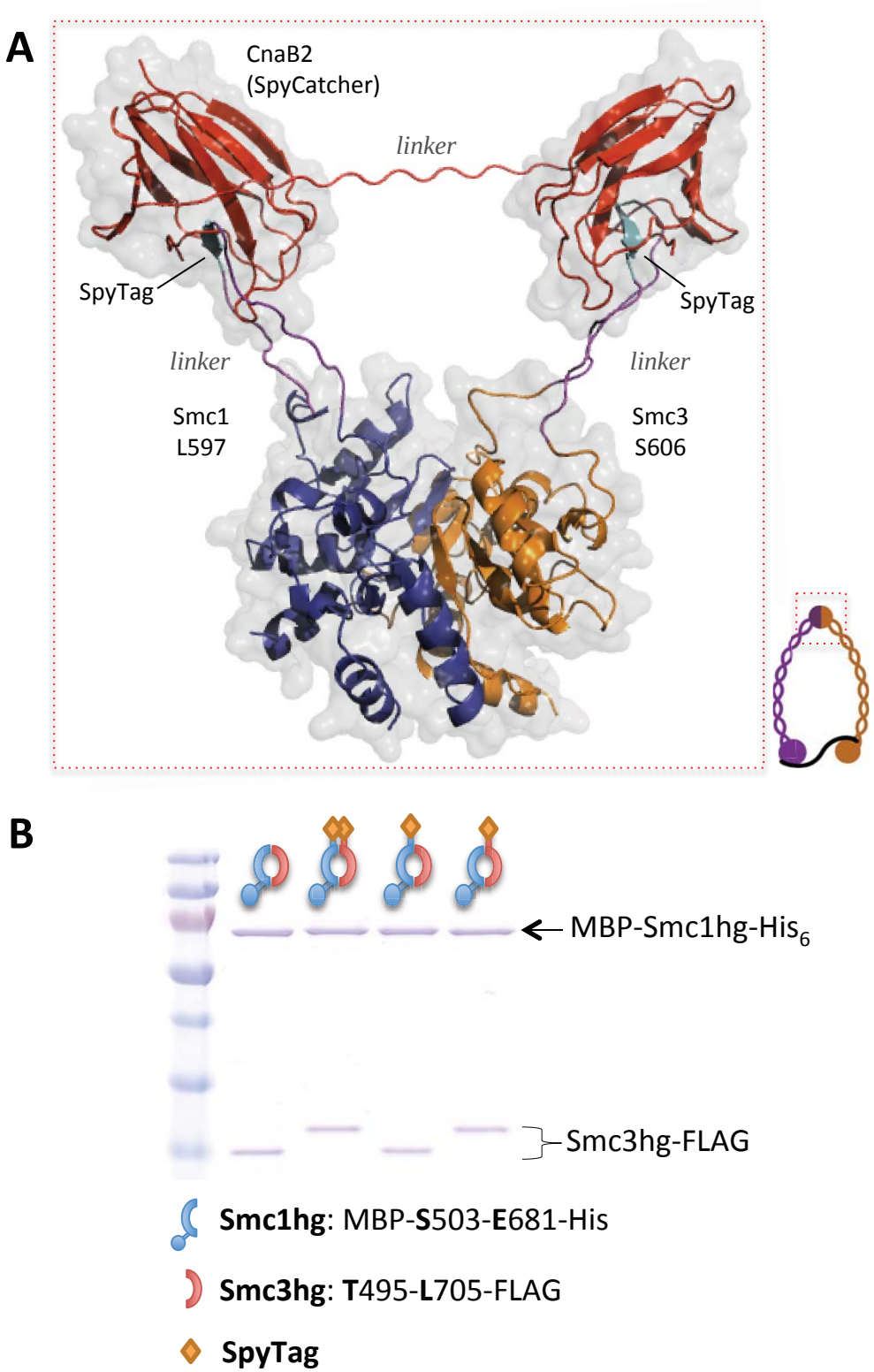


Figure III-4. Smc1/3 hinge domains with SpyTag insertions.

A) Representation of the linker-SpyTag-linker insertions within Smc1 and Smc3 hinge domains of *S. cerevisiae*, structure homology-modelled using *M. musculus* hinge structures as templates (PDB 2WD5). For the linker length used, SpyCatcher domains (based on CnaB2; PDB 2X5P) are small enough to bind each SpyTag without steric interference either between domains or with hinges.

B) Purification of hinge domains from *E. coli.* MBP-Smc1hg-His_6 with or without SpyTag revealed the same stoichiometry of associated Smc3hg (+/- SpyTag) after affinity purification to homogeneity (~600ng of each dimer was analysed by SDS-PAGE).

very minor degree of aggregation where SpyTag insertions were present in Smc3hg (Fig. III-5). Otherwise, each of the proteins eluted at a volume consistent with their expected size.

In vitro cross-linking of Smc1 and Smc3 hinge domains

SpyTag has been characterised as reacting with SpyCatcher in the order of minutes and with high efficiency in a simple *in vitro* assay (Zakeri et al., 2012). Here, the efficiency of cross-linking was investigated in the context of the hinge. On account of the covalent nature of the cross-linking, any reactions taking place are detectable by SDS-PAGE. Addition of a 3-fold molar excess of SpyCatcher[7] to the hinge dimers formed higher molecular weight bands corresponding to SpyCatcher-cross-linked products (Fig. III-6A). The formation of these species was dependent on the presence of the SpyTag moiety and also correlated with a reduction in the amount of unreacted protein (approximately 90%). This represents a high efficiency and specificity in the cross-linking reaction.

To determine whether the Smc1hg and Smc3hg SpyTag insertions allow efficient cross-linking to one another by a tandem version of the SpyCatcher, the assay was repeated under the same conditions and the products were assigned identities based on agreement between estimated and observed molecular weights, as judged by SDS-PAGE (Fig. III-6B). As seen with monomeric

[7] SpyCatcher refers to the monomeric protein construct, i.e. containing a single reactive domain. Where Tandem SpyCatcher protein was used it is referred to as such.

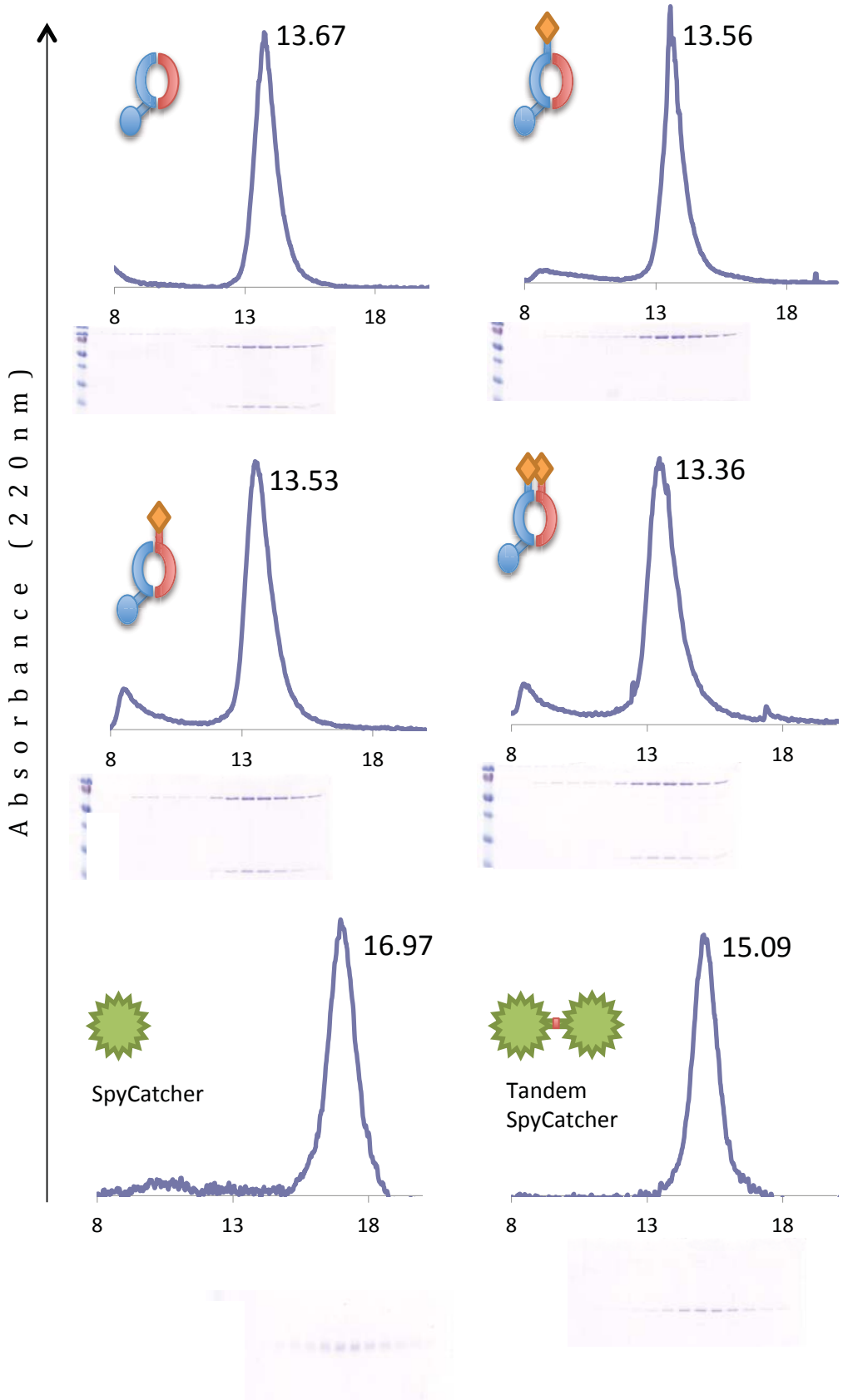


Figure III-5. SpyTagged hinge domains and SpyCatcher constructs are predominantly monodisperse with minimal aggregation.
10μM of each complex was run on the analytical gel-filtration column (Superdex 200 10/300) to test the associations between proteins. Each complex eluted at a size in accordance with a dimer between Smc1 and Smc3 hinge domains. Purified monomeric and Tandem SpyCatcher proteins were also seen to elute as expected for non-aggregated, soluble proteins.

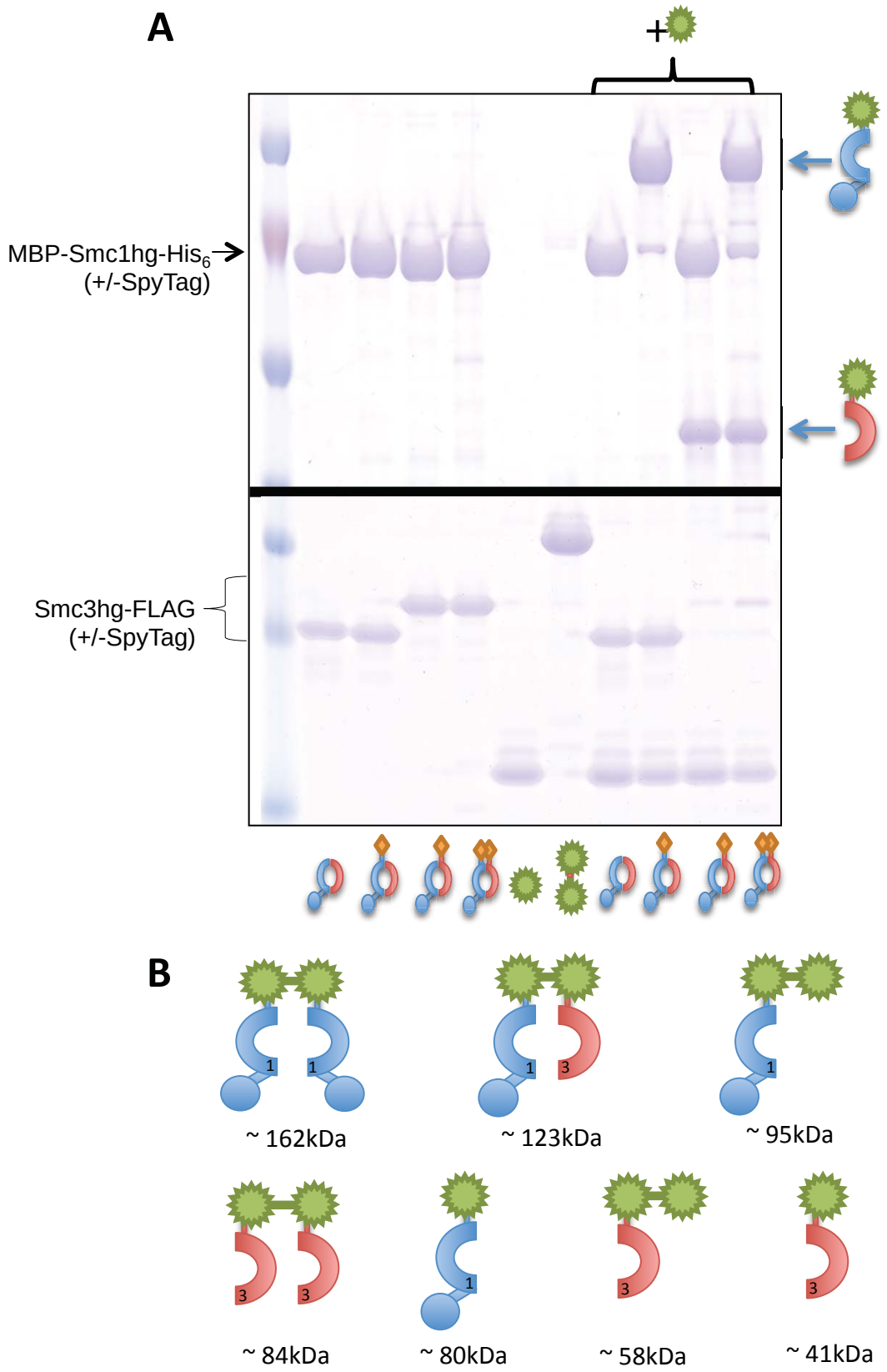


Figure III-6. SpyTagged hinge domains react efficiently with SpyCatcher *in vitro*.

(A) SDS-PAGE of covalent species produced upon incubation of 4μM hinge complex (+/- SpyTag as indicated) with 12μM SpyCatcher for 1.5hrs at 23°C. The image is horizontally bisected into the two gels used to resolve proteins of all sizes (8% acrylamide, top; 12.5% acrylamide, bottom).

(B) Predicted molecular masses for all possible cross-linked species produced upon addition of SpyCatcher or Tandem SpyCatcher.

SpyCatcher, the Tandem SpyCatcher reacted extensively only when SpyTag was present on either or both of the Smc hinge proteins (Fig. III-7).

However, in addition to the expected intra-complex cross-linking witnessed, i.e. between Smc1hg and Smc3hg bound to one another as a heterodimer, extra bands appeared that most likely represent cross-linking occurring between two molecules of the same Smc protein. This may be due to high concentrations of protein providing the opportunity for the two reactive domains of Tandem SpyCatcher to cross-link with SpyTags on two separate heterodimers of Smc1 and Smc3, simply by virtue of their concentration and thus proximity in solution. Secondly, some species are cross-links between SpyTagged hinge and just one domain of the Tandem SpyCatcher; the other domain remains free and unreacted. Also of note is the fact that each cross-linked species runs as a doublet. This may be due to asymmetry within the Tandem SpyCatcher such that the conformations of cross-linked products are subtly different even after denaturation depending on the orientation of SpyCatcher at the time of bond formation. SpyCatcher itself tends not to run as a clear, single band on account of its ability to partially fold even in the presence of detergent.

It was reasoned that decreasing the concentration of reactants and manipulating the ratio of SpyCatcher to SpyTag may favour the formation of certain products and minimise others associated with less specific cross-linking without compromising efficiency. The assay was repeated using up to a 40-fold lower concentration of hinge dimers and a molar ratio of SpyCatcher to hinge ranging from as much as 50:1 to 2:1. However, at each range of concentrations

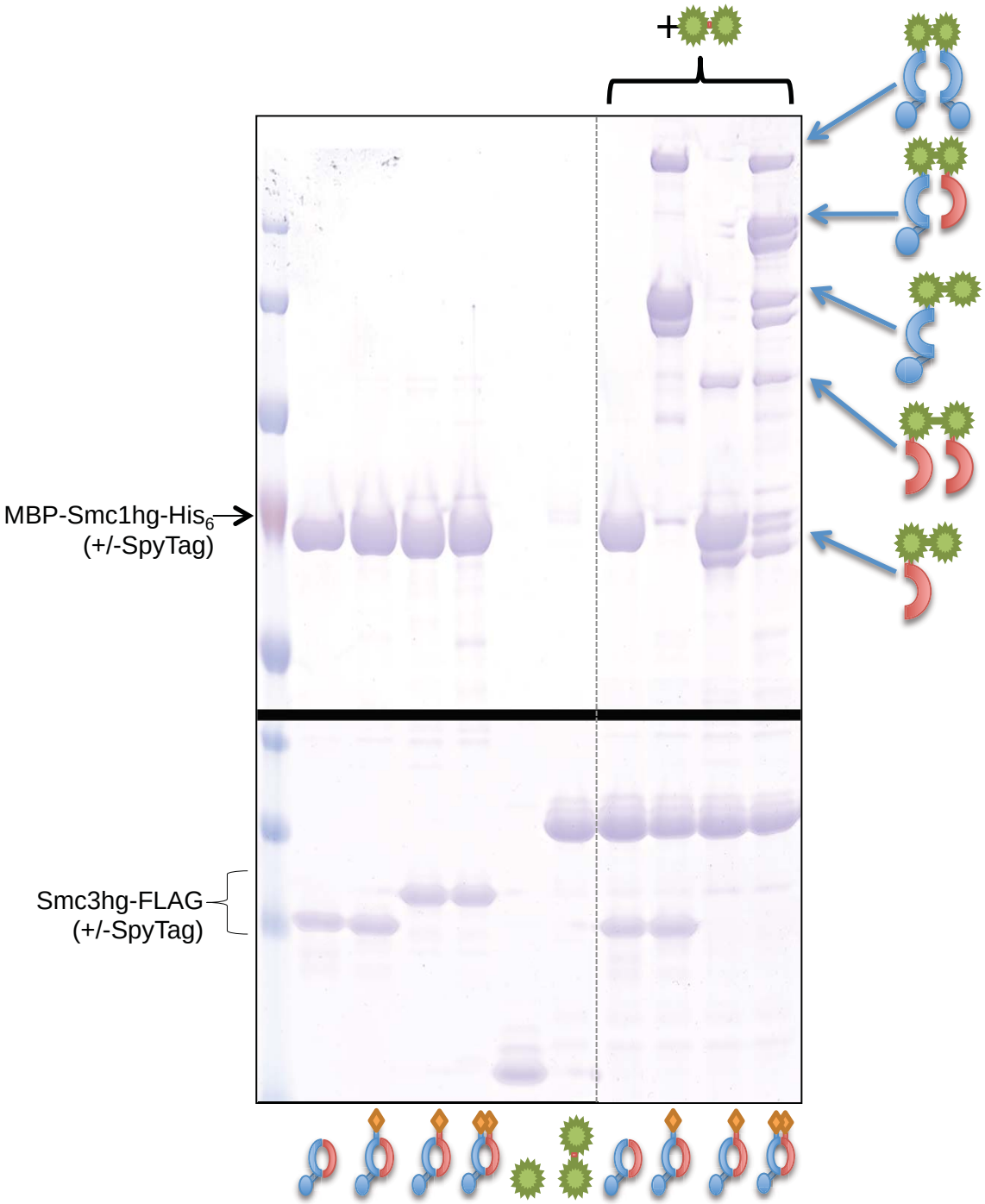


Figure III-7. Tandem SpyCatcher generates intradimeric cross-links between Smc1 and Smc3 hinge domains *in vitro*.
SDS-PAGE of covalent species produced upon incubation of 4μM hinge complex (+/- SpyTag as indicated) with 12μM Spycatcher for 1.5hrs at 23°C. The image is horizontally bisected into the two gels used to resolve proteins of all sizes (8% acrylamide, top; 12.5% acrylamide, bottom). The dashed grey line indicates where intervening lanes were omitted to simplify comparison.

tested, the proportions of each product were not significantly different (Fig. III-8A). Increasing the incubation time from 15mins to 120mins (21°C) did not greatly increase the extent of cross-linking, although the relative amounts of singly reacted Tandem SpyCatcher were marginally decreased.

Another variable predicted to influence the specificity of cross-linking was the length of the linkers used both to insert SpyTag into the hinge domains and to separate the two domains of Tandem SpyCatcher. In fact, shortening the linker length between SpyTag and Smc sequence seemed to slightly decrease the proportion of the favoured Smc1-Smc3 species, whilst changing the linker length between SpyCatcher domains had no effect (Fig. III-8B).

Although the *in vitro* cross-linking assay demonstrated the formation of some unexpected products, the conditions are not truly representative of what might occur *in vivo*. Inside the cell, full-length Smc proteins form the backbone of the cohesin complex. The probability of inter-complex cross-linking is thought to be significantly less *in vivo* due to a much lower local concentration of substrate. Moreover, in the context of the intact cohesin ring, hinge domains cannot dissociate and recombine to the same extent.

Toxicity, reactivity, expression and localisation of SpyCatcher in S. cerevisiae.

Although SpyCatcher has been reported as cross-reacting with unidentified proteins when expressed in the cytoplasm of *E. coli*, albeit to a minor degree (J. Fierer, pers. comm.), the propensity for side-reactions to occur between SpyCatcher and *S. cerevisiae* proteins is unknown and may be more significant. To address this, purified SpyCatcher protein was added to yeast

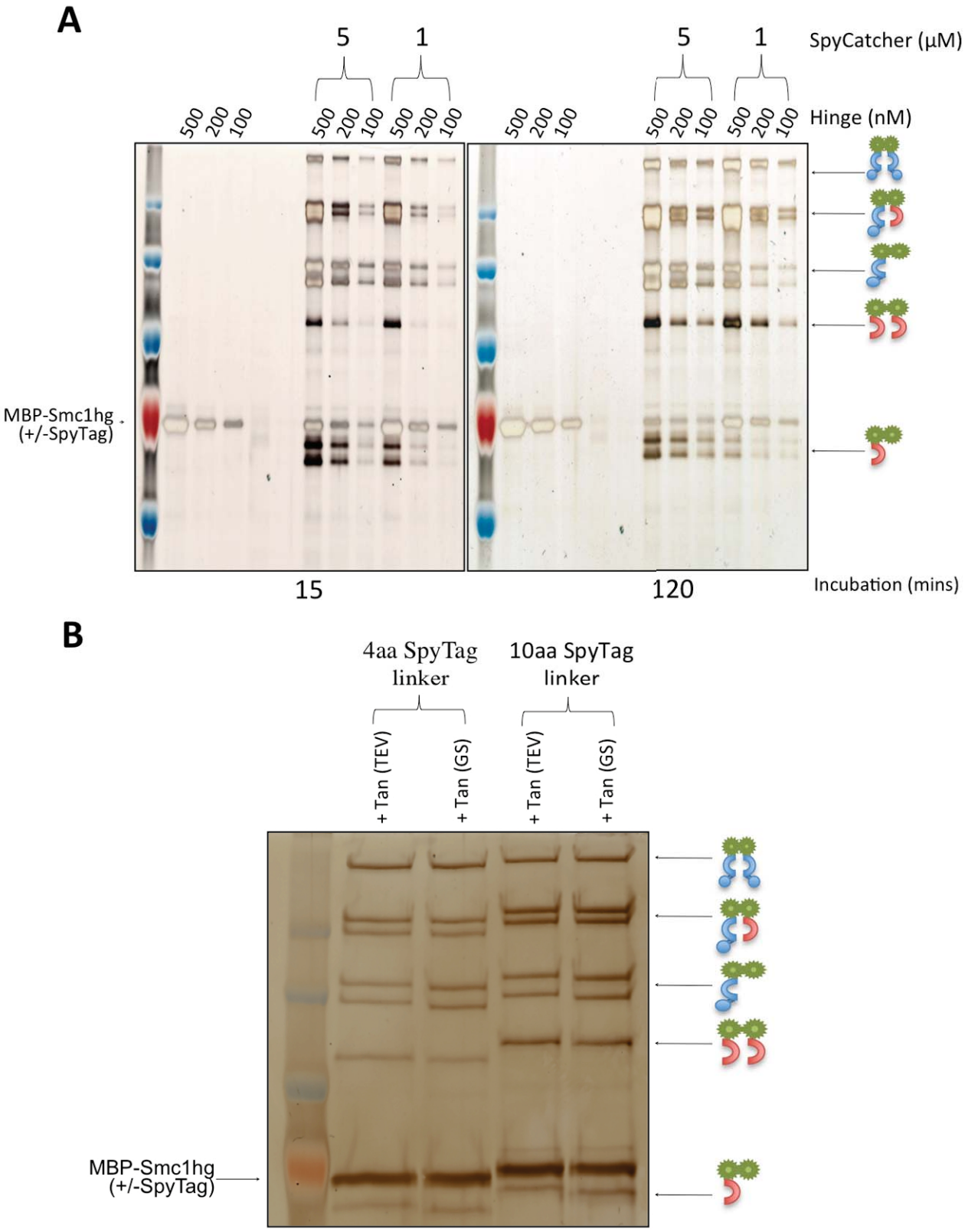


Figure III-8. The effects of reactant concentration, hinge molar ratio to Tandem SpyCatcher, incubation time and linker lengths on the proportions of each product.

(A) Silver-staining of covalent species produced with different concentrations of hinges and SpyCatcher after 15mins (left) and 120mins (right).

(B) 100nM of hinge constructs containing either 4aa or 10aa of GS-linker length either side of SpyTag at insertion sites were mixed with 150nM of Tandem SpyCatcher (Tan) with a 10aa linker (GS) or 14aa linker (TEV) between domains. Proteins were incubated for 1.5hrs prior to analysis by silver-staining.

whole cell extract and incubated at 30°C for 6hrs. Potential adducts were assayed for by western blot against the His_6 tag present on SpyCatcher (Fig. III-9A). Several bands in addition to the primary SpyCatcher band were evident but these were present equally when SpyCatcher was in buffer or yeast lysate. This suggests adducts formed in *E. coli* during the expression and purification procedure. There was, however, a more prominent band specific to SpyCatcher in the yeast lysate and this could represent either an unspecific covalent reaction between an endogenous protein and SpyCatcher, or post-translational modification of SpyCatcher itself. The levels of these species represented a small fraction of the unmodified SpyCatcher added to the lysate and were not investigated further. However, the importance of these species may be contingent on the *in vivo* expression level of SpyCatcher: if the induced level of SpyCatcher is significantly less than that added *in vitro*, then there may no longer be an excess of reactive, free or unmodified SpyCatcher.

Having established that SpyCatcher does not participate in side-reactions with yeast proteins to a prohibitive degree, it is important to ascertain whether components in the yeast proteome may instead prevent or inhibit the reactivity of SpyCatcher towards its target substrate, SpyTag. As such, purified, SpyTagged Smc1 and Smc3 hinge domains were incubated together with purified SpyCatcher in either buffer or yeast whole cell extract (at a concentration of 3mg/mL; Fig. III-9B). The proportions of cross-linked products (X) produced under both conditions were seen to be similar.

Although SpyCatcher appeared to participate minimally in unspecific cross-reactions with other molecules, it was important to assess whether the

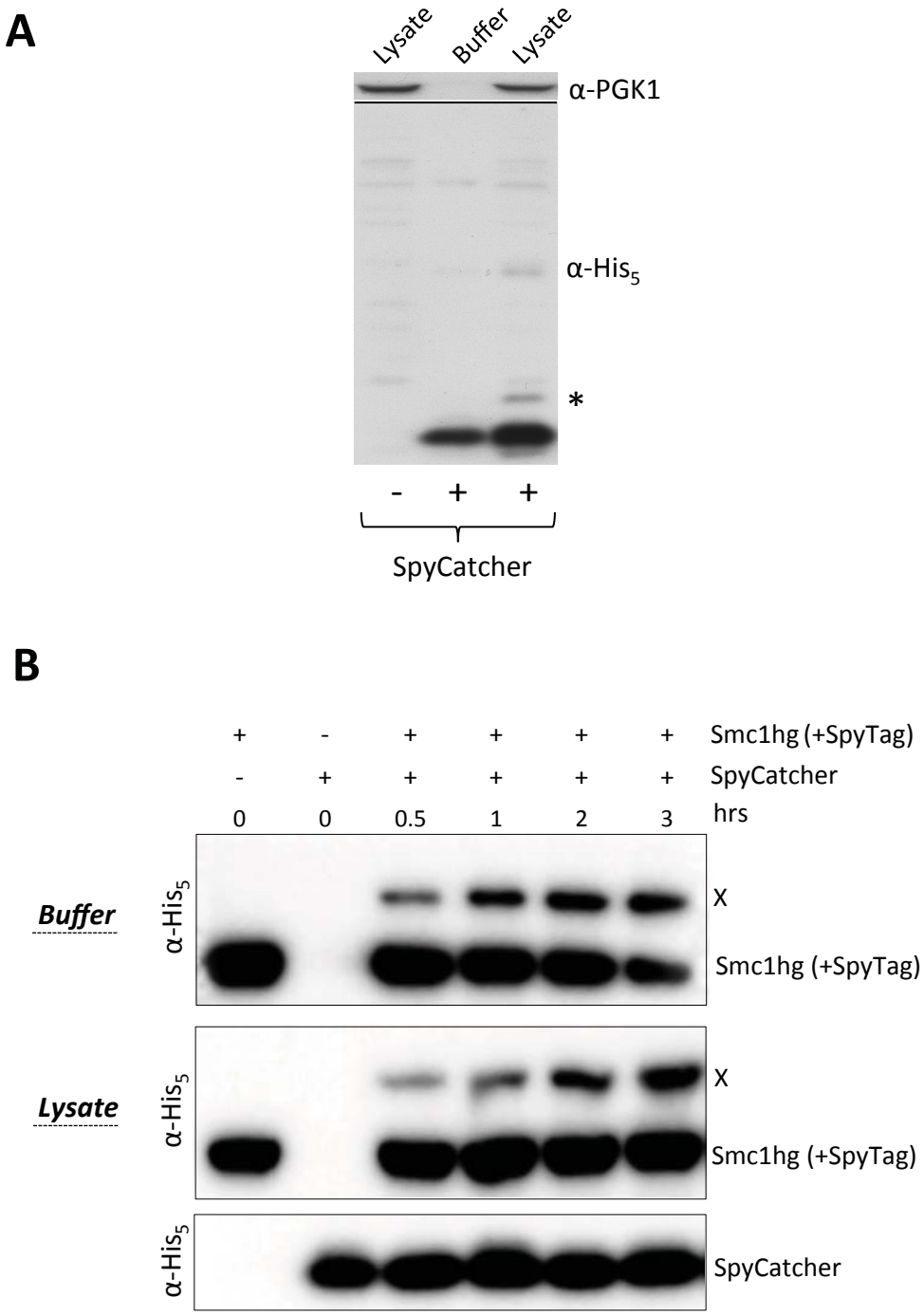


Figure III-9. SpyCatcher shows negligible cross-reactivity with budding yeast proteins and is similarly uninhibited by them.

(A) SpyCatcher was added to yeast lysate to a concentration of 1μM and incubated for 3hrs at 4°C. * represents a side-reaction product seen only upon addition of SpyCatcher.

(B) 400nM of SpyTagged Smc1/3hg complex was added to an equal volume of either buffer (Extract buffer, see Ch.VI:Materials and methods) or yeast lysate at 3mg.mL^{-1} in buffer. Samples were incubated at 30°C for up to 3hrs and levels of products (X) were assessed by α-His$_5$ western blot.

expression of SpyCatcher was toxic to *S. cerevisiae*, as non-covalent interactions with other proteins may be equally deleterious. To address this, both monomeric and Tandem SpyCatcher constructs were integrated at the *URA3* locus under control of the galactose-inducible promoter (*GAL1pro*) in wild-type background strains. Cell growth appeared normal on galactose relative to wild type and an empty vector background (Fig. III-10A). Analysis of induced populations by FACS was in agreement, with no cell cycle defects discernible (Fig. III-10B).

The kinetic profile of SpyCatcher expression was established by immunoblotting and protein was detectable after 45mins in galactose and approaching a plateau of expression by 300mins (Fig. III-10C). As the ultimate aim is to cross-link the DNA-bound fraction of cohesin, i.e. the population already loaded, a nuclear localisation sequence (NLS) was added to each SpyCatcher construct. Live-cell imaging of a C-terminally GFP-tagged version of SpyCatcher demonstrated accumulation of this fusion protein to a high level within the nucleus (Fig. III-10D).

Optimisation of SpyCatcher expression

To improve the speed and strength of response of the galactose-inducible SpyCatcher protein, the background wild type strain was transformed with a greater concentration of the integrative vector to increase the number of tandem repeats of the construct. Strains carrying multiple integrations of each construct expressed greater levels of the protein within a given period of time.

Further enhancement of expression was achieved by crossing SpyCatcher

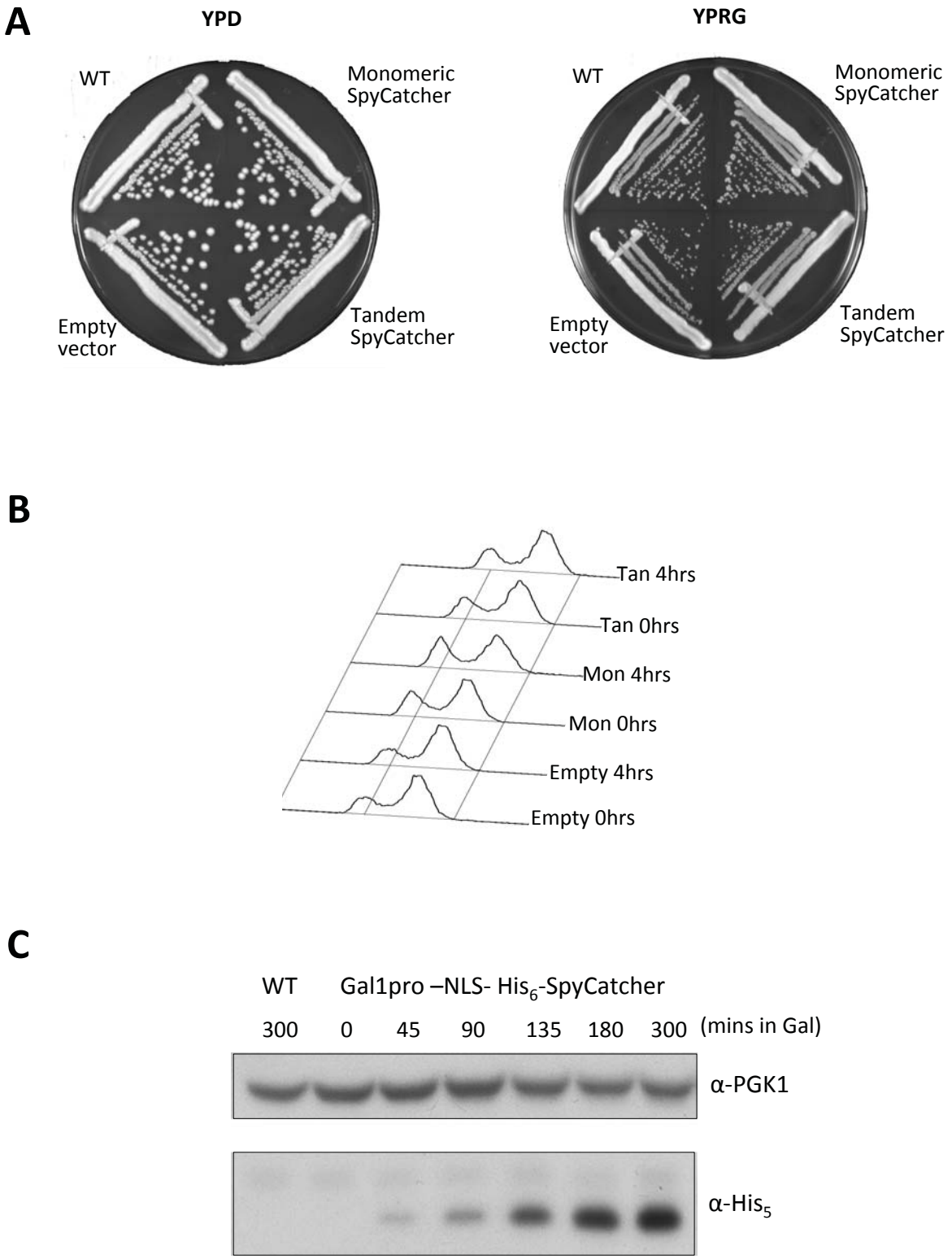


Figure III-10. Strain growth and expression of SpyCatcher *in vivo* (*continued overleaf*).

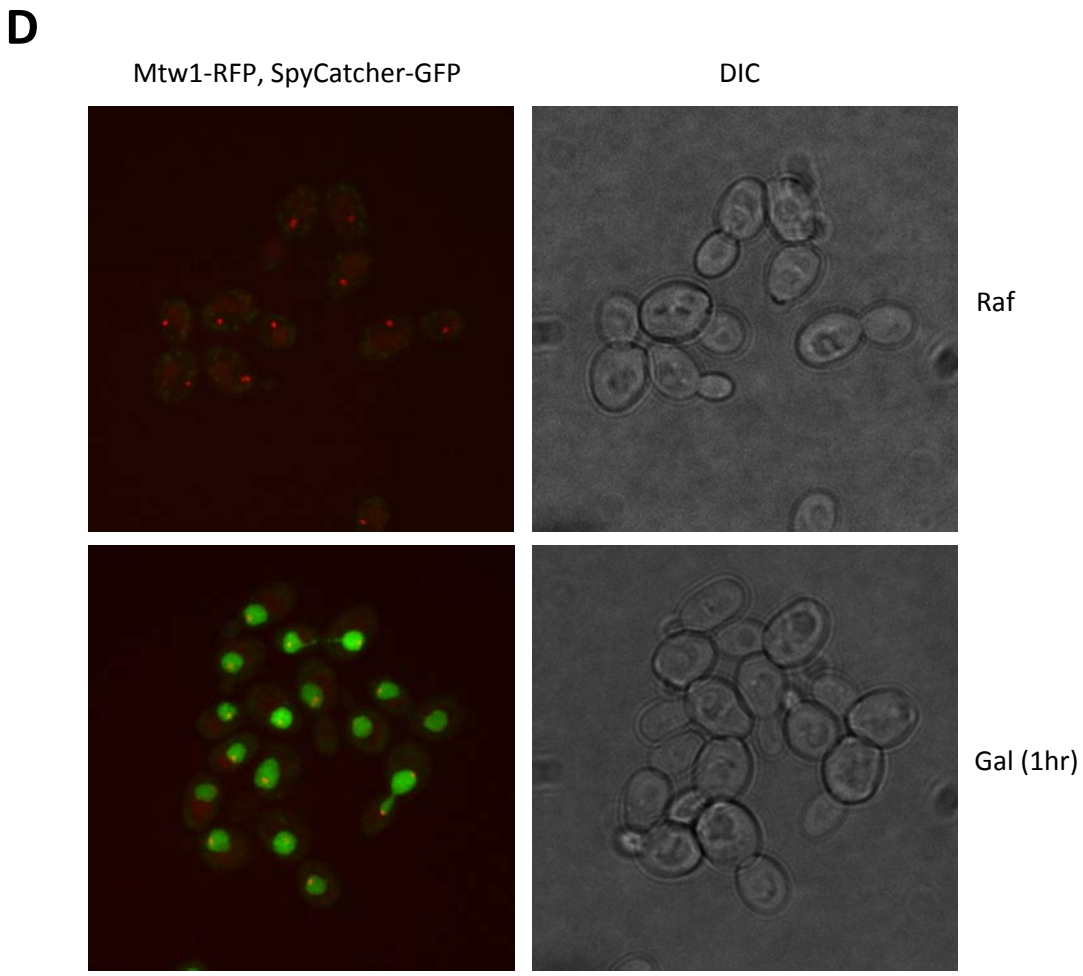


Figure III-10. Strain growth and expression of SpyCatcher *in vivo*. SpyCatcher is non-toxic, expressed and localises to the nucleus as directed.

(A) Strains were constructed with monomeric or Tandem SpyCatcher under control of the *GAL1* promoter at the *URA3* locus. All strains grew as wild-type on glucose or galactose medium.

(B) FACS analysis of SpyCatcher-expressing strains (Mon or Tan) revealed no defects in growth.

(C) His_6-NLS-SpyCatcher is expressed upon induction with galactose, as assessed by immunoblot of TCA preps.

(D) His_6-NLS-SpyCatcher-GFP accumulates in the nucleus. Live, exponentially-growing diploid cell populations were imaged in the presence of either raffinose or galactose.

strains with GAL*plus*, a strain harbouring *HIS3pro-GAL1*, *HIS3pro-GAL2* and *GAL10pro-GAL4* at the *LEU2* locus (Matsuyama et al., 2011). Constitutive expression of Gal1 and Gal2 from the *HIS3* promoter primes the cell for the metabolic response to galactose whilst the levels of Gal4, a transcriptional activator, are boosted through positive feedback. Tetrad spores viable on both –LEU and –URA media, produced from sporulation and dissection of this parent strain, were grown in culture and the expression was seen to be greatly accelerated and increased relative to a strain lacking the GAL*plus* cassette (Fig. III-11A). The empty vector PGK1 loading control allowed comparison. Levels of Tandem SpyCatcher were visible by western blot within 30mins and had reached a high concentration by 90mins, approximately equivalent to that achievable after 4hrs with standard galactose induction in a wild-type background.

Whilst beneficial for short-term protein expression, this system appears to be lethal to the cell when galactose is present over longer periods of time, even in the absence of a heterologous protein being expressed from the *GAL1* promoter. The GAL*plus* background strain is dead on galactose plates (Fig. III-11B). The reason for such lethality is unknown but may involve transcriptional stress. The demand for RNA polymerase II by the galactose-responsive genes may be so high as to quench transcription from other areas of the genome involved in essential processes.

Even though exogenous SpyCatcher added to yeast lysate was able to react with SpyTag (as demonstrated in Fig. III-9B), the subsequent addition of the N-terminal NLS and the conditions of endogenous, *de novo* expression may

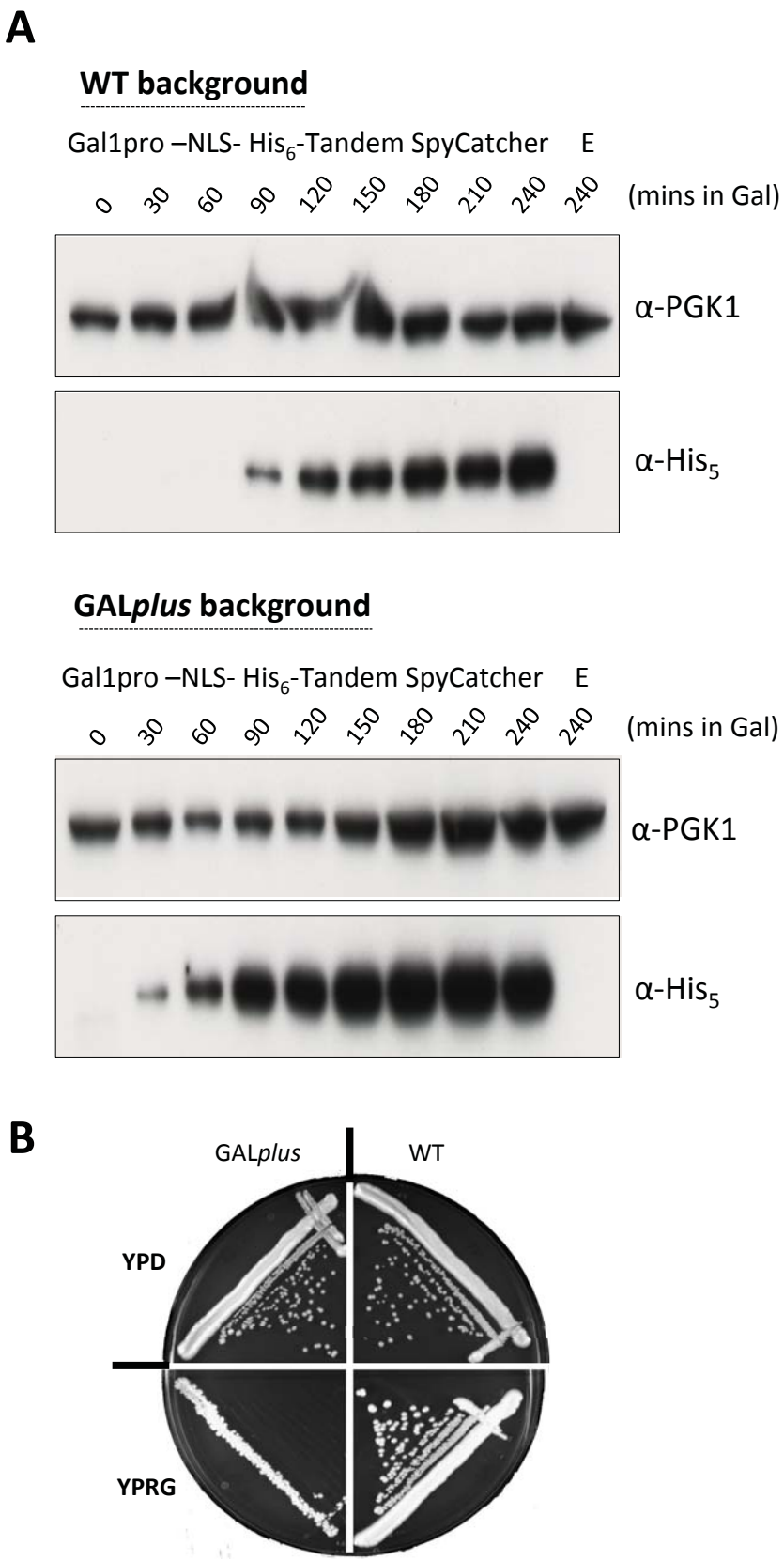


Figure III-11. Optimization of SpyCatcher expression

(A) Expression over time of Tandem SpyCatcher was compared between a wild-type background and one harboring the GAL*plus* cassette (see text). TCA preps were prepared from each time point and membranes were immunoblotted and exposed simultaneously, with the empty vector loading control acting as a reference.

(B) The GAL*plus* background causes lethality in the presence of galactose.

negatively affect its ability to react chemically. Therefore, SpyCatcher was endogenously expressed and a cell lysate prepared to which a substrate was added (hinge dimers containing SpyTags). The formation of covalently cross-linked products was monitored over a period of 6hrs by western blot and found to reach proportions in keeping with what was observed previously for NLS-free SpyCatcher translated in *E. coli* (Fig. III-12).

In vivo cross-linking of Smc proteins

Strains with *SMC1* and *SMC3* at their endogenous loci were engineered to contain SpyTag insertions at the same positions described above for the *in vitro* cross-linking experiments with isolated hinge domains. Initially, both Smc1 and Smc3 were C-terminally tagged (PK6 and HA6, respectively) for detection by western blot, however the presence of the PK6 tag on Smc1 caused sickness that appeared synthetic with the SpyTag insertion in Smc3. For this reason, the PK6 tag on Smc1 was removed. Each parent strain and the double Smc1 and Smc3 SpyTagged strain (1s3s) were healthy. Interestingly, when 1s3s was combined with Tandem SpyCatcher (1s3s +Tan), this strain failed to grow on galactose, whereas on glucose, colony growth appeared as wild type (Fig. III-13A). In contrast, the 1s3s strain expressing monomeric SpyCatcher (1s3s +Mon) was healthy. Similarly, a diploid strain heterozygous for the SpyTag insertions within each Smc protein was also healthy, indicating that the lethality from Tandem SpyCatcher induction represents a loss of function and is not a dominant-negative effect (Fig. III-13A).

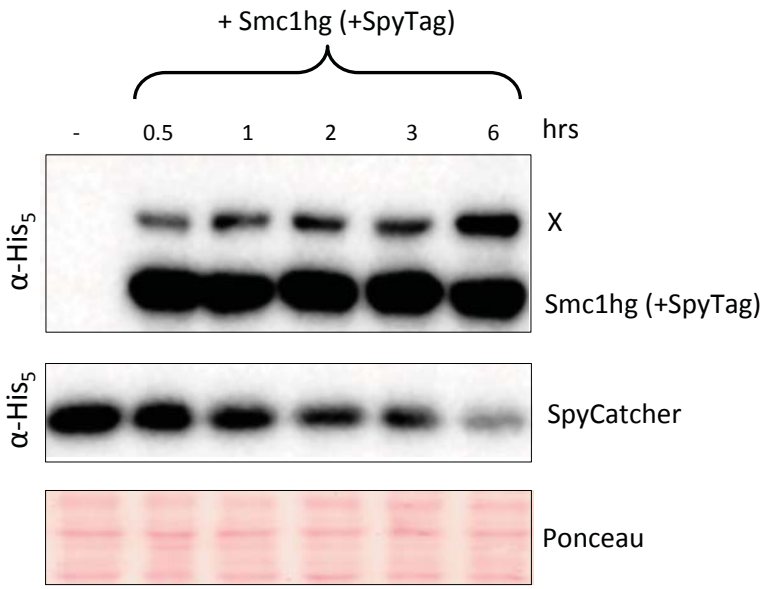


Figure III-12. Endogenously-translated and NLS-tagged SpyCatcher can react with exogenous SpyTagged Smc1hg added to whole cell extracts.
SpyCatcher expression was induced by growth of cells in galactose for 4hrs. Cells were lysed and the SpyTagged Smc1hg complex was added as substrate (to a concentration of 400nM) to asses the reactivity of SpyCatcher produced *in vivo* by *S. cerevisiae.* Equal loading of lysate in each sample was assessed by ponceau-staining of the membrane blotted with α-His_5 to detect the formation of higher molecular weight species. The experiment was carried out in parallel with those depicted in Figure III-9B. Levels of SpyCatcher in the extracts decrease with time, corresponding with the appearance of a cross-linked protein band (X).

Importantly, lethality also depends on the presence of a SpyTag moiety in both Smc proteins; strains containing just one SpyTagged Smc protein (i.e. either 1s3 or 13s) were viable in the presence of Tandem SpyCatcher (1s3 +Tan shown in Fig. III-13B). Furthermore, co-expression of TEV protease with Tandem SpyCatcher prevented lethality (Fig. III-13B).

The physiological effect of cross-linking at the hinges was investigated further by analysing the growth of strains in liquid culture in rich medium over a 26hr period after the addition of galactose. Measurements were begun 3hrs after induction of cultures with galactose upon dilution in YPRG to an optical density (OD_{600nm}) of ~0.015. The growth profile of 1s3s +Mon was comparable to wild type, whereas 1s3s +Tan showed a reduction in the rate of growth after approximately 9hrs in the presence of galactose (Fig. III-14). A strain harbouring the GAL*plus* modification showed rapid inhibition in the rate of growth (within 3hrs). However, as this strain expressed the normally healthy combination of monomeric SpyCatcher construct with SpyTagged 1s3s (1s3s +Mon), the inhibition of growth is seemingly a result of metabolic dysfunction in galactose-supplemented medium. The timing in suppression of growth of the 1s3s +Tan culture correlated well with cell cycle defects manifest in FACS analysis (Fig. III-14). In comparison, the 1s3s +Mon strain showed normal profiles over the measured time-span.

The specific lethality observed in the 1s3s +Tan strain may be attributable to cross-linking of the interface between the two Smc proteins, thereby preventing the entrapment of DNA by the cohesin complex. If this is the case, cross-linking between Smc1 and Smc3 should be visible by western blot. To this

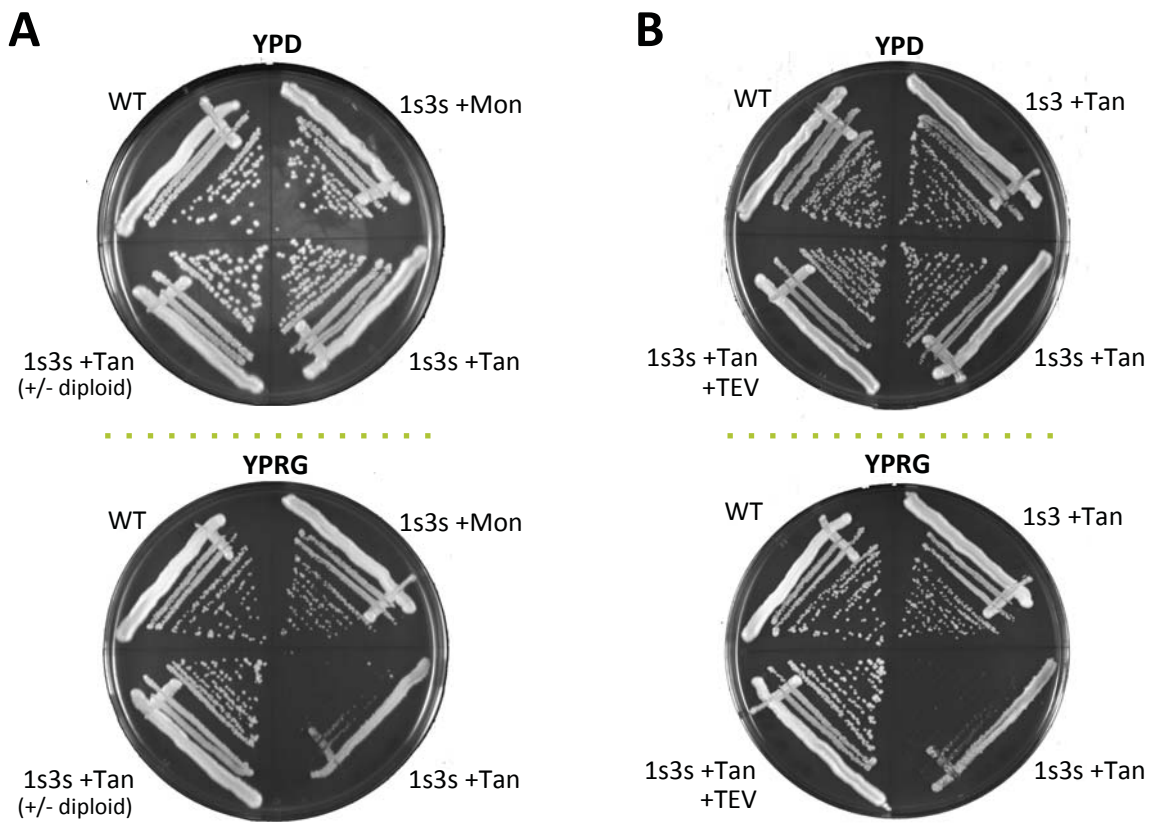


Figure III-13. Covalent closure of the Smc1-Smc3 hinge interface is lethal.
Genetic analysis of SpyCatcher (+Mon or +Tan) expression in strains with SpyTagged full-length Smc1 and Smc3 proteins.

(A) On YPD (top), when the expression of SpyCatcher constructs is repressed, all strains grow as wild-type. Thus, SpyTag insertions themselves do not disrupt protein function. When SpyCatcher expression is induced by galactose, the Tandem construct causes death of doubly SpyTagged (i.e. within both Smc1 and Smc3; 1s3s), whereas the monomeric construct does not (lower). A diploid strain heterozygous for SpyTag insertions is viable and unaffected.

(B) Co-expression of TEV protease rescues the lethality from inducing Tandem SpyCatcher (TEV site within linker domain) when both Smc1 and Smc3 contain SpyTags. SpyTag within just one Smc has no negative effect on growth, as demonstrated by 1s3 +Tan (13s +Tan not shown).

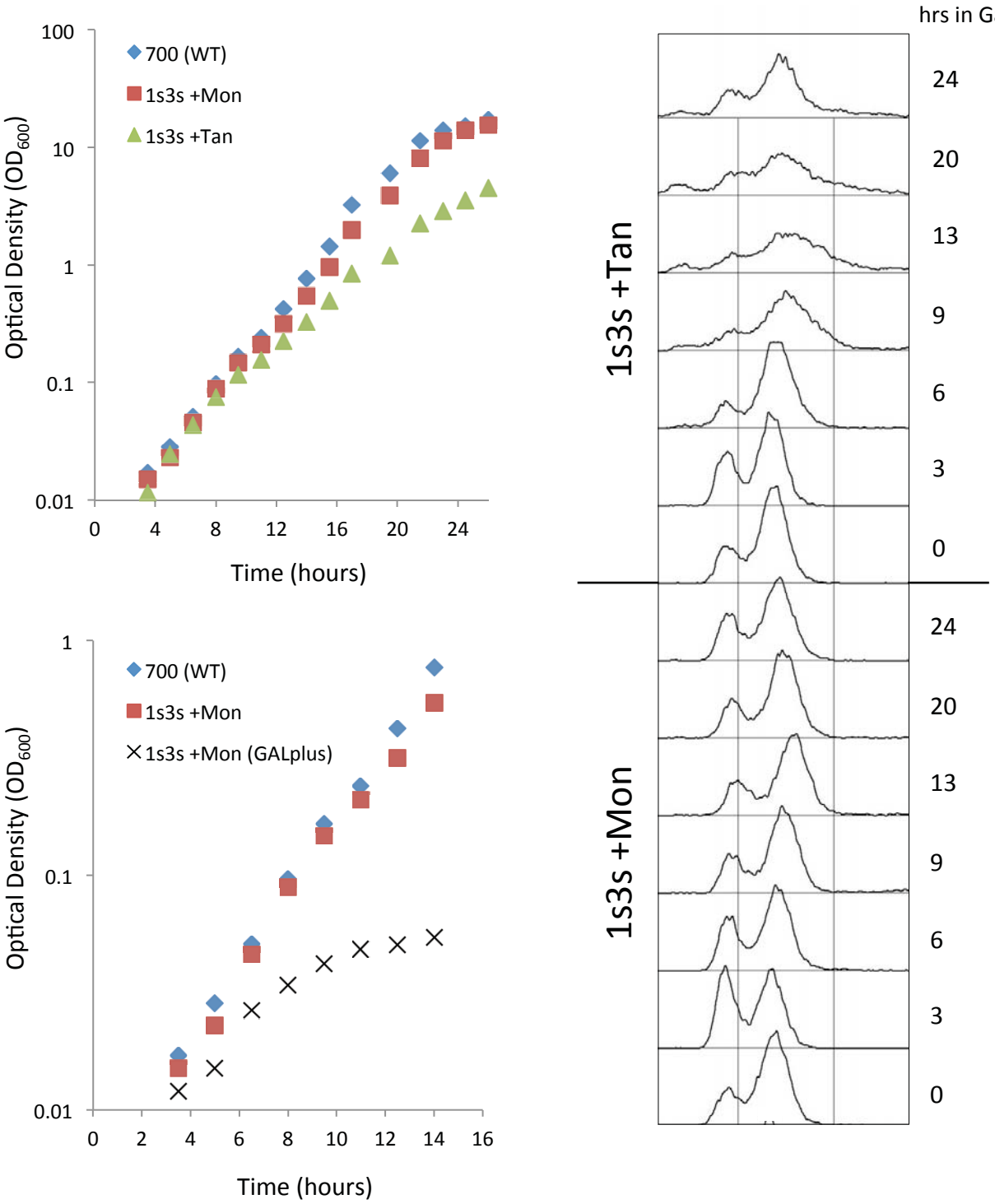


Figure III-14. Covalent closure of the Smc1-Smc3 hinge produces a phenotype consistent with cohesion defects.

Logarithmic growth curves of each strain are based on duplicate OD_{600nm} measurements taken every 90mins in the presence of galactose. 1s3s +Mon (red) grows at a rate and to a final density comparable to WT (blue; OD_{600nm} ~16), whilst 1s3s +Tan (green) begins to show longer doubling times after approximately 8hrs in galactose (final OD_{600nm} ~4.5). This was also reflected in the FACS analysis. The otherwise healthy 1s3s +Mon strain showed growth defects more rapidly than 1s3s +Tan when combined with GAL*plus*. The rapid induction system is thus unsuitable for studying the phenotype of Tandem SpyCatcher cross-linking of the hinges.

end, cells were grown in the presence of galactose to induce either monomeric or Tandem SpyCatcher, and samples were taken at regular time intervals to probe for the appearance of high molecular weight adducts indicative of cross-linking with SpyCatcher and/or the partnered Smc protein.

Cross-linking of SpyTagged Smc3 with monomeric SpyCatcher was seen to take place readily (within 3hrs) and efficiently; approximately 40-50% of detected full length Smc3 was shifted to a higher molecular weight (Fig. III-15A). In contrast, cross-linking by Tandem SpyCatcher showed a greatly reduced efficiency in giving rise to a very high molecular weight species of a size in accordance with cross-linked Smc1/3 heterodimer. This band represents at most 10% of the total substrate, most prominent between 6-9hrs after expression of SpyCatcher. Also visible is a cross-linking intermediate representing SpyTagged Smc3 covalently bound to a molecule of Tandem SpyCatcher with a free, unreacted domain. Both SpyCatcher constructs (Mon and Tan) were expressed and were not used up in the cross-linking reactions, remaining in excess throughout. Between 9 and 24hrs the levels of Smc3, both cross-linked and unreacted, appeared to diminish, despite equal loading per well. This was a reproducible observation and not likely to represent anomalies in transfer efficiency (see Appendix Fig. A-2 for similar transfer). The loss of protein correlates temporally with the manifestation of slowed growth.

To support the assignment of the highest molecular weight band as protein cross-linked by Tandem SpyCatcher, the time-course was repeated and whole cell extracts were created instead of TCA preparations. Samples from each time point were adjusted to an equivalent concentration and treated with

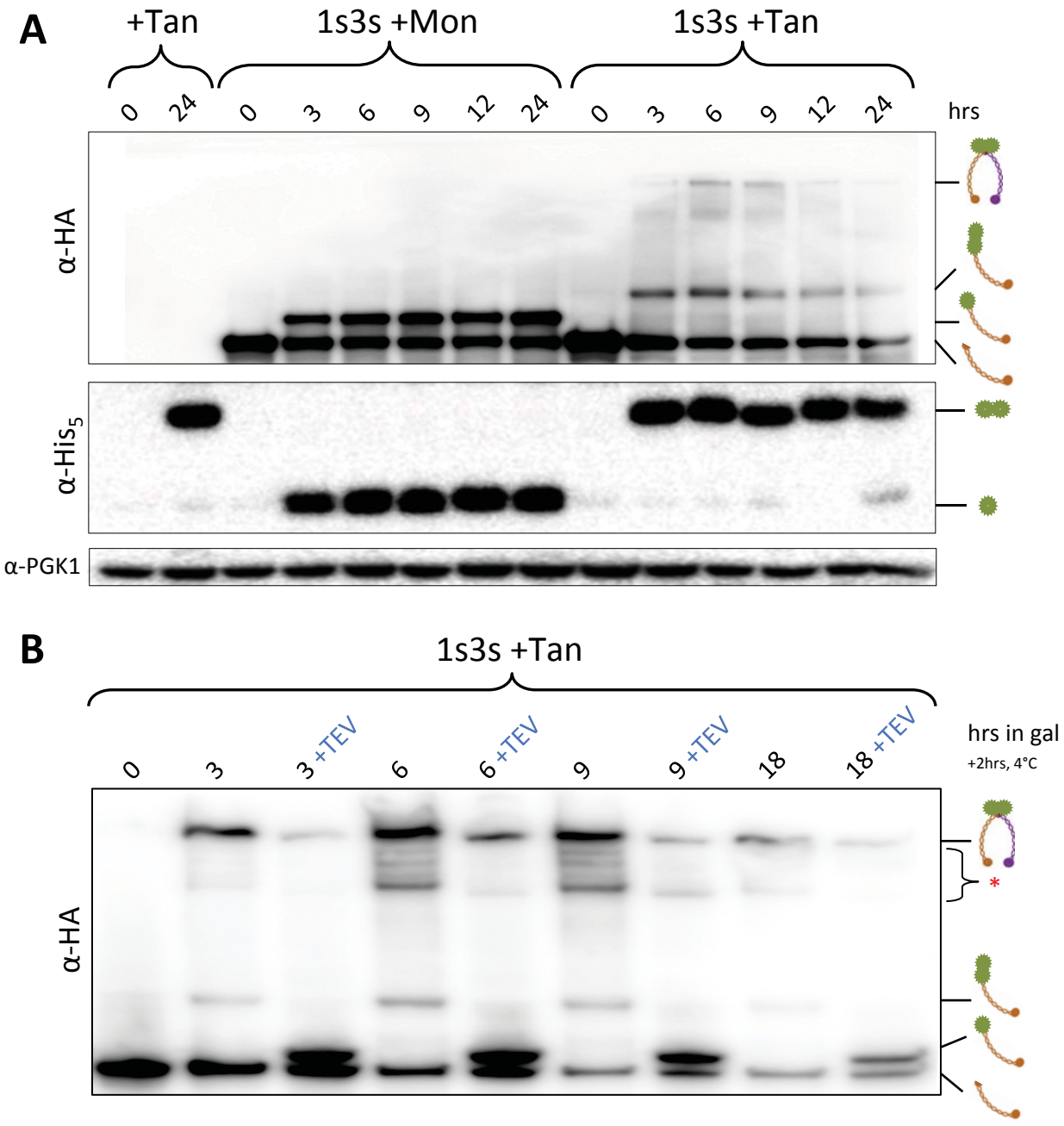


Figure III-15. Cross-linking of cohesin hinge domains *in vivo.*

(A) Western blots of TCA preps of strains grown in YPR to which galactose has been added to the indicated time. Monomeric SpyCatcher is induced and cross-links to the SpyTag in the Smc3 hinge, producing a higher molecular weight product that is stable and approximately 50% of total 3s-HA6 by 24hrs. Tandem SpyCatcher induction produces 2 cross-linked bands assigned the identities of cross-linked heterodimer, and singly-reacted 3s-HA6. Levels of both unreacted and cross-linked proteins decrease over time when Tan is induced.

(B) Whole cell extracts were created from cells grown in the presence of galactose to induce Tan for time indicated. Lysates were adjusted to equivalent concentrations and 30U of TEV protease was added for 2hrs at 4°C. Untreated samples were incubated in parallel to assess the potential contribution of degradation to changes in protein levels.

TEV protease to cleave the linker separating the two SpyCatcher domains of Tan. Following TEV treatment, there is a clear drop in the intensity of the band assigned as cross-linked heterodimer (Fig. III-15B). A band representing a more mobile protein appears concomitantly with TEV treatment and this is assumed to be 3s cross-linked to the SpyCatcher domain remaining after proteolytic cleavage. Both treated and untreated samples were incubated in parallel at 4°C for 2hrs to allow any contribution from unspecific protein degradation to be assessed. The apparently enhanced efficiency of cross-linking seen in these samples may be attributable to the fact that the reaction can continue to occur after sample lysis and may even be enhanced by this process. Non-physiological interactions may occur. Also, lysates consist of predominantly soluble pools of protein and are thus not completely representative of all cellular protein. TCA precipitated protein blots are thought to depict a more accurate snapshot of *in vivo* protein states. Notably, in both preparations the level of protein declined with time suggesting degradation of the product. A ladder of bands was visible that could either represent this degradation or, alternatively, variable mobility of the cross-linked product.

In the strains used above, 1s lacks a tag for detection by western blot and so was not probed for in these samples. As previously mentioned, strains that do have 1s tagged (with PK6) grew very slowly and therefore were unsuitable for assessing cross-linking efficiency *in vivo*. However, it was necessary to check that both the Smc1 and Smc3 SpyTag sites reacted with comparable efficiency to support the identification of the slowest migrating band as a 1s3s cross-linked heterodimer rather than a 3s3s cross-linked homodimer. Therefore, lysates

were created from strains containing single SpyTag insertions in either Smc1 (1s3) or Smc3 (13s) that were fused to PK6 and HA6 epitopes, respectively. Addition of pure Tandem SpyCatcher to these lysates showed no discernible differences in the levels of 1s-Tan or 3s-Tan cross-linked species (Fig. III-16A). Therefore, Tandem SpyCatcher appears not to favour a reaction with SpyTag within one hinge versus another in the context of the cohesin complex. This experiment was repeated with doubly SpyTagged strains to detect the presence of both 1s-PK6 and 3s-HA6 in the heaviest cross-linked product (Fig. III-16B). Therefore, the purported dimeric cross-linked product is likely to represent the capture of 1s3s heterodimers rather than any inter-complex species. This is substantiated by the fact that cohesin complexes with just one SpyTagged Smc, e.g. 13s, do not reveal a shift to a similarly high molecular weight as does the 1s3s strain (Fig. III-16C). Thus, Tandem SpyCatcher appears able to cross-link heterodimers as would be expected by the Smc1/Smc3 composition of cohesin complexes. However, the reaction appears slower than expected *in vivo* and appears to deplete cross-linked proteins.

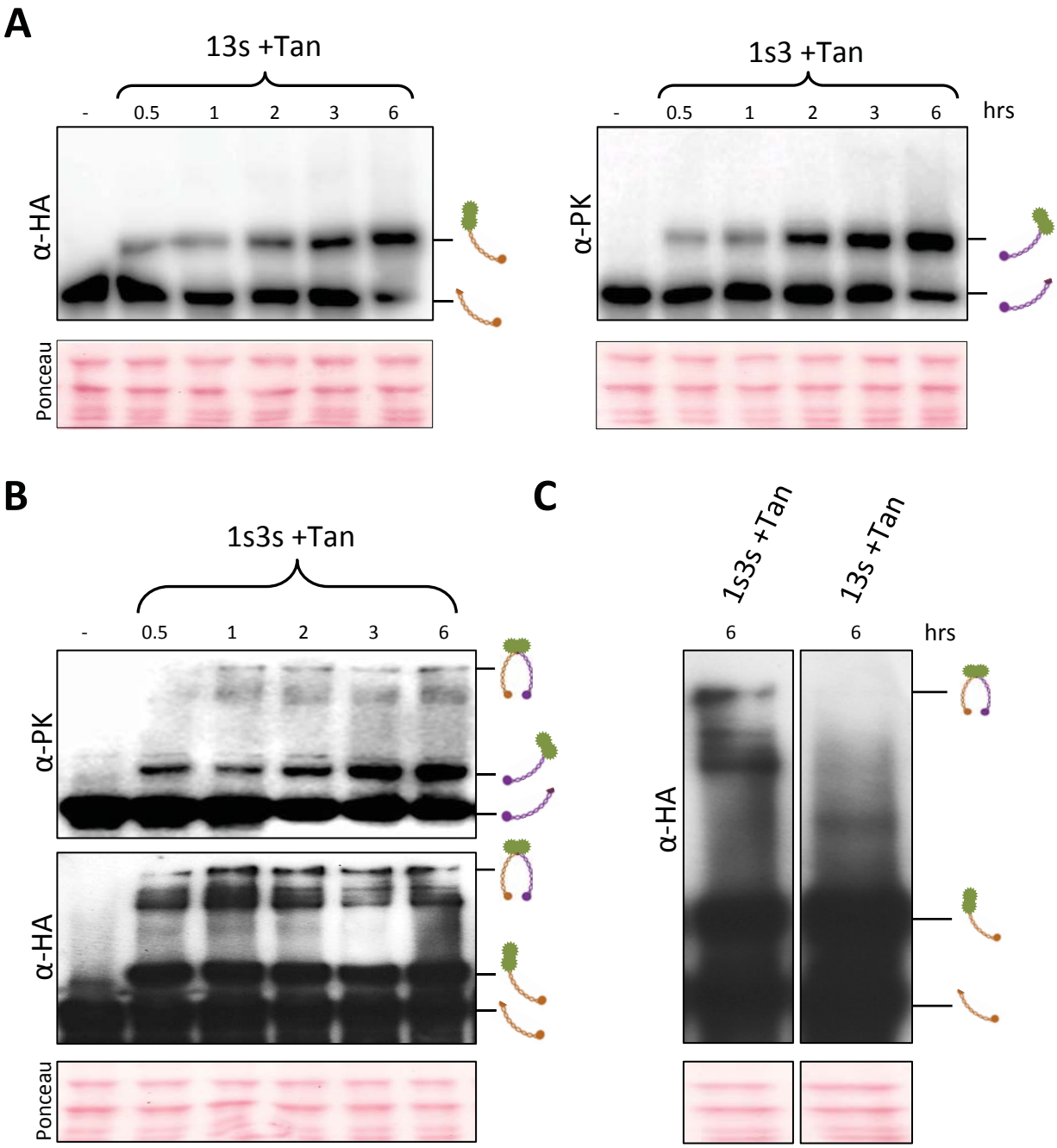


Figure III-16. SpyTagged Smc1 and Smc3 react equally with Tandem SpyCatcher.

(A) Western blots of lysates from the indicated strains to which 1μM Tan was added at 30°C. 1s(PK6) and 3s(HA6) show no considerable differences in the ability to react with Tandem SpyCatcher. Additionally, no higher molecular weight products indicative of homodimeric cross-links were formed, as shown in (C).

(B) High molecular weight, dimeric cross-linked species are seen to be formed by both 1s(PK6) and 3s(HA6)upon *in vitro* incubation of full length SpyTagged proteins in lysates with pure Tandem SpyCatcher.

(C) Strains with only one Smc containing a SpyTag do not form suspected dimeric cross-links. Here, comparison is made between the products formed in 1s3s and 13s. After incubation of lysates with Tan, 13s lacks the slowest migrating band, suggesting that cross-linking only occurs within SpyTag-containing Smc complexes, and not between separate complexes. The two samples depicted were treated, blotted and imaged identically, however intervening bands have been removed for clarity.

DISCUSSION

The topological encirclement of sister chromatids is widely accepted as the most well supported model for how cohesin facilitates bi-orientation. Currently, however, there is no consensus on how the interfaces of the tripartite ring might be modified to allow the association and dissociation of the complex from DNA. As such, new techniques that can be applied to study cohesin's distinctive activity as a structural DNA concatenase are invaluable.

Here, the chemical biology of the SpyTag-SpyCatcher system, as developed by the Howarth lab (Zakeri et al., 2012), was adopted as a potential tool to study cohesin. The unique properties of this particular cross-linking technique were, in theory, ideal for studying the chromosomal association of an otherwise unperturbed population of cohesin. The *S. cerevisiae* hinge domains were chosen as a model interface for the validation of SpyCatcher as a prospective tool. Indeed, initial experiments with isolated hinge domains revealed a promising efficiency. Although some inter-hinge complex reactions were observed in addition to those within hinges, it was reasoned this might be an artefact of the *in vitro* conditions. Both SpyTag and SpyCatcher could be expressed without issue in *S. cerevisiae.* Surprisingly, the formation of heterodimeric cross-links, whilst demonstrable, was very slow to occur. Single cross-linking of hinges to monomeric SpyCatcher was comparatively efficient. In both cases, the levels of SpyCatcher constructs were seen to be in excess at all times, and thus expression of the cross-linker was unlikely to be the rate-limiting step. As the cross-linking reaction is essentially a two-step process, it

was possible that intermediate, partially reacted species might cross-link with SpyTags on other cohesin complexes (depicted in Fig. III-17). It was shown, however, that dimerisation and cross-linking is not forced to occur *in vivo* where it would not otherwise do so.

Viability has been reported even with a very large quantized reduction in the amount of cellular cohesin; that is to say, functionality is maintained at just 13% of wild type cohesin levels (Heidinger-Pauli et al., 2010a). It is therefore difficult to reconcile the findings that the proportion of cross-linked 1s3s heterodimer appears so low, yet causes lethality. Closer examination revealed an apparent disappearance of Smc proteins over time, suggesting cross-linking was occurring but leading to proteolytic degradation of the fusion. It is speculated that the protein-protein cross-links formed by Tandem SpyCatcher between Smc1 and Smc3 lead to their degradation. Covalent closure of the Smc1/3 hinge domain has not been observed in live cells. Nonetheless, replacement with heterologous dimerisation domains has been reported without degradation (Gruber et al., 2006). As these are highly stable with nanomolar affinity, the constitutive closure of the hinge is unlikely to cause degradation *per se*. Furthermore, SpyCatcher and SpyTag are not proteolytic targets themselves, as individually and together following reconstitution, they appear stable. Testing cross-linking between other protein partners such as the model FRB-FKBP12 pair might elucidate whether stability is indeed an issue caused by covalent linkage of multiple proteins via isopeptide bonds. The other possibility is that the methods to assess the true extent of cross-linking have been inadequate. It was thought that TCA-precipitated, cross-linked protein

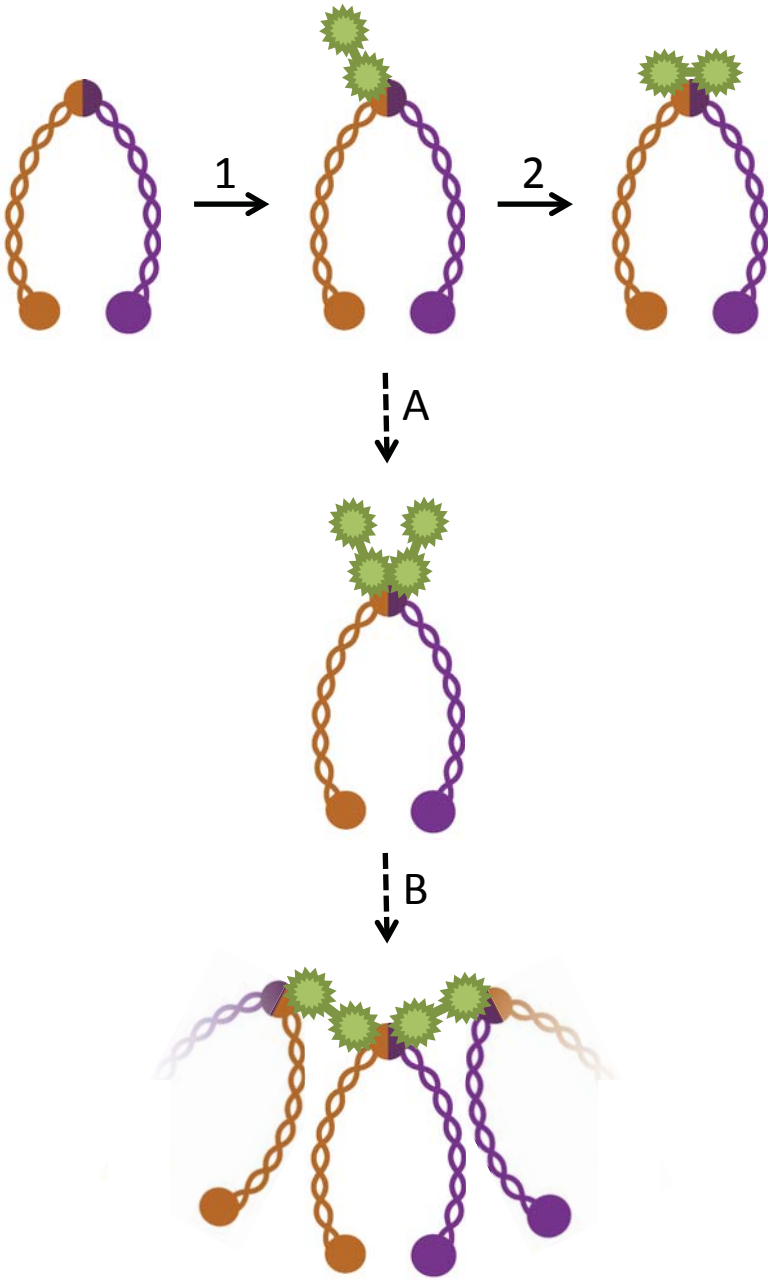


Figure III-17. Tandem SpyCatcher cross-linking is a two-step process.
To cross-link two SpyTagged proteins, one domain of Tandem SpyCatcher must react first (1). The adjacent, free SpyTag is then presumably favored for reaction with the other SpyCatcher domain (2). It is possible, however, that a separate molecule of Tandem SpyCatcher can react with the free SpyTag first (A). Free SpyCatcher moieties may plausibly react with SpyTags on other Smcs (B), however no evidence of these unspecific cross-links were observed *in vivo*.

might not be equivalently resolubilsed in SDS sample buffer. This would result in a misrepresentation of the true ratio of cross-linked to free protein species produced *in vivo*. This proves unsatisfactory, however, as lysing cells with beads directly in SDS sample buffer did not affect ratios. The specific lethality caused by closing the hinge with this system cannot be fully explained on available evidence. Regardless, it warrants further investigation, which was limited here by time constraints. An inherent problem with assessing efficacy by cross-linking the hinge is that it represents a loss of function and is therefore more open to interpretation. A more unequivocal assessment of SpyCatcher efficacy may be had by exploiting recent data concerning the Smc3-Scc1 interface to insert SpyTags at appropriate positions therein. If SpyCatcher can suitably cross-link this interface, then this would provide a gain of function, namely the ability to bypass *Eco1* mutation, e.g. *eco1-1*. Even if the current efficiency of cross-linking is too low to support this growth, then randomly mutagenizing SpyCatcher may create suppressors of *eco1-1,* representing a version of the cross-linker enhanced for use in *S. cerevisiae.*

Finally, there are alternative approaches to using a Tandem SpyCatcher cross-linker. For example, a recently reported system meets the specified requirements for conditional cross-linking (Rutkowska et al., 2011). Using only minimally disruptive insertions, just 12 amino acids in length, proteins can be covalently cross-linked with a dimeric biarsenic compound that has the added advantage of exhibiting fluorescence upon binding. The cross-link is reversible using reducing agents. This technique and others like it will be important for

future progress in understanding the association and dissociation of cohesin from chromosomes.

CHAPTER IV

Biochemical analysis of Pds5

Biochemical analysis of Pds5

INTRODUCTION

The means by which cohesin dissociates from chromatin non-proteolytically remains an unanswered question. In the previous chapter, a technique was adopted to try and address the fundamental nature of this removal, i.e. by what interface does the complex allow the exit of DNA. This approach was unsuccessful. The process remains of interest from another perspective: how are the regulatory factors themselves promoting release?

As discussed, the protein Wapl is identified as being directly involved in releasing activity. The dissociation of cohesin from DNA by Wapl appears to be an activity not reliant on this protein alone, but the combined action of several other factors. One of these is the Pds5 protein, with which Wapl has been shown to bind (Keung et al., 2006; Rowland et al., 2009; Shintomi and Hirano, 2009). This suggests Pds5 has an important role in promoting the removal of cohesin, both throughout the cell cycle as well as during the metazoan prophase pathway (Shintomi and Hirano, 2009). Although critical to this dissociating process, Pds5 is also required at the same time to build and maintain cohesion. Pds5 is required to facilitate the acetylation of Smc3 by the acetyltransferase, which in turn renders the complex refractory to Wapl activity. This alone, however, is insufficient for cohesion, as inactivating Pds5 in G_2/M leads to a loss of cohesin's stable association with chromosomes (Tóth et al., 1999; Vaur et al., 2012). The

Pds5 protein thus serves a number of distinct functions that at times work in opposition to one another.

Pds5 is a HEAT repeat protein

The large size and predicted secondary structure of Pds5 suggests it has an important role in the regulation of cohesin. Containing no conserved domain indicative of enzymatic activity, it is instead calculated to be highly helical (Panizza et al., 2000). Specifically, these helices are thought to arrange in a repeating structural motif consisting of interacting pairs in a hairpin-like orientation, called a HEAT repeat (Andrade and Bork, 1995). These then stack together into a super helix. HEAT repeats or the similar ARM motif repeat are present in several other large cohesin-associated proteins including Scc2, Wapl, and also Scc3 (Neuwald and Hirano, 2000, Andrade et al., 2001). The HEAT repeats are thought to serve as a scaffold upon which interacting partners can bind and may thus be important for establishing the in vivo architecture of the cohesin complex.

The interaction between the cohesin ring and Pds5 is most likely of lower affinity. Immunoprecipitation (IP) of the ring via tags on any of the core subunits shows a ready co-IP of the Scc3 subunit, whereas Pds5 appears at a sub-stoichiometric level that is reduced further by higher salt concentrations (Mc Intyre et al., 2007; Sumara et al., 2000). Nevertheless, additional factors in vivo might strengthen this interaction. In budding yeast, it was seen that Wapl was not essential and WaplΔ mutants are viable, whereas the same was not true for Pds5 deletion (Rowland et al., 2009). This suggests that although Pds5

interacts with Wapl, it must also bind the ring independently for its role in promoting cohesion. This is consistent with evidence showing that Scc1 is required for the interaction, principally the N-terminus, both in budding yeast and other organisms (Shintomi and Hirano, 2009; Kulemzina et al., 2012; Chan et al., 2012). Another study used a FRET-based assay, revealing that while no signal was detected between fluorophores concurrently fused to Scc1 and Pds5, there was significant, albeit weak, FRET between the Smc1 and Smc3 hinge domains and Pds5 (Mc Intyre et al., 2007). This raises the interesting possibility that Pds5 regulates the hinge association as well as that between Smc3 and Scc1, in combination with Wapl.

Insight into these varying functions of Pds5 might be gleaned from its structure, which is currently unverified. What makes Pds5 a particularly attractive target for structural determination is the fact that it seems to have at least three separate, and at times opposing, roles. It is required for the acetylation of Smc3 leading to establishment, for the releasing activity of Wapl, and for the maintenance of cohesion during G2/M. In vertebrates, it is also the site at which sororin, a cohesion maintenance factor, associates with the complex (Nishiyama et al., 2010). After Scc2, budding yeast Pds5 is the largest cohesin-associated protein known, and like Scc2, it has been estimated to interact with multiple subunits to form a protein-protein interaction scaffold. Therefore, Pds5 is likely a pivotal protein that in combination with other subunits undergoes conformational change to regulate the dynamics of cohesin. Herein, it was pursued as a target for crystallisation after optimising its purification. While unyielding to these attempts so far, the protein obtained

could be used to substantiate two of its critical interactions, with the N-terminus of Scc1, and with Wapl.

RESULTS

Expression and purification of Pds5 from S. cerevisiae.

Due to its considerable size (1277 amino acids and 148kDa), it was thought most amenable to over-expression in its natural environment in S. cerevisiae; as translating foreign proteins of such size in E. coli is often problematic. Furthermore, the purification of Pds5 from yeast had been described previously (Rowland et al., 2009; T. Nishino, pers. comm.). The constructs previously employed a FLAG-tagged protein under control of the hybrid PGK1-GAL1 promoter on a 2μ vector with LEU marker. Surprisingly, in a protease-deficient mutant strain, the expression was seen to be very low upon induction with galactose, which necessitated very large culture volumes for downstream handling. As detailed in the materials and methods, cells were broken and cleared lysates incubated with anti-FLAG agarose to capture Pds5-FLAG. The elution was enriched in Pds5-FLAG that ran largely as a single peak on gel-filtration, with some minor aggregates (Fig. IV-1). Purification through this method was prohibitive due to the cost of reagents (raffinose and anti-FLAG affinity matrix) and the difficulty in processing such large volumes of culture with sufficient speed to avoid degradation or aggregation. These are all important considerations for producing the large amounts of protein required for crystallisation. Therefore, *E. coli* was explored as a potential expression host for ease of handling and economy.

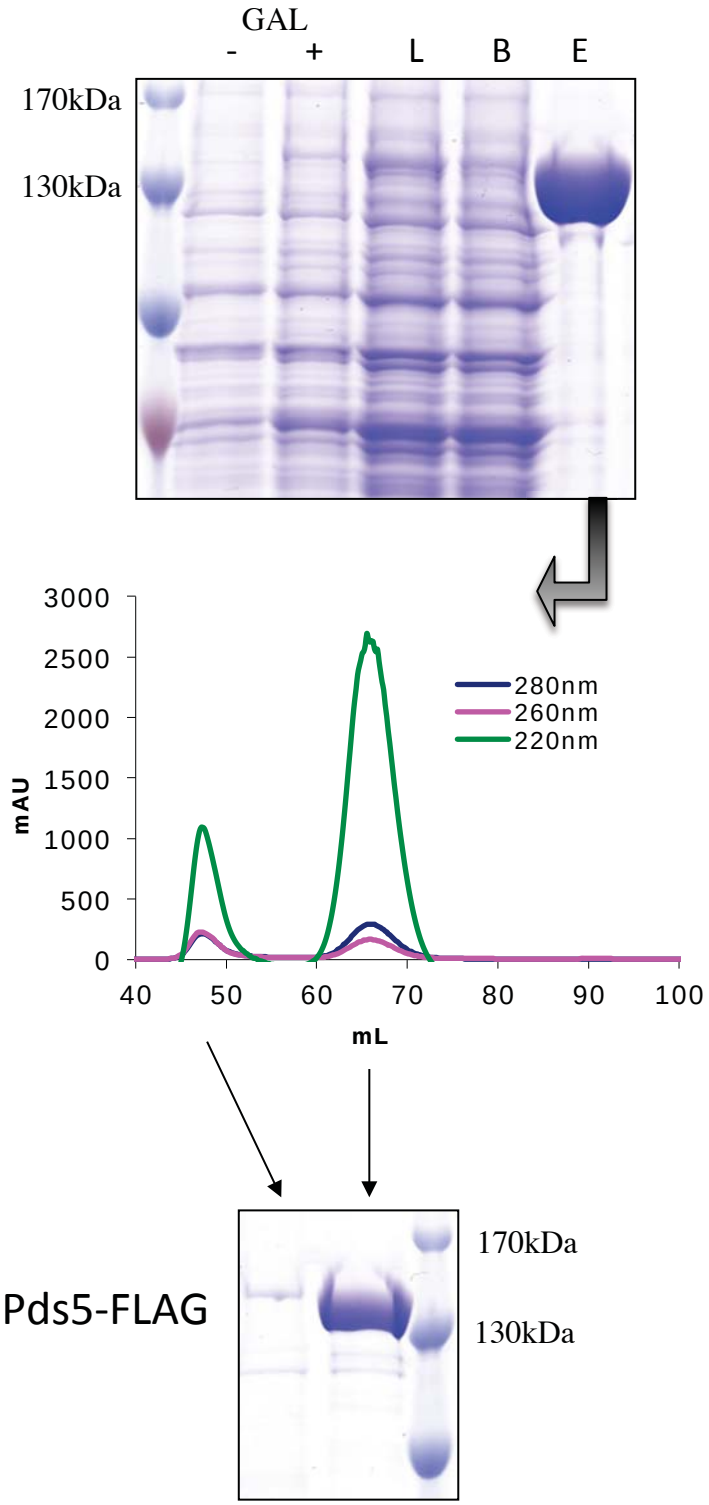


Figure IV-1. Expression and purification of Pds5 in *S. cerevisiae*.
Protease-negative strains carrying a plasmid for Pds5-FLAG expression and LEU-selective growth were grown to saturation in –LEU medium and then shifted into YEP + raffinose for 3hrs before induction with galactose for 6hrs. Pds5-GLAG was expressed at a low level, but could be enriched from a large volume of culture by affinity purification followed by gel-filtration. This yielded a homogenous, pure preparation.

Expression and purification of Pds5 from E. coli

Previous pilot tests had revealed a low expression of Pds5 in *E. coli* (not shown), even in strains designed for production of difficult heterologous proteins. However, it was reasoned that codon-optimisation of the sequence would prove beneficial to minimise the requirement for rare tRNAs in *E. coli.* Thus, a sequence of Pds5 was synthesised (Epoch Lifesciences) codon-optimised for expression in *E. coli.* In an expression plasmid, this was transformed into three different commonly used laboratory strains. To compare the expression to non-codon-optimised (WT), the relevant construct was tested in parallel. Induction was accompanied by the appearance of two distinct but specific bands separable by SDS-PAGE (Fig. IV-2A). It was found that codon-optimisation had a profound effect in enhancing the expression of FLAG-Pds5-His_6. This was seen even in Rosetta Blue and BL21 DE3 RIPL strains, which both carry an additional plasmid to supply the less common tRNAs needed to translate heterologous proteins. In this situation, compared to codon-optimisation, these expression systems are evidently inadequate to fully compensate for uncommon codon usage. Initial induction tests were performed at 37°C, which seemed superior in terms of protein level compared to conditions of lower temperate and longer duration (Fig. IV-2B). This however, may come at the expense of decreased solubility and this was assessed by examining the proportion of Pds5 present in the supernatant fraction of lysates. Increased expression at 37°C did not compromise solubility.

To optimise expression further, a time-course monitoring the target protein levels was performed. Coomassie staining of whole cell extracts at each

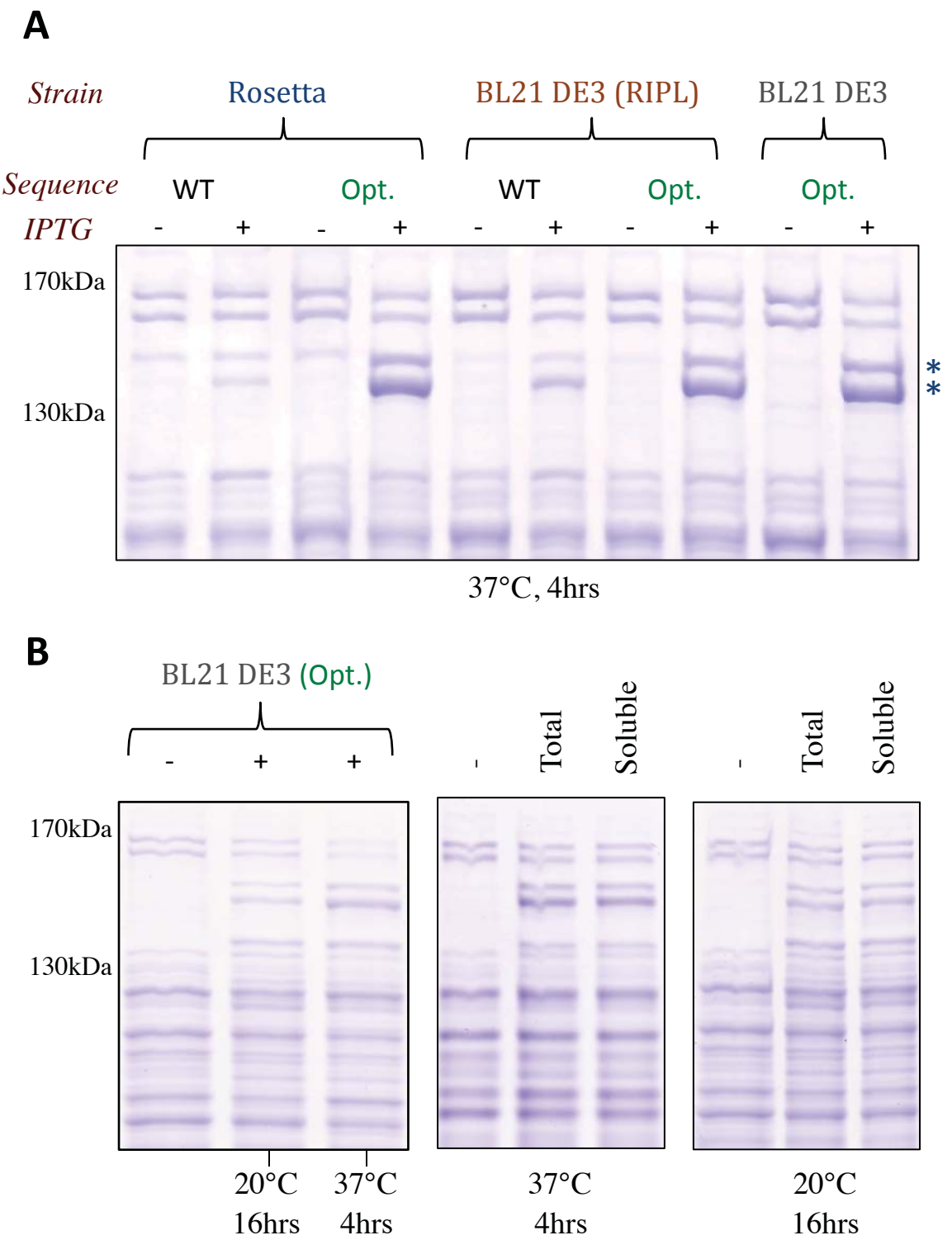


Figure IV-2. Comparison of wild-type versus optimised codon sequence for Flag-Pds5-His$_6$ induction in *E. coli* protein-expression strains.

(A) Rosetta, BL21 DE3 (RIPL), and BL21 DE3 strains were transformed with pET21a plasmids with either WT *S. cerevisiae* Pds5 sequence or the equivalent with codon usage optimised (Opt.) for *E. coli*, as indicated. Strains were induced to express protein with 1mM IPTG for 4hrs at 37°C and WCEs were compared by SDSPAGE. In each strain tested, Opt. plasmids expressed at a much higher level than WT. * Pds5

(B) Expression and solubility were compared between overnight induction at 20°C and short-term induction at 37°C for the optimally-expressing BL21 DE3 strain. Cells were lysed by sonication and the supernatant was used as a measure of the soluble fraction. Induction at 37°C for 4hrs was superior without compromising solubility.

time point revealed maximal induction within 1-2hrs (Fig. IV-3). The two bands were verified to be Pds5 by western blot against the His_6 and FLAG epitopes. The lower band, due to its large size and absence of anti-His signal, represents a truncation close to the C-terminus confirmed by mass spectrometry. This was thought to be attributable to either mistranslation or cleavage of the full-length protein. As levels of full-length protein appear to diminish beyond 4hrs induction, this may be representative of cleavage taking place. Having established ideal conditions for the expression of Pds5, its large-scale purification was performed as described in the materials and methods. Briefly, induced cultures were lysed, sonicated and cleared by centrifugation before incubation with beads for affinity capture of the His-tag. A final gel-filtration step was employed to separate any aggregated material or unspecifically-bound proteins.

The value of codon-optimisation for expression in *E. coli* is made apparent by a comparison of the yields from each system tested (Fig. IV-4A). Using the best possible parameters for expression, sub-cloned protein constructs could be purified to high concentrations from *E. coli*. The exact amino acid sequences were chosen on the basis of predicted secondary structure and were at sites corresponding to breaks within the HEAT repeat motifs and helices. These constructs were produced with a view to obtaining useable crystals for X-ray diffraction. Despite pure preparations, neither full length Pds5 or fragments thereof were able to crystallise (Fig. IV-4B). A number of additional constructs were designed but these either failed to express, were insoluble, or aggregated and eluted in the void peak of the gel filtration column. In each construct, the N-

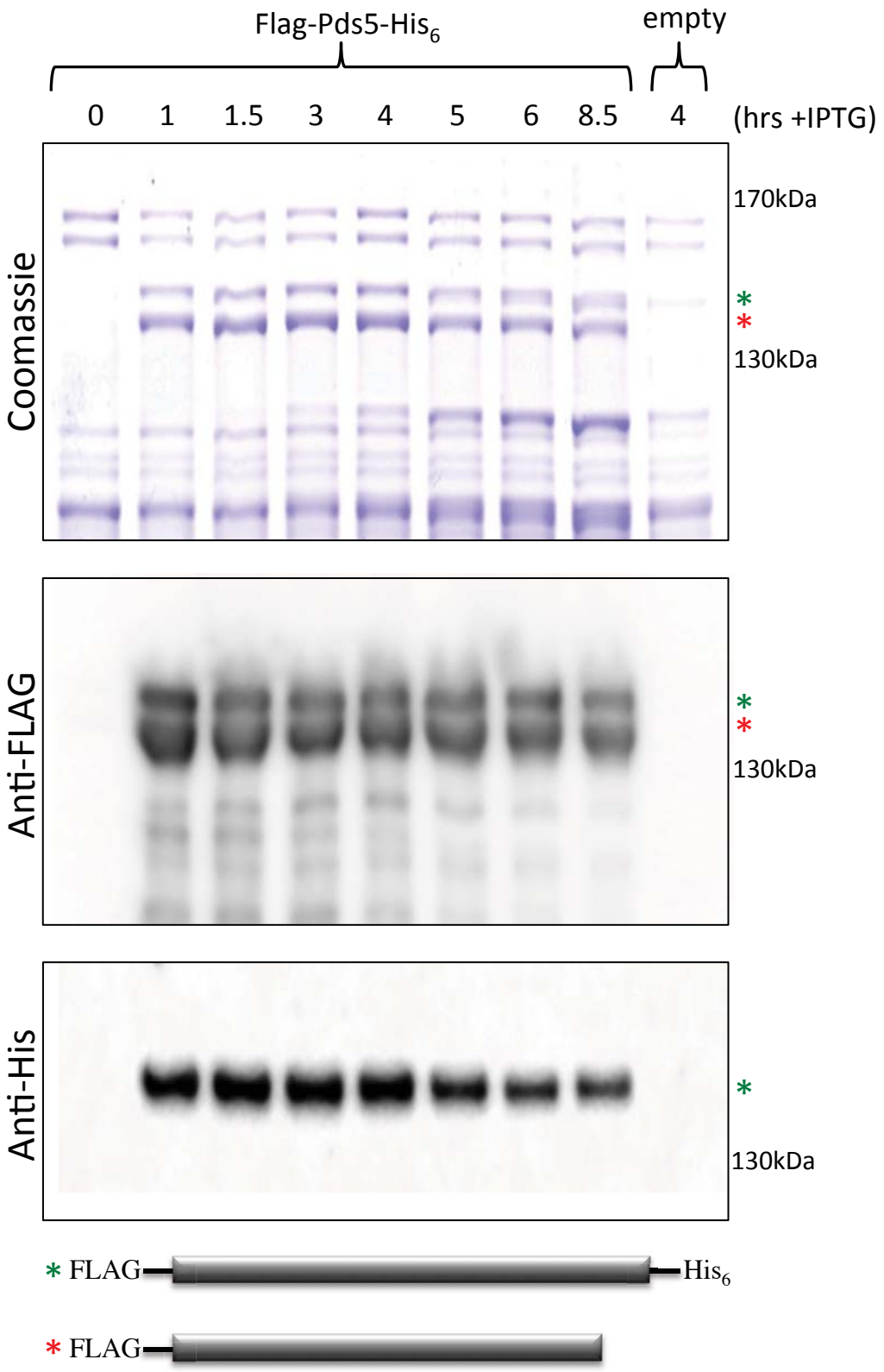


Figure IV-3. Time-course of Flag-Pds5-His$_6$ expression at 37°C.
BL21 DE3 carrying pET21a:FLAG-Pds5-His$_6$ (Opt.) was grown to log phase and induced with 1mM IPTG and samples were taken at the indicated time-points. Coomassie-stained SDS-PAGE revealed maximal induction within 1-2hrs. The two bands** were verified to be Pds5 by western blot against the His$_6$ and FLAG epitopes. The lower band, due to its large size and absence of anti-His signal, represents a truncation close to the C-terminus. This was thought to be attributable to either mistranslation or cleavage of the full-length protein. Levels of full-length protein appear to diminish beyond 4hrs induction.

A

Flag-Pds5-His_6

Organism	Purification	Yield µg per L of culture
E. coli (WT)	His_6, gel filtration	15
S. cerevisiae	α-FLAG, gel filtration	80
E. coli (Opt.)	His_6, gel filtration	265

B

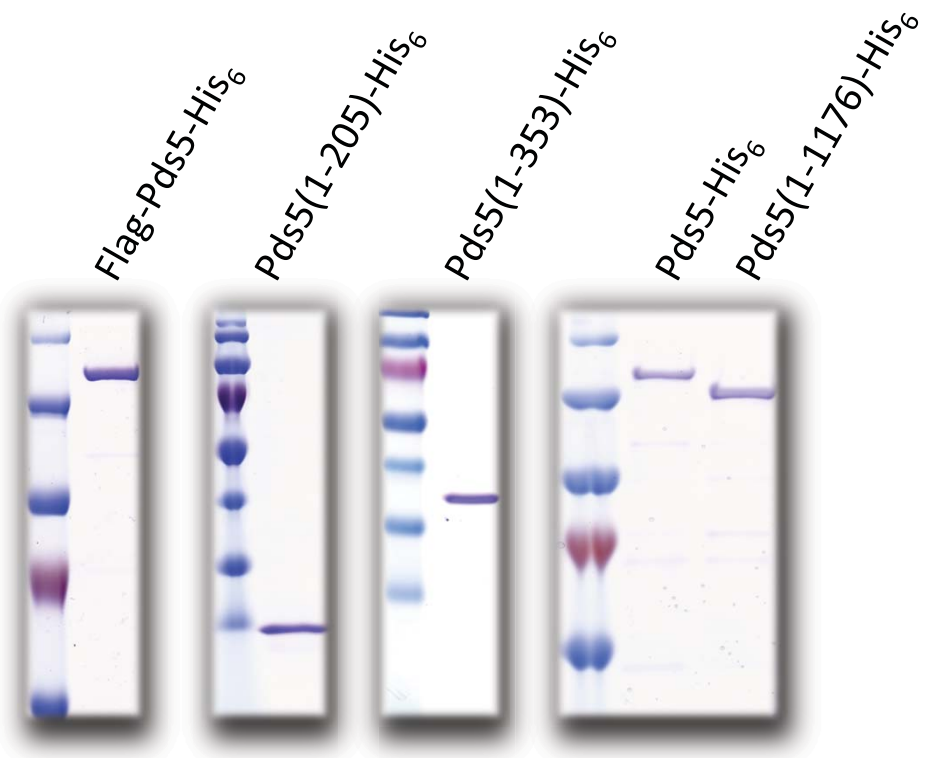


Figure IV-4. Pds5 purification from *E. coli*.

(A) The final quantity of protein achievable was far greater using codon-optimised sequences. Compared to wild-type sequence, combining codon-optimised constructs with suitable expression conditions was able to increase yields ten-fold. Additionally, the ease of handing of *E. coli* versus *S. cerevisiae* as a host allowed greater yields.

(B) Constructs able to be purified to high concentrations from *E. coli*. The exact amino acid sequences were chosen on the basis of predicted secondary structure and were at sites corresponding to breaks within the HEAT repeat motifs and helical structure.

terminal sequence of Pds5 was retained as this is most conserved and has been demonstrated as necessary for the interaction with Wapl *in vivo*. The C-terminal truncation (1-1176) removed a region that was predicted to be disordered, poorly conserved and highly dense in acidic and basic residues. These factors might have inhibited crystallization, but eliminating this domain was not productive in this regard.

Pds5 protein-protein interactions

Pre-existent evidence of an *in vivo* association between the hinge and Pds5 was addressed using protein purified here (Pds5) and previously (Smc1/3 hinge domains, Chapter II). A standard binding assay detecting complex-dependent shifts in size by gel-filtration failed to detect any such shift when Pds5 and the hinge domains were mixed (Fig. IV-5).

The interaction between full length Pds5 and purified MBP-Wapl was examined using a pull-down assay. Pds5 has been established as binding to Wapl previously (Kueng et al., 2006; Rowland et al., 2009). Recently the N-terminal domain of Pds5 has been suggested as the critical region for this interaction. Both were tested and found to weakly associate with MBP-Wapl (Fig. IV-6).

To verify an interaction between Pds5 and the N-terminus of Scc1, fluorescence polarization was chosen as an assay. This employs a fluorescein-conjugated peptide (in this case Scc1) as the probe for binding to a larger partner, as described in Figure IV-7. Fluorescence polarization was chosen as the assay due to the fact that until recently the N-terminal domain of Scc1 was

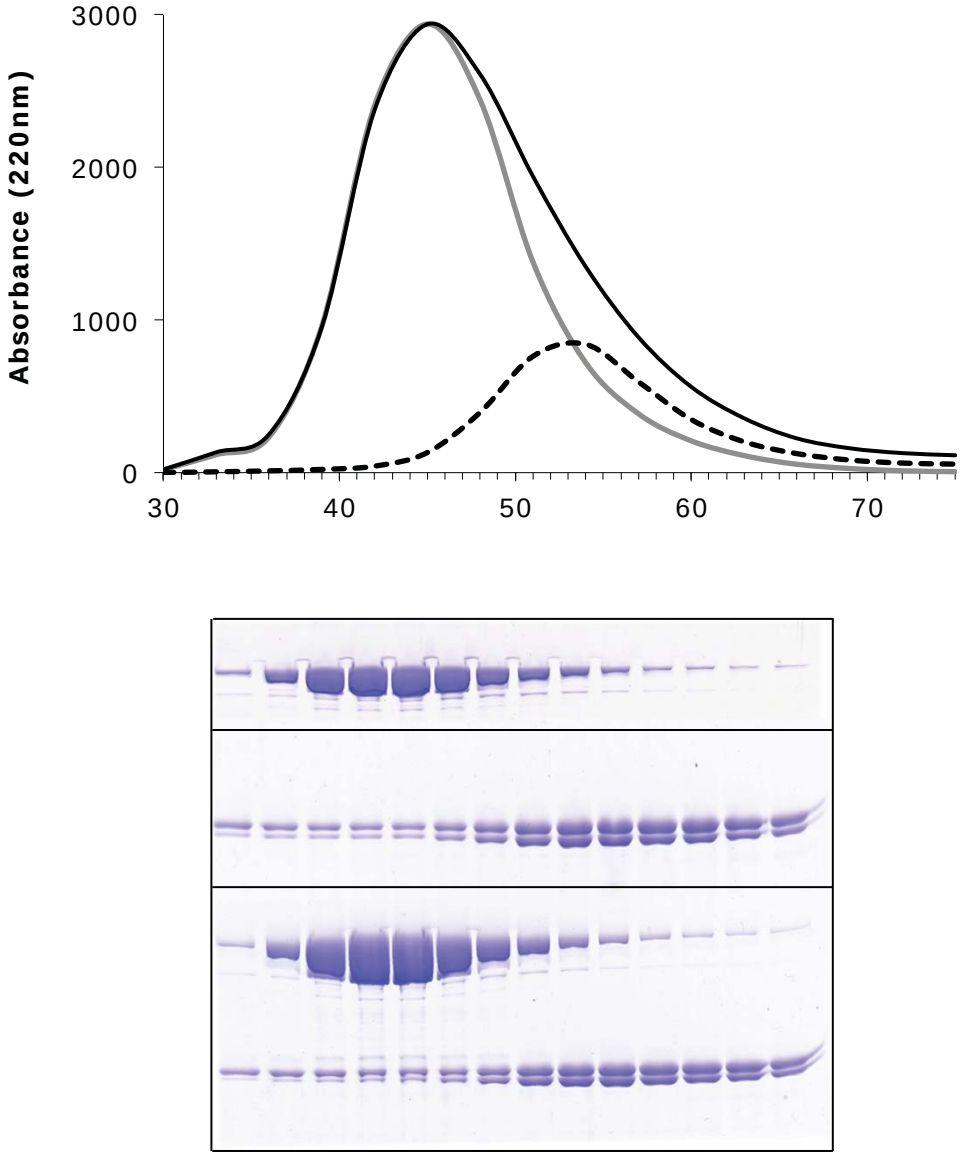


Figure IV-5. Pds5 does not directly bind the Smc1/3 hinge.
Reports of an interaction between Pds5 and the hinge *in vivo* were addressed using the purified proteins obtained here (Mc Intyre et al., 2007). Equimolar amounts of Pds5 and Smc1/3 hinge complex were incubated together prior to injection onto an analytical gel filtration column. A comparison of the chromatograms of individual proteins and their elution within each fraction with that of the interaction experiment showed no discernible difference suggesting no complex is formed.

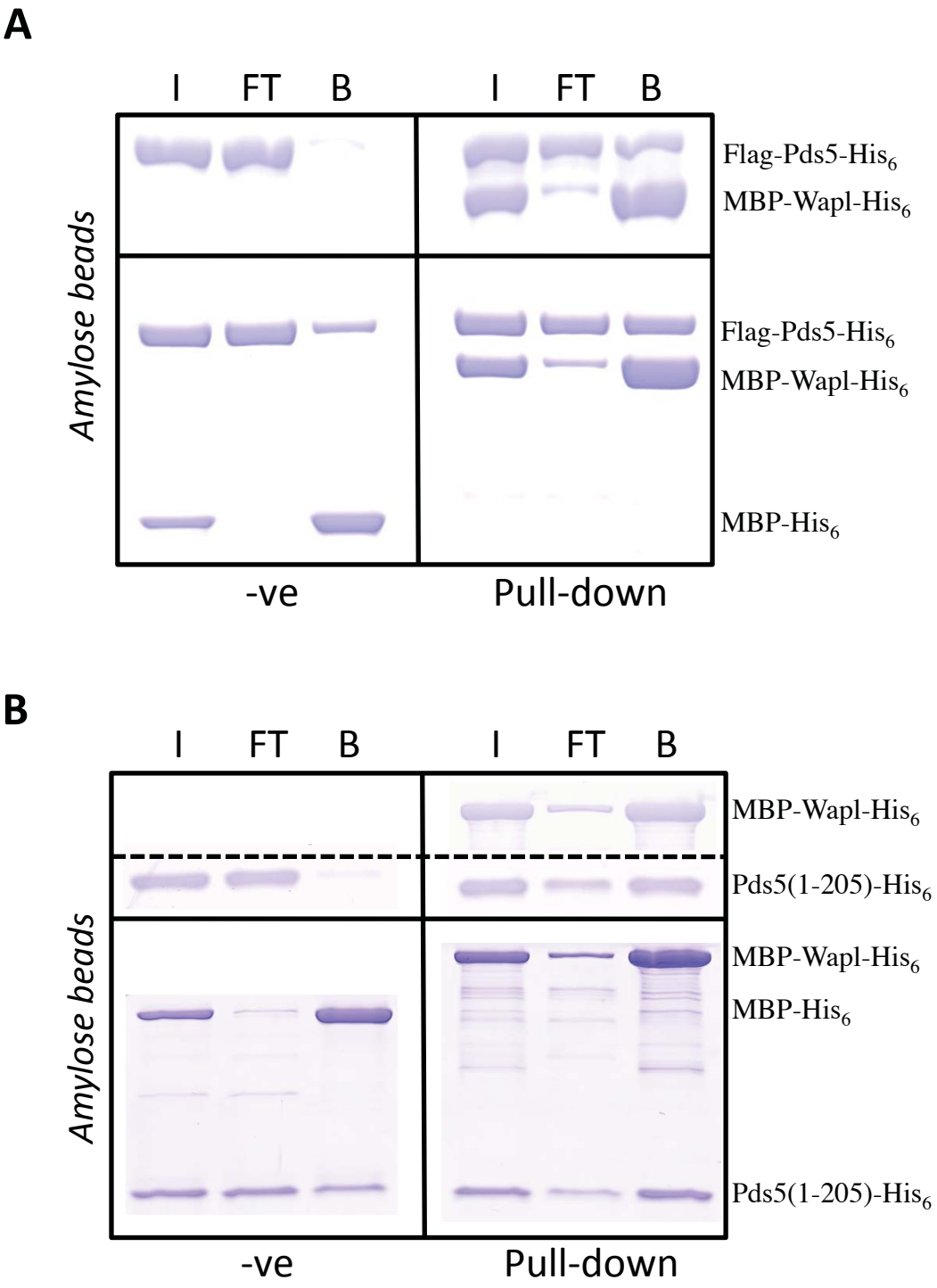


Figure IV-6. Pull-down of Pds5 by Wapl.
Purified MBP-Wapl was assessed for its ability to pull-down Pds5 both full length (A) and an N-terminal fragment (B). In the top panel of each, binding to the beads alone was assessed, while in the bottom panel, the potentially unspecific interaction with MBP was addressed by including MBP-His_6 as part of the negative control. The MBP domain itself was found to contribute to background binding of Pds5.
Conditions: Buffer 150mM NaCl, 50mM Tris 7.5, 1mM ß-ME. 2hrs at 4°C.

unable to be isolated. A short peptide was soluble, however, and thus most amenable to study *in vitro*. The sequence of this peptide corresponds to a region of high conservation amongst fungi that, on the basis of work by others, was implicated in binding Pds5 (Fig. IV-8A). Chan et al. showed that this region (and specifically the residue V137) was necessary for co-localisation of Pds5 and cohesin *in vivo* (2012). Likewise, surrounding residues were found to cross-link to Pds5 when replaced with a photo-reactive artificial residue (Chan et al., submitted manuscript). The binding was confirmed here *in vitro*, as demonstrated by an increase in the polarisation signal of the fluorescein-linked peptide with increasing concentrations of Pds5 (Fig. IV-8B). Fluorescence polarization occurred to a much lesser extent when a mutant peptide – identical but for the critical single amino acid substitution, V137K – was used in the same assay. Furthermore, the wild-type peptide did not appear to associate with protein indiscriminately, as no such signal was observed when BSA was tested for binding with the labelled probe.

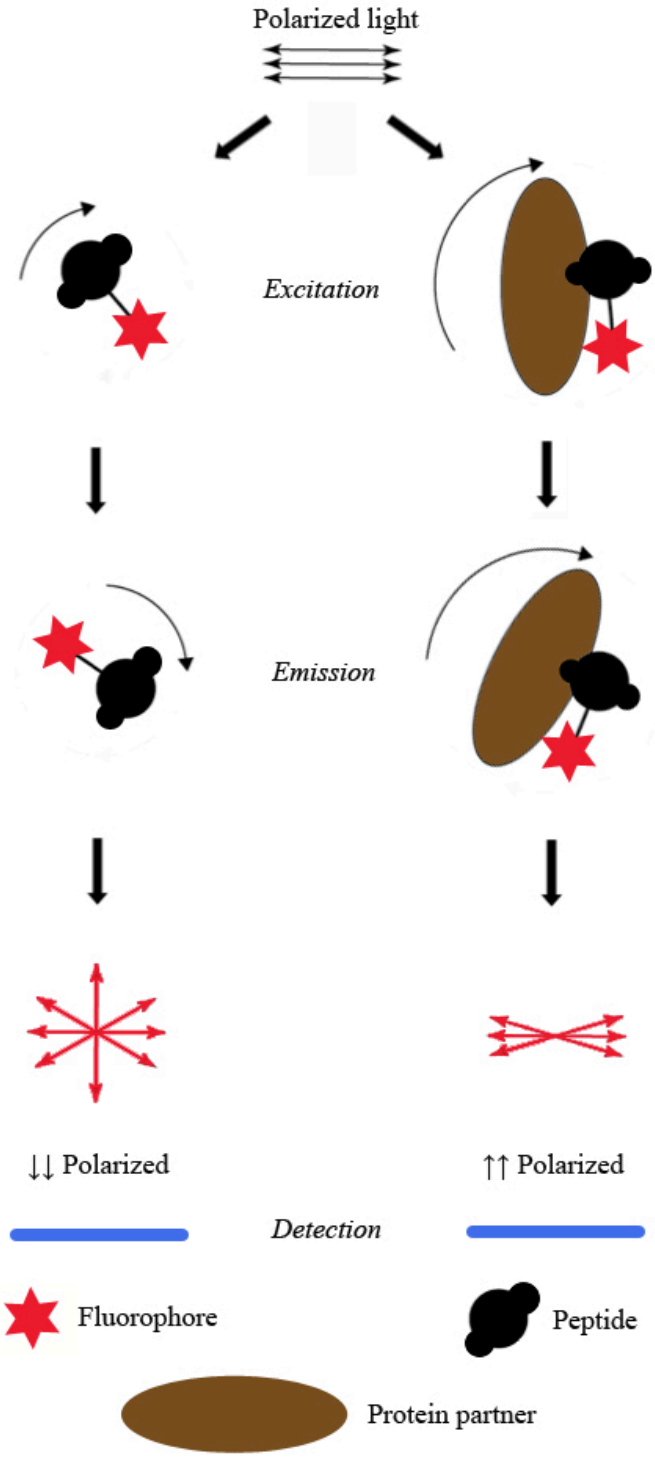


Figure IV-7. Fluorescence polarization assay
Fluorescence polarization allows protein interactions to be detected based on changes in the speed of rotation of a fluorescently-conjugated peptide upon binding to a protein partner. Excitation of a static fluorophore with polarized light leads to emission of polarized light, however, rotation of a fluorophore in solution during the excited state lifetime scatters this signal. Rotational movement is reduced by binding to a larger protein and thus a greater amount of polarized light is detected.

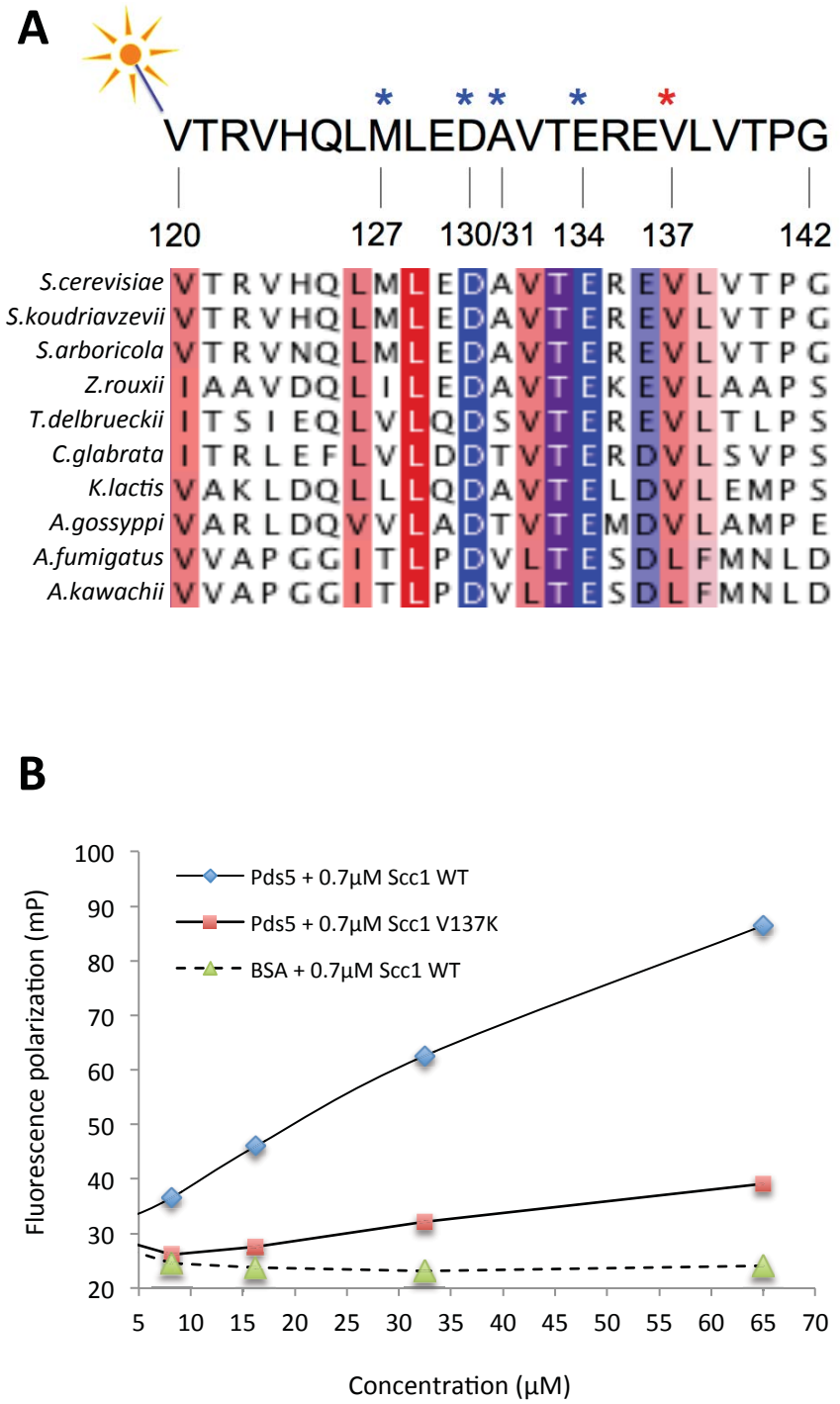


Figure IV-8. The conserved N-terminal region of Scc1 binds to Pds5

(A) A region highly conserved amongst fungi was identified as being required for recruitment of Pds5 *in vivo* and a single amino acid mutation of V137 to K abolished this (K-L. Chan; Chan et al., submitted manuscript). Photo-crosslinking to Pds5 could be detected at the * indicated residues (T. Gligoris). Wild-type and mutant peptides corresponding to this region were synthesized and coupled to fluoroscein for use in the fluorescence polarization assay.

(B) Serial dilutions of Pds5 showed a large concentration-dependent increase in the polarized light signal when mixed with wild-type peptide (Scc1 WT) but not mutant peptide (Scc1 V137K). The same concentration of an unrelated protein (BSA) was unable to produce polarization.

DISCUSSION

Relatively little is known about the function of Pds5. Its predicted structure is likely to be a central determinant in the variety of functions it appears to perform. Thus, elucidation of the structure of Pds5 was an aim of this work. Purification of this protein was optimised, allowing for a large number of truncations to be prepared, in addition to homogeneous intact full-length protein. To date, none of the constructs have yielded crystals. These protein preparations have been employed, however to demonstrate *in vitro* binding to two of its crucial partners.

The binding of Pds5 to Scc1 (as characterised below) and in turn the binding of Wapl to Pds5, are central structural requirements for the anti-establishment or releasing activity and the maintenance of cohesion. How Pds5 might also lead to acetylation is less clear. Direct recruitment of Eco1 is a potential mechanism, although an *in vitro* interaction with purified Eco1 was not seen here in a pull-down assay (not shown).

An interaction between Pds5 and Wapl was verified and maintained, albeit somewhat weakened, despite the truncation of all but the first 205 amino acids of Pds5. The fact that neither full-length nor the N-terminal fragment (1-205) of Pds5 are pulled down by Wapl in a 1:1 ratio is consistent with the fact that Wapl turns over rapidly and is never found stably bound to pericentromeric cohesin (Chan et al., 2012). That said, a more stoichiometric interaction has been reported with metazoan proteins, which may reflect a number of differences in the regulation of Wapl activity in these organisms (Shintomi and Hirano, 2009).

The fluorescence polarization assay was used to show the specific binding of an N-terminal fragment of Scc1 with Pds5. A similar probe containing the V137K mutation demonstrated the high specificity of this interaction. Given a sufficient range of protein concentrations, the binding of the fluorescent probe can be saturated and the resultant curve can be used to calculate the affinity of the interaction. The fluorescence polarization assay is suitable for studying interactions as the technique eliminates the potential complications of using immobilised proteins. Having established its validity here, it may also be used to pinpoint the site within Pds5 required to bind Scc1.

Crystallography

To date, no crystals have formed from the constructs tested. It is exceedingly difficult to predict the impact of the numerous parameters at play, nevertheless, there are a number of strategies that can be employed here to aid with attaining suitably diffracting material. They are summarised as follows:

- The presence of affinity tags.

Different tags vary in their capacity to negatively impact the quality of crystals. Bucher et al. illustrated the influence of different tags using *Pfu*MBP as a crystallographic model protein (2002). The FLAG-tag allowed equivalent crystal growth and diffracted to the same resolution. Despite being the same length (8 amino acids), the Strep II-tagged *Pfu*MBP failed to crystallise whatsoever. These are both small peptide tags that can nevertheless strongly influence protein behaviour. Larger tags are even more likely to cause problems and this is not just limited to effects on the tendency to form crystals. In fact, as

was seen in the pull-downs between Pds5 and Wapl, binding was abolished outright when Pds5 was bead-bound via its N-terminal FLAG-tag. The N-terminal region of Pds5 binds Wapl at an as-yet-unknown site. This site could also be at the N-terminus of Wapl and thus close to the MBP-domain. Therefore, the interaction between the two might be sterically hindered in the particular arrangement created when the binding site is nearer to the affinity beads. That is to say, the large MBP group on Wapl prevents its access to the N-terminus of Pds5 anchored at the spatially crowded surface on the affinity bead. If true, this serves to highlight the limitations of the pull-down technique for establishing protein-protein interactions. Regardless of whether the goal is the study of binding or the acquisition of X-ray diffractive material, complete removal of affinity tags is optimal. Of course, the proteolytic removal of an affinity domain necessitates a further step in the purification sequence and comes at the expense of yield; in the case of Pds5, this was limiting until its expression was optimised.

- Binding partners:

Another frequently effective technique for obtaining crystals is the inclusion of a binding partner, even if the affinity is weak, as the interaction will be enhanced in the high concentration environment used to produce crystals. This can enhance crystal packing and bury protein surfaces that interfere with the process. Importantly, the binding studies performed here using Pds5 might facilitate the acquisition of co-crystals of fragments with its partner proteins: Scc1 and/or Wapl.

- Sequence composition:

When a protein proves particularly resistant to crystal growth, changing the construct can make the critical difference. One or more of the following modifications can achieve this:

1) Truncation of N- and C-termini.
2) Removal of predicted regions of disordered secondary structure and low conservation.
3) Changing sequence entirely to an ortholog from another organism with high sequence homology.

Approximately 40% sequence identity and 70% similarity is desirable as a minimum to ensure that the sequence differs enough whilst remaining structurally relevant should crystals be produced (J. Löwe, pers. comm.). Recent success using this approach was had with structural determination of Wapl from *A. gossypii* and of Scc3 from *Z. rouxii* (M. Singleton, unpublished results; M.B Roig, unpublished results). As such, *Z. rouxii* Pds5 (~60% identity and ~80% similarity) is currently being trialled as a candidate.

A combination of the above processes might be necessary for obtaining a structure of Pds5. Achieving useful material for diffraction is a considerable challenge, but will be nevertheless worthwhile given the potential it provides for understanding the basis of the disparate roles played by Pds5.

CHAPTER V

Summary, discussion, and future directions

Summary, discussion, and future directions

SUMMARY

The original goal of this work was to characterise the cohesin complex loading and releasing mechanisms by examining the biochemical requirements for these processes. Although the identity of a chromosomal cofactor could not be assigned, the loading reaction was found to necessitate engagement of the conserved ABC-like ATPase domains in an ATP-dependent manner as evinced *in vitro.* Likewise, notions of direct effects of Smc3 acetylation on ATP binding and NBD engagement were discredited. Furthermore, a direct and unassisted, stable, physical association of hinge domains with dimerised NBDs was shown to be an unlikely conformational intermediate of a reaction that brings about hinge opening, one of the two suspected mechanisms for regulating the association of cohesin with DNA. The other interface is assumed to be between Smc3 and Scc1.

A novel protein-protein cross-linking system was adapted for use in *S. cerevisiae*, with a view to (1) confirming this well-founded role of hinge dissociation in topological entrapment, and (2) validating the Smc3-Scc1 interface as the recently conjectured exit gate. Preliminary SpyTag-SpyCatcher cross-linking kinetics were promising *in vitro.* Furthermore, SpyCatcher could be expressed within *S. cerevisiae* and possessed no intrinsic toxicity. The SpyTag-SpyCatcher system appeared considerably less efficient *in vivo*. The time-scale required for Smc1-Smc3 cross-linking complicated the interpretation

of the specific lethality seen in this situation; proteolysis of cross-linked heterodimers may be a contributing factor. The SpyCatcher system was thus deemed unsuitable in its current format for investigating the process of interface dissociation *in vivo*. Nevertheless, the chemistry of SpyCatcher affords a useful tool for probing protein-protein interactions and could be modified for other purposes such as targeted substrate degradation. Creation of a heterotypic version could in theory improve cross-linking efficiency.

Finally, the cohesin complex-associated protein Pds5 was chosen for biochemical study. Pds5 has been either speculated or shown to be involved in all of the aforementioned processes, potentially modulating the hinge and NBD-kleisin interfaces. Acting as a regulatory node in cohesin function, it represents an attractive target for structural studies. Its purification was optimised allowing for high yield from *E. coli*. An interaction with the N-terminal region of Scc1 was confirmed *in vitro* as the critical site at which the protein binds the cohesin complex. Likewise, Wapl, responsible for inducing cohesin release from DNA, relies on the conserved N-terminal region of Pds5 for binding. All attempts to obtain useable crystals were unsuccessful. Nevertheless, recent developments (vide infra) will provide an opportunity for insight into its activity, and indeed the regulation of cohesin's association with DNA in general.

DISCUSSION & FUTURE DIRECTIONS

Cohesion is controlled by an ever-growing list of processes and factors. These share in common the ability to affect, at some level, the propensity of the cohesin complex to associate with DNA. In this regard the loading factor Scc2/4

has a positive effect by recruiting DNA to chromosomal loci, whereas Wapl negatively affects association and promotes removal (Nasmyth, 2011). The equilibrium between these two opposing forces is shifted when cohesion is generated during DNA replication; acetylation neutralises the releasing activity of Wapl (Chan et al., 2012). In vertebrates, it appears the balance is altered once more during prophase when all cohesin, save for a centromeric fraction, reverts to a Wapl-sensitive form and is removed from chromosome arms (Gandhi et al., 2006; Kueng et al., 2006).

The cohesin loading reaction is poorly understood. Fully addressing the mechanisms will likely necessitate complete reconstitution of this reaction *in vitro*, where each factor can be systematically manipulated. Both biochemical and biophysical assays may be employed. For example, AFM imaging of the sequence of steps might be feasible, as has been reported for the MRN complex (Moreno-Herrero et al., 2005). Finally, studying these processes in multiple species will allow progress towards a unified model of the steps and requirements for loading. These are clearly long-term and ambitious goals and will require many incremental steps forward. Although I fell short of identifying the chromosomal receptor protein for Scc2/4 and cohesin recruitment, recent progress in studying loading is encouraging. For example, Onn and Koshland document the conversion of cohesin associated with closed DNA from an easily dissociated form to one highly salt-resistant (>0.5M KCl; 2011). Importantly, this *in vitro* system mirrors requirements demonstrated *in vivo*, including the need for the Scc2/4 complex and ATP binding by Smc3. The salt-resistant mode of association takes time to arise, whereas the salt-sensitive form is bound from

the outset, pointing towards the occurrence of a stabilising reaction. This maturation cannot be reproduced when the DNA is linear. This is a further validation of the hypothesis that cohesin must be remodelled from a physically and transiently DNA-bound structure into one that is topological. For this transition, the combination of Scc2/4 binding and ATPase activity seems paramount. In separate work, the Scc2/4 complex has been demonstrated to interact with the cohesin complex at the hinge domain. Purified complex interacts solely with heterodimers of the Smc proteins and not either of the individual proteins (Bermudez et al., 2012). Given the large size of the Scc2/4 complex, it may interact at structurally separate sites within each Smc protein, or it may require association of an interface to form a complete surface with which to bind tightly. A recent scc4 *ts*-suppressor screen from our lab has corroborated the relationship between the loading complex and the hinge, showing that scc4 deletion can be completely bypassed by a single amino acid substitution in the Smc1 hinge that exacerbates its dissociation from Smc3 (M. Demidova and S. Dixon, unpublished data). Although budding yeast Scc2 and Scc4 proteins proved refractory to purification (not shown), sequences from other species may allow *in vitro* studies. Notably, a method for the purification of intact kinetochores has been developed, providing the means to investigate the contribution of this protein assemblage as well (Akiyoshi et al., 2010).

The body of data demonstrates quite convincingly that cohesin loading mandates hinge dissociation, although what generates the force required for this process remains uncertain. Analogy to ABC transporters suggests either

NBD engagement or ATP hydrolysis can power conformational change (Newstead et al., 2009).

Surprisingly, cohesin release from DNA is now thought to utilise a different gate within the ring structure, explicitly that of the Smc3-Scc1 binding site (Nasmyth, 2011; Chan et al., 2012). How this association is regulated is equally intriguing. Wapl is brought to the ring via an interaction with the conserved N-terminal domain of Pds5 (this work, and Chan et al., 2012) to promote DNA exit. Again, the biochemistry of this releasing activity is unknown. Binding between the Smc3hd and Scc1 might involve competitive binding with Wapl. Another possibility that cannot be discounted is that the anti-establishment/releasing activity works to change the activity of the ATPase domains. In both the bacterial MukBEF and Smc/ScpAB complexes, ATP-dependent engagement of the NBDs is incompatible with the kleisin binding to both simultaneously, at least in a symmetrical manner (Woo et al., 2009; Bürmann et al., submitted manuscript). Whether an analogous dissociation of the cohesin kleisin occurs is an interesting question. If so, then NBD engagement and/or ATP hydrolysis could represent the common engine to both loading and release of cohesin. The differential regulation of this activity by Scc2/4 and Wapl could account for the distinct effects. It was found that mutating the critical lysine necessary for bypass of Eco1Δ does not affect NBD engagement *per se*. This might not be surprising given the complex interplay of other proteins now known to be involved in the establishment process surrounding this interface. It does, however, rule out a direct allosteric effect on the capacity of the Smc3hd to bind ATP and co-operatively engage with the Smc1hd. This is

also consistent with the fact that DNA damage-induced cohesion is established without Smc3hd acetylation but with Scc1 acetylation instead (Heidinger-Pauli et al., 2009). Until very recently, the paucity of knowledge regarding the Smc3-Scc1 interface was the limiting factor in allowing study of the relationships between the kleisin, the ATPase domains, and pro- and anti-establishment factors; in vitro studies are of questionable physiological pertinence in the absence of all these elements. It has now been discovered that the critical binding site for the N-terminus of Scc1 is the coiled-coil of Smc3, rather than the NBD as suspected on the basis of the Smc1hd/Scc1-C structure (Gligoris et al., submitted manuscript). Despite consisting of homodimeric Smc complexes, bacterial condensins also appear to bind the corresponding kleisin in this asymmetrical manner (Bürmann et al., submitted manuscript). These discoveries have enabled the purification of a stably interacting complex between Smc3 and the Scc1 N-terminal region, including the conserved site required for recruitment to the ring (K.L. Chan), BPA cross-linking (T. Gligoris), and *in vitro* binding (this work) of Pds5. A crystal structure of the protein assembly at this interface might illuminate how the Smc3-Scc1 interaction is controlled.

If both loading and releasing activities were to be modulated via changes in ATPase activity, a cell cycle-dependent difference in the requirement for this activity would be evident. This would correspond to the shift from an unstable to a stable population of cohesin. Whether continued hydrolysis of ATP occurs during G_2/M (i.e. after establishment) is unknown and addressing this is technically difficult. Specific, conditional obstruction of the ATPase activity by

small molecule inhibitors is problematic, as the active site does not form a well-defined pocket. Thus designing or screening for analogues/mutations to block hydrolysis is complicated (Gregan et al., 2007).

Arumugam et al. showed that the C-terminal domain of Scc1 could stimulate ATP hydrolysis whilst the N-terminus could not, but this was demonstrated *in vitro*, without consideration of the role of the Wapl/Pds5/Scc3 subcomplex. The individual and combined effects of these proteins on ATPase activity and kleisin-NBD association are worth revisiting for study. Now that the components of these processes are obtainable, separable and beginning to be defined by their structural inter-relationships, deciphering the underlying mechanisms is a promising prospect.

The biology of cohesin is still in its infancy, but given the relatively recent discovery of this protein complex, our understanding has advanced at an extraordinary rate. The last decade or so has been punctuated by several profound breakthroughs, which together with steady increments of knowledge paint an elegant portrait of this chromosomal machine. Furthermore, the insight gained by studying cohesin is not applicable to this family member alone, but can be used to infer commonalities amongst all SMC complexes. Their roles are many, and ultimately, a greater understanding of these complexes will open further avenues of research with far-reaching theoretical and clinical outcomes.

CHAPTER VI

Materials and methods

Materials and methods

Plasmid construction for protein expression

Smc3hd (M1-E190, 18 amino acid linker, S1022-V1230) and Scc1-C (F451-A566) were fused to a His6 tag at the C-terminus in pET21 vectors (Novagen). Smc1hd (M1-Q177, 31 amino acid linker, N1036-E1225) was cloned into a pET28a vector (Novagen). The Smc1hd and Smc3hd constructs were designed and created by Dr Christian Häring. Dr Maurici Brunet made the Scc1-C construct. The hydrolysis and signature motif mutations (Smc1hdE1158Q, S1130R, Smc3hdE1155Q, S1127R) were introduced by site-directed mutagenesis PCR.

MBP-Smc1hg(S503-E681)-His6 was cloned into the pMALc2x vector (NEB) by Dr Ajay Mishra. The Smc3 hinge domain (T495-L695) was cloned into a pET28a vector with a 3xFLAG tag at the C-terminus by Dr Alexander Kurze. SpyTag sequences were introduced by overlap-extension PCR. For expression and purification from E. coli, SpyCatcher was cloned and purified by Jacob Fierer as described in Zakeri et al. 2012. For expression in yeast, SpyCatcher constructs were placed under the GAL1 promoter in the YIplac211 vector for integration at the URA3 locus. Wapl sequence was cloned into a pMALc2x vector along with His6 to create the MBP-Wapl fusion construct. Wild-type Pds5 sequence was cloned into a pET21a vector with a C-terminal His6 tag with or without an N-terminal FLAG tag. Codon-optimised Pds5 (Epoch Life Science) was synthesised and cloned as per wild type, as were all fragments tested (M1-I205, M1-E245,

M1-E353, M1-E533, M1-M625, M1-Q1081, M1-I1176, M625-T1277). For expression in S. cerevisiae, Pds5 was cloned with a C-terminal FLAG tag into the 2µ pPM231 vector containing a galactose-inducible hybrid GAL-PGK promoter with a TRP1 selection marker.

PCR conditions

Polymerases used were either Phusion® (New England Biolabs), Taq polymerase (Qiagen), or KAPA HiFi HotStart ReadyMix (Kapa Biosystems) as stated. Unless otherwise indicated, the final concentration of each forward and reverse primer used in reactions was 0.3µM.

Site-directed mutagenesis PCR

Plasmid DNA for mutagenesis was used at a final concentration of 50ng in the following 50µL mix:

SDM PCR mix	
Phusion HF reaction buffer	1x
dNTPs	200µM
Forward primer	0.5µM
Reverse primer	0.5µM
Phusion® polymerase	1µL (1 unit)
Plasmid DNA	50ng
DMSO	1.5µL (3%)
MilliQ®	to 50µL

The PCR conditions used were as follows:

Reaction parameters		
Initial denaturation	98°C, 30s	
Denaturation	98°C, 10s	
Annealing	Primer T_m, 30s	33 cycles
Extension	72°C, 30s/kb	
Final extension	72°C, 5mins	
Hold	4°C	

Overlap-extension PCR

Overlap-extension PCR was used to insert sequences within genes. Two PCR products (A and B) with approximately 20bp of overlapping homology (at the 3' end of product A and the 5' end of product B) were created in separate reactions using primer pairs containing the sequence for insertion, in addition to template specific homology. Each product was run on an agarose gel, purified, and mixed in a 1:1 molar ratio in a final PCR containing the forward primer from reaction A and the reverse primer from reaction B.

Yeast strains and growth conditions

Strains used and modified are derived from the W303 background and grown at 30°C unless otherwise indicated. Rich medium consists of YEP supplemented with 2% glucose (YPD) or 2% raffinose (YPR) ± 2% galactose (YPRG). Liquid cultures were agitated at 200rpm (Multitron Standard, Infors HT).

Yeast transformation

Following overnight growth to saturation, strains were diluted and grown in liquid culture at 30°C. Cultures were grown to log phase, corresponding to an OD_{600nm} between 0.3-0.8 and approximately 4 OD units were harvested by centrifugation and washed with 10mL of 0.1M lithium acetate. Following centrifugation, pellets were resuspended in 360µL transformation buffer:

Transformation buffer	
LiAc	0.1M
Salmon sperm DNA	1.5mg/mL
PEG-3500	33% (w/v)
DNA construct	1µg
MilliQ®	to 360µL

In the case of single integration with integrative vectors targeted to endogenous auxotrophic marker loci, approximately 200ng of endonuclease pre-digested plasmid DNA was used as described. Samples in transformation buffer were heat-shocked at 42°C for 20mins before plating on selective drop-out plates. Where the marker for successful transformation was a heterologous drug resistance cassette, heat-shocked samples were recovered in YPD for 3hrs at 30°C prior to plating on YPD agar supplemented with the appropriate compound.

Genomic integration

For integration of plasmids (YIplac128, YIplac204, YIplac211) at auxotrophic marker loci (*LEU2*, *TRP1*, *URA3*, respectively), restriction digest was performed with *EcoRV* (New England Biolabs) to create linear DNA for recombination upon transformation. Following extraction of genomic DNA from colonies grown on selective media, integration at the desired locus was confirmed by PCR with specific primer sets. The correct size and presence of PCR products indicated either successful single-copy integration, tandem-copy integration or no integration.

For C-terminal tagging of genes at their endogenous loci, pFA6a plasmids (with varying heterologous selection markers and epitope tags) were used as templates in PCRs producing 50bp of homology either side of the ORF stop codon, flanking the insertion cassette. Purified PCR material was transformed as normal.

For insertion of sequences within genes at their endogenous loci, plasmids were constructed containing the gene of interest with a selection marker 3' to the ORF, followed by additional homology to the chromosomal sequence. Insertions were created by overlap-extension PCR and cloned into the plasmid at unique restriction sites flanking the point of insertion. Integration was carried out by transformation of PCR amplification from the plasmid template with at least 1kb of homology upstream of the insertion site, and an equivalent length of homologous sequence after the marker cassette. Primers flanking the insertion

site were used to confirm that the desired sequence was introduced in the genomes of positive transformants.

Tetrad dissection

Diploid strains were grown on rich medium agar before sporulation on SpoVB plates at 25°C until tetrads were visible by microscopy. A patch of cells was suspended in dissection buffer:

Dissection buffer	
Sorbitol	1M
Zymolase 20T (Seikagau Biobusiness)	1mg/mL

Tetrads were incubated for 15mins at 30°C and 10µL was spread onto YEP plates for microscopic dissection (Singer MSM).

Yeast genomic DNA extraction

Strains were patched onto rich medium agar and allowed to grow overnight before being transferred into 200µL of SCE buffer:

SCE buffer	
Sorbitol	1M
NaAc	0.1M
EDTA pH 7.0	60mM
ß-mercaptoethanol	0.12M
Zymolase 20T	1.5mg/mL

Cells in SCE buffer were incubated at 37°C with shaking at 900rpm for 1hr. Spheroplasting was confirmed microscopically before the addition of 200µL SDS lysis buffer:

SDS lysis buffer	
SDS	1% (w/v)
Tris	50mM
EDTA pH 8.0	50mM

Samples were incubated at 65°C for 5mins before the addition of 200µL of 5M KAc to induce precipitation. Precipitate was centrifuged at 16000x *g* for 10mins and 350µL of the supernatant was added to 800µL of ethanol. Samples were centrifuged at 16000 x *g* and the ethanol supernatant was discarded. The pellet was dried briefly at 65°C to remove remaining ethanol before resuspension of the nucleic acid in 400µL of deionised water. Concentration of DNA was determined by absorbance at 260nm using the NanoDrop-1000™ (Thermo Fisher Scientific) and samples were stored at -80°C until analysis.

Trichloroacetic acid protein precipitation

Approximately 4 OD units of each strain undergoing exponential growth (e.g. 10mL of culture at an OD_{600nm} of ~ 0.4) in liquid culture were harvested by centrifugation and washed in 1mL of 20% TCA before resuspension in 200µL of 20% TCA in screw-cap 2mL polypropylene microvials. A bed volume of 100µL acid-washed glass beads (Sigma) was added before cell lysis was carried out at 4°C using a FastPrep®-24 automated homogenizer (MP Biomedicals). Samples were agitated 6 x 30s at 4.5 M/s. Lysis was confirmed microscopically and

precipitated protein was suspended in a further 100µL of 20% TCA and separated from the beads. Glass beads were washed with 400µL of 5% TCA and supernatants were combined before pelleting the insoluble protein by centrifugation. Supernatant was discarded and the pellet was resuspended in SDS sample buffer supplemented with 0.5M Tris pH 8.8 to neutralise acidity.

Preparation of whole cell extracts

Approximately 4 OD units of each strain undergoing exponential growth (e.g. 10mL of culture at an OD_{600} of ~ 0.4) in liquid culture were harvested by centrifugation and washed in 1x PBS. Samples were resuspended in 200µL of extract buffer:

Extract buffer	
NaCl	150mM
Tris pH 7.5	50mM
NP-40	0.1% (w/v)
Protease inhibitor cocktail (Roche)	1 tablet/50mL
EDTA	2mM
DTT	1mM
PMSF (Fischer Scientific)	1mM

Upon transfer to screw-cap 2mL polypropylene microvials. A bed volume of 100µL acid-washed glass beads (Sigma) was added before cell lysis was carried out at 4°C using a FastPrep®-24 automated homogenizer (MP Biomedicals). Samples were agitated 6 x 30s at 4.5 M/s. Lysis was confirmed microscopically before centrifugation at 16000x *g* for 15mins to sediment insoluble cellular

debris. Supernatants were decanted and the concentration of protein was measured.

Quantitation of lysate protein concentrations

The concentrations of protein in extract samples were ascertained through the Bradford protein assay (Bradford, 1976). A standard curve was created using 1mg/mL bovine serum albumin (BSA) serially diluted to a range of concentrations (2µg/mL, 5µg/mL, 10µg/mL, 15µg/mL, 20µg/mL, 30µg/mL) within which the concentration of proteins in extracts was predicted to fall. Bradford® reagent (dye concentrate; Bio-Rad) diluted 1:5 was added to 10µL of each standard to a total of 1mL and allowed to stand at room temperature for 5mins. The absorbance at 595nm was measured for each standard relative to a blank (10µL MilliQ® water, 990µL Bradford® reagent) using the Ultrospec 3300 pro spectrophotometer (Amersham Biosciences). A standard curve was generated by the spectrophotometer. An aliquot of each sample was diluted 1:10 with MilliQ® water to 10µL and made up to 1mL with diluted Bradford® reagent as described above for standards. Absorbance was measured relative to a buffer-only blank and the concentration of protein in extracts was calculated from the standard curve by the spectrophotometer.

SDS-PAGE and gel staining

Samples prepared as above were loaded onto Tris-glycine acrylamide gels, pre-cast NuPAGE® Tris-acetate gels (3-8%; Invitrogen), or pre-cast NuPAGE® Bis-Tris gels (4-12%; Invitrogen) and separated using appropriate current for electrophoresis. For coomassie staining, gels were washed 3x 5mins in

deionised water before incubation in Imperial protein stain (Thermo Scientific). For silver staining, the SilverQuest™ silver staining kit (Invitrogen) was used and manufacturer instructions followed.

Western blotting

Protein samples separated by SDS-PAGE were transferred to a 0.2µm PVDF membrane (Bio-Rad) using a wet transfer set-up. Transfer was assessed by ponceau staining, removed by incubation in phosphate buffered saline + 0.05% Tween 20 (PBST). The membrane was blocked for 1hr in PBST + 5% milk (w/v) before incubation for 1hr with the primary antibody diluted as below in PBST +5% milk:

Primary antibody	Origin	Dilution
Penta-His (Qiagen)	Mouse	1:2000
V5/Pk (Serotec)	Mouse	1:1000
HA (3F10; Roche)	Rat	1:500
PGK1 (Abcam)	Mouse	1:5000
FLAG (Sigma)	Mouse	1:1000

Blots were washed three times with PBST for 5mins to remove unbound primary antibody before incubation with HRP-conjugated secondary antibody (anti-mouse/rat IgG; Amersham) diluted 1:5000 in PBST +5% milk. Following 1hr of incubation with secondary antibody, membranes were washed three times more with PBST before a final wash in PBS. Bands were visualised using Immobilon™ chemiluminescent HRP substrate (Millipore).

Fluorescence-activated cell sorting

1mL of culture was spun down at high speed and resuspended in 70% (v/v) ethanol and stored at 4°C overnight. Samples were centrifuged and the pellet resuspended in 1mL of 50mM Tris pH 7.5 containing RNAse A (Roche) at a concentration of 200µg/mL. Cells were incubated for RNAse treatment at 37°C for >2hrs with agitation at 900rpm. Samples were pelleted and resuspended in 400µL of FACS buffer:

FACS buffer	
Tris pH 7.5	180mM
$MgCl_2$	70mM
Propidium iodide	50µg/mL
NaCl	190mM

Resuspended cells were sonicated with a Vibra-cell sonicator (VCX 130FSJ; Sonics & Materials Inc.) for 5secs at amplitude of 40% to eliminate cellular aggregation. Of this, 50µL was diluted into 1mL of 50mM Tris pH 7.5 in FACS tubes and analysis was carried out using the FACSCalibur flow cytometer (BD Biosciences) and 10000 events were measured per sample.

Live-cell imaging and analysis

Exponentially growing cells were aliquoted onto 2% agarose pads mounted on glass slides and all imaging was carried out at 25°C. Imaging of SpyCatcher-GFP expressing strains was performed using a spinning disk confocal system (PerkinElmer UltraVIEW) with an Olympus IX8 microscope. Volocity software

was used for image acquisition. For strains used to screen for Scc2 interaction partners, a wide-field Olympus IX 70 microscope coupled with Deltavision Core imaging system (Applied Precision) was used. The microscope was controlled via softWoRx software (Applied Precision) and image deconvolution was carried out using softWoRx. Kinetochore fusion strains were growing exponentially in YPD.

Transformation of E. coli

Competent *E. coli* strains, XL1 Blue (for cloning), BL21 DE3, and BL21 DE3 (RIPL; for protein expression; Stratagene) stored at -80°C were thawed on ice and 100µL was transferred to a pre-chilled 14mL round-bottomed falcon tube (BD Biosciences) on ice. Up to 50ng of plasmid DNA was added to the cell suspension, mixed gently on ice and allowed to stand for 30mins. Each aliquot was heat-shocked at 42°C for 40secs and returned to ice for 2mins before addition of 900µL of 2xTY medium. Transformants were allowed to recover at 37°C for 1hr before plating on 2xTY agar plates supplemented with antibiotic corresponding to the resistance marker carried in the transformed plasmid. Plates were incubated at 37°C for >12hrs.

Preparation of competent E. coli

For transformation with more than one vector, the protein expression strain BL21 DE3 was transformed as described and made competent for transformation with the second plasmid as follows. Strains were grown in LB medium containing the antibiotic selection for the transformed plasmid to saturation overnight at 25°C. This culture was diluted into 50mL fresh,

antibiotic free LB medium and grown at 37°C until the optical density reached a reading of OD_{600nm} = 0.4. At this point, cells were harvested by centrifugation at 3000 x *g* for 10mins at 4°C. The supernatant was discarded and the pellet was gently resuspended in 12.5mL of 100mM $MgCl_2$ pre-cooled to 4°C. This solution was mixed gently on ice for 5mins before centrifugation at 3000x *g* at 4°C for 10mins. The supernatant was again discarded and the bacterial pellet resuspended in 25mL of 100mM $CaCl_2$ pre-cooled to 4°C. This suspension was allowed to stand on ice for 20mins with occasional agitation before centrifugation at 3000 x *g* for 10mins at 4°C. Finally, the pellet was resuspended in 1mL of 85mM $CaCl_2$ supplemented with 15% (w/v) glycerol. Cells were aliquoted and stored at -80°C unless transformed immediately.

Processing of lysate from E. coli

Cultures containing were grown at 37°C until reaching an OD_{600} of 0.6, and expression was induced by addition of IPTG to a concentration of 1mM, for 12-20hrs at 20°C unless otherwise indicated. Cells were harvested and mixed with lysis buffer until a homogeneous suspension was achieved:

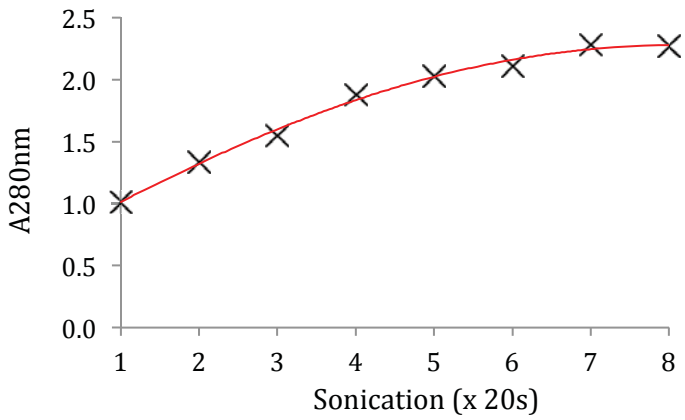
Lysis buffer	
NaCl	250mM
Tris pH 7.5	50mM
Protease inhibitor cocktail (Roche)	1 tablet/50mL
EDTA	1mM
ß-mercaptoethanol	1mM
Glycerol	5% (v/v)

In the case where downstream purification involves immobilized metal ion affinity chromatography (IMAC), EDTA was omitted from the lysis buffer.

Lysis of cells was achieved upon one passage through a cell disrupter (Constant Systems) at 20 kpsi and PMSF (Fisher) in isopropanol was added to the lysate to a final concentration of 1mM. Crude lysate was sonicated with a Vibra-cell sonicator (VCX 130FSJ; Sonics & Materials) for 2x 30s/50mL at amplitude 80% to reduce viscosity and shear DNA. Lysates were centrifuged at 20000x *g* in an Avanti J-26S XP centrifuge (Beckman Coulter) for 15mins to pellet cellular debris before transfer of the supernatant to 50mL polypropylene tubes (Nalgene) for further centrifugation at 50000x *g* for 90mins. Cleared lysates were transferred to 50mL Falcon tubes (BD Biosciences) for affinity purification.

Protein solubility assay

Small, 20-50mL cultures were grown under appropriate conditions, induced with IPTG, and harvested by centrifugation at 3000x *g* in 50mL Falcon tubes. Cell pellets were resuspended in 3 volumes of lysis buffer and transferred to ultracentrifuge tubes. Cell suspensions were sonicated in 20s intervals on ice to prevent overheating, at amplitude 80%. After each treatment, samples were centrifuged at high speed for 2mins. The absorbance at 280nm of clarified supernatant was measured using the NanoDrop-1000™ (Thermo Fisher Scientific). Once the absorbance no longer increased with subsequent bouts of sonication, lysis was deemed to be complete, as represented graphically below:



Finally, a sample of homogenate was taken as a total cell extract before high-speed centrifugation for 2mins. A sample of equal volume from the resulting supernatant was designated as containing soluble protein. Both samples were analyzed in parallel by SDS-PAGE to gauge the proportion of protein of interest in the soluble fraction relative to the total amount.

Affinity purification

Cleared lysates containing proteins of interest were incubated with a suspension of the appropriate affinity resin pre-equilibrated in lysis buffer. Approximately 1.5mL bed volume of Talon Superflow beads (Clontech) per 50mL of lysate was used for purification of His_6-tagged proteins. Amylose Resin High Flow (3mL bed volume per 50mL lysate; New England Biolabs) was used for binding MBP-tagged proteins, while for Strep-tagged proteins 3mL bed volume per 50mL lysate of Strep-Tactin Sepharose (IBA) was used. Flag-tagged proteins were bound with 2mL bed volume per 50mL lysate of anti-FLAG M2 affinity gel (Sigma). All lysates were incubated with beads for 1hr before

sedimentation at 700x *g* for 5mins. Supernatants were decanted and resins washed 3x in this manner with 50mL of lysis buffer there being transferred to chromatography columns (Bio-Rad) and washed once more with 2 column volumes of lysis buffer. In the case of IMAC with Talon Superflow beads (Clontech), wash buffer was supplemented with 10mM imidazole. Proteins were eluted with 2x the column volume of lysis buffer supplemented with the following eluents for each system:

Talon Superflow beads	250mM imidazole
Amylose resin	20mM maltose
Strep-Tactin sepharose	2.5mM d-desthiobiotin
Anti-FLAG agarose	100µg/mL 3XFLAG peptide

Gel-filtration chromatography purification

Following affinity purification and elution, up to 5mL of eluent was injected onto a Superdex 200 16/60 size-exclusion chromatography column (GE Healthcare) equilibrated with Buffer A.

The chromatography column was connected to the ÄKTApurifier 100 purification system controlled by UNICORN software (GE Healthcare). Peak fractions were collected and concentrated using Vivaspin columns (Sartorius Stedim Biotech) with a molecular weight cut-off of 10kDa, 30kDa, or 100kDa.

Concentration measurements of purified protein

Protein purified to homogeneity was quantified based on the absorbance at 280nm using the NanoDrop-1000™ (Thermo Fisher Scientific). In each case, measurements were made relative to a buffer-only blank and adjusted based on the calculated extinction coefficient (ExPasy ProtParam) and molecular weight to provide a reading in milligrams per millilitre.

Analytical gel-filtration for binding analysis

A Superdex 200 10/300 GL size-exclusion chromatography column (GE Healthcare) was used to investigate size, purity and/or dimerisation between pre-purified proteins constructs. For testing the dimerisation between Smc head constructs, proteins were first incubated together at room temperature for 15 minutes in buffer. 100µl of 30µM Smc1hd/Scc1-C and Smc3hd were loaded together in the presence or absence of ATP at 0.5mM in a buffer consisting of 200mM NaCl, 50mM Tris pH 7.5, 0.5mM $MgCl_2$. Peak fractions were collected and analysed by SDS-PAGE.

Pull-down assay

Equimolar amounts of protein (5µM) were mixed in Buffer A and allowed to stand at room temperature for 15mins. 20µL was taken as an input sample and 90µL was added to 90µL of prewashed (3x) affinity matrix 50% slurry in microcentrifuge tubes. Samples were incubated at 4°C for 2hrs with agitation before centrifugation at 300x *g* to pellet bead-bound material. Beads were

resuspended and washed 5x in Buffer A before a final resuspension in 45µL of 4x SDS lysis buffer. Samples were analysed by SDS-PAGE or immunoblot.

In vitro cleavage of SpyCatcher by TEV protease

Whole cell extracts were diluted to an equivalent concentration of protein in extract buffer and 30 units of AcTEV™ protease (Invitrogen) was added to each sample. Incubation at 4°C was continued for 2hrs before the addition of SDS sample buffer to stop the reaction.

Fluorescence polarization

Fluorescence polarization measurements were made using the PHERAstar FS platereader (BMG Labtech) with fluorescein polarization module. Protein 2-fold dilution series were pipetted in triplicate (20µL per well) in black, low protein-binding 384-well plates (Corning). Peptide concentration was 700nM and the receptor was 0µM-65µM. Gain was calibrated against the zero-point and the starting fluorescence was set to 23mP. Intensities were recorded using 200 flashes per point. The buffer used was 150mM NaCl, 50mM Tris 7.5, 1mM ß-mercaptoethanol.

Structure-based homology-modelling and sequence alignments

A structure-based homology model of the S. cerevisiae hinge domains was created using SWISS-MODEL (Peitsch et al., 1995; Arnold et al., 2006; Kiefer et al., 2009). Multiple sequence alignments were generated using the Clustal Omega program (www.ebi.ac.uk/Tools/msa/clustalo) and viewed and colored according to conservation using Jalview (Sievers et al., 2011).

Table of strains used

Strain	Genotype
17454	Mat A, prb1-1122, pep4-3, prc1-407, pM343-TRP Pds5-FLAG
17456	Mat A, leu2, ura3-52, prb1-1122, pep4-3, prc1-407, pM346-LEU2D Pds5-FLAG YEPlac195 URA
17675	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Chl4-TetR:hph, ADE2 wild type
17676	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Iml3-TetR:hph, ADE2 wild type
17677	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Ybp2-TetR:hph, ADE2 wild type
17678	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Ctf3-TetR:hph, ADE2 wild type
17679	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Mcm16-TetR:hph, ADE2 wild type
17680	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Mcm22-TetR:hph, ADE2 wild type
17681	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Ndc10-TetR:hph, ADE2 wild type
17874	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Ctf19-TetR:hph, ADE2 wild type
17875	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Mif2-TetR:hph, ADE2 wild type
17876	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Ctf13-TetR:hph, ADE2 wild type
17877	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Nkp1-TetR:hph, ADE2 wild type
17878	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Nkp2-TetR:hph, ADE2 wild type
17879	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Cep3-TetR:hph, ADE2 wild type
17880	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Spc25-TetR:hph, ADE2 wild type
17881	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Okp1-TetR:hph, ADE2 wild type
17882	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Nnf1-TetR:hph, ADE2 wild type
17883	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Nsl1-TetR:hph, ADE2 wild type
17884	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Mtw1-TetR:hph, ADE2 wild type
17887	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Srb7-TetR:hph, ADE2 wild type
17888	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc2-GFP:KanMX, Mtw1-RFP:KanMX, Spt5-TetR:hph, ADE2 wild type
16516	Mat A/Mat ∝, ura3::3XURA3 tetO112, Scc4-GFP:KanMX, Mtw1-RFP:KanMX, Scc2-TetR:KanMX, ADE2 wild type
19259	Mat A, Smc3-HA6::HIS3
19260	Mat A Smc1-PK6::KanMX
19261	Mat A/Mat ∝, Smc1-PK6::KanMX, Smc3HA6::HIS3 (heterozygous)

19507	Mat A, ura3::Gal1p-NLS-His6-Spycatcher::URA3
19508	Mat A, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19509	Mat A, ura3::Gal1p-NLS-His6-Spycatcher::URA3, leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19511	Mat A, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19648	Mat A, Smc1-PK6::KanMX, Smc3HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3
19649	Mat A, Smc1 L595(10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3-HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3
19650	Mat ∝, Smc1-PK6::KanMX, Smc3 S606(10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3
19651	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3
19652	Mat ∝, Smc1-PK6::KanMX, Smc3-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19653	Mat A, Smc1 L595(10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19654	Mat ∝, Smc1-PK6::KanMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19655	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)-PK6::KanMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19656	Mat ∝, Smc1-PK6::KanMX, Smc3-HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19657	Mat ∝, Smc1 L595(10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19658	Mat ∝, Smc1-PK6::KanMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3, leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19659	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)-PK6::KanMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19660	Mat ∝, Smc1-PK6::KanMX, Smc3-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19661	Mat ∝, Smc1 L595(10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3-HA6::HIS3, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19662	Mat ∝, Smc1-PK6::KanMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19663	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)-PK6::KanMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19776	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3
19778	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3

19779	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3
19780	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19782	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19784	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19844	Mat A, ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integration)
19846	Mat A, ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integration) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19848	Mat A, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant)
19849	Mat ∝, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19850	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant)
19852	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant), leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19854	Mat A, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant)
19856	Mat A, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19858	Mat ∝, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant)
19859	Mat ∝, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19860	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant)
19862	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19949	Mat ∝, Smc3HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant)
19950	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3-HA6::HIS3, ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant) leu2::His3p-Gal1/His3p-Gal2/Gal1p-Gal4::Leu2 (single copy)
19951	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3-HA6::HIS3 ura3::Gal1p-NLS-His6-Spycatcher::URA3 (multiple integrant)
19952	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant)

19953	Mat A, Smc3-HA6::HIS3, ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant)
19956	Mat ∝, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-NLS-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (multiple integrant) Gal1p-NLS-myc9-TEVprotease-NLS2::TRP1 (10-fold integrant by southern)
20130	Mat A, Smc1 L595 (10GSlinker-Spytag-10GSlinker)::hphMX, Smc3 S606 (10GSlinker-Spytag-10GSlinker)-HA6::HIS3 ura3::Gal1p-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (NLS removed)
20211	Mat A, ura3::Gal1p-His6-Spycatcher::URA3 (NLS removed)
20212	Mat A, ura3::Gal1p-His6-TandemSpycatcher(TEV-cleavable linker)::URA3 (NLS removed)

References

Aasland, R., Abrams, C., Ampe, C., Ball, L.J., Bedford, M.T., Cesareni, G., Gimona, M., Hurley, J.H., Jarchau, T., Lehto, V.P.*, et al.* (2002). Normalization of nomenclature for peptide motifs as ligands of modular protein domains. FEBS Letters *513*, 141-144.

Adolphs, K.W., Cheng, S.M., Paulson, J.R., and Laemmli, U.K. (1977) Isolation of a protein scaffold from mitotic HeLa cell chromosomes. Proc Natl Acad Sci U S A *74*, 4937-4941.

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

Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L.*, et al.* (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science *323*, 1718-1722.

Almedawar, S., Colomina, N., Bermúdez-López, M., Pociño-Merino, I., and Torres-Rosell, J. (2012). A SUMO-Dependent Step during Establishment of Sister Chromatid Cohesion. Curr Biol *22*, 1576-1581.

Ambudkar, S.V., Kim, I.-W., Xia, D., and Sauna, Z.E. (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*, 1049-1055.

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. J Cell Biol *156*, 419-424.

Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. Nat Genet *11*, 115-116.

Andrade, M.A., Petosa, C., O'donoghue, S.I., Müller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats1. Journal of Molecular Biology *309*, 1-18.

Andrulis, E.D., Guzman, E., Doring, P., Werner, J., and Lis, J.T. (2000). High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. Genes Dev *14*, 2635-2649.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195-201.

Arumugam, P. (2003). ATP Hydrolysis Is Required for Cohesin's Association with Chromosomes. Current Biology *13*, 1941-1953.

Bailis, J.M., Bernard, P., Antonelli, R., Allshire, R.C., and Forsburg, S.L. (2003). Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. Nat Cell Biol *5*, 1111-1116.

Bazett-Jones, D.P., Kimura, K., and Hirano, T. (2002). Efficient supercoiling of DNA by a single condensin complex as revealed by electron spectroscopic imaging. Mol Cell *9*, 1183-1190.

Beckouët, F., Hu, B., Roig, M.B., Sutani, T., Komata, M., Uluocak, P., Katis, V.L., Shirahige, K., and Nasmyth, K. (2010). An Smc3 acetylation cycle is essential for establishment of sister chromatid cohesion. Molecular Cell *39*, 689-699.

Ben-Shahar, T.R., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M., and Uhlmann, F. (2008). Eco1-Dependent Cohesin Acetylation During Establishment of Sister Chromatid Cohesion. Science *321*, 563-566.

Bermudez, V.P., Farina, A., Higashi, T.L., Du, F., Tappin, I., Takahashi, T.S., and Hurwitz, J. (2012). In vitro loading of human cohesin on DNA by the human Scc2-Scc4 loader complex. Proceedings of the National Academy of Sciences, 1-6.

Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. Science *294*, 2539-2542.

Bernard, P., Schmidt, C.K., Vaur, S., Dheur, S., Drogat, J., Genier, S., Ekwall, K., Uhlmann, F., and Javerzat, J.-P. (2008). Cell-cycle regulation of cohesin stability along fission yeast chromosomes. EMBO J *27*, 111-121.

Bhat, M.A., Philp, A.V., Glover, D.M., and Bellen, H.J. (1996). Chromatid segregation at anaphase requires the barren product, a novel chromosome-associated protein that interacts with Topoisomerase II. Cell *87*, 1103-1114.

Bigay, J., Deterre, P., Pfister, C., and Chabre, M. (1987). Fluoride complexes of aluminium or beryllium act on G-proteins as reversibly bound analogues of the gamma phosphate of GTP. The EMBO Journal *6*, 2907-2913.

Biggins, S., Severin, F.F., Bhalla, N., Sassoon, I., Hyman, A.A., and Murray, A.W. (1999). The conserved protein kinase Ipl1 regulates microtubule binding to kinetochores in budding yeast. Genes Dev *13*, 532-544.

Blat, Y., and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. Cell *98*, 249-259.

Blatch, G.L., and Lässle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays *21*, 932-939.

Borges, V., Lehane, C., Lopez-Serra, L., Flynn, H., Skehel, M., Rolef Ben-Shahar, T., and Uhlmann, F. (2010). Hos1 deacetylates Smc3 to close the cohesin acetylation cycle. Mol Cell *39*, 677-688.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry *72*, 248-254.

Bressan, D.A., Baxter, B.K., and Petrini, J.H. (1999). The Mre11-Rad50-Xrs2 protein complex facilitates homologous recombination-based double-strand break repair in Saccharomyces cerevisiae. Mol Cell Biol *19*, 7681-7687.

Britton, R.A., Lin, D.C., and Grossman, A.D. (1998). Characterization of a prokaryotic SMC protein involved in chromosome partitioning. Genes Dev *12*, 1254-1259.

Bucher, M.H., Evdokimov, A.G., and Waugh, D.S. (2002). Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrin-binding protein. Acta crystallographica Section D, Biological crystallography *58*, 392-397.

Chan, K.-L., Roig, M.B., Hu, B., Beckouët, F., Metson, J., and Nasmyth, K. (2012). Cohesin's DNA Exit Gate Is Distinct from Its Entrance Gate and Is Regulated by Acetylation. Cell.

Chavez, A., Agrawal, V., and Johnson, F.B. (2011). Homologous recombination-dependent rescue of deficiency in the structural maintenance of chromosomes (Smc) 5/6 complex. J Biol Chem *286*, 5119-5125.

Cheeseman, I.M., Drubin, D.G., and Barnes, G. (2002). Simple centromere, complex kinetochore: linking spindle microtubules and centromeric DNA in budding yeast. J Cell Biol *157*, 199-203.

Chen, J., Sharma, S., Quiocho, F.A., and Davidson, A.L. (2001). Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport. Proc Natl Acad Sci U S A *98*, 1525-1530.

Chiu, A., Revenkova, E., and Jessberger, R. (2004). DNA interaction and dimerization of eukaryotic SMC hinge domains. J Biol Chem *279*, 26233-26242.

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

Ciosk, R., Zachariae, W., Michaelis, C., Shevchenko, A., Mann, M., and Nasmyth, K. (1998). An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. Cell *93*, 1067-1076.

Cobbe, N., and Heck, M.M. (2004). The evolution of SMC proteins: phylogenetic analysis and structural implications. Mol Biol Evol *21*, 332-347.

Cohen-Fix, O., Peters, J.M., Kirschner, M.W., and Koshland, D. (1996). Anaphase initiation in Saccharomyces cerevisiae is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. Genes & Development *10*, 3081-3093.

Cooper, G. (2000). The Cell: A Molecular Approach, 2nd edition edn (Boston University, Sunderland (MA): Sinauer Associates).

Cuylen, S., and Haering, C.H. (2010). A new cohesive team to mediate DNA looping. Cell Stem Cell *7*, 424-426.

Cuylen, S., Metz, J., and Haering, C.H. (2011). Condensin structures chromosomal DNA through topological links. Nature structural & molecular biology.

D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. Trends Biochem Sci *28*, 655-662.

Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., and Possoz, C. (2007). MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. Mol Microbiol *65*, 1485-1492.

De la Rosa, M.B., and Nelson, S.W. (2011). An interaction between the Walker A and D-loop motifs is critical to ATP hydrolysis and cooperativity in bacteriophage T4 Rad50. J Biol Chem *286*, 26258-26266.

De Piccoli, G., Cortes-Ledesma, F., Ira, G., Torres-Rosell, J., Uhle, S., Farmer, S., Hwang, J.-Y., Machin, F., Ceschia, A., McAleenan, A.*, et al.* (2006). Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. Nat Cell Biol *8*, 1032-1034.

De Wulf, P. (2003). Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. Genes & Development *17*, 2902-2921.

Deardorff, M.A., Bando, M., Nakato, R., Watrin, E., Itoh, T., Minamino, M., Saitoh, K., Komata, M., Katou, Y., Clark, D.*, et al.* (2012). HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature *489*, 313-317.

Deardorff, M.A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J., Gil-Rodriguez, C., Arnedo, M., Loeys, B., Kline, A.D.*, et al.* (2007). Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet *80*, 485-494.

Dekker, J. (2002). Capturing Chromosome Conformation. Science *295*, 1306-1311.

Detweiler, C.S., and Li, J.J. (1997). Cdc6p establishes and maintains a state of replication competence during G1 phase. Journal of Cell Science *110 (Pt 6)*, 753-763.

Diaz-Martinez, L.A., Gimenez-Abian, J.F., and Clarke, D.J. (2008). Chromosome cohesion - rings, knots, orcs and fellowship. Journal of Cell Science *121*, 2107-2114.

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

Donovan, S., Harwood, J., Drury, L.S., and Diffley, J.F. (1997). Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. Proceedings of the National Academy of Sciences of the United States of America *94*, 5611-5616.

Ebmeier, C.C., and Taatjes, D.J. (2010). Activator-Mediator binding regulates Mediator-cofactor interactions. Proc Natl Acad Sci U S A *107*, 11283-11288.

Eckert, C.A., Gravdahl, D.J., and Megee, P.C. (2007). The enhancement of pericentromeric cohesin association by conserved kinetochore components promotes high-fidelity chromosome segregation and is sensitive to microtubule-based tension. Genes & Development *21*, 278-291.

Efferth, T. (2001). The human ATP-binding cassette transporter genes: from the bench to the bedside. Curr Mol Med *1*, 45-65.

Esnault, C., Ghavi-Helm, Y., Brun, S., Soutourina, J., Van Berkum, N., Boschiero, C., Holstege, F., and Werner, M. (2008). Mediator-dependent recruitment of TFIIH modules in preinitiation complex. Mol Cell *31*, 337-346.

Fang, G., Yu, H., and Kirschner, M.W. (1998a). The checkpoint protein MAD2 and the mitotic regulator CDC20 form a ternary complex with the anaphase-promoting complex to control anaphase initiation. Genes Dev *12*, 1871-1883.

Fang, G., Yu, H., and Kirschner, M.W. (1998b). Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. Mol Cell *2*, 163-171.

Farcas, A.-M., Uluocak, P., Helmhart, W., and Nasmyth, K. (2011). Cohesin's Concatenation of Sister DNAs Maintains Their Intertwining. Molecular Cell *44*, 97-107.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Fernius, J., and Marston, A.L. (2009). Establishment of Cohesion at the Pericentromere by the Ctf19 Kinetochore Subcomplex and the Replication Fork-Associated Factor, Csm3. PLoS Genet *5*, e1000629.

Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M., and Rayment, I. (1995). X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4-. Biochemistry *34*, 8960-8972.

Furuya, K., Takahashi, K., and Yanagida, M. (1998). Faithful anaphase is ensured by Mis4, a sister chromatid cohesion molecule required in S phase and not destroyed in G1 phase. Genes Dev *12*, 3408-3418.

Gandhi, R., Gillespie, P.J., and Hirano, T. (2006). Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. Curr Biol *16*, 2406-2417.

Geourjon, C., Orelle, C., Steinfels, E., Blanchet, C., Deléage, G., Di Pietro, A., and Jault, J.M. (2001). A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. Trends Biochem Sci *26*, 539-544.

Gerlich, D., Koch, B., Dupeux, F., Peters, J., and Ellenberg, J. (2006). Live-Cell Imaging Reveals a Stable Cohesin-Chromatin Interaction after but Not before DNA Replication. Current Biology *16*, 1571-1578.

Gillespie, P.J., and Hirano, T. (2004). Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. Curr Biol *14*, 1598-1603.

Glynn, E.F., Megee, P.C., Yu, H.-G., Mistrot, C., Unal, E., Koshland, D.E., Derisi, J.L., and Gerton, J.L. (2004). Genome-Wide Mapping of the Cohesin Complex in the Yeast Saccharomyces cerevisiae. Plos Biol *2*, e259.

Gregan, J., Zhang, C., Rumpf, C., Cipak, L., Li, Z., Uluocak, P., Nasmyth, K., and Shokat, K.M. (2007). Construction of conditional analog-sensitive kinase alleles in the fission yeast Schizosaccharomyces pombe. Nat Protoc *2*, 2996-3000.

Griese, J.J., Witte, G., and Hopfner, K.-P. (2010). Structure and DNA binding activity of the mouse condensin hinge domain highlight common and diverse features of SMC proteins. Nucleic Acids Res, 1-12.

Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K., and Nasmyth, K. (2006). Evidence that Loading of Cohesin Onto Chromosomes Involves Opening of Its SMC Hinge. Cell *127*, 523-537.

Gruber, S., and Errington, J. (2009). Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in B. subtilis. Cell *137*, 685-696.

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

Gullerova, M., and Proudfoot, N.J. (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell *132*, 983-995.

Gustafsson, C.M., Myers, L.C., Beve, J., Spahr, H., Lui, M., Erdjument-Bromage, H., Tempst, P., and Kornberg, R.D. (1998). Identification of new mediator subunits in the RNA polymerase II holoenzyme from Saccharomyces cerevisiae. J Biol Chem *273*, 30851-30854.

Haber, J.E., Thorburn, P.C., and Rogers, D. (1984). Meiotic and mitotic behavior of dicentric chromosomes in Saccharomyces cerevisiae. Genetics *106*, 185-205.

Haering, C. (2004). Structure and Stability of Cohesin's Smc1-Kleisin Interaction. Molecular Cell *15*, 951-964.

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

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

Hagan, R.M., Bjornsson, R., McMahon, S.A., Schomburg, B., Braithwaite, V., Buhl, M., Naismith, J.H., and Schwarz-Linek, U. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. Angew Chem Int Ed Engl *49*, 8421-8425.

Hamza, A., and Baetz, K. (2011). The iron-responsive transcription factor Aft1 interacts with the kinetochore protein Iml3 and promotes pericentromeric cohesin. J Biol Chem.

Harvey, S.H., Sheedy, D.M., Cuddihy, A.R., and O'Connell, M.J. (2004). Coordination of DNA damage responses via the Smc5/Smc6 complex. Mol Cell Biol *24*, 662-674.

Hauf, S., Roitinger, E., Koch, B., Dittrich, C.M., Mechtler, K., and Peters, J.-M. (2005). Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. PLoS Biol *3*, e69.

Hauf, S., Waizenegger, I.C., and Peters, J.M. (2001). Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. Science *293*, 1320-1323.

Hayashi, M.T., Takahashi, T.S., Nakagawa, T., Nakayama, J.-I., and Masukata, H. (2009). The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. Nat Cell Biol *11*, 357-362.

Heidinger-Pauli, J.M., Mert, O., Davenport, C., Guacci, V., and Koshland, D. (2010a). Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. Curr Biol *20*, 957-963.

Heidinger-Pauli, J.M., Onn, I., and Koshland, D. (2010b). Genetic Evidence that the Acetylation of the Smc3p Subunit of Cohesin Modulates its ATP-bound State to Promote Cohesion Establishment in Saccharomyces cerevisiae. Genetics.

Heidinger-Pauli, J.M., Unal, E., and Koshland, D. (2009). Distinct Targets of the Eco1 Acetyltransferase Modulate Cohesion in S Phase and in Response to DNA Damage. Molecular Cell *34*, 311-321.

Hirano, M., and Hirano, T. (2002). Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. EMBO J *21*, 5733-5744.

Hirano, M., and Hirano, T. (2004). Positive and negative regulation of SMC-DNA interactions by ATP and accessory proteins. EMBO J *23*, 2664-2673.

Hirano, T. (2005). SMC proteins and chromosome mechanics: from bacteria to humans. Philosophical transactions of the Royal Society of London Series B, Biological sciences *360*, 507-514.

Hirano, T., and Mitchison, T.J. (1994). A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. Cell *79*, 449-458.

Hohl, M., Kwon, Y., Galván, S.M., Xue, X., Tous, C., Aguilera, A., Sung, P., and Petrini, J.H.J. (2011). The Rad50 coiled-coil domain is indispensable for Mre11 complex functions. Nat Struct Mol Biol *18*, 1124-1131.

Hollenstein, K., Dawson, R.J.P., and Locher, K.P. (2007). Structure and mechanism of ABC transporter proteins. Curr Opin Struct Biol *17*, 412-418.

Hooke, R. (1665). Micrographia: or, Some physiological descriptions of minute bodies made by magnifying glasses (London: J. Martyn and J. Allestry).

Hopfner, K.-P., Craig, L., Moncalian, G., Zinkel, R.A., Usui, T., Owen, B.A.L., Karcher, A., Henderson, B., Bodmer, J.-L., McMurray, C.T.*, et al.* (2002). The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. Nature *418*, 562-566.

Hopfner, K.-P., and Tainer, J.A. (2003). Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. Curr Opin Struct Biol *13*, 249-255.

Hopfner, K.P., Karcher, A., Craig, L., Woo, T.T., Carney, J.P., and Tainer, J.A. (2001). Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell *105*, 473-485.

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

Horsfield, J.A., Print, C.G., and Mönnich, M. (2012). Diverse Developmental Disorders from The One Ring: Distinct Molecular Pathways Underlie the Cohesinopathies. Front Gene *3*, 1-15.

Hu, B., Ito, T., Mishra, A., Katoh, Y., Chan, K.-L., Upcher, W., Godlee, C., Brunet, M., Shirahige, K., and Nasmyth, K. (in press). Identification of an intermediate step in the cohesin chromosome loading reaction. Molecular Cell.

Hu, B., Itoh, T., Mishra, A., Katoh, Y., Chan, K.-L., Upcher, W., Godlee, C., Roig, M.B., Shirahige, K., and Nasmyth, K. (2011). ATP Hydrolysis Is Required for Relocating Cohesin from Sites Occupied by Its Scc2/4 Loading Complex. Current Biology *21*, 12-24.

Huang, C.E., Milutinovich, M., and Koshland, D. (2005). Rings, bracelet or snaps: fashionable alternatives for Smc complexes. Philosophical Transactions of the Royal Society B: Biological Sciences *360*, 537-542.

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

Irniger, S., Piatti, S., Michaelis, C., and Nasmyth, K. (1995). Genes involved in sister chromatid separation are needed for B-type cyclin proteolysis in budding yeast. Cell *81*, 269-278.

Ivanov, D., and Nasmyth, K. (2005). A Topological Interaction between Cohesin Rings and a Circular Minichromosome. Cell *122*, 849-860.

Ivanov, D., Schleiffer, A., Eisenhaber, F., Mechtler, K., Haering, C.H., and Nasmyth, K. (2002). Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion. Curr Biol *12*, 323-328.

Jackson, J.R., Patrick, D.R., Dar, M.M., and Huang, P.S. (2007). Targeted anti-mitotic therapies: can we improve on tubulin agents? Nat Rev Cancer *7*, 107-117.

Janas, E., Hofacker, M., Chen, M., Gompf, S., van der Does, C., and Tampé, R. (2003). The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p. J Biol Chem *278*, 26862-26869.

Jessberger, R. (2010). Deterioration without replenishment--the misery of oocyte cohesin. Genes & Development *24*, 2587-2591.

Joglekar, A.P., Bouck, D.C., Molk, J.N., Bloom, K.S., and Salmon, E.D. (2006). Molecular architecture of a kinetochore–microtubule attachment site. Nat Cell Biol *8*, 581-585.

Jones, P.M., and George, A.M. (1999). Subunit interactions in ABC transporters: towards a functional architecture. FEMS Microbiol Lett *179*, 187-202.

Jones, P.M., and George, A.M. (2002). Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. Proc Natl Acad Sci USA *99*, 12639-12644.

Jones, P.M., and George, A.M. (2011). Molecular-dynamics simulations of the ATP/apo state of a multidrug ATP-binding cassette transporter provide a structural and mechanistic basis for the asymmetric occluded state. Biophys J *100*, 3025-3034.

Jones, P.M., O'Mara, M.L., and George, A.M. (2009). ABC transporters: a riddle wrapped in a mystery inside an enigma. Trends Biochem Sci *34*, 520-531.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature.

Kenna, M.A., and Skibbens, R.V. (2003). Mechanical link between cohesion establishment and DNA replication: Ctf7p/Eco1p, a cohesion establishment factor, associates with three different replication factor C complexes. Mol Cell Biol *23*, 2999-3007.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Res *37*, D387-392.

Kiermaier, E., Woehrer, S., Peng, Y., Mechtler, K., and Westermann, S. (2009). A Dam1-based artificial kinetochore is sufficient to promote chromosome segregation in budding yeast. Nat Cell Biol *11*, 1109-1115.

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

Krantz, I.D., McCallum, J., DeScipio, C., Kaur, M., Gillis, L.A., Yaeger, D., Jukofsky, L., Wasserman, N., Bottani, A., Morris, C.A*., et al.* (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet *36*, 631-635.

Ku, B., Lim, J.H., Shin, H.C., Shin, S.Y., Oh, B.H. (2010). Crystal structure of the MukB hinge domain with coiled-coil stretches and its functional implications. Proteins *78,* 1483-1490.

Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.-M. (2006). Wapl controls the dynamic association of cohesin with chromatin. Cell *127*, 955-967.

Kulemzina, I., Schumacher, M.R., Verma, V., Reiter, J., Metzler, J., Failla, A.V., Lanz, C., Sreedharan, V.T., Rätsch, G., and Ivanov, D. (2012). Cohesin rings devoid of scc3 and pds5 maintain their stable association with the DNA. PLoS Genetics *8*, e1002856.

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

Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? Genes & Development *24*, 1208-1219.

Lacefield, S., Lau, D.T., and Murray, A.W. (2009). Recruiting a microtubule-binding complex to DNA directs chromosome segregation in budding yeast. Nat Cell Biol *11*, 1116-1120.

Lafont, A.L., Song, J., and Rankin, S. (2010). Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. Proc Natl Acad Sci USA, 1-6.

Laloraya, S., Guacci, V., and Koshland, D. (2000). Chromosomal addresses of the cohesin component Mcd1p. J Cell Biol *151*, 1047-1056.

Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010). The MRN complex in double-strand break repair and telomere maintenance. FEBS Letters *584*, 3682-3695.

Lammens, A., Schele, A., and Hopfner, K.-P. (2004). Structural biochemistry of ATP-driven dimerization and DNA-stimulated activation of SMC ATPases. Curr Biol *14*, 1778-1782.

Lammens, K., Bemeleit, Derk J., Möckel, C., Clausing, E., Schele, A., Hartung, S., Schiller, Christian B., Lucas, M., Angermüller, C., Söding, J*., et al.* (2011). The Mre11:Rad50 Structure Shows an ATP-Dependent Molecular Clamp in DNA Double-Strand Break Repair. Cell *145*, 54-66.

Larionov, V.L., Karpova, T.S., Kouprina, N.Y., and Jouravleva, G.A. (1985). A mutant of Saccharomyces cerevisiae with impaired maintenance of centromeric plasmids. Curr Genet *10*, 15-20.

Lee, B.-K., and Iyer, V.R. (2012). Genome-wide Studies of CCCTC-binding Factor (CTCF) and Cohesin Provide Insight into Chromatin Structure and Regulation. Journal of Biological Chemistry *287*, 30906-30913.

Leipe, D.D., Koonin, E.V., and Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. Journal of Molecular Biology *333*, 781-815.

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

Li, X., and Nicklas, R.B. (1995). Mitotic forces control a cell-cycle checkpoint. Nature *373*, 630-632.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., and Hartzog, G.A. (2003). Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. Mol Cell Biol *23*, 1368-1378.

Ling, G., Wei, Y., and Ding, X. (2007). Transcriptional regulation of human CYP2A13 expression in the respiratory tract by CCAAT/enhancer binding protein and epigenetic modulation. Molecular pharmacology *71*, 807-816.

Linton, K.J. (2007). Structure and function of ABC transporters. Physiology (Bethesda) *22*, 122-130.

Linton, K.J., and Higgins, C.F. (2007). Structure and function of ABC transporters: the ATP switch provides flexible control. Pflugers Arch *453*, 555-567.

Liu, D., Vader, G., Vromans, M.J.M., Lampson, M.A., and Lens, S.M.A. (2009). Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. Science *323*, 1350-1353.

Liu, J., and Krantz, I.D. (2008). Cohesin and human disease. Annual review of genomics and human genetics *9*, 303-320.

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

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

Lyons, Nicholas A., and Morgan, David O. (2011). Cdk1-Dependent Destruction of Eco1 Prevents Cohesion Establishment after S Phase. Molecular Cell *42*, 378-389.

Matsuyama, T., Yamanishi, M., and Takahashi, H. (2011). Improvement of galactose induction system in Saccharomyces cerevisiae. JBIOSC *111*, 175-177.

Mc Intyre, J., Muller, E.G.D., Weitzer, S., Snydsman, B.E., Davis, T.N., and Uhlmann, F. (2007). In vivo analysis of cohesin architecture using FRET in the budding yeast Saccharomyces cerevisiae. EMBO J *26*, 3783-3793.

McAleenan, A., Cordon-Preciado, V., Clemente-Blanco, A., Liu, I.-C., Sen, N., Leonard, J., Jarmuz, A., and Aragón, L. (2012). SUMOylation of the α-Kleisin Subunit of Cohesin Is Required for DNA Damage-Induced Cohesion. Curr Biol *22*, 1564-1575.

McGuinness, B.E., Hirota, T., Kudo, N.R., Peters, J.-M., and Nasmyth, K. (2005). Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. PLoS Biol *3*, e86.

McNairn, A.J., and Gerton, J.L. (2008). Cohesinopathies: One ring, many obligations. Mutat Res *647*, 103-111.

Megee, P.C., and Koshland, D. (1999). A functional assay for centromere-associated sister chromatid cohesion. Science *285*, 254-257.

Megee, P.C., Mistrot, C., Guacci, V., and Koshland, D. (1999). The centromeric sister chromatid cohesion site directs Mcd1p binding to adjacent sequences. Molecular Cell *4*, 445-450.

Melby, T.E., Ciampaglio, C.N., Briscoe, G., and Erickson, H.P. (1998). The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. J Cell Biol *142*, 1595-1604.

Michaelis, C., Ciosk, R., and Nasmyth, K. (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell *91*, 35-45.

Mishra, A., Hu, B., Kurze, A., Beckouet, F., Farcas, A.-M., Dixon, S.E., Katou, Y., Khalid, S., Shirahige, K., and Nasmyth, K. (2010). Both Interaction Surfaces within Cohesin's Hinge Domain Are Essential for Its Stable Chromosomal Association. Current Biology, 1-11.

Moerke, N.J. (2009) Fluorescence polarization assays for monitoring peptide-protein or nucleic acid-protein binding. Curr Protoc Chem Biol 1: 1-15,

Moldovan, G., Pfander, B., and Jentsch, S. (2006). PCNA Controls Establishment of Sister Chromatid Cohesion during S Phase. Molecular Cell *23*, 723-732.

Moncalian, G., Lengsfeld, B., Bhaskara, V., Hopfner, K.-P., Karcher, A., Alden, E., Tainer, J.A., and Paull, T.T. (2004). The rad50 signature motif: essential to ATP binding and biological function. Journal of Molecular Biology *335*, 937-951.

Moody, J.E., Millen, L., Binns, D., Hunt, J.F., and Thomas, P.J. (2002). Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. J Biol Chem *277*, 21111-21114.

Moreno-Herrero, F., de Jager, M., Dekker, N.H., Kanaar, R., Wyman, C., and Dekker, C. (2005). Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. Nature *437*, 440-443.

Morgan, D.O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. Cell and Developmental Biology *13*, 261-291.

Morgan, D.O. (1999). Regulation of the APC and the exit from mitosis. Nat Cell Biol *1*, E47-53.

Morgan, D.O. (2007). The cell cycle: principles of control. (Oxford: Oxford University Press).

Musacchio, A., and Salmon, E.D. (2007). The spindle-assembly checkpoint in space and time. Nat Rev Mol Cell Biol *8*, 379-393.

Musio, A., Selicorni, A., Focarelli, M.L., Gervasini, C., Milani, D., Russo, S., Vezzoni, P., and Larizza, L. (2006). X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat Genet *38*, 528-530.

Nakayama, J., Klar, A.J., and Grewal, S.I. (2000). A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. Cell *101*, 307-317.

Nasmyth, K. (2005). How Do so Few Control so Many? Cell *120*, 739-746.

Nasmyth, K. (2011). Cohesin: a catenase with separate entry and exit gates? Nat Cell Biol *13*, 1170-1177.

Nasmyth, K., and Haering, C.H. (2005). The structure and function of SMC and kleisin complexes. Annu Rev Biochem *74*, 595-648.

Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.-M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet *5*, e1000739.

Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nat Chem Biol *4*, 232-234.

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

Newberger, D.S. (2000). Down syndrome: prenatal risk assessment and diagnosis. Am Fam Physician *62*, 825-832, 837-828.

Newstead, S., Fowler, P.W., Bilton, P., Carpenter, E.P., Sadler, P.J., Campopiano, D.J., Sansom, M.S.P., and Iwata, S. (2009). Insights into how nucleotide-binding domains power ABC transport. Structure *17*, 1213-1222.

Ng, T.M., Waples, W.G., Lavoie, B.D., and Biggins, S. (2009). Pericentromeric sister chromatid cohesion promotes kinetochore biorientation. Mol Biol Cell *20*, 3818-3827.

Nicklas, R.B., and Koch, C.A. (1969). Chromosome micromanipulation. 3. Spindle fiber tension and the reorientation of mal-oriented chromosomes. J Cell Biol *43*, 40-50.

Nicklas, R.B., Ward, S.C., and Gorbsky, G.J. (1995). Kinetochore chemistry is sensitive to tension and may link mitotic forces to a cell cycle checkpoint. J Cell Biol *130*, 929-939.

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

Nishiyama, T., Ladurner, R., Schmitz, J., Kreidl, E., Schleiffer, A., Bhaskara, V., Bando, M., Shirahige, K., Hyman, A.A., Mechtler, K., *et al.* (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. Cell *143*, 737-749.

Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nat Cell Biol *4*, 89-93.

Oke, M., Carter, L., Johnson, K., Liu, H., McMahon, S., Yan, X., Kerou, M., Weikart, N., Kadi, N., Sheikh, M.A., *et al.* (2010). The Scottish Structural Proteomics Facility: targets, methods and outputs. J Struct Funct Genomics *11*, 167-180.

Onn, I., and Koshland, D. (2011). In vitro assembly of physiological cohesin/DNA complexes. Proceedings of the National Academy of Sciences of the United States of America.

Opitz, J.M. (1985). The Brachmann-de Lange syndrome. American journal of medical genetics *22*, 89-102.

Orelle, C., Dalmas, O., Gros, P., Di Pietro, A., and Jault, J.-M. (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. J Biol Chem *278*, 47002-47008.

Panizza, S., Tanaka, T., Hochwagen, A., Eisenhaber, F., and Nasmyth, K. (2000). Pds5 cooperates with cohesin in maintaining sister chromatid cohesion. Curr Biol *10*, 1557-1564.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H., Jarmuz, A., Canzonetta, C., Webster, Z., and Nesterova, T. (2008). Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. Cell *132*, 422-433.

Pauli, A., Althoff, F., Oliveira, R., Heidmann, S., Schuldiner, O., Lehner, C., Dickson, B., and Nasmyth, K. (2008). Cell-Type-Specific TEV Protease Cleavage Reveals Cohesin Functions in Drosophila Neurons. Developmental Cell *14*, 239-251.

Peitsch, M.C., Wells, T.N., Stampf, D.R., and Sussman, J.L. (1995). The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. Trends Biochem Sci *20*, 82-84.

Pellegrino, S., de Sanctis, D., McSweeney, S., and Timmins, J. (2012). Expression, purification and preliminary structural analysis of the coiled-coil domain of Deinococcus radiodurans RecN. Acta Crystallogr Sect F Struct Biol Cryst Commun *68*, 218-221.

Peters, J., Tedeschi, A., Schmitz, J. (2008). The cohesin complex and its roles in chromosome biology. Genes and Development *22*, 3089-3114.

Pié, J., Gil-Rodriguez, M.C., Ciero, M., López-Viñas, E., Ribate, M.P., Arnedo, M., Deardorff, M.A., Puisac, B., Legaretta, J., Karam, J.C.d, *et al.* (2010). Mutations and variants in the cohesion

factor genes NIPBL, SMC1A and SMC3 in a cohort of 30 unrelated patients with Cornelia de Lange Syndrome. Am J Med Genet A *152*, 924-929.

Pot, I., Measday, V., Snydsman, B., Cagney, G., Fields, S., Davis, T.N., Muller, E.G.D., and Hieter, P. (2003). Chl4p and iml3p are two new members of the budding yeast outer kinetochore. Mol Biol Cell *14*, 460-476.

Procko, E., Ferrin-O'Connell, I., Ng, S.-L., and Gaudet, R. (2006). Distinct structural and functional properties of the ATPase sites in an asymmetric ABC transporter. Molecular Cell *24*, 51-62.

Rankin, S., Ayad, N.G., and Kirschner, M.W. (2005). Sororin, a substrate of the anaphase-promoting complex, is required for sister chromatid cohesion in vertebrates. Molecular Cell *18*, 185-200.

Rees, D.C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to change. Nat Rev Mol Cell Biol *10*, 218-227.

Renshaw, M.J., Ward, J.J., Kanemaki, M., Natsume, K., Nédélec, F.J., and Tanaka, T.U. (2010). Condensins Promote Chromosome Recoiling during Early Anaphase to Complete Sister Chromatid Separation. Developmental Cell *19*, 232-244.

Revenkova, E., Eijpe, M., Heyting, C., Gross, B., and Jessberger, R. (2001). Novel meiosis-specific isoform of mammalian SMC1. Mol Cell Biol *21*, 6984-6998.

Revenkova, E., Herrmann, K., Adelfalk, C., and Jessberger, R. (2010). Oocyte cohesin expression restricted to predictyate stages provides full fertility and prevents aneuploidy. Curr Biol *20*, 1529-1533.

Rollins, R.A., Morcillo, P., and Dorsett, D. (1999). Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics *152*, 577-593.

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

Rutkowska, A., Haering, C.H., and Schultz, C. (2011). A FlAsH-based cross-linker to study protein interactions in living cells. Angew Chem Int Ed Engl *50*, 12655-12658.

Saka, Y., Sutani, T., Yamashita, Y., Saitoh, S., Takeuchi, M., Nakaseko, Y., and Yanagida, M. (1994). Fission yeast cut3 and cut14, members of a ubiquitous protein family, are required for chromosome condensation and segregation in mitosis. EMBO J *13*, 4938-4952.

Sakai, A., Hizume, K., Sutani, T., Takeyasu, K., and Yanagida, M. (2003). Condensin but not cohesin SMC heterodimer induces DNA reannealing through protein-protein assembly. EMBO J *22*, 2764-2775.

Santaguida, S., and Musacchio, A. (2009). The life and miracles of kinetochores. *28*, 2511-2531.

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

Schmidt, C.K., Brookes, N., and Uhlmann, F. (2009). Conserved features of cohesin binding along fission yeast chromosomes. Genome Biol *10*, R52.

Schmitz, J., Watrin, E., Lénárt, P., Mechtler, K., and Peters, J.-M. (2007). Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase. Curr Biol *17*, 630-636.

Seitan, V.C., Banks, P., Laval, S., Majid, N.A., Dorsett, D., Rana, A., Smith, J., Bateman, A., Krpic, S., Hostert, A.*, et al.* (2006). Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance. Plos Biol *4*, e242.

Sharma, S., and Davidson, A.L. (2000). Vanadate-induced trapping of nucleotides by purified maltose transport complex requires ATP hydrolysis. Journal of Bacteriology *182*, 6570-6576.

Sherratt, D.J. (2003). Bacterial chromosome dynamics. Science *301*, 780-785.

Sheu, Y.-J., and Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. Molecular Cell *24*, 101-113.

Shintomi, K., and Hirano, T. (2009). Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl-Pds5 and Sgo1. Genes & Development.

Shirayama, M., Tóth, A., Gálová, M., and Nasmyth, K. (1999). APC(Cdc20) promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5. Nature *402*, 203-207.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., Mcwilliam, H., Remmert, M., SÖDing, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology *7*, 1-6.

Skibbens, R.V., Corson, L.B., Koshland, D., and Hieter, P. (1999). Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. Genes Dev *13*, 307-319.

Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J., and Hunt, J.F. (2002). ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Molecular Cell *10*, 139-149.

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Splinter, E., Heath, H., Kooren, J., Palstra, R.-J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & Development *20*, 2349-2354.

Stead, K., Aguilar, C., Hartman, T., Drexel, M., Meluh, P., and Guacci, V. (2003). Pds5p regulates the maintenance of sister chromatid cohesion and is sumoylated to promote the dissolution of cohesion. J Cell Biol *163*, 729-741.

Strachan, T. (2005). Cornelia de Lange Syndrome and the link between chromosomal function, DNA repair and developmental gene regulation. Curr Opin Genet Dev *15*, 258-264.

Ström, L., Karlsson, C., Lindroos, H.B., Wedahl, S., Katou, Y., Shirahige, K., and Sjögren, C. (2007). Postreplicative formation of cohesion is required for repair and induced by a single DNA break. Science *317*, 242-245.

Strom, L., Lindroos, H.B., Shirahige, K., and Sjogren, C. (2004). Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair. Mol Cell *16*, 1003-1015.

Strunnikov, A.V., Larionov, V.L., and Koshland, D. (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. J Cell Biol *123*, 1635-1648.

Sullivan, N.L., Marquis, K.A., and Rudner, D.Z. (2009). Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation. Cell *137*, 697-707.

Sumara, I., Vorlaufer, E., Gieffers, C., Peters, B.H., and Peters, J.M. (2000). Characterization of vertebrate cohesin complexes and their regulation in prophase. J Cell Biol *151*, 749-762.

Sundin, O., and Varshavsky, A. (1981). Arrest of segregation leads to accumulation of highly intertwined catenated dimers: dissection of the final stages of SV40 DNA replication. Cell *25*, 659-669.

Taatjes, D.J. (2010). The human Mediator complex: a versatile, genome-wide regulator of transcription. Trends Biochem Sci *35*, 315-322.

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

Takahashi, T.S., Basu, A., Bermudez, V., Hurwitz, J., and Walter, J.C. (2008). Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in Xenopus egg extracts. Genes & Development *22*, 1894-1905.

Takahashi, T.S., Yiu, P., Chou, M.F., Gygi, S., and Walter, J.C. (2004). Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. Nat Cell Biol *6*, 991-996.

Tanaka, T., Cosma, M.P., Wirth, K., and Nasmyth, K. (1999). Identification of cohesin association sites at centromeres and along chromosome arms. Cell *98*, 847-858.

Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J., and Strachan, T. (2004). NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet *36*, 636-641.

Toth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. (1999). Yeast Cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. Genes & Development *13*, 320-333.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P.*, et al.* (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature *403*, 623-627.

Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. Nature *400*, 37-42.

Uhlmann, F., and Nasmyth, K. (1998). Cohesion between sister chromatids must be established during DNA replication. Curr Biol *8*, 1095-1101.

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

Unal, E., Arbel-Eden, A., Sattler, U., Shroff, R., Lichten, M., Haber, J.E., and Koshland, D. (2004). DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain. Mol Cell *16*, 991-1002.

Unal, E., Heidinger-Pauli, J.M., Kim, W., Guacci, V., Onn, I., Gygi, S.P., and Koshland, D.E. (2008). A Molecular Determinant for the Establishment of Sister Chromatid Cohesion. Science *321*, 566-569.

Unal, E., Heidinger-Pauli, J.M., and Koshland, D. (2007). DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7). Science *317*, 245-248.

Urbatsch, I.L., Sankaran, B., Weber, J., and Senior, A.E. (1995). P-glycoprotein is stably inhibited by vanadate-induced trapping of nucleotide at a single catalytic site. J Biol Chem *270*, 19383-19390.

van den Ent, F., Lockhart, A., Kendrick-Jones, J., and Löwe, J. (1999). Crystal structure of the N-terminal domain of MukB: a protein involved in chromosome partitioning. Structure *7*, 1181-1187.

Vaur, S., Feytout, A., Vazquez, S., and Javerzat, J.-P. (2012). Pds5 promotes cohesin acetylation and stable cohesin-chromosome interaction. EMBO Rep *13*, 645-652.

Vega, H., Waisfisz, Q., Gordillo, M., Sakai, N., Yanagihara, I., Yamada, M., van Gosliga, D., Kayserili, H., Xu, C., Ozono, K.*, et al.* (2005). Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. Nat Genet *37*, 468-470.

Verdon, G., Albers, S.-V., van Oosterwijk, N., Dijkstra, B.W., Driessen, A.J.M., and Thunnissen, A.-M.W.H. (2003). Formation of the productive ATP-Mg2+-bound dimer of GlcV, an ABC-ATPase from Sulfolobus solfataricus. Journal of Molecular Biology *334*, 255-267.

Waizenegger, I.C., Hauf, S., Meinke, A., and Peters, J.M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. Cell *103*, 399-410.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (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*, 945-951.

Wang, L., Xie, J., and Schultz, P.G. (2006). EXPANDING THE GENETIC CODE. Annu Rev Biophys Biomol Struct *35*, 225-249.

Wäsch, R., and Cross, F.R. (2002). APC-dependent proteolysis of the mitotic cyclin Clb2 is essential for mitotic exit. Nature *418*, 556-562.

Watrin, E., Schleiffer, A., Tanaka, K., Eisenhaber, F., Nasmyth, K., and Peters, J. (2006). Human Scc4 Is Required for Cohesin Binding to Chromatin, Sister-Chromatid Cohesion, and Mitotic Progression. Current Biology *16*, 863-874.

Weber, S.A., Gerton, J.L., Polancic, J.E., DeRisi, J.L., Koshland, D., and Megee, P.C. (2004). The kinetochore is an enhancer of pericentric cohesin binding. Plos Biol *2*, E260.

Weibezahn, J. (2003). Characterization of a Trap Mutant of the AAA+ Chaperone ClpB. Journal of Biological Chemistry *278*, 32608-32617.

Weitzer, S. (2003). A Model for ATP Hydrolysis-Dependent Binding of Cohesin to DNA. Current Biology *13*, 1930-1940.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Westermann, S., Drubin, D.G., and Barnes, G. (2007). Structures and functions of yeast kinetochore complexes. Annu Rev Biochem *76*, 563-591.

Whelan, G., Kreidl, E., Wutz, G., Egner, A., Peters, J.-M., and Eichele, G. (2011). Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. EMBO J *31*, 71-82.

Williams, G.J., Williams, R.S., Williams, J.S., Moncalian, G., Arvai, A.S., Limbo, O., Guenther, G., Sildas, S., Hammel, M., Russell, P.*, et al.* (2011). ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. Nat Struct Mol Biol *18*, 423-431.

Wong, J., Nakajima, Y., Westermann, S., Shang, C., Kang, J.-S., Goodner, C., Houshmand, P., Fields, S., Chan, C.S.M., Drubin, D.*, et al.* (2007). A protein interaction map of the mitotic spindle. Mol Biol Cell *18*, 3800-3809.

Woo, J., Lim, J., Shin, H., Suh, M., Ku, B., Lee, K., Joo, K., Robinson, H., Lee, J., and Park, S. (2009). Structural Studies of a Bacterial Condensin Complex Reveal ATP-Dependent Disruption of Intersubunit Interactions. Cell *136*, 85-96.

Wu, N., Kong, X., Ji, Z., Zeng, W., Potts, P.R., Yokomori, K., and Yu, H. (2012). Scc1 sumoylation by Mms21 promotes sister chromatid recombination through counteracting Wapl. Genes & Development *26*, 1473-1485.

Xia, G., Luo, X., Habu, T., Rizo, J., Matsumoto, T., and Yu, H. (2004). Conformation-specific binding of p31(comet) antagonizes the function of Mad2 in the spindle checkpoint. The EMBO Journal *23*, 3133-3143.

Xiao, Z., McGrew, J.T., Schroeder, A.J., and Fitzgerald-Hayes, M. (1993). CSE1 and CSE2, two new genes required for accurate mitotic chromosome segregation in Saccharomyces cerevisiae. Mol Cell Biol *13*, 4691-4702.

Xiong, B., Lu, S., and Gerton, J.L. (2010). Hos1 Is a Lysine Deacetylase for the Smc3 Subunit of Cohesin. Current Biology, 1-6.

Xu, Z., Cetin, B., Anger, M., Cho, U.S., Helmhart, W., Nasmyth, K., and Xu, W. (2009). Structure and function of the PP2A-shugoshin interaction. Mol Cell *35*, 426-441.

Yamamoto, A., Guacci, V., and Koshland, D. (1996). Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). J Cell Biol *133*, 99-110.

Ye, H., Chen, T.C., Xu, X., Pennycooke, M., Wu, H., and Steegborn, C. (2004). Crystal structure of the putative adapter protein MTH1859. J Struct Biol *148*, 251-256.

Yoshimura, S.H., Hizume, K., Murakami, A., Sutani, T., Takeyasu, K., and Yanagida, M. (2002). Condensin architecture and interaction with DNA: regulatory non-SMC subunits bind to the head of SMC heterodimer. Curr Biol *12*, 508-513.

Yuan, Y.R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P.J., and Hunt, J.F. (2001). The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. J Biol Chem *276*, 32313-32321.

Zaitseva, J., Jenewein, S., Oswald, C., Jumpertz, T., Holland, I.B., and Schmitt, L. (2005). A molecular understanding of the catalytic cycle of the nucleotide-binding domain of the ABC transporter HlyB. Biochem Soc Trans *33*, 990-995.

Zakeri, B., Fierer, J.O., Celik, E., Chittock, E.C., Schwarz-Linek, U., Moy, V.T., and Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc Natl Acad Sci USA *109*, E690-697.

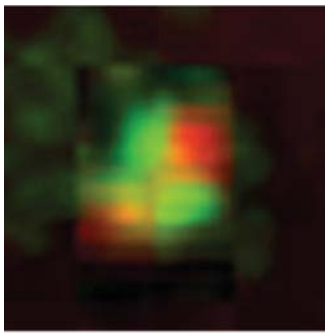
Zhang, N., Kuznetsov, S.G., Sharan, S.K., Li, K., Rao, P.H., and Pati, D. (2008). A handcuff model for the cohesin complex. J Cell Biol *183*, 1019-1031.

Zhang, N., and Pati, D. (2012). Sororin is a master regulator of sister chromatid cohesion and separation. Cell cycle (Georgetown, Tex) *11*.

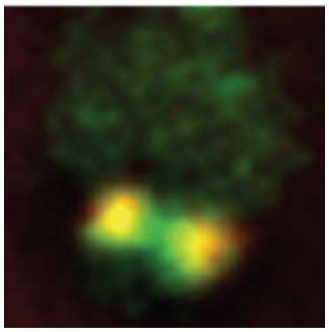
Zoghbi, M.E., Krishnan, S., and Altenberg, G.A. (2012). Dissociation of ATP-binding cassette nucleotide-binding domain dimers into monomers during the hydrolysis cycle. J Biol Chem *287*, 14994-15000.

Appendix

A

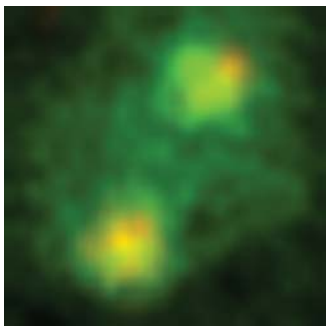


Smc3(WT)-GFP
Mtw1-RFP



Smc3(E1155Q)-GFP
Mtw1-RFP

B



Smc1(E1158Q)-GFP
Mtw1-RFP
Scc2-TetR

Figure A-1. Localisation of hydrolysis-defective cohesin in live diploid cells.

(A) Wild type GFP-tagged Smc3 forms fluorescent metaphase barrels corresponding to pericentromeric cohesion, whereas the Walker B mutant forms a focus at the kinetochore, marked by Mtw1-RFP.

(B) Recruitment of Scc2-TetR to the TetO is insufficient to colocate the hydrolysis-defective Smc1(E1158Q)-GFP.

The above experiments were performed by Dr. Bin Hu at the University of Oxford.

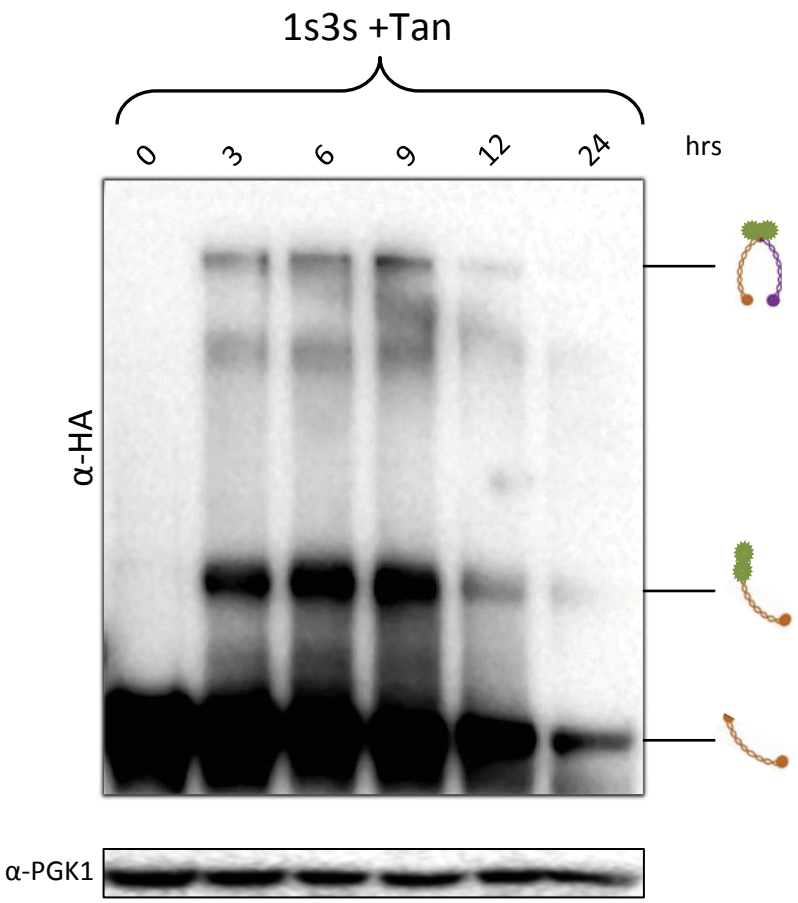


Figure A-2. Cross-linking of cohesin hinge domains *in vivo.*
Western blots of TCA preps of strains grown in YPR to which galactose has been added to the indicated time. Monomeric SpyCatcher is induced and cross-links to the SpyTag in the Smc3 hinge, producing a higher molecular weight product that is stable and approximately 50% of total 3s-HA6 by 24hrs. Tandem SpyCatcher induction produces 2 cross-linked bands assigned the identities of cross-linked heterodimer, and singly-reacted 3s-HA6. Levels of both unreacted and cross-linked proteins decrease over time when Tan is induced.

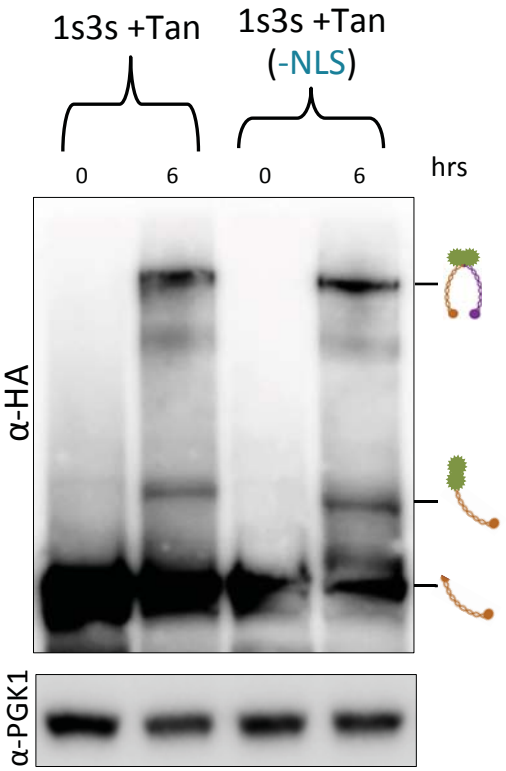


Figure A-3. The nuclear localisation sequence (NLS) of Tandem SpyCatcher has no effect on cross-linking of 1s3s.
Tandem SpyCatcher (+/- NLS) expression was induced by galactose and the extent of cross-linking between 1s and 3s after 6hrs was determined by western blot against HA-tagged 3s. The similar proportions of each product indicate the NLS does not affect cross-linking.

Acknowledgements

I would like to express my gratitude to a number of people who contributed to this work. I want to thank Kim Nasmyth for his stimulating intellectual input and for giving me the great opportunity to join his group, where I've learned a huge amount about how to conduct research. The wonderful members of the Nasmyth lab surely make it one of the most formative, interesting and lively environments in which to work. In particular I would like to thank: Fred Beckouët, for being a great help, guide, and friend over the years; Maurici Brunet-Roig, for being so friendly and generous with his time teaching me; and Sarah Dixon, for her dear friendship and sympathy as a fellow PhD student. There are many other past and present members of the lab that I would like to thank for their encouragement and discussion of research – I'm sure these people know who they are and how much they improved my time here.

I would also like to thank the Clarendon Fund and Cancer Research UK for generously funding this work, as well as those with whom I collaborated at various times: Mark Howarth, Jan Löwe, Susan Hancock, and Sheila Wang.

I am indebted to my collegiate friends (Derek, Craig, Mike, Robin, and Steve) for their kindness and for providing me with many happy memories of Oxford.

Finally, and most of all, I want to thank Monic and my family. Their love and care put everything in perspective and kept me going when things were most difficult.