The SUMOylation of the kinetochore and its effects on the chromosome segregation

Joseph Bareille

UNIVERSITÉ PARIS V DESCARTES

Année 2016

THÈSE

Pour obtenir le grade de

Docteur de l'UNIVERSITÉ PARIS DESCARTES V
Spécialité : Biochimie et Biologie Moléculaire

Ecole doctorale de rattachement : MTCE

Présentée et soutenue publiquement par :

Joseph BAREILLE

le 28 Novembre 2016

Intitulée

La SUMOylation du kinétochore et ses effets sur la ségrégation chromosomique

Directeurs de thèse : M. Nicolas Leulliot et Mme. Mary Dasso

Jury :

Mme Katja Wassmann : Rapporteur
M. Jean-Pierre Quivy : Rapporteur
Mme Mary Dasso : Examinateur
M. Julien Dumont : Examinateur
M. Nicolas Leulliot : Examinateur
M. Nicolas Leulliot : Directeur de thèse

All models are wrong but some are useful
George E. P. Box

TABLE DES MATIERES

Acknowledgments

First, I would like to sincerely thank the members of the comittee to honor me by agreeing to participate to my PhD thesis defense. The time they spent to read my manuscript led to relevant comments and questions that definitely increased the quality of my work.

I would like to thank Dr. Hinnebusch to be an active member of the CBMP branch.

I would like to greatly thank Dr. Dasso for let me join your team and your mentoring. Thanks to you, I worked into the prestigious NIH and enjoyed unbelievable times. I learned new field, new methods and became a better scientist. Thanks for these 3 wonderful years.

I greatly thank our collaborators for their help.
I would like to thank the Pr. Musacchio and Dr. Weir to provide enough cenp-H/I/K/M to perform the SUMOylation assays even if the conclusions of the experiments were not outstanding.
I would like to thank the Dr. Stuckenberg and his team to give us the plasmids carrying every proteins of the inner kinetochore. I assume that it was such a pain to do the monstruct.

Je ne remercierais jamais assez le Pr. Leulliot. Il m'a accepté dans son laboratoire alors que je n'étais qu'étudiant. Il m'a aidé à devenir un professionnel de la Science. Quelques années plus tard, il est venu me “débaucher”, persuadé de mes qualités. Et encore une fois grâce à lui, j'ai pu franchir une nouvelle étape en mûrissant et m'affirmant. Pendant tout ce temps, il a su me canaliser (…) et m'aiguiller. Lorsque je suis venu lui parler de mon désir de quitter un poste pourtant enviable dans son équipe et de réaliser une thèse aux États-Unis, il m'a soutenu. Même séparé par un océan, il a toujours été là quand j'en avais besoin malgré le décalage horaire, ses nombreuses responsabilités et sa vie de famille. Nicolas, le respect et la gratitude que j'ai pour toi sont sans bornes.

Tout naturellement, je remercie mes parents, Martine et Gabriel. J’éviterais la très convenue liste des raisons pour lesquelles je dois vous remercier. Je me contenterais d'un merci profond et sincère.
Vient ensuite la dream team, aussi appelée famille Le Cunff. Christophe et Alexandrine, je vous remercie. En des temps incertains, vous avez été apaisant et un exemple de force tranquille, toujours prêt à aider. Le casting ne serait pas complet sans les loulous. Mélanie, Vincent, merci d'avoir fait les jeunes et en avoir fait voir de toutes les couleurs à vos parents, j'ai bien rigolé ! Votre avenir est devant vous, il sera ce que vous en ferez. Quelque soit la voie que vous choisissez, allez y à fond.

Je remercie la famille Barrère, (beaucoup) trop nombreuse pour permettre des remerciements individuels. Je m'en excuse. Je vous remercie tous. Une pensée émue a Mamie.

Je remercie la famille Bareille. Mamie Romaine, merci pour tout. Le temps ne semble pas avoir prise sur toi, ton esprit toujours alerte m'impressionne.

I would like to thank my colleagues for these 3 years. Thanks to Maria who showed me the lab and methods. It was a pleasure to see you in and out of the lab. Congrats for your green card and good luck for your next step. Thanks to Maia. Always smilingly, ready to help and able to order things for yesterday, I really enjoyed this time in your presence. Ming-Ta, beyond your quiet attitude, I found a wonderful person who is the first when it is about helping people. I really regret to not work with you longer. I wish you the best for your new life ! Hope to see you soon. Thanks to Vasilisa, my dear bench neighbor. Thanks to teach me cell biology methods. I really appreciate to speak and share my space with you. My only regret, there was not enough happy hours. Good luck for the end of your post-doc (and kisses to Mikhail). Thanks to Alex. Thanks to you, I went deeper into the american culture. It seems almost not weird now. Good luck for your PhD. Hope it will be a productive one. To conclude, I would like to thank Saroj and Shane. You are my favorite indian and korean, the new Laurel and Hardy, ready to storm the lab. Thanks for the beers (to Saroj) and no thanks to skip them (to Shane). Be a dad is not an excuse, you could bring Amelia with you ! It was short but intense. I greatly enjoyed this year. Best wishes to both of you.

I thank every members of the team of Dr. Serpe, Dr. Levin, Dr. Shi and Dr. Yun-Bo. I hope you enjoy as I did working in the marvelous building 18T … Thanks to the post-docs/PhD students of these team. Hope to see you all at least one more time. I greatly thank Dr. Serpe for her advices. Merci Caroline pour avoir partagé ce bâtiment durant ses 3 années. Félicitations et bonne chance pour la suite. Vous êtes la bienvenue à Lyon. (la bise à Alex).

I thank all people that I met at the NIH. I greatly enjoyed speaking with you, it opens your mind to meet people from every corner of the world !

I would like to thank the Dr. Azuma. Our talk was amazingly inspiring scientifically and personally. Thanks to share your experience and knowledge. And of course, thank to give the right protocol for E1 purification.

I would like to thank the persons from NIH that shared their knowledge. Vincent Schram, je te remercie d'avoir été disponible pour m'enseigner ton art de la microscopie. On ne devient pas chef par hasard.
Valentin Jaumouillet, je te remercie pour avoir partagé ton expérience. Tes compétences et tes talents de chercheur ne font aucun doute, je suis certain que tu trouveras enfin un poste à ta juste valeur.
Je remercie Mme Basseville de m'avoir enseigner les rudiments de la biologie cellulaire. Grâce à elle, culture de cellule HeLa ne rime pas avec *Escherichia*.

Je profite de cet espace pour remercier les personnes avec lesquelles j'ai eu l'occasion de travailler ou que j'ai côtoyé avant mon aventure américaine.
Une pensée pour l'équipe du Pr. Van Tilbeurgh et les autres du feu IBBMC. Merci Pr. Van Tilbeurgh de m'avoir accueilli dans votre équipe. On peut difficilement imaginer meilleure première expérience professionnelle.
Je vous remercie tous. Par manque de place, je ne citerais que ceux avec qui j'ai directement collaboré, Bruno Collinet, Nourredine Lazar, Wenhua Zhang et l'équipe Forterre. Ce fut un plaisir et un honneur. Une petite pensée à celle qui a l'œil du tigre et à la fumeuse adepte de high fives et autres checks.

Merci au LCRB et à l’équipe du Pr. Leulliot. J'ai passé d'excellent moments aussi bien en M2 que par la suite. Merci a Michèle, Patrice, Magali, Lila, Medhi, Stéphane et Julie. Mention spéciale pour Stéphane Réty. Tu m'as initié à ton art du dénigrement. Malheureusement, le monde n'est pas encore prêt. JRP, je te remercie tant que tu es encore en liberté. Le monde n'est pas prêt pour toi aussi. Merci aux étudiants que j'ai rencontré et encadré. Mention spéciale à Florian Chardon. À force de maltraitance, je pense avoir fait quelque chose de toi.
En espérant assister à une autre réunion d’équipe au Pantalon !

Merci aux autres équipes. Désolé vous êtes trop nombreux pour être tous citer. Les 2 ans passés à vos cotés sont inoubliables.

Qui de mieux pour conclure que ma tendre et chère. Cathy, je te remercie. Ça a été compliqué mais on l'a fait. Sans toi, tout ça n'aurait pas été possible. Merci pour tout. Cette these est un peu la tienne aussi. Gloire au MC.

List of Abbreviations

ATP: Adenosine TriPhosphate

CATD: cenp-A Targeting Domain
CCAN: Constitutive Centromere Associated Network
cenp: CENtromere Protein
CIN: Chromosome Instability

DNA: Deoxyribo Nucleic Acid
DTT: DiThioThreitol

EDTA: EthyleneDiamineTetraacetic Acid
EM: Electron Microscopy

GAP: GTPase-Activating Protein
GF: Gel Filtration
GFP: Green Fluorescent Protein
GST: Glutathion S-Transferase

HIC: Hydrophobic Interaction Chromatography
HOR: High Order Repeat

IEC: Ion Exchange Chromatography
IMAC: Immobilized Metal ion Affinity Chromatography
IPP: Inorganic PyroPhosphate

kDa: kilo Dalton

LB: Luria-Bertani

MBP: Maltose Binding Proteins

MW: Molecular Weight

NICHD: National Institute of Child health and Human Development

OD: Optical Density

ORF: Open Reading Frame

PBS: Phosphate-Buffered Saline

PCR: Polymerization Chain Reaction

PIAS: Protein Inhibitor of Activated STAT

PFA: ParaFormAldehyde

RCF: Rotating Centrifuge Force

RNA: Ribo Nucleic Acid

RT: Room Temperature

SAC: Spindle Assembly Checkpoint

SDS: Sodium Dodecyl-Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate-Poly Acrylamide Gel Electrophoresis

SEC: Size Exclusion Chromatography

SENP: SENtrin Protease

SPB: Spindle Pole Body

SUMO: Small Ubiquitin-like Modifier

TAE: Tris-Acetate-Ethylenediaminetetraacetic acid

TE: Tris Ethylenediamine tetra acetic acid

TNT: Tris NaCl Tween20

XEE: Xenopus Egg Extract

List of Figures

.

List of Tables

General Introduction

The cell cycle is divided in two phases, the interphase and the mitosis, or cell division. During interphase, the cell grows and duplicates its genetic information sustained by the chromosomal DNA. During mitosis, a set of chromosomes is transferred to each daughter cell, a critical step called chromosome segregation. The chromosome missegregation leads to an inappropriate number of chromosomes and subconsequent defects such as cell death.

The interface between the chromosome and the microtubules is a key point in order to regulate the accuracy of chromosome segregation. More than one hundred proteins are involved in this interface, so called kinetochore. A higher level of regulation is driven by Post Translational Modifications (PTM), which tightly control its architecture and function. The SUMOylation is one of the PTM identified in the kinetochore and alteration of the SUMOylation pathway leads to severe defects during mitosis. SUMO is an Ubiquitin-like protein that can conjugate to lysines of other cellular proteins. These conjugated species can be cleaved by SUMO proteases, called SENPs in vertebrates, which ensure a rapid turnover and the specificity of SUMOylated species. Human SENP6 and SENP7 act specifically on the disassembly of poly-SUMO chains. PolySUMOylated species are also substrates of SUMO-targeted ubiquitin ligases, leading to their proteasomal degradation. SENP6-depleted HeLa cells have defects in metaphase chromosome congression and alignments due to the loss of kinetochore components such as the CENP-H/I/K/M complex by a proteasome-mediated proteolysis. The removal of the CENP-H/I/K/M complex leads to the disruption of the kinetochore.

The objective of my thesis was to determine the role(s) of the poly-SUMOylation of the CENP-H/I/K/M on the inner kinetochore regulation. We hypothesized that this SUMO-dependent degradation is a time or a quality control step, disrupting lagging kinetochore and thus restarting the process, or removing excess/unproper loaded CENP-H/I/K/M complex and finally revealing the mechanisms of this poly-SUMOylation.

To demonstrate that the SUMOylation pathway may promote kinetochore assembly by enhancing the proteinaceous interactions, my first goal was to understand the roles and the

mechanisms of protein poly-SUMOylation in the kinetochore. To reveal the roles and effects of poly-SUMOylation of CENP-I, I tried to identify the SUMOylated lysines of CENP-I and investigated the effects of mutations. Studying these mechanisms *in vivo* allowed us to work as close as possible of the endogenous levels of protein expression as well as the normal cell physiology. Nevertheless, due to technical issues, the identification could not be performed by immuno-precipitation coupled to Mass Spectrometry analysis. Thus, I performed the SUMOylation of CENP-I *in vitro*. CENP-I is a member of the CENP-H/I/K complex, involved in the Constitutive Centromere Associated Network (CCAN), a protein complex mandatory for cell viability. I tried to produce the entire and parts of the CCAN in bacterial system and purify them. These complexes remain insoluble or too instable for biochemical assays. At this time, the Pr. Musacchio's team identified a fourth partner of the CENP-H/I/K complex, CENP-M. He determined purification protocol and kindly provided us the complex. I reconstituted the SUMO pathway, performed a SUMOylation assay of the CENP-H/I/K/M complex and products were analyzed by Mass Spectrometry analysis. Even if the SUMOylation of CENP-I was confirmed by SDS-PAGE, no SUMOylated lysines were identified by Mass Spectrometry analyses.

In parallel, in order to study the SUMOylation of the kinetochore, I attempted to reconstitute kinetochores *in vitro* according to a published protocol (A. Guse, Fuller, and Straight 2012). These *in vitro* kinetochores were too instable to be suitable for biochemical assays. To circumvent this problem, I adapted the method used by the teams of Dr. Fukagawa and Dr. Cheeseman. I fused the kinetochore protein CENP-T to the lactose repressor, targeted this protein to beads coupled to lactose operon DNA templates and determined if this complex can recruit the kinetochore proteins as CENP-T does *in vivo*.

In a first time, I will describe the mitosis and the kinetochore that is the interface between microtubules and centromeres. This structure is mandatory to ensure proper chromosome segregations. After description of its architecture that is composed of two main layers, the inner and the outer kinetochore, I will highlight the SUMOylation that is one of the post translation modification occurring in the kinetochore. I will summarize a publication showing that the mis-regulation of the SUMOylation leads to kinetochore defects. If this publication links the poly-SUMOylation to the degradation of kinetochore protein, the underlying mechanisms are unknown.

My goals are to understand them. In a second part, I will describe the methods used during my Thesis. Then, I will show the results of my experiments. This part is subdivided in two parts. The first is the SUMOylation of the cenp-H/I/K/M complex and the identification of the lysines involved in the SUMOylation. The second is the attempt to reconstitute *in vitro* kinetochores suitable for biochemical assays by using a published protocol and by adapting one used for *in vivo* studies.

I The chromosome segregation

1) The chromosome composition

In eukaryotes, the genetic information is encoded by the DeoxyriboNucleic Acid (DNA), which is located in the nucleus, a highly specialized compartment delineated from the cytoplasm by a lipid bilayer.

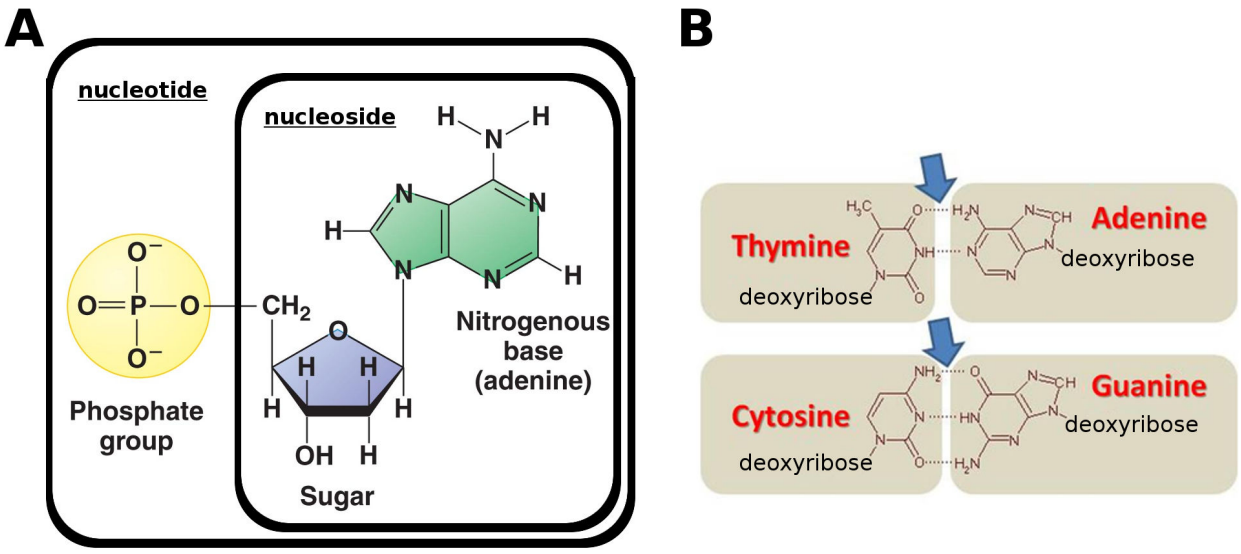
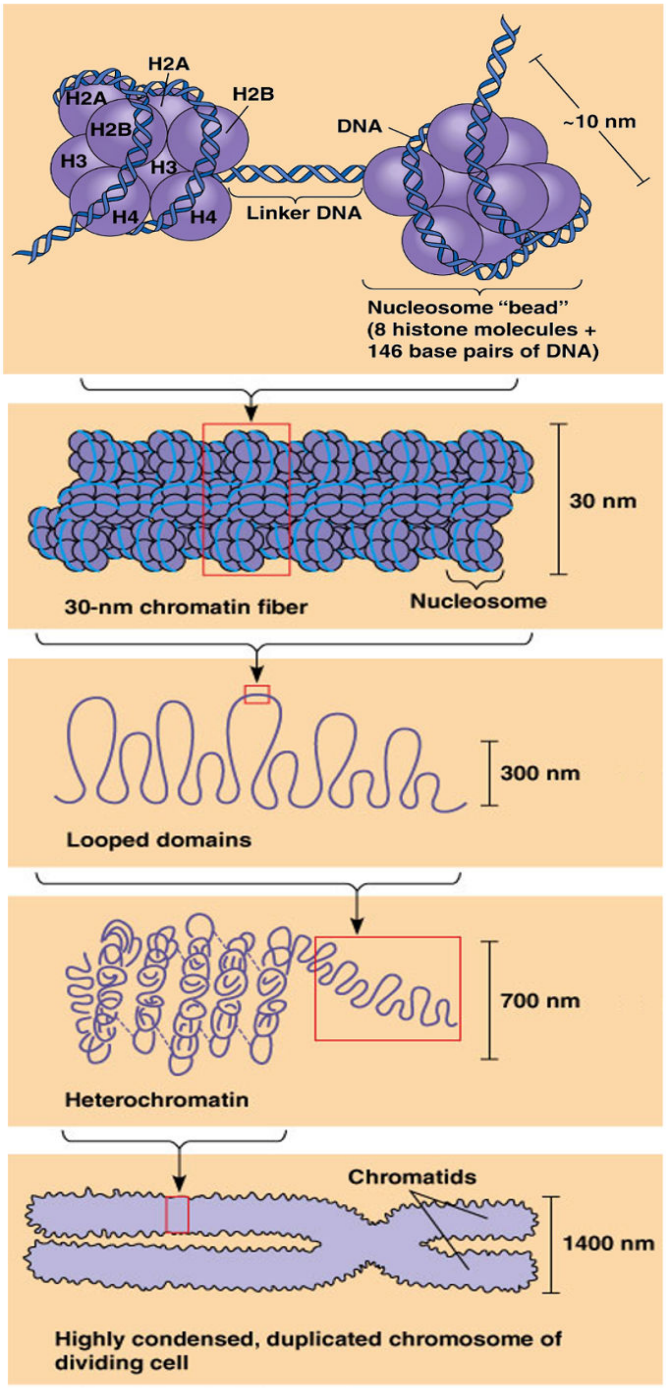


Figure 1: The DNA structure.
A) Bases and nucleotides. B) Base pairing.
From (dnarnanews.blogspot.com) and (Rosalind.info/glossary/base-pair)

2) The chromosome structure

If the human genomic DNA was unwound, its length would be around 3-meters long. Indeed, the genomic DNA total size is around 3 billion bases *(Lander et al. 2001).* Most of the DNA is contained in a packaged and organized structure called the chromosome. In order to constrain the whole genomic sequence into the nucleus, the DNA is wrapped around protein complexes called nucleosomes (Figure 2A). Nucleosomes are hetero-octamers that are composed of two units of proteins histone 2A and histone 2B, and a single unit of histone 3 and histone 4 (Figure 2B) (Arents et al. 1991). Histones 1 proteins increase the condensation of the nucleosome/DNA complex to form fibers of 30nm diameter *(Bednar et al. 1998)*. The scaffold proteins, such as condensins add another level of condensation to form the chromatin, leading to the final structure of the chromosomes *(Chen et al. 2014)*. The condensation of the chromosome varies during the

cell cycle. The variation can be local during interphase, depending on the specific regulations of genes or global during the mitosis, when the entire chromosomes are extremely condensed, inhibiting the transcription of any genes *(Antonin and Neumann 2016)*.



<u>Figure 2:</u> The chromosome structure.

Adapted from (https://mcb.illinois.edu/faculty/profile/cmizzen/) and (Pearson Education Inc., 2012)

The chromatin remodeling process is highly regulated and dynamic, involving transcription factors and chromatin associated proteins. The post-translational modifications such as histone acetylation regulate of the chromatin condensation/ decondensation state, thus allowing or not the transcription of a gene (Figure 3).

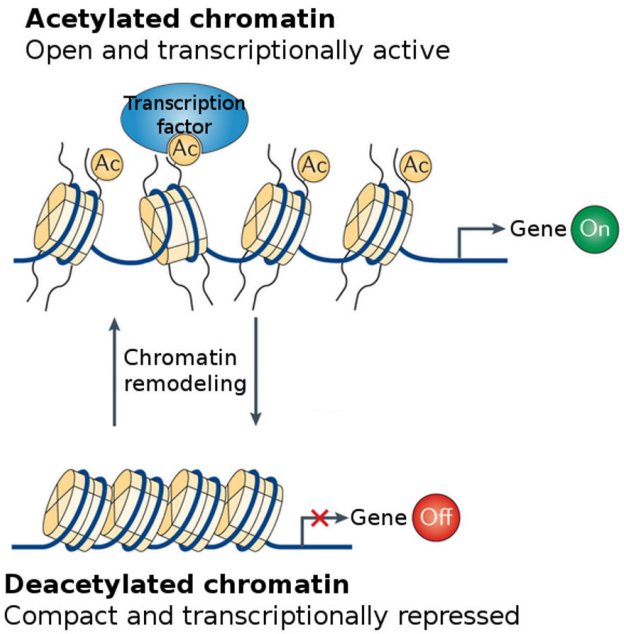


Figure 3: Example of chromatin condensation/decondensation driven by post-translational modifications of histones.
From (Verdin and Ott 2015)

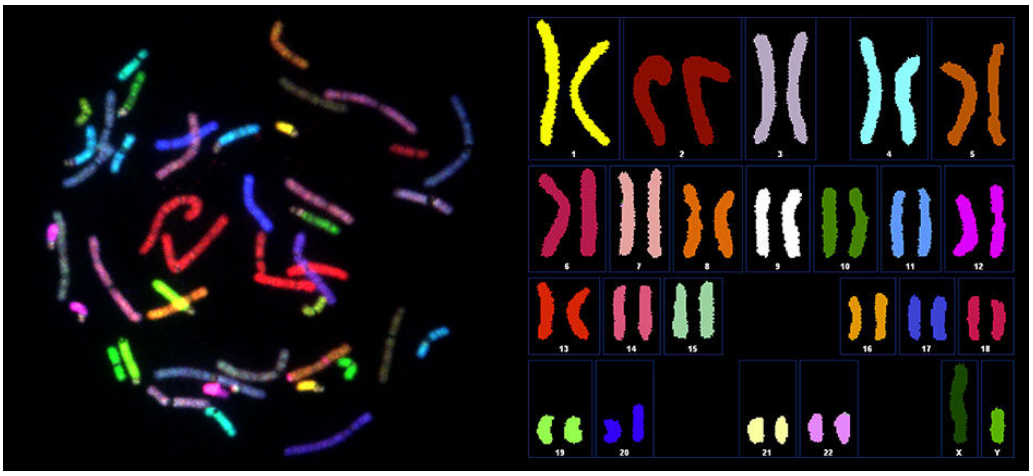


Figure 4: Karyogram of a human male.
From (http://www.humanlongevity.com/science-technology/science/human-genomics/)

Eukaryotic species can be characterized through the number and the appearance of their chromosomes, so called the karyotype (Figure 4). Human cells have twenty-three chromosomes pairs: twenty-two non sexual chromosome, called autosomes, and one sexual chromosome pairing, the allosomes; the mother and the father provide each of the two chromatids. Vertebrates have monocentric chromosomes that can be divided in three parts, the arms, the telomeres and the centromere. The arms contain most of the genes and are the main location of the condensation/decondensation processes and transcription events. The telomeres are the end of the chromosomes and are important marker for cell survival.

The definition of the centromere varies between geneticists, cell biologists or biochemists. Nonetheless, it could be defined as the place where the spindles bind chromosome during mitosis. Finally, in genetics, centromeres can also be defined as the regions excluded from the meiotic recombination process. The structure of the centromere is described in the kinetochore chapter (page 32).

The composition of the set of chromosomes is not random; it is a requirement for cell and organism survival. The disruption of the twenty-three pairs of chromosomes occurs during cell division leads to severe defects.

3) A proper set of chromosomes is mandatory for cell and organisms survival

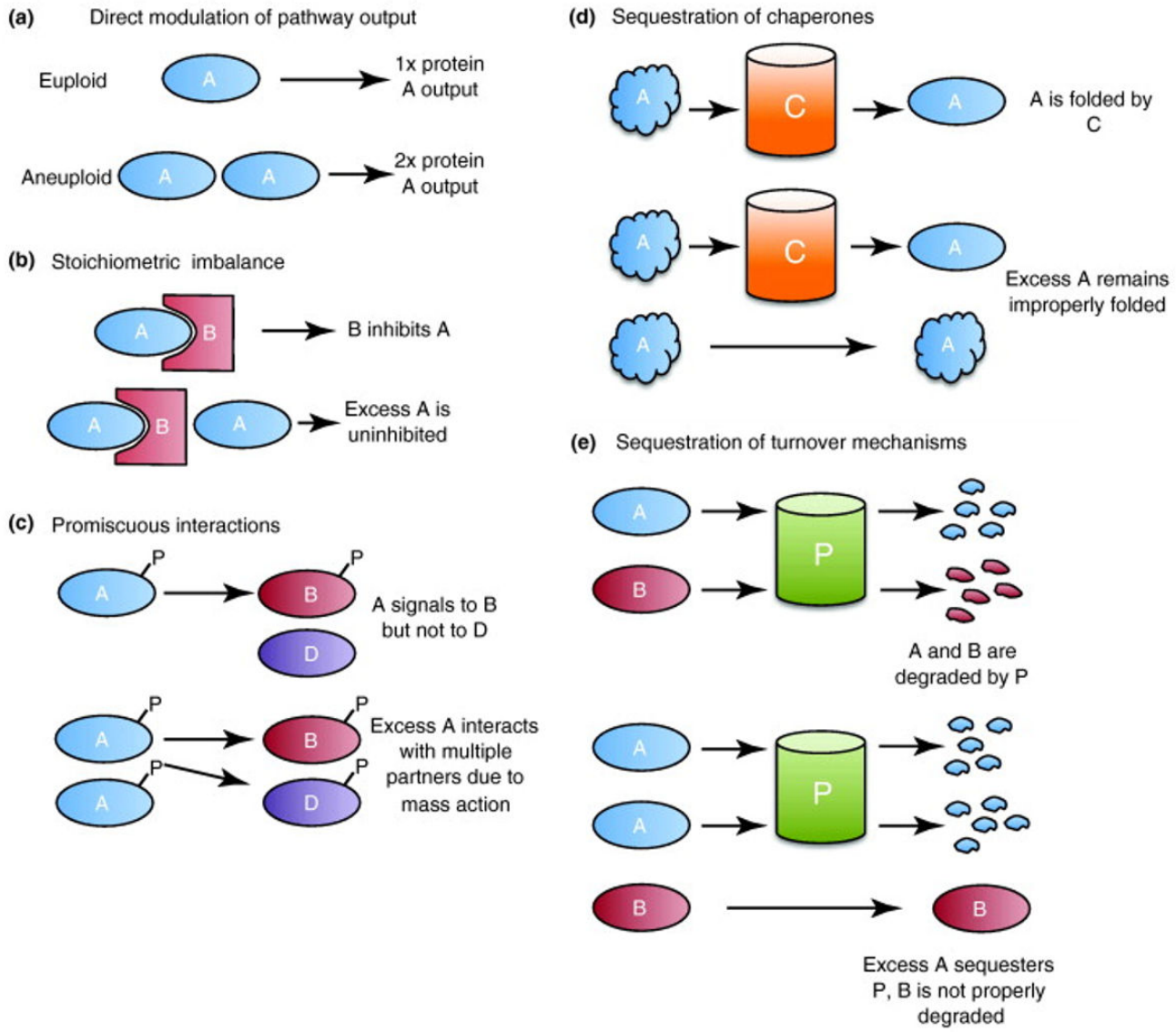
Chromosome mis-segregation is the non-proper separation between the sister chromatids. In that case, the number of chromosomes is disrupted in both daughter cells, a phenomenon called aneuploidy, leading to the misbalance of genes *(Torres et al. 2007)*. Sometimes only a part of a chromosome is mis-segregated but the consequences of aneuploidy are severe for the cells and the whole organism development.

The defects caused by aneuploidy can be explained by several hypotheses: a) the changes in the concentration of a specific protein may directly modulate its cellular functions (Figure 5A); b) the impairment of the copy number of a given gene can affect the formation or the normal functions of complexes that are sensitive to stoichiometry (Figure 5B); c) a protein excess could lead to its mis-localization (Figure 5C) through the binding of an additional target, leading to the mis-regulation of the associated pathway; d) the overwhelming of chaperones could lead to the improper folding of proteins and thus to the release of cytotoxic aggregates (Figure 5D); e) the

disruption in protein balance leads to a competition for processes, such as the ubiquitin-pathway (Figure 5E).

4) Aneuploidy in cancer cells

While aneuploidy leads to a decreased rate of cell proliferation in normal cells, this is also a hallmark of cancer, a disease of enhanced proliferative capacity *(Weaver and Cleveland 2006)*. Most cancers have an aneuploidy frequency higher than healthy cells, and more than 90% of solid tumors and 75% of hematopoietic cancers have gained or lost entire chromosomes *(Cleveland, Mao, and Sullivan 2003) (Weaver and Cleveland 2006) (Beroukhim et al. 2010) (Sheltzer and Amon 2011) (Gordon, Resio, and Pellman 2012) (Knouse et al. 2014)*("Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer" 2016). Despite the dramatic effects of the huge change in gene dosage, aneuploidy appears to confer proliferative advantages in certain circumstances.



<u>Figure 5:</u> Consequences of aneuploidy and its change in gene copy number.

A) An increased dosage of a single gene can increase the output or function of a cellular pathway. B) Altered gene dosage can interfere with the function of stoichiometry-sensitive complexes. C) Protein–protein interactions depend on the concentration of each binding partner. Altered expression of some proteins, such as signal-transduction kinases, might cause promiscuous molecular interactions that alter cellular phenotypes. D) Many proteins require chaperones to fold correctly. If aneuploidy overwhelms cellular chaperones, then misfolded proteins that escape chaperone- dependent folding might form insoluble and potentially cytotoxic aggregates. It is also possible that other essential clients of these chaperones remain unfolded. E) Quality-control mechanisms, including the ubiquitin–proteasome pathway, ensure that misfolded or improperly expressed proteins are rapidly turned over. Regulated protein degradation is also utilized to trigger various cellular programs, including cell cycle progression. The overabundance of certain proteins might interfere with the folding or turnover of other client proteins. *From (Sheltzer and Amon 2011)*

a) The relationship between aneuploidy and Chromosomal Instability

While the physiological consequences linked to aneuploidies are well defined, the real causes are still debated. An improper environment, a deleterious mutation in a protein essential for mitosis regulation or genes misbalances lead to Chromosomal Instability (CIN) during segregation *(Wu, Ooi, and He 2004) (Liu et al. 2009)*. It is to note that CIN and aneuploidy are related but different. Indeed, while aneuploidy results in an improper set of chromosome within a cell, CIN is defined as the instability of chromosome segregation, involving a higher probability of mis-segregation.

Aneuploidy is frequently associated to cancer and is considered as a hallmark of cancer, thus aneuploidy confers advantages to these cells. Aneuploidy leads to heterogeneity inside this cell population. This heterogeneity increases the probability of cancerous cells resistant to treatment.

b) Proteins/pathway leading to aneuploidy/CIN

The exact mechanisms leading to aneuploidy or CIN are still debated. However teams identified that several proteins are often mutated in aneupoid cancer cells. The depletion of these proteins leads to aneuploidy or increase CIN. Most of these proteins are associated with the mitosis process or with the cell cycle *(Orr and Compton 2013).*

To conclude, it is still unknown if aneuploidy is a cause or a consequence of cancer. However, some studies have identified molecular mechanisms leading to CIN persistence *(Thompson, Bakhoum, and Compton 2010) (Orr and Compton 2013).* They described that some of the key signaling proteins are the main actors of mitosis, such as Mad2 our Mps1, that are both involved in the Spindle Assembly Checkpoint. This protein complex is in charge of blocking the cell division until the conditions required for proper chromosome segregation is attained.

5) Cell division, from one mother to two daughter cells

Two types of divisions occur in eukaryotes: mitosis and meiosis, but I will not describe the latter that is a particular case of mitosis occurring during the gametogenesis. The human body is composed of billions of cells, all emerging from a unique cell coming from the fecundation. The cells need to transfer to each daughter cell one set of chromosome as well as a functional set of

proteins. This transfer has to be extremely accurate, if not it leads to aneuploidy and its associated defects.

The cell division process involves two phases, the interphase and the mitosis, and different cellular checkpoints (Figure 6). The interphase is sub-divided in 3 phases: the Gap1 phase, the Synthesis phase and the Gap2 (G2) phase. During the G1 phase, cells grow and synthesize mRNA and proteins that are required for chromosome duplication. The Synthesis phase is specifically dedicated to DNA duplication. The duplicated chromatins are called chromatids or sister chromatids. The goal of the Gap 2 phase (G2 phase) is to prepare the cell for its equal division. During this phase, the cell grows fast and massively synthesizes proteins. Once is reached, the cells move to the mitosis *(Ford and Pardee 1999)* and *(Hershko 1999).*

6) Steps of the cell division

Mitosis is challenging and is composed of a lot of critical steps allowing the proper transmission of a defined set of chromosome that is absolutely required for the viability and functionality of the daughter cell. These processes involve many steps of regulation, so called checkpoints, whose description is beyond the scope of my thesis. Mitosis is composed of the prophase, metaphase, anaphase and telophase. The *prophase*, *metaphase*, and *anaphase* were initially used in 1884 in the first description of the process of chromosome distribution during cell division *(Strasburger 1884)*.

a) The prophase

In order to support the mechanic stress of the segregation process, chromosomes have to become rigid. During the first mitotic step, the chromatin is extremely condensed (3% of its average volume). The evolutionary conserved condensin complexes and its regulation modulate the chromosome shape and the condensation during mitosis *(Hirano, Kobayashi, and Hirano 1997)* and *(Aragon, Martinez-Perez, and Merkenschlager 2013)*. The condensin complexes form loop wrapping the chromosome DNA. Consequently, the chromosomes become visible through wide-field microscope. Each centrosome is pushed apart to the opposite pole of the cell.

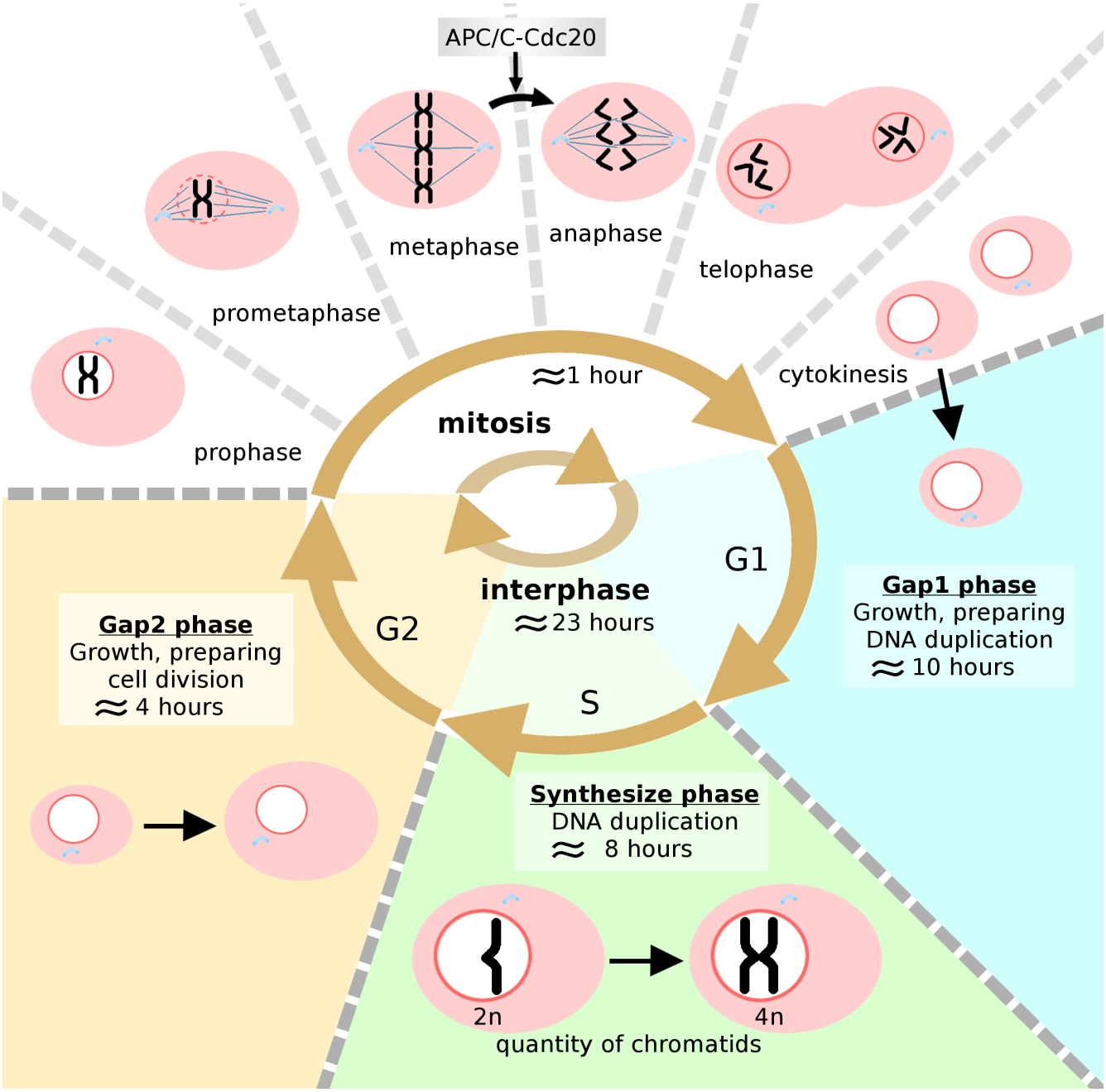


Figure 6: The different phases of the cell division.

b) The prometaphase

During the prometaphase, the nuclear membrane breaks down and forms vesicles, allowing the radiation of microtubules that are bound to the centromeres, from the centrosomes into the vacant space of the nucleus. The prometaphase is often considered as a transition state between prophase and the next phase, the metaphase, rather than a real mitotic phase. However, this step

is a major time point since the different key factors that are required for mitotic progression have now access to the chromosomes and its centromeres.

c) The metaphase

Chromosomes are pulled and pushed apart until they reach the equator of the cell. Once all chromosomes are properly aligned in the middle of the cell, they form the metaphase plate. The tension of spindles is crucial for their alignment and it implies that the sister chromatids are firmly attached. Only one misaligned chromosome is able to block mitosis. The cohesin complex ensures this attachment. The cohesin complex forms a ring that encircles both chromatids together *(Hirano 2016)*. The sensitivity to chromosome misalignment is ensuring by sensing of the microtubule tension through a protein complex called Spindle Assembly Checkpoint (SAC) localized at the interface between centromere and the spindles called kinetochore. It is to note that this step is extremely important. Chromosome misalignment leads to chromosome missegregation leading to aneuploidy and its severe associated defects in daughter cells. The sensing of chromosome misalignment that is the key point is done through the SAC.

d) Spindle Assembly Checkpoint, the keeper of the chromosome alignment

The Spindle Assembly Checkpoint is recruited at the centromere/microtubule interphase, called the kinetochore. The phosphorylation of the MELT residues of knl1 by the kinase mps1 leads to the recruitment of bub3. Bub3 functions as a platform to recruit and regulate the SAC.

The SAC controls the metaphase-anaphase transition. The SAC proteins are recruited to the tensionless kinetochore leading to the SAC activation (for review, *Musacchio and Salmon 2007)*. Activated SAC inhibits the complex promoting the release of sister chromatids attachment. To perform this inhibition, the SAC sequestrates cdc20, a protein required to form the APC/C-cdc20 complex that is an E3 ubiquitin ligase that degrade the cyclinB and securin (Figure 7). The securin inhibits the separase that is a protein degrading the cohesin complexes (Hornig et al. 2002). That leads to the removal of link between sister chromatids (Figure 7) and to the entry in the anaphase, allowing each set of chromosomes to target each pole. Defects of SAC recruitment or function leads to severe defects in mitosis, such as chromosome mis-segregation or mitosis delay. These mechanisms are extremely sensitive and dynamic. Indeed, the SAC is able to block the metaphase-anaphase transition, even if only one kinetochore is not bound to microtubule.

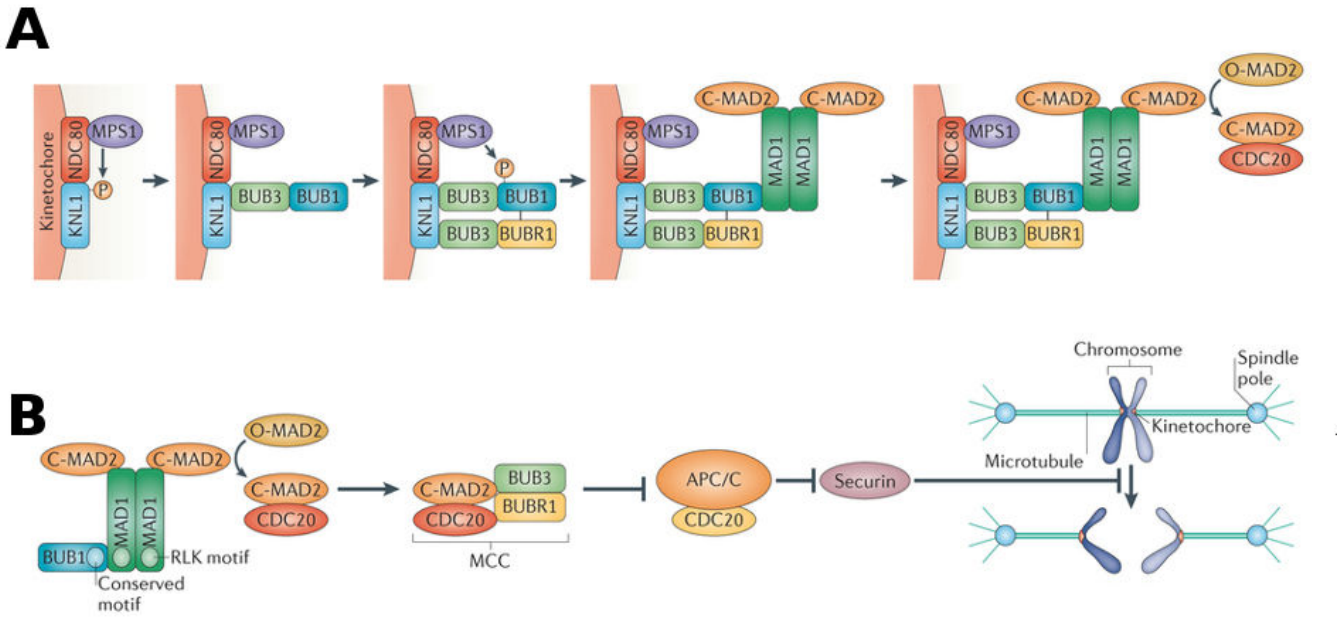


Figure 7: The Spindle Assembly Checkpoint.

A) Assembly of a functional checkpoint-signalling complex on kinetochores is stepwise. B) Once activated, the spindle assembly checkpoint sequestrates cdc20 and blocks the metaphase-anaphase transition through the inhibition of the E3 ubiquitin-ligase APC/C.
From (London and Biggins 2014)

e) End of mitosis

Once the chromosomes are perfectly aligned, the SAC is inactivated leading to the release of the Cell division cycle 20 (Cdc20) protein. The free Cdc20 activates the Anaphase Promoting Complex or Cyclosome (APC/C) that is an E3 ubiquitin ligase. APC/C-Cdc20 complex poly-ubiquitinates the securin leading to its degradation via the proteasome. Once the securin is degraded, the separase degrades the ring of cohesin. Thus, the chromatids are not bound anymore, a new phase of the mitosis starts, the anaphase and one set of chromosome can segregate to each pole. This regulation by the SAC is critical. Once each chromosome set moves to the cell pole, the cell moves to the telophase while the nuclear membrane is reconstituted around the chromosomes and a nucleus is formed in each daughter cell. At this time, the chromosomes are decondensed. Finally, the mitosis ends with the cytokinesis that describes the physical separation between each daughter cells, leads to the entry in G1 phase for both daughter cells. The kinetochore has several mandatory functions such as the microtubule attachment or the SAC recruitment.

II The kinetochore

In order to finely tune these tension forces, the kinetochore is the protein interface between spindle and centromere that is in charge of the regulation of the microtubules attachment.

1) The overall structure of the kinetochore

The kinetochore is the protein interface between the centromere and the microtubules, which assembles just after the nuclear membrane breakdown and persists until the end of mitosis *(K. E. Gascoigne and Cheeseman 2013).* While the core of the kinetochore is composed of the same proteins during the mitosis, their abundance varies through the process and some proteins are recruited or released at precise time points. The centromeres represents around 2% of the chromosome size and are transient complexes, making biochemical studies hard to perform.

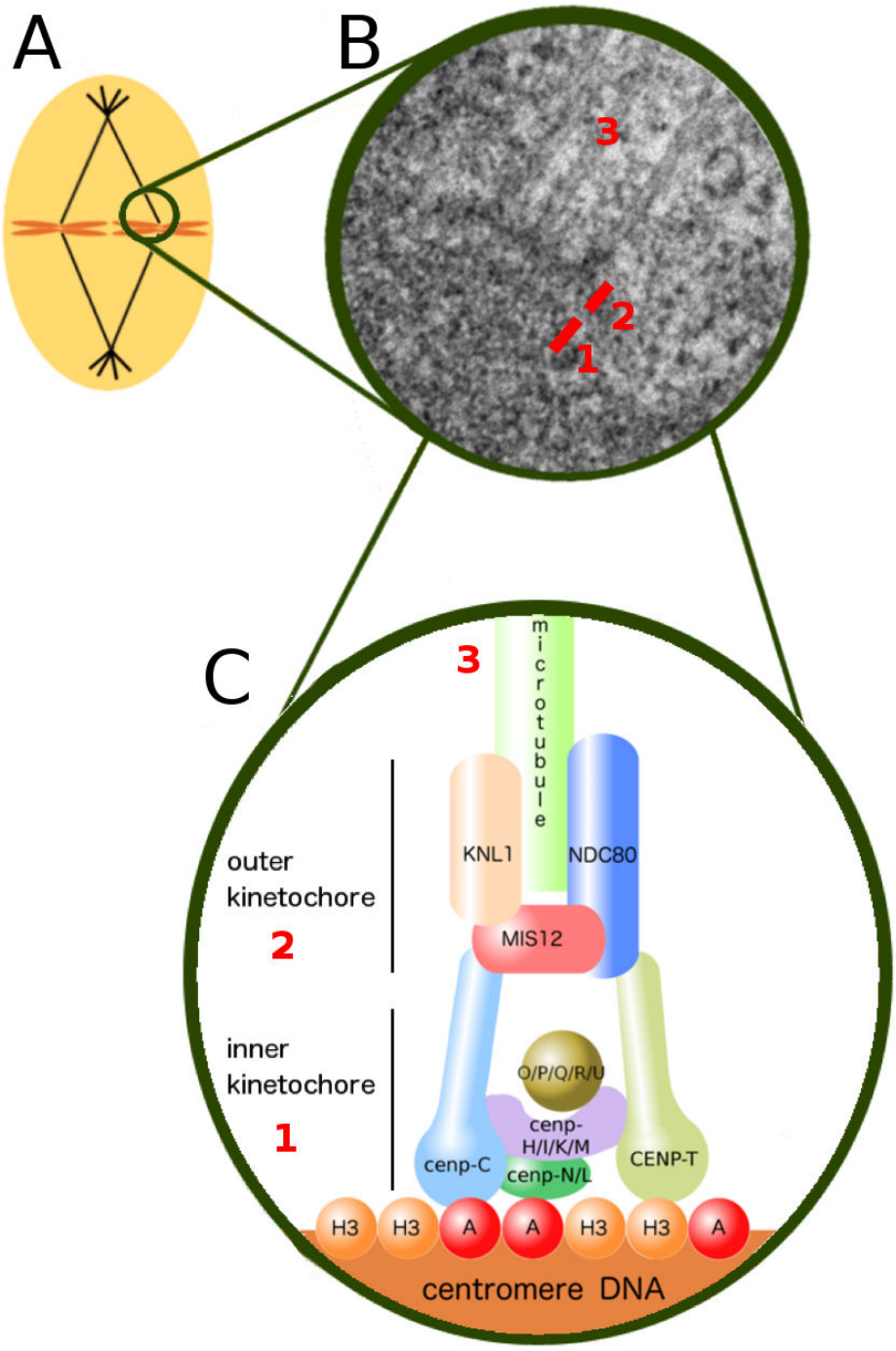
In 1985, studies revealed that the CREST disease (Calcinosis / Raynaud's phenomenon / Esophageal dysmotility / Sclerodactyly/Telangiectasia), a variant of scleroderma, is mostly due to a high frequency of auto-immune antibodies that are specifically targeted to the centromere during the whole cell cycle *(Moroi et al. 1980)*. The identification of these antibodies led to the identification of three CENtromere Proteins (cenp), cenp-A, cenp-B and cenp-C, with a molecular weight of 17, 80 and 140 kDa respectively *(Earnshaw and Rothfield 1985)*. The other kinetochore proteins were identified by co-Immuno-Precipitation (co-IP) and Pull-Down experiments. The proteins are mostly conserved among eukaryotes *(Schleiffer et al. 2012)*.

The kinetochore is composed of about a hundred proteins in vertebrate cells *(Iain M. Cheeseman and Desai 2008) (Karen E Gascoigne and Cheeseman 2011)* that belong to two main layers (figure 8):

- the **inner kinetochore** that is bound to the centromere. This layer anchors the kinetochore onto the centromere DNA to sustain the tension provided by the spindles. Some components of this layer are always bound onto centromeres during the cell cycle *(Moroi et al. 1980) (Fukagawa et al. 2001) (Okada et al. 2006).*

- the **outer kinetochore** that is bound to the spindles. This layer is the interface between the microtubules and the inner kinetochore. It controls the microtubules attachment and defines the location of some major checkpoints *(Musacchio and Salmon 2007)*.

The qualitative description of such a huge structure (around 100 proteins involved and dynamically interacting with each other) and also its quantitative characterization in terms of regulation, function and evolution through the cell cycle is challenging. I am starting the description with the centromere that is the chromosome part sustaining the kinetochore, moving to the inner kinetochore that is bound to the centromere through the cell cycle and finishing with the outer kinetochore binding both the inner kinetochore and the spindles.



<u>Figure 8:</u> The human kinetochore at metaphase.
A) Cell at metaphase. B) Electron micrograph of a kinetochore, the attachment of spindles onto the centromere. The kinetochore has 3 layers with different electron density. The inner kinetochore, bound to the centromere (1), the interstice (2), the outer kinetochore, bound to the spindles and the fibrous corona ("Biomedical EM Unit: Gallery" 2016). C) Scheme of the human kinetochore. H3: histone 3, A: cenp-A, 1: inner kinetochore, 2: outer kinetochore, 3: microtubule

2) The centromere

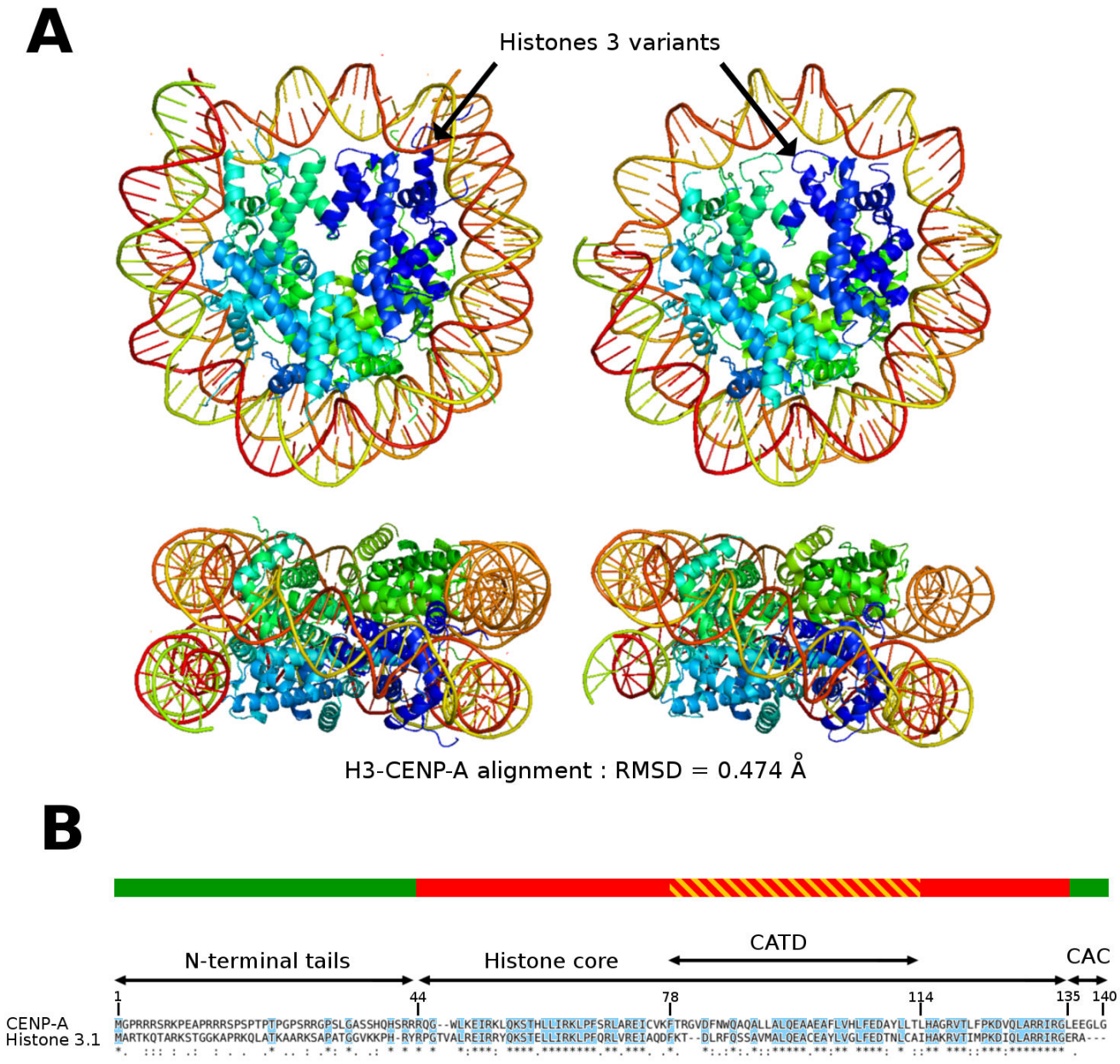
The centromere is defined as the place where spindles bind chromosome. This region contains specific proteins and repetitions of specific DNA sequences called the satellites sequences. The presence of these DNA sequences in all species strongly suggests that these sequences are important for centromere functions *(Melters et al. 2013)*. Nonetheless, the role of satellite sequences on the centromere varies through species.

*S. cerevisia*e chromosomes contain a single centromere that is composed of a 125 base pair sequence that is mandatory and sufficient to recruit the centromere proteins and to ensure its function *(Meluh and Koshland 1997).* In contrast, human centromeres that are composed by tandem repeats of DNA sequence, mostly the α-satellite DNA sequence represent approximately 2% of chromosome sequence *(Cleveland, Mao, and Sullivan 2003)*. The conservation of α-satellite sequences in all natural human centromeres suggests that repeats are required for centromere identity and/or function *(Schueler et al. 2001)*. These sequences are neither mandatory nor sufficient for accurate chromosome segregation *(Cleveland, Mao, and Sullivan 2003) (Fukagawa and Earnshaw 2014)*.

It appeared that epigenetic markers have more important roles than genetic factors in the regulation of the centromere (Shang et al. 2013; Karen E Gascoigne and Cheeseman 2011). Indeed, some of them are able to recruit a functional kinetochore such as the CENtromere Protein A (cenp-A). cenp-A is conserved among eukaryotes and is absolutely required for viability in animals (Stoler et al. 1995; Howman et al. 2000).

a) The cenp-A nucleosomes

The cenp-A protein is a histone 3 variant interacting with histones 2A, 2B and 4, to form a nucleosome that is only present at the centromere (figure 9.A). The presence of cenp-A in the centromere nucleosome, together with its mandatory requirement for cell survival, places the cenp-A protein as a master key of the centromere and is considered as the centromere marker. Thus, understanding the properties of the cenp-A nucleosome is essential to understand the kinetochore.



<u>Figure 9:</u> The structure of cenp-A nucleosome and domains of cenp-A.

A) Structure of the Histone 3 and cenp-A nucleosomes. Pdb : 1AOI and 3AN2. The figures are visulaized by the software pymol. B) Parts of cenp-A and the alignment of the sequences of histone 3.1 and cenp-A sequences. The text above the arrows are the name of each functional parts of cenp-A. The results of alignment through the software clustal is shown. The blue parts represent the identical residues. CATD: Cenp-A Targeting Domain. CAC: Cenp-A C-terminal tail.

The cenp-A protein shares 51% of sequence identity with the histone 3.1 in humans (figure 9.B) and their respective nucleosomes are structural homologues (Tachiwana et al. 2011). cenp-A is composed of three main functional domains, the histone core which is involved in the DNA binding, the cenp-A Targeting Domain (CATD) that recruits the critical component cenp-N and the cenp-A C-terminal tail (CAC) that is directly involved in the recruitment of a major protein of the kinetochore, cenp-C (Fachinetti et al. 2013; kato et al. 2013).

b) cenp-A recruitment

In vertebrate, cenp-A is deposited into the centromere at late telphase, early G1 phase in a two-step process. First, the dimer cenp-A/H4 binds the protein Holliday Junction Recognition Protein (HJURP) (Bernad et al. 2011; Tachiwana et al. 2015). Then, the HJURP/cenp-A/H4 complex binds the mis18 complex to achieve the deposition of cenp-A on the centromere by replacing H3.3, a cenp-A placeholder, forming a new nucleosome (Dunleavy et al. 2011). The mis18 complex is then released (Nardi et al. 2016; Stellfox et al. 2016). The loading of cenp-A occurs upon exit from mitosis. Interestingly, human HJURP can replace endogenous HJURP in Xenopus egg extracts, which demonstrates the conservation of the mechanism (Bernad et al. 2011; Tachiwana et al. 2015).

According to Dr. Nechemia-Arbely, ectopic cenp-A is removed from the chromosome arms during the S phase (communication during the ASCB annual meeting 2015) (Nechemia-Arbely, n.d.). The chromosome duplication may be a step to control the localization of cenp-A.

3) The inner kinetochore

The inner kinetochore links the centromere and the outer kinetochore that binds the spindles. It has a main core of 16 proteins that form the Centromere Constitutively Associated Network (CCAN). This complex remains at the centromere through the cell cycle. Some proteins such as chaperones or kinases such as Mis18 or Plk1 are also recruited into the inner kinetochore.

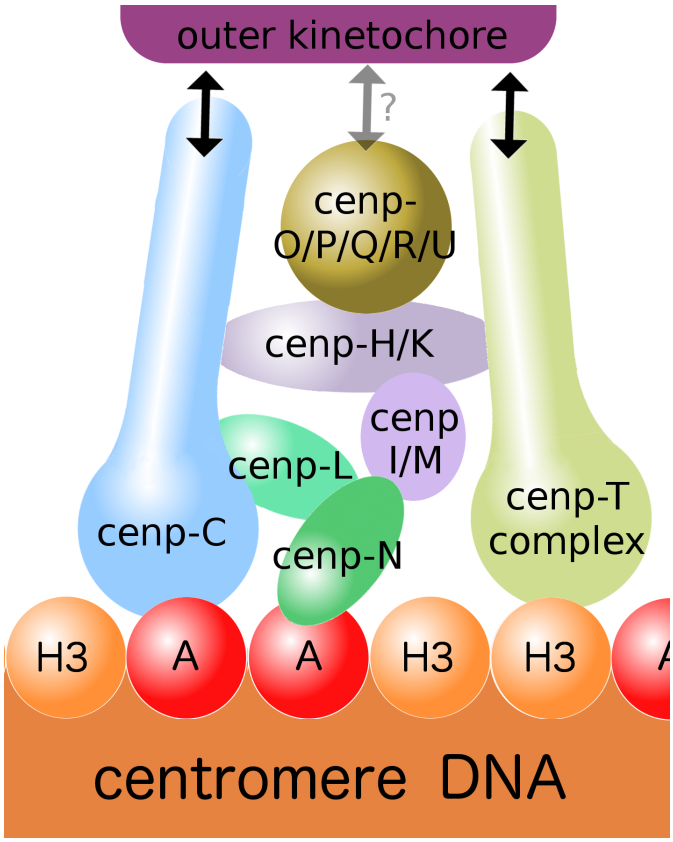


Figure 10: The Human inner kinetochore

a) Overall structure

The CCAN is composed of 1) two core proteins, cenp-C and cenp-T that bind the outer kinetochore to anchor it to the centromere and by 2) sub-complexes such as cenp-N/L and cenp-H/I/K/M involved in the kinetochore loading and cenp-R/O/P/Q/U involved in the regulation of the segregation process (I. M. Cheeseman 2014) (Figure 10). Its architecture could be considered as hierarchical but the depletion of a protein of each of these sub-complexes but cenp-R/O/P/Q/U leads to kinetochore disruption and blockade at metaphase. During the chromosome segregation, the force applied to the chromosome is transmitted by cenp-C and cenp-T, defining them as critical components of the kinetochore structure. Two proteins of the CCAN directly bind the cenp-A nucleosomes and are important for the kinetochore integrity: cenp-N and cenp-C.

b) The cenp-C protein

The cenp-C is conserved through eukaryotes and localized at the centromere through the cell cycle (Earnshaw and Rothfield 1985). The cenp-C protein has three known functions: 1) binding to cenp-A nucleosomes (kato et al. 2013), 2) interaction with the inner kinetochore (Karen E. Gascoigne et al. 2011), and 3) interaction with the outer kinetochore (Screpanti et al. 2011).

cenp-C binds the cenp-A nucleosomes (kato et al. 2013). During the mitosis, the N-terminal part of cenp-C recruits the Mis12 complex, a component of the outer kinetochore (Screpanti et al. 2011). Thus, cenp-C protein anchors the outer kinetochore to the centromere. It also interacts with the components of the inner kinetochore (McKinley et al. 2015) but little is known about these interactions.

The depletion of cenp-C leads to severe defects such as the chromosome misalignment and the presence of multipolar spindles in mitotic cells. The consequence of its depletion is the presence of micronuclei in daughter cells (McKinley et al. 2015). The depletion of cenp-C in mice leads to embryonic lethality (Fukagawa and Brown 1997; Kalitsis et al. 1998). During prometaphase, the cenp-C protein recruit by tethering the Mis12 complex in an Aurora B-dependent manner, a master key kinase involved in the regulation of the centromere, the kinetochore and the spindle midzone (Kim and Yu 2015; Emanuele et al. 2008).

c) The cenp-N and cenp-L proteins

Another mandatory kinetochore protein that binds cenp-A is cenp-N. cenp-N is a scaffold protein with no enzymatic activity that is required for the structural integrity of the kinetochore. The depletion of cenp-N leads to kinetochore disruption and chromosome mis-segregation (Carroll et al. 2009; Foltz et al. 2006; McClelland et al. 2007). The cenp-N protein specifically binds the cenp-A nucleosome in a DNA sequence independent manner. This interaction occurs mainly with two residues of the cenp-A Targeting Domain (CATD), arginine 80 and glycine 81, also called the RG loop. This domain is sufficient to recruit cenp-N (Fang et al. 2015).

When the affinity of cenp-N to cenp-A is reduced by mutation of the RG loop, cenp-N localization is disrupted, leading to defects similar to the cenp-N depletion by siRNA (Carroll et al. 2009). cenp-N forms a heterodimer complex with cenp-L that binds cenp-C (Hinshaw and

Harrison 2013; McKinley et al. 2015). These cenp-A, cenp-C and cenp-N/L network could coordinate or strengthen the kinetochore base.

d) The cenp-HIKM complex

While the cenp-N/L complex is bound to the centromere, the complex of cenp-H, cenp-I, cenp-K and cenp-M (cenp-H/I/K/M) is far more distant to the nucleosome (Sugata, Eisuke, and Todokoro, n.d.; Nishihashi et al. 2002; Okada et al. 2006; Okada et al. 2009) and (Basilico et al. 2014). For years, this complex was described as a heterotrimer composed of cenp-H, cenp-I and cenp-K and the experiments performed failed to reconstitute this complex *in* vitro. The studies led by Pr. Musacchio demonstrated elegantly that cenp-M is a partner of the cenp complex (figure 11) (Basilico et al. 2014).

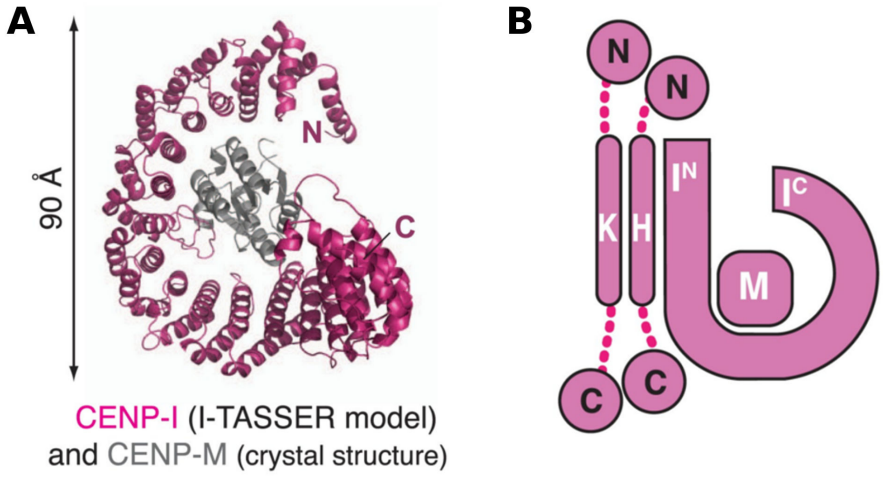


Figure 11: The heterotetrameric complex of cenp-H, -I, -K and -M.
A) Structural model of the cenp-I/M complex. cenp-M's pdb:4POT. B) Structural model of the cenp-H/I/K/M complex.
From (Basilico et al. 2014)

The depletion of one of these proteins leads to the mislocalization of the three others, the disruption of the kinetochore and the metaphase lagging in a similar way to cenp-N depletion (Iain M. Cheeseman and Desai 2008; Basilico et al. 2014). The depletion of cenp-N leads to aberrant metaphase and anaphase with a majority of cells presenting bipolar misalignment,

monopolar and multipolar alignments (Fukagawa et al. 2001; Nishihashi et al. 2002; Okada et al. 2006).

The cenp-H/I/K/M complex contains two blocks, cenp-I/M in one hand (figure 11.A) and cenp-H/K in the other (Basilico et al. 2014; McKinley et al. 2015). The structure of cenp-M is close to small GTPases but exhibits no enzymatic activity and is thus considered as a "pseudo G-protein". According to structural predictions, cenp-I shares structural homologies with the family of β-karyopherins, such as the Importin-β. The cenp-I/M interactions may be similar to those reported between Importin-β and the small GTPase Ran and that would suggest that cenp-I surrounds cenp-M, thus limiting its access to the solvent (Basilico et al. 2014). The cenp-H/I/K/M complex interacts with the cenp-N, cenp-T and with cenp-C (McKinley et al. 2015). It has been shown that the quantity of cenp-H/I/K/M varies during prometaphase and metaphase (Amaro et al. 2010; McAinsh and Meraldi 2011).

e) The cenp-O/P/Q/U/R complex

The cenp-O/P/Q/U/R complex is at the kinetochore during the whole cell cycle and is closer to the outer kinetochore than the cenp-H/I/K/M complex. These proteins are interdependent for their localization at the kinetochore (Okada et al. 2006; Tetsuya Hori, Okada, et al. 2008). While, the depletion of cenp-O/P/Q/U/R does not affect the kinetochore localization of other proteins of the CCAN, it leads to different defects in different cell lines (Kagawa et al. 2014). Embryonic stem cells of mice are unable to divide without cenp-O and the complete depletion of the complex leads embryonic death at the stage 7.5 day in mice. On the other hand, chicken DT40 cells and mouse dermal cells are viable without cenp-O (Kagawa et al. 2014; Okada et al. 2006). The exact roles of the cenp-O/P/Q/U/R complex remain unknown. It has been suggested that cenp-O/P/Q/U/R recruits kinases such as Plk1 (Kagawa et al. 2014).

f) The cenp-T/W/S/X complex

The cenp-T protein is a major component of the CCAN, anchoring the outer kinetochore to the centromere. The depletion of this protein leads to severe defects such as disruption of the kinetochore and blockade in mitosis. This protein shows two distinct functional parts (Figure 12), a C-terminal domain that is a histone-like core domain binding to the centromere DNA and a N-

terminal domain that is in charge of the binding with the outer kinetochore and also with the components of the CCAN previously described.

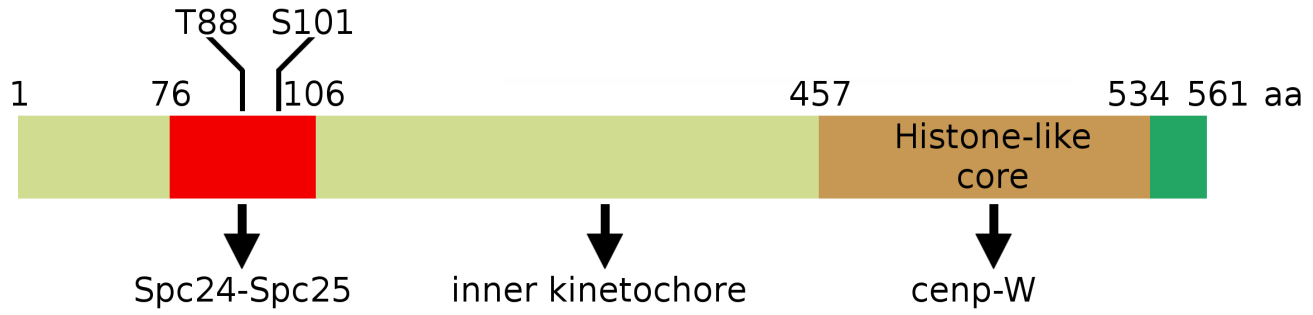


Figure 12: Human cenp-T and its domains.
The domain between aa 76-106 (in red) interacts with Spc24-25 of the outer kinetochore. The central part (light green) interacts with the inner kinetochore. cenp-T has an histone-like core (in brown) at its C-terminal part (aa 457-534) that binds cenp-W and then the centromere DNA. At the C-terminal tail, there is a conserved peptide among eukaryotes (in dark green).

Through its histone-like core, cenp-T forms a heterodimer with cenp-W that binds to the cenp-S/X complex and then together recruits the centromere DNA (Amano et al. 2009; Tetsuya Hori, Amano, et al. 2008; Nishino et al. 2012; Nishino et al. 2013; Takeuchi et al. 2014). This cenp-T/W/S/X complex is structurally related to nucleosomes (Nishino et al. 2012). The DNA binding activity of cenp-T/W is required for kinetochore localization (Tetsuya Hori, Amano, et al. 2008).

cenp-T localization is disrupted in cenp-W deficient cells (Tetsuya Hori, Amano, et al. 2008). The histone-core domain of cenp-T is sufficient for the localization at the centromere during the interphase and the mitosis. However, it cannot rescue the depletion of full length cenp-T. The recruitment of the cenp-T/W/S/X is poorly understood. However, it has been shown *in vitro* that the cenp-T/W/S/X complex binds with the same affinity the DNA surrounding by cenp-A and the histone 3 nucleosomes (Takeuchi et al. 2014).

cenp-T is involved in the recruitment of the outer kinetochore. This anchoring is mediated by interactions with the ndc80 complex that is mandatory for the structural integrity of the outer kinetochore and consists of four proteins, Ndc80, Nuf2, Spc24 and Spc25 (Janke et al. 2001).

The N-terminal part of cenp-T is involved in the recruitment of the outer kinetochore through the binding of Spc24 and Spc25, part of the NDC80 complex (Bock et al. 2012; Karen E. Gascoigne et al. 2011; Schleiffer et al. 2012; Nishino et al. 2013). The phosphorylation by Cdk1 of the two residues 85/99 in human and 72/88 in chicken increases the binding of the Ndc80 10 folds (Schleiffer et al. 2012; Malvezzi et al. 2013). The structure of the interaction shows that cenp-T acts as a hook, tightly anchoring the Ndc80 complex to cenp-T (Nishino et al. 2013). The N-terminal part also interacts with the rest of the inner kinetochore according to Pull-Down experiments (Basilico et al. 2014). However, the details of these interactions are still unknown.

g) The inner kinetochore links between the centromere to the outer kinetochore

To conclude, to link the outer kinetochore to the centromere, cenp-C and cenp-T function coordinately but provide distinct pathways. Indeed, cenp-C is recruited to the kinetochore through cenp-A, while the recruitment of cenp-T is cenp-A independent (Karen E. Gascoigne et al. 2011). The depletion of cenp-C or cenp-T in cell culture shows a partial removal of kinetochore proteins that is sustained the other root protein. Leading to severe defects for the cell viability in both cases. The depletion of cenp-C and cenp-T completely disrupts the inner and outer kinetochores, demonstrating that they are the roots (Karen E. Gascoigne et al. 2011).

4) The outer kinetochore

The inner kinetochore acts as a hook anchored to the centromere and recruiting the outer kinetochore. Thus, the outer kinetochore has to be strong enough to sustain the force applied by the microtubules.

The outer kinetochore has a core composed of three sub-complexes, the Kinetochore NulL 1 (knl1) complex, the MISsegregation 12 (mis12) complex and the Nuclear Division Cycle 80 (ndc80) complex. This core, also called Knl1-Mis12-Ndc80 (KMN) complex is recruited to the kinetochore during mitosis (Santaguida and Musacchio 2009). The composition of these sub-complexes is described in the Table 1.

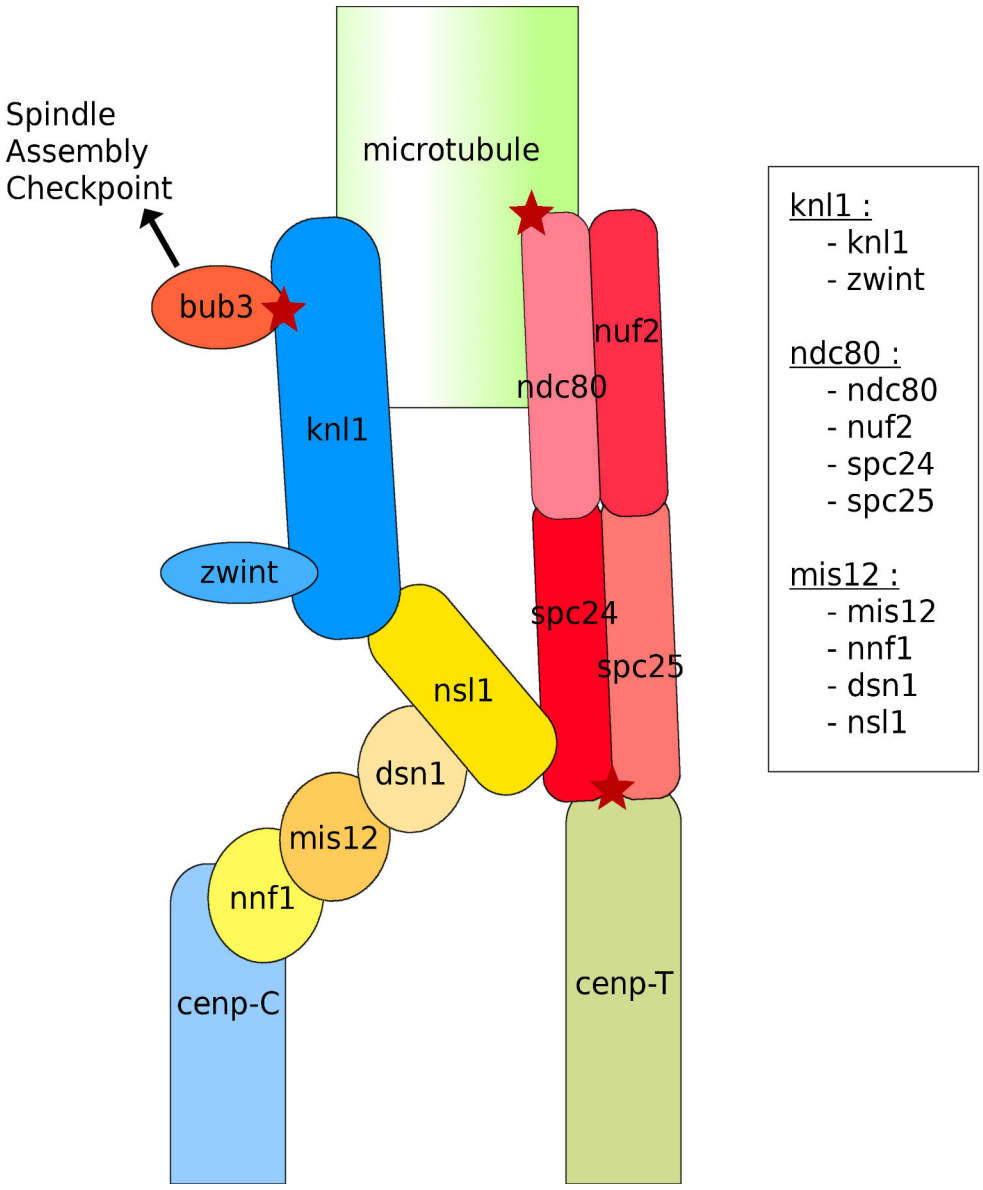


Figure 13: Representative scheme of the outer kinetochore protein interactions.

Stars highlight phospho-dependant interactions. Bold proteins give their names to their respective complex.

Complex name	Full name	Proteins in this complex
ndc80	Nuclear Division Cycle 80	ndc80, nuf2, spc24, spc25
mis12	MISsegregation 12	mis12, dsn1, nnf1, nsl1
knl1	Kinetochore NulL 1	knl1, zwint

Table 1: The KMN complex is composed of three sub-complexes.

The KMN complex forms an inter-connected complex (Figure 13) (Petrovic et al. 2010; Santaguida and Musacchio 2009) and its main functions are :

1) Binding to the inner kinetochore
2) Binding to the microtubules and their associated motor proteins
3) Sensing of the microtubule tension and of the surrounding environment
4) Regulation of these attachments through post-translational modifications
5) Fast blocking or activation of the mitosis through different checkpoints

a) The binding to the inner kinetochore

To ensure the proper chromosome segregation, the outer kinetochore has to be properly anchored to the chromosome. This attachment occurs through the cenp-T and cenp-C proteins. These interactions occur only during the mitosis. The ndc80 complex is localized in the cytoplasm during the interphase and is transferred into the kinetochore at the Nuclear BreakDown (NBD) (Santaguida and Musacchio 2009). Cenp-T binds the spc24 and spc25 proteins (spc24-25) belonging to the Ndc80 complex (Figure 13). The structure and the mechanisms of the cenp-T/spc24/spc25 binding is known (Nishino et al. 2013). The phosphorylation of cenp-T by the cdk1 kinase greatly increases the affinity cenp-T/spc24-25 (Malvezzi et al. 2013).

The N-terminal tail of cenp-C that is highly conserved among mammals binds nnf1 of the Mis12 complex (Screpanti et al. 2011; Milks, Moree, and Straight 2009; Goshima et al. 2003; Kline et al. 2006) (Figure 13). The mis12 complex is localized in the nucleus during interphase and is recruited to the kinetochore during mitosis (Goshima et al. 2003). The ndc80, mis12 and knl1 complexes interact with each other (Figure 13). These cross-interactions increase the stiffness of the KMN complex and its affinity for microtubules (Kudalkar et al. 2015).

The ndc80 complex can be recruited by cenp-T or the mis12 complex. The mis12 complex can be recruited by cenp-C or the ndc80 complex. But the binding of the outer kinetochore to both cenp-C and cenp-T is mandatory to perform a proper chromosome segregation.

b) Interactions with the microtubules

The outer kinetochore is also in charge of the binding to the microtubules and their associated proteins such as motor proteins. These interactions occur through the knl1 complex and the ndc80 complex (Iain M. Cheeseman et al. 2006).

The ndc80 complex is composed of two heterodimers: ndc80/nuf2 dimer and spc24/spc25 that interacts with cenp-T and nsl1 and the. ndc80 contains an unstructured 80-amino acid N-terminal tail that is mandatory for the microtubule binding (Wei, Al-Bassam, and Harrison 2007). Nine phosphorylation sites were identified on this tail and have be shown to tune the affinity to microtubules (Zaytsev et al. 2015). The knl1 complex is composed of knl1 and its constitutive partner zwint that recruits the RZZ complex. The RZZ complex recruits dynein and is mandatory for the proper recruitment of mad2 and mad1 (Karess 2005).

The sensing of tension provided by microtubules is important for the recruitment of the SAC and the regulation of the microtubules attachment. The underlying mechanism is still debated. Nevertheless, it seems the tension leads to stretch proteins allowing the access to motifs previously hidden and thus the recruitment of further protein partners.

5) Regulation of kinetochore function through post translation modification

The kinetochore is regulated by post translational modifications to ensure the fast and efficient regulation of its architecture or function. Some modifications by phosphorylations are identified and described (Funabiki and Wynne 2013; Welburn et al. 2010). For example, the recruitment of bub3, the binding of microtubule by ndc80 and the recruitment of the ndc80 complex by cenp-T are regulated by phosphorylation processes.

III Post-translational modifications of the kinetochore : the SUMOylation

Among the kinetochore post-translational modifications of the kinetochore, a couple of kinetochore proteins were identified detected by mass spectrometry analysis as SUMOylated (Denison 2005; Hannich et al. 2005; Panse et al. 2004; Wohlschlegel 2004). This modification could be affect protein and cell functions by various way. To reveal the impact of the SUMOylation on the kinetochore function, a study disbalance the SUMO pathway to understand the role of the SUMO pathway into mitosis and kinetochore (D. Mukhopadhyay, Arnaoutov, and Dasso 2010).

1) The SUMO pathway

The Small Ubiquitin-like Modifier (SUMO) is an ubiquitin-like protein conserved in all eukaryotes, initially identified in yeast by a genetic screening (Meluh and Koshland 1997). There are 4 different SUMO, SUMO1, 2, 3 and 4 in human but only one in *Saccharomyces cerevisiae* (Table 2). However, SUMO4 is expressed few tissues such as kidney and immune cells (Bohren et al. 2004) and its function remain elusive, I will therefore focus on SUMO1, 2 and 3. SUMO2 and SUMO3 differ by one residue and are considered similar in terms of enzymatic activities and functions, explaining the SUMO2/3 denomination. This leaves two sorts of SUMO in human, SUMO1 and SUMO2/3 that can provide different properties to the protein substrate.

Protein	Species	Sequence identity with yeast	Sequence idendity with SUMO1
smt3	*Saccharomyces cerevisiae*	-	51%
SUMO1	*Homo sapiens*	51%	-
SUMO2	*Homo sapiens*	44%	47%
SUMO3	*Homo sapiens*	45%	48%

Table 2: Sequence identity between human SUMOs and yeast smt3

SUMO proteins have been shown to function in different ways (Figure 14). They can recruit new partners (Balakirev et al. 2015), disrupt protein complex (Ryu et al. 2014), alter enzyme activities (Fernandez-Miranda et al. 2010), target the substrate to specific pathway such as the proteasome (D. Mukhopadhyay, Arnaoutov, and Dasso 2010) or be used as a solubility tag by the cells (Pinto et al. 2012).

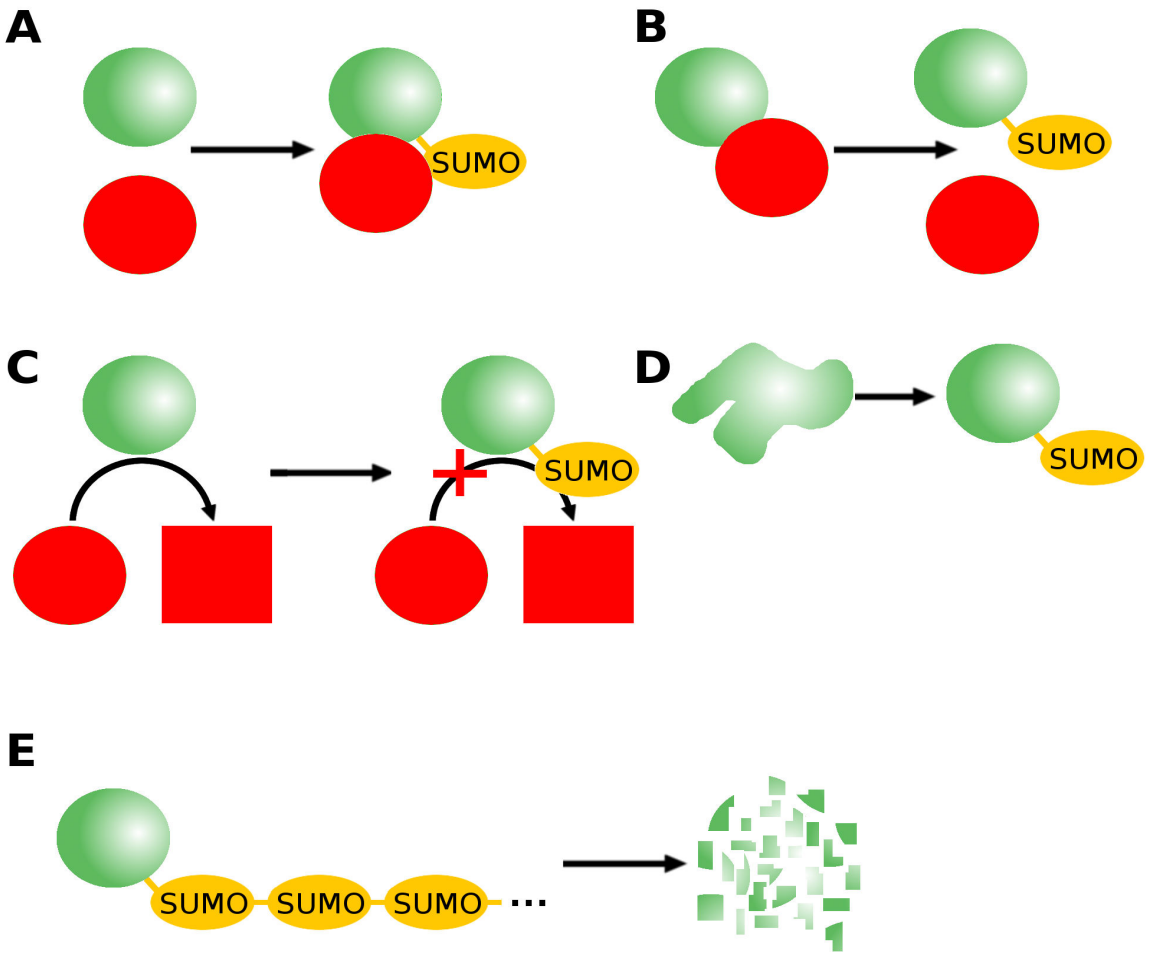


Figure 14: The pleiotropic roles of SUMOylation on protein substrates.
A) Recruitment of new partners. B) Disruption of protein complexes. C) Inhibition of an enzymatic activity. D) Solubility tag. E) The poly-SUMOlyation leads to protein degradation via the proteasome.

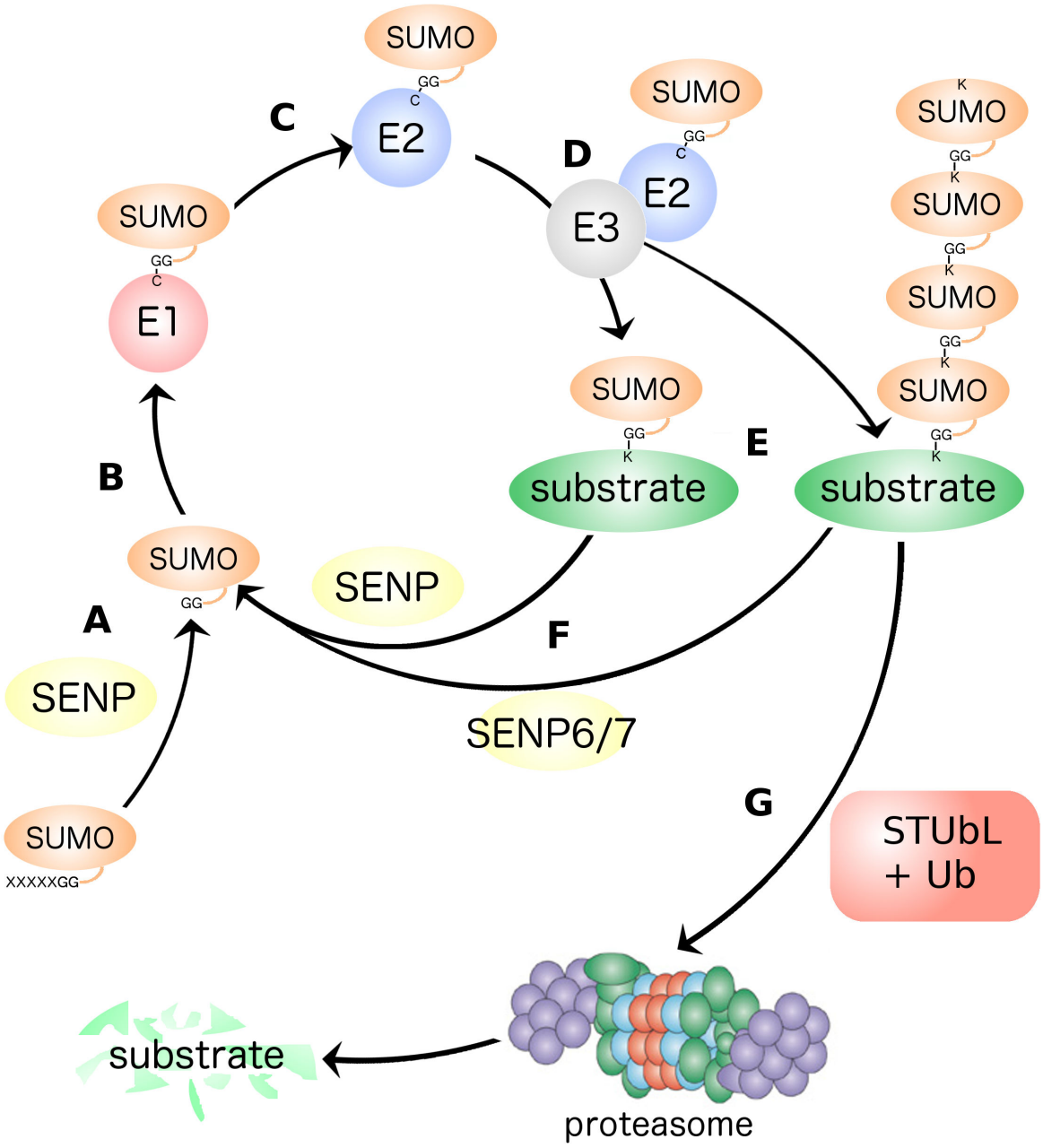


Figure 15: The SUMOylation pathway.
A) The C-terminal part of SUMO is cleaved by SENP. B) Activated SUMO is recruited by E1. C) SUMO is transferred to E2 and E) then linked to substrate, forming a mono-SUMOylation. D) E3 is involved in the process specificity. F) SUMO can be the substrate forming a poly-SUMOylation chain. G) SUMO can be cleaved by SENP. H) Poly-SUMO chain leads to poly-ubiquitination and targeting to the proteasome.

The SUMO pathway is mechanistically similar to ubiquitin pathway (Figure 15). SUMO is first cleaved by isopeptidases, the SENtrin specific Protease (SENP) that generate a functional

SUMO with a C terminal di-glycine motif used for the linkage. SUMO is then bound to the cysteine of a first enzyme, E1 (Uba2/Aos1 in humans) in an ATP dependent mechanism. Then, E1 transfer its linked SUMO to a second enzyme, E2 (Ubc9 in humans). SUMO is tethered to the cysteine 93 of E2 in human (Knipscheer et al. 2008) and bound a lysine of the substrate. The link between SUMO and the substrate is a pseudo peptide bond. Indeed, the carboxyl group of the last glycine of SUMO is linked to the amino group of the side chain of the lysine. There are one E1, one E2 and few E3 in all organisms studied. The SUMOylation preferentially occurs on the consensus site [FILMV]Kx[DE] (Figure 16) (Impens et al. 2014; Rodriguez 2001; Sampson, Wang, and Matunis 2001).

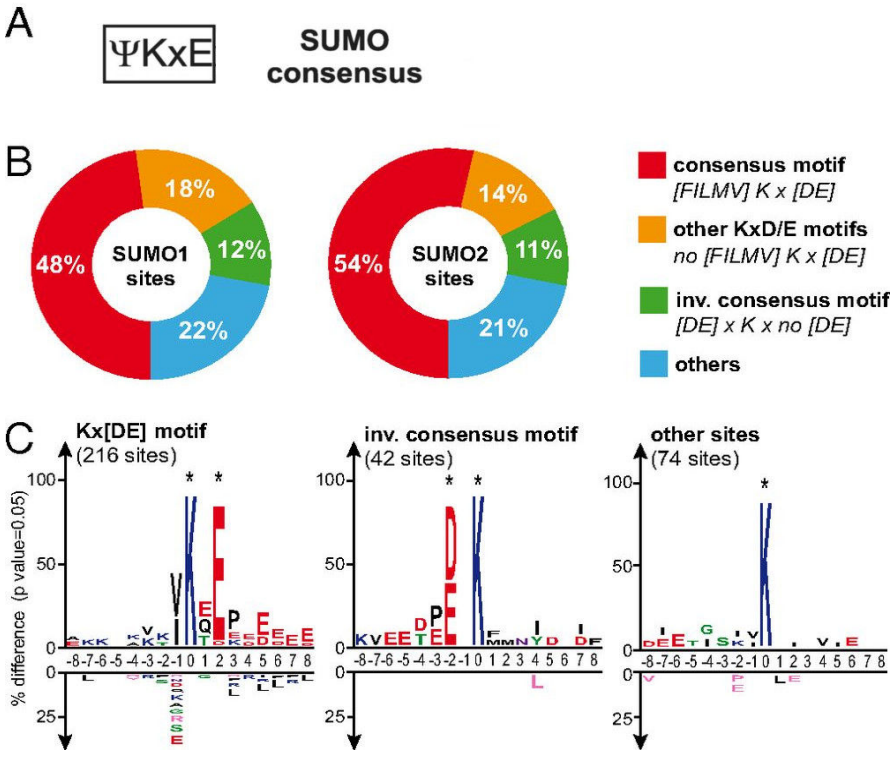


Figure 16: The SUMOylation sites identified in human cells.

A) The consensus site of the SUMOylation. B) Ratio of the SUMOylated sites identified in the paper. C) IceLogo representations showing the amino acids surrounding the SUMO-conjugated lysine residue in different motifs. Imposed amino acids for each type of motif are marked by an asterisk. The frequency of nonimposed amino acids at every subsite is compared with sampled frequencies in the human proteins stored in the UniProt/Swiss-Prot database (negative control). Residues that never were observed at specific positions are shown in pink.
From (Impens et al. 2014)

A competition occurs among all SUMOylable proteins. However, a couple of proteins, called E3 SUMO ligase increase the specificity of the substrate. It is important to note that SUMO2/3 itself can be substrate, leading to poly-SUMO chain (Tatham et al. 2001). SUMO1 generates a mono-SUMOylation or blocks the growth of poly-SUMO chain, demonstrating a highly dynamic process. The SUMO tag can be cleaved by specific isopeptidases such as SENP (Li and Hochstrasser 1999; Li and Hochstrasser 2000).

2) Removal of SUMO

Saccharomyces cerevisiae could be considered as an interesting model to understand the cleavage of SUMO. In this species, there are only two main SUMO-specific isopeptidases, the Ubiquitin-Like specific cysteine Proteases (ULP1 and ULP2) (Li and Hochstrasser 1999; Li and Hochstrasser 2000). Ulp1 is specific to the processing of Smt3 and the removal of mono-SUMOylation, while Ulp2 is specific to poly-SUMO chain (Bylebyl, Belichenko, and Johnson 2003; Li and Hochstrasser 1999). The molecular mechanism of SUMO cleavage by Ulp1 has been described (Mossessova and Lima 2000).

Humans have six homologous proteins of ULPs divided in two families, those related to Ulp1 (SENP1, 2, 3 and 5) or to Ulp2 (SENP6 and 7) (Debaditya Mukhopadhyay and Dasso 2007; Reverter and Lima 2006; Lima and Reverter 2008). Biochemical assays have demonstrated that only SENP2 is efficient for SUMO processing, not SENP6 or SENP7.

Several isopeptidase are specific to SUMO but not related to SENP in term of sequence (Shin et al. 2012) and their depletion has no major effects on cell viability. The DeSUMOylating Isopeptidase (DeSI) depolymerizes the poly-SUMO chain and seems more specific than SENP proteins (Shin et al. 2012).

3) The SUMO interacting motif

The SUMO proteins are able to bind their target protein via an hydrophobic core of 4 amino acids called Sumo Interacting Motif (SIM) (consensus sequence: V/-V/I-x-V/I) (Minty et al. 2000; Bohren et al. 2004). This process is a key mechanism of SUMO-dependent protein

recruitment and of the SUMO pathway such as the recognition of poly-SUMO chain by E3 ubiquitin ligases SUMO dependent (Kung et al. 2014). Although the SIM is the canonical motif of SUMO binding, other motifs allow the binding of SUMO (Pilla et al. 2012). For this reason, the *in silico* analysis cannot identifiy all SUMO-dependant protein interaction.

IV Goals

An initial publication from the team focused on the roles of the poly-SUMOylation in mitosis, precisely on chromosome segregation (D. Mukhopadhyay, Arnaoutov, and Dasso 2010). In order to reveal the poly-SUMOylation effects, SENP6 was depleted using siRNA in HeLa cells. This knockdown led to defects such as chromosome misalignments, changes of the chromosome compaction during the metaphase and alteration of the kinetochore integrity. The phenotypes of SENP6 depletion in HeLa cells are similar to depletion of members of the CENP-H/I/K/M complex. Indeed, SENP6 depleted cells show a decreased amount of outer kinetochore proteins and a removal of CENP-H/I/K/M and CENP-R/O/P/Q/U complexes from the inner kinetochore. Importantly, all proteins of these complexes are mislocalized except CENP-I that is degraded during metaphase after SENP6 depletion. Thus our team identified the mechanisms of degradation CENP-I.

The length of the poly-SUMO chain is a result of the equilibrium between the addition and the removal of SUMO2/3 peptides. After SENP6 depletion, the removal process is disrupted, leading to an increased poly-SUMOylation state of CENP-I. The long poly-SUMO chains recruit rnf4 that is a protein previously identified as a SUMO-dependent E3 ubiquitin ligase leading to the poly-ubiquitination of CENP-I (Häkli et al. 2004; Hakli et al. 2005). The affinity of rnf4 for poly-SUMO chain greatly increases when the SUMO poly-chain consists of at least 4 SUMO molecules. Finally, the poly-ubiquitinated CENP-I is degraded by the proteasome.

The CENP-H/I/K/M is central and is known as a scaffold complex but the roles of the CENP-I poly-SUMOylation during mitosis remain unknown. I wanted to determine how does SUMOylation regulate the presence of the CENP-H/I/K/M complex within the kinetochore, thus leading to its cellular maintenance or degradation. Indeed, the SUMO-dependent CENP-I degradation could control the loading efficiency of proteins on the kinetochore or be a time-checkpoint for kinetochore disruption that induces lagging in chromosome segregation. I thus

hypothesize that the CENP-H/I/K/M complex loading could be affected by SUMO dependent interactions.

To reveal the roles and effects of CENP-I poly-SUMOylation, I identified CENP-I SUMOylated lysines and generated mutated residues. Due to technical issues, their identification could not be performed using immunoprecipitation coupled to mass spectrometry analysis.

To further study the CENP-I-dependent regulation of kinetochores, I decided to perform *in vitro* SUMOylation assays of CENP-I. CENP-I is involved in the Constitutive Centromere Associated Network (CCAN). I tested the CCAN purification from overexpression in bacterial system using Ion Exchange Chromatography and Size Exclusion Chromatography.
Nevertheless, the CENP-H/I/K complex is instable and not suitable for biochemical assays. Following the same protocol, the purification CENP-I from overexpression in bacterial system was performed. Unfortunately, CENP-I is degraded or remains insoluble. At this time, the Pr. Musacchio's team identified a fourth partner of the CENP-H/I/K complex, determined a protocol for its purification and provided us several micrograms of the complex. After production, purification and characterization of the SUMO pathway proteins, a SUMOylation assay for the CENP-H/I/K/M complex was set up. The results were analyzed by Mass Spectrometry analysis.

In parallel, in order to study the kinetochore SUMOylation, I reconstituted *in vitro* kinetochore according to a published protocol (D. Mukhopadhyay, Arnaoutov, and Dasso 2010). However, it appeared that these kinetochores complexes are highly instable and non suitable for biochemical assays. To circumvent these issues, I adapted the method used by Dr. Fukagawa and Dr. Cheeseman. I fused the kinetochore protein CENP-T to the lactose repressor, targeted this protein to beads coupled to lactose operon DNA templates and determined if this complex can recruit kinetochore proteins as CENP-T does *in vivo.*

Material and Methods

I The molecular biology

1) Buffers and reagents

- Tris EDTA buffer (TE buffer): 10mM Tris pH7.4, 1mM EDTA.

- Tris Acetate EDTA (TAE buffer): 40mM Tris pH7.4, 20mM Acetate, 1mM EDTA.

- Lysogeny Broth (LB media): Trypton 10g/L, yeast extract 5g/L, NaCl 5g/L.

For plasmid production, *Escherichia coli* DH5 alpha bacterial system from New England Biolabs^{tm} (NEB) was used. The phenotype of this strain is *E. coli fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*. To produce plasmid DNA, the transformed cells grew overnight (o/n) in LB medium at 37˚C under a 180 rpm shaking. These bacteria were then used to perform mini or midi prep following the instructions provided by the QIAgen^{tm} company.

2) Bacterial transformations

The protocol used for bacterial transformation with the plasmid(s) is the same for all the bacterial strains. Commercial competent cells are used from aliquots stored at -80˚C. For the bacterial transformation with a single plasmid, an aliquot of 100µL of cells is thawed on ice for 10min. Then, 100ng of plasmid is added to the competent bacteria and gently mixed by pipetting. The mix is left on ice for 30min, incubated at 42˚C for 30sec and incubated on ice for 2min. Then, 300µL of pre-warmed LB (37˚C) is added to the mix and left at 37˚C for 30min. Half of the mix is spread on LB-agar plates with the appropriate antibiotics.

The remaining mix is inoculated to LB for liquid culture with appropriate antibiotic. For plasmid preparation, 10mL or 1L of LB with antibiotic is used. For protein production, 50mL of LB with antibiotic is used. The transformed cells grow o/n in LB medium at 37*˚C* under a 180rpm shaking. Then I use these bacteria to perform mini or midi prep by following the instructions provided by the QIAgen^{tm} company.

3) Primers design

The primers used during my Thesis were designed as follows :

- The length of primer confirms this equation: 4^n > number of base pairs of the plasmid/genome, where n is the length of the primer.
- The very last base in 3' is an A/T and the 2 bases before are G/C.
- The theoretical hybridization temperature of primers is between 55 and 60°C *(http://www6.appliedbiosystems.com/support/techtools/calc/).*
- The A/T percentage is between 25 and 75%.
- High complexity or highly repeated sequences were avoided.

4) Polymerization Chain Reaction

The Polymerization Chain Reaction (PCR) is a method used to amplify specific DNA sequences. Since its optimization using the *Thermus aquaticus* (*Taq*) DNA polymerase (Saiki et al. 1988), this method is considered as a very basic method of molecular biology. I used two polymerases during my doctoral work:

- **Taq DNA polymerase (NEB):** This DNA polymerase is considered as the standard for PCR. It is a very robust enzyme with a high rate of polymerization. However, it has a relatively high error rate (1.10^{-4}), leading to mutations in the amplified DNA sequence. Thus, this enzyme was used when the reaction does not required a high fidelity amplification.
- **Herculase II Fusion DNA polymerase (Agilent Genomics):** This DNA polymerase combines the high polymerization rate of Taq DNA polymerase (15s/kb) with the error rate of *Pyrococcus furiosus* (*Pfu*) DNA polymerase (1.10^{-6}). This enzyme was used to amplify DNA sequences requiring the highest amplification fidelity.

The reaction mix is:

- 250µM of each dNTP
- 0.25µM of each primer
- 100ng of plasmid
- DMSO: 0% usually or 4% if mentioned
- Buffer provided by the company
- 0.5µL of Herculase II Fusion DNA polymerase

NB: The polymerase is not added into the negative control.

The standard protocol is:

1) 95˚C: 2 or 5min for plasmid or bacteria respectively.
2) 95˚C: 20sec
3) 53˚C: 45sec
4) 68˚C: 30sec/kb
5) The steps 2), 3) and 4) are repeated 30 times.
6) 68˚C: 10min
7) 4˚C holds forever.

To confirm the amplification of the gene or the plasmid, 1% agarose gels are used in TAE buffer. For visualization, a mix of 5µL of PCR products and 1.5µL of loading dye (Gel Loading Dye Purple, NEB) is loaded onto the gel and the migration is performed at 90V for 30min. The PCR products are then purified by the QIAquick PCR purification kit (Qiagen). The PCR products are suspended in pure water and the DNA concentration is determined by optical density. The final DNA products are stored at -20˚C.

5) <u>Digestion by restriction enzyme and cloning</u>

Enzymes and buffers were used as described in NEB protocols *(https://www.neb.com/tools-and-resources/usage-guidelines/double-digests)* *(https://www.neb.com/tools-and-resources/interactive-tools/double-digest-fin).*

All cloning experiments are performed by the Single-Tube Restriction-based Ultrafiltration method (STRU method) (Bellini, Fordham-Skelton, and Papiz 2011). In the STRU method, the plasmid and the insert are digested in the same tube. The small fragments resulting from the reactions are removed using centrifugal filter units containing a membrane with a suitable cut off. Once the reaction is cleaned of all DNA fragments, the mix is transferred into a tube for ligation. The cloning method has several advantages. Less enzymes are necessary for the restriction digestion, the step of purification by agarose gel of insert and plasmid after digestion is skipped, reducing the number of steps and the required time. In my hand, this method is more efficient than the standard protocol. This method is “fast, inexpensive and efficient”.

If the insert is a PCR product, 500ng of template and five fold more amounts of times insert were added in a final volume of 20µL. In the case of a plasmid, 200ng of template and three times more inserts are mixed in a final volume of 20µL. All the constructs generated are verified by molecular sequencing.

6) Directed mutagenesis

I generated mutants of several genes by amplification of the plasmid of interest with specific primers carrying the desired mutation (Fisher and Guo 2016). The whole plasmid is amplified using two primers, P1 and Pm, with Pm carrying the designed mutation. The PCR products obtained are blunt end linear plasmids with the mutation. The PCR results in a mix of initial plasmid and PCR products carrying the mutation. The mutated linear DNA is then ligated and digested by the enzyme Dpn1, which cuts methylated DNA. All the plasmids generated from *E. coli* are methylated but not the DNA strands synthesized by PCR. Thus, the presence of the mutation will avoid the plasmid digestion.

The reaction mix of PCR is:

- 250µM of each dNTP
- 0.25µM of each primer
- 200ng of plasmid from *Dam+* bacteria
- DMSO: 0% usually or 4% if mentioned
- Herculase II Reaction buffer (Agilent Technologies)
- 0.5µL of Herculase II Fusion DNA polymerase (Agilent Technologies)

Standard PCR protocol is as follows:

1) 95˚C: 2 min
2) 95˚C: 30 sec
3) 53˚C: 45 sec
4) 68˚C: 1min per kb + 1min
5) The steps 2), 3) and 4) are repeated 18 times
6) 68˚C: 10min

The PCR products are visualized with 1% agarose gel to confirm the amplification and purified using the *QIAquick PCR purification kit* (Qiagen). The DNA concentration of purified plasmids is determined by optical density.

To perform the circularization of the PCR products, the T4 DNA ligase (NEB) was used *(https://www.neb.com/protocols/1/01/01/dna-ligation-with-t4-dna-ligase-m0202)*. After ligation, five units of DpnI are added to the tube for incubation at 37˚C for 30min. The samples are incubated at 80˚C for 20min and then purified using the QIAquick PCR purification kit (NEB) and suspended in water for further bacterial transformation.

II Biochemistry

1) Protein overexpression in bacterial system

For protein production, *Escherichia coli* BL21 (DE3) bacterial system was used. This *E. coli* B strain is infected with DE3, a λ prophage carrying the T7 RNA polymerase gene and *lacI*q. The genotype is *E. coli* str. B F− *ompT gal dcm lon hsdSB*(*rB*−*mB*−) λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) [*malB*+]K-12(λS). *E. coli* BL21 (DE3) RIL and *E. coli* BL21 (DE3) RIPL were also used when specified in the text (Annexes). These strains carry a plasmid overexpressing rare tRNA in *E. coli*. It increases the speed of translation of heterologous proteins. The *E. coli* BL21 (DE3) RIL contains extra copies for argU, ileW and leuW tRNA. The *E. coli* BL21 (DE3) RIPL contains extra copies for argU, ileW, proL and leuW tRNA.

Once transformed with the plasmids, the bacteria are spread on LB plates. A couple of single colonies are picked from LB plates and incubated in 50ml of LB broth with the proper antibiotic at 37˚C o/n. The antibiotics and the concentration used during my work are ampicillin 0.1mg/ml, kanamycin 50µg/ml and chloramphenicol 25µg/ml. The day after, the LB complemented with antibiotic is inoculated to bacteria from o/n culture to obtain an optical density (OD) of 0.1. The bacteria culture is incubated at 37˚C until the OD reaches 0.55, leading to the cells induction by two different protocols depending of the nature of the protein of interest:

- production of proteins at 37˚C: IsoPropyl β-D-1-Thio-Galactopyranoside (IPTG) is added to a final concentration of 0.5mM and the cells are incubated at 37ºC for 3h.

- production of proteins at 19˚C: The cells are incubated at 4˚C for 10min. IPTG is then added to a final concentration of 0.5mM and the cells are incubated at 19˚C for 18h.

Finally, the bacteria are centrifuged at 5,000rcf for 30min at 4˚C and the pellet is suspended in the proper buffer and stored at -20˚C.

2) Protein extraction and purification

Buffers:

- **Extraction Buffer:** 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH7.4), 137mM NaCl, 2.7mM KCl, 1mM DTT.
- **Phosphate Buffered Saline (PBS):** 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH7.4), 137mM NaCl, 2.7mM KCl.
- **Loading Blue Solution:** 62.5mM Tris-HCl (pH 6.8), 2.5% SDS, 0.002% Bromophenol Blue, 0.7135M (5%) β-mercaptoethanol, 10% glycerol.
- **Tris Glycine SDS buffer (TGS):** 3g/L Tris, 14.4g/L glycine, 1g/L SDS.
- **SDS-PAGE destaining buffer:** 45% ethanol, 10% acetic acid.
- **SDS-PAGE Staining buffer:** 5% ethanol, 5% acetic acid, 0.02% Coomassie blue G250.

To perform protein extraction, the cells are thawed in a water bath at room temperature (RT). The samples are homogenized and kept on ice. Then, sonication is performed on ice as a cycle of 20% with an output power of 20W for 1min per 15ml of samples. The bacteria lysates are centrifuged at 13,000rcf for 40min at 4˚C.

For purification of soluble proteins, the supernatant is used, whereas for purification of insoluble proteins, the pellet is processed for a further step of solubilization.

a) Immobilized Metal Affinity Chromatography-Ni^{2+} (IMAC- Ni^{2+})

The standard protocol for IMAC-Ni^{2+} purification is:

1) pre-chill the buffer on ice

2) pre-wash the Ni^{2+}-NTA beads (#30230, Qiagen) with the extraction buffer and packed the matrix in the plastic column Econo-Pac® Chromatography Columns (#7321010, Bio-Rad).

3) add imidazole to a final concentration of 20mM to the soluble fraction of bacterial lysate

4) add the sample on the Ni^{2+}-NTA column

5) wash the beads twice with 10 volumes of column with extraction buffer complemented with 40mM imidazole to remove the non specific interactions

6) add 1 volume of elution buffer (PBS with 400mM imidazole).

7) cap the Ni^{2+}-NTA column and resuspend the beads.

8) wait for 5min

9) uncap the column

10) collect the elution fraction.

11) repeat the steps 6) to 10), twice

12) analyze the samples by SDS-PAGE

b) Ion Exchange Chromatography

The standard protocol for IEC (anion or cation) is:

Buffer A: low ion force buffer (100mM or 150mM NaCl)

Buffer B: low ion force buffer (1M and 2M NaCl)

1) equilibrate the column with buffer A.

2) resuspend the sample in buffer A and load onto the column

3) wash the column with 5 volumes of buffer A.

4) perform a gradient of buffer B from 0% to 100% on the column

5) check the OD of the fraction at 280 nm and analyze by SDS PAGE gel.

Once the salt concentration required for protein elution is known, the gradient of elution can be adjusted for a higher purification.

c) Size Exclusion Chromatography

Two main types of columns were used:

- a preparative column, the Superdex 200 16/600 (GE Healthcare) with a constant flow of 1 ml/min.
- a semi-preparative column, the Superdex 200 10/200 (GE Healthcare) with a constant flow of 0.5 ml/min.

3) Protein revelation

a) Samples preparation

The samples are mixed with loading buffer (Laemmli buffer): 2% SDS, 62.5mM Tris/HCl (pH6.8), 10% Glycerol, 0.01% bromophenol blue; 5% Beta-2-mercaptoethanol and then boiled at 95C for 10 min. They are loaded on Novex^{tm} 4-20% Tris Glycine gels, 15 wells (Invitrogen). Commercial "Kaleidoscope^{tm} Prestained Standards" (Bio-Rad) were used as markers of molecular weight.

For most of the experiments, the proteins are stained by Coomassie blue. After their migration, the gel is unmolded and incubated in SDS-PAGE destaining buffer for 5min under shaking. The gel is stained with SDS-PAGE staining buffer for 5min under shaking and the buffer is replaced for another incubation. The gel is heated for 20 sec by microwave at maximum power and incubated for 10 min. Then, the gel is washed three times with distilled water and incubated in SDS-PAGE destaining buffer diluted twice for 30 min under shaking. If the protein of interest amount is under 0.1µg per band, a silver staining (#24600, Thermo Fischer Scientific) or an immunoblot (see the following part) is performed.

b) Analysis by Western Blot

Buffers:

Transfer buffer: 25mM Tris, 192mM glycine and 0.1% SDS (TGS buffer) complemented with 10% methanol.

PBST: PBS complemented with 0.1% Tween-20 (PBST).

Blocking solution: 5% milk powder diluted in PBS complemented with 0.1% Tween-20 (PBST).

After gel migration, the transfer cassette is washed with transfer buffer and placed into a Hoefer^{tm} TE77XP semi-dry transfer unit (#TE70XP, Hoefer). A membrane of PolyVinyliDene Fluoride (PVDF) with a porosity of 0.45 µm was used. The transfer is performed using a constant current of 45 mA per gel for 80 min at RT. The constant current, applied laterally, leads to the migration of proteins that are charged negatively due to the SDS from the gel to the membrane. The fixation of proteins is due to hydrophobic and ionic interactions between the membrane and the proteins.

To minimize the non-specific interactions of antibodies, the membrane is incubated for 1h at RT in the blocking solution. The membrane is washed twice for 5min in PBST and then incubated with primary antibodies (Table 3). The primary antibodies are diluted in PBST and their dilution varies according to their specificity and affinity (the average concentration is 1 µg/ml). The membrane is incubated overnight at 4˚C with a gentle shaking.

Target	Target species	Host	Generation	Molecular Weight expected (kDa)
AuroraB	xenope	rabbit	Custom	40
Bub1	xenope	chicken	Custom	128
BubR1	xenope	rabbit	Custom	118
Cdc20	xenope	rabbit	Custom	56
cenp-A	human	rabbit	Commercial (USbiological c2615)	16
cenp-C	human	goat	Custom	105
cenp-H	human	mouse	Commercial (Bethyl BL1112)	29
cenp-I	human	rabbit	Commercial (Bethyl A303-374A)	61
cenp-K	human	rabbit	Commercial (Abcam Ab56370)	32
cenp-M	human	rabbit	Custom	20
	xenope	rabbit	Custom	20
cenp-O	human	rabbit	Custom	34
	xenope	rabbit	Custom	37
cenp-P	xenope	rabbit	Custom	33
	human	rabbit	Custom	31
cenp-S	human	rabbit	Custom	16
CREST	human	rabbit	Commercial	16 / 65 / 105

(See the next page)

Target	Target species	Host	Generation	Molecular Weight expected (kDa)
GFP	*Aequorae victoria*	mouse	Commercial (Abcam Ab6556)	27
HA peptide	*Influenza*	rat	Commercial (Roche 11867423001)	
LacI	*Escherichia coli*	mouse	Commercial (Abcam Ab33832)	35
Mad1	xenope	rabbit	Custom	83
Mad2	xenope	rabbit	Custom	24
Mis12	human	rabbit	Commercial (A300-775A)	26
Ndc80	xenope	rabbit	Custom	74
PIASy	human	rabbit	Custom	63
Ska1	xenope	rabbit	Custom	29
Spc24	xenope	rabbit	Custom	23
SUMO2	human	Rabbit	Commercial (Abcam AB3742)	14
Topo II	xenope	rabbit	Custom	180

Table 3: List of primary antibodies used during my work.

The day after, the membrane is then washed three times in PBST for 5min at RT and incubated with 0.1µg/ml of secondary antibodies in PBST for 2h at 4°C under gentle shaking (Table 3).

The proteins were revealed using the Enhanced ChemiLuminescence that is a method based on the coupling of the HorseRadish Peroxydase (HRP) to the secondary antibodies.

The membrane is washed three times with PBST at RT. The proteins were revealed using the commercial kit SuperSignal West Pico Substrate Substrate Supersignal™ West Pico Chemiluminescent, #PI34080, Fischer Scientific) and visualized using the imager FluorChem R (FluorChemR, ProteinSimple).

Target proteins	Host	Origin	Type of detection
Mouse IgG	sheep	Commercial (GE Healthcare NA931V)	ECL
Rat IgG	goat	Commercial (Sigma A9037)	ECL
Rabbit IgG	donkey	Commercial (GE Healthcare NA934V)	ECL

Table 4: List of the secondary antibodies used during my work.

4) Reconstitution *in* vitro of the SUMOylation pathway

a) Production and purification of the SUMOylation pathway proteins

I produced and purified all the proteins of the SUMO pathway that are used for the *in vitro* SUMOylation assays. Ubc9, SUMO2 and its variants were cloned with a 6-His tag provided by the pET28a+ backbone and were produced in *Escherichia coli* BL21 (DE3) codon+ RIPL following the standard protocol with an induction temperature of 37˚C.

The protein PIASy with a 6-His tag carried by a pET28a+ backbone is produced in *E. coli* BL21 (DE3) codon+ RIL following the standard protocol with an induction temperature of 19˚C. These proteins are purified by an IMAC-Ni^{2+} followed by a SEC. The activity of the purified proteins is confirmed by SUMOylation assays. The pure and active proteins are frozen in liquid nitrogen and stored at -80˚C for further experiments.

Buffers:

- **SUMO-column Binding Buffer:** 100mM NaCl, 20mM HEPES (pH7.8), 5mM MgCl2, 1mM ATP, 1unit/ml of inorganic pyrophosphate (iPP), 0.1mM DTT and 5% glycerol.
- **Elution Buffer:** 100mM NaCl, 1mM MgCl2, 20mM Tris (pH8.8), 10mM DTT, 5% glycerol.
- **Size Exclusion Chromatography buffer:** 20mM Tris (pH 8), 200mM NaCl, 1mM DTT
- **SUMOylation reaction buffer:** 40mM Tris-HCl (pH7.5), 100mM NaCl, 2mM DiThioThreitol (DTT), 4mM MgCl2, 2mM ATP, 0.05% Tween-20, 5% glycerol.

This method was developped by Dr. Azuma. Basically, this protocol is based on two-step purification. First, E1 is purified by IMAC-Ni^{2+} to remove endogenous bacterial proteins according to the standard protocol. However the buffer used are different, there is no imidazole in the lysis/wash buffer (20mM Tris pH8, 100mM NaCl). The elution is performed with an elution buffer of 20mM Tris pH 8, 100mM NaCl and 30mM imidazole. The samples are analyzed by SDS-PAGE stained by Coomassie blue. Once the E1 complex (Uba2/Aos1) is purified, the active E1 protein is purified with a SUMO-affinity column. To prepare the SUMO-affinity column, pure SUMO2 is covalently linked to N-HydroxySuccinimide (NHS) beads following the company instructions (#17-0906-01, GE Healthcare).

The NHS-SUMO column is equilibrated with SUMO-column binding buffer. 1mM ATP is added to the E1 purified by IMAC-Ni^{2+}. This sample is applied to SUMO-NHS column by using gravity flow (without needle). The column is washed with the binding buffer for 5 volumes of column. The column is washed with the elution buffer without DTT. E1 is eluted with 5 volumes of column with Elution Buffer.

b) *In vitro* SUMOylation assays

This is the standard protocol that is optimized based on the substrate. In a final volume of 60µL of buffer (40mM Tris (pH7.5), 100mM NaCl, 2mM DTT, 4mM MgCl2, 2mM ATP, 0.05% Tween-20, 5% glycerol):

150nM E1

200nM E2

50nM PIASy

10µM of SUMO2

The reaction is incubated at 30˚C. To monitor the experiment, 15µL of each reaction is mixed with loading blue and heat at 95˚C for 3min, processed by SDS-PAGE and revealed by silver nitrate staining or Coomassie blue.

5) Production of kinetochore on beads

a) cenp-A nucleosome beads

The generation of cenp-A nucleosomes *in vitro* was performed using the protocol described by Dr. Guse (A. Guse, Fuller, and Straight 2012).

b) Production of DNA arrays

The DNA array is a 19-repeat of the 601 sequence that is the DNA sequence with the highest affinity for the cenp-A nucleosome (Lowary and Widom 1998). To briefly summarize the protocol described in this paper, the DNA arrays are carried by a pUC18 backbone (from D. Rhodes) and are produced in *Escherichia coli* DH5 α (NucleoBond Xtra Midi EF, Macherey nagel). They are extracted by restriction enzymes digestion, purified by PolyEthylene Glycol precipitation and finally washed by ethanol precipitation. The pure DNA arrays are biotinylated.

c) Purification of proteins

The human histones (2A, 2B, 3, 4) and cenp-A are produced in *Escherchia coli* (BL21) DE3 codon+ RIPL following the standard protocol with an induction temperature of 37°C. Each protein is re-solubilized after lysis and purified by IEC.

d) Nucleosome beads

The pure proteins are mixed with pure DNA arrays to form two types of nucleosomes: the canonical one including histones 2A, 2B, 3 and histone 4, and the centromeric nucleosome consisting of cenp-A and the histones 2A, 2B and 4. The protocols were described in (Guse, Fuller, and Straight 2012). The nucleosomes are then bound to Dynabeads M-280 Streptavidin magnetic beads (#11205D, Life Technologies) following the company's recommendations. The nucleosome beads are then used for further experiments. All along this part, I will speak about streptavidin, but each time it will be the biotin binding core of the streptavidin.

6) Production of LacI/LacO beads

Buffers

- Phosphate Buffered Solution (PBS): 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH7.4), 137mM NaCl, 2.7mM KCl.

- Tris EDTA buffer (TE buffer): 10mM Tris (pH7.4), 1mM EDTA.

a) Plasmid digestion

To generate the LacO arrays, the plasmid pSV2 containing 256 repeats of LacO sequence (#33143, Addgene) was used. The plasmids are produced in bacterial systems and purified by the midi prep kit NucleoBond XtraMidi EF (#740410, Macherey-Nagel). In a volume of 100µL of 10X NEB Buffer 3 and 10µL of BSA 100X (NEB), a quantity of 1mg of pSV2 is added. Distilled water is added to a volume of 960µL. The reaction is homogenized by pipetting. A volume of 10µL of each BamHI, SalI, DraI and EcoRV-HF is added and homogenized by pipetting. The reaction is incubated at 37˚C ideally o/n.

The running of 1% agarose gels confirms the plasmid digestion. The LacO repeats are precipitated by a PolyEthylene Glycol gradient (PEG 3350 from 4% to 7%) and purified as described in the paper from Dr. Guse (Guse, Fuller, and Straight 2012). The efficiency of precipitation is monitored by the samples loading on a 1% agarose gel.

b) Purification of DNA template

The 256 repeats of LacO arrays are purified by Poly-Ethylene Glycol (PEG) 3350 precipitation. Increasing the PEG concentration in presence of 0.5M NaCl leads to pellet DNA fragments according to their size. The protocol is similar than published by Dr. Guse except the use of PEG 3350. The repeats of LacO arrays precipitates around 4% of PEG 3350. The precipitate of LacO arrays is resuspended in water and then dialyzes o/n with TE buffer at 4˚C to remove any traces of PEG 3350. An ethanol precipitation is performed to pellet the LacO array, avoiding the presence of PEG to interfere with the biotinylation reaction, and allows its resuspension in the desired volume and buffer.

c) Tagging the LacO arrays

The standard protocol is performed as follows by mixing:

- 450µg of pure LacO repeats
- 32.5µl of Standard Taq Mg^{2+} Free Buffer (NEB)
- 17.5µL of a 50µM mix of dGTP/dCTP/dTTP
- 87.5µL of 0.4mM biotin-14-dATP (#19524-016, Invitrogen)
- 17.5µL of 0.1M $MgCl_2$
- 17.5µL of Taq DNA polymerase (#M0273S, NEB)

This mix is completed with distilled water to reach a final volume of 325µL. The reaction mix is incubated for 1h at 65˚C and then washed by ethanol precipitation as described (Guse, Fuller, and Straight 2012). The pure LacO repeats are dialyzed o/n in TE buffer at 4˚C. The pure and biotinylated LacO arrays are further targeted to Dynabeads M-280 Streptavidin magnetic beads (#11205D, Life Technologies) following the company's recommendations.

d) Recruitment of LacO arrays on streptavidin beads

The biotinylated LacO arrays bind to streptavidin that is coated on magnetic beads. A volume of 80µL of magnetic beads is washed three times in 800µL of PBS + EDTA 0.5mM at RT. Using magnetic racks, 40µL of beads (at 50% slurry) is pelleted and then incubated for 20min at RT in a 800µL solution of biotinylated LacO arrays at 200ng/µL. The beads are pelleted and washed three times with 800µL of PBS + EDTA 0.5mM. It is to note that the quantity of DNA loaded on beads cannot be quantified by spectrophotometer (results are aberrant) or by spectral properties (due to the presence of the iron atoms of the beads). The most reliable method to determine the linking of LacO arrays to streptavidin magnetic beads is to measure the ability of LacO arrays beads to recruit LacI.

e) Purification of LacI

Two constructs of LacI were designed, one for the production of LacI alone and the other one for the fusion with cenp-T. The LacI gene is inserted into the templates pET28 and pET23 between the restriction sites Eag1 and Hind3 (see Supplementary part).

The LacI production is performed using the *Escherichia coli* BL21 (DE3) RIL strain with the standard production protocol. The LacI purification is performed using the standard lysis protocol, followed by an ion exchange chromatography with a Q-sepharose Hi-Trap (#17-1154-01, GE Healthcare) and then a further treatment by size exclusion chromatography with a Superdex200 16/60 column (#28-9893-35, GE Healthcare). Each step of the purification process is analyzed by SDS-PAGE. The fractions containing the pure proteins are pooled and are concentrated at 10mg/mL and frozen on liquid nitrogen for further long-term storage.

f) Purification of the fusion protein cenp-T-LacI

The gene encoding the fusion protein cenpT-LacI was extracted by Xba1 and Eag1 from the backbone pET28a and inserted into a pET23 template previously digested by Nhe1 and Eag1 (Supplementary). The cenpT-LacI protein is expressed in *Escherichia coli* BL21 (DE3) RIL strain following the standard protocol for an induction temperature of 19°C. The purification of cenp-T/LacI was tried by IEC, HIC, IMAC-Ni^{2+}, SEC or by ammonium sulfate precipitation. For each of these, I followed the standard protocol written here or provided by the company GE Healthcare. The ammonium sulfate is performed by increasing the concentration of ammonium sulfate. The proteins will precipitate according to their tolerance to the ionic force. Cenp-T/LacI precipitate around 32% of ammonium sulfate in PBS. The purification protocols are more discussed in the results part.

g) Recruitment of LacI/LacO beads

To target LacI to LacO beads, 100µg of pure LacI is incubated with magnetic beads linked or not to LacO arrays in PBS at RT for 15min. After incubation, the beads are washed three times with 800µL of PBS at RT. The beads are resuspended in 20µL of SDS-PAGE loading buffer, boiled at 95°C for 5 min and 8µL of each sample are loaded in each well. After analysis of the samples by SDS-PAGE stained by Coomassie blue, a band at 40kDa corresponding to LacI appears only in the samples of beads with LacO.

h) Recruitment of cenp-T/LacI/LacO beads

The equivalent of 50 ml of bacteria culture expressing cenpT/LacI is thawed on ice to prepare for cell lysis. The lysis is performed by sonication on ice as a cycle of 20% with an output power of 20W for 30sec per sample. The bacterial lysates are centrifuged at 25,000rcf for 20min at 4˚C. Then, 800µL of the soluble fraction of the cell lysis is mixed with 20µL of LacO arrays beads and incubated for 20min at RT. The beads are pelleted using a magnetic rack and the solution is discarded. The beads are washed three times with PBS + EDTA 0.5mM. Finally, the beads are suspended in 20µL of PBS + EDTA 0.5mM and kept on ice. For analysis, the beads are resuspended in 20µL of SDS-PAGE loading buffer, boiled for 5min at 95˚C and the 8µL of samples are analyzed by SDS-PAGE stained by Coomassie blue.

i) Recruitment of kinetochore on cenp-T/LacI beads

The recruitment of the kinetochore proteins on LacI/LacO beads has been extensively optimized. After the recruitment of cenp-T/LacI on LacO arrays beads, the beads are incubated in 200µL of Xenopus Egg Extract (XEE) for 1h at 21˚C. The samples are slightly hand tapped every 5min. The beads are washed three times in a volume of 800µL of the same buffer than used for XEE. The beads are resuspended in 30µL of SDS-PAGE loading buffer. Thus, it will be discussed in the chapter of Results.

III Cell Biology

1) Reagents

- Fetal Bovine Serum (FBS) (#S01520, Biowest)
- Dulbeccco's Modified Eagle Medium (DMEM) (#10313-021, Gibco)
- Glutamax (#35050-061, Gibco)
- 0.25% Trypsin-EDTA (#25200-056, Gibco)
- Penicillin and Streptomycin (Pen Strep) (#15140-122, Gibco)
- geneticin (G418) sulfate (#400-113, Gemini)
- 100mm cell culture dishes (#T1110, Denville)

- Phosphate Buffered Solution (PBS): 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH7.4), 137mM NaCl, 2.7mM KCl.

- Complete growth media: DMEM, 10% FBS, 1X Glutamax, 1X Pen Strep.

- Selection growth media: DMEM, 10% FBS, 1X Glutamax, 1X Pen Strep, 400mg/L geneticin (G418).

2) HeLa cell line

a) Maintenance of HeLa cells

The cells are incubated in 10mL plates at 37˚C in 5% CO_2 and wet atmosphere. The division time is around 20-24h under optimal conditions. The HeLa cells are adherent. However, their mutual contact inhibits their growth. Thus, once they reached 100% confluency (proportion of the surface covered by cells), they have to be split. The confluency directly influences the gene expression and growth of HeLa cells.

The splitting protocol is:

1) Aspirate off the cell media
2) Wash the cells with 5mL PBS
3) Aspirate off the PBS
4) Add 2mL of Trypsin/EDTA (just enough to cover the surface of the 10mL plate)
5) Incubate the cells at 37˚C for ~2min
6) Add 8mL of complete growth media to the dish to quench the trypsinization
7) Pipette media up and down and squirt it around the surface of the dish to remove adherent cells from the surface
8) Put the 10mL of cells into a 15mL Falcon canonical
9) Add the appropriate concentration of cells into new 10 cm dishes and/or six well plates for further experiments
10) Spread the cells evenly by rocking the dish/plate back and forth.

b) Transfection of HeLa cells

Transfection is the process allowing the introduction of nucleic acids into eukaryotic cells by a non-viral method. HeLa cells are transfected with Viafect (#E4981, Promega) when their confluency is around 60-70%, following the instructions of the company. The selection of transfected cells is performed using G418, an antibiotic related to the gentamicin.

IV Microscopy

1) Buffers and reagents

- Phosphate Buffered Solution (PBS): 10mM Na_2HPO_4, 2mM KH_2PO_4 (pH7.4), 137mM NaCl, 2.7mM KCl.

- Tris NaCl Tween-20 buffer (TNT buffer): 10mM Tris (pH7.5), 150mM NaCl, 0.1% Tween20.

2) Cell imaging on cover slips

To determine the presence and the localization of proteins in the cells, I used specific antibodies and a fluorescent probe, the 4', 6-diamidino-2-phénylindole (DAPI) to reveal DNA. A Zeiss Axio Observer.Z1 inverted microscope was used for the imaging with the following objectives: LD A-plan 20x/0.3 (#441241-9910, Zeiss) and NeoFLUAR 40x/1.30 oil (#420462-9900, Zeiss).

The standard protocol to stain cells for cell imaging:

1) grow the cells on a cover slip localized on the bottom of the plate
2) wash cells with PBS
3) remove the PBS
4) add a solution of 4% ParaFormAldehyde (PFA) and incubate for 10min
6) remove and wash twice with PBS
7) add PBS + 0.5% Triton X-100 and incubate for 10min
8) remove the solution, add TNT buffer + 10% Serum and incubate for 20min
9) remove the solution
10) add primary antibodies mix with TNT buffer + 10% serum
11) incubate at RT for 1-2h or o/n at 4˚C

12) remove the solution

13) wash three times with TNT buffer

14) mix secondary antibodies (1:1000) with TNT buffer

15) incubate at RT for 1-2h

16) wash three times with TNT buffer

17) remove the media

18) add 5µL of commercial DAPI solution on the slide

19) put the coverslips on the slide

The slides are now ready and can be kept for years at 4˚C or -20˚C

3) <u>Live imaging</u>

The cell imaging experiments are performed with the Zeiss LSM 780 confocal microscopy (inverted) with a culture chamber at the NICHD MIC facility under the direction of Dr. Vincent Schram.

Results

I Objectives

When I joined the laboratory, cenp-M was not identified as a member of the cenp-H/I/K complex. For this reason, my experiment focused on the cenp-H/I/K complex. To determine the role and effects of the poly-SUMOylation of cenp-I on the kinetochore, a mandatory tool is the use of mutant cenp-I unable to be poly-SUMOylated. To do so, the first step is to identify the SUMOylation sites to generate an unSUMOylable cenp-I mutant and thus determine the phenotypes on cell behaviors. There are 58 lysines in human cenp-I, thus the systematic mutagenesis of each residue, alone or in combination, would have been time consuming and could have led to misinterpretations due to artifacts. Due to technical limits, an immuno-precipitation coupled to silver nitrate staining and mass spectrometry analysis could not be performed. So to identify the SUMOylated lysines of the cenp-H/I/K complex, I decided to perform an *in vitro* SUMOylation assay of the cenp-H/I/K complex.

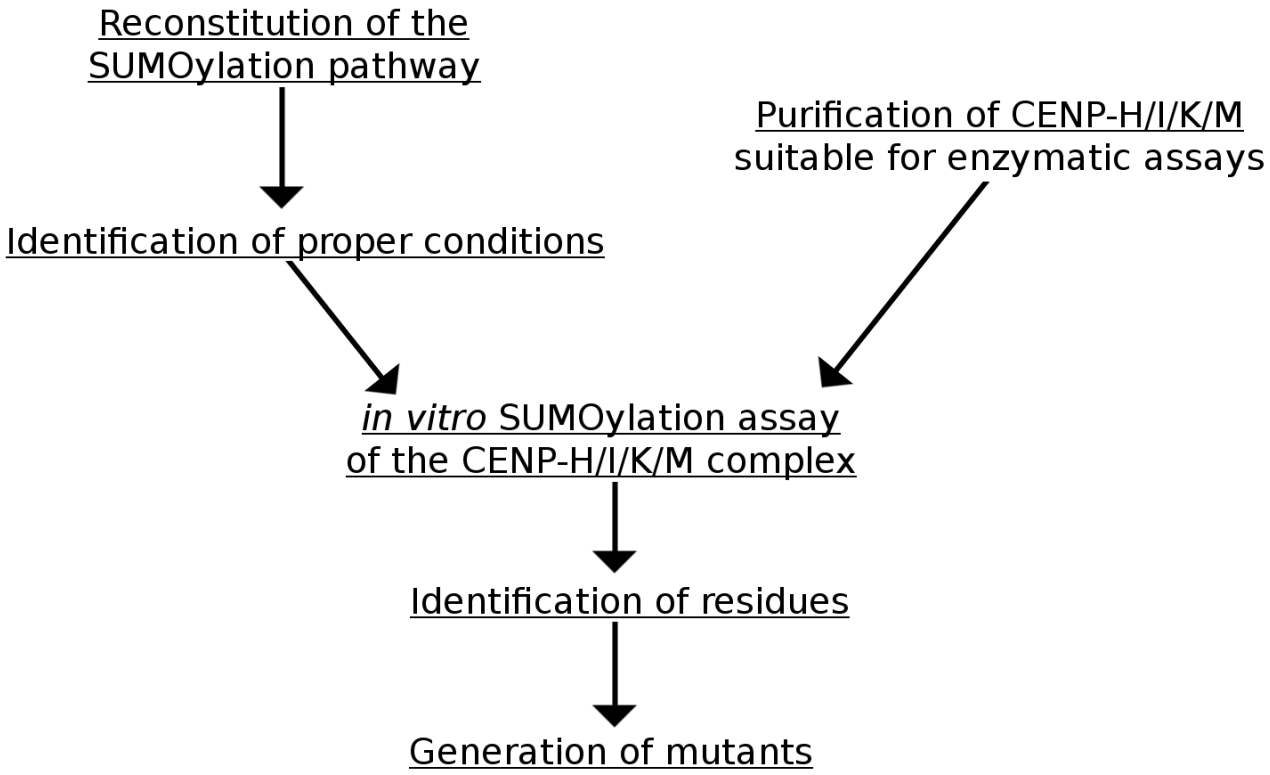


Figure 17: Experiment plan

At that time, several teams attempted to purify recombinant sub-complexes of the kinetochore. Since cenp-H/I/K is known to be insoluble or very instable, they tried to over-express and purify each inner kinetochore sub-complexe one by one. We chose two different approaches. Since the molecular environment provided by the inner kinetochore could affect the SUMOylation process of cenp-I, we first tried to perform SUMOylation assays of cenp-I within soluble parts of the inner kinetochore produced in bacterial system. The second approach was to reconstitute *in vitro* kinetochore suitable for biochemical assays such as SUMOylation. We used a published protocol that recruits kinetochore from xenopus egg extracts onto magnetic beads. This protocol requires advanced skills in biochemistry and weeks of experiments. Unfortunately, we found that the kinetochore produced with this protocol is not suitable for biochemical assays. We tried to reconstitute *in vitro* kinetochore by adapting the method used by Dr. Cheeseman and Dr. Fukagawa (Karen E. Gascoigne et al. 2011; Tetsuya Hori, Amano, et al. 2008). They fused kinetochore proteins with the Lactose repressor (LacI) to target them to chromosome arms and determine which kinetochore proteins are able to reconstitute a functional kinetochore.

II SUMOylation of the cenp-H/I/K/M complex

To identify the lysines involved in the SUMOylation of cenp-I, I first focused on the purification of cenp-I to perform *in vitro* SUMOylation assays. However, the SUMOylation of the cenp-H/I/K complex can vary depending of the environment. The SUMOylation of the cenp-H/I/K complex could be greatly different to the endogenous one.

1) Expression and purification of the CCAN

My first approach was to overexpress and purify the whole Constitutive Centromere Associated Network (CCAN) from a bacterial system. The constructs were gratefully provided by the Stuckenberg team. The genes encoding cenp-C/B/N/H/I/K/M/T/W/S/X/P/O/U/L/Q/R, corresponding to the CCAN, are inserted in a pST44 backbone (see Supplementary data). This plasmid confers an Ampicillin resistance to bacteria and cenp genes are under a T7 promoter. I will name this construct as Dr. Stuckenberg does, the monstruct.

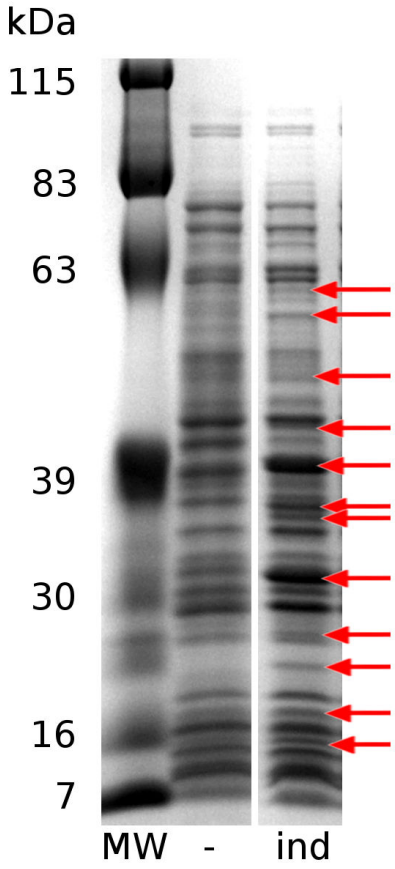
a) Production/Optimization

An expression test was done as describes into the part "protein overexpression" in the chapter Material and Methods. Decreasing the induction temperature reduces the translation speed and thus increases the time of expression of the nascent protein, allowing the protein to be fold properly. Different conditions of induction were tested such as 0.5mM IPTG with a 180rpm shaking at 37˚C for 3h, 28˚C for 6h and 19˚C for 15h, with and without 2% Ethanol + 5% glycerol. According to previous experiments in the laboratory, the addition of 2% ethanol + 5% glycerol prevents aggregation of proteins. We assume that the glycerol increases the solubility of proteins, increases the viscosity and the molecular dynamics leading to increase the folding time. The adding of 2% ethanol leads to stress of bacteria that produce chaperones. These chaperones help to refold the proteins. To analyze the proteins in samples, Sodium DodecylSulfate-PolyAcrylamide Gel Electrophoresis (SDS-PAGE) was used. This method allows for the separation of charged molecules regarding their molecular weight through a poly-acrylamide gel when a constant electric current is applied. The poly-acrylamide acts as a sieve and the presence SDS leads to denaturating conditions (Shapiro, Viñuela, and Maizel 1967; Laemmli 1970).

The production pattern is similar in all induction conditions tested (data not shown). In the figure 18, the results of an induction at 18˚C are shown. Between the control (non-induced) and the experiment (induced) lanes, several bands with different molecular weights appear after induction (Figure 18.A). We expect to see a band for each of the 17 proteins encoded by the monstruct, listed in the Figure 18.B. Each band that appears after induction is consistent with the expected molecular weights of a protein listed in the Figure 18.B. However, the bands of induced proteins have different intensities suggesting that the proteins of the CCAN are differently expressed. The identification of proteins in the gel could be performed by Mass Spectrometry analysis. However the identification of 17 proteins is resource consuming, especially for performing routine identification. I am looking for a pattern of bands that is similar than expected after purification to send the samples to MS analysis (Figure 18.B).

In order to identify the SUMOylation sites, we needed to optimize the production of the monstruct and the purification of the CCAN. I started from bacteria expressing the monstruct at 18˚C o/n with 2% Ethanol + 5% glycerol.

A



B

Gene (CENP)	Molecular Weight (kDa)
C	107
B	65
T	60
I	60
U	47
L	44
N	40
O	34
P	33
K	32
Q	31
H	29
R	20
M	20
S	16
W	10
X	7

Figure 18: Test of Overexpression of CCAN proteins.
A) Production of monstruct in *E. coli* after induction. The SDS-PAGE is stained by Coomassie blue.
B) List of genes present in the monstruct with their theoretical molecular weights.

b) Purification of the monstruct

Biochemical assays requires that the protein is soluble and functional in the experiment conditions, and that a negligible amount of contaminants can be found in the sample. Usually, the proteins are purified and characterized after purification to confirm they fit to the requirements. To purify proteins, I used two approaches, 1) the use of purification tag and 2) the use of the intrinsic properties of the protein.

The purification by tag uses the affinity between a peptide and a non-peptidic molecules. The peptide is added to the interest protein by molecular biology while the ligand is covalently fixed on the stationary phase (agarose/sepharose). The samples with the tagged protein (lysate) are incubated with the stationary phase. After an incubation time, the flow through is discarded, the matrix washed and only the proteins interacting with the ligand remain bound on the stationary phase. The protein of interest can be eluted by competition (Table 5).

Column	Principle	Elution Process	Comments
Size Exclusion Chromatography	Size differences, presence of complexes (dimer, trimer…)	Constant flow	Good to discriminate aggregate and multimers
Ion Exchange Chromatography	Ionic interactions	Increase of the salt concentration	No tag required
Immobilized Metal Affinity Chromatography- Ni^{2+} ($IMAC\text{-}Ni^{2+}$)	Histidine interacts with Ni^{2+} through its imidazole group	Addition of competitive imidazole	Requires an affinity tag
Hydroxyapatite	Anionic, cationic and hydrophobic interactions	Increase of the salt concentration	No tag required
Hydrophobic Interaction Chromatography	Hydrophobic interactions	Decrease of the salt concentration	No tag required
SUMO-affinity column	Interaction with SUMO	Addition of DTT	No tag

Table 5: Chromatographic systems used during my doctoral work.
The chromatographic system used is an Akta Purifier (GE Healthcare).

Most of the time, in order to circumvent the intrinsic properties of the protein that are often not known and challenging to determine, a specific sequence of amino acids (tag) is added to its sequence. Several tags are available and present various biochemical properties that are useful for purification such as the poly-histidines tag, the maltose binding protein or the thioredoxin.

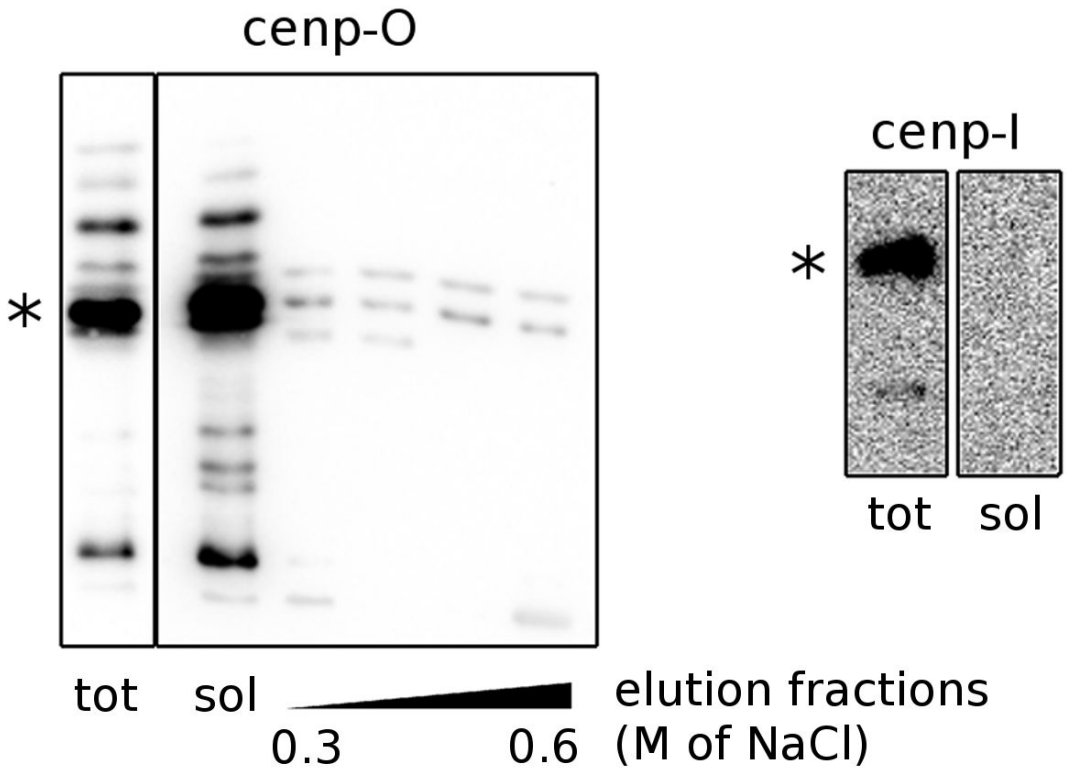
c) Purification of the monstruct by Ion Exchange Chromatography

I first performed a purification test of the CCAN using an Ion Exchange Chromatography (IEC). This method of chromatography separates proteins as a function of their charge, an intrinsic property provided by their composition, structure and buffer conditions. The principle of IEC is to separate ions and polar molecules based on their affinity with the ion exchanger that is linked to the stationary phase of the column. At a given pH, the protein residues show an ionic charge that can be positive or negative. For example, the lysine, arginine and histidine residues are positively charged under physiological conditions (pH7.4), while the aspartate and glutamate are negatively charged. According to the presence of multiple charged residues, the protein has a global charge, the isoelectric point (pI). Nonetheless, the repartition of the same type of charged residues (+ or -) along the protein is not homogenous and can lead to specific charged "patches". Thus, although a protein that is globally negatively charged, it can strongly interact with cation exchange chromatography.

A 0.5L bacterial culture is pelleted and sonicated following the protocols described in the chapter Material and Methods. The soluble fraction is then purified by an Anion Exchange Chromatography using a Q sepharose HiTrap 5 ml (GE healthcare) with a gradient between 0 and 1M NaCl. In the elution fractions, the predicted protein bands are not found. However, since Coomassie blue staining has low sensitivity, we identified some of the kinetochore proteins by western blot (Figure 19). Briefly, the Western Blot is an analytic technique to detect specific proteins in a sample using specific antibodies following their transfer onto a membrane. The primary antibodies recognize a specific epitope on the protein and secondary antibodies, which are coupled to a fluorophore or an enzyme, target the primary antibody, allowing the amplification of the signal. The revelation occurs using the properties of the fluorophore or the enzymatic activities.

To circumvent the lack of antibodies for each protein of the monstruct, I took advantage of the network of the inner kinetochore and its sub-complexes. According to *in vivo* studies, each sub-complex cannot be targeted into the inner kinetochore if one of its members are absent. For this reason, I consider that the presence of a member of a sub-complex means that the whole sub-complex is present. Thus, I tested the presence of cenp-O proteins from the cenp-R/O/P/Q/U complex, cenp-I from the cenp-H/I/K/M complex.

After Western Blot, the bands corresponding to cenp-I is present in the total fraction at the expected MW but not in the soluble fraction (Figure 19), suggesting that the protein is fully expressed but not soluble. However, cenp-I is neither soluble alone nor within a soluble complex. Concerning cenp-O, the protein is detected in the soluble fraction and in several elution fractions. That could be due to interactions with multiple partners such as other proteins of the cenp-R/O/P/Q/U complex and/or the presence of several multimeric forms.



<u>Figure 19:</u> Western Blot analysis of monstruct purification by anion exchange chromatography. tot: fraction of cells after induction. Sol: soluble fraction after lysis. Asterisks show the expected molecular weight. The antibodies used are written at the top of the gels

d) <u>Precipitation of the bacterial genome and its associated proteins</u>

The cenp-H/I/K complex, which is required for the kinetochore integrity *in vivo,* is not present in the soluble fraction. The kinetochore proteins cenp-C/T/W/X/S are responsible for recruitment of the kinetochore on the centromere and therefore strongly bind DNA. We reasoned that the absence of soluble cenp-H/I/K complex could be due to the DNA binding properties of the

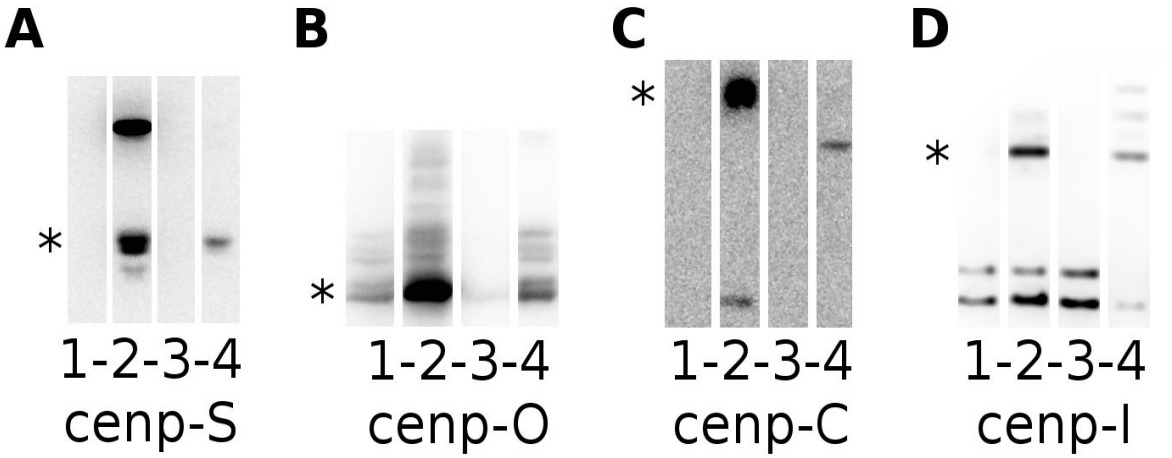
cenp-C/T/W/X/S proteins. These proteins, could bind the bacterial chromosome or other cellular DNA and recruit the rest of CCAN. Since the lysis and centrifugation steps are harsh conditions for proteins and genomic DNA, most of the bacterial genome and its associated proteins are in the insoluble fraction. In conclusion, the CCAN was potentially not loaded onto the column on this first purification test.

To determine if these proteins are localized on the genomic DNA, I attempted to extract and precipitate the nucleic acids and its associated proteins. The bacterial pellet is thawed, sonicated and centrifuged at low speed (500rcf) to specifically remove high-density components and insoluble molecules. The nucleic acid precipitation is done by adding a final concentration of 5% of streptomycin to the soluble fraction. The streptomycin preferentially binds the 16S rRNA but at high concentration, streptomycin binds to all nucleic acid and forms precipitates that can be pelleted and analyzed by western blot to check for interactions with CCAN proteins.

I tested for the proteins cenp-C, cenp-S, cenp-I (cenp-H/I/K/M complex) and cenp-O (cenp-R/O/P/Q/U complex) (Figure 20) in non-induced bacteria to determine the specificity of the antibodies, in the soluble fraction and the precipitate after precipitation by streptomycin of bacteria expressing the CCAN to determine the protein expression and/or degradation.

The western blots analyses reveal bands in the induced lanes and not in the non-induced, suggesting that the proteins are expressed during induction and that the antibodies are specific. The proteins cenp-C/I/O/S are present in the streptomycin precipitate (Figure 20.lanes 4) but not in the soluble fractions (Figure 20.lanes 3). A band with a molecular weight slightly lower than expected (90kDa compared to the expected 105kDa) is present in the precipitate lane of cenp-C experiment (Figure 20.C.lane 4). That could be assigned to a form of cenp-C due to the lack of other bands in the sample and in the soluble fraction.

According to the results of the streptomycin precipitation, the kinetochore proteins tested are loaded on nucleic acids. However, the proteins could be also be insoluble in presence of high concentration of streptomycin. Thus, I cannot conclude that the CCAN-like complex structural and functional integrity is conserved during this reconstitution.



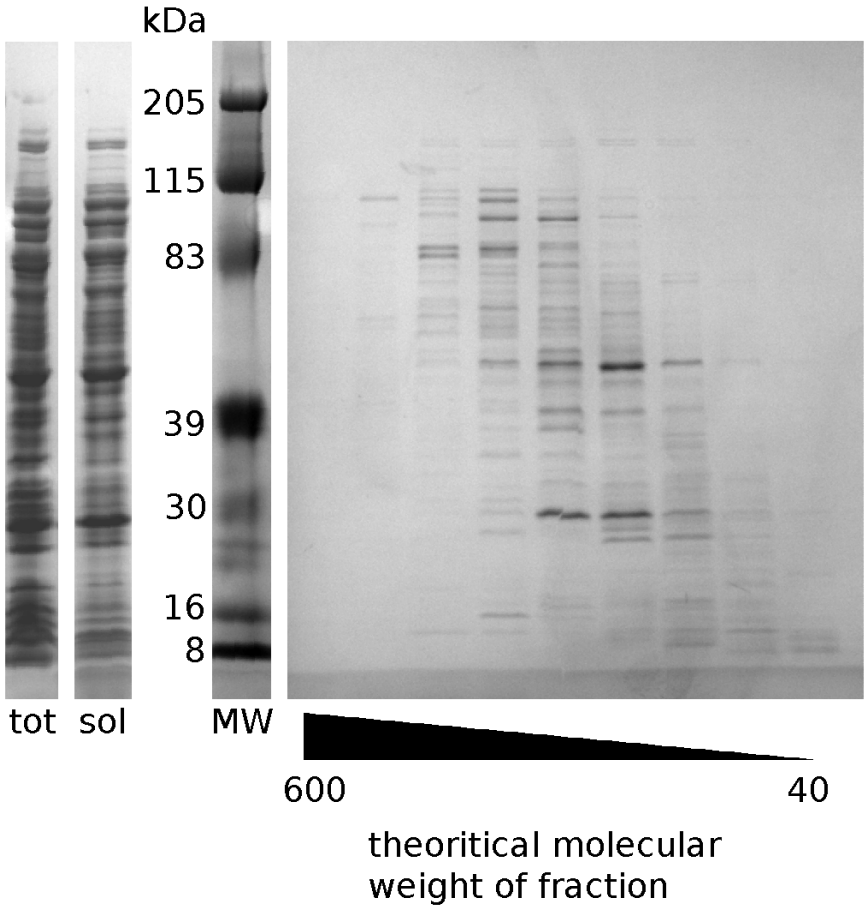
<u>Figure 20:</u> Western Blot of nucleic acid precipitation by streptomycin.

A) Detection by cenp-S antibody. B) Detection by cenp-O antibody. C) Detection by cenp-C antibody. D) Detection by cenp-I antibody.
For each gel, 1 is the non induced bacteria, 2 is the induced bacteria, 3 is the soluble fraction, 4 is the nucleic acid bound protein. The asterisk show the expected molecular weight for each protein.

If the CCAN is loaded on DNA, it would be interesting to purify it together with its native DNA. Thus, the fraction containing the nucleic acids is incubated with DNaseI enzyme, that digests the nucleic acid that are not protected by proteins. The product of digestion is then purified by Size Exclusion Chromatography (SEC). The SEC, also called Gel Filtration (GF), separates the molecules based on their radius of gyration and their hydrodynamic radius. For these reasons, it is considered that the SEC allows the separation of proteins and complexes according to their molecular weights. In theory, a complex with one copy of each protein of the monstruct has a molecular weight of 650kDa. Thus, the CCAN is expected to be eluted at high molecular weight.

After the lysis of the bacteria expressing the monstruct, a low speed centrifugation is performed to keep the nucleic acid soluble. After centrifugation, 100µg/ml of DnaseI is added to the soluble fraction and the sample is incubated at 30˚C for 30 min. Aggregates sticks to the SEC and decrease its separation ability, a centrifugation is performed at 13,000g for 10 min at 4˚C after digestion. The sample is then loaded onto a HiLoad 16/600 Superdex 200 with a constant flow of 1mL/min of Tris-HCl 20mM pH8, NaCl 150mM. The elution pattern is analyzed by SDS-PAGE stained by Coomassie blue and by WB. Bands appear in the elution fraction between

150kDa and 40kDa. No bands appear in the first two elution fractions that are the elution fraction of complexes with an approximate molecular weight of 600kDa and 250kDa.



<u>Figure 21:</u> Purification of CCAN complexes using SEC.
The SDS-PGE are stained by Coomassie blue. A) The soluble fraction of lysate of bacteria expressing the monstruct is digested by DNaseI for 30min and purified by a Superdex S200 10/300 SEC. The fractions corresponds to theoretical elution of complexes with a respective molecular weight of 600, 250, 150, 100, 85, 75, 60, 50 and 40kDa. B) Western Blot analysis of fractions. Tot = bacteria fractions after unfreezing, sol = soluble fraction after the lysis process.

To confirm that the kinetochore and its support DNA is released after digestion by DNase1, western blot analysis are performed with antibodies against cenp-I and cenp-O (data not shown). cenp-I is only detected in the total fraction, and not in the soluble fraction or the elution fractions. It suggests that cenp-I is insoluble in this condition. cenp-O is detected in the total fraction, in the

soluble fraction and in several several elution fractions. cenp-O is expressed, soluble and the detection of cenp-O in several fraction after SEC suggests that it is involved in several complexes. However, with these results, I cannot determine the nature of these complexes.

cenp-I is present in the insoluble fraction but I do not know if it is due to cenp-I forming aggregates that are soluble or if the complex is highly instable. A CCAN containing the cenp-H/I/K/M complex cannot be reconstituted by this approach. Since the production and the purification of a 17-mere complex led with a lot of experimental and technical issues to solve, we decided to try a different approach.

e) The purification test of cenp-I protein

Since the reconstitution of the CCAN to study cenp-I SUMOylation is not possible, we focused on the cenp-I associated sub-complexes. When I started my doctoral work, cenp-M was not known as a member of the cenp-H/I/K complex. The cenp-H/I/K complex is too unstable to perform SUMOylation assays, so I tried to purify cenp-I alone by using IMAC-Ni^{2+}. This is a widely used method for protein purification that is based on the six coordination sites of the Ni^{2+} molecule. In the NitriloTriAcetic (NTA) acid column matrix, each NTA molecule interacts with four coordination sites of the Ni^{2+}. The two remaining sites of the nickel interact with the N7 of the imidazol ring of the histidine. The addition of a poly-histidine (6-H) tag by molecular biology to the protein sequence offers an easy method to perform protein purification. It is to note that a poly-histidines tag with 8 or 10 histidines is required for eukaryote expression.

A 6-histidine tag was added to the N-terminal tail of cenp-I in a pET28 backbone (see Supplementary). The expression patterns are similar in all conditions tested: induction with 0.5mM IPTG with a 180 rpm shaking at 37˚C for 3h, 28˚C for 6h and 19˚C for 15h, with and without 2% Ethanol + 5% glycerol (data not shown).

Protein expression was revealed by SDS-PAGE followed by Coomassie blue staining. A low intensity band at the expected MW and a band with a higher intensity at a small MW were observed (Figure 22.C). The low MW band could be the result of the degradation of cenp-I. To test that hypothesis, the degradation rate was monitored (Figure 22.A.B). The amount of cenp-I in the

cells remains constant over the time (Figure 22.B), while the intensity of the low MW band increases (Figure 22.C). In conclusion, the overexpression of cenp-I in *E. coli* leads to its degradation.

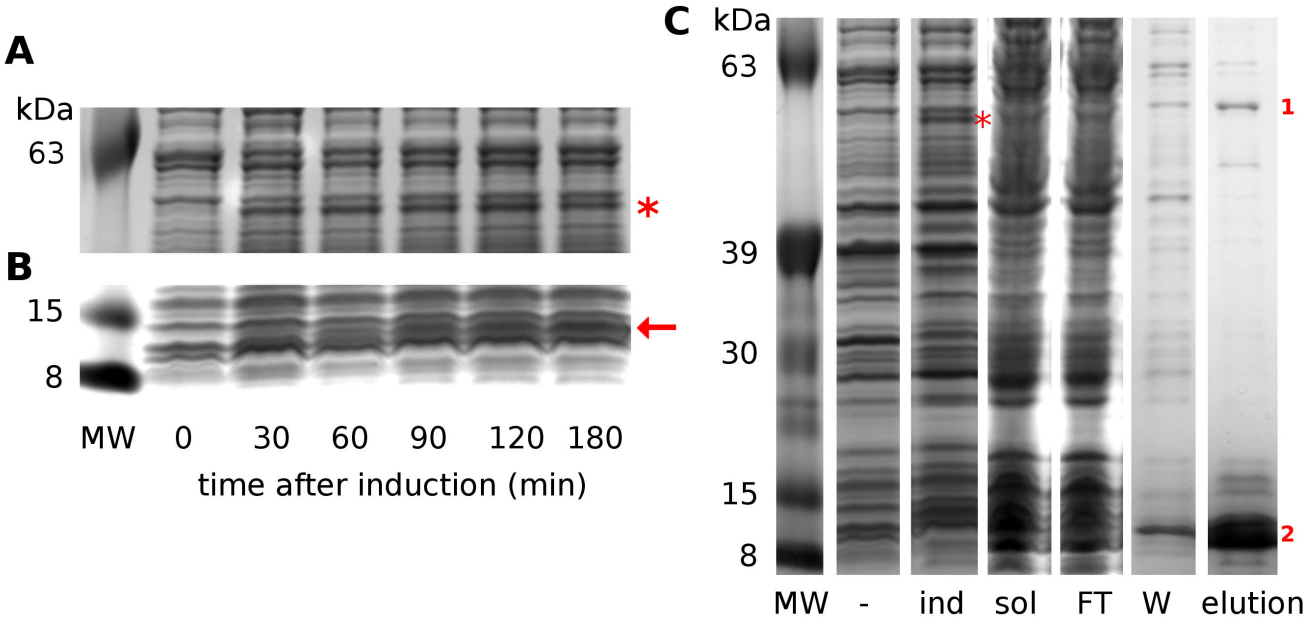


Figure 22: Expression and purification test of cenp-I in *Escherichia coli*.
The SDS-PAGE are stained by Coomassie blue. A) Evolution of cenp-I expression after induction. B) Evolution of the degradation product of cenp-I after induction. C) Purification of cenp-I using an IMAC-Ni^{2+}.
Star marks the FL cenp-I. The arrow marks the low molecular weight that appears after induction.

Since it is not a full degradation of cenp-I, a purification test was performed using IMAC-Ni^{2+}, leading to the purification of two bands (Figure 22.C1 and 2). The SDS-PAGE analysis shows a band with a low intensity at the expected MW for full length cenp-I (Figure 22.C.1 ~65kDa) and a intense band at a smaller MW (Figure 22.C.2, ~10kDa), which could be a product of degradation of cenp-I. Further WB analysis demonstrated that the upper band is not related to cenp-I (data not shown) but a contaminant, probably GlmS (Bolanos-Garcia and Davies 2006).

The small band after purification has a molecular weight identical to the degradation product in the induction lane. This protein is efficiently purified by IMAC-Ni^{2+}. Because there is no

histidine cluster in cenp-I wild type, it suggests that the small band contains the 6-Histidines tag of the recombinant cenp-I protein released after degradation. To note that there is no identified domain in the N-terminal part of cenp-I. Even if the cenp-I recombinant protein is expressed in the bacteria (Figure 22.A.*), it cannot be purified by IMAC-Ni^{2+} with the standard protocol.

2) SUMOylation assays on the cenp-H/I/K/M complex

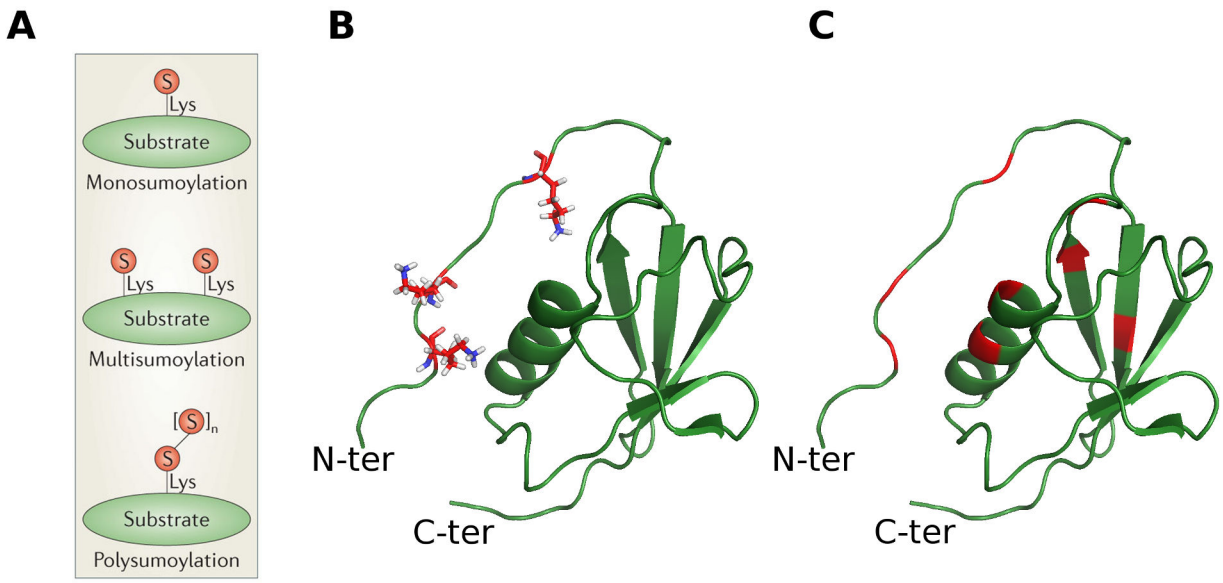
Meanwhile, a paper from the Pr. Musacchio's team revealed a fourth component of the cenp-H/I/K complex, the cenp-M protein (Basilico et al. 2014). While the cenp-H/I/K complex is unstable, the cenp-H/I/K/M with a six-histidines tag added at the N-terminal tail of cenp-H forms a complex stable enough for biochemical assays. We engaged collaboration with Pr. Musacchio. They provided the pure cenp-H/I/K/M complex and we were in charge of the SUMOylation assays to identify the sites of SUMOylation by Mass Spectrometry analysis. Several micrograms of SUMOylated cenp-I is required to readily identify every SUMOylation sites by MS analysis.

a) The generation of the SUMOylation pathway

In order to mimic cenp-H/I/K/M poly-SUMOylation, I am using SUMO2/3 that compose the poly-SUMO chains (Tatham et al. 2001), but we are expecting SUMOylation on multiple sites on cenp-I, and not by the poly-SUMO chains on single sites. Every SUMO2/3 added leads to a 10kDa mass increment of the substrate. After analysis by SDS-PAGE of SUMOylation reaction with SUMO2/3 wild type, a ladder pattern appears but it is impossible to discriminate between multi and poly-SUMOylation.

This issue led me to generate SUMO 2/3 mutants that cannot form poly-SUMO chains, thus ensuring that the increment number reflects the state of multi-SUMOylation of the protein. Since the poly-SUMOylation occurs at the N-terminal part of SUMO2/3, two mutants of human SUMO2 were generated: 1) the lysines 5, 7 and 11 were mutated to arginine ($SUMO2^{K5,7,11R}$) (Figure 23.B) and 2) all lysines were mutated to arginine ($SUMO2^{KR}$) (Figure 23.C).

I purified all proteins belonging to the SUMOylation pathway. The proteins used are SUMO2 wild type and its mutants, E1, E2 and two different E3 (PIASy and truncated RanBP2). SUMO2 and its mutants and E2 are purified by IMAC-Ni^{2+} following by a SEC (see Material and Methods). E3 are purified by IMAC-Ni^{2+} following by an IEC and finally by A SEC. Unpublished results from our lab suggest that PIASy could be responsible for the SUMOylation of the kinetochore.

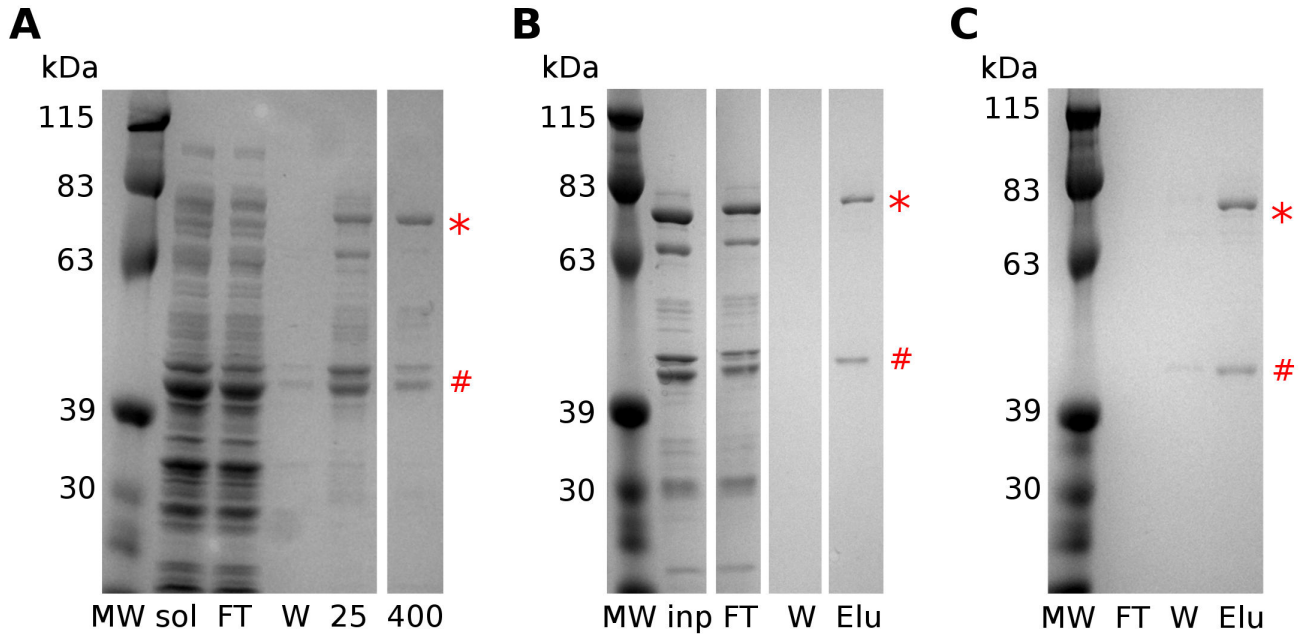


<u>Figure 23:</u> The SUMOylation types of SUMOylation.
A) Substrates could be mono, multi or poly-SUMOylated. From (Hickey, Wilson, and Hochstrasser 2012). B) Structure of human SUMO2 (pdb 1WZ0). The residues in red show the lysines 5, 7 and 11. C) Structure of human SUMO2 (pdb 1WZ0). The red parts show all lysines.

A 6-histidines tagged of each uba2 (E1), E2, E3, SUMO2 and its mutants are inserted into the backbone pET28a. For E2 and versions of SUMO2, the production is performed at an induction temperature of 37°C. The purification is performed by IMAC-Ni^{2+} followed by SEC. The fractions with pure proteins are pooled and concentrated to 20 mg/mL, then frozen in liquid nitrogen and stored at -80°C. The SUMO2 ^{KR} mutant is less soluble and is concentrated at 10 mg/mL before freezing and storage processes. To avoid any interference of the histidine tag in the following experiments, I decided to remove the tag. Thus, the histidine tag of the E2 product was digested by thrombin for 30min at 30°C. The product of the digestion was purified by SEC.

The purification of PIASy was more challenging. An additional step of purification by anion exchange chromatography is inserted between the IMAC-Ni^{2+} and the SEC to purify PIASy. Until the SEC, the purifications have to be done quickly cause PIASy is degraded during time. After IEC and SEC, the samples were analyzed by WB using a specific antibody. Indeed, several endogenous bacterial proteins exhibit a molecular weight close to PIASy. The SEC leads to the purification of two forms of PIASy, a full length (65kDa) and degraded one (55kDa). I pooled and used the fractions with the full length PIASy.

The E1 enzyme is an heterodimer composed of Uba2 and Aos1. The purification of E1 is the most difficult among proteins of the SUMO pathway. After months to attempt the purification of this complex, I met Dr. Azuma who gratefully shared a working protocol. Basically, this protocol is based on two-step purification. First, E1 is purified by IMAC-Ni^{2+} to remove endogenous bacterial proteins. According to Dr. Azuma, the ratio is about 90% of inactive E1 form after IMAC-Ni^{2+} purification. Then, the active E1 protein is purified with a SUMO-affinity column.



<u>Figure 24:</u> Purification of recombinant human E1.

The SDS-PAGE are stained by Coomassie blue. A) Purification of E1 by IMAC-Ni^{2+} from bacterial expressing the recombinant E1. A six-histidines tag is added to Uba2. B) Purification by NHS-SUMO. C) Purification by Ion Exchange Chromatography. For each, MW: Molecular Weight. Sol:

soluble fraction. FT: Flow Through. W: Wash. 25: 25mM imidazole. 400MM: 400mM imidazole. Elu: Elution.

The recombinant E1 complex, Aos1 wild type (42kDa) and uba2 with a six histidines tag (80kDa) is overexpressed in *E. coli* BL21 (DE3) RIPL. After cell lysis, E1 is purified by IMAC-Ni^{2+} as describes in the chapter Material and Methods except that the buffers used are different. There is no imidazole and no DTT in the lysis/wash buffer (20mM Tris pH8, 100mM NaCl). The elution is performed with an elution buffer of 20mM Tris pH 8, 100mM NaCl and 30mM imidazole. The samples are analyzed by SDS-PAGE (Figure 24.A). The lanes 30 and 400 correspond to the concentration of imidazole used to elute the proteins. In the elution at 30mM imidazole, four main bands appear at 42, 45, 63 and 80kDa. The cross and the hashtag show the bands corresponding respectively to Uba2 and Aos1. There are also bands of lower intensity.

The purification by IMAC-Ni^{2+} leads to a mix of inactive and active forms. Dr Azuma took advantage of the biochemical properties of active E1 to covalently bind SUMO as a tag in an ATP dependent process. To prepare the SUMO-affinity column, pure SUMO2 is covalently linked to N-HydroxySuccinimide (NHS) beads. The NHS molecules forms covalent bound with primary amine groups. The E1 is then applied on the SUMO-NHS column in the presence of ATP. The E1 covalently links SUMO through its catalytic cysteine. Then, the SUMO-E1 links is broken by the addition of DTT. The results of this purification is shown in the Figure 24.B. The input of this assay is the fraction of elution at 30mM imidazole of the previous IMAC-Ni^{2+} (lane inp) that contains Uba2 and Aos1 but also contaminants. In the elution fraction (lane elu), there are only two bands at 42 and 80kDa, respectively corresponding to Aos1 and Uba2. No contaminants are detected by SDS-PAGE stained by Coomassie blue. By this method, I am able to both 1) separate active form of E1 (lane Elu) and inactive form that are still in the flow through (lane FT) and 2) to remove all contaminants.

The active E1 is finally purified by anion exchange chromatography to remove the DTT (Figure 24.C). Since E1 is relatively unstable, the purification by IEC allows us to change the buffer quickly and without dilution of the sample, and was preferred over SEC. The activities of the E1 enzyme were validated prior to experiments.

b) Reconstitution of a SUMO pathway *in vitro*

Since the cenp-H/I/K/M complex was a valuable gift from Pr. Musacchio's team, the SUMOylation reaction was first optimized with Psmd1, a protein associated to the proteasome (Figure 24). The SUMOylation of the C-terminal part regulates the activity of this protein. PIASy is the E3 SUMO ligase that greatly increases the reaction rate (Ryu et al. 2014). Psmd1 is used as positive control to first determine the optimal conditions.

To confirm that every protein of the SUMOylation pathway works, a SUMOylation assay with Psmd1 with and without PIASy was performed. In this experiment, Psmd1 is added to a reaction mix that contains 150nM E1, 200nM E2, with or without 50nM E3 (PIASy) and 35µM SUMO2 ^{K5,7,11R}. The reaction is done in 40mM Tris (pH8), 100mM NaCl, 0.05% Tween-20, 5% glycerol, 2mM DTT, 4mM MgCl2, 2mM ATP in a final volume of 60µL.

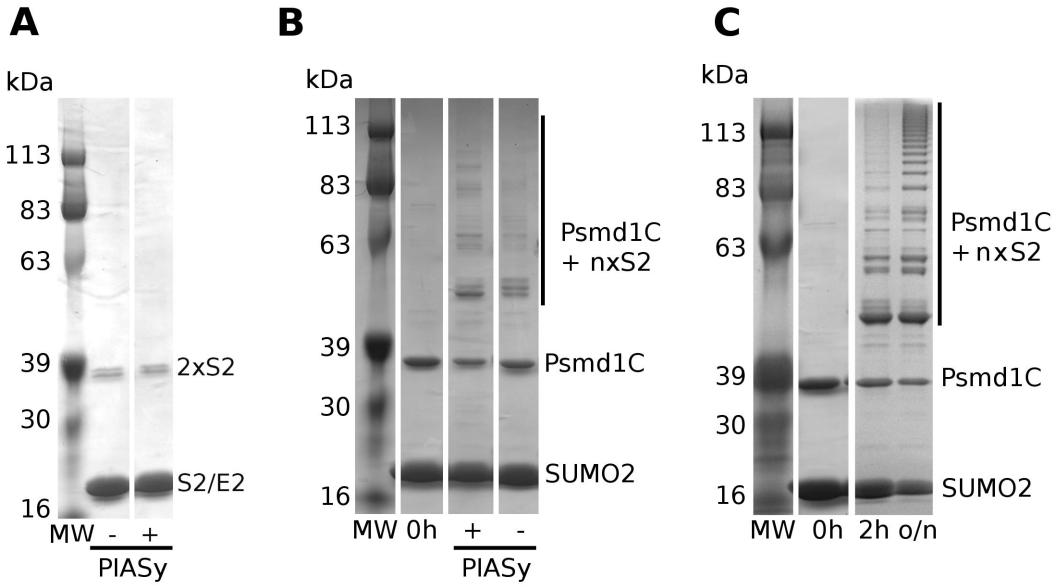


Figure 25: Optimization of SUMOylation reaction.
The SDS-PAGE are stained by Coomassie blue. A) Test of inability of SUMO2 ^{K5,7,11R} to form poly-SUMO chain. SUMO2 ^{K5,7,11R} is incubated without E3 and substrat 2h at 30°C. B) Comparison of the SUMOylation of psmd1 after 1h at 30°C with or without PIASy as E3 SUMO ligase. SUMO2 WT is used. The reaction is displayed at 0min, 120min and o/n at 30°C. SUMO2 ^{K5,7,11R} is used. C) Test of the SUMOylation state after 2h and o/n. The reaction containing 50nM PIASy is incubated at 30°C.

In the Figure 25.A, I tested the ability of SUMO2 ^{K5,7,11R} to form poly-SUMO chain. The reaction contains 150nM E1, 200nM E2, 50nM E3 (PIASy) and 35µM SUMO2 ^{K5,7,11R} and no substrate. The reaction is incubated o/n at 30˚C. After reaction, there are three visible bands, one at 19kDa and two at 38kDa. The bands at 19kDa corresponds to SUMO2 ^{K5,7,11R} alone. The bands around 38kDa correspond to the di-SUMO2^{K5,7,11R} peptide and its intensity is very low, even after an o/n reaction. No band corresponding to tri-SUMO2^{K5,7,11R} is observed. It means that the SUMO2^{K5,7,11R} cannot initiate poly-SUMO chain.

In the Figure 25.B, I tested the ability of PIASy to act as an E3 on the SUMOylation of Psmd1. The results of a 1h reaction are analyzed by SDS-PAGE stained by Coomassie blue. Before incubation, there are 2 major bands, one around 18kDa which is SUMO2 ^{K5,7,11R} and one around 38kDa which is unSUMOylated psmd1. The enzymes are not detected due to their low concentration in the mix. After a reaction at 30˚C for 1h, the pattern drastically changes. Compared to the input, we can see bands migrating at higher molecular weight. In the presence of PIASy, we can detect bands at molecular weights of 38kDa, 50kDa, 70kDa, 85kDa and 95kDa (Figure 25.A.lane +). In the absence of PIASy, we detect three bands around 45kDa (Figure 25.A.lane -). Each SUMO added to the substrate leads to a mass increment of 14kDa. In the presence of PIASy, the band at 38kDa correspond to unmodified psmd1, while the bands at 45kDa, 65kDa, 80kDa and 95kDa respectively correspond to psmd1 with one, two, three and four SUMO2 ^{K5,7,11R}. In the absence of PIASy, there are three bands at around 45kDa. They are mono-SUMOylated psmd1. While there is only mono-SUMOylated psmd1 after 1h reaction without PIASy, there are up to four SUMO tags added to psmd1 in presence of PIASy. From these results, I can conclude that the SUMOylation of psmd1 works and that PIASy increases the reaction but 1h reaction is not enough to SUMOylate all psmd1.

For identification of residues by mass spectrometry analysis, a relative large amount of SUMOylated protein is needed. For this reason, I tried to determine if it is possible to use this assay for a longer time to, increase the amount of SUMOylated proteins. The reaction is repeated and is analyzed at 0h, 2h and 16h (o/n) (Figure 25.C). There are 2 major bands, one around 18kDa which is SUMO2 ^{K5,7,11R} and one around 38kDa which is unSUMOylated psmd1. The enzymes are not detected due to their low concentration in the mix. After 2h of incubation, we can detect bands at higher molecular weights (Figure 25.C.lane 2h). The intensity of the band at 38kDa is lower

than in the input. After an o/n reaction, the intensity of the band at 38kDa is even lower and the upper bands form a perfect ladder at molecular weight greater than 80kDa (Figure 25.C.lane o/n). The band at 38kDa corresponds to the unmodified psmd1. The intensity of this band decreases with time, meaning that the amount of unmodified psmd1 decreases. The upper bands are the SUMOylated forms of psmd1. The intensities of these bands greatly increase between 2h and o/n incubation. The SUMOylation machinery is able to work for a long time. The SUMOylation assay can therefore be used o/n to drastically increase the amount of SUMOylated species.

c) The SUMOylation of the cenp-H/I/K/M does not rely on PIASy

In the initial SUMOylation test of the cenp-H/I/K/M complex, I used the conditions optimized with Psmd1C using PIASy as E3 (Figure 26.A). To characterize the impact of PIASy on the SUMOylation of the cenp-H/I/K/M complex, each reaction performed are done with and without PIASy. I compare the SUMOylation of the cenp-H/I/K/M complex with a reaction with psmd1. The reactions last o/n at 30°C using SUMO2 ^{K5,7,11R}.

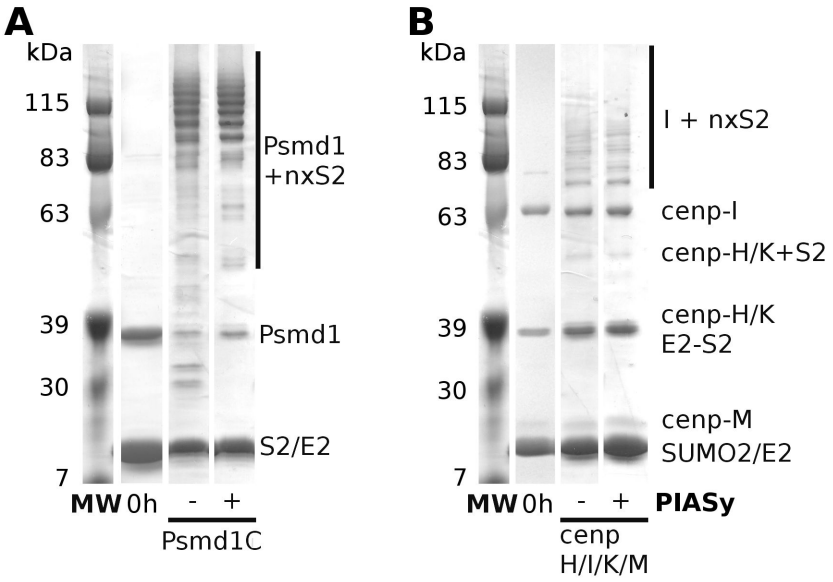
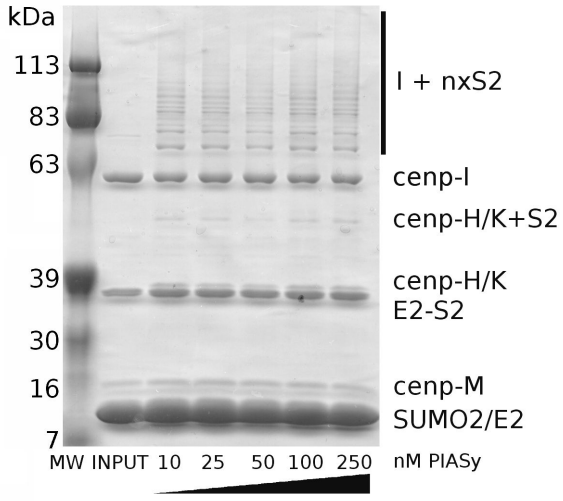


Figure 26: SUMOylation assays of the cenp-H/I/K/M complex.
The assays are performed with 150nM E1, 200nM E2 and 35µM SUMO2 ^{K5,711R} at 30°C o/n and stained by Coomassie blue. A) The SUMOylation of Psmd1C is performed and used as positive control. B) SUMOylation assay of the cenp-H/I/K/M complex.

A SUMOylation assay of the cenp-H/I/K/M complex was performed to determine if the SUMOylation occurs on any protein of the cenp-H/I/K/M complex and if this process is PIASy dependant (Figure 26.A.lane cenp-H/I/K/M). The conditions are those determined with psmd1. 2µg of the cenp-H/I/K/M complex is added to the reaction mix and incubated o/n with or without PIASy. Before incubation, four bands appear. The upper one is cenp-I (80kDa). Under cenp-I is the bands corresponding to cenp-H tagged with a sox-histidines tag and cenp-K that have almost the same the same molecular weight (33kDa and 35kDa). The intense band at 18kDa is SUMO2 ^{K5,7,11R}. A small band is migrating higher than SUMO is cenp-M (20kDa). After o/n incubation, the pattern of bands is the same in the presence or not of PIASy. Bands are present higher than the unmodified cenp-I (around 95kDa, 115kDa, 130kDa and 150kDa) and a band with a low intensity are around 50kDa. The band of 50kDa could correspond to a SUMOylated form of cenp-H or cenp-K. The bands higher than unmodified cenp-I could correspond to SUMOylated forms of cenp-I. However, the intensities of these bands are low. The unmodified form of cenp-I is predominant even after o/n incubation. It means that the SUMOylation of cenp-I in this condition is very low. Because the pattern of bands are similar with or without PIASy, it suggests that PIASy does not act as E3 for the SUMOylation of cenp-I. The reaction was optimized for Psmd1C and the SUMOylation of cenp-I might require a different PIASy concentration.

d) <u>Titration of PIASy in the SUMOylation assay</u>



<u>Figure 27:</u> SUMOylation assays of the cenp-H/I/K/M complex using a titration of PIASy. The assays are performed with 150nM E1, 200nM E2 and 35µM SUMO2 ^{K5,711R} at 30°C o/n and stained by Coomassie blue.

To determine if this result is due to technical issues or to the inability of PIASy to act as E3, a titration of PIASy was done (Figure 27). The intensity of the bands corresponding to SUMOylated cenp-I does not vary in the PIASy titration after an o/n reaction (Figure 27.B). The PIASy concentration does not affect the SUMOylation rate even after a 2h reaction (data not shown). The PIASy concentration does not affect the SUMOylation of cenp-I. Thus, we conclude that PIASy is not the E3 SUMO ligase of the cenp-H/I/K/M complex that raises the question of which E3 SUMO ligase is involved in the cenp-H/I/K/M complex SUMOylation?

e) The SUMOylation of cenp-I is independent of RanBP2

In order to determine the E3 SUMO ligase is involved in the cenp-H/I/K/M complex SUMOylation, I tried to produce and purify several proteins known as E3 SUMO ligases such as Pc2, PIAS1 and PIAS2 but they were all expressed in an insoluble state (data not shown). Only a fragment of RanBP, a nucleoporin described as an E3 SUMO ligase in the SUMOylation of RanGAP1, a GTPase-activating protein for small GTPase Ran (Pichler et al. 2002), could be produced. In my experiments, I use this construct of RanBP2, named BP2E3 fragment, which harbors E3 SUMO ligase activity.

Following the previous conditions, a cenp-H/I/K/M SUMOylation assay was performed using increasing concentrations of BP2E3. The samples are analyzed by SDS-PAGE and stained by Blue Coomassie (Figure 28). After the reaction, a band appears at a higher MW (50kDa) than cenp-H/K (35kDa) and a ladder pattern appears higher than BP2E3 and cenp-I proteins MW. The intensity of SUMOylated smears increases, the intensity of SUMO2 bands decreases in parallel of the addition of BP2E3 and the intensity of the cenp-I bands remains the same. The BP2E3 fragment has no effect on the final amount of SUMOylated cenp-I (Figure 28). The SUMOylated smears are due to the multi-SUMOylation of BP2E3. The part of RanBP2 that harbors E3 SUMO ligase activity has no E3 activity on the SUMOylation of cenp-I.

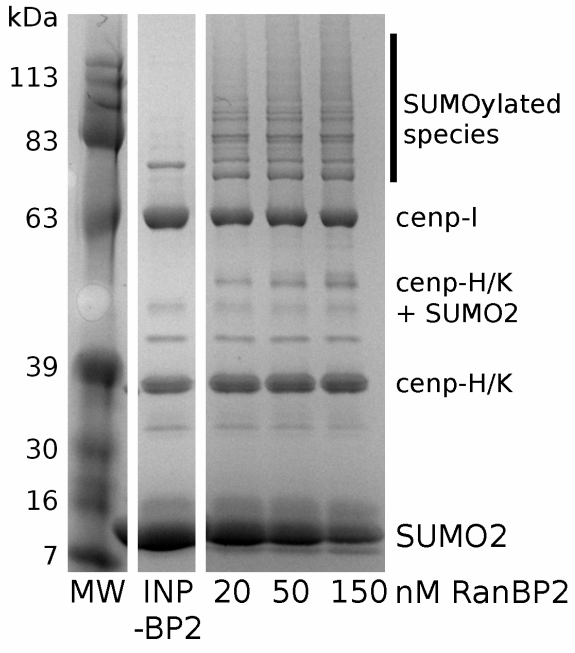


Figure 28: SUMOylation assay of the cenp-H/I/K/M complex using RanBP2 as E3 SUMO ligase incubated o/n.

INP-BP2 : Input before the add of the RanBP2 fragment . The SDS-PAGE is stained by Coomassie blue.

f) Titration of E2 for the SUMOylation of the cenp-H/I/K/M complex

The specific E3 SUMO ligase involved in the SUMOylation of the CENP-H/I/K/M complex remains unknown. The role of the E3 enzyme is to increase the SUMOylation rate, specificity, or both. Thus, it is perhaps possible to mimic *in vitro* some effects of the E3 SUMO ligase by increasing the concentration of SUMO enzymes. To this end, the E2 titration was performed to identify a condition where the SUMOylation rate is increased (Figure 28.B).

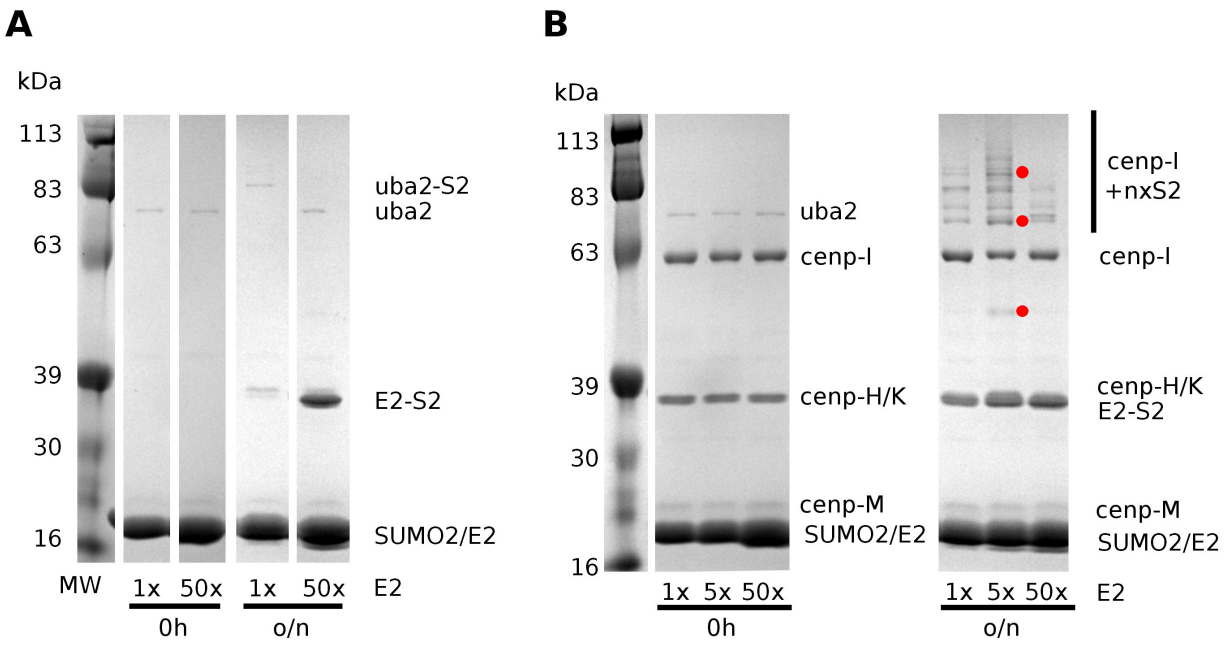


Figure 29: SUMOylation assays of the CENP-H/I/K/M complex analyzed by SDS-PAGE stained by Coomassie blue.

A) SUMOylation assays without substrate incubated o/n at 30°C. B) Titration of E2 for SUMOylation assays incubated o/n at 30°C.

To confirm the inability of the SUMO2 ^{K5,7,11R} to form poly-SUMO chains, the reaction is performed without E3 and substrate (Figure 28.A). Two concentrations of E2 have been tested, one similar to the previous assays (lane 1x) and one with a fifty fold increase of E2 concentration (lane 50x). At t=0h, there are two bands visible, one very intense at 18kDa and another at 75kDa. The SUMO2 ^{K5,7,11R} and E2 have similar MW and thus both migrate around 18kDa using SDS-PAGE. Also, the 18kDa-band is a mix of SUMO ^{K5,7,11R} and E2, while the 75kDa-band is Uba2. After an o/n incubation, without increased E2 (lane 1x, panel o/n), three bands appeared at 18kDa, 38kDa and 90kDa that correspond respectively to the mix of mono-SUMO2 ^{K5,7,11R} and E2, to di-SUMO2 ^{K5,7,11R} peptide or SUMO2 ^{K5,7,11R} linked to E2, and to Uba2 linked to SUMO ^{K5,7,11R}. The E1 recruits the SUMO2 through covalent link with Uba2 but is not able to transfer SUMO2 to E2. In this condition, there is no formation of poly-SUMO chains.

After an o/n incubation, in the reaction with increased E2 (lane 50x, panel o/n), three bands are revealed at 18kDa, 38kDa and 75kDa corresponding respectively to the mix of mono-SUMO2 ^{K5,7,11R} and E2, to di-SUMO2 ^{K5,7,11R} peptide or SUMO2 ^{K5,7,11R} linked to E2, and to Uba2. Compared to the condition lacking E2 increase, the band intensity is stronger and this difference

could be explained by the enhancement of E2 coupled to SUMO2 ^{K5,7,11R}. Since E2 is not able to transfer the linked SUMO2 ^{K5,7,11R} to the substrate, this mutant is not able to form poly-SUMO chain, even if the enzymes concentration is largely increased.

In the Figure 28.B, a titration of E2 was performed. Three concentrations of E2 were tested 1) the E2 concentration determined with Psmd1, 2) a five fold increase of E2 concentration and 3) a fifty fold of E2 concentration. The reaction mix is incubated o/n at 30˚C. Before incubation (panel 0h), several bands are detected. Specifically, the bands at 23kDa, 60kDa and at 75kDa correspond respectively to CENP-M, CENP-I and Uba2. The band at 18kDa corresponds to both SUMO2 ^{K5,7,11R} and E2 since they show similar MW. The 38kDa-band corresponds to CENP-K and CENP-H tagged with six-histidines since they migrate at the same MW.

After o/n incubation (panel o/n), new bands appear at 50kDa at MW higher than 63kDa. The band at 50kDa corresponds to SUMOylated CENP-H or CENP-K. The pattern of bands at high MW corresponds to SUMOylated state of CENP-I and forms a ladder characteristic to several multi-SUMOylation states of a same substrate. In the sample with a five-fold increase of E2 concentration, the bands intensities of the corresponding to SUMOylated substrates are stronger, while the intensity of the unmodified CENP-I band is lighter. These results show that the five-folds increase of the E2 concentration (Figure 28.B.lane 5x) greatly increases the amount of mono and multi-SUMOylated CENP-I and mono-SUMOylated CENP-H/K compared to the standard concentration of E2. However, when E2 concentration is increased fifty fold (Figure 28.B.lane 50x), the amount of SUMOylated CENP-I is lower than other tested conditions. Thus, E2 could act as a dominant negative protein when it is not bound to SUMO.

In the sample with a five-fold increase of E2 concentration (Figure 28.B.lane 5x), the amount of unmodified CENP-I is significantly reduced compared to the condition previously used (Figure 28.B.lane 1x). That suggests that we obtained a sufficient amount of SUMOylated CENP-I for identification of the SUMOylation sites using mass spectrometry analysis.

g) Analysis of the samples by mass spectrometry

The bands marked with a red dot of the Figure 28 were cut and sent to our collaborator Dr. Robinson for Mass Spectrometry analysis. The protein extraction and digestion as well as the

peptide analysis were performed by the Mass Spectrometry and Proteomics Resource laboratory of Harvard University, according to their protocol.

After analysis, no cenp-I peptides and only few cenp-H peptides were identified. This absence of identified peptide could be a consequence of the cenp-I SUMOylation. Indeed, this method has two limits, the first is the low efficiency of the protein extraction from the gel and the second is the low ionization of big peptide fragment.

The gel extraction of SUMOylated proteins has low efficiency. The branches created by the SUMOylation process shift the substrate from a two-dimension (a peptide wire) to three-dimensions, leading to their retention into the gel. This problem mainly occurs with SUMO and not with ubiquitin. After trypsin digestion, the ubiquitin only adds a di-glycin motif while the SUMO adds 32 residues (FDGQPINETDTPAQLEMEDEDTIDVFQQQTGG). The MS ionizes more readily the small peptide and is more sensitive to detect small peptides. Again the di-glycine motif of the ubiquitination is convenient in m with the 32 residues tag of the SUMO. A team designed SUMO compatible with MS identification. Even if their approach is interesting and promising, it unfortunately leads to modification of the SUMOylation kinetic and asks about the relevance of the identified residues (Tammsalu et al. 2014).

To circumvent this gel extraction issue, the SUMOylation assay of the cenp-H/I/K/M complex was repeated, the liquid samples were frozen in liquid nitrogen just after reaction and sent to the collaborator. The sample was digested by trypsin, the product of digestion is precipitated by ammonium sulfate and then analyzed by MS. The results of protein identification by MS analysis are shown in the Table 6.

Protein	coverage	SUMOylation sites	Sites of SUMOylation
cenp-H	89%	5	K58, 66, 75, 80, 182
cenp-I	21%	0	-
SUMO2	84%	1	K21

Table 6: Summary of the mass spectrometry analysis.

Compared to the previous experiments, parts of the protein sequence was recovered by MS. They identified respectively 84%, 89% and 21% percent of SUMO2 ^{K5,7,11R}, cenp-H and cenp-I. We consider the identification coverage of SUMO2 mutant and cenp-H good enough to detect potential SUMOylation sites. One SUMOylation site was identified on SUMO ^{K5,7,11R}, but according to the previous results, I assume this is negligible and that the SUMO mutant unable to form a poly-SUMO chain. Five sites of SUMOylation were determined on cenp-H. Unfortunately, the coverage of cenp-I is too low and no SUMOylation sites could be identified (Table 6). The protein sequences and the peptides identified are shown in the Supplementary part.

The identification of SUMOylated lysines of cenp-I could be performed by systematic mutations of the 58 lysines of cenp-I and determine which one affect the SUMOylation state, which is a long and challenging project. We could also use the SUMO designed for MS identification despite artifacts. This SUMO puts a lysine right before the di-glycine C-terminal tail. After trypsin digestion, a di-glycine motif is exhibited as ubiquitin potentially fixing the gel extraction troubles and eventually increasing the ionization of the peptides. This method leads to a decreased rate of SUMOylation and the disruption of the substrate recognition (Tammsalu et al. 2014). Thus, these artifacts coupled to the absence of a specific E3 in my *in vitro* assays lead to results that could be far from the endogenous cellular pathways.

My work was focused on the poly-SUMOylation of cenp-I but several kinetochore proteins are SUMOylated. The studies of these SUMOylations on the kinetochore architecture and functions could be performed more easily *in vitro*. For the next part of my Thesis, I moved from the SUMOylation of cenp-I to the reconstitution of the kinetochore *in vitro* to study its SUMOylation.

III The *in vitro* kinetochores

To study the kinetochore, the purification of the chromatin is possible but the centromere represents only 2% of the chromosome sequence. After purification, the chromatin proteins are overepresented while the kinetochore proteins have low abundance. Teams tried to reconstitute or purified kinetochore *in vitro*. The SUMOylation assay on *in vitro* kinetochore is the best approach to mimic the endogenous condition.

1) The *in vitro* reconstitution of kinetochore using cenp-A nucleosomes

A protocol describing the reconstitution the reconstitution of *in vitro* kinetochore developed by the Dr. Straigh's team was published when I joined the team (Supplementary part). In this protocol, they recreated *in vitro* centromere nucleosomes linked to magnetic beads and mixed them with Xenopus Egg Extracts (XEE) to recruit kinetochore proteins. In theory, this protocol is able to reconstitute both centromere and kinetochore.

a) A published protocol to reconstitute cenp-A nucleosomes on beads

In collaboration with another team member, Shaofei Zhang, we reconstitute kinetochores *in vitro* according to he protocol from Dr. Guse's team to perform biochemical assays including SUMOylation assays. To reconstitute kinetochore *in vitro*, kinetochore proteins from xenopus egg extract are recruited onto cenp-A nucleosomes linked magnetic beads. The main steps of this protocol are listed in the Figure 30. Two sets of nucleosome proteins, one with Histone3 and one with cenp-A are produced and purified (Figure 30.1). The DNA template used for nucleosomes reconstitution is a repeat of the Widon 601 sequence, that is the DNA sequence with the highest affinity for cenp-A (Lowary and Widom 1998). Once purified, the DNA template is biotinylated (Figure 30.2), mixed with the sets of nucleosome proteins to reconstitute the nucleosomes (Figure 30.3) and targeted to the beads to form the nucleosome beads (Figure 30.4). These nucleosome beads are incubated into Xenopus Egg Extracts to recruit kinetochore proteins (Figure 30.5). The beads are quickly washed and prepared for WB analysis or immunofluorescence (Figure 30.6). The cenp-A nucleosomes beads are compared to the Histone 3 nucleosomes that are the canonical nucleosomes, mostly present in the chromosome arms recruiting proteins that are not specific to the centromere.

This system provides a major advantage: the manipulation of the composition of the chromatin template and the protein composition of the xenopus egg extract through immunodepletion to study the assembly of kinetochores. A second advantage is the release of the extract from metaphase into interphase with the addition of calcium. Indeed, by cycling the extract through multiple rounds of interphase and metaphase, this technique could allow the study of cell cycle-associated changes to centromere and kinetochore components.

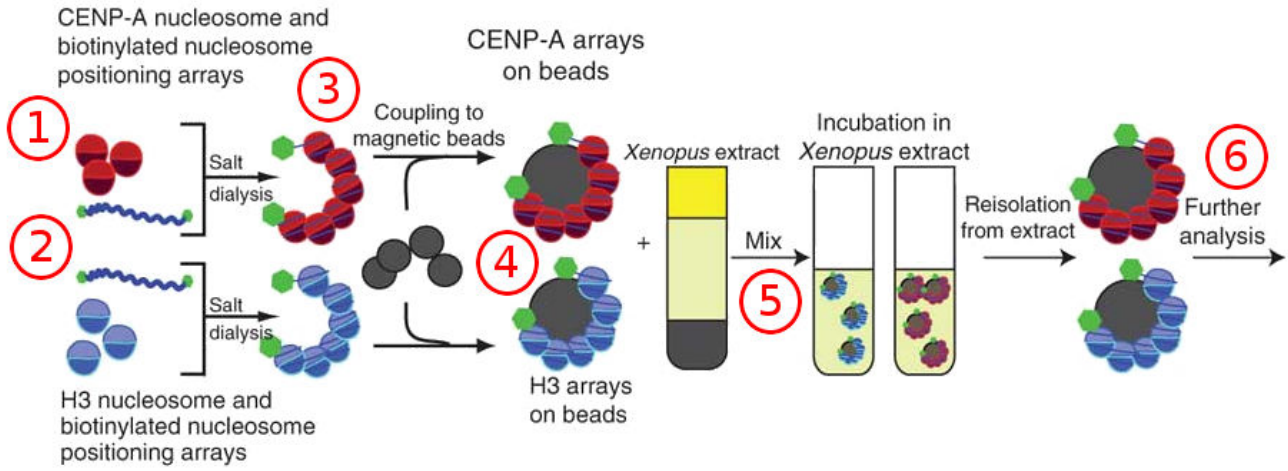


Figure 30: The protocol of production of nucleosome beads.
1) Purificartion of histone proteins. 2) Purification and biotinylation of DNA templates. 3) Reconstitution of nucleosomes on biotinylated DNA template. 4) Targeting of the nuclesomes on magnetic beads. 5) Mix of the nucleosome beads into Xenopus Egg Extract. 6) Analysis
Adapted from (A. Guse, Fuller, and Straight 2012)

b) Recruitment and release of centromere proteins

Shaofei Zhang used first the nucleosomes beads and was more focusing on centromere issues. While we performed the experiments to create the nucleosome beads together, he did the analysis. By western blot analysis, he tested on cenp-A nucleosome beads the presence of the centromere proteins AuroraB, a master key kinase localized in the centromere, and sgo1, a protein protecting the cohesin complex from premature cleavage by separase (Figure 31). 4µL of cenp-A nucleosomes beads are loaded in each lane. Bands appear at the expected molecular weight without contamination meaning that the antibodies seem specific. He was able to recruit auroraB and sgo1 (Figure 31).

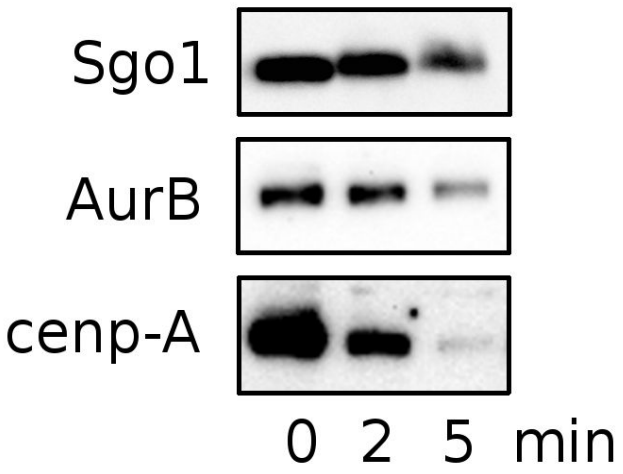


Figure 31: Western Blots of recruitment of centromere and kinetochore proteins by cenp-A nucleosome beads.
The antibodies used are written in front of the panels.

In the published protocol, the beads have to be quickly washed and analyzed, or fixed by formaldehyde. Both of us were interested in biochemical assays requiring hours of incubation, thus the stability of the kinetochore was tested. After the first wash, the cenp-A beads are incubated in several buffers. While the intensities of band are strong in the sample without incubation (Figure 31.lanes 0), the intensities decrease over the time and are almost undetectable (even cenp-A !) after only 5 min of incubation (Figure 30.lane 5). cenp-A nucleosomes is the root of the kinetochore, that recruit cenp-C and cenp-N, both mandatory for the kinetochore architecture. Removal of cenp-A from beads leads to removal of cenp-C and cenp-N and disruption of the kinetochore-like structure. After several trials, we concluded that with this protocol we can produce with unstable kinetochores which are not suitable for biochemical assays, a conclusion that was implicitly suggested in the initial protocol.

c) The kinetochore recruited are not stable enough for biochemical assays

This method has another limitations. The amount of cenp-C is low compared to the amount of cenp-A (Annika Guse et al. 2011). It suggests that the affinity between cenp-A and cenp-C is not that strong. It could be due to the absence of mandatory post translation modifications that strengthens the interaction or due to the lack of centromere partner. Moreover, the difference of

band intensities using Western Blot between control (histone 3 nucleosomes) and experiment (cenp-A nucleosomes) is not obvious for several proteins such as tubulin (A. Guse, Fuller, and Straight 2012). It suggests that a part of the recruitment is non-specific. These results were repeated by several teams, including the Straigh's one who also faced the release of cenp-A. According to their results, the interaction between cenp-C and cenp-A nucleosomes is highly unstable. The team of Dr. Straigh is currently trying to reveal the mechanism of the cenp-A eviction.

The reasons of these limitations are still unknown but several hypotheses are possible. First, the removal of cenp-A could be due to duplication of the DNA array. Indeed, after fecundation of XEE, the cell cycle goes through the duplication and the mitosis phases, in which the G1 and G2 steps are almost nonexistent. Thus, when the incubation with XEE is too long or during the washing steps, some proteins involved in the duplication could release cenp-A from the DNA array. Second, the centromere and kinetochore structures are extremely complex; some post-translation modifications that are mandatory for the kinetochore functions could be absent, especially the essential PTM of the histones (Bannister and Kouzarides 2011). Third, the gradients of proteins or biomolecules such as GTP are important for centromere or kinetochore regulation (Clarke and Zhang 2008; Carmena et al. 2012). The set up of this experiment could be not able to restore or mimic this gradient and its downstream consequences. Fourth, in these experiments the cenp-A nucleosomes provided could not reconstitute or mimic an endogenous centromere. Despite all these limitations, this protocol is a breakthrough to recreate kinetochores *in vitro* and is the basis of the optimal assay for *in vitro* kinetochores.

2) Reconstitution *in vitro* of kinetochore by using cenp-T as anchor :

Several teams tried to determine which proteins initiate the kinetochore reconstitution. A very interesting approach was performed by the teams of Dr. Cheeseman and Dr. Fukagawa. In their studies, they targeted kinetochore proteins on the chromosome arms and checked their ability to reconstitute a functional ectopic kinetochore using the LacO/LacI system. The LacI/LacO system is used to target protein onto DNA. The fusion of proteins to LacI localizes them on the LacO sequence that can be placed at exogenous loci. Fukagawa's team first generated a Z chromosome in chicken depleted of its centromere but providing LacO arrays in its arms (T. Hori et al. 2012). Then, by expressing the kinetochore proteins fused to LacI, they screened proteins

able to rescue the loss of the centromere by forming ectopic kinetochores. cenp-T is one of the protein able to rescue the loss of the centromere. The Cheeseman team performed a similar approach in human cell lines by inserting LacO arrays on chromosome arms but they do not deplete the centromere and only checked the ability of a particular protein to recruit kinetochore proteins (Karen E. Gascoigne et al. 2011). They demonstrated that cenp-T and cenp-C recruit the outer kinetochore and some proteins of the inner kinetochore.

My objective is to adapt *in vitro* the approach of cenp-T targeted by LacI. Using this method, we could recruit kinetochore and centromere proteins through cenp-T while the DNA template could stabilize the kinetochore structure through binding of proteins such as cenp-N or cenp-C and then cenp-A. The Straigh's approach leads to the release of cenp-A from nucleosome beads while the LacI/LacO system provides a more stable protein/DNA interaction (Figure 32). Moreover, the LacI/LacO system is simpler and fast to produce compared to the production of cenp-A nucleosomes. The LacO DNA arrays is tightly linked to the magnetic beads through the high affinity biotin- streptavidin interaction.

To recruit cenp-T to magnetic beads, a biotinylated template containing repeats of the DNA sequence LacO is targeted to streptavidin magnetic beads, while the histone-like domain of cenp-T is replaced by LacI. The fusion of LacI to cenp-T allows the binding to the LacO arrays beads and the kinetochore recruitment through cenp-T. The LacI binding to LacO arrays beads was used as a negative control to specifically identify which proteins are recruited by cenp-T and not by the DNA arrays or by LacI.

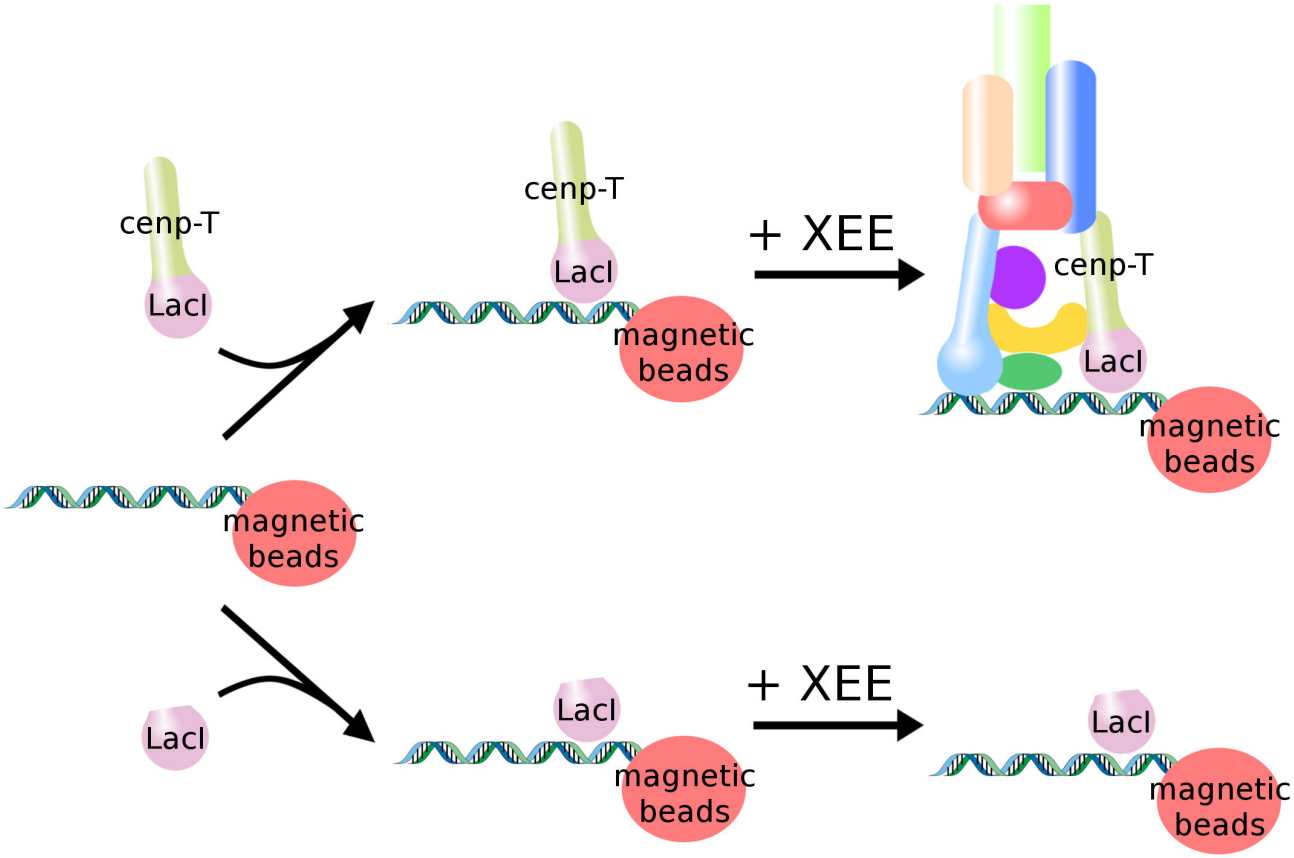


Figure 32: Outline of cenp-T/LacI recombinant protein used for kinetochore recruitment. The LacI based proteins are targeted to the LacO arrays beads. The LacI/LacO beads are mixed with Xenopus Egg Extract. The cenp-T/LacI/LacO beads recruit kinetochore proteins.

a) DNA array preparation

To generate the DNA arrays, the protocol is described in the Material and Method chapter. The DNA template is generated in bacterial cultures, purified, extracted from the template and purified, biotinylated and finally linked to streptavidin magnetic beads. The main steps are shown in the Figure 33 and detailed in the Material and Methods chapter.

The digestion of the plasmid pVS1 leads to the release of a 256 repeats of LacO DNA sequence (Figure 33.B.LacOx256). The 256 repeats of LacO is purified by Poly-Ethylene Glycol (PEG) 3350 precipitation (Figure 33.C), by dialysis and by ethanol precipitation, avoiding the presence of PEG to interfere with the biotinylation reaction, and allows its resuspension in the TE buffer at a concentration of 3µg/µL.

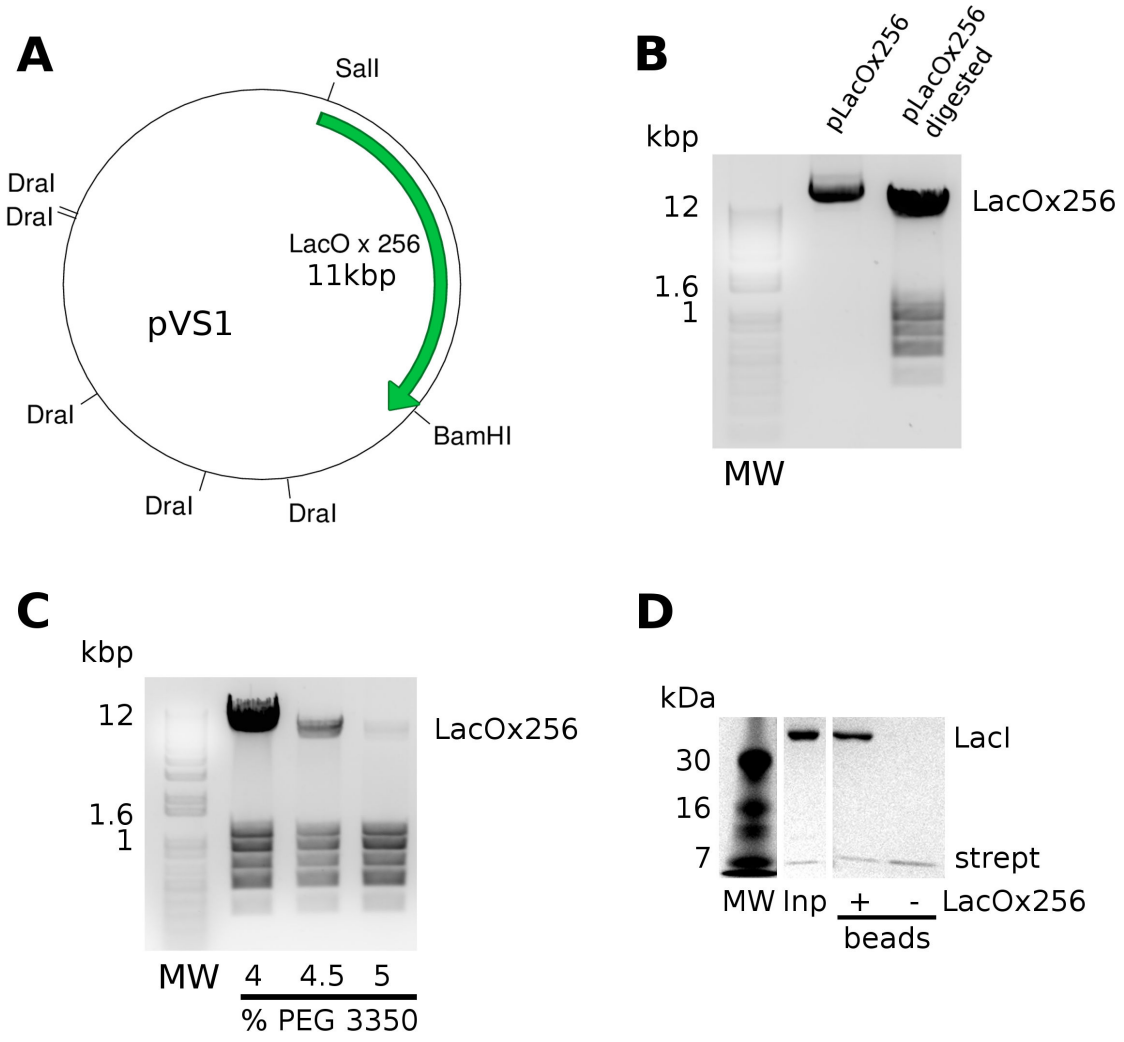


Figure 33: Preparation of the LacOx256 magnetic beads.
A) Plasmid containing the 256 repeats of LacO (pLacO256). B) The results of digestion of the plasmid pLacO256 by the restriction enzymes DraI, BamHI, EcoRV and SalI from NEB. C) The purification of the LacO repeats by PEG precipitation. D) The biotinylation of LacO array is confirmed by the ability of LacO array beads to bind LacI. Beads are incubated with LacI, washed and then boiled in SDS-PAGE Loading Buffer and analyzed by SDS-PAGE

To biotinylate the LacO array, I first follow the protocol used by Dr. Guse who uses the Klenow fragment and a biotin-14-dATP to fulfill the non-blunt end of the LacO and insert the biotin (Annika Guse et al. 2011). While this protocol works perfectly for the biotinylation of the DNA array composed of the 19 repeats of the 601 sequence, the efficiency of the reaction greatly varies with the LacO arrays. In order, to fix this, I use the DNA polymerase from *Thermus aquaticus* (Taq) (#M0273, NEB) that fills the non-blunt ends and adds an extra adenosine at the 3'

termini. The use of biotin-14-dATP leads to the biotin tagging of the LacO arrays. This approach is less expensive and leads to a complete and highly reproducible tagging of the 256 repeats of LacO. The biotinylated arrays are purified by two ethanol precipitations, resuspended, dialyzed with TE buffer o/n at 4˚C and submitted to a second ethanol precipitation. The LacO arrays are resuspended in TE buffer at a final concentration of 200ng/µL to be linked to the magnetic beads as described in the Material and Methods.

b) Production and purification of LacI

The protocol to produce and purify LacI is described in the Method part. The LacO binding ability of the purified LacI is tested by incubation with LacO beads and the results are shown in the Figure 34.D. A band corresponding of LacI appears in the sample with beads linked to LacO arrays (lane +) and not in beads without LacO arrays (lane -). LacI is recruited to these beads through the LacO arrays. The streptavidin of the magnetic beads appears at an approximate molecular weight of 10kDa. The intensity of the streptavidin bands is the best way to quantify the amount of beads loaded in each lane. The intensity of the streptavidin band is drastically lower than the LacI band. The company does not explain what is the streptavidin construct and how they are linked to the magnetic beads. For this reason, I assume that their streptavidin forms a quadrimer as the wild type and that only one subunit is linked to beads. Each recruited DNA array has 256 LacO repeats explaining the difference of intensity between the streptavidin and LacI bands.

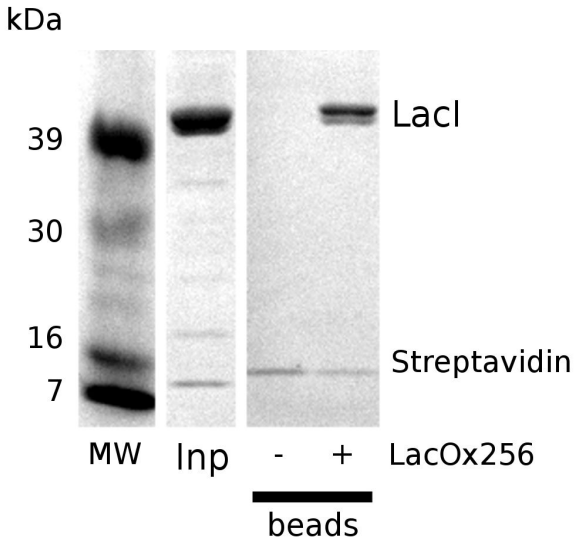


Figure 34: Test of recruitment of LacI on LacO array beads.
The SDS-PAGE is stained by Coomassie blue. Pure LacI is incubated at RT for 15min with magnetic beads coupled to streptavidin with (+) or without (-) LacO arrays. The beads are washed three times, resuspended in SDS-PAGE Loading Buffer and boiled.

c) Production of cenp-T/LacI

The production of cenp-T/LacI/LacO magnetic beads that are suitable for biochemistry assays involves the absence of contaminants and artifacts such as non-specific recruitment of partners or steric clash due to inadequate construct. Therefore, the quality and the quantity of pure cenp-T/LacI on LacO array beads are important to ensure the proper functionality of cenp-T to recruit the kinetochore. The fusion protein must be able to specifically bind LacO through its LacO part and remain the function of the cenp-T part.

To design the protein resulting of the fusion of LacI and cenp-T, I replaced the histone-like core of cenp-T that targets the protein to the centromere by LacI. In order to preserve the LacI and cenp-T protein integrities, no recombinant tag was added and the linker was designed as a short peptide (KLMVN). The cenp-T fragment used (from amino acid 1 to 380) was those previously used by the Dr. Cheeseman's team (Karen E. Gascoigne et al. 2011) (Figure 35.A).

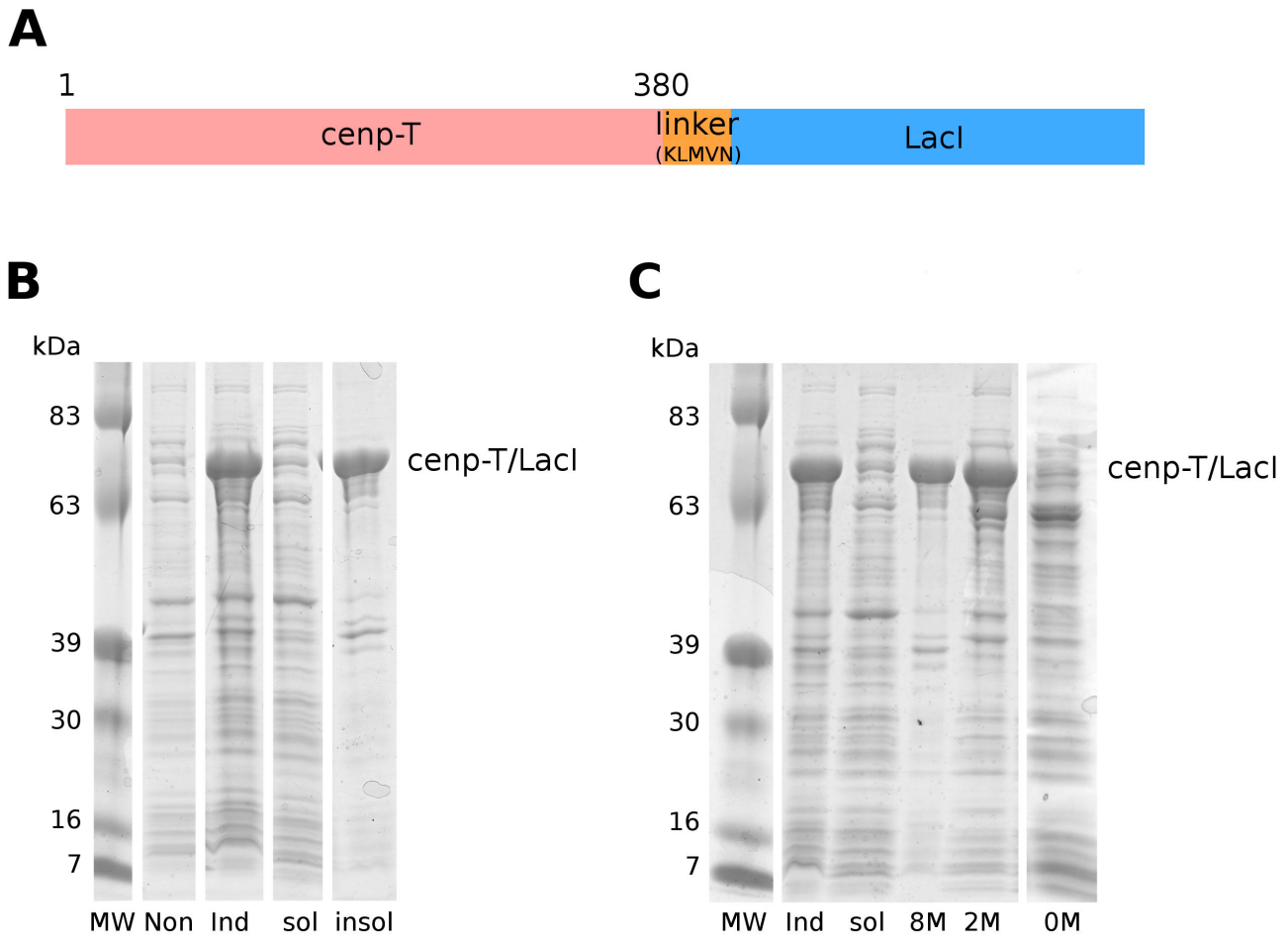


Figure 35: Production test of cenp-T/LacI.
A) Scheme of the cenp-T/LacI construct. B) Expression and solubility tests of cenp-T/LacI in bacterial system. C) Refolding tests of cenp-T/LacI. Ind: induced. Sol: soluble. 8M: urea 8M. 2M: urea 2M+Tris-HCl 20mM pH8. 0M: PBS.

The coding DNA sequence (CDS) of LacI is amplified by PCR from pET28 and the short linker is added by the primers during this step. The CDS is inserted between Hind3 and Not1 restriction sites and then, the cenp-T peptide is added using the same sites. The theoretical molecular weight of the fusion protein cenp-T/LacI is 82.115Da. The product of cloning is sequenced and transformed in *E. coli* (BL21) DE3 RIPL. The results of the expression and purification tests are shown in the Figure 35. An expression test was performed at an temperature induction of 37°C as described in the Material and Methods chapter (Figure 35.B). After lysis, the soluble fraction (lane sol) and the pellets (lane insol) are also analyzed. Comparion of the non-induced and the induced lanes reveals specific bands after induction: a smear of

proteins with various molecular weights are detected in the non induced lane, while in the induced lane, an intense band appears around 75kDa. WB analysis with antibodies against LacI confirms that this band contains LacI. Because LacI has been added at the C-terminal part of the fusion protein, it strongly suggests that the cenp-T part is fully translated (Figure 35.A.ind lane), and that the cenp-T/LacI fusion is expressed. Unfortunately, the intense band corresponding to cenp-T/LacI is not in the soluble fraction but is present in the pellet or the insoluble fraction.

Following the lysis of bacteria expressing the cenp-T/LacI protein, the fusion protein is mostly in the insoluble fraction (Figure 35.A.sol and insol lanes). Several parameters were tested in order to increase the production of the cenp-T/LacI fusion in the soluble fraction such as the temperature of expression, the bacterial strains during the induction step and the use of detergent during the lysis. Despite our efforts of optimization, the production of the fusion protein cenp-T/LacI remained in the insoluble fraction (data not shown).

d) <u>Resolubilization of the cenp-T/LacI protein</u>

I tried to resolubilize the cenp-T/LacI protein from the insoluble fraction to be able to perform a further purification steps. cenp-T/LacI is the most abundant protein in the insoluble fraction, but can be resolubilized in a buffer of 8M urea (Figure 36.B). All the further dialysis steps are performed o/n to allow slow but correct refolding and analyzed by SDS-PAGE. The sample is dialyzed o/n against urea 2M, Tris-HCl 20mM pH8.. After dialysis against 2M urea and centrifugation to remove insoluble molecules, an intense band corresponding to cenp-T/LacI is soluble in urea 2M (Figure 36.B). The urea 2M condition is used for further solubility experiments. After dialysis o/n at 4°C from urea 2M, Tris-HCl 20mM pH8, cenp-T/LacI is neither soluble in PBS (Figure 36.B), PBS + Tween 20 0.05% or Tris 20mM pH8, NaCl 0.2M. The guanidinium 6M buffer was also tested for the resolubilization process from cenp-T/LacI but it led to the same results (data not shown).

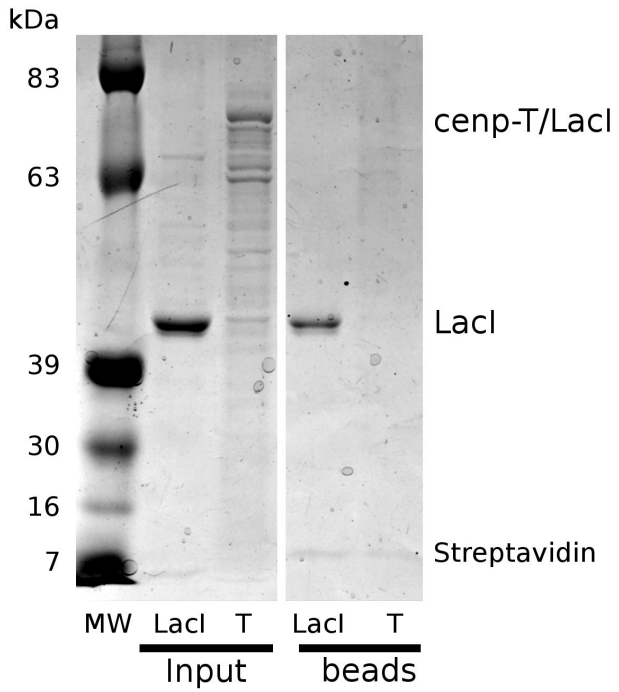


Figure 36: Recruitment test of resolubilized cenp-T/LacI.
Pure LacI and resolubilized cenp-T/LacI are incubated with LacO arrays beads. The beads are washed and boiled in SDS-PAGE loading buffer and 8µL are loaded per well. T: cenp-T/LacI. The SDS-PAGE is stained by Coomassie blue.

From the 2M urea initial buffer containing resolubilized cenp-T/LacI, quick dilution by direct adding of PBS were tried. Surprisingly, cenp-T/LacI does not precipitate in final urea 0.1M. Even after a 20.000 rcf centrifugation for 30min, some of the cenp-T/LacI proteins remain in the soluble fraction. A relatively pure cenp-T/LacI was obtained and tested the ability of the resolubilized cenp-T/LacI fusion proteins to bind LacO arrays beads. Using a protocol similar than for the LacI, the resolubilized cenp-T/LacI is not able to bind LacO arrays beads (Figure 36). I mixed 10µL of LacO arrays beads with 100µg of proteins of pure LacI or of the fraction containing resolubilized cenp-T/LacI (Figure 36.lane Input). There is only one band in the input of LacI. In the input of the cenp-T/LacI experiment, there is a smear of protein but full-length cenp-T/LacI is present in the sample. In the sample of beads, the band around 10kDa corresponds to the streptavidin. In the LacI sample, the band of LacI appears around 39kDa. However in the beads of cenp-T/LacI sample, there is no band detected. The LacI is detected on the beads, so the LacO arrays are linked to the beads and in this condition,LacI is able to bind LacO. But the absence of proteins bound to the beads in the cenp-T/LacI sample means that LacI domain is not properly

folded after the quick resolubilization. Probably, after quick resolubilization, the proteins form soluble aggregates that are not pelleted by ultracentrifugation.

e) Recruitment of cenp-T/LacI onto LacO beads from soluble fraction

I was not able to resolubilize a functional cenp-T/LacI. However, according to the SDS-PAGE, I cannot determine if there is any low concentration of soluble and functional cenp-T/LacI in the soluble fraction after lysis. To test the presence of functional cenp-T/LacI after lysis, I performed a recruitment test on LacO arrays beads using two samples, of LacI (positive control) and the cenp-T/LacI fusion. For this second, the soluble fraction from lysis of bacteria expressing cenp-T/LacI is incubated for 30min in PBS + EDTA 0.5mM at 21˚C with 10µL LacO arrays beads. The beads are washed three times with PBS and are then resuspended in SDS-PAGE loading buffer, boiled at 95˚C for 5min, analyzed by SDS-PAGE. The results are shown in the Figure 37. In the LacI samples, there is one main band in the input that corresponds to LacI. On the boiled beads with linked LacO arrays, there are two bands, one with a low intensity at a molecular weight around 10kDa and one more intense at 40kDa, corresponding to respectively streptavidin and LacI. There is only one band without LacO arrays, those corresponding to streptavidin. LacI is able to bind LacO arrays in these conditions.

In the input of the sample from lysate of bacteria expressing cenp-T/LacI, there is a smear of proteins that corresponding the endogenous proteins and the eventual exogenous soluble proteins. In the absence of LacO arrays, two bands proteins of approximate MW 25kDa and 10kDa are retained on the beads. With LacO arrays beads, two additional proteins respectively at 40kDa and 75kDa are pull-downed. The band at 10kDa corresponds to the streptavidin in each sample. The band at 25kDa is recruted by the beads even without LacO, its recruitment is non specific. The band at 40kDa corresponds to LacI. There are several bands around 75kDa. The main band at 75kDa could be cenp-T/LacI. To confirm that these bands are LacI and cenp-T/LacI, a WB with antibodies against LacI was performed (Figure 37.B). In the WB, bands appear at 40kDa in the boiled beads with LacO arrays (lane LacI/+). Several bands appears in the cenp-T/LacI with LacO arrays beads, from 75kDa to 40kDa. The band with the higher band has a molecular weight similar than cenp-T/LacI. This confirms that soluble cenp-T/LacI is present in the soluble fraction from lysate, even if its concentration is very low. However, the other bands stained by antibodies

contain LacI and correspond to degradation products. The band stained by Coomassie blue but not by WB are contaminants and not cenp-T/LacI with a LacI part degraded. Indeed, the degradation of the LacI part will disrupt the ability of the protein to bind LacO.

The incubation of LacO arrays beads with soluble fraction from lysate of bacteria expressing cenp-T/LacI leads to the specific recruitment of cenp-T/LacI to the LacO arrays beads. Although this demonstrates that there are soluble cenp-T/LacI remaining in the soluble fraction after lysis, the cenp-T might not be functional. In addition, using this direct approach leads to the recruitment of many contaminants.

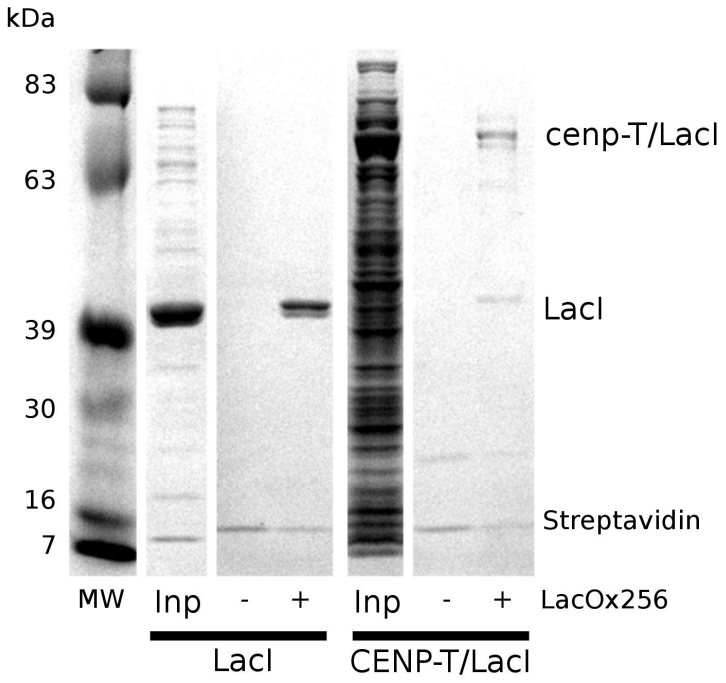


Figure 37: Recruitment of LacI and cenp-T/LacI on LacO arrays beads.
LacI and soluble fraction from lysate bacteria expressing cenp-T/LacI are incubated 15min at RT with magnetic beads linked (+) or not (-) to LacO arrays. Inp: input; -: absence of LacO arrays; +:presence of LacO arrays; The SDS-PAGE is stained by Coomassie blue.

f) Purification of cenp-T/LacI

In order to remove contaminants, I tried to purify cenp-T/LacI from the soluble fraction. Since no tag is inserted in the fusion cenp-T/LacI protein, I used a chromatographic system based on physico-chemical properties such as IEC or hydroxyapatite. After several trials, none of these

technics allowed me to purify cenp-T/LacI. Indeed, the fusion protein was not in the flow through or in the elution fractions, suggesting that cenp-T/LacI is highly unstable, precipitates and/or forms soluble aggregates that stick to the columns.

cenp-T/LacI was successfully purified by sulfate ammonium precipitation from soluble fraction after cell lysis. Unfortunately, cenp-T/LacI is highly instable after resuspension and precipitates entirely after 5min at 4˚C (data not shown). Several assays were performed to purify cenp-T/LacI using its ability to bind LacO but these tests failed due to protein aggregation.

To circumvent the inability to purify cenp-T/LacI using protein intrinsic properties, I tried the purification by IMAC-Ni^{2+}. cenp-T is the N-terminal part of the protein, so a 6 His-tag was inserted at the C-terminal tail of the fused protein just after the LacI part to avoid any artifact with the recruitment of kinetochore proteins by cenp-T. The production, lysis and purification by IMAC-Ni^{2+} are performed as described in Material and Methods. The results of the purification process are shown in the Figure 38.A. In the elution lane, several bands appear at 80kDa and 40kDa, corresponding respectively to cenp-T/LacI and to LacI. The proteins present in the eluate are highly unstable and precipitate after 5min on ice. The aggregates are pelleted and the sample is applied on a SEC to separate the full-length cenp-T/LacI to the degraded forms. The extinction coefficient at 280nm of the LacI part is 23045 $M^{-1}.cm^{-1}$ for a MW of 38,500 Da. The purification system absorbs no light at 280nm after the SEC column and no bands are detected by SDS-PAGE. These results suggest that all proteins precipitate in the chromatographic system.

The next obvious step was to test for the ability of cenp-T/LacI/6His to bind LacO arrays. After the purification of cenp-T/LacI/6xHis by IMAC-Ni^{2+}, the purified cenp-T/LacI/6His was incubated with LacO array beads as described in the standard protocol. The results are shown in the Figure 38.B. Pure LacI and the products of IMAC-Ni^{2+} are incubated with LacO arrays beads (lanes input). After three washes with PBS, the beads are boiled in 20µL of SDS-PAGE loading buffer for 5min and 8µL are loaded per well. In the LacI sample, a band at 40kDa corresponding to LacI is present both in the input and on the beads. The bands present in the input of 6xHis/cenp-T/LacI are not detected on the beads, meaning that the recombinant fusion protein is not recruited by the LacO arrays. As mentioned after the purification, the 6xHis/cenp-T/LacI is instable and it is possible that the proteins aggregate during the incubation time or that the LacI

part is not folded properly. Another reason could be the creation of steric hindrance due to the poly-His tag. The tag is added at the C-terminal tail of LacI, which is important for its dimerization. However, I designed the LacI protein as previously described by various structural biology teams and they did not report any change in the protein activities.

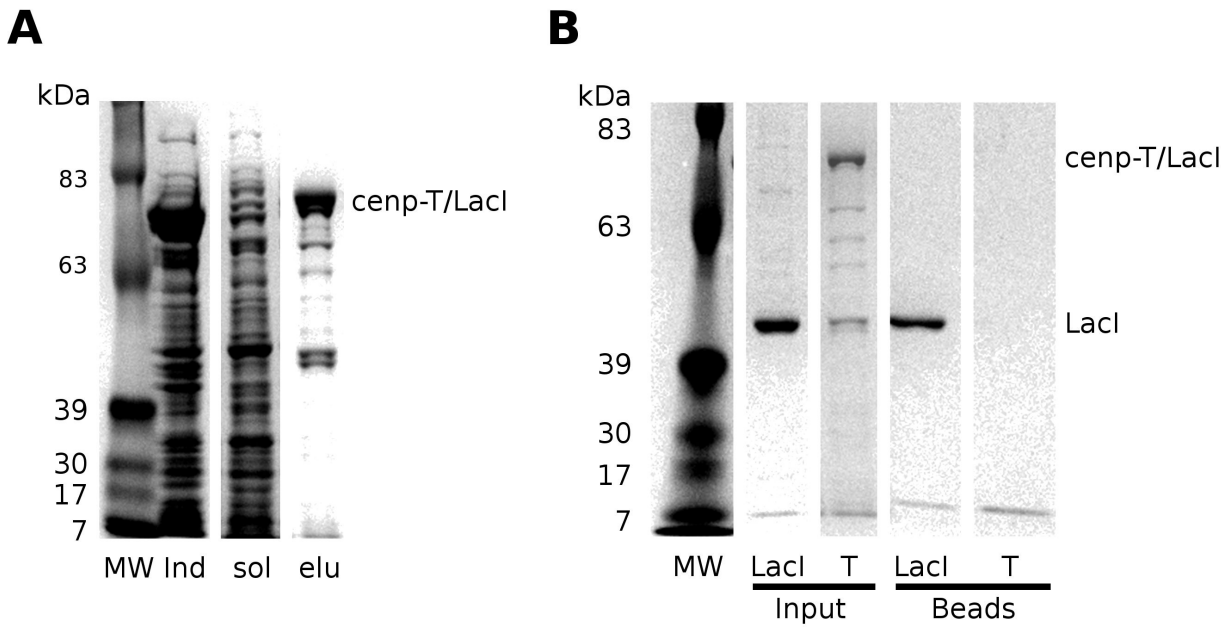


Figure 38: Purification and recruitment of cenp-T/LacI/6xHis on LacO array beads.
The SDS-PAGE are stained by Commassie blue. A) IMAC-Ni^{2+} purification of cenp-T/LacI. Ind: Induced. Sol: soluble fraction. Elu: elution. B) Recruitment of cenp-T/LacI on LacO array beads. T: cenp-T/LacI

g) Recruitment of kinetochore *in vitro* by cenp-T/LacI

Finally, to generate the cenp-T/LacI beads, I decided to skip the purification step and mix the LacO arrays beads with the soluble fraction of lysed bacteria expressing cenp-T/LacI as shown in the Figure 39. This method is the best compromise between the presence of contaminants and degraded forms of cenp-T/LacI, and the recruitment of the full-length cenp-T/LacI.

A volume of 20µL (50% v/v) of LacO beads with LacI or cenp-T/LacI is incubated with 100µL of XEE for 1h at 21˚C. The XEE are generated according to the standard protocol ("Preparation and Use of Xenopus Egg Extracts to Study DNA Replication and Chromatin Associated Proteins" 2014). The beads are quickly washed three times with the XEE buffer and resuspended in SDS-PAGE loading buffer. A volume of 13µL of resuspended beads is loaded on

SDS-PAGE for further silver staining and 8µL are used for WB analysis. The results are shown in the Figure 39.

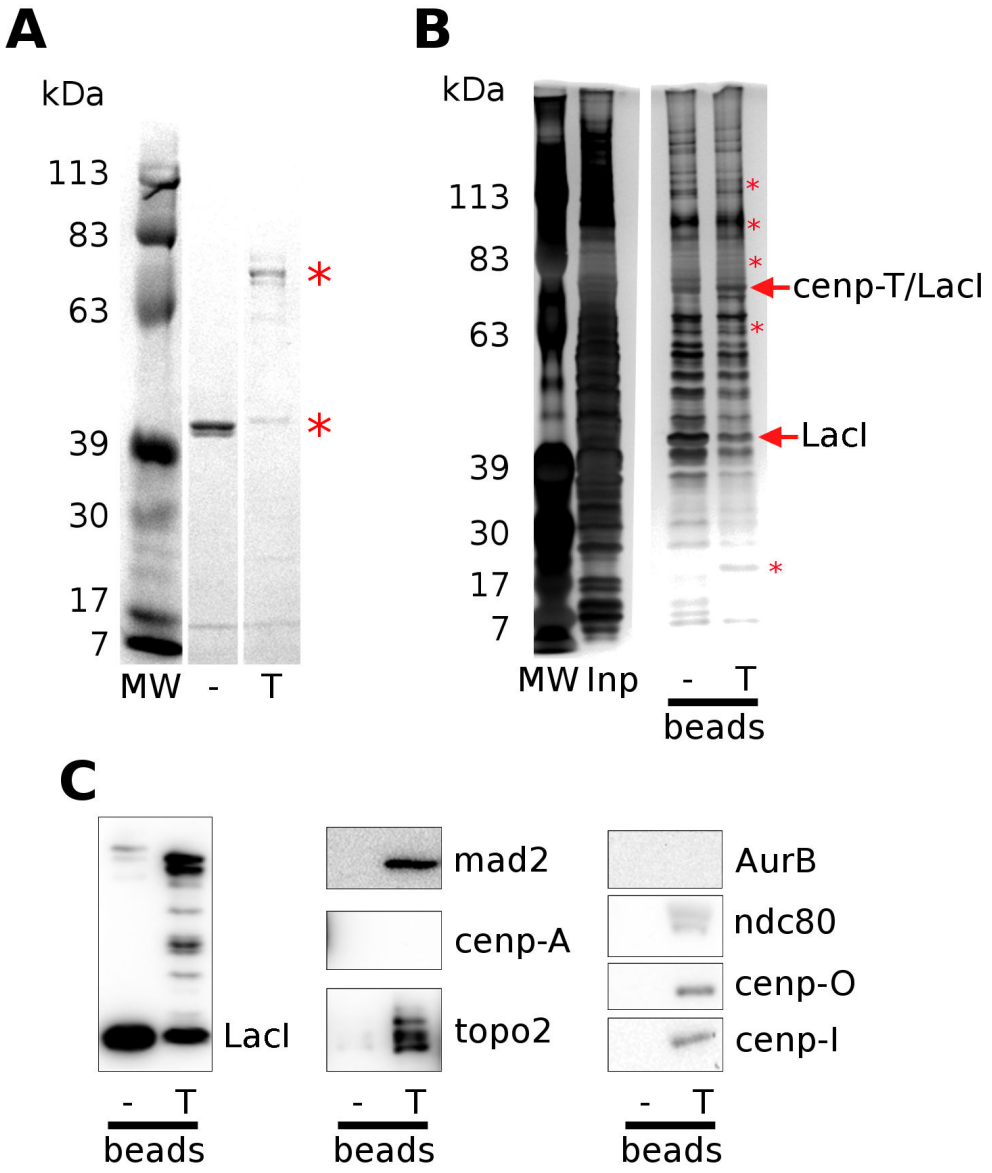


Figure 39: Test of kinetochore recruitment on cenp-T beads analyzed by SDS-PAGE.
A) The composition of beads before mixing with xenopus egg extract. A volume of 2µL of beads is loaded on gel stained by Coomassie blue. The star at 40kDa corresponds to LacI. The star at 75kDa corresponds to cenp-T/LacI. The SDS-PAGE is stained by Coomassie blue. B) The proteins bound to the beads after an 1h incubation with XEE. A volume of 13µL are loaded per well and stained by silver staining. Inp: XEE diluted ten times. The SDS-PAGE is silver stained. C) Western blot analysis of beads after a 1h incubation with XEE. The antibodies used are written on the side on each panel. A volume of 4µL are loaded per well and analyzed using various antibodies. -: LacI beads. T: cenp-T/LacI beads

The beads used for the kinetochore recruitment are analyzed by SDS-PAGE and then stained by Coomassie blue before the incubation (Figure 39.A). Two samples are used, the LacI bound to LacO arrays beads as a negative control (lane -) and the cenp-T/LacI bound to LacO arrays beads (lane T). Two bands are present in the LacI lane at 10kDa and 40kDa, which correspond respectively to streptavidin and LacI. Multiple bands are present in the cenp-T/LacI lane, one at 10kDa and several between 40kDa and 80kDa, corresponding to the streptavidin and the cenp-T/LacI proteins respectively. The smear of band reflects the degradation state of cenp-T/LacI as mentioned previously. However, it is to note that the full-length cenp-T/LacI is present in the experiment.

After incubation with XEE, the beads are analyzed by SDS-PAGE and further stained with silver nitrate (Figure 39.B). Because XEE is highly enriched in proteins, it is diluted ten times before analysis (Figure 39.B.lane Inp). The patterns of LacI/LacO beads (Figure 39.B.lane -) are similar to the cenp-T/LacI/LacO beads (Figure 39.B.lane T) except for several bands marked by an asterisk. In the cenp-T sample, the two bands around 75kDa that are absent in the LacI sample correspond to cenp-T/LacI. The presence of LacI and cenp-T/LacI is confirmed by WB (Figure 39.B.lane LacI). In the WB, antibodies against LacI target one band in the LacI sample and several in the cenp-T/LacI sample from 40kDa to 75kDa. The WB pattern on cenp-T/LacI with antibodies directed against LacI is quite similar to the beads before incubation (Figure 39.A). The cenp-T/LacI protein is still present on the beads after incubation. In conclusion, even if it is possible to detect a couple of bands specifically in the cenp-T/LacI sample, there is too much contamination to detect precise differences between both samples.

To determine the presence of proteins, WB analysis was used (Figure 39.C). A volume of 4µL of sample is loaded per well. In each panel, the samples with LacI or cenp-T/LacI are shown. Except for the detection of LacI and TopoII, the intensity of bands is low and I have to expose the membranes several minutes. It is too expensive and time consuming to detect the whole kinetochore by WB, that is why I focused on several specific proteins: ndc80, cenp-I, cenp-O, cenp-A, auroraB, topoII and mad2.

mad2 is a member of the spindle assembly checkpoint. The SAC is recruited by knl1 phosphorylated by the kinase mps1. The recruitment of the SAC suggests that knl1 is present and

phosphorylated by mps1. cenp-O is a member of the most downstream complex of the inner kinetochore. AuroraB is a key point kinase present at the core centromere.

The bands corresponding to the expected proteins are shown in the Figure 39. The pattern of cenp-T/LacI is similar the one previously observed, several forms are detected from the full-length cenp-T/LacI to a band with a molecular weight similar to LacI alone. The proteins ndc80, cenp-I, cenp-O, topoII and mad2 are detected only in the cenp-T/LacI sample. cenp-A, auroraB and mad1 (data not shown) are detected in none of these samples. However, according to the low intensity of the band, the amount of proteins is extremely low and hard to distinguish between the contaminant bands.

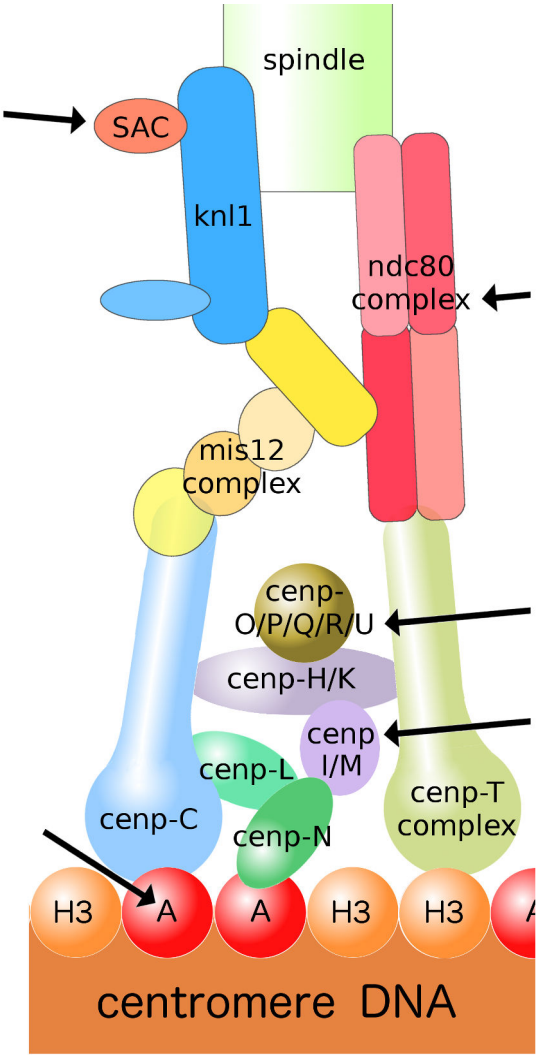


Figure 40: The kinetochore proteins tested by Western Blot analysis.
The arrows show the proteins tested.

h) Elution of kinetochore proteins

A lot of contaminants are bound to the LacO array beads. I attempted to separate the proteins loaded onto the DNA arrays to the beads and their contaminants using two approaches: remove the DNA arrays from beads or elute the proteins loaded on the DNA arrays. I chose to test the elution of proteins using a solution of glycine 0.2M pH 2.3, which disrupts biochemical interactions that are not due to hydrophobic interactions. It is important to verify that glycine (1) does not elute proteins bound to the magnetic beads without DNA arrays and (2) elutes proteins loaded on the DNA arrays.

To confirm that glycine does not elute proteins non-specifically bound to the beads. streptavidin magnetic beads are mixed with XEE for 30min at 21˚C, washed three times with PBS at RT and incubated for 10min at RT in 20µL four different buffers :

- PBS, the buffer used to wash the beads.
- urea 8M solution that will release all proteins on beads.
- urea 2M solution leading to partial release of proteins.
- glycine elution buffer that does not disrupt proteins binding based on hydrophobic interactions or aggregations.

The beads are pelleted and the elution buffer is analyzed by SDS-PAGE stained by Coomassie blue (Figure 41.A).

i) <u>Specific elution of proteins bound onto the DNA arrays</u>

For all buffers tested, the elution panel shows the protein eluted and the beads panel shows the protein bound on the beads after elution and wash. In the elution panel, there are no bands in the PBS lane. Bands with low intensity appears in the urea 2M lane, while an intense smear appears in the urea 8M. Urea solutions lead to elution of contaminants and therefore cannot be used to specifically elute proteins loaded onto DNA arrays. No bands appears in the glycine elution.

The second step is to determine whether the glycine elution buffer elutes proteins loaded onto LacO arrays. LacI/LacO array bound beads were prepared as indicated in the Method part. The LacI/LacO beads are incubated in elution buffer for 10min at RT. The elution is discarded and the beads are washed 3 times with PBS. The beads are then boiled in SDS-PAGE loading buffer for 5 min at 95˚C. The results are shown in the Figure 41.B.

In the PBS lane (negative control), almost no LacI is eluted and stayed on beads, as expected. As positive control, I used a competition elution by IPTG (lanes 2) to disrupt the LacI/LacO interaction. While I expect to detect a high intensity band in the elution, all LacI stays on beads, meaning that IPTG solution was not able to elute it. Sticking of LacI on beads due to aggregation remains possible. The DNA/protein interactions are mainly based on ionic or hydrogen interactions. For this reason, I tried a solution with IPTG with a higher salt

concentration (lane 3). In this condition, most of LacI is eluted and a little remains on LacO array beads. It appears that the addition of salt to IPTG is mandatory to disrupt the LacI/LacO interactions. The glycine elution buffer (lane 4) successfully elutes LacI from LacO DNA array beads. To analyze the recruitment of kinetochore proteins on magnetic beads, I therefore used a glycine elution buffer to specifically elute proteins of interest without most of the contaminants.

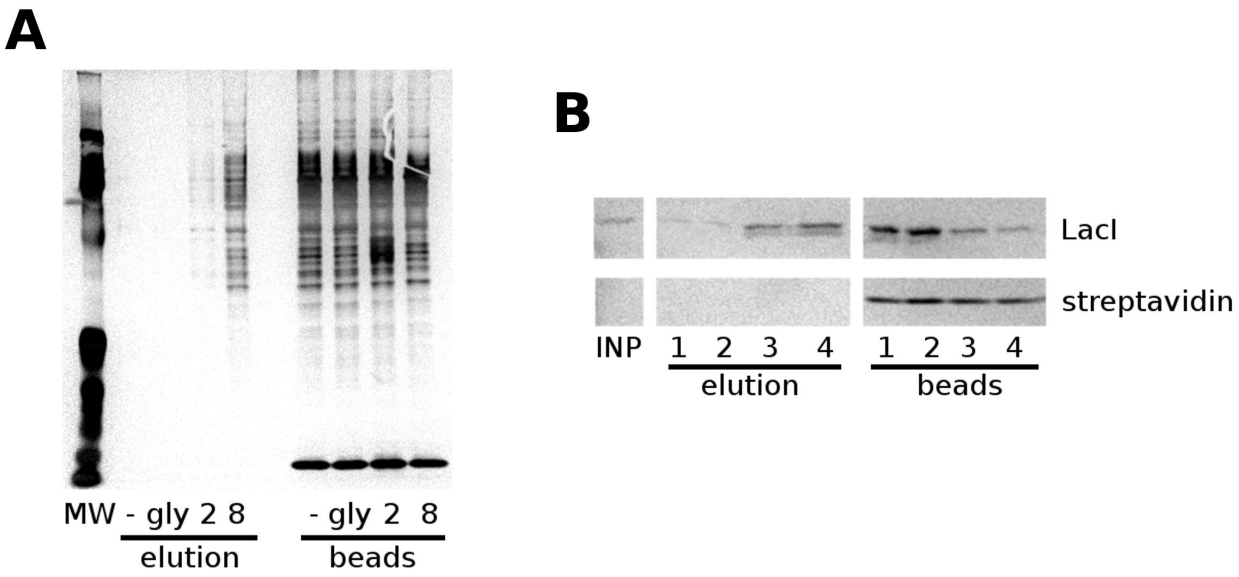


Figure 41: Tests of elution of LacI from LacO arrays beads.
The SDS-PAGE are stained by Coomassie blue. A) Identification of buffers removing proteins from streptavidin beads without DNA arrays. -: PBS. gly: 0.2M glycine pH2.3. 2: 2M urea. 8: 8M urea B) Elution of LacI loaded on LacO DNA arrays. INP: input. 1: PBS. 2: PBS + 1mM IPTG. 3: PBS + 0.5M NaCl + 1mM IPTG. 4: 0.2M glycine pH 2.3

j) Glycine Elution of protein recruited by cenp-T/LacI

The experiment of kinetochore recruitment is repeated. LacO beads with LacI or cenp-T/LacI are incubated with 100µL of XEE for 1h at 21˚C, quickly washed three times with the XEE buffer and incubated in 20µL of 0.2M glycine pH 2.3 for 10min at RT. The beads are pelleted by magnetic racks, washed two times with PBS and resuspended in 30µL of SDS-PAGE loading buffer. The results are shown in the Figure 42. A volume of 13µL per well is loaded for SDS-PAGE further stained by Coomassie blue and 4µL are used per well for WB analysis.

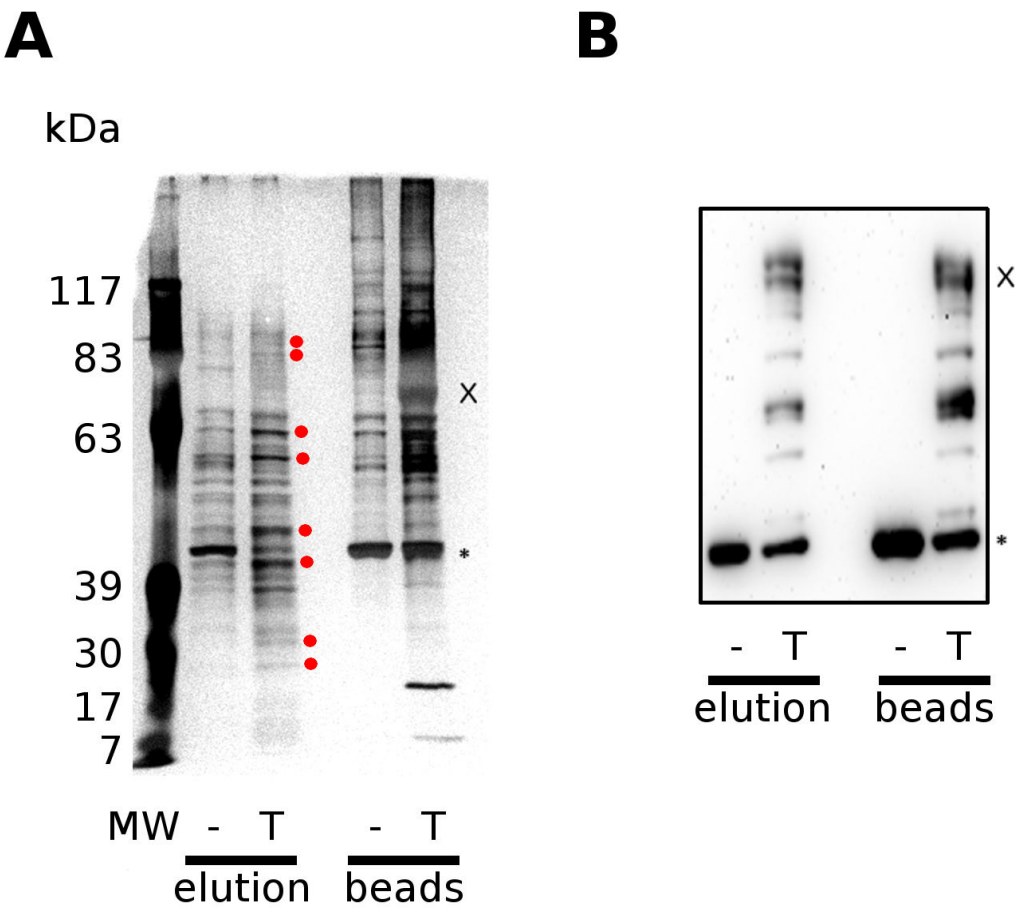


Figure 42: Recruitment of kinetochore proteins.
A) Elution of proteins recruited on LacO beads analyzed by SDS-PAGE stained by Coomassie blue. B) WB analysis of the samples with an antibody against LacI. –: LacI coupled to LacO beads. T: cenp-T/LacI coupled to LacO beads. Red circles mark LacI and crosses mark full-length cenp-T/LacI.

In the elution sample of LacI (Figure 42.A.panel elution, lane -), a smear of bands is present higher than 30kDa but lower than 100kDa with a higher-intensity band around 40kDa. The band at 40kDa corresponds to LacI. In the beads sample (Figure 42.A.panel beads, lane -T), a smear of protein is present higher than 39kDa. The most intense band in this sample is around 40kDa. The band at 10kDa (corresponding to streptavidin) is not visible in this figure but can be visualized by eye due to its low intensity and the limits of the imager.

In the elution sample of cenp-T/LacI (Figure 42.A.panel elution, lane T), a smear of bands is present approximately from 7kDa to 110kDa. There are no high intensity bands but several at 39kDa, 45kDa, 55kDa and 63kDa. In the beads sample, there is a smear of bands from upper than 39kDa and two bands at 7 and 20kDa. The patterns of bands are different between the elution and the beads samples, meaning that the glycine buffer is able to elute proteins loaded on DNA arrays and leave contaminants on beads.

To both confirm that the bands at 40kDa are LacI and the presence of cenp-T/LacI, WB analyses are performed using antibodies against LacI (Figure 42.B.lanes -). In both lanes of LacI in elution and beads sample, a unique band appears at an approximate MW of 40kDa. LacI is present in elution but not all LacI proteins are eluted and some remains on LacO beads. It is to note that the band corresponding to cenp-T/LacI is poorly stained by Coomassie blue. The presence and quantity of cenp-T/LacI is determined by WB analysis (Figure 42.B.lanes T). The pattern of bands present in the elution and beads samples is quite similar. There are bands at 40, 50, 72 and 75kDa. The bands are more intense in the beads sample and comparing to the volume of SDS-PAGE loading buffer used to analyze the sample, most of the cenp-T/LacI remains on the beads.

By comparing the patterns of bands between the elution of LacI and cenp-T/LacI samples, the proteins present in the LacI sample are present in the cenp-T/LacI sample. However, in the cenp-T/LacI sample, a few bands are more intense compared to LacI sample and few bands additional bands are present only in the cenp-T/LacI sample. Following a scientific discussion with the staff scientist, it was considered that the patterns of band are not different enough to further investigate by mass spectrometry analysis and protein identification. Only the WB results were used to identify proteins with antibodies.

Discussion

This approach recruits too much contaminant to clearly distinguish between proteins bound to cenp-T and proteins bound to the DNA arrays or to the beads. Even if the glycine elution allows a better identification of proteins non-specifically bound to the beads, the difference between the patterns of LacI and cenp-T sample is not clear enough. One of the reasons is the recruitment of bacterial proteins binding DNA that are eluted by glycine. The absence of MS analysis limits the ability to identify proteins in the sample.

Another issue is the low amount of proteins loaded on cenp-T, since the amount of contaminant is higher than the kinetochore proteins. The bands corresponding to LacI and cenp-T/LacI proteins are quite intense. The WB suggests that the amount of kinetochore proteins is low and has to be increased. Nevertheless, increasing the amount of beads will also increase the amount of contaminants, and it has to be considered that the beads preparation is quite expensive. There are 256 repeats of LacO per DNA arrays, thus increasing even more the quantity of LacO per arrays seems unsuitable. Thus, the beads protocol still has to be optimized

- Post Translation Modifications

The post translation modifications such as phosphorylations are important for the kinetochore functions. For example, the recruitment of the ndc80 complex by cenp-T is phospho-regulated through the cdk1 kinase. A vast numbers of enzymes are required and each one has different prerequisites that the LacI system may not necessary ensure. For example, the kinase AuroraB is a key kinase of the structural regulation of the centromere and the kinetochore. This enzyme is a member of the Chromosome Passenger Complex (CPC) that is located into the core of the centromere (Carmena et al. 2012) and a gradient of phosphorylation by auroraB is essential for the mitosis (Kaláb et al. 2006). AuroraB is not detected on the cenp-T beads, meaning that the phosphorylation state of the kinetochore proteins and its associated regulation is probably wrong.

- absence of centromere proteins

The absence of auroraB on our beads is probably due to the absence of centromere proteins. The cenp-A nucleosome leads to the recruitment of centromere proteins. Indeed, from the release of the protocol, this assay was used in a couple of publications about centromere (Westhorpe, Fuller, and Straight 2015) but, as far as I know, not for kinetochore studies. In the

cenp-T beads, the centromere proteins are not recruited, thus their associated functions are not absent totally or maybe partially rescued by some proteins in the solution.

- the folding of cenp-T

As mentioned during the purification test, cenp-T/LacI is highly insoluble. It is unknown if the cenp-T part loaded on the LacO arrays remains active. It would be interesting to check its phosphorylation state, to then test for its phosphorylation by cdk1.

- Use of a DNA template

I developed this assay in order to place cenp-T in a suitable DNA environment that I thought would be mandatory for the recruitment and the function of the kinetochore and centromere proteins. However, according to the WB results, it seems that centromere proteins are not recruited, leading to the conclusion that the challenging development of this method to reconstitute *in vitro* kinetochore has to be further optimized.

Conclusions and future directions

My initial objective of my thesis was to mutate the lysines of the cenp-H/I/K complex involved in its poly-SUMOylation and determine the associated phenotypes. The consequences of the disruption of the cenp-I poly-SUMOylation may be the increase of improperly loaded cenp-H/I/K/M complex onto the kinetochore or the lag of metaphase. It could be possible that the SUMOylation of cenp-I increases its incorporation into the kinetochore through SUMO-dependent interaction.

A standard method to identify post translation modification isimmuno-precipitation coupled to mass spectrometry analysis. But due to the low amount of cenp-I in the cell, the immuno-precipitation was challenging and does not allow mass spectrometry analysis. To obtain enough SUMOylated cenp-H/I/K/M complex for mass spectrometry analysis, I tried to purify the cenp-H/I/K complex to perform *in vitro* SUMOylation assay. The purification of the whole inner kinetochore from bacterial system has to be performed in conditions as close as possible to endogenous environment. Nevertheless, after optimization of the experimental conditions I was not able to purify cenp-H/I/K complex due to the high instability of cenp-I alone or within its complex.

During my Ph.D., the team of Dr. Stuckenberg identified a fourth member of this complex and developed a protocol for its purification. He kindly provided us pure cenp-H/I/K/M and I performed *in vitro* SUMOylation assays. While I successfully reconstituted the SUMOylation pathway *in vitro*, the SUMOylation of the cenp-H/I/K/M complex is very low without E3. Indeed, the E3 associated to the SUMOylation of the cenp-H/I/K/M complex is still unknown. However, I bypassed this limitation by increasing the E2 concentration. The products of SUMOylation were sent to mass spectrometry analyses for identification. Unfortunately, our collaborators were only able to identify the SUMOylated lysines for cenp-H but not for cenp-I.

To identify the lysines involved in the SUMOylation of cenp-I, systematic mutations could be performed. Once these lysines are identified, the identification of the role of the cenp-I SUMOylation can be determined. However, this approach is time-consuming. To identify the SUMOylated residues, the SUMOylation assay could also be performed with a mutated SUMO with a lysine surrounding the di-glycine motif (Tammsalu et al. 2014). Even this method leads to

modification of the kinetic of the SUMOylation pathway, the reliability of the identified lysines could be confirmed *in vivo.*

If no residues were identified for cenp-I, five SUMOylated residues were identified for cenp-H and it could be interesting to determine the effects of mutations of these residues. Because expression of eGFP-cenp-H by transient transfection leads to cenp-H aggregates (data not shown), stable transfections with low expression or CRISPR have be performed. However these methods are time-consuming and I was not able to do it during my Thesis.

To study the SUMOylation of the kinetochore Shaofei Zhang and I followed a protocol from the team of Dr. Straigh to reconstitute the kinetochore *in vitro*. In this protocol, the centromere nucleosome containing cenp-A is formed *in vitro* and linked to magnetic beads. These nucleosome beads are mixed to xenopus egg extracts (XEE), washed and analyzed. We successfully reproduced the protocol and recruited centromere proteins. However, when the beads are removed from the XEE, the cenp-A is immediately released. Thus, this protocol is not suitable for biochemical assays. Optimizations to increase the stability were tried but all the tests failed.

To determine which kinetochore proteins could reconstitute kinetochores at ectopic places, the teams of Dr. Cheeseman and Dr. Fukagawa inserted LacO DNA sequence into chromosome arms and fused kinetochore proteins to LacI. Some kinetochore proteins including cenp-T are able to reconstitute functional kinetochores. I tried to adapt this protocol for *in vitro* reconstitution. Among the proteins able to reconstitute ectopic kinetochores, I focused on cenp-T that is a root of the inner kinetochore, linking the outer kinetochore to the centromere. A 256 repeats of LacO DNA arrays was coupled to magnetic beads. The production and purification tests of the fusing protein cenp-T/LacI were performed to load pure and homogenous sample on beads. However, this protein is highly instable and I was not able to purify it.

To reconstitute *in vitro* kinetochores, we consider two directions: a) find a way to strengthen the cenp-A interaction with beads. If cenp-A nucleosomes recruit centromere and kinetochore proteins, these interactions could be transient and this assay could not be suitable for biochemical studies even in the case of the presence of cenp-A on beads; b) target proteins with the LacI system as I did but complete this assay with proteins able to recruit centromere proteins.

If the cenp-T/LacI is able to recruit some kinetochore proteins, the absence of centromere proteins is a huge limit. In all cases, the recruitment of centromere proteins is a requirement.

In conclusion, we faced a lot of technical and conceptual challenges during my thesis. I determined that PIASy and the E3 fragment of RanBP2 do not act as an E3 for the cenp-I SUMOylation. Also, cenp-T is able to recruit several kinetochore proteins but this method is not convenient compared to chromatin purification or the cenp-A nucleosomes of the Dr. Straigh's team. My work was a modest step to the understanding of the molecular mechanisms regulating the complex function of the kinetochores.

Bibliography

Amano, M., A. Suzuki, T. Hori, C. Backer, K. Okawa, I. M. Cheeseman, and T. Fukagawa. 2009. “The CENP-S Complex Is Essential for the Stable Assembly of Outer Kinetochore Structure.” *The Journal of Cell Biology* 186 (2): 173–82. doi:10.1083/jcb.200903100.

Amaro, Ana C., Catarina P. Samora, René Holtackers, Enxiu Wang, Isabel J. Kingston, Maria Alonso, Michael Lampson, Andrew D. McAinsh, and Patrick Meraldi. 2010. “Molecular Control of Kinetochore-Microtubule Dynamics and Chromosome Oscillations.” *Nature Cell Biology* 12 (4): 319–29. doi:10.1038/ncb2033.

Antonin, Wolfram, and Heinz Neumann. 2016. “Chromosome Condensation and Decondensation during Mitosis.” *Current Opinion in Cell Biology* 40 (June): 15–22. doi:10.1016/j.ceb.2016.01.013.

Aragon, Luis, Enrique Martinez-Perez, and Matthias Merkenschlager. 2013. “Condensin, Cohesin and the Control of Chromatin States.” *Current Opinion in Genetics & Development* 23 (2): 204–11. doi:10.1016/j.gde.2012.11.004.

Arents, Gina, Rufus W. Burlingame, Bi-Cheng Wang, Warner E. Love, and Evangelos N. Moudrianakis. 1991. “The Nucleosomal Core Histone Octamer at 3.1 A Resolution: A Tripartite Protein Assembly and a Left-Handed Superhelix.” *Proceedings of the National Academy of Sciences* 88 (22): 10148–10152.

Balakirev, Maxim Y., James E. Mullally, Adrien Favier, Nicole Assard, Eric Sulpice, David F. Lindsey, Anastasia V. Rulina, Xavier Gidrol, and Keith D. Wilkinson. 2015. “Wss1 Metalloprotease Partners with Cdc48/Doa1 in Processing Genotoxic SUMO Conjugates.” *Elife* 4: e06763.

Bannister, Andrew J., and Tony Kouzarides. 2011. “Regulation of Chromatin by Histone Modifications.” *Cell Research* 21 (3): 381–395.

Basilico, Federica, Stefano Maffini, John R. Weir, Daniel Prumbaum, Ana M. Rojas, Tomasz Zimniak, Anna De Antoni, et al. 2014. “The Pseudo GTPase CENP-M Drives Human Kinetochore Assembly.” *eLife* 3: e02978.

Bednar, Jan, Rachel A. Horowitz, Sergei A. Grigoryev, Lenny M. Carruthers, Jeffrey C. Hansen, Abraham J. Koster, and Christopher L. Woodcock. 1998. “Nucleosomes, Linker DNA, and Linker Histone Form a Unique Structural Motif That Directs the Higher-Order Folding and Compaction of Chromatin.” *Proceedings of the National Academy of Sciences* 95 (24): 14173–14178.

Bellini, Dom, Anthony P. Fordham-Skelton, and Miroslav Z. Papiz. 2011. “STRU-Cloning: A Fast, Inexpensive and Efficient Cloning Procedure Applicable to Both Small Scale and Structural Genomics Size Cloning.” *Molecular Biotechnology* 48 (1): 30–37.

Bernad, R., P. Sanchez, T. Rivera, M. Rodriguez-Corsino, E. Boyarchuk, I. Vassias, D. Ray-Gallet, et al. 2011. “Xenopus HJURP and Condensin II Are Required for CENP-A Assembly.” *The Journal of Cell Biology* 192 (4): 569–82. doi:10.1083/jcb.201005136.

Beroukhim, Rameen, Craig H. Mermel, Dale Porter, Guo Wei, Soumya Raychaudhuri, Jerry Donovan, Jordi Barretina, et al. 2010. “The Landscape of Somatic Copy-Number Alteration across Human Cancers.” *Nature* 463 (7283): 899–905. doi:10.1038/nature08822.

“Biomedical EM Unit: Gallery.” 2016. April 13. http://pcwww.liv.ac.uk/~emunit/gallerycell.html.

Bock, Lucy J., Cinzia Pagliuca, Norihiko Kobayashi, Ryan A. Grove, Yusuke Oku, Kriti Shrestha, Claudio Alfieri, et al. 2012. “Cnn1 Inhibits the Interactions between the KMN Complexes of the Yeast Kinetochore.” *Nature Cell Biology* 14 (6): 614–24. doi:10.1038/ncb2495.

Bohren, K. M., V. Nadkarni, J. H. Song, K. H. Gabbay, and D. Owerbach. 2004. “A M55V Polymorphism in a Novel SUMO Gene (SUMO-4) Differentially Activates Heat Shock Transcription Factors and Is Associated with Susceptibility to Type I Diabetes Mellitus.” *Journal of Biological Chemistry* 279 (26): 27233–38. doi:10.1074/jbc.M402273200.

Bolanos-Garcia, Victor Martin, and Owen Richard Davies. 2006. “Structural Analysis and Classification of Native Proteins from E. Coli Commonly Co-Purified by Immobilised Metal Affinity Chromatography.” *Biochimica et Biophysica Acta (BBA) - General Subjects* 1760 (9): 1304–13. doi:10.1016/j.bbagen.2006.03.027.

Bylebyl, G. R., I. Belichenko, and E. S. Johnson. 2003. “The SUMO Isopeptidase Ulp2 Prevents Accumulation of SUMO Chains in Yeast.” *Journal of Biological Chemistry* 278 (45): 44113–20. doi:10.1074/jbc.M308357200.

Carmena, Mar, Michael Wheelock, Hironori Funabiki, and William C. Earnshaw. 2012. “The Chromosomal Passenger Complex (CPC): From Easy Rider to the Godfather of Mitosis.” *Nature Reviews Molecular Cell Biology* 13 (12): 789–803. doi:10.1038/nrm3474.

Carroll, Christopher W., Mariana C.C. Silva, Kristina M. Godek, Lars E.T. Jansen, and Aaron F. Straight. 2009. “Centromere Assembly Requires the Direct Recognition of CENP-A Nucleosomes by CENP-N.” *Nature Cell Biology* 11 (7): 896–902. doi:10.1038/ncb1899.

Cheeseman, I. M. 2014. “The Kinetochore.” *Cold Spring Harbor Perspectives in Biology* 6 (7): a015826–a015826. doi:10.1101/cshperspect.a015826.

Cheeseman, Iain M., Joshua S. Chappie, Elizabeth M. Wilson-Kubalek, and Arshad Desai. 2006. “The Conserved KMN Network Constitutes the Core Microtubule-Binding Site of the Kinetochore.” *Cell* 127 (5): 983–97. doi:10.1016/j.cell.2006.09.039.

Cheeseman, Iain M., and Arshad Desai. 2008. “Molecular Architecture of the Kinetochore–microtubule Interface.” *Nature Reviews Molecular Cell Biology* 9 (1): 33–46. doi:10.1038/nrm2310.
Chen, Zhu Hong, Minglu Zhu, Jingyi Yang, Hui Liang, Jinxue He, Shiming He, Pan Wang, et al. 2014. “PTEN Interacts with Histone H1 and Controls Chromatin Condensation.” *Cell Reports* 8 (6): 2003–14. doi:10.1016/j.celrep.2014.08.008.
Clarke, Paul R., and Chuanmao Zhang. 2008. “Spatial and Temporal Coordination of Mitosis by Ran GTPase.” *Nature Reviews Molecular Cell Biology* 9 (6): 464–77. doi:10.1038/nrm2410.
Cleveland, Don W., Yinghui Mao, and Kevin F. Sullivan. 2003. “Centromeres and Kinetochores: From Epigenetics to Mitotic Checkpoint Signaling.” *Cell* 112 (4): 407–421.
Denison, C. 2005. “A Proteomic Strategy for Gaining Insights into Protein Sumoylation in Yeast.” *Molecular & Cellular Proteomics* 4 (3): 246–54. doi:10.1074/mcp.M400154-MCP200.
Earnshaw, William C., and Naomi Rothfield. 1985. “Identification of a Family of Human Centromere Proteins Using Autoimmune Sera from Patients with Scleroderma.” *Chromosoma* 91 (3–4): 313–321.
Emanuele, Michael J., Weijie Lan, Miri Jwa, Stephanie A. Miller, Clarence S.M. Chan, and P. Todd Stukenberg. 2008. “Aurora B Kinase and Protein Phosphatase 1 Have Opposing Roles in Modulating Kinetochore Assembly.” *The Journal of Cell Biology* 181 (2): 241–54. doi:10.1083/jcb.200710019.
Fachinetti, Daniele, H. Diego Folco, Yael Nechemia-Arbely, Luis P. Valente, Kristen Nguyen, Alex J. Wong, Quan Zhu, et al. 2013. “A Two-Step Mechanism for Epigenetic Specification of Centromere Identity and Function.” *Nature Cell Biology*, July. doi:10.1038/ncb2805.
Fang, Junnan, Yuting Liu, Yun Wei, Wenqiang Deng, Zhouliang Yu, Li Huang, Yan Teng, et al. 2015. “Structural Transitions of Centromeric Chromatin Regulate the Cell Cycle-Dependent Recruitment of CENP-N.” *Genes & Development* 29 (10): 1058–73. doi:10.1101/gad.259432.115.
Fernandez-Miranda, G., I. P. de Castro, M. Carmena, C. Aguirre-Portoles, S. Ruchaud, X. Fant, G. Montoya, W. C. Earnshaw, and M. Malumbres. 2010. “SUMOylation Modulates the Function of Aurora-B Kinase.” *Journal of Cell Science* 123 (16): 2823–33. doi:10.1242/jcs.065565.
Fisher, and Guo. 2016. “Modification of a PCR-Based Site-Directed Mutagenesis Method.” April 5. http://www.biotechniques.com/multimedia/archive/00009/97234bm01_9763a.pdf.
Foltz, Daniel R., Lars E. T. Jansen, Ben E. Black, Aaron O. Bailey, John R. Yates, and Don W. Cleveland. 2006. “The Human CENP-A Centromeric Nucleosome-Associated Complex.” *Nature Cell Biology* 8 (5): 458–69. doi:10.1038/ncb1397.
Ford, Heide L., and Arthur B. Pardee. 1999. “Cancer and the Cell Cycle.” *Journal of Cellular Biochemistry* 75 (s 32): 166–172.
Fukagawa, Tatsuo, and William RA Brown. 1997. “Efficient Conditional Mutation of the Vertebrate CENP-C Gene.” *Human Molecular Genetics* 6 (13): 2301–2308.
Fukagawa, Tatsuo, and William C. Earnshaw. 2014. “The Centromere: Chromatin Foundation for the Kinetochore Machinery.” *Developmental Cell* 30 (5): 496–508. doi:10.1016/j.devcel.2014.08.016.
Fukagawa, Tatsuo, Yoshikazu Mikami, Ai Nishihashi, Vinciane Regnier, Tokuko Haraguchi, Yasushi Hiraoka, Naoko Sugata, Kazuo Todokoro, William Brown, and Toshimichi Ikemura. 2001. “CENP-H, a Constitutive Centromere Component, Is Required for Centromere Targeting of CENP-C in Vertebrate Cells.” *The EMBO Journal* 20 (16): 4603–4617.
Funabiki, Hironori, and David J. Wynne. 2013. “Making an Effective Switch at the Kinetochore by Phosphorylation and Dephosphorylation.” *Chromosoma* 122 (3): 135–58. doi:10.1007/s00412-013-0401-5.
Gascoigne, K. E., and I. M. Cheeseman. 2013. “CDK-Dependent Phosphorylation and Nuclear Exclusion Coordinately Control Kinetochore Assembly State.” *The Journal of Cell Biology* 201 (1): 23–32. doi:10.1083/jcb.201301006.
Gascoigne, Karen E, and Iain M Cheeseman. 2011. “Kinetochore Assembly: If You Build It, They Will Come.” *Current Opinion in Cell Biology* 23 (1): 102–8. doi:10.1016/j.ceb.2010.07.007.
Gascoigne, Karen E., Kozo Takeuchi, Aussie Suzuki, Tetsuya Hori, Tatsuo Fukagawa, and Iain M. Cheeseman. 2011. “Induced Ectopic Kinetochore Assembly Bypasses the Requirement for CENP-A Nucleosomes.” *Cell* 145 (3): 410–422.
Gordon, David J., Benjamin Resio, and David Pellman. 2012. “Causes and Consequences of Aneuploidy in Cancer.” *Nature Reviews Genetics*, January. doi:10.1038/nrg3123.
Goshima, Gohta, Tomomi Kiyomitsu, Kinya Yoda, and Mitsuhiro Yanagida. 2003. “Human Centromere Chromatin Protein hMis12, Essential for Equal Segregation, Is Independent of CENP-A Loading Pathway.” *The Journal of Cell Biology* 160 (1): 25–39. doi:10.1083/jcb.200210005.
Guse, A., C. J. Fuller, and A. F. Straight. 2012. “A Cell-Free System for Functional Centromere and Kinetochore Assembly.” *Nature Protocols* 7 (10): 1847–1869.
Guse, Annika, Christopher W. Carroll, Ben Moree, Colin J. Fuller, and Aaron F. Straight. 2011. “In Vitro Centromere and Kinetochore Assembly on Defined Chromatin Templates.” *Nature* 477 (7364): 354–58. doi:10.1038/nature10379.

Hakli, M, U Karvonen, O Janne, and J Palvimo. 2005. “SUMO-1 Promotes Association of SNURF (RNF4) with PML Nuclear Bodies.” *Experimental Cell Research* 304 (1): 224–33. doi:10.1016/j.yexcr.2004.10.029.
Häkli, M, K.L Lorick, A.M Weissman, O.A Jänne, and J.J Palvimo. 2004. “Transcriptional Coregulator SNURF (RNF4) Possesses Ubiquitin E3 Ligase Activity.” *FEBS Letters* 560 (1–3): 56–62. doi:10.1016/S0014-5793(04)00070-5.
Hannich, J. Thomas, Alaron Lewis, Mary B. Kroetz, Shyr-Jiann Li, Heinrich Heide, Andrew Emili, and Mark Hochstrasser. 2005. “Defining the SUMO-Modified Proteome by Multiple Approaches in *Saccharomyces Cerevisiae*.” *Journal of Biological Chemistry* 280 (6): 4102–10. doi:10.1074/jbc.M413209200.
Hershko, Avram. 1999. “Mechanisms and Regulation of the Degradation of Cyclin B.” *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 354 (1389): 1571–1576.
Hickey, Christopher M., Nicole R. Wilson, and Mark Hochstrasser. 2012. “Function and Regulation of SUMO Proteases.” *Nature Reviews Molecular Cell Biology* 13 (12): 755–66. doi:10.1038/nrm3478.
Hinshaw, Stephen M., and Stephen C. Harrison. 2013. “An Iml3-Chl4 Heterodimer Links the Core Centromere to Factors Required for Accurate Chromosome Segregation.” *Cell Reports* 5 (1): 29–36. doi:10.1016/j.celrep.2013.08.036.
Hirano, Tatsuya. 2016. “Condensin-Based Chromosome Organization from Bacteria to Vertebrates.” *Cell* 164 (5): 847–57. doi:10.1016/j.cell.2016.01.033.
Hirano, Tatsuya, Ryuji Kobayashi, and Michiko Hirano. 1997. “Condensins, Chromosome Condensation Protein Complexes Containing XCAP-C, XCAP-E and a Xenopus Homolog of the Drosophila Barren Protein.” *Cell* 89 (4): 511–521.
Hori, T., W.-H. Shang, K. Takeuchi, and T. Fukagawa. 2012. “The CCAN Recruits CENP-A to the Centromere and Forms the Structural Core for Kinetochore Assembly.” *The Journal of Cell Biology* 200 (1): 45–60. doi:10.1083/jcb.201210106.
Hori, Tetsuya, Miho Amano, Aussie Suzuki, Chelsea B. Backer, Julie P. Welburn, Yimin Dong, Bruce F. McEwen, et al. 2008. “CCAN Makes Multiple Contacts with Centromeric DNA to Provide Distinct Pathways to the Outer Kinetochore.” *Cell* 135 (6): 1039–52. doi:10.1016/j.cell.2008.10.019.
Hori, Tetsuya, Masahiro Okada, Katsumi Maenaka, and Tatsuo Fukagawa. 2008. “CENP-O Class Proteins Form a Stable Complex and Are Required for Proper Kinetochore Function.” *Molecular Biology of the Cell* 19 (3): 843–854.
Hornig, Nadine CD, Philip P. Knowles, Neil Q. McDonald, and Frank Uhlmann. 2002. “The Dual Mechanism of Separase Regulation by Securin.” *Current Biology* 12 (12): 973–982.
Howman, Emily V., Kerry J. Fowler, Ainsley J. Newson, Saara Redward, Andrew C. MacDonald, Paul Kalitsis, and KH Andy Choo. 2000. “Early Disruption of Centromeric Chromatin Organization in Centromere Protein A (Cenpa) Null Mice.” *Proceedings of the National Academy of Sciences* 97 (3): 1148–1153.
Impens, F., L. Radoshevich, P. Cossart, and D. Ribet. 2014. “Mapping of SUMO Sites and Analysis of SUMOylation Changes Induced by External Stimuli.” *Proceedings of the National Academy of Sciences* 111 (34): 12432–37. doi:10.1073/pnas.1413825111.
Janke, Carsten, Jennifer Ortiz, Johannes Lechner, Anna Shevchenko, Andrej Shevchenko, Maria M. Magiera, Carolin Schramm, and Elmar Schiebel. 2001. “The Budding Yeast Proteins Spc24p and Spc25p Interact with Ndc80p and Nuf2p at the Kinetochore and Are Important for Kinetochore Clustering and Checkpoint Control.” *The EMBO Journal* 20 (4): 777–791.
Kagawa, Naoko, Tetsuya Hori, Yuko Hoki, Osamu Hosoya, Kimiko Tsutsui, Yumiko Saga, Takashi Sado, and Tatsuo Fukagawa. 2014. “The CENP-O Complex Requirement Varies among Different Cell Types.” *Chromosome Research* 22 (3): 293–303. doi:10.1007/s10577-014-9404-1.
Kaláb, Petr, Arnd Pralle, Ehud Y. Isacoff, Rebecca Heald, and Karsten Weis. 2006. “Analysis of a RanGTP-Regulated Gradient in Mitotic Somatic Cells.” *Nature* 440 (7084): 697–701. doi:10.1038/nature04589.
Kalitsis, Paul, Kerry J. Fowler, Elizabeth Earle, Joanne Hill, and KH Andy Choo. 1998. “Targeted Disruption of Mouse Centromere Protein C Gene Leads to Mitotic Disarray and Early Embryo Death.” *Proceedings of the National Academy of Sciences* 95 (3): 1136–1141.
Karess, Roger. 2005. “Rod–Zw10–Zwilch: A Key Player in the Spindle Checkpoint.” *Trends in Cell Biology* 15 (7): 386–92. doi:10.1016/j.tcb.2005.05.003.
kato, H., U. Ohto, T. Hiramatsu, M. Urabe, Y. Uchida, Y. Satow, and Y. Bai. 2013. “A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C.” *Science* 340 (6136): 1106–10. doi:10.1126/science.1233508.
Kim, S., and H. Yu. 2015. “Multiple Assembly Mechanisms Anchor the KMN Spindle Checkpoint Platform at Human Mitotic Kinetochores.” *The Journal of Cell Biology* 208 (2): 181–96. doi:10.1083/jcb.201407074.
Kline, Susan L., Iain M. Cheeseman, Tetsuya Hori, Tatsuo Fukagawa, and Arshad Desai. 2006. “The Human Mis12 Complex Is Required for Kinetochore Assembly and Proper Chromosome Segregation.” *The Journal of Cell Biology* 173 (1): 9–17. doi:10.1083/jcb.200509158.
Knipscheer, Puck, Annette Flotho, Helene Klug, Jesper V. Olsen, Willem J. van Dijk, Alexander Fish, Erica S. Johnson, Matthias Mann, Titia K. Sixma, and Andrea Pichler. 2008. “Ubc9 Sumoylation Regulates SUMO Target Discrimination.” *Molecular Cell* 31 (3): 371–382.

Knouse, Kristin A., Jie Wu, Charles A. Whittaker, and Angelika Amon. 2014. “Single Cell Sequencing Reveals Low Levels of Aneuploidy across Mammalian Tissues.” *Proceedings of the National Academy of Sciences* 111 (37): 13409–14. doi:10.1073/pnas.1415287111.

Kudalkar, Emily M., Emily A. Scarborough, Neil T. Umbreit, Alex Zelter, Daniel R. Gestaut, Michael Riffle, Richard S. Johnson, Michael J. MacCoss, Charles L. Asbury, and Trisha N. Davis. 2015. “Regulation of Outer Kinetochore Ndc80 Complex-Based Microtubule Attachments by the Central Kinetochore Mis12/MIND Complex.” *Proceedings of the National Academy of Sciences*, October, 201513882. doi:10.1073/pnas.1513882112.

Kung, Camy, Mandar Naik, Szu-Huan Wang, Hsiu-Ming Shih, Che-Chang Chang, Li-Ying Lin, Chia-Lin Chen, Che Ma, Chi-Fon Chang, and Tai-Huang Huang. 2014. “Structural Analysis of Poly-SUMO Chain Recognition by RNF4-SIMs Domain.” *Biochemical Journal*, May. doi:10.1042/BJ20140521.

Laemmli, U K. 1970. “Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4.” *Nature* 227 (5259): 680–85.

Lander, Eric S., Lauren M. Linton, Bruce Birren, Chad Nusbaum, Michael C. Zody, Jennifer Baldwin, Keri Devon, et al. 2001. “Initial Sequencing and Analysis of the Human Genome.” *Nature* 409 (6822): 860–921.

Li, Shyr-Jiann, and Mark Hochstrasser. 1999. “A New Protease Required for Cell-Cycle Progression in Yeast.” 02. http://www.nature.com/nature/journal/v398/n6724/pdf/398246a0.pdf.

———. 2000. “The Yeast ULP2 (SMT4) Gene Encodes a Novel Protease Specific for the Ubiquitin-like Smt3 Protein.” *Molecular and Cellular Biology* 20 (7): 2367–2377.

Lima, C. D., and D. Reverter. 2008. “Structure of the Human SENP7 Catalytic Domain and Poly-SUMO Deconjugation Activities for SENP6 and SENP7.” *Journal of Biological Chemistry* 283 (46): 32045–55. doi:10.1074/jbc.M805655200.

Liu, Qiang, Jia Cao, Ke Qiu Li, Xu Hong Miao, Guang Li, Fei Yue Fan, and Yong Cheng Zhao. 2009. “Chromosomal Aberrations and DNA Damage in Human Populations Exposed to the Processing of Electronics Waste.” *Environmental Science and Pollution Research* 16 (3): 329–38. doi:10.1007/s11356-008-0087-z.

London, Nitobe, and Sue Biggins. 2014. “Signalling Dynamics in the Spindle Checkpoint Response.” *Nature Reviews Molecular Cell Biology* 15 (11): 736–48. doi:10.1038/nrm3888.

Lowary, P. T., and Jonathan Widom. 1998. “New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning.” *Journal of Molecular Biology* 276 (1): 19–42.

Malvezzi, Francesca, Gabriele Litos, Alexander Schleiffer, Alexander Heuck, Karl Mechtler, Tim Clausen, and Stefan Westermann. 2013. “A Structural Basis for Kinetochore Recruitment of the Ndc80 Complex via Two Distinct Centromere Receptors.” *The EMBO Journal*. http://www.nature.com/emboj/journal/vaop/ncurrent/full/emboj2012356a.html.

McAinsh, Andrew D., and Patrick Meraldi. 2011. “The CCAN Complex: Linking Centromere Specification to Control of Kinetochore–microtubule Dynamics.” *Seminars in Cell & Developmental Biology* 22 (9): 946–52. doi:10.1016/j.semcdb.2011.09.016.

McClelland, Sarah E., Satyarebala Borusu, Ana C. Amaro, Jennifer R. Winter, Mukta Belwal, Andrew D. McAinsh, and Patrick Meraldi. 2007. “The CENP-A NAC/CAD Kinetochore Complex Controls Chromosome Congression and Spindle Bipolarity.” *The EMBO Journal* 26 (24): 5033–5047.

McKinley, Kara L., Nikolina Sekulic, Lucie Y. Guo, Tonia Tsinman, Ben E. Black, and Iain M. Cheeseman. 2015. “The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface.” *Molecular Cell* 60 (6): 886–98. doi:10.1016/j.molcel.2015.10.027.

Melters, Daniël P., Keith R. Bradnam, Hugh A. Young, Natalie Telis, Michael R. May, J. Graham Ruby, Robert Sebra, et al. 2013. “Comparative Analysis of Tandem Repeats from Hundreds of Species Reveals Unique Insights into Centromere Evolution.” *Genome Biology* 14 (1): 1.

Meluh, Pamela B., and Douglas Koshland. 1997. “Budding Yeast Centromere Composition and Assembly as Revealed by in Vivo Cross-Linking.” *Genes & Development* 11 (24): 3401–3412.

Milks, Kirstin J., Ben Moree, and Aaron F. Straight. 2009. “Dissection of CENP-C–directed Centromere and Kinetochore Assembly.” *Molecular Biology of the Cell* 20 (19): 4246–4255.

Minty, A., X. Dumont, M. Kaghad, and D. Caput. 2000. “Covalent Modification of p73 by SUMO-1: TWO-HYBRID SCREENING WITH p73 IDENTIFIES NOVEL SUMO-1-INTERACTING PROTEINS AND A SUMO-1 INTERACTION MOTIF.” *Journal of Biological Chemistry* 275 (46): 36316–23. doi:10.1074/jbc.M004293200.

“Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer.” 2016. April 4. http://cgap.nci.nih.gov/Chromosomes/Mitelman.

Moroi, Yasuoki, Carol Peebles, Marvin J. Fritzler, J. Steigerwald, and Eng M. Tan. 1980. “Autoantibody to Centromere (Kinetochore) in Scleroderma Sera.” *Proceedings of the National Academy of Sciences* 77 (3): 1627–1631.

Mossessova, Elena, and Christopher D. Lima. 2000. “Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast.” *Molecular Cell* 5 (5): 865–876.

Mukhopadhyay, D., A. Arnaoutov, and M. Dasso. 2010. “The SUMO Protease SENP6 Is Essential for Inner Kinetochore Assembly.” *The Journal of Cell Biology* 188 (5): 681–92. doi:10.1083/jcb.200909008.

Mukhopadhyay, Debaditya, and Mary Dasso. 2007. “Modification in Reverse: The SUMO Proteases.” *Trends in Biochemical Sciences* 32 (6): 286–95. doi:10.1016/j.tibs.2007.05.002.

Musacchio, Andrea, and Edward D. Salmon. 2007. “The Spindle-Assembly Checkpoint in Space and Time.” *Nature Reviews Molecular Cell Biology* 8 (5): 379–93. doi:10.1038/nrm2163.

Nardi, Isaac K., Ewelina Zasadzińska, Madison E. Stellfox, Christina M. Knippler, and Daniel R. Foltz. 2016. “Licensing of Centromeric Chromatin Assembly through the Mis18α-Mis18β Heterotetramer.” *Molecular Cell* 61 (5): 774–87. doi:10.1016/j.molcel.2016.02.014.

Nechemia-Arbely. n.d. “Communication during the Annual Meeting of the ASCB 2015.”

Nishihashi, Ai, Tokuko Haraguchi, Yasushi Hiraoka, Toshimichi Ikemura, Vinciane Regnier, Helen Dodson, William C. Earnshaw, and Tatsuo Fukagawa. 2002. “CENP-I Is Essential for Centromere Function in Vertebrate Cells.” *Developmental Cell* 2 (4): 463–476.

Nishino, Tatsuya, Florencia Rago, Tetsuya Hori, Kentaro Tomii, Iain M. Cheeseman, and Tatsuo Fukagawa. 2013. “CENP-T Provides a Structural Platform for Outer Kinetochore Assembly.” *The EMBO Journal*. http://www.nature.com/emboj/journal/vaop/ncurrent/full/emboj2012348a.html.

Nishino, Tatsuya, Kozo Takeuchi, Karen E. Gascoigne, Aussie Suzuki, Tetsuya Hori, Takuji Oyama, Kosuke Morikawa, Iain M. Cheeseman, and Tatsuo Fukagawa. 2012. “CENP-T-W-S-X Forms a Unique Centromeric Chromatin Structure with a Histone-like Fold.” *Cell* 148 (3): 487–501. doi:10.1016/j.cell.2011.11.061.

Okada, Masahiro, Iain M. Cheeseman, Tetsuya Hori, Katsuya Okawa, Ian X. McLeod, John R. Yates, Arshad Desai, and Tatsuo Fukagawa. 2006. “The CENP-H–I Complex Is Required for the Efficient Incorporation of Newly Synthesized CENP-A into Centromeres.” *Nature Cell Biology* 8 (5): 446–57. doi:10.1038/ncb1396.

Okada, Masahiro, Katsuya Okawa, Toshiaki Isobe, and Tatsuo Fukagawa. 2009. “CENP-H–containing Complex Facilitates Centromere Deposition of CENP-A in Cooperation with FACT and CHD1.” *Molecular Biology of the Cell* 20 (18): 3986–3995.

Orr, Bernardo, and Duane A. Compton. 2013. “A Double-Edged Sword: How Oncogenes and Tumor Suppressor Genes Can Contribute to Chromosomal Instability.” *Frontiers in Oncology* 3. doi:10.3389/fonc.2013.00164.

Panse, V. G., U. Hardeland, T. Werner, B. Kuster, and E. Hurt. 2004. “A Proteome-Wide Approach Identifies Sumoylated Substrate Proteins in Yeast.” *Journal of Biological Chemistry* 279 (40): 41346–51. doi:10.1074/jbc.M407950200.

Petrovic, Arsen, Sebastiano Pasqualato, Prakash Dube, Veronica Krenn, Stefano Santaguida, Davide Cittaro, Silvia Monzani, et al. 2010. “The MIS12 Complex Is a Protein Interaction Hub for Outer Kinetochore Assembly.” *The Journal of Cell Biology* 190 (5): 835–52. doi:10.1083/jcb.201002070.

Pichler, Andrea, Andreas Gast, Jacob S. Seeler, Anne Dejean, and Frauke Melchior. 2002. “The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity.” *Cell* 108 (1): 109–120.

Pilla, E., U. Moller, G. Sauer, F. Mattiroli, F. Melchior, and R. Geiss-Friedlander. 2012. “A Novel SUMO1-Specific Interacting Motif in Dipeptidyl Peptidase 9 (DPP9) That Is Important for Enzymatic Regulation.” *Journal of Biological Chemistry* 287 (53): 44320–29. doi:10.1074/jbc.M112.397224.

Pinto, Manuel P., Andreia F. Carvalho, Cláudia P. Grou, José E. Rodríguez-Borges, Clara Sá-Miranda, and Jorge E. Azevedo. 2012. “Heat Shock Induces a Massive but Differential Inactivation of SUMO-Specific Proteases.” *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1823 (10): 1958–66. doi:10.1016/j.bbamcr.2012.07.010.

“Preparation and Use of Xenopus Egg Extracts to Study DNA Replication and Chromatin Associated Proteins.” 2014. August 27. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3437562/?report=printable.

Reverter, David, and Christopher D Lima. 2006. “Structural Basis for SENP2 Protease Interactions with SUMO Precursors and Conjugated Substrates.” *Nature Structural & Molecular Biology* 13 (12): 1060–68. doi:10.1038/nsmb1168.

Rodriguez, M. S. 2001. “SUMO-1 Conjugation in Vivo Requires Both a Consensus Modification Motif and Nuclear Targeting.” *Journal of Biological Chemistry* 276 (16): 12654–59. doi:10.1074/jbc.M009476200.

Ryu, Hyunju, Steven P. Gygi, Yoshiaki Azuma, Alexei Arnaoutov, and Mary Dasso. 2014. “SUMOylation of Psmd1 Controls Adrm1 Interaction with the Proteasome.” *Cell Reports* 7 (6): 1842–48. doi:10.1016/j.celrep.2014.05.009.

Sampson, D. A., M. Wang, and M. J. Matunis. 2001. “The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and Is Essential for SUMO-1 Modification.” *Journal of Biological Chemistry* 276 (24): 21664–69. doi:10.1074/jbc.M100006200.

Santaguida, Stefano, and Andrea Musacchio. 2009. “The Life and Miracles of Kinetochores.” *The EMBO Journal* 28 (17): 2511–2531.

Schleiffer, Alexander, Michael Maier, Gabriele Litos, Fabienne Lampert, Peter Hornung, Karl Mechtler, and Stefan Westermann. 2012. “CENP-T Proteins Are Conserved Centromere Receptors of the Ndc80 Complex.” *Nature Cell Biology* 14 (6): 604–613.

Schueler, Mary G., Anne W. Higgins, M. Katharine Rudd, Karen Gustashaw, and Huntington F. Willard. 2001. “Genomic and Genetic Definition of a Functional Human Centromere.” *Science* 294 (5540): 109–115.

Screpanti, Emanuela, Anna De Antoni, Gregory M. Alushin, Arsen Petrovic, Tiziana Melis, Eva Nogales, and Andrea Musacchio. 2011. “Direct Binding of Cenp-C to the Mis12 Complex Joins the Inner and Outer Kinetochore.” *Current Biology* 21 (5): 391–98. doi:10.1016/j.cub.2010.12.039.

Shang, Wei-Hao, Tetsuya Hori, Nuno M.C. Martins, Atsushi Toyoda, Sadahiko Misu, Norikazu Monma, Ichiro Hiratani, et al. 2013. “Chromosome Engineering Allows the Efficient Isolation of Vertebrate Neocentromeres.” *Developmental Cell* 24 (6): 635–48. doi:10.1016/j.devcel.2013.02.009.

Shapiro, A L, E Viñuela, and J V Maizel. 1967. “Molecular Weight Estimation of Polypeptide Chains by Electrophoresis in SDS-Polyacrylamide Gels.” *Biochemical and Biophysical Research Communications* 28 (5): 815–20.

Sheltzer, Jason M., and Angelika Amon. 2011. “The Aneuploidy Paradox: Costs and Benefits of an Incorrect Karyotype.” *Trends in Genetics* 27 (11): 446–53. doi:10.1016/j.tig.2011.07.003.

Shin, Eun Ju, Hyun Mi Shin, Eori Nam, Won Seog Kim, Ji-Hoon Kim, Byung-Ha Oh, and Yungdae Yun. 2012. “DeSUMOylating Isopeptidase: A Second Class of SUMO Protease.” *EMBO Reports* 13 (4): 339–46. doi:10.1038/embor.2012.3.

Stellfox, Madison E., Isaac K. Nardi, Christina M. Knippler, and Daniel R. Foltz. 2016. “Differential Binding Partners of the Mis18α/β YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres.” *Cell Reports*, May. doi:10.1016/j.celrep.2016.05.004.

Stoler, Sam, Kevin C. Keith, Kathie E. Curnick, and Molly Fitzgerald-Hayes. 1995. “A Mutation in CSE4, an Essential Gene Encoding a Novel Chromatin-Associated Protein in Yeast, Causes Chromosome Nondisjunction and Cell Cycle Arrest at Mitosis.” *Genes & Development* 9 (5): 573–586.

Strasburger, Eduard. 1884. “Die Controversen Der Indirecten Kerntheilung.” *Archiv Für Mikroskopische Anatomie*.

Sugata, Naoko, Munekata Eisuke, and Kazuo Todokoro. n.d. “Characterization of a Novel Kinetochore Protein, CENP-H.” http://www.jbc.org/content/274/39/27343.full.html.

Tachiwana, Hiroaki, Wataru Kagawa, Tatsuya Shiga, Akihisa Osakabe, Yuta Miya, Kengo Saito, Yoko Hayashi-Takanaka, et al. 2011. “Crystal Structure of the Human Centromeric Nucleosome Containing CENP-A.” *Nature* 476 (7359): 232–35. doi:10.1038/nature10258.

Tachiwana, Hiroaki, Sebastian Müller, Julia Blümer, Kerstin Klare, Andrea Musacchio, and Geneviève Almouzni. 2015. “HJURP Involvement in De Novo CenH3CENP-A and CENP-C Recruitment.” *Cell Reports* 11 (1): 22–32. doi:10.1016/j.celrep.2015.03.013.

Takeuchi, K., T. Nishino, K. Mayanagi, N. Horikoshi, A. Osakabe, H. Tachiwana, T. Hori, H. Kurumizaka, and T. Fukagawa. 2014. “The Centromeric Nucleosome-like CENP-T-W-S-X Complex Induces Positive Supercoils into DNA.” *Nucleic Acids Research* 42 (3): 1644–55. doi:10.1093/nar/gkt1124.

Tammsalu, Triin, Ivan Matic, Ellis G. Jaffray, Adel FM Ibrahim, Michael H. Tatham, and Ronald T. Hay. 2014. “Proteome-Wide Identification of SUMO2 Modification Sites.” *Science Signaling* 7 (323): rs2.

Tatham, M. H., E. Jaffray, O. A. Vaughan, J. M. P. Desterro, C. H. Botting, J. H. Naismith, and R. T. Hay. 2001. “Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9.” *Journal of Biological Chemistry* 276 (38): 35368–74. doi:10.1074/jbc.M104214200.

Thompson, Sarah L., Samuel F. Bakhoum, and Duane A. Compton. 2010. “Mechanisms of Chromosomal Instability.” *Current Biology* 20 (6): R285–95. doi:10.1016/j.cub.2010.01.034.

Torres, Eduardo M., Tanya Sokolsky, Cheryl M. Tucker, Leon Y. Chan, Monica Boselli, Maitreya J. Dunham, and Angelika Amon. 2007. “Effects of Aneuploidy on Cellular Physiology and Cell Division in Haploid Yeast.” *Science* 317 (5840): 916–924.

Verdin, Eric, and Melanie Ott. 2015. “50 Years of Protein Acetylation: From Gene Regulation to Epigenetics, Metabolism and beyond.” *Nature Reviews Molecular Cell Biology* 16 (4): 258–264.

Weaver, Beth AA, and Don W Cleveland. 2006. “Does Aneuploidy Cause Cancer?” *Current Opinion in Cell Biology* 18 (6): 658–67. doi:10.1016/j.ceb.2006.10.002.

Wei, Ronnie R, Jawdat Al-Bassam, and Stephen C Harrison. 2007. “The Ndc80/HEC1 Complex Is a Contact Point for Kinetochore-Microtubule Attachment.” *Nature Structural & Molecular Biology* 14 (1): 54–59. doi:10.1038/nsmb1186.

Welburn, Julie P.I., Mathijs Vleugel, Dan Liu, John R. Yates, Michael A. Lampson, Tatsuo Fukagawa, and Iain M. Cheeseman. 2010. “Aurora B Phosphorylates Spatially Distinct Targets to Differentially Regulate the Kinetochore-Microtubule Interface.” *Molecular Cell* 38 (3): 383–92. doi:10.1016/j.molcel.2010.02.034.

Westhorpe, F. G., C. J. Fuller, and A. F. Straight. 2015. “A Cell-Free CENP-A Assembly System Defines the Chromatin Requirements for Centromere Maintenance.” *The Journal of Cell Biology* 209 (6): 789–801. doi:10.1083/jcb.201503132.

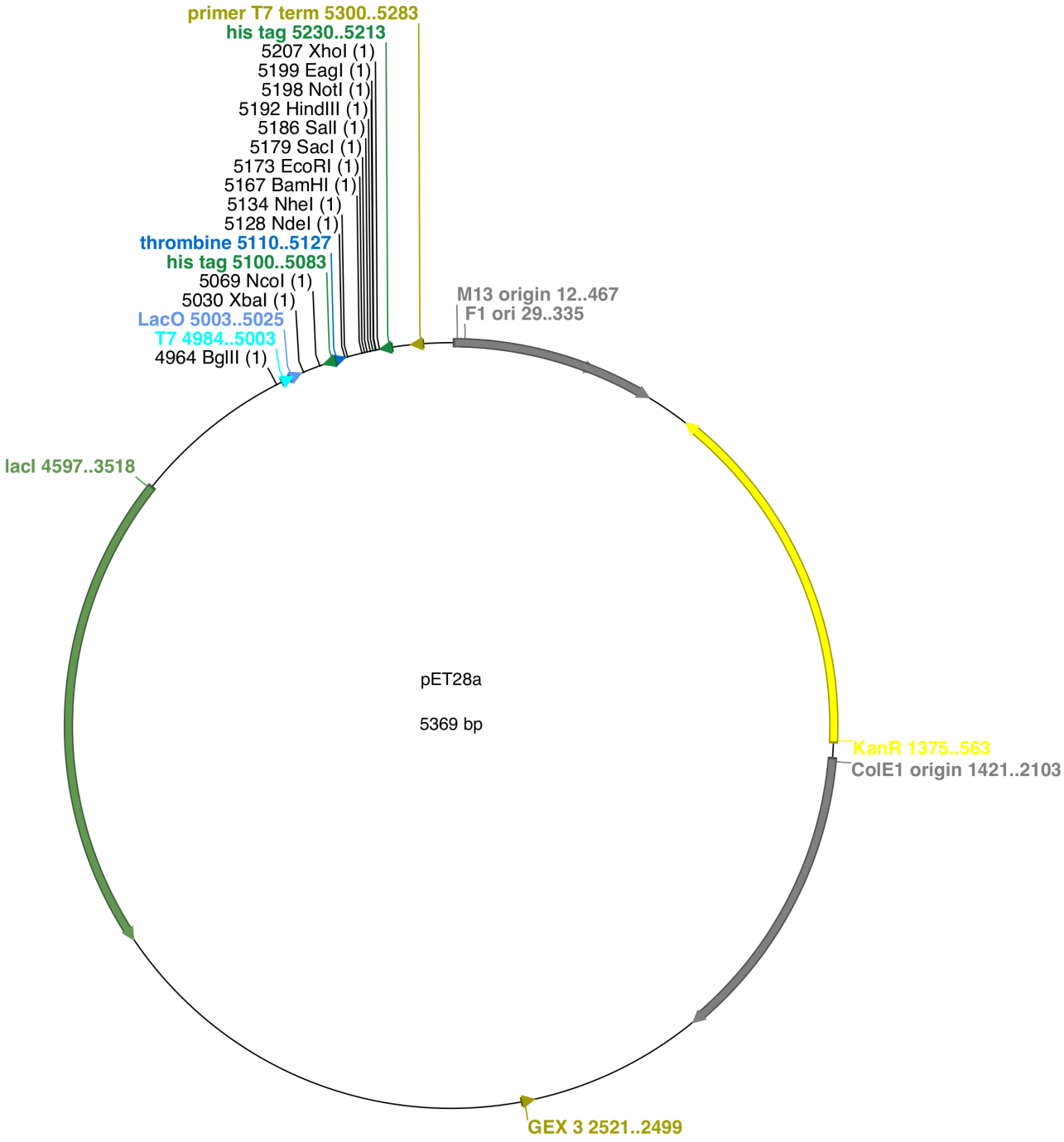
Wohlschlegel, J. A. 2004. “Global Analysis of Protein Sumoylation in Saccharomyces Cerevisiae.” *Journal of Biological Chemistry* 279 (44): 45662–68. doi:10.1074/jbc.M409203200.

Wu, Bing, Teng Leng Ooi, and Zijiang J. He. 2004. “Perceiving Distance Accurately by a Directional Process of Integrating Ground Information.” *Nature* 428 (6978): 73–77. doi:10.1038/nature02350.

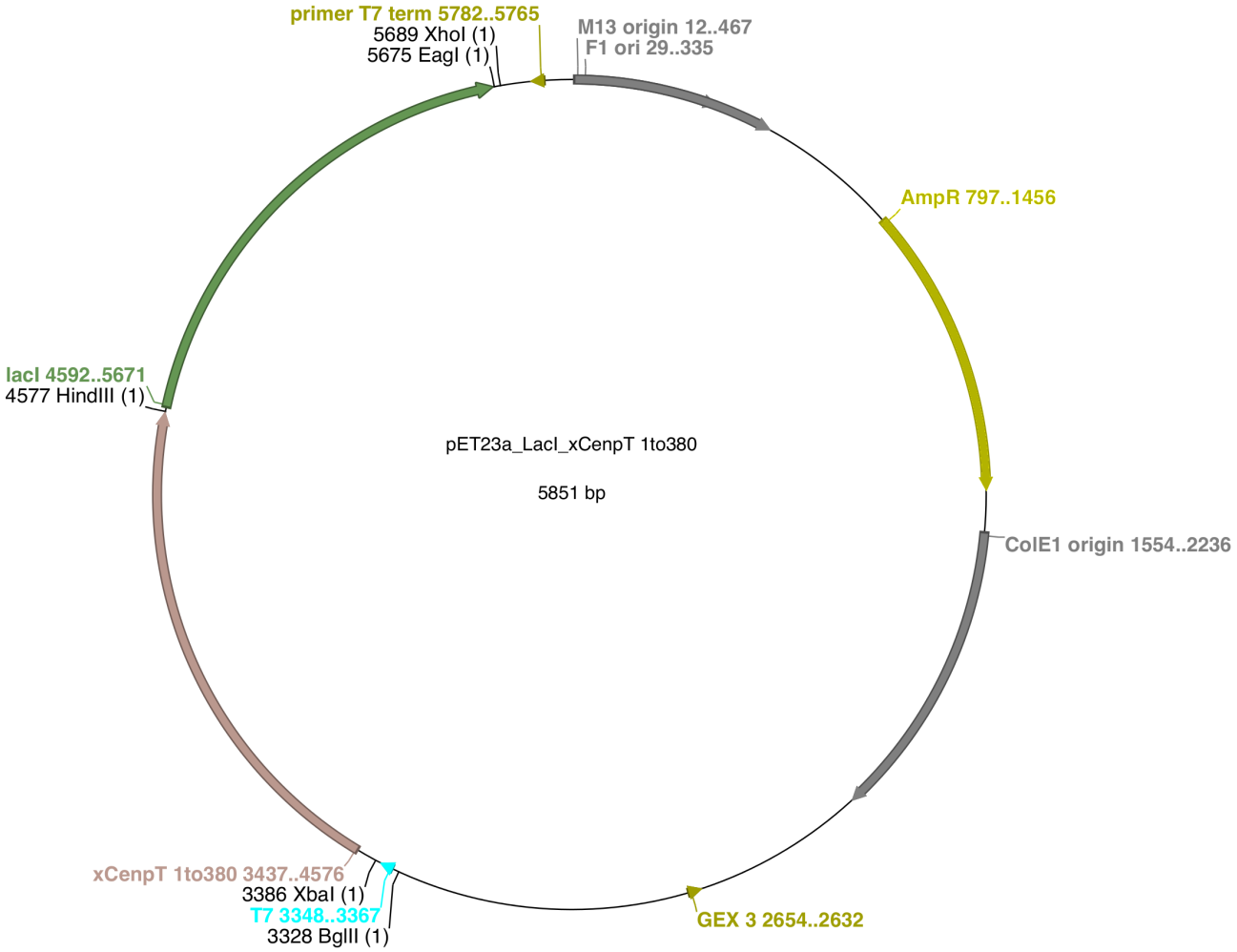
Zaytsev, Anatoly V., Jeanne E. Mick, Evgeny Maslennikov, Boris Nikashin, Jennifer G. DeLuca, and Ekaterina L. Grishchuk. 2015. “Multisite Phosphorylation of the NDC80 Complex Gradually Tunes Its Microtubule-Binding Affinity.” *Molecular Biology of the Cell* 26 (10): 1829–1844.

Supplementary

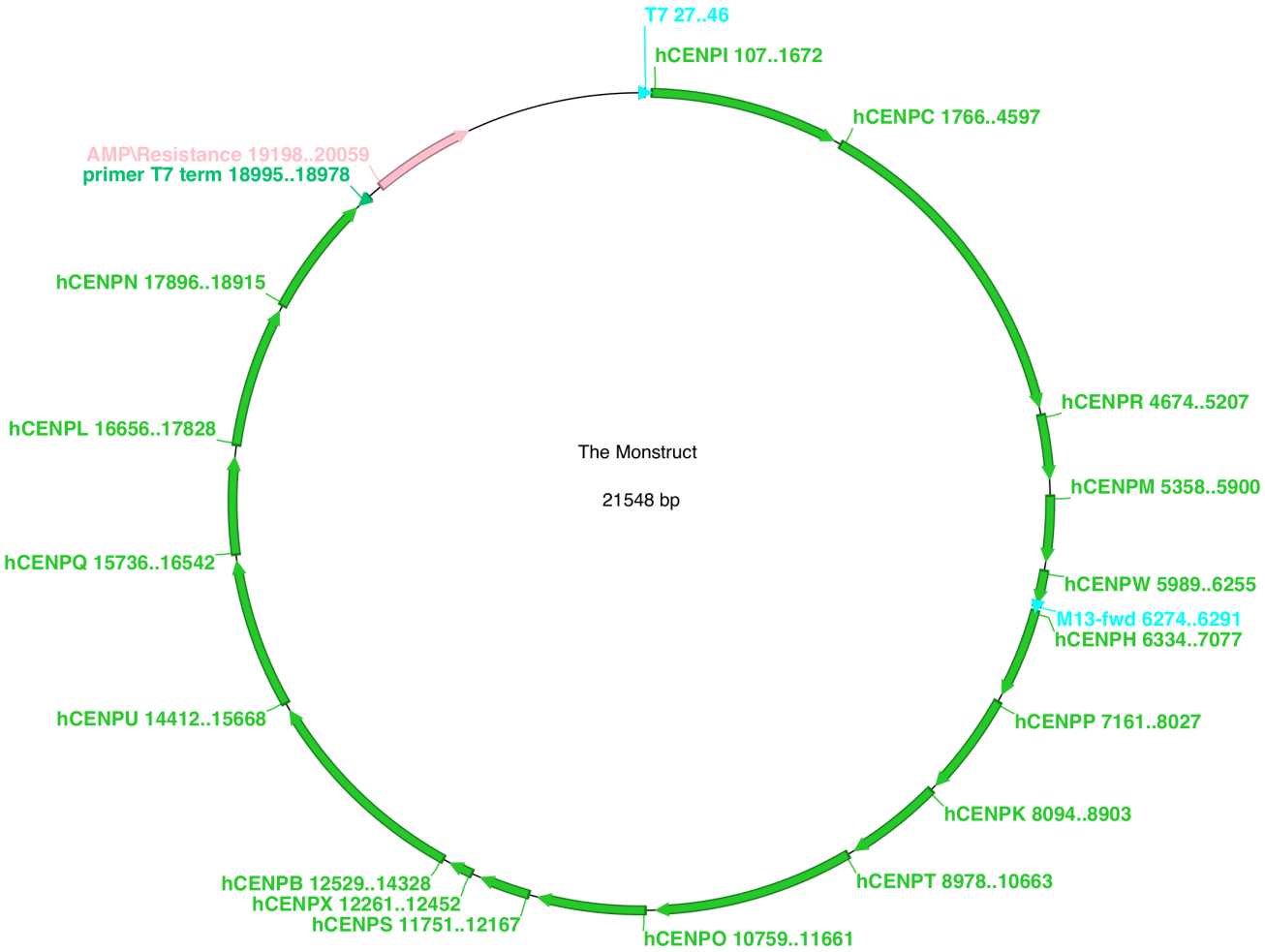
pET28a+


primer T7 term 5300..5283
his tag 5230..5213
5207 XhoI (1)
5199 EagI (1)
5198 NotI (1)
5192 HindIII (1)
5186 SalI (1)
5179 SacI (1)
5173 EcoRI (1)
5167 BamHI (1)
5134 NheI (1)
5128 NdeI (1)
thrombine 5110..5127
his tag 5100..5083
5069 NcoI (1)
5030 XbaI (1)
LacO 5003..5025
T7 4984..5003
4964 BglII (1)
M13 origin 12..467
F1 ori 29..335
lacI 4597..3518
pET28a
5369 bp
KanR 1375..563
ColE1 origin 1421..2103
GEX 3 2521..2499

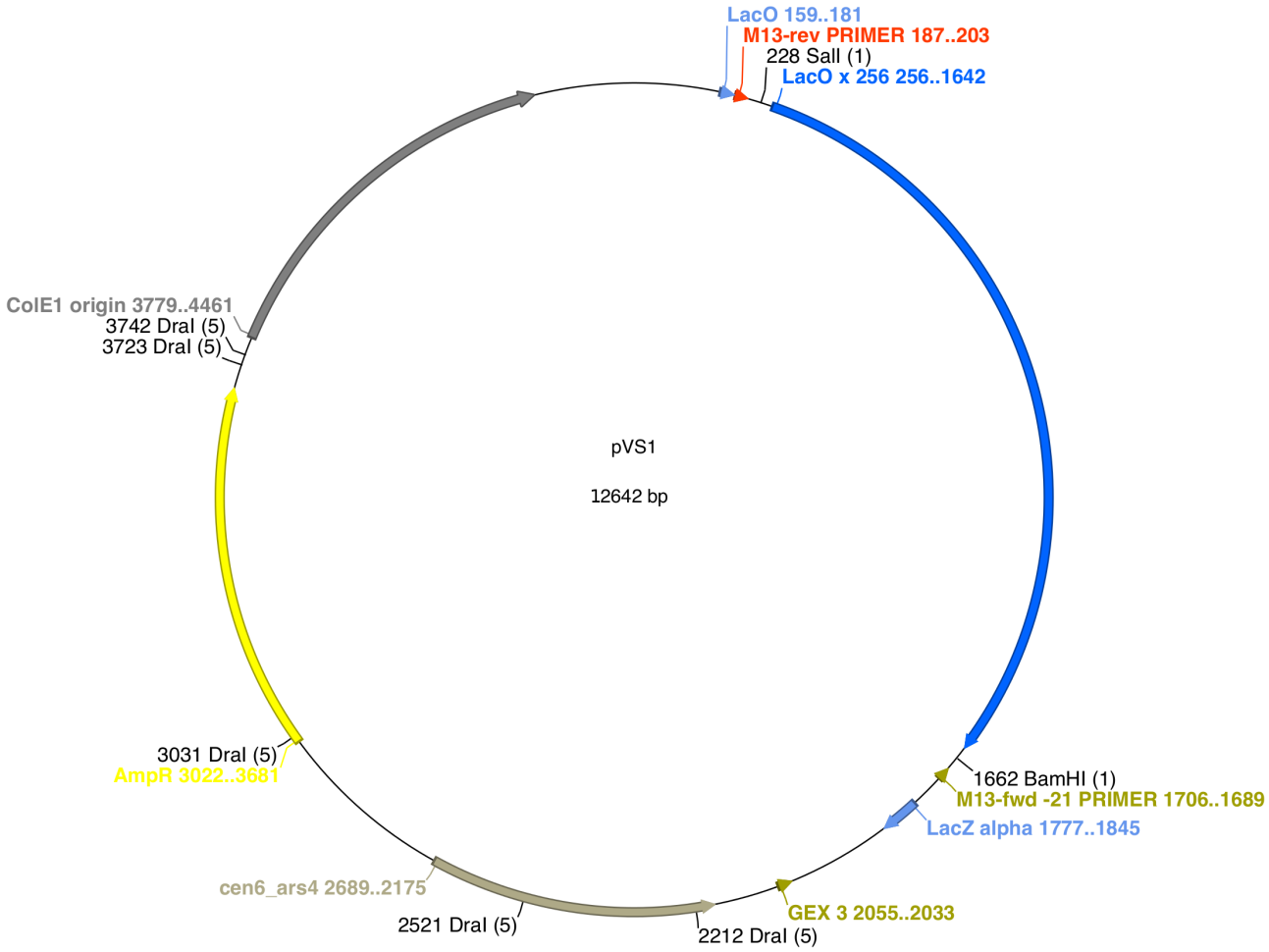
pET23-cenp-T/LacI



pST44 (monstruct)



pVS1


LacO 159..181
M13-rev PRIMER 187..203
228 SalI (1)
LacO x 256 256..1642
ColE1 origin 3779..4461
3742 DraI (5)
3723 DraI (5)
pVS1
12642 bp
3031 DraI (5)
AmpR 3022..3681
1662 BamHI (1)
M13-fwd -21 PRIMER 1706..1689
LacZ alpha 1777..1845
cen6_ars4 2689..2175
2521 DraI (5)
GEX 3 2055..2033
2212 DraI (5)

> SUMO2

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGS**M**ADERPREGVRTENNDHINLKVAGQDGSVVQFKIKRHTP
LSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG

> cenp-I

MADDQAEEDALQMAVGYFEKGPIKASQNKDKTLEKHLKTVENVAWKNGLASEEIDILLNIALSGKFGNAVNTRIL
KCMIPATVISEDSVVKAVSWLCVGKCSGSTKVLFYRWLVAMFDFIDRKEQINLLYGFFFASLQDDALCPYVCHLLY
LLTKKENVKPFRVRKLLDLQAKMGMQPHLQALLSLYKFFAPALISVSLPVRKKIYFKNSENLWKTALLAVKQRNR
GPSPEPLKLMLGPANVRPLKRKWNSLSVIPVLNSSSYTKECGKKEMSLSDCLNRSGSFPLEQLQSFPQLLQNIHCLEL
PSQMGSVLNNSLLLHYINCVRDEPVLLRFYYWLSQTLQEECIWYKVNNYEHGKEFTNFLDTIIRAECFLQEGFYSCE
AFLYKSLPLWDGLCCRSQFLQLVSWIPFSSFSEVKPLLFDHLAQLFFTSTIYFKCSVLQSLKELLQNWLLWLSMDIH
MKPVTNSPLETTLGGSMNSVSKLIHYVGWLSTTAMRLESNNTFLLHFILDFYEKVCDIYINYNLPLVVLFPPGIFYSA
LLSLDTSILNQLCFIMHRYRKNLTAAKKNELVQKTKSEFNFSSKTYQEFNHYLTSMVGCLWTSKPFGKGIYIDPEILE
KTGVAEYKNSLNVVHHPSFLSYAVSFLLQESPEERTVNVSSIRGKKWSWYLDYLFSQGLQGLKLFIRSSVHHSSIPR
AEGINCNNQY

> cenpH

MEEQPQMQDADEPADSGGEGRAGGPPQVAGAQAACSEDRMTLLLRLRAQTKQQLLEYKSMVDASEEKTPEQIMQ
EKQIEAKIEDLENEIEEVKVAFEIKKLALDRMRLSTALKKNLEKISRQSSVLMDNMKHLLELNKLIMKSQQESWDLE
EKLLDIRKKRLQLKQASESKLLEIQTEKNKQKIDLDSMENSERIKIIRQNLQMEIKITTVIQHVFQNLILGSKVNWAED
PALKEIVLQLEKNVDMM

The sequences of the proteins analysed by MS.

The sequences colored in green are the peptides identified by MS. The sequences in red are the identified SUMOylated lysines. The bold methionines are the initial methionine of the wild type proteins.