Reassembly and biochemical characterization of the human Smc5/6 complex.

Université de Montréal

Reconstitution et caractérisation biochimique du complexe humain Smc5/6

par Gustavo Segundo Cordero Guzmán

Institut de recherche en immunologie et en cancérologie (IRIC),
Faculté de médecine.

Mémoire présenté
en vue de l'obtention du grade de Master
en Biologie Moléculaire

30 août 2017

Résumé

Les membres de la famille de protéines de maintenance structurelle des chromosomes (SMC) sont des régulateurs essentiels de la stabilité génomique. Le complexe Smc5-6 est indispensable pour la réparation de l'ADN, la maintenance des télomères et le redémarrage des fourches de réplication bloquées. La façon dont le complexe Smc5-6 remplit ses fonctions pour favoriser la stabilité du génome est encore incertaine. Ici, nous avons développé une nouvelle stratégie de purification pour isoler un complexe réassemblé fonctionnel. Cette approche nous a permis d'effectuer d'importantes analyses biochimiques et structurelles du complexe humain Smc5-6. Nous montrons que le complexe humain se lie avec une affinité plus faible aux substrats ADNdb par rapport à ADNsb. Ce complexe a également été caractérisé par une séparation zonale en gradient continu pour déterminer la masse moléculaire du complexe heteropentamerique après filtration sur gel. Nous avons utilisé la méthode "Gradient Fixation" (GraFix) pour stabiliser le complexe Smc5/6 afin de visualiser celui-ci par microscopie électronique (EM). En conclusion, nous avons identifié avec succès les conditions natives pour la purification d'un complexe Smc5-6 humain entièrement assemblé et fonctionnel. En outre, nous avons montré que ce complexe est biochimiquement actif et lie le ADNsb et le ADNdb avec différentes affinités. L'achèvement de notre analyse structurale éclairera le mécanisme d'action du complexe Smc5-6 lors de la réparation de l'ADN.

Mots-clés : Complexe Smc5-6 humain, liaison ADN, entretien structurel du chromosome.

Abstract

Members of the structural maintenance of chromosomes (SMC) family of proteins are essential regulators of genomic stability. The Smc5-6 complex is indispensable for DNA repair, telomere maintenance and restart of stalled replication forks. How the Smc5-6 complex performs its functions to promote genome stability is still unclear. Here, we developed a novel purification strategy to isolate a reassembled complex. This approach allowed us to perform extensive biochemical and structural analyses of the human Smc5-6 complex. We show that the human complex binds with lower affinity to dsDNA substrates relative to ssDNA. This complex was also characterized by rate-zonal centrifugation to determine the molecular mass of the heteropentameric complex after gel filtration. We took advantage of the "Gradient Fixation" (GraFix) method to stabilize macromolecular complexes for single particle electron microscopy (EM). In conclusion, we have successfully identified native conditions for the purification of a fully assembled and functional human Smc5-6 complex. In addition, we showed that this complex is biochemically active and binds ssDNA and dsDNA with different affinities. Completion of our structural analysis will shed light on the mechanism of action of the Smc5-6 complex during DNA repair.

Keywords: Human Smc5-6 complex, DNA-binding, structural maintenance of chromosome.
.

Table des matières

Liste des tableaux

Table 1: A list of strains used in this study.

Table 2: A list of primers used in this study.

Liste des figures

Liste des sigles et abréviations

3D: Tridimensionnelle

ABC: «ATP binding casette»

ADN: Acide désoxyribonucléique

ADNsb (ssDNA): Acide désoxyribonucléique simple brin

ADNdb (dsDNA) : Acide désoxyribonucléique double brins

ALT: «Alternative lengthening of telomeres»

ATP: Adénosine triphosphate

bp: Paire de base

CDK: Protéine kinase dépendante d'une cycline

dHJ: «double-holliday junction»

DSB: Bris double brins

EMSA: «Electrophoretic mobility shift assay»

FL : «Full length»

HR: réparation par homologie de séquence

MAGE: «Melanoma antigen gene»

NHEJ : réparation par jonction de brins non homologues

Nse: «Non-SMC elements»

nts: nucléotides

OB fold: «Oligosaccharide/Oligonucleotide binding fold»

ORF: open reading frame

SMC: «Structural maintenance of chromosome»

SP-RING: «SIZ/PIAS RING»

SUMO: «Small ubiquitin like modifier»

Á Dieu, mes parents et Sandra

« En réalité, rien finisse, seulement c'est le principe d'une autre chose plus grand »

Gustavo Cordero, 2017

Remerciements

J'aimerais premièrement remercier mon directeur de recherche, Dr Damien D'Amours, de m'avoir donné l'opportunité de m'intégrer à son équipe de travail. Je vous remercie aussi pour tous vos conseils professionnels qui m'ont permis d'approfondir mes connaissances scientifiques et d'améliorer mes compétences professionnelles. De même, je profite de l'occasion pour dire un gros merci au Dr Diego Serrano pour ses recommandations tout au long de ma recherche, mais surtout pour sa contribution durant la révision de ma mémoire. Merci encore à vous deux de toujours me faire confiance.

Je voudrais remercier les Drs John Pascal, Marie Frances et Aleksandr Sverzhinsky pour toute leur collaboration dans le projet de recherche Smc5/6.

Des remerciements spéciaux à mon équipe de laboratoire, particulièrement à Mirela, Guillaume, Albert, Yoghitta, Dhana, Rim, Laurence et Roger.

Merci aux professeurs Dr Gerardo Ferbeyre et Dr James G Omichinski d'avoir accepté être membres du jury de ma thèse. J'ai beaucoup apprécié tous vos commentaires et suggestions.

Finalement, je profite de l'occasion pour remercier mes parents, ma famille et particulièrement ma Sandra pour leur compression et leur soutien en tout temps.

À tous ceux qui ont été avec moi.

Merci beaucoup.

1 Introduction

1.1 Les protéines de maintenance structurelle des chromosomes.

Les protéines de maintenance structurelle des chromosomes (*i.e., structural maintenance of chromosomes,* ou SMC) sont connues pour être essentielles dans l'assemblage ainsi que la stabilité structurelle des chromosomes, chez les humains et pratiquement tous les organismes vivants. Il s'agit d'une grande famille de protéines intervenant dans deux événements majeurs de l'organisation structurelle chromosomique requis pour la ségrégation mitotique des chromosomes : la condensation des chromosomes et la cohésion des chromatides sœurs. De plus, elles sont aussi impliquées dans la recombinaison, la réplication et la réparation de l'ADN (Hirano 2006).

Les protéines SMC partagent des homologies de séquence et une structure à cinq domaines, avec des domaines globulaires N- et C-terminaux séparés par un long segment de bobine enroulé (coiled-coil) (environ 100 nm ou 900 résidus) au centre duquel se trouve un domaine "charnière" globulaire (i.e., *hinge*). Toutes les protéines SMC semblent former des dimères, soit en formant des homodimères avec eux-mêmes ou des hétérodimères entre des protéines SMC différentes mais apparentées, ce qui donne au noyau du complexe Smc une symétrie caractéristique "à deux bras", architecture essentielle pour son fonctionnement dans les cellules (Lowe, Cordell et al. 2001). Par ailleurs, elles présentent une structure commune incluant un domaine ATPase composé de deux motifs « Walker A et B », qui participent à la liaison de l'ATP et du Mg2+, et deux régions α-hélicoïdales (Hirano 2006).

Ces protéines sont des ATPases chromosomiques qui contiennent entre 1 000 et 1 400 résidus d'acides aminés. Elles possèdent deux motifs canoniques « *Walker* A et B » (situés aux extrémités N- et C- terminaux de la protéine, respectivement), une paire de motifs de liaison nucléotidique couramment trouvée dans la plupart des protéines de liaison nucléotidique. Ces deux motifs forment une ATPase fonctionnelle de la famille ABC (cassette de liaison à l'ATP) qui constitue un domaine clé des protéines SMC. Des études de microscopie électronique ont établi que ces protéines se replient de manière antiparallèle entre les deux domaines principaux, en formant de longs motifs bobinés en spirale (*i.e.*, *coiled-coil*) séparés par une séquence non

hélicoïdale nommée les domaines de charnière (*i.e., hinge domain*). Les deux protéines SMC qui constituent le noyau du complexe interagissent au niveau de leur domaine d'articulation (Hirano 2006). Le domaine *coiled-coil* présente une flexibilité remarquable, probablement due à la présence de perturbations conservées, permettant aux complexes Smc d'adopter une grande variété de conformations. Par ailleurs, chaque complexe de protéine Smc contient des sous-unités non Smc essentielles à leur fonction, qui seront discutées plus bas dans cette révision.

Chez les eucaryotes, trois complexes protéiques Smc ont été identifiés : le complexe Smc5/6 et deux ensembles distincts d'hétérodimères Smc constituant les noyaux de deux complexes multiprotéiques appelés « cohésine » et « condensine », chacun spécialisé dans la cohésion et la condensation, respectivement. L'hétérodimère Smc1/Smc3 (noyau de la cohésine) permet entre autres de maintenir les chromatides sœurs ensemble après la réplication de l'ADN et au cours de la mitose, jusqu'à ce que la transition métaphase-anaphase ait lieu (Haering, Farcas et al. 2008). L'hétérodimère Smc2/Smc4 (noyau de la condensine) intervient dans la condensation et ainsi dans la ségrégation des chromosomes mitotiques et méiotiques (Stray, Crisona et al. 2005). Le troisième complexe, simplement appelé complexe Smc5/6, est le plus récemment identifié et semble être le plus grand de cette famille de protéines. Il est composé par l'hétérodimère Smc5/Smc6 et des sous-unités régulatrices portant le nom de « non-SMC elements » ou « Nse ». Même si la fonction précise du complexe Smc5/6 reste inconnue, il existe des preuves qu'il est impliqué dans la réparation de l'ADN (Lehmann 2005).

1.2 L'architecture générale du complexe Smc5/6

Le complexe humain Smc5/6 est formé par l'hétérodimère Smc5/Smc6, et quatre protéines régulatrices (non-SMC elements) : Nse1, Nse2, Nse3 et Nse4 (Taylor, Copsey et al. 2008). Par ailleurs, ce complexe est composé d'aux moins six sous-unités non Smc (Nse1 à Nse6) chez la levure boulangère, *Saccharomyces cerevisiae*. Les composants du complexe Smc5/6 sont conservés tout au long de l'évolution et requis pour la survie cellulaire. De même, les formes mutées des sous-unités rendent les cellules hypersensibles aux dommages à l'ADN induits par la lumière UV et l'irradiation ionisante (Hirano 2006).

En tant que protéines de maintenance structurelle des chromosomes, Smc5 et Smc6 partagent les mêmes caractéristiques de celles-ci qui ont été décrites plus en détail dans la section de cette

révision destinée à ce sujet (voir section « **Les protéines de maintenance structurelle des chromosomes** »). Les sous-unités régulatrices portant le nom de « non-SMC elements » ou « Nse » forment avec l'hétérodimère Smc5/Smc6 un complexe (Figure 1.1 A). De plus, l'interaction entre les deux protéines Smc à travers leurs régions charnières dans les hétérodimères est indispensable pour la fonction du complexe ainsi que la viabilité de la cellule. Des études montrent que la mutation d'une glycine conservée dans le domaine charnière de Smc6 perturbe l'interaction entre les protéines Smc5 et Smc6, ce qui confère une sensibilité à la température chez *Schizosaccharomyces pombe* (Sergeant, Taylor et al. 2005).

Selon les observations de Ampatzidou et al., Smc6 semble être impliqué dans le redémarrage des fourches de réplication effondrées ainsi qu'avoir des fonctions dépendantes et indépendantes de la *HR* au cours de la réplication de l'ADN. Lorsque les fourches de réplication s'effondrent, Smc6 est nécessaire pour la réparation efficace de l'ADN, mais pas pour le recrutement de protéines de recombinaison (Ampatzidou, Irmisch et al. 2006). Par ailleurs, cette protéine joue un rôle important dans la prévention des événements aberrants de recombinaison entre les régions péricentromériques au cours de la première prophase méiotique, qui autrement causeraient des aberrations chromosomiques menant la cellule à l'apoptose, l'arrêt méiotique ou les aneuploïdies (Verver, van Pelt et al. 2013).

La protéine régulatrice Nse1 contient un domaine structurel caractéristique des ubiquitine-ligases E3, indiquant qu'elle fonctionne comme celles-ci dans l'ubiquitylation. Il s'agit d'un motif de doigt RING sur la partie C-terminale de la protéine qui contient un domaine riche en cystéine et qui forme une structure croisée avec deux ions zinc (Borden 2000, Fujioka, Kimata et al. 2002). Bien que le motif de doigt RING de Nse1 ne soit pas impliqué dans l'interaction avec le complexe Smc5/6, il a été démontré que les mutations dans cette région affectent l'intégrité du complexe en causant une hypersensibilité à différents produits génotoxiques (Pebernard, Perry et al. 2008).
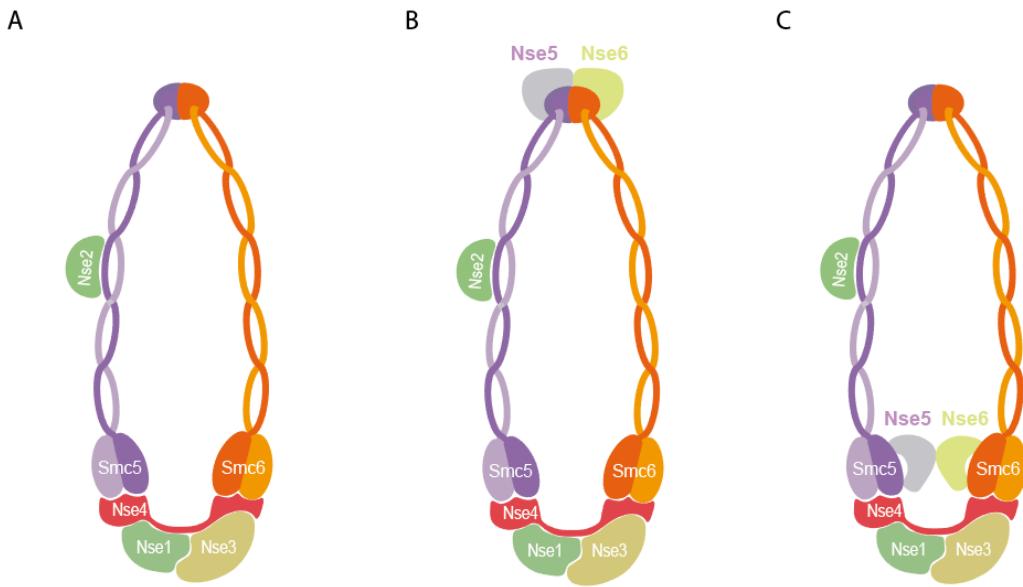
Figure 1.1 : Représentation des différents modèles du complexe Smc5/6 **A :** Chez l'humain, ce complexe est composé par deux sous-complexes : le premier est formé par les protéines Smc5, Smc6 et Nse2/Mms21 ; et le deuxième par les protéines Nse1-MAGEGI/Nse3-Nse4/EID. Nse2/Mms21 se lie au domaine coiled-coil de Smc5. L'interaction entre les protéines MAGE-G1/Nse3-Nse1 stabilise l'interaction avec Nse4/EID, ce dernier permet de maintenir ensemble les domaines ATPase des protéines Smc5 et Smc6. Chez *S. cerevisiae et chez S. pombe,* il existe un autre sous-complexe composé par les protéines Nse5 et Nse6. **B :** Ce sous-complexe est lié à la région charnière de l'hétérodimère Smc5/Smc6 chez *S. cerevisiae*. **C :** Tandis qu'il est plus proche des domaines ATPase des protéines Smc5 et Smc6 chez *S. pombe*.

Nse2, communément appelé Mms21, constitue aussi une sous-unité régulatrice du complexe Smc5/6. Il s'agit d'une SUMO ligase (*i.e., Small Ubiquitin-like Modifier ligase*) dont la partie N-terminale interagit avec le domaine « *coiled-coil* » de Smc5. Par ailleurs, le C-terminus de Nse2 présente un motif RING qui n'est pas nécessaire pour la liaison à Smc5, mais il est associé à son activité E3 SUMO ligase (Andrews, Palecek et al. 2005, Potts and Yu 2005). Même si la fonction précise de Nse2 n'est pas claire présentement, cette protéine peut fonctionner comme un commutateur moléculaire et aussi être impliquée dans la conformation structurelle du complexe Smc5/6. Grâce à son activité SUMO ligase, Nse2 est nécessaire pour résoudre les structures X des molécules d'ADN sur les fourches de réplication endommagées, dans la réparation de l'ADN et dans l'organisation de la chromatine. De plus, il a été démontré que

l'activité SUMO ligase de Nse2 n'est pas essentielle pour l'intégrité de Smc5 et donc, du complexe Smc5/6 (Zhao and Blobel 2005, Lee and O'Connell 2006).

La protéine Nse3 fait partie de la famille des protéines MAGE (*i.e., Melanoma Antigen Gene*) (Katsura and Satta 2011). Les membres de la superfamille MAGE partagent un domaine d'homologie MAGE (*i.e., MHD*) composé de deux sous-domaines à hélice ailée (WH / A et WH / B) (Chomez, De Backer et al. 2001). Chez l'homme, les gènes MAGE sont classés, sur la base de leur expression, soit comme gènes MAGE de type I ou de type II. Les protéines MAGE de type II sont relativement conservées ; elles sont exprimées dans les cellules humaines et jouent des rôles importants dans le retrait du cycle cellulaire, la différenciation neuronale et l'apoptose (Barker and Salehi 2002). La sous-unité MAGE-G1/Nse3 est présente dans le complexe humain Smc5/6 (Taylor, Copsey et al. 2008) quoique sa fonction n'est pas encore claire, toutefois elle semble être essentielle pour la structure et la conformation de ce complexe. Il a été démontré que la poche hydrophobe conservée dans le sous-domaine WH/B de Nse3 interagit avec la protéine Nse4/EID durant la formation du sous-complexe Nse1-Nse3-Nse4 (Hudson, Bednarova et al. 2011) ; tandis que le sous-domaine WH/A se lie à Nse1 (Hudson, Bednarova et al. 2011). Des études montrent que l'association de MAGE-G1/hNse3 avec hNse1 améliore l'activité ubiquitine ligase de hNse1. Cependant, aucun substrat pour l'activité de l'hNse3/MAGEG1-hNse1 E3-ubiquitine ligase n'a été rapporté jusqu'à présent. De façon similaire, il a été observé que MAGEA1 augmentait l'activité E3 ubiquitine ligase de TRIM31 (Kozakova, Vondrova et al. 2015). Dans le chapitre 1.4, je décrirai plus en détail les caractéristiques de Nse3 et de ses homologues MAGE.

Nse4 est un membre de la famille de protéines *E1A-like inhibitor of differentiation* (*EID*) (Bavner, Matthews et al. 2005), qui constitue la composante *kleisin* du complexe Smc5/6 ainsi que le pont reliant les têtes ATPases de Smc5 et Smc6 (Palecek, Vidot et al. 2006). Cette sous-unité régulatrice est connue pour être impliquée dans la réparation des aberrations chromosomiques apparaissant au cours de la progression normale du cycle cellulaire, spécifiquement dans la transition métaphase - anaphase (Hu, Liao et al. 2005). Chez l'humain, il existe deux homologues de la protéine Nse4 : hNse4a et hNse4b. Ces deux isoformes font partie du complexe humain Smc5/6, mais seulement hNSE4a est exprimé dans les cellules

somatiques, alors que hNSE4b est exprimé exclusivement dans les testicules (Taylor, Copsey et al. 2008).

Plusieurs études corroborent l'existence de deux sous-complexes Smc5-Smc6-Nse2 et Nse1-Nse3-Nse4 chez l'humain (Taylor, Copsey et al. 2008, Hudson, Bednarova et al. 2011). Bien que la morphologie réelle du complexe Smc5/6 reste inconnue jusqu'à présent, il existe des preuves montrant que le complexe peut adopter une structure en forme d'anneau. Comme dans les autres complexes Smc, les protéines Smc5 et Smc6 interagissent à travers leurs domaines charnière (Hirano 2006). Par ailleurs, les domaines ATPase de Smc5 et Smc6 sont reliés par Nse4, ce qui crée une forme annulaire, comme pour la cohésine, dont la fermeture est médiée par Scc1 (Palecek, Vidot et al. 2006). Finalement, la structure en forme d'anneau du complexe serait cohérente avec les études précédentes montrant que Smc5/6 maintient les chromatides sœurs ensemble au cours de la *HR* après la formation de *DSB* ou de fourches de réplication effondrées (Potts, Porteus et al. 2006, Kegel and Sjogren 2010).

1.3 Les sous-complexes Nse1/MAGEG1-Nse3/Nse4

L'étude faite par Hudson et all en 2011 (Hudson, Bednarova et al. 2011) révèle que la surface hydrophobe située sur le domaine C-terminal de Nse3 (aa200-307) interagit avec la partie N-terminale de Nse4 (aa1-150) ; cette dernière interagit en même temps avec le C-terminus de Nse1 (aa170-232). Par ailleurs, la région N-terminale de Nse3 (aa80-207) interagit avec le N-terminus de Nse1, ce qui indique que le motif doigt RING de Nse1 n'est pas donc impliqué dans aucune des interactions décrites ci-dessous. **(Figure 1.2)**.

La modélisation moléculaire de la sous-unité régulatrice Nse4, basée sur l'alignement de séquences avec Scc1 et sa structure cristalline, a montré l'existence des homologies élevées entre les structures Nse4 et Scc1 (**Figure 1.3**) Des observations confirment que Nse4 et Scc1 (sous-unité régulatrice de la cohésine) partagent un motif *Kleisin* hautement conservé qui est situé à leur partie C-terminale. Cette analogie structurale suggère que Nse4 se lie à l'extrémité N-terminale de la protéine Smc5, spécifiquement à travers le motif *Kleisin* (aa1011-1065) de sa région C-terminale, telle que l'interaction entre Scc1 et Smc1 (Haering, Schoffnegger et al. 2004, Palecek, Vidot et al. 2006) (**Figure 1.2**).
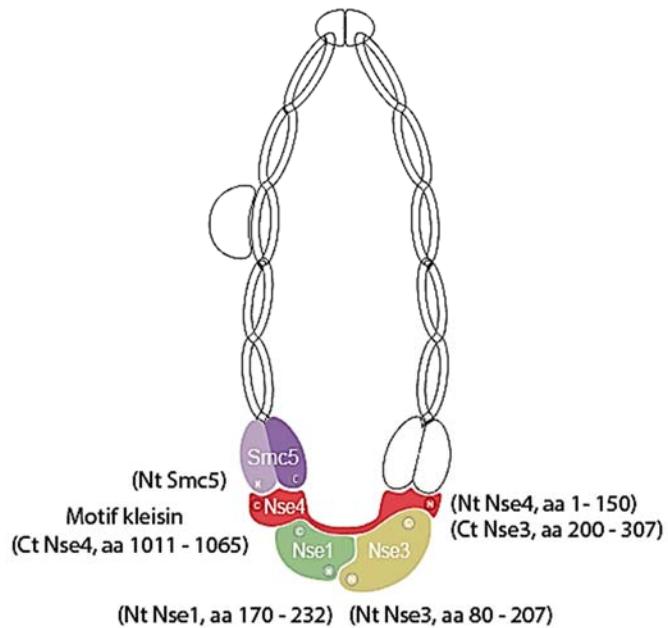
Figure 1.2 : Représentation des interactions parmi les « non-SMC elements » : Nse1, Nse3 et Nse4. Cette figure a été modifiée à partir de la représentation faite par Hudson et al en 2011 (Hudson, Bednarova et al. 2011). On a aussi représenté l'interaction entre le motif *Kleisin* de la sous-unité régulatrice Nse4 et le domaine ATPase dans la région N-terminale de la protéine Smc5.

Dans un essai de stabilité de protéine, lorsque Scc1 a été dégradé à un certain point, le niveau de la protéine Nse4 est resté constant durant la progression à travers le cycle cellulaire. Cela suggère que, dans le cas du complexe Smc5/6 avec une forme d'anneau, l'ouverture du complexe peut ne pas être régulée par la dégradation de Nse4 (Palecek, Vidot et al. 2006). On pourrait donc penser que l'activation du complexe Smc5/6 dépend des modifications faites sur la sous-unité régulatrice Nse4 ; telle la sumoylation décrite sur Scc1 lors que la cohésine est recrutée au *DSB* (Wu, Kong et al. 2012, Bermudez-Lopez and Aragon 2017).

En 2010, Doyle et al., ont résolu la structure cristalline du dimère de la protéine hNSE3/MAGEG1-hNSE1 (Doyle, Gao et al. 2010). Cette étude a révélé que hNse1 contient deux domaines à hélice ailée (WH-1/WH-2) dans sa région N-terminale, lesquels se trouvent liés au motif WH-A situé dans la partie N-terminale de MAGE-G1/hNSE3. Ils ont aussi découvert que cette interaction est établie grâce à la présence d'un motif conservé de deux

Leucines (L96 et L97) dans les protéines MAGE, faisant partie d'un noyau hydrophobe. Plus récemment, d'autres chercheurs ont observé que l'interaction parmi les sous-complexes Nse1-3-4 et l'ADN, est médiée par le domaine WH-B de la sous-unité Nse3, via des résidus basiques fortement conservés dans l'hélice H3. Par ailleurs, Nse1 peut également interagir avec l'ADN lorsqu'elle est jumelée à Nse3 ; alors que Nse4 peut être impliqué (directement ou indirectement) dans les interactions avec l'ADN potentiellement par un changement de conformation (Zabrady, Adamus et al. 2016).
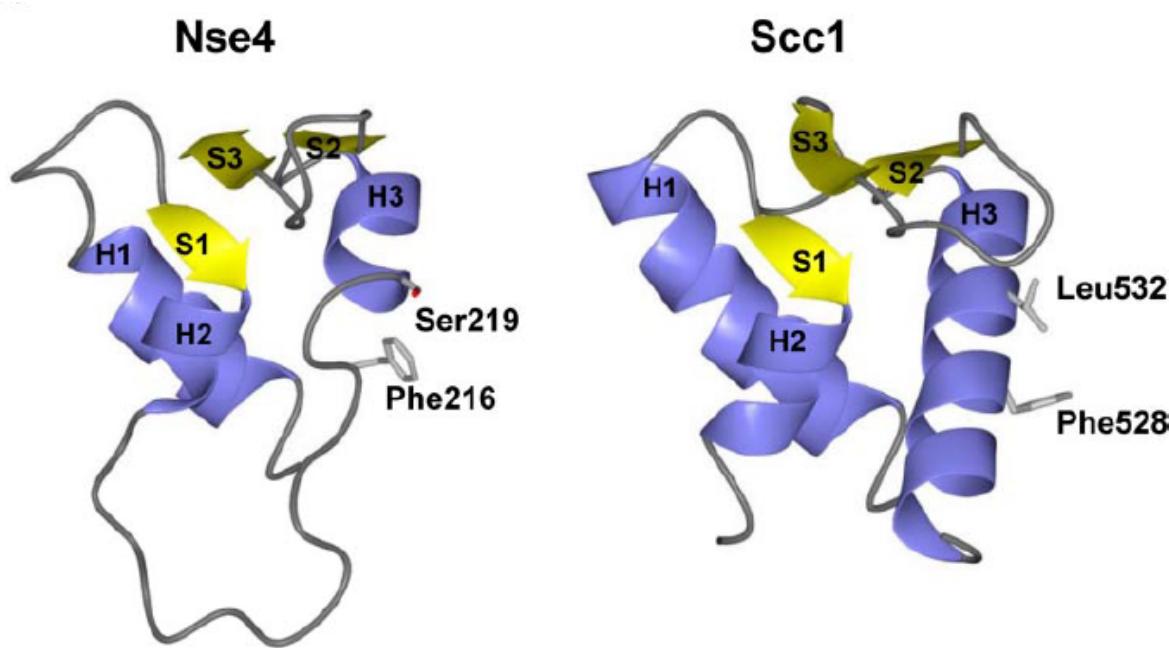
Figure 1.3 : Représentation de la structure cristalline de Scc1 et la structure prédite de l'extrémité C de Nse4. Cette figure a été modifiée à partir de la représentation faite par Palecek, Vidot et al en 2006 (Palecek, Vidot et al. 2006). Ces structures reperesentent le motif *kleisin* composé par les domaines (H1, H2, H3, S1, S2 et S3) qui font contacts physiques avec des résidus de la tête ATPase de Smc1 et Smc5 respectivement.

Les protéines MAGE sont exprimées dans une grande variété de cancers et aussi dans les cellules somatiques normales, mais MAGE-G1/hNSE3 est le seul membre de la superfamille MAGE qui fait une partie du complexe Smc5/6 (Taylor, Copsey et al. 2008). Les autres protéines MAGE montrent une grande diversité et, malheureusement, la pertinence de son évolution et de sa diversification en dehors du complexe Smc5/6 n'a pas été encore clarifiée (Katsura and Satta 2011). On abordera plus sur ce sujet dans la section à continuation.

1.4 Les classifications et les fonctions des protéines MAGEs.

La famille des gènes MAGE (*i.e., Melanoma Antigen Gene*) est composée de gènes qui partagent un domaine d'homologie MAGE d'environ 200 résidus d'acides aminés. Chez l'homme, cette famille contient 37 gènes qui encodent différentes protéines et qui sont classés, sur la base de leur expression, soit comme gènes MAGE de type I ou de type II. Les protéines MAGE partagent un domaine hautement conservé appelé le domaine d'homologie MAGE (*i.e., MHD*) (Chomez, De Backer et al. 2001) composé par deux sous-domaines, nommés WH/A et WH/B. Cette superfamille de protéines est connue par être impliquée dans la régulation du cycle cellulaire, dans la différenciation cellulaire et dans la croissance cellulaire (Barker and Salehi 2002). Les protéines MAGE de type I comprennent MAGE-A, MAGE-B et MAGE-C ; alors que celles de type II comprennent MAGE-D, MAGE-E, MAGE-F, MAGE-G, MAGE-H, les sous-familles MAGE-L et Necdin (Chomez, De Backer et al. 2001).

Même si la fonction des protéines MAGE de type I dans les cellules tumorales reste incertaine, il existe des études qui montrent qu'elles jouent un rôle actif dans la survie ou la mort des cellules cancéreuses. Il a été démontré que ces protéines sont aussi exprimées dans des cellules germinales mâles, mais malheureusement leur fonction est également inconnue. Cependant, les modèles d'expression suggèrent qu'elles peuvent être impliquées dans le développement et la prolifération cellulaire (De Smet, Lurquin et al. 1999).

La sous-famille MAGE-A appartient aux CTA (*cancer/testis antigens* en anglais) exprimés dans une grande variété de tumeurs malignes, mais pas dans les tissus adultes normaux, à l'exception du testicule (De Plaen, Arden et al. 1994). Des études récentes montrent que les protéines MAGE-A ont une interaction avec la région de la protéine p53 qui interagit avec l'ADN. Ces données suggèrent qu'elles peuvent inhiber p53 en interférant avec sa capacité à interagir avec la chromatine, et ainsi protéger les cellules cancéreuses de l'apoptose médiée par p53 (Marcar, Maclaine et al. 2010). Il a été observé que les protéines MAGE-B et MAGE-C interagissent et inhibent KAP1/TRIM28, un répresseur de la protéine p53 qui participe aussi à la survie des cellules tumorales (Yang, O'Herrin et al. 2007).

Plusieurs stratégies pour traiter les cancers sont en cours de développement et les protéines MAGE-As sont considérées comme des cibles attrayantes (Sang, Lian et al. 2011). Des études

où l'expression de l'antigène MAGE-A a été diminuée ont montré qu'il était dispensable pour le cycle cellulaire, mais nécessaire pour la survie des cellules de myélome en prolifération. La perte de MAGE-A conduit à une apoptose médiée par l'activation p53, ce qui confirme le rôle essentiel de MAGE-A dans la pathogenèse et la progression du myélome multiple en inhibant l'apoptose dans les cellules prolifératives de myélome (Nardiello, Jungbluth et al. 2011). De manière similaire, l'administration d'inhibiteurs de l'ADN méthyltransférase et de l'histone désacétylase conduit à l'expression plus élevée des protéines MAGEs en améliorant la réponse immunitaire. Cependant, l'utilisation de l'antigène MAGE dans l'immunothérapie est encore préliminaire (Meek and Marcar 2012).

Les protéines MAGE de type II sont presque omniprésentes, particulièrement dans les tissus adultes en cours de développement. Elles sont impliquées dans la régulation du cycle cellulaire, dans la croissance des cellules musculaires ou dans la différenciation cellulaire (Barker and Salehi 2002). Necdin et MAGE-D1 ont été les protéines MAGE de type II les mieux caractérisées jusqu'à présent (Kuwako, Taniura et al. 2004, Sasaki, Hinck et al. 2005). Des études indiquent que Necdin et MAGE-G1 partagent des caractéristiques biochimiques et fonctionnelles communes, ce qui suggère qu'elles agissent de manière complémentaire dans le développement du cerveau (Kuwako, Taniura et al. 2004). Il a été suggéré que Necdin interagit avec le facteur de transcription E2F1, en réprimant la transcription dépendante de ce régulateur majeur du cycle cellulaire et en supprimant la croissance spécifique aux neurones déficients en *Riboblastine* (Rb). La surexpression de Necdin et MAGE-G1 à proximité de la membrane plasmique réduit leur association avec E2F1. De plus, ces deux protéines interagissent avec le récepteur de la neurotrophine p75 via ses domaines intracellulaires (Taniura, Taniguchi et al. 1998).

Plus récemment, il a été observé que plusieurs protéines MAGE se lient aux protéines RING avec différentes affinités et spécificités. Ces complexes MAGE-RING sont localisés dans des compartiments subcellulaires distincts, tels que MAGE-B18-LNX1 dans le cytoplasme, MAGE-A2/C2-TRIM28 et MAGE-G1/hNse3-hNse1 dans le noyau, et MAGE-D1-Praja-1 au cytosol (Doyle, Gao et al. 2010). L'étude effectuée par Taylor, Copsey et collègues en 2008 a démontré que l'interaction hNse3/MAGEG1-hNSE1 est particulièrement intéressante, car elle a lieu dans le complexe humain Smc5/6 (Taylor, Copsey et al. 2008). Cette association améliore

l'activité ubiquitine ligase de hNSE1 in vitro (Doyle, Gao et al. 2010) et aussi stabilise le sous-complexe Nse1/3/4.

1.5 Le rôle du complexe Smc5/6 dans la stabilité du génome.

Bien que la fonction précise du complexe Smc5/6 ne soit pas claire, il est retrouvé chez pratiquement tous les eucaryotes de la levure à l'humain. Ce complexe joue un rôle essentiel dans la réparation des dommages dans l'ADN induits par des agents externes, en assurant l'intégrité génomique grâce à ses fonctions dans la ségrégation et la réparation des chromosomes (Lehmann 2005, Kegel and Sjogren 2010). La localisation chromosomique du complexe Smc5/6 de *Saccharomyces cerevisiae* révèle des rôles dans plusieurs mécanismes de duplication du génome. Le premier mécanisme contrôle l'association aux centromères et aux bras chromosomiques, le second régule l'association aux ruptures d'ADN et le troisième dirige le complexe vers le bras chromosomique qui héberge les répétitions d'ADN ribosomique. Dans l'ensemble, il a été rapporté que le complexe Smc5/6 est recruté pendant la réplication de l'ADN en phase S jusqu'à la mitose, afin d'empêcher l'accumulation de structures réplicatives toxiques pour la cellule (Lindroos, Strom et al. 2006).

En plus d'être impliqué dans la maintenance structurelle des chromosomes, il existe des indications que le complexe Smc5/6 possède d'autres fonctions qui le rendent indispensable pour la survie de la cellule. Par exemple, il est nécessaire pour maintenir l'arrêt du cycle cellulaire en présence de lésions dans l'ADN, et une défaillance de celle-ci conduit à des mitoses hautement aberrantes (Verkade, Bugg et al. 1999). Les observations confirment qu'il est aussi essentiel dans la prévention de la mauvaise ségrégation chromosomique au cours de l'anaphase (Bermudez-Lopez, Ceschia et al. 2010). Par ailleurs, le complexe est fortement exprimé dans les testicules chez les mammifères (Verver, Langedijk et al. 2014), et semble important pour la recombinaison chromosomique tout au long de la méiose (Copsey, Tang et al. 2013). De plus, il joue un rôle indispensable dans l'élimination des liaisons médiées par l'ADN afin de prévenir la mauvaise ségrégation des chromosomes, l'aneuploïdie et l'arrêt du cycle cellulaire en métaphase (Bermudez-Lopez, Ceschia et al. 2010).

DNA, les bris double brin
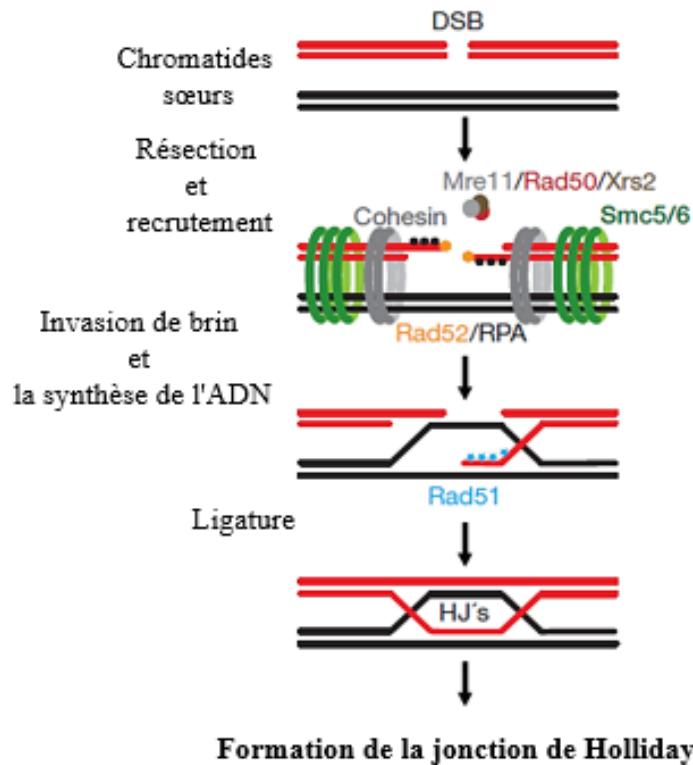
Figure 1.4 : Recrutement du complexe Scm5/6 et Smc1/3 (cohésine) sur les bris double brin (*DSBs,* abréviation anglaise) (Kegel and Sjogren 2010). La protéine de réplication A (*RPA*, abréviation anglaise) joue un rôle essentiel dans de nombreux aspects du métabolisme des acides nucléiques, y compris la réplication de l'ADN, la réparation de l'excision nucléotidique et la recombinaison homologue (Cox 2007). Par ailleurs, la protéine Rad52 est très importante dans la réparation des *DSBs* et dans la recombinaison homologue (*HR*, abréviation anglaise) ; elle se lie aux extrémités d'ADN monocaténaire et favorise le recrutement de Rad51. La formation de la jonction de Holliday (*HJ*, abréviation anglaise) garantie la réparation de l'ADN via la *HR*.

Également, le complexe Scm5/6 est important dans la stabilité ou la récupération des fourches de réplication bloquées ou effondrées. Dans ce contexte, il est requis dans le recrutement de la protéine de réplication A (*RPA*, abréviation anglaise) et Rad52 sur les les bris double brin (*DSBs,* abréviation anglaise) (Wu and Yu 2012, Lilienthal, Kanno et al. 2013). Une autre étude suggère que le complexe Smc5/6 favorise le recrutement de la cohésine (complexe Smc1/3) aux *DSBs*

afin de maintenir les chromatides sœurs à proximité l'une de l'autre, ce qui facilite la formation de la jonction de Holliday (*HJ*, abréviation anglaise) et la réparation de l'ADN par l'ensemble des protéines de la recombinaison homologue (protéines *HR)* (Potts, Porteus et al. 2006, Kegel and Sjogren 2010, McAleenan, Cordon-Preciado et al. 2012, Wu and Yu 2012, Bermudez-Lopez and Aragon 2017) (**Figure 1.4).** Même si le complexe Smc5/6 ne recrute pas directement les protéines *HR*, il semble crucial pour conserver la fourche de réplication bloquée dans une conformation permettant le recrutement de toutes les protéines *HR* ainsi que le redémarrage de la fourchette de réplication (Ampatzidou, Irmisch et al. 2006, Kegel and Sjogren 2010) (**Figure 1.5)**. En outre, les expériences effectuées avec des versions mutées de Smc6 indiquent que les cellules ne réussissent pas à recruter de protéines *HR* dans des fourches de réplication effondrées (Ampatzidou, Irmisch et al. 2006), induites par l'absence de régulateur du cycle cellulaire Cds1 (Miyabe, Morishita et al. 2009).
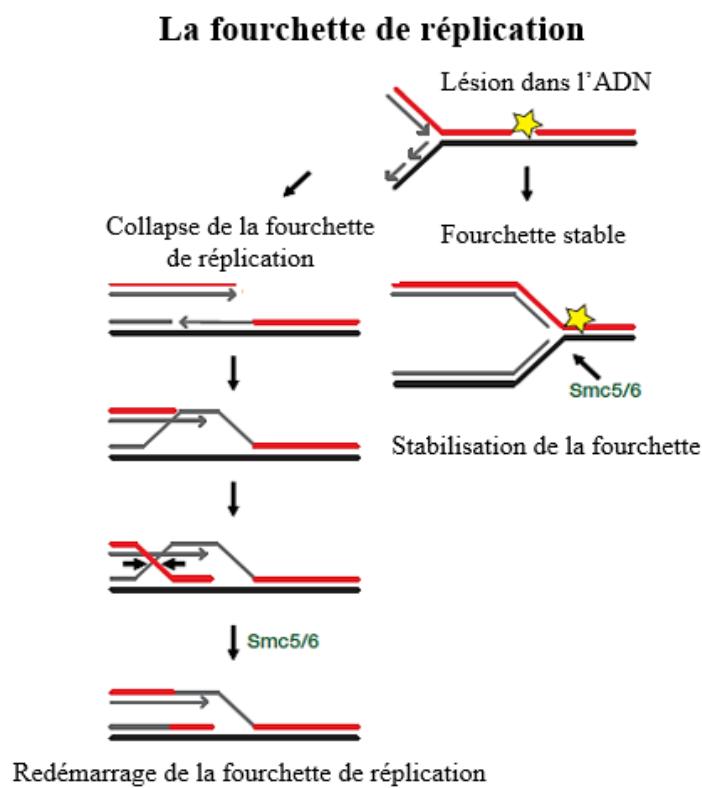
Figure 1.5 Le complexe Smc5/6 est nécessaire pour maintenir la fourche de réplication bloquée dans une conformation permettant le recrutement de toutes les protéines *HR* ainsi que le redémarrage de la fourchette de réplication (Kegel and Sjogren 2010).

Jusqu'à maintenant, le complexe Smc5/6 a été étudié principalement du point de vue phénotypique ; mais il reste encore à comprendre plus en détail ses propriétés biochimiques et avoir une vision plus claire de son mécanisme d'action moléculaire impliqué directement ou indirectement dans les processus énumérés ci-dessus. Plusieurs études ont été récemment focalisées sur la structure et les interactions du complexe Smc5/6 dans le but de fournir de nouvelles perspectives quant à sa fonction et ses mécanismes d'action moléculaires.

Contributions

Experiments were designed, and results were analyzed by Gustavo Cordero (GC), Dr. Diego Serrano (DS) and Dr. Damien D'Amours (DD). Yeast strains were created by GC and DS. Purification of Nse1 and Nse3 dimer was carried out by GC and DS. Subcomplex Nse1/Nse3/Nse4 was purified by GC. Purification of Nse4 under native conditions was performed by GC while denaturing conditions were optimized by DS. Fusion proteins were conceptualized by GC, DS and DD. Cloning strategy for Nse4-Smc5 fusion protein was designed and implemented by DS, whereas Smc6-Nse4 by GC. Smc5/6 complex (Nse4-Smc5 fusion) was purified by GC. DS purified complex carrying the Smc6-Nse4 fusion. TEV protease was purified by DS. DNA binding experiments were carried out by GC. Sucrose density gradient and GraFix were implemented by GC. Electron microscopy of the Smc5/6 complex was performed by Dr. Aleksandr Sverzhinsky.

Reassembly and biochemical characterization of the human Smc5/6 complex.

Gustavo Cordero, Diego Serrano and Damien D'Amours

Institute for Research in Immunology and Cancer (IRIC).

Faculty of Medicine, Université de Montréal.

P.O. Box 6128, Succursale Centre-Ville,

Montréal, QC, H3C 3J7

CANADA

Abstract

Members of the structural maintenance of chromosomes (SMC) family of proteins are essential regulators of genomic stability. The Smc5-6 complex is indispensable for DNA repair, telomere maintenance and restart of stalled replication forks. How the Smc5-6 complex performs its functions to promote genome stability is still unclear. Here, we developed a novel purification strategy to isolate a reassembled complex. This approach allowed us to perform extensive biochemical and structural analyses of the human Smc5-6 complex. We show that the human complex binds with lower affinity to dsDNA substrates relative to ssDNA. This complex was also characterized by rate-zonal centrifugation to determine the molecular mass of the heteropentameric complex after gel filtration. We took advantage of the "Gradient Fixation" (GraFix) method to stabilize macromolecular complexes for single particle electron microscopy (EM). In conclusion, we have successfully identified native conditions for the purification of a fully assembled and functional human Smc5-6 complex. In addition, we showed that this complex is biochemically active and binds ssDNA and dsDNA with different affinities. Completion of our structural analysis will shed light on the mechanism of action of the Smc5-6 complex during DNA repair.

2.1 Introduction

The structural maintenance of chromosomes (SMC) family of proteins plays a fundamental role in the integrity of our genome and in the repair of DNA. These proteins act within 3 complexes that are evolutionarily conserved among eukaryotes and play distinct roles in the chromosome biosynthesis, in the establishment of their "X" condensed structure during mitosis, as well as in the segregation of the genome during of cell division (Hirano 2006) and DNA damage repair. Specifically, the Smc1-3 complex (cohesin), is essential to maintain cohesion between sister chromatids in mitosis. The Smc2-4 complex (condensin) is responsible for the condensation of chromosomes in mitosis (Nasmyth and Haering 2005, Hirano 2006, Onn, Heidinger-Pauli et al. 2008). Finally, a complex composed of Smc5-6 proteins has recently been identified and its precise function in maintaining the integrity of chromosomes is still unknown (Kegel and Sjogren 2010).

Despite the fact that we do not know the precise role played by the Smc5-6 complex in the integrity of the genome, we know that its activity is essential for cell survival and that genetic inactivation of its components is associated with spraying of the chromosomes (Verver, Hwang et al. 2016). The main reason we do not know the mode of action of this complex comes from the fact that the physical and enzymatic properties of the complex are not completely understood. The organizational structure of the SMC proteins represents a first hint of the functions that Smc5-6 complex could potentially play *in vivo*. Specifically, SMC proteins share a common structure including an ATPase domain composed of two "Walker A and B" motifs (located at the N- and C-terminal ends of the protein) that participate in the hydrolysis of ATP, and two α-helical regions separated by a hinge domain. In addition, SMC proteins dimerize via their hinge domain serving as a scaffold for the attachment of various regulatory subunits, called non-Smc elements, or NSE (Fousteri and Lehmann 2000, Sergeant, Taylor et al. 2005). We can deduce from these observations that the ATPase activity of the complex probably plays a key role in the performance of its functions. In addition, the fact that the complex acts directly on the integrity of the chromosomes, it also suggests that the Smc5-6 components contain DNA and / or chromatin binding domains. This observation is consistent with the involvement of the Smc5-6 complex in some DNA repair pathways (double-strand breaks, or DSB) (De Piccoli, Cortes-Ledesma et al. 2006, Murray and Carr 2008, Sollier, Driscoll et al. 2009). However, the role played by Smc5-6 ATPase activity in DNA binding and DSB repair is still unknown.

Another indication of the potential role played by the Smc5-6 complex in chromosome integrity comes from the regulatory NSE subunits of the complex. Indeed, the human complex Smc5/6 contains four NSE subunits (Nse1, Mms21 / Nse2, Nse3 and Nse4). Nse1 is a "RING finger" type protein with ubiquitin ligase activity. Nse3 is linked to the family of MAGE proteins (antigen associated with melanomas). The Nse4 protein is the component that bridges the ATPase domains of Smc5 and Smc6 to the rest of the NSE proteins. Finally, the Nse2 / Mms21 subunit associates with the α-helical regions of Smc5. This protein has a "Sumo ligase" domain which facilitates the addition of the small Sumo modifying protein to the target proteins (Kegel and Sjogren 2010, Wu and Yu 2012, Verver, Hwang et al. 2015). It has been observed that Nse2/Mms21 is essential for the alternative lengthening of the telomeres (ALT) (Noel and Wellinger 2011). We can deduce from these observations that the Smc5-6 complex acts in part by the post-translational modification of proteins, specifically by ubiquitylation and sumoylation. How these activities of NSE regulatory proteins contribute to the overall function of the Smc5-6 complex is still unknown to date.

Hypothesis and specific objectives: A careful biochemical characterization of the human Smc5-6 complex will serve to uncover its mechanisms of action. For this reason, the main objective of my master thesis is to purify the human Smc5-6 complex, to characterize its enzymatic actives, and to analyse the structure of the complex. With this information, we will help to understand the molecular mechanisms used by the Smc5-6 complex to promote DNA repair.

Specific objectives:

1. Reassemble the human Smc5-6 complex.
2. Study the structure of the human Smc5-6 complex.
3. Analyse the DNA binding properties of human Smc5-6 complex.

2.2 Materials and Methods

2.2.1 *Saccharomyces cerevisiae* strains growth conditions and media.

All yeast strains used in this study are derivatives of W303. *S. cerevisiae* strain Table 1.

Table 1: A list of strains used in this study.

Strain	Genotype	Reference
D5462	*MATa ade2-1 can1-100 his3-11,112 trp1-1 ura3-1 (PSI+) rad5-535 lys2::Pgal1:GAL4::lys2 cir+ tdh3::KanMX6 (2u HIS3 Pgal-hNSE4a)*	This study
D5735	*MATa ade2-1 can1-100 his3-11,112 leu2-3,112 trp1-1 ura3-1 (PSI+) rad5-535 lys2::Pgal1:GAL4::lys2 cir+ tdh3::KanMX6 (2u HIS3 Pgal-hNSE4a) (URA3 LEU2 Pgal hSMC5 hSMC6)*	This study
D5819	*MATa ade2-1 can1-100 his3-11,112 leu2-3,112 trp1-1 ura3-1 (PSI+) rad5-535 lys2::Pgal1:GAL4::lys2 cir+ tdh3::KanMX6 (LEU2 URA3 Pgal hNSE4a-hSMC5 hSMC6) (LEU2 TRP1 Pgal-hNSE1 hNSE3)*	This study

2.2.1.1 Media

Yeast media were prepared according to standard recipes. Yeast extract peptone medium (YP; 1% yeast extract, 2% peptone) or synthetic complete minimal dropout medium (SC; 0.17% yeast nitrogen base without amino-acids and $(NH4)_2SO_4$ (YNB-AA/AS), 0.5% $(NH4)_2SO_4$, 0.13% dropout) were used. Solid plates were prepared by adding powder agar to a final concentration of 2%. Media were autoclaved at 121ºC for 30 min before use. Yeast strains were grown on different carbon sources, depending on the experimental requirements: glucose (gluc), galactose (gal), raffinose (raff) or glycerol and lactic acid.

2.2.2 Protein electrophoresis

Denaturing electrophoreses were performed on Biorad Criterion™ XT Bis-Tris Precast Gels. Samples were boiled 5 minutes at 100 ° C in the loading buffer (1% SDS, 2.5% β-mercaptoethanol [2-βME], 20% glycerol, 0.0025% blue bromophenol, 125 mM Tris-HCl, pH 6.8). Gels were stained with 0.1% Coomassie Blue R250 in a 10% acetic acid, 50% methanol, and 40% H_2O solution for 10 min. Gels were distained by shaking in a rotating platform for at least 2 to 4 hours in 20% methanol and 80% H_2O with at least two changes of this solvent.

2.2.3 Polymerase chain reaction (PCR)

The High Fidelity PFU PCR System was generally used. 2.6 U (1 μl) enzyme per 50 μl reaction were used with, 200 μM dNTPs and 0.2 μM primers. The PCR used were: initial denaturation

step of 95 ºC for 2 min; followed by denaturalization at 95 ºC for 30 sec, annealing at 52-65 ºC (depending on the melting temperature of the primers) for 30 sec, 1 minute/kb either at 72 ºC (for PCR products up to 3 kb) or at 68 ºC (for PCR products larger than 3 kb. This cycle was repeated 29 times. A final elongation step was performed at 72 ºC for 10 min.

2.2.4 Site-directed mutagenesis

Q5$^®$ Site-Directed Mutagenesis Kit (from Agilent) was used to perform site-directed mutagenesis. Briefly, the desired mutation was incorporated into primers (designed using the DNA 2.0 website). PCR was performed using the standard reaction mix containing 50 ng of template DNA. The thermal cycles used were: initial denaturation step of 95 ºC for 30 sec; followed by amplification cycles at 95 ºC for 1 min (denaturation), 55 ºC for 1 min (annealing), 1 min/kb at 65 ºC (elongation) repeated 29 times. Following PCR amplification, parental DNA (not containing the mutation) was digested by the addition of 1 µl DpnI (10 U/µl), an enzyme that specifically digests methylated and hemimethylated DNA, and incubated for 1 hours at 37 ºC. The nicked vector containing the desired mutation was then transformed into *Escherichia coli* 10X Gold. 3 µl of the digested PCR mix were added to 45 µl of competent cells that had been previously treated with 2-βME. The mixture was incubated on ice for 30 min, and then heat shocked at 42 ºC for 30 sec. Afterwards, cells were immediately chilled on ice for 2 minutes before being resuspended in 150 µl of LB media and allowed to grow for 1 hour at 37 ºC. Finally, 10 µl, 30 µl and 150 µl of cells were plated in 3 different LB plates containing the specific antibiotic for each plasmid. The cells were allowed to grow overnight at 37 ºC.

2.2.5 Reassembling of human subcomplex Nse1, MAGE-G1/Nse3 and Nse4a.

2.2.5.1 Expression and purification of human Nse1 and Nse3 dimer.

The bacterial strain *Escherichia coli* Rosetta BL21, designed to enhance the expression of eukaryotic proteins, was used as the host for recombinant plasmids p1334, kindly donated by Doyle's lab, (Doyle, Gao et al. 2010) carrying the genes encoding human *NSE1* and human *MAGE-G1/NSE3*. The transformed cell was cultured in LB (containing yeast extract 5 g/L, tryptone 10 g/L, NaCl 10 g/L), supplemented with 30 µg/mL kanamycin and 30 µg/mL Cloramphenicol. The cells were grown at 37 °C up to OD_{600} ~ 0.5 and protein overexpression was induced by 1 mM isopropoyl-thiogalactopyranoside (IPTG) for 4 h at 22 °C. The cells were harvested and lysed by sonication in Lysis buffer containing 50 mM K_2HPO_4, 50 mM Tris-HCl

pH 8, 150 mM NaCl, 10 % Glycerol, 0.2% Triton X-100, 2 mM [2-βME], lysozyme and 10mM imidazole. The lysate was centrifuged at 16 500 G for 15 min at 4 °C. The proteins were then incubated with Ni-NTA agarose resin (Qiagen). Finally, the human His-Nse1 tagged and Nse3 were eluted using 25 mM K_2HPO_4, 200 mM NaCl, 10% Glycerol, 0.1% Tween 20, 2 mM [2-βME] buffer supplemented with 300 mM imidazole.

Eluted protein (Nse1 and Nse3) from the Ni-NTA column was diluted 10-fold with CE buffer (50 mM K_2HPO_4/KH_2PO_4, pH 6.5, 10% glycerol, 0.2% Tween 20, 2 mM [2-βME]) and they were loaded on an SP-Sepharose FF 5mL column® (GE Healthcare), washed with 10 CVs of (CE buffer supplemented with 50mM NaCl) and eluted with a linear gradient of 10 CVs of CE buffer and CE buffer with 1 M NaCl). The protein eluted in a fraction containing approximately 230 mM NaCl. The fractions containing the purified proteins were concentrated to ∼0.5 mg/mL with Amicon Ultra filtration units (10K NMWL; Millipore), frozen on dry ice and stored at -80 °C.

2.2.5.2 Purification of human Nse4a under native conditions.

Human Nse4a was overexpressed in strain yeast D5462, transformed with plasmid p1309 using the LiAc/SS carrier DNA/PEG method (Gietz and Schiestl 2007). Yeast cultures were grown in 6 L in YED-Raf at 30 °C to an A_{600} of 0.6 OD and induced overnight with galactose (2 % final) at 18 °C. The lysis of human Nse4a- overexpressing cell was performed by 3 cycles of freezer mill. The broken cells were washed two time with 25 mL of wash pellet buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 % glycerol supplemented with protease inhibitors (EASBF, E64 and PEPA). Lysate was then centrifuged at 50 000 G for 30 min at 4 °C. Finally, the pellet was resuspended in solubilization buffer (SB) (50 mM K_2HPO_4/ KH_2PO_4 pH 7.0, 1 M NaCl, 10 % glycerol, 1 % triton X100, 2 mM [BME] and 20 mM imidazole) and it was centrifuged at 16 500 G for 15 min at 4°C. The human Nse4a was immobilised in HisTrap HP 1 ml column (GE Amersham). The column was washed with 10 column volumes (CVs) of Ni-NTA buffer (NB) (50 mM K_2HPO_4/ KH_2PO_4 pH 7.0, 1 M NaCl, 10% glycerol, 2 mM [2-βME]) supplemented with 100 mM imidazole. The human Nse1 and Nse3 proteins from (2.2.5.1) were diluted three time in 25 mM K_2HPO_4/ KH_2PO_4, 200 mM NaCl, 10% Glycerol, 0.1 % Tween 20, 2 mM 2BME and in was passed at 0.3 mL/min through the column with immobilized the human Nse4a. The column was washed again with 10 CVs of NB supplemented with 100 mM imidazole. The subcomplex formed by human Nse1, Nse3 and

human Nse4a was eluted by an imidazole gradient (0.1 M to 1 M) in NB. Elutions were concentrated by ultrafiltration using Amicon Ultra filtration units (50 K NMWL; Millipore). In order, to isolate stoichiometric Nse1/Nse3/Nse4a subcomplex, we performed size exclusion chromatography on a Superose 6 10/300 column (GE Healthcare) in NB. All the samples were analyzed by electrophoresis on Bio-Rad® Polyacrylamide gel Criterion™ XT Bis-Tris Precast Gels Biorad electrophoresis.

2.2.5.3 Glyceraldehyde-3-phosphate dehydrogenase (TDH3) gene deletion.

The glyceraldehyde-3-phosphate dehydrogenase (Tdh3) was found as an important contaminant in the purification of Nse4a (2.2.5.2). Insertions of epitope tags were performed as described by Janke et al. (Janke, Magiera et al. 2004): Briefly, cassettes that confers resistance to G418, and an appropriate epitope tag, were amplified by PCR from the plasmid pFA6a-   using primers (3029 and 3030) that will direct integration by homologous recombination at a *TDH3* locus, see below for list of primers Table 2. The deletions of the *TDH3* gene and the correct insert orientation of *KAN-MX6* cassette were verified by PCR. PCR amplification of the plasmid using an insert specific primer (3031) paired with a vector specific primer (141) were designed to produce an amplicon of a specific size only if the insert is in the correct orientation. In all experimental designs, presence or absence of a PCR amplicon and size of the product are determined by electrophoresis alongside a DNA size marker on an agarose gel. Finally, the two positives yeast colonies *TDH3Δ* were selected from G418 plates and store at -80 °C in 15% glycerol.

Table 2: A list of primers used in this study.

Number	Description	Sequence
3029	TDH3 U1 5'- 3'	**CAAGAACTTAGTTTCGAATAAACACACATAAACAAACAA**ACGGATCCCCGG
3030	TDH3 L1 5'- 3'	**AAGAAAATTTATTTAAATGCAAGATTTAAAGTAAATTCA**CGAATTCGAGCTCGTTTAAAC
3031	Upstream TDH3	GGTTGAAACCAGTTCCCTGA
141	pFA6-L3	CAAATGACAAGTTCTTGAAAACAAGAATC
3086	hSMC6_XmaI	GTCGTTGCTAATTCTCTGATTGATATGCCCGGGATTGAAACAGTTTTGTTGATCAAAAAC

3037	SMC6-NSE4 AmpR L	CAGTGTTTTGCATGCGTGCAGGTGA
2842	SMC5_NSE4a	AGATGTCTTTTTAGATGGTG
2835	hNSE4a AgeI	CACCACCACCACCGGTAGCTGAAGGCTTTTGTTGTCTCTGG
3087	hNSE4a_SphI	ttgccaagGCATGCCGGAGGGGGCACAGGAGGCGGTACC
2836	hNSE4a KpnI	CGGAGGGGGCACAGGAGGCGGTACC
2673	hSMC5/6 seq 7	CTCTTTTCGGTAGCTTCCAAAAGTT

2.2.6 Expression and purification of human Smc5, Smc6 and Nse4a.

Smc5, Smc6 and Nsea4a were co-expressed in strain yeast D5735 transformed with plasmid p1479 using the LiAc/SS carrier DNA/PEG method (Gietz and Schiestl 2007). Yeast cultures were grown in 6 L in YED-Raf at 30 °C to an A_{600} of 0.6 OD and induced overnight with galactose (2 % final) at 18 °C. Human Smc5, Smc6 and Nse4a- overexpressing cells were lysed and it were resuspended in lysis buffer (LyB) 50 mM K_2HPO_4/ KH_2PO_4 pH 8, 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 % glycerol, 0.5 % triton X100, 2 mM [BME] supplemented with 20 mM imidazole as well as phosphatase inhibitors and protease inhibitors (EASBF, E64 and PEPA). The sample was loaded onto a HisTrap HP 5mL column. After wash in 10 CVs of LyB supplemented with 50 mM imidazole, the human Smc5, Smc6 and Nse4a trimer was eluted in Strep Buffer (StB) 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 % glycerol, 0.5 % tween20, 2 mM [BME] supplemented with 500 mM imidazole. The elution was then loaded on a 1mL Strep-Tactin II column (GE Amersham) and washed with 10 CVs of StB supplemented with 0.5% Triton. The trimer was eluted with 5 CVs of S6-buffer 50 mM K_2HPO_4/ KH_2PO_4 pH 8.0, 1 M NaCl, 10 % glycerol, 1 mM EDTA and 2 mM [2-BME] supplemented with 20 mM of desthiobiotin. Finally, the trimer complex was concentrated by ultrafiltration using Amicon Ultra filtration units (100K NMWL; Millipore) and further purified by size exclusion chromatography on a 10/300 Superose 6 column (GE Healthcare) in S6-buffer.

2.2.7 Reassembling of human Smc5/6 complex.

2.2.7.1 Plasmids for human fusion Nse4a-Smc5 and Smc6 overexpression

The human Nse4a and Smc5 proteins were co-expressed as fusion protein to obtain a ratio 1:1:1 between Nse4a, Smc5 and Smc6. The carboxy-terminus (Ct) of Nse4a was connected to amino-terminus (Nt) of Smc5 using a flexible linker as done described for Scc1-Smc1 (Gruber, Arumugam et al. 2006). In order, to purify these three proteins together, 6xHistidine-tag was added to the amino-terminus of human Nse4a and 3xStrep-tagII, to carboxy-terminus of human Smc6. The fusion (Nse4a-Smc5) and Smc6 were downstream of the GAL1-Gal10 promoter in a 2μ URA3 leu2-d containing plasmid for expression in *S. cerevisiae.*

Two plasmids were used to fuse Nse4a to Smc5. The plasmid p1420 was mutagenized using the primer 3086 to introduce an Xmal restriction site in SMC5 ORF. The positive plasmid showed three bands on agarose gel after digestion with XmaI/ScaI-HF. DNA fragments encoding the open reading frame (ORF) of Linker-Smc5 from plasmid p1409 and plasmid p1420- Xmal were digested with AvrII/XmaI. The digested products were purified from the agarose gel using MiniElute ^{TM} Gel Extraction Kit (Qiagen®, USA) following the manufactures recommendations. The ligation was performed with a 5:1 ratio of insert to vector. The new plasmid p1420-Xmal-Linker-Smc5 was identified by PCR using the primers 3037/2842 and it was purified from *Escherichia Coli* using the MiniElute^{TM} Purification Kit (Qiagen®, USA) according to the manufacturer's instructions. DNA fragments encoding the functional regions of human Nse4a from the plasmid p1309 were amplified by PCR using the following primers: 2835 and 3087. The PCR product was purified with a PCR^{TM} Purification Kit (Qiagen®, USA) and digested with AgeI/SphI to obtain as single band of 1224 bp. New plasmid p1420-Xmal-Linker-Smc5 was digested with AgeI/SphI and in this case two bands 10585 bp and 1746 bp were obtained. Positive plasmid obtained from fragments of 10585 bp and 1224 bp was screened by PCR using the following primers 2836/2673. This new plasmid that contains the ORF of human Nse4a-linker-Smc5 was digested with KasI/KpnI and a new plasmid was obtained after the ligation of 4876 bp and 11373 bp (p1420) respectively. Finally, the new plasmid p1473 was selected following digestion with AgeI/SphI.

As an alternative, we also created an alternative fusion in order, to overexpress and solubilize Smc5, Smc6 and Nse4a in a stoichiometric relation. In this case, the gene of Nse4a was fused to the C terminus of Smc6. The purification yield of Smc6-Nse4a fusion was much less in comparison to the Nse4a-Smc5 fusion (data no shown). At this time, we choose to continue using the Nse4a-Smc5 fusion.

2.2.7.2 Culture conditions in fermenter and overexpression of human Smc5/6 complex.

The human Smc5/6 complex was overexpressed in yeast strain D5819 transformed with plasmid p1482, containing the human NSE1 and NSE3 genes; using the LiAc/SS carrier DNA/PEG method (Gietz and Schiestl 2007). The batch fermentations cultures were carried out in a BIOFLO®410 fermentor/Bioreactor with a working volume of 12 L. The following parameters were set and recorded throughout the experiment using the NBS Local Controller Interface. The pH set to 5.0 and adjusted adding NaOH (15 % [v / v]) or H_3PO_4 (5 % [v / v]), temperature 30 °C. Dissolved oxygen was maintained at 30% by varying the stirring between 200 and 400 rpm, O_2 and aeration between 0.2 and 25 L min-1. The initial volume of media was 11.5 L of medium (Yeast extract 11 g L^{-1}, Peptone 22 g L^{-1}, NaOH 10 g L^{-1}, Adenine 55 mg L^{-1}, Lactic Acid 1%, Glycerol 3% were used as carbon sources). The fermentor was inoculated with 500 mL of the inoculum grown in a shaking incubator over night at 30 °C. Yeast cultures were grown at 30 °C to an A_{600} of 0.6 OD and induced with galactose (2 % final) to overexpress the human Smc5/6 complex over night at 20 °C.

2.2.7.3 Purification of human Smc5/6 Complex.

The yeast strain D5819 was lysed and it was resuspended in LyB supplemented with 20 mM imidazole and phosphatase inhibitors and protease inhibitors (EASBF, E64 and PEPA). The cell lysis was passed twice over a Ni-NTA resin and then was washed under 10 CVs of LyB supplemented with 60mM imidazole. The human Smc5/6 complex was then eluted from the Ni-NTA resin (SIGMA) in StB supplemented with 500 mM imidazole. The eluted complex was then loaded on a 5 mL Strep-Trap column (GE Amersham). Washed with 10 CVs of StB supplemented with 0.5 % Triton; the complex was eluted with 5 CVs of GST buffer (GB) 25 mM K_2HPO_4/ KH_2PO_4 pH 8, 500 mM NaCl, 10 % glycerol and 2 mM [2-βME] supplemented with 20 mM of desthiobiotin. Next, the eluted complex was loaded on a 5 mL GSTrap column (GE Healthcare) and then washed with 10 CVs of GB. Finally, the human Smc5/6 complex was

eluted with 5 CVs of 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 % glycerol and 2 mM [2-βME] supplemented with 10mM of Glutathione. The linker, 6xHis, 3xStrep and GST were cleaved by overnight digestion with 1 mg/mL of S-TEV protease per 4 mg/mL of fusion complex in presence of 1 mM of Dithiothreitol (DTT). The cleaved complex was applied to a Superose 6 10/300 column gel filtration column (GE Healthcare). The resulting human Smc5/6 complex was isolated from the TEV enzyme and it was concentrated to 4.67 mg/ml in gel filtration buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol and 2mM [2-βME]) using (10K NMWL; Millipore). Finally, 160 pmoles of the complex was kept in ice in order, to perform the Sucrose density gradient and GraFix. The remain protein was frozen on dry ice and stored at -80°C until required for other experiments.

2.2.8 Purification by zonal ultracentrifugation and cross-linking

2.2.8.1 Sucrose density gradient

Sucrose density gradient ultracentrifugation is a powerful technique for fractionating a mixture containing macromolecules of different size. The proteins are separated according to its size and density by increasing the gradient of sucrose from top to bottom in addition to a centrifugal force (Raschke, Guan et al. 2009).

The gradient was performed using two different concentrations of sucrose, 5 % and 20 %, in 50 mM Tris-HCl pH8, 500 mM NaCL, 0.05 mM EDTA, 6.25 mM $MgCl_2$, 0.5 mM [2-βME]. To prepare the gradient, the different-density solutions were layered in a 12.5 ml polyallomer centrifuge tube, first by adding 5.5 ml of buffer with 5% sucrose to the tube. Next, 5.5 ml of buffer with 20% sucrose was carefully placed at the bottom of the tube using a stainless-steel needle. The solution was slowly expelled, so that the lighter solution is displaced upwards. Tubes were then closed with a cap. To create a continuous density gradient between 5% and 20 %, the tubes were placed into a BioComp® gradient station machine, configured according to the manufactures recommendation. 80 pmoles of concentrated human Smc5/6 complex was loaded at the top of the gradient. Ultracentrifugation was performed over 22 hours at 32 000 RPM at 4 °C using a Beckman Coulter SW 41 Ti Rotor on a Sorvall WX100 ultracentrifuge. The next day the gradient was carefully separated into 24 fractions of 0.5 mL, starting from top to bottom. Fractions were loaded on a SDS-PAGE acrylamide gel and visualized using SilverQuest™ Silver Staining Kit (Invitrogen®).

2.2.8.2 Gradient Fixation (GraFix).

The fixation and purification of human Smc5/6 complex with GraFix, help prevent the protein heterogeneity and drastically reduced the disruption of complex during the preparatory steps for Electronic Microscopy (EM) studies. In this protocol, the human Smc5-6 complex are exposed to a cross-linker gradient of glutaraldehyde and sucrose by ultracentrifugation (Stark 2010).

The gradient was made using two different concentrations of sucrose, 5% and 20%, diluted in 50 mM HEPES pH 8, 500 mM NaCl, 0.05 mM EDTA, 6.25 mM MgCl2, 0.5 mM 2 BME. The buffer with 20 % of sucrose was supplement with 0.15% of glutaraldehyde. The cross-linking was performed using the same protocol described section 2.2.8.1. The cross-linking of the complex was stopped by diluting the protein in quenching buffer (100 mM Tris-HCl pH8, 500 mM glycine, 300 mM NaCl and 2 % glycerol) (Stark 2010) . Each collected fraction was loaded on a SDS-PAGE acrylamide gel and visualized using SilverQuest™ Silver Staining Kit (Invitrogen®). The fractions corresponding to cross-linked human Smc5-6 complex were concentrated and sent to our collaborators in the biochemistry department of the University of Montreal for analysis.

2.2.9 DNA-binding activity of human Smc5/6 complex

The DNA binding activity of human Smc5-6 complex was analysed by electrophoretic mobility shift assay (EMSA) following the protocol used in Roy *et al* (Roy, Siddiqui et al. 2011). DNA substrates used were phiX174 (ssDNA substrate; 5386 bp) and EcoRI-digested pBluescript II KS1 (dsDNA substrate; 2961 bp). Protein–DNA binding reactions were performed in 25 µl of (10mM Hepes pH 7.5, 160 mM NaCl, 7 mM MgCl2, 20% glycerol and 2 mM 2-βME) containing 50 ng of dsDNA or 50 ng of ssDNA substrate in each of the reactions and a different molar fold-excess of human Smc5-6 complex. After incubation at 30 °C for 30 min, the reactions were terminated by addition of an equal volume of 1.6% low melting point (LMP) agarose containing loading buffer (0.6% glycerol, 0.005% bromophenol blue, 0.005% xylene cyanol). The mixture was then loaded on 0.8% TAE-agarose gels and the DNA was resolved by electrophoresis for 16 h at 4 °C. DNA bands were visualized after staining the gels with SYBR Gold (ssDNA; Molecular Probes) or SYBR Green I (dsDNA; Molecular Probes); and they were visualized using the Thyphoon FLA 9500 fluorescence imaging system (General electric). Quantification of free DNA was performed using ImageJ software version 5.7. Curve fitting of affinity

constants and Hill coefficients were determined using GraphPad Prism 7 (GraphPad Software Inc.). All DNA-binding experiments were performed at least in triplicates.

2.3 Results

2.3.1 Reassembling of human Nse1/Nse3/Nse4 subcomplex

Full-length human Nse3 and a 6x Histidine-tagged Nse1 were co-expressed in *E. coli*. A purification scheme using nickel-affinity chromatography followed by cationic exchange, on SP Sepharose fast flow column, revealed that human Nse1 and Nse3 were co-purified as a single peak (**Figure 2.1 A**). Examination of eluted fractions by SDS-PAGE showed that the two proteins co-migrate, demonstrating that Nse1 and Nse3 form a heterodimer.

The human 6x Histidine-tagged Nse4a is a hydrophobic protein, forming part of high density aggregates called inclusion bodies. Normally these aggregates are not properly folded, and it is very difficult get them in a soluble form. In a first attempt to solubilize the human Nse4a from the inclusion body, we first a purified the protein under denaturing conditions, the protein was refolded getting it very pure after one step of Ni-NTA purification (data no shown). Despite the high yield of the denaturalization and refolding protocol; the purified Nse4a showed a weak interaction with Nse1 and Nse3. For this reason, we tried to purify this protein under native conditions.

The human Nse4a was successfully solubilized under native condition in sodium phosphate buffer at pH 7 and 1M sodium chloride after trying different pH and salt concentrations. Although approximately 40% of Nse4a was solubilized, the remaining 60% was retained in the pellet after 3 steps of resuspension. The novel protocol described in Materials and Methods allows us to obtain Nse4a almost pure after its nickel-affinity purification. However, the protein contaminated with Tdh3, even after washing with 100mM imidazole with 10 CV and eluting with a linear gradient of Imidazole (100mM to 1M) (**Figure 2.1 B**). Therefore, we decided to delete Tdh3 gen because the protein molecular mass (36 kDa) is very similar to human Nse3, (34.3 KDa). The gene of *TDH3* was disrupted from the yeast genome, following the procedure described in section 2.2.5.3 of materials and method. After creation of the ΔTDH strain without Tdh3, it was necessary to find another way to stabilize the Nse4a protein. For this reason, we attempted to reassemble the human Nse1/Nse3/Nse4 subcomplex, keeping the Nse4a protein immobilized on to the HisTrap column, and loading purified the Nse1 and Nse3.
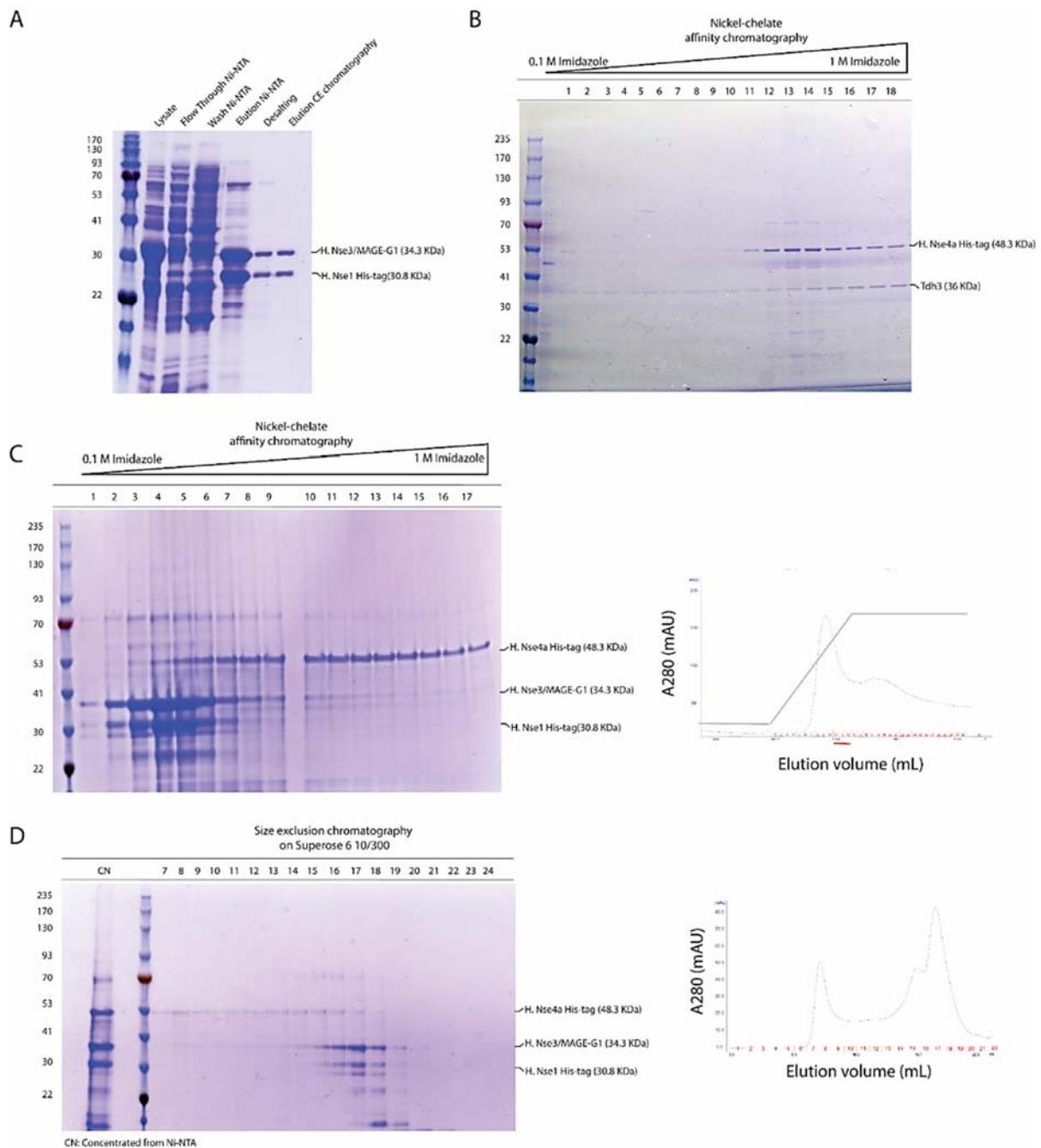
Figure 2.1: Coomassie-stained gels of reconstitution in vitro of human Nse1/Nse3/Nse4a subcomplex. **A:** Purification of full-length human Nse3/MAGE-G1 and a 6x Histidine-tagged Nse1 by nickel-affinity chromatography followed by a cationic exchange. **B:** Purification of human Nse4a by nickel-affinity chromatography. **C:** Reassembling of human Nse1/Nse3/Nse4a complex in His-trap column. **D:** Gel filtration chromatography of of human Smc5/6 sub-complex. According to its molecular weigh, ~115 kDa, the human Smc5/6 subcomplex should be collected in the fraction 14 as a single peak.

In a first attempt to reconstitute the Nse1/Nse3/Nse4a subcomplex, we loaded the concentrated human Nse1, Nse3 from cationic exchange chromatography onto the HisTrap column where Nse4a was immobilized. The concentrated elution from the previous step, was loaded on Superose 6 10/300 chromatography column. According to its molecular weigh, ~115 kDa, the Nse1/Nse3/Nse4a subcomplex should be elute from a Superose 6 purification in the fraction 14 as a single peak. Instead, two overlapped peaks were obtained demonstrating that the subcomplex formation did not form in the HisTrap column. We hypothesise that the pH changes prior to the Nse1/Nse3 by cationic exchange step, could change the three-dimensional structure of these proteins; affecting the adequate exposure of the protein domains necessary for the interaction with Nse4a (data not shown). In a second attempt, the Nse1 and Nse3 dimer was directly used from the Ni-NTA elution. In this case the concentration of imidazole was diluted under 100mM. This dilution prevented the elution of the immobilized Nse4a protein from the HisTrap column. In addition, low imidazole concentration should facilitate the interaction between Nse1, Nse3 and Nse4. The Nse1, Nse3 and Nse4 were eluted from Ni-NTA using a gradient of imidazole (**Figure 2.1 C**). In this case the pH was not changed, maintaining the subcomplex in a more stable confirmation. A small peak was observed that corresponded 14th fraction in gel filtration chromatography (**Figure 2.1 D**), mediating the possibility of forming the sub-complex *in vitro*.

2.3.2 Purification of the trimer formed by Smc5, Smc6 and Nse4a.

Next, we tested whether or not human Smc5, Smc6 and Nse4a form a complex. Full-length His6-tagged Smc5 and, Strep-tagged Smc6 and His6-tagged Nse4a were co-expressed in yeast. A two-step purification scheme using nickel-affinity chromatography followed by Strep-tactin affinity chromatography revealed that human Smc5, Smc6 and Nse4a co-purified. Each fraction from the chromatography steps was examined by SDS-PAGE and a co-migrating Smc5-Histag, Smc6-Streptag and Nse4-Histag was found, demonstrating Smc5, Smc6 and Nse4a form a heterotrimeric complex (**Figure 2.2 A**). The intensity of the Coomassie Blue staining showed that the ratio of Smc5, Smc6 and Nse4a eluate in ratio of approximately 1:1:1; indicating that they form a stoichiometric complex when these three proteins are co-expressed in the same yeast
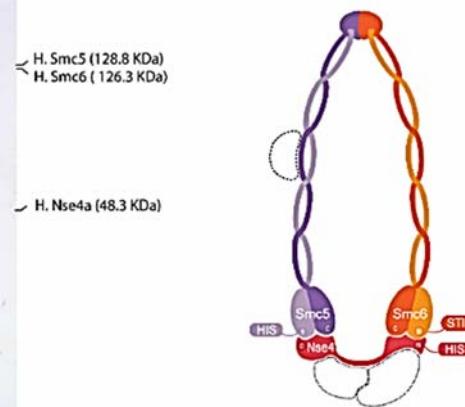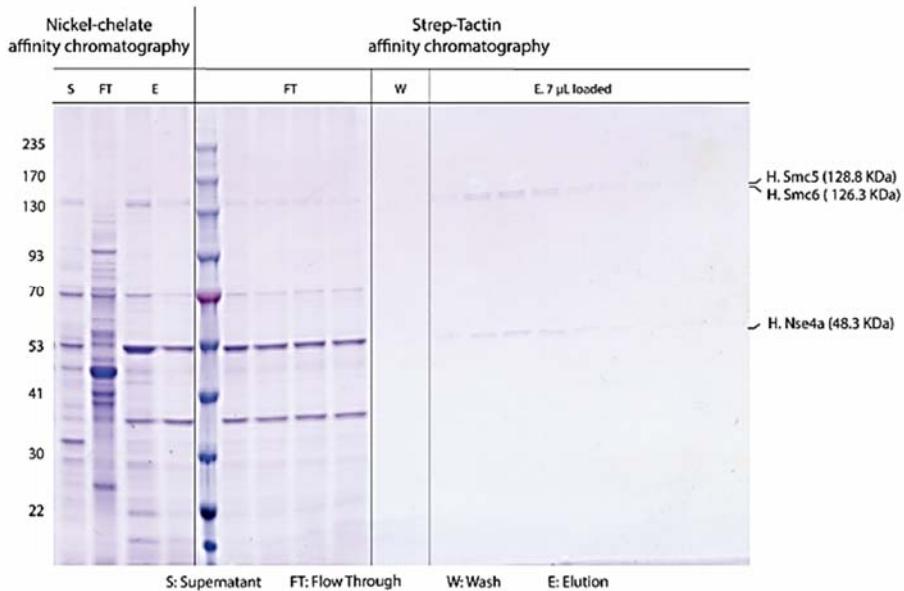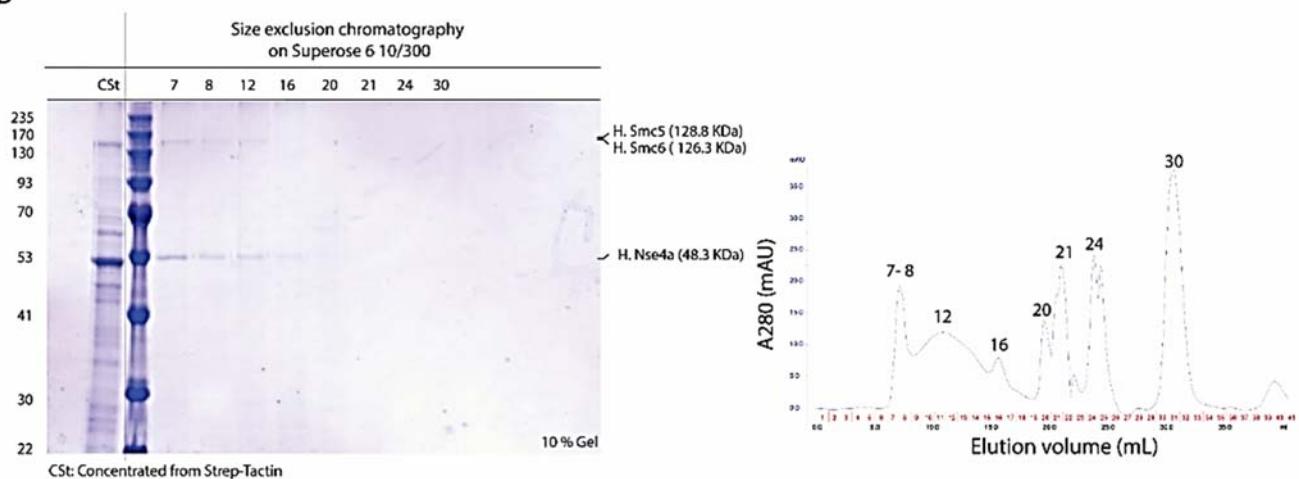
Figure 2.2: Coomassie-stained gels of purification of the trimer formed by Smc5, Smc6 and Nse4a proteins. **A:** Purification of full-length human 6x Histidine-tagged Smc5, Strep-tagged Smc6 and His6-tagged Nse4a by nickel-affinity chromatography followed by Strep-Tactin affinity chromatography. **B:** Gel filtration chromatography of human Smc5/6 trimer. The elution profile of gel filtration column is shown in the graph on the right.

Finally, we examined whether Nse1 and Nse3 directly binds to the Nse4a when the heterotrimeric complex (Smc5/Smc6/Nse4a) was immobilized in Strep-tactin. We found that the heterotrimeric complex did not interact with the heterodimer Nse1/Nse3 (data not shown), suggesting that it may not be possible to reassemble the human Smc5/6 complex *in vitro*.

2.3.3 Purification of human Smc5/6 complex.

We previously demonstrated that when expressed from yeasts, His6-tagged Smc5, Strep-tagged Smc6, and His-tagged Nse4a can be co-purified in stochiometric retio after Strep-tactin affinity chromatography. This suggests that we can reassemble and purify the entire complex by overexpressing the human His6-tagged Nse4a-Smc5, Strep tactin-tagged Smc6, untagged Nse3 and GST-tagged Nse1 in the same strain.

To characterize the biochemical properties of the full human Smc5/6 complex, we first needed to purify large amounts of the complex. To achieve this, we fused three affinity purifications tags in different proteins of the complex to facilitate its purification. Specifically, we co-expressed a human His-tag Nse4a-Smc5 fusion in yeast with Strep-tagged Smc6, Nse3 and GST-tagged Nse1. We able to purify the human Smc5/6 complex to greater than 95 % homogeneity using consecutive steps of metal-chelate, streptactin and GST-affinity chromatography (**Figure 2.3 A**). An additional step of size exclusion chromatography on a 10/300 Superose 6 was necessary after cleaving the Nse4a-Smc5 fusion whit TEV. The linker between Nse4a and Smc5 was removed and two bands were visible on the gel; one at 127 KDa corresponding to human Smc5 and another to human Nse4a at 44.3KDa. In addition, another band corresponding to human Nse1 appeared at 30.4 KDa after cleaving the GST-tag. Finally, the cleaved version of human Smc5/6 complex was eluted as a single peak and it was easily separated from the TEV enzyme, the His-tag, the Strep-tag and the GST-tag (**Figure 2.3 B**). Using this procedure, we obtained a yield of 12.8 mM of pure human Smc5/6 complex per 400 g of biomass. This purification was sufficient to conduct key biochemical assays like DNA binding assay, to obtain images of electron microscopy imagen and to perform the structural interaction analysis between the proteins by Mass Spectrometry (MS).
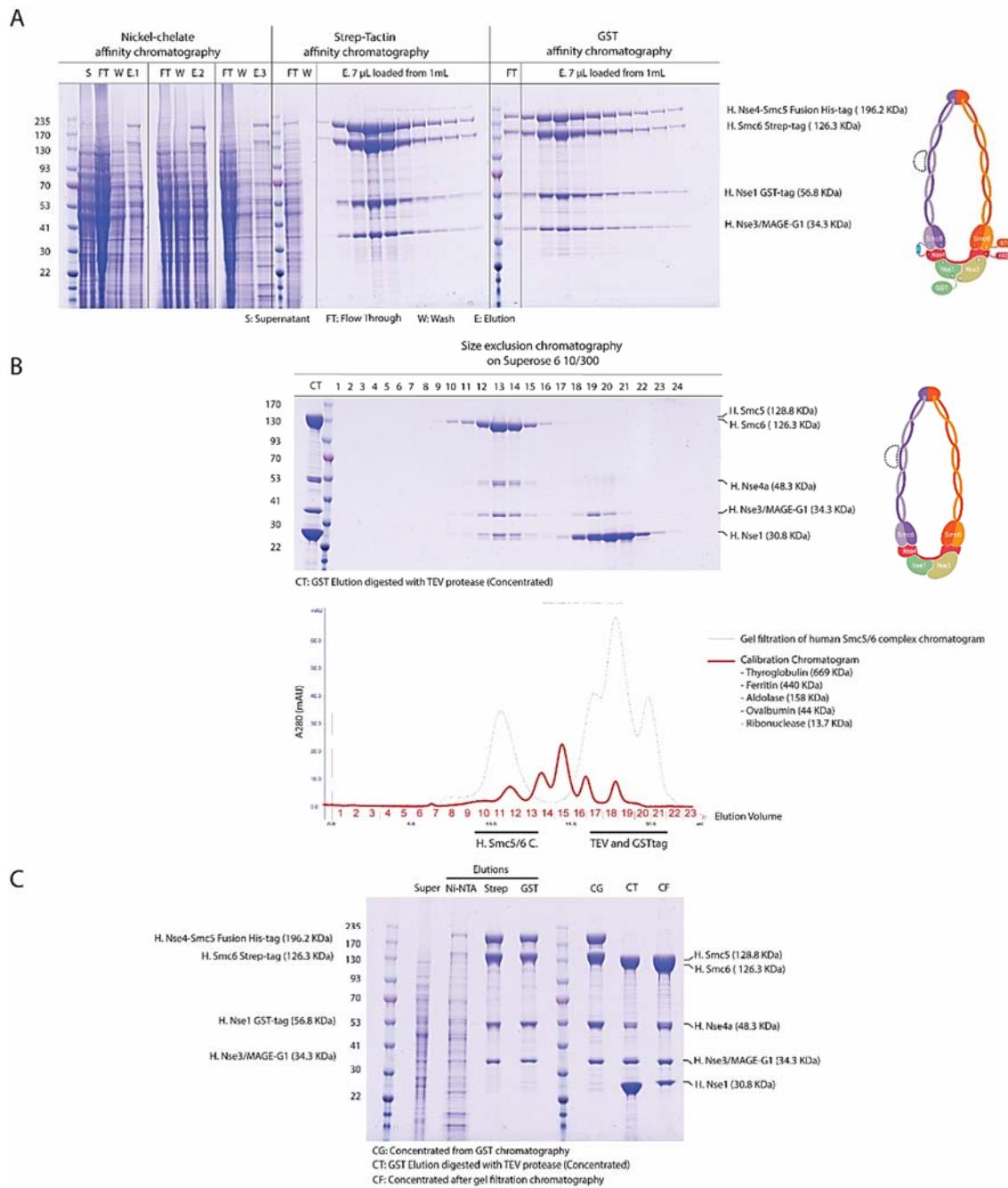
Figure 2.3: Coomassie-stained gels of purification of human Smc5/6 complex. **A:** Purification of full-length human 6x Histidine-tagged Nse4a-Smc5 fusion, strep tactin-tagged Smc6 and His6-tagged, Nse3/MAGE-G1 and GST-tag human Nse1 by nickel-affinity chromatography followed by Strep-Tactin affinity chromatography and GST affinity chromatography. **B:** Gel filtration chromatography of human Smc5/6 complex digested with TEV protease to cleave the fusion protein. The elution profile of gel filtration column is shown in the graph. **C:** Representative gel of the different purification steps.

Strikingly, after gel filtration chromatography, we found that Smc5/6 complex eluted in fractions 10 to13 corresponding to a higher MW complex than expeted. For this reason, we confirmed its conformation and MW by sucrose density gradient ultracentrifugation.

2.3.4 Sucrose density gradient and GraFix

The human Smc5/6 complex and thyroglobulin (667 KDa; used as a control) were analysed under the same conditions on sucrose density gradient (**Figure 2.4 A**). The human Smc5/6 complex was observed in a fraction with a lower concentration of sucrose than thyroglobulin; demonstrating that the molecular weight of the Smc5/6 complex correspond to a monomeric conformation. Next, we decided to analyse the human Smc5/6 complex by Electron Microscopy (EM) (**Figure 2.4 A**). The images obtained from EM, showed a heterogeneous conformation of complex particles, a high background due to protein aggregation, and poor distribution of human Smc5/6 complex (**Annexe 1 A**). For this reason, we decided to use the GraFix method to prevent the disruption of the complex during grid preparation for Electron Microscopy. Only a single band was visible on the SDS–PAGE gel and we have now successfully stabilized the complex after GraFix. This confirmed that the cross-linking was perfectly done (**Figure 2.4 B**). At this moment, we have been able to maintain the integrity of the complex and this has enables us to obtain a very good imagen by EM of the monomeric complex (**Annexe 1 B**).

2.3.5 The human Smc5/6 complex binds to ssDNA and dsDNA.

It has been previously reported that the yeast Smc5/6 dimer, Nse1/Nse3/Nse4 subcomplex and Smc5/6 hinge domains all have the capacity to bind DNA (Roy, Dhanaraman et al. 2015) (Zabrady, Adamus et al. 2016) (Alt, Dang et al. 2017). In our initial analysis, we determined that the relative affinities of the full human Smc5/6 complex for ssDNA and dsDNA (**Figure 2.5 A**). The DNA binding was analysed using EMSA and increasing progressively the concentrations of human Smc5-6 complex at physiological concentrations of salt (160mM). The human Smc5/6 complex binds with lower affinity to dsDNA $K_d = 82.48 \pm 4$ nM than to ssDNA $K_d = 27.32 \pm 0.76$ nM (**Figure 2.5 B**). These results strongly suggest that the human Smc5/6 complex can directly bind to ssDNA and dsDNA when such structure is formed at stalled replication forks (Lindroos, Strom et al. 2006).

2.4 Discussion

Structural and functional studies of a complexes require a considerable quantity of protein in the native conformation. It is sometimes impossible to get this concentration from natural sources. Thus, it is necessary to overexpress it in another organism. *Saccharomyces cerevisiae* is one of the most commonly host used for protein overexpression. This model organism is characterized by its rapid growth, the simple and inexpensive culture media, and importantly, many of the cellular and metabolic processes found in higher eukaryotes are conserved in yeast. While protein overexpression in yeast does not usually present big difficulties, purification a protein complex under native conditions with many subunits may be a challenge. Reassembling the human Smc5/6 complex was the main goal of this work. As a first attempt to reassemble the complex, we tried to purify all 6 subunits of the complex separately, for subsequent *in vitro* reconstitution. Overexpressing each protein separately, increases the difficulties of the complex reassembly because is not always possible to control the correct folding and the final conformation of each protein.

Due to the difficulties in obtaining a good yield for the *in vitro* refolding of Nse4a and its aggregation in native conditions, we decided to approach purifying the complex using with a different strategy. This was accomplished fusing the carboxy-terminus of human Nse4a to the amino-terminus of Smc5 by a flexible linker, as done previously for structure of the Scc1-Smc1 (Gruber, Arumugam et al. 2006, Palecek and Gruber 2015)

2.4.1 Reconstitution *in vitro* of human Smc5/6 complex.

2.4.1.1 Reassembling the human Smc5/6 subcomplex

The Nse4a subunit contains a conserved hydrophobic pattern characteristic of the entire kleisin family (Palecek, Vidot et al. 2006). When working with highly hydrophobic proteins, the use of detergents and high salt concentrations is inevitable, to prevent the formation of aggregates and/or precipitation. Although the first part of the IMAC purification of the human Nse4a protein under native conditions was successful, as the process progressed, difficulties arose which threatened the assembly of the human Smc5/6 subcomplex. Tdh3 as a contaminant in the Ni-NTA elutions was the first complication that need to be solved before attempting to assemble the subcomplex. Cation exchange chromatography was used in unattempt eliminate Tdh3, but the low salt concentration required to perform this type of chromatography affected the

solubility of the Nse4 protein. Next, we decided to delete the Tdh3 gene from the yeast genome. There are not reports that explain the presence of Tdh3 in Ni-NTA elutions; but we believe that this protein could interact in an unspecific way with the union sites for Nse1 and Nse3.
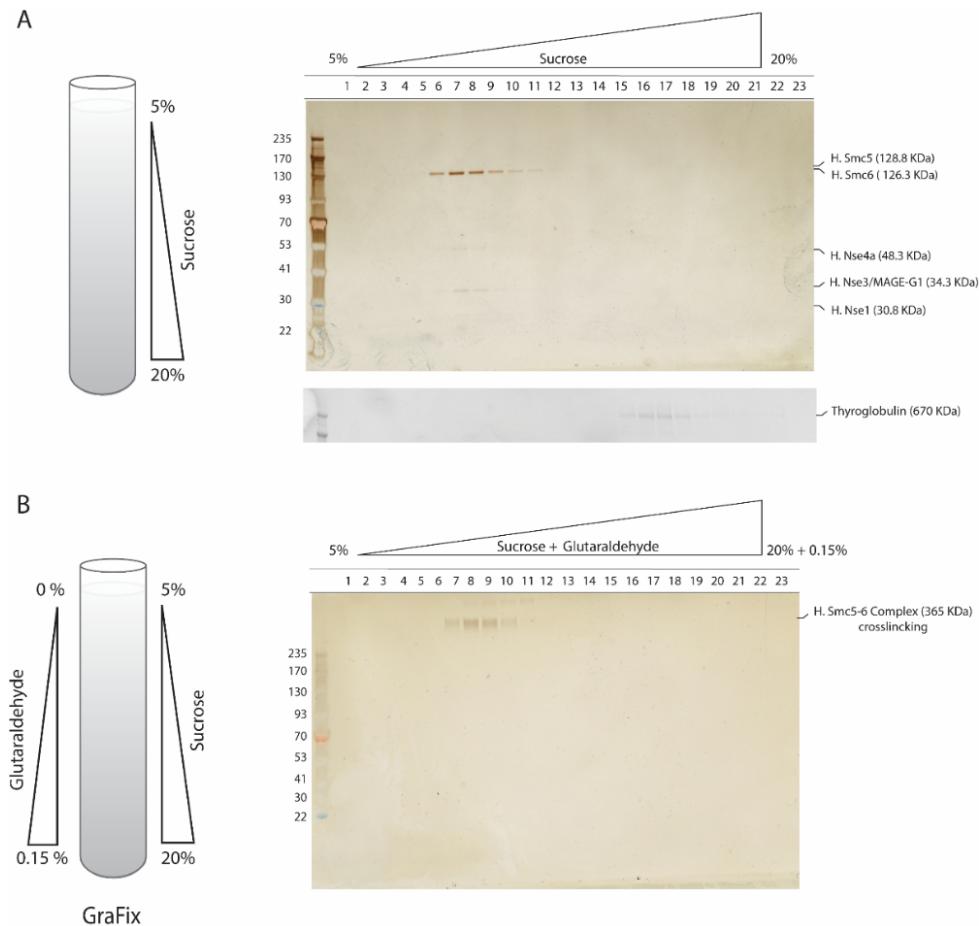

Figure 2.4: Sucrose density gradient and GraFix of Smc5/6 complex. **A:** After ultracentrifugation sucrose density gradient, the complex was collected in the MW range corresponding to 300 to 400 kDa (predicted MW 365kDa). Smc5/6 complex is stable after purification and is in a monomer. **B:** GraFix, sucrose density gradient in presence of 0.15 % glutaraldehyde. The human Smc5/6 complex was stabilized in a monomeric conformation for the electron microscopy studies.

As described in the results section, we attempted several strategies for the *in vitro* reconstitution the Nse1/Nse3/Nse4a subcomplex. Unfortunately, none of them were to remarkably success. Changes in pH, salt concentration and imidazole concentration to some extent affected of the conformations of Nse1 and Nse3, and reducing their ability to interact with Nse4a during the reassembly on the His-Trap column. Recent studies have shown that it has been possible to purify the Nse1/Nse3/Nse4a subcomplex and to characterize it. Using a combined approach

assays and *in vitro*, it was shown that N-terminal half (aa 1-150) of the yeast S. pombe Nse4 protein is sufficient for binding to the C-terminal of Nse3 and the C-terminal of Nse1 (Hudson, Bednarova et al. 2011). In other studies, it was shown that the Nse1, Nse3 and Nse4, co-expressed in *E. coli*, could be purified in a single peak after gel filtration (Duan, Yang et al. 2009).
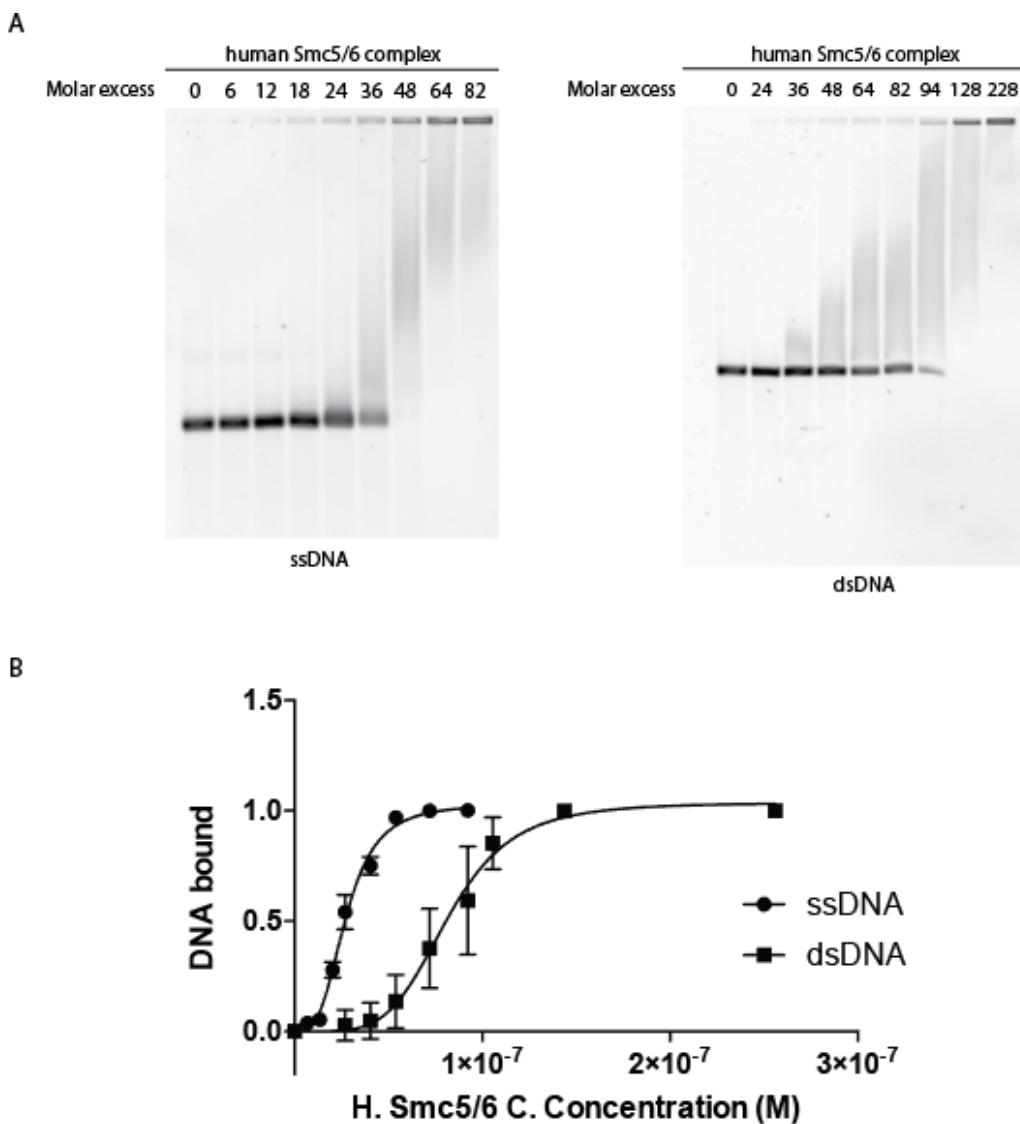
Kd (ssDNA) = 27.32 ± 0.76 nM   Kd (dsDNA) = 82.48 ± 4 nM

Figure 2.5: DNA binding activity of human Smc5/6 complex. **A:** The saturations DNA-binding activity experiments of human Smc5/6 complex. Increased amounts of protein in molar excess were added to ssDNA and dsDNA plasmids. **B:** The human Smc5/6 complex is bind with lower affinity to dsDNA (Kd = 82.48 ± 4 nM) than to ssDNA (Kd = 27.32 ± 0.76) nM. The binding curves fit the experimental values with a correlation coefficient of 0.9901 and 0.9243 for ssDNA and dsDNA, respectively.

Solubilized proteins often lose stability over time by aggregating to each other. The reasons for aggregation are not exactly known, but it is easy to think that the environment and buffer conditions, although necessary to keep the protein in suspension in aqueous solution, do not simulate the in natural environment.

2.4.1.2 Purification of Smc5, Smc6 and Nse4a trimer.

In view our unsuccessful results in our attempts to reconstitute the human Nse1/Nse3/Nse4 subcomplex *in vitro*. We decided to attempt the purification of the trimer formed by Smc5, Smc6 and Nse4a.

Smc5, Smc6 and Nse4a were overexpressed in the same yeast and were subsequently purified using two steps of purification. The result of the Smc5, Smc6 and Nse4a assembly was remarkably different from our results with the reconstitution of the human Nse1/Nse3/Nse4a subcomplex. A stoichiometric reassembly of the Smc5, Smc6 and Nse4 trimer was clearly observed after the Streptactin purification step (**Figure 2.2 A**). Although we never had success reassembling the full complex *in vitro* by the combination of Smc5, Smc6 and Nse4a trimer and Nse1 and Nse3 dimer. The purification of the trimer showed that it was possible to reassemble and purify in a stoichiometric relation, proteins co-overexpressed in yeast. This important result led us to co-overexpress all subunits of the Smc5/6 complex in the same yeast and try to purify them then as a single using a series of affinity chromatography steps.

2.4.2 Reconstitution of human Smc5/6 complex *in vivo.*

2.4.2.1 Purification of human Smc5/6 complex.

Our study also represents a tentative approach towards a comprehensive biochemistry characterization of the human Smc5/6 complex. As previously mentioned, the Smc5/6 complex is required for telomere recombination and telomere-maintenance in the absence of telomerase (Noel and Wellinger 2011), it enhances sister chromatid alignment after DNA damage, facilitating correct DSB repair by HR (Xu, Yuan et al. 2013) and it is also required for the loading of RPA and Rad52 onto stalled replication forks to maintain them in a recombination competent configuration (Irmisch, Ampatzidou et al. 2009).

To help with this characterization we have also developed an affinity purification protocol for the rapid isolation of assembled human Smc5/6 complex using three affinity chromatography

selection steps. The high binding capacity of Ni-NTA resin in the first step of purification facilitates the processing of a large volumes of lysate and ensures an efficient recovery. The subsequent purification steps (Streptrap and GSTrap) improve the purity of the complex and help to guarantee the stochiometric relation ship between the human Smc4-Smc5, Smc6, Nse1 and Nse3. The 1:1:1:1 ratio is clear evident after GST purification, because in this step the complex is purified using human Nse1, the limiting protein.

This work constitutes the first evidence of purification of human Smc5/6 complex. Purification of the entire complex reduced considerably the protein aggregation and increased its stability for futures biochemical experiments. The stability was also enhanced using a higher salt concentration in the buffers and very low concentrations of non-ionic detergents (0.1% Triton X-100 and 0.5% Tween20).

2.4.3 Sucrose density gradient and GraFix

In gel filtration chromatography, the human Smc5/6 complex appeared to have a higher molecular weight. For this reason, we decided to do a more precis analyse the molecular weight using a different technique. By sucrose density gradient, a molecular mass of 365 kDa was determined, which matches the predicted molecular mass of a monomeric complex with a stochiometric ratio of the individual subunits. Examination of collected fractions using SDS-PAGE showed that Smc5, Smc6, Nse4, MAGE-G1/Nse3 and Nse1 proteins, eluted in the same fraction (**Figure 2.4 A**), and this demonstrates that the human Smc5/6 complex is still stable after four steps of purification.

In addition, the cross-linking by ultracentrifugation through a sucrose density gradient drastically reduced complex degradation and helped maintain the integrity of the human Smc5-6 complex during grid preparation for analysis by EM. (**Annexe 1 B**).

2.4.4 New evidence of human Smc5-6 complex binding properties to ssDNA and dsDNA.

Several studies have examined the DNA-binding properties of Smc5 and Smc6 dimer and Nse1/MAGEG1-Nse3/Nse4 subcomplex (Roy, Dhanaraman et al. 2015) (Zabrady, Adamus et al. 2016) We have studied the DNA-binding of the full human Smc5/6 complex and shows that the DNA-binding activity of Smc5/6 complex occurs at physiological concentrations of salt

(160mM). This results strongly suggest that the Smc5/6 complex can directly associate with ssDNA and dsDNA intermediates created during DNA repair. The human Smc5-6 complex binds with lower affinity to dsDNA than the ssDNA. Our results provide biochemical evidence that the human Smc5/6 complex contribute to the initiation of HR at stalled replication forks by binding to the recombination intermediates (Lilienthal, Kanno et al. 2013, Xue, Choi et al. 2014).

Conclusion

The Co-translational system allows *in vivo* complex assembly and facilitates full complex purification as monomer conformation. GraFix reduce the heterogeneity of complex conformations and to stabilize for the electron microscopy. Finally, the human Smc5/6 complex has the ability to directly associate with ssDNA and dsDNA intermediates created during DNA repair. These results will have a critical impact on our understanding of the basic biochemical processes that govern DNA repair and maintenance of human genome integrity. As genomic instability is one of the main causes of tumorigenesis. The results we will achieve in this project will have a significant impact on human health and the possible development of new drugs targeting human complex activity Smc5/6.

Bibliographies

Alt, A., H. Q. Dang, O. S. Wells, L. M. Polo, M. A. Smith, G. A. McGregor, T. Welte, A. R. Lehmann, L. H. Pearl, J. M. Murray and A. W. Oliver (2017). "Specialized interfaces of Smc5/6 control hinge stability and DNA association." Nat Commun **8**: 14011.

Ampatzidou, E., A. Irmisch, M. J. O'Connell and J. M. Murray (2006). "Smc5/6 is required for repair at collapsed replication forks." Mol Cell Biol **26**(24): 9387-9401.

Andrews, E. A., J. Palecek, J. Sergeant, E. Taylor, A. R. Lehmann and F. Z. Watts (2005). "Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage." Mol Cell Biol **25**(1): 185-196.

Barker, P. A. and A. Salehi (2002). "The MAGE proteins: emerging roles in cell cycle progression, apoptosis, and neurogenetic disease." J Neurosci Res **67**(6): 705-712.

Bavner, A., J. Matthews, S. Sanyal, J. A. Gustafsson and E. Treuter (2005). "EID3 is a novel EID family member and an inhibitor of CBP-dependent co-activation." Nucleic Acids Research **33**(11): 3561-3569.

Bermudez-Lopez, M. and L. Aragon (2017). "Detection of Cohesin SUMOylation In Vivo." Methods Mol Biol **1515**: 55-64.

Bermudez-Lopez, M., A. Ceschia, G. de Piccoli, N. Colomina, P. Pasero, L. Aragon and J. Torres-Rosell (2010). "The Smc5/6 complex is required for dissolution of DNA-mediated sister chromatid linkages." Nucleic Acids Res **38**(19): 6502-6512.

Borden, K. L. (2000). "RING domains: master builders of molecular scaffolds?" J Mol Biol **295**(5): 1103-1112.

Chomez, P., O. De Backer, M. Bertrand, E. De Plaen, T. Boon and S. Lucas (2001). "An overview of the MAGE gene family with the identification of all human members of the family." Cancer Res **61**(14): 5544-5551.

Copsey, A., S. Tang, P. W. Jordan, H. G. Blitzblau, S. Newcombe, A. C. Chan, L. Newnham, Z. Li, S. Gray, A. D. Herbert, P. Arumugam, A. Hochwagen, N. Hunter and E. Hoffmann (2013). "Smc5/6 coordinates formation and resolution of joint molecules with chromosome morphology to ensure meiotic divisions." PLoS Genet **9**(12): e1004071.

Cox, M. M. (2007). "Regulation of bacterial RecA protein function." Crit Rev Biochem Mol Biol **42**(1): 41-63.

De Piccoli, G., F. Cortes-Ledesma, G. Ira, J. Torres-Rosell, S. Uhle, S. Farmer, J. Y. Hwang, F. Machin, A. Ceschia, A. McAleenan, V. Cordon-Preciado, A. Clemente-Blanco, F. Vilella-Mitjana, P. Ullal, A. Jarmuz, B. Leitao, D. Bressan, F. Dotiwala, A. Papusha, X. Zhao, K. Myung, J. E. Haber, A. Aguilera and L. Aragon (2006). "Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination." Nat Cell Biol **8**(9): 1032-1034.

De Plaen, E., K. Arden, C. Traversari, J. J. Gaforio, J. P. Szikora, C. De Smet, F. Brasseur, P. van der Bruggen, B. Lethe, C. Lurquin and et al. (1994). "Structure, chromosomal localization, and expression of 12 genes of the MAGE family." Immunogenetics **40**(5): 360-369.

De Smet, C., C. Lurquin, B. Lethe, V. Martelange and T. Boon (1999). "DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter." Molecular and Cellular Biology **19**(11): 7327-7335.

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

Duan, X., Y. Yang, Y. H. Chen, J. Arenz, G. K. Rangi, X. Zhao and H. Ye (2009). "Architecture of the Smc5/6 Complex of Saccharomyces cerevisiae Reveals a Unique Interaction between the Nse5-6 Subcomplex and the Hinge Regions of Smc5 and Smc6." J Biol Chem **284**(13): 8507-8515.

Fousteri, M. I. and A. R. Lehmann (2000). "A novel SMC protein complex in Schizosaccharomyces pombe contains the Rad18 DNA repair protein." EMBO J **19**(7): 1691-1702.

Fujioka, Y., Y. Kimata, K. Nomaguchi, K. Watanabe and K. Kohno (2002). "Identification of a novel non-structural maintenance of chromosomes (SMC) component of the SMC5-SMC6 complex involved in DNA repair." J Biol Chem **277**(24): 21585-21591.

Gietz, R. D. and R. H. Schiestl (2007). "High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method." Nat Protoc **2**(1): 31-34.

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

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

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

Hirano, T. (2006). "At the heart of the chromosome: SMC proteins in action." Nat Rev Mol Cell Biol **7**(5): 311-322.

Hu, B., C. Liao, S. H. Millson, M. Mollapour, C. Prodromou, L. H. Pearl, P. W. Piper and B. Panaretou (2005). "Qri2/Nse4, a component of the essential Smc5/6 DNA repair complex." Mol Microbiol **55**(6): 1735-1750.

Hudson, J. J., K. Bednarova, L. Kozakova, C. Liao, M. Guerineau, R. Colnaghi, S. Vidot, J. Marek, S. R. Bathula, A. R. Lehmann and J. Palecek (2011). "Interactions between the Nse3 and Nse4 components of the SMC5-6 complex identify evolutionarily conserved interactions between MAGE and EID Families." PLoS One **6**(2): e17270.

Hudson, J. J. R., K. Bednarova, L. Kozakova, C. Liao, M. Guerineau, R. Colnaghi, S. Vidot, J. Marek, S. R. Bathula, A. R. Lehmann and J. Palecek (2011). "Interactions between the Nse3 and Nse4 Components of the SMC5-6 Complex Identify Evolutionarily Conserved Interactions between MAGE and EID Families." PLoS One **6**(2).

Irmisch, A., E. Ampatzidou, K. Mizuno, M. J. O'Connell and J. M. Murray (2009). "Smc5/6 maintains stalled replication forks in a recombination-competent conformation." EMBO J **28**(2): 144-155.

Janke, C., M. M. Magiera, N. Rathfelder, C. Taxis, S. Reber, H. Maekawa, A. Moreno-Borchart, G. Doenges, E. Schwob, E. Schiebel and M. Knop (2004). "A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes." Yeast **21**(11): 947-962.

Katsura, Y. and Y. Satta (2011). "Evolutionary history of the cancer immunity antigen MAGE gene family." PLoS One **6**(6): e20365.

Kegel, A. and C. Sjogren (2010). "The Smc5/6 complex: more than repair?" Cold Spring Harb Symp Quant Biol **75**: 179-187.

Kozakova, L., L. Vondrova, K. Stejskal, P. Charalabous, P. Kolesar, A. R. Lehmann, S. Uldrijan, C. M. Sanderson, Z. Zdrahal and J. J. Palecek (2015). "The melanoma-associated antigen 1 (MAGEA1) protein stimulates the E3 ubiquitin-ligase activity of TRIM31 within a TRIM31-MAGEA1-NSE4 complex." Cell Cycle **14**(6): 920-930.

Kuwako, K., H. Taniura and K. Yoshikawa (2004). "Necdin-related MAGE proteins differentially interact with the E2F1 transcription factor and the p75 neurotrophin receptor." J Biol Chem **279**(3): 1703-1712.

Lee, K. M. and M. J. O'Connell (2006). "A new SUMO ligase in the DNA damage response." DNA Repair (Amst) **5**(1): 138-141.

Lehmann, A. R. (2005). "The role of SMC proteins in the responses to DNA damage." DNA Repair (Amst) **4**(3): 309-314.

Lilienthal, I., T. Kanno and C. Sjögren (2013). "Inhibition of the Smc5/6 Complex during Meiosis Perturbs Joint Molecule Formation and Resolution without Significantly Changing Crossover or Non-crossover Levels." PLoS Genet **9**(11).

Lindroos, H. B., L. Strom, T. Itoh, Y. Katou, K. Shirahige and C. Sjogren (2006). "Chromosomal association of the Smc5/6 complex reveals that it functions in differently regulated pathways." Mol Cell **22**(6): 755-767.

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

Marcar, L., N. J. Maclaine, T. R. Hupp and D. W. Meek (2010). "Mage-A cancer/testis antigens inhibit p53 function by blocking its interaction with chromatin." Cancer Res **70**(24): 10362-10370.

McAleenan, A., V. Cordon-Preciado, A. Clemente-Blanco, I. C. Liu, N. Sen, J. Leonard, A. Jarmuz and L. Aragon (2012). "SUMOylation of the alpha-kleisin subunit of cohesin is required for DNA damage-induced cohesion." Curr Biol **22**(17): 1564-1575.

Meek, D. W. and L. Marcar (2012). "MAGE-A antigens as targets in tumour therapy." Cancer Lett **324**(2): 126-132.

Miyabe, I., T. Morishita, H. Shinagawa and A. M. Carr (2009). "Schizosaccharomyces pombe Cds1Chk2 regulates homologous recombination at stalled replication forks through the phosphorylation of recombination protein Rad60." J Cell Sci **122**(Pt 20): 3638-3643.

Murray, J. M. and A. M. Carr (2008). "Smc5/6: a link between DNA repair and unidirectional replication?" Nat Rev Mol Cell Biol **9**(2): 177-182.

Nardiello, T., A. A. Jungbluth, A. Mei, M. Diliberto, X. Huang, A. Dabrowski, V. C. Andrade, R. Wasserstrum, S. Ely, R. Niesvizky, R. Pearse, M. Coleman, D. S. Jayabalan, N. Bhardwaj, L. J. Old, S. Chen-Kiang and H. J. Cho (2011). "MAGE-A inhibits apoptosis in proliferating myeloma cells through repression of Bax and maintenance of survivin." Clin Cancer Res **17**(13): 4309-4319.

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

Noel, J. F. and R. J. Wellinger (2011). "Abrupt telomere losses and reduced end-resection can explain accelerated senescence of Smc5/6 mutants lacking telomerase." DNA Repair (Amst) **10**(3): 271-282.

Onn, I., J. M. Heidinger-Pauli, V. Guacci, E. Unal and D. E. Koshland (2008). "Sister chromatid cohesion: a simple concept with a complex reality." Annu Rev Cell Dev Biol **24**: 105-129.

Palecek, J., S. Vidot, M. Feng, A. J. Doherty and A. R. Lehmann (2006). "The Smc5-Smc6 DNA repair complex. bridging of the Smc5-Smc6 heads by the KLEISIN, Nse4, and non-Kleisin subunits." J Biol Chem **281**(48): 36952-36959.

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

Pebernard, S., J. J. Perry, J. A. Tainer and M. N. Boddy (2008). "Nse1 RING-like domain supports functions of the Smc5-Smc6 holocomplex in genome stability." Mol Biol Cell **19**(10): 4099-4109.

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

Potts, P. R. and H. Yu (2005). "Human MMS21/NSE2 is a SUMO ligase required for DNA repair." Mol Cell Biol **25**(16): 7021-7032.

Raschke, S., J. Guan and G. Iliakis (2009). "Application of alkaline sucrose gradient centrifugation in the analysis of DNA replication after DNA damage." Methods Mol Biol **521**: 329-342.

Roy, M. A., T. Dhanaraman and D. D'Amours (2015). "The Smc5-Smc6 heterodimer associates with DNA through several independent binding domains." Sci Rep **5**.

Roy, M. A., N. Siddiqui and D. D'Amours (2011). "Dynamic and selective DNA-binding activity of Smc5, a core component of the Smc5-Smc6 complex." Cell Cycle **10**(4): 690-700.

Sang, M., Y. Lian, X. Zhou and B. Shan (2011). "MAGE-A family: attractive targets for cancer immunotherapy." Vaccine **29**(47): 8496-8500.

Sasaki, A., L. Hinck and K. Watanabe (2005). "RumMAGE-D the members: structure and function of a new adaptor family of MAGE-D proteins." J Recept Signal Transduct Res **25**(3): 181-198.

Sergeant, J., E. Taylor, J. Palecek, M. Fousteri, E. A. Andrews, S. Sweeney, H. Shinagawa, F. Z. Watts and A. R. Lehmann (2005). "Composition and Architecture of the Schizosaccharomyces pombe Rad18 (Smc5-6) Complex." Mol Cell Biol **25**(1): 172-184.

Sollier, J., R. Driscoll, F. Castellucci, M. Foiani, S. P. Jackson and D. Branzei (2009). "The Saccharomyces cerevisiae Esc2 and Smc5-6 Proteins Promote Sister Chromatid Junction-mediated Intra-S Repair." Mol Biol Cell **20**(6): 1671-1682.

Stark, H. (2010). "GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM." Methods Enzymol **481**: 109-126.

Stray, J. E., N. J. Crisona, B. P. Belotserkovskii, J. E. Lindsley and N. R. Cozzarelli (2005). "The Saccharomyces cerevisiae Smc2/4 condensin compacts DNA into (plus) chiral structures without net supercoiling." Journal of Biological Chemistry **280**(41): 34723-34734.

Taniura, H., N. Taniguchi, M. Hara and K. Yoshikawa (1998). "Necdin, a postmitotic neuron-specific growth suppressor, interacts with viral transforming proteins and cellular transcription factor E2F1." J Biol Chem **273**(2): 720-728.

Taylor, E. M., A. C. Copsey, J. J. Hudson, S. Vidot and A. R. Lehmann (2008). "Identification of the proteins, including MAGEG1, that make up the human SMC5-6 protein complex." Mol Cell Biol **28**(4): 1197-1206.

Verkade, H. M., S. J. Bugg, H. D. Lindsay, A. M. Carr and M. J. O'Connell (1999). "Rad18 Is Required for DNA Repair and Checkpoint Responses in Fission Yeast." Mol Biol Cell **10**(9): 2905-2918.

Verver, D. E., G. H. Hwang, P. W. Jordan and G. Hamer (2015). "Resolving complex chromosome structures during meiosis: versatile deployment of Smc5/6." Chromosoma.

Verver, D. E., G. H. Hwang, P. W. Jordan and G. Hamer (2016). "Resolving complex chromosome structures during meiosis: versatile deployment of Smc5/6." Chromosoma **125**(1): 15-27.

Verver, D. E., N. S. Langedijk, P. W. Jordan, S. Repping and G. Hamer (2014). "The SMC5/6 complex is involved in crucial processes during human spermatogenesis." Biol Reprod **91**(1): 22.

Verver, D. E., A. M. van Pelt, S. Repping and G. Hamer (2013). "Role for rodent Smc6 in pericentromeric heterochromatin domains during spermatogonial differentiation and meiosis." Cell Death Dis **4**: e749.

Wu, N., X. Kong, Z. Ji, W. Zeng, P. R. Potts, K. Yokomori and H. Yu (2012). "Scc1 sumoylation by Mms21 promotes sister chromatid recombination through counteracting Wapl." Genes Dev **26**(13): 1473-1485.

Wu, N. and H. Yu (2012). "The Smc complexes in DNA damage response." Cell Biosci **2**: 5.

Xu, P., D. Yuan, M. Liu, C. Li, Y. Liu, S. Zhang, N. Yao and C. Yang (2013). "AtMMS21, an SMC5/6 complex subunit, is involved in stem cell niche maintenance and DNA damage responses in Arabidopsis roots." Plant Physiol **161**(4): 1755-1768.

Xue, X., K. Choi, J. Bonner, T. Chiba, Y. Kwon, Y. Xu, H. Sanchez, C. Wyman, H. Niu, X. Zhao and P. Sung (2014). "Restriction of replication fork regression activities by a conserved SMC complex." Mol Cell **56**(3): 436-445.

Yang, B., S. M. O'Herrin, J. Wu, S. Reagan-Shaw, Y. Ma, K. M. Bhat, C. Gravekamp, V. Setaluri, N. Peters, F. M. Hoffmann, H. Peng, A. V. Ivanov, A. J. Simpson and B. J. Longley (2007). "MAGE-A, mMage-b, and MAGE-C proteins form complexes with KAP1 and suppress p53-dependent apoptosis in MAGE-positive cell lines." Cancer Res **67**(20): 9954-9962.

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

Zhao, X. and G. Blobel (2005). "A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization." Proc Natl Acad Sci U S A **102**(13): 4777-4782.

Annexe 1
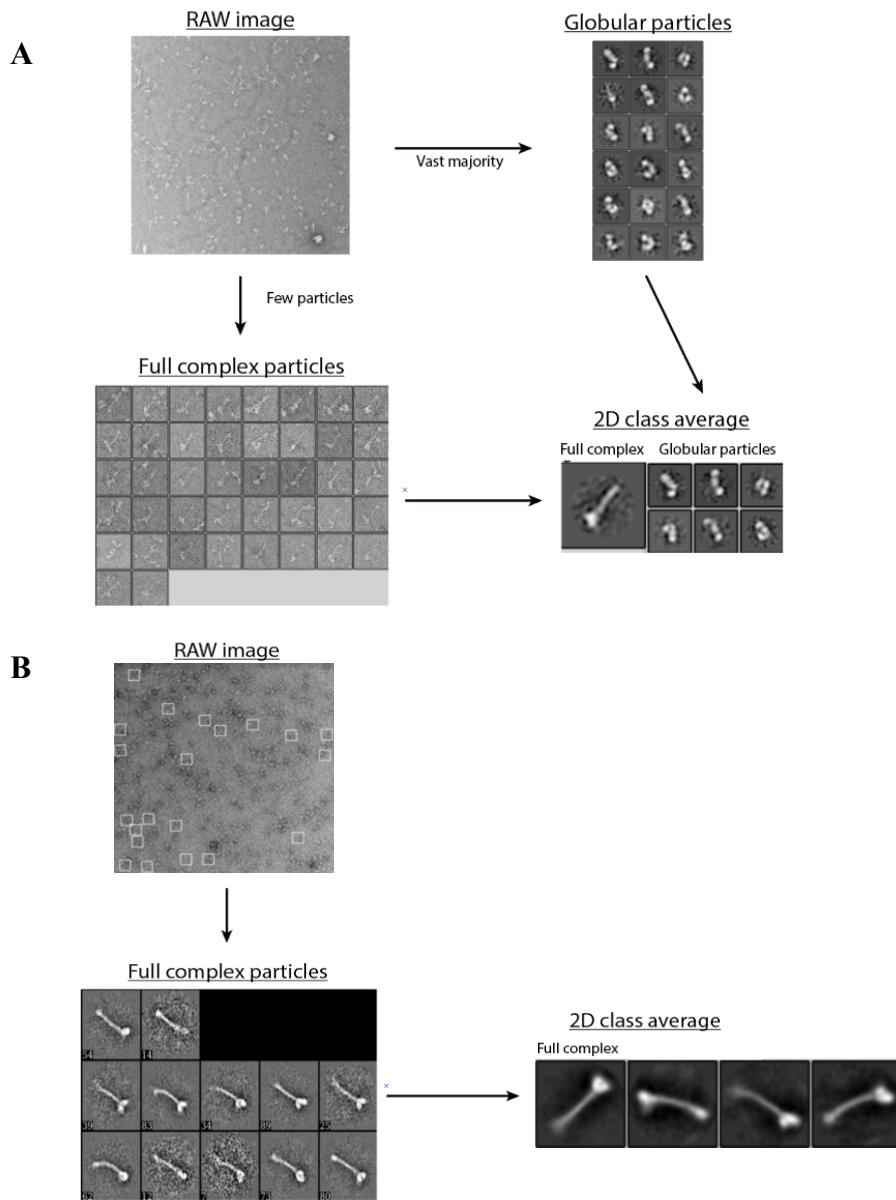
A. Negative staining electron microscopy. Sucrose density gradient purified Smc5/6 complex was processed to capture electron micrographs. Vast majority of the particles were globular collapsed complexes and thus, were discarded from analysis. A small fraction was selected as full complex particles and a 2D class average was calculated. Circular mask in full complex corresponds to 150 Å. **B.** Sucrose density gradient in in presence of glutaraldehyde, GraFix reduced the heterogeneity of complex conformations and to stabilize for the electron microscopy.