Investigation of Structure and Function of Esco1 and Esco2 Acetyltransferases

Dissertation

for the award of the degree

"**Doctor rerum naturalium**"

of the Georg-August-Universität Göttingen

within the doctoral program "Biology"

of the Georg-August University School of Science (GAUSS)

submitted by

Tahereh Ajam

from

Iran

Göttingen, 2018

Thesis Committee

Prof. Dr. Gregor Eichele, Genes and Behavior, Max-Planck-Institute for Biophysical Chemistry, Göttingen (**Mentor and Reviewer**)

Dr. Martin Kollmar, Systems Biology of Motor Proteins, Max-Planck-Institute for Biophysical Chemistry, Göttingen (**Co-reviewer**)

Members of the Extended Examination Board

Dr. Vladimir Pena, Macromolecular Crystallography, Max-Planck-Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Ahmed Mansouri, Molecular Cell Differentiation, Max-Planck-Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Henning Urlaub, Bioanalytical Mass Spectrometry, Max-Planck-Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Steven Johnsen, Translational Cancer Research, University Medical Center Göttingen, Göttingen

Date of oral examination: 22.11.2018

Affidavit

I hereby declare that the dissertation entitled "**Investigation of Structure and Function of Esco1 and Esco2 Acetyltransferases**" is submitted in support of my application for the degree of **Doctor rerum naturalium** to Georg-August-Universität Göttingen within the doctoral program Biology of the Georg-August University School of Science (GAUSS)**.** This thesis has been written independently and with no sources and aids other than quoted within texts, references and acknowledgments. This thesis has not been submitted in any previous application for any degree, elsewhere.

Göttingen, October 2018   Tahereh Ajam

To my mom

      for her boundless love

Table of Contents

Summary

Cohesin is a protein complex, whose core subunits are assembled into a ring-like structure encircling the DNA. In this manner, cohesin traps DNA molecules and plays key roles in expression, repair, and segregation of eukaryotic genomes. Esco1 and Esco2 regulate the function of cohesin by acetylation of Smc3, a subunit of the complex. Esco1 and Esco2 show distinct patterns of expression during the cell cycle. Esco1 is present constantly during the cell cycle; however, Esco2 is highly abundant during the S-phase. Both enzymes have important implications in human diseases. Mutations of Esco1 have been linked with bladder and endometrial cancer while mutations in Esco2 have been associated to Roberts syndrome (RBS), a developmental disorder with defective sister chromatid cohesion. Esco1 and Esco2 belong to the GCN5-related N-acetyltransferases (GNAT) family. While the N-terminal parts of Esco1 and Esco2 are highly divergent and likely account for the functional differences, the C-terminal GNAT acetyltransferase domains of these enzymes are conserved.

Comprehensive investigation of the catalytic mechanism of the Esco1 and Esco2 acetyltransferases is essential to understand the role of cohesin acetylation in different cellular functions. Here, we combined the *Mus musculus* Esco2 (MmEsco2) structure, *in vitro* biochemistry, and cell-based studies to identify the catalytic residues of Esco1 and Esco2 and to gain insights into the functions of these residues in catalysis.

We determined the structure of the acetyltransferase domain of MmEsco2, natively in complex with coenzyme A (CoA) at 1.8 Å resolution. To characterize the active site of Esco2, a number of potential catalytic residues were chosen considering their proximity and side chain orientation toward the CoA in the MmEsco2/CoA complex structure. Next, the functional role of these potential catalytic residues, S566, D567, E491, and S527 was investigated. For this purpose, *in vitro* mutational analysis using incubation of human recombinant cohesin complex with different variants of full-length human ESCO1 and assessment of SMC3 acetylation was performed (sufficient amounts of either mouse or human Esco2 could not be purified due to low expression and solubility). To complement the *in vitro* assay, *in vivo* mutational analysis was performed by transfection of Esco1- and Esco2-deficient mouse embryonic fibroblasts (MEFs) with different mutants of Esco1 and Esco2. Smc3 acetylation for various mutants of MmEsco1 and Smc3 acetylation along with sister chromatid cohesion for various mutants of

MmEsco2 were used as readouts. *In vivo* results from mutational analysis differed from the *in vitro* results. The single mutants were catalytically inactive in the *in vitro* assays, while the same mutants exhibited detectable activity *in vivo*. This inconsistency could be due to the absence of required cohesin regulatory factors *in vitro*. Such limitation makes it clear that an *in vivo* activity assessment of various Esco1 and Esco2 mutants is crucial in addition to *in vitro* analysis.

Taken together, the results from *in vitro* and *in vivo* mutational analysis reveal that the four conserved catalytic residues S566, D567, E491, and S527 in the active site of MmEsco2 and corresponding residues in MmEsco1 cooperatively play a role in deprotonation of the lysine substrate. In this proposed mechanism, the general bases, namely aspartate and glutamate, abstract the proton of substrate lysine *via* serine residues.

1 Introduction

1.1 Functions and regulation of cohesin

1.1.1 Cohesin and the cell cycle

The accurate distribution of the genetic material into the two identical daughter cells during cell division is crucial. Thereby, dividing eukaryote cells go through a series of phases known collectively as the cell cycle, which is controlled by numerous mechanisms and checkpoints ensuring accurate cell division (Hartwell and Weinert, 1989). Failure in this process can be the underlying cause of various human diseases. The cell cycle consists of four distinct phases: G1 (first gap phase), S (synthesis phase), G2 (second gap phase), and mitosis. During G1-phase, the cell copies organelles, grows physically larger, and makes the various component that are required for DNA synthesis. Subsequently, the cell enters S-phase and synthesizes a copy of the DNA. It also duplicates the centrosome which organizes the microtubules and is essential for DNA separation during mitosis. After completing the S-phase, the cell enters the G2-phase that is characterized by protein synthesis in preparation for mitosis. At the end of the G2-phase when the checkpoints are satisfied, the cell progresses into mitosis, including prophase, metaphase, anaphase and telophase. During prophase, the nuclear membrane breaks down to a number of small vesicles and the nucleolus fragments. The coiled chromatin begins to condense into compact chromosomes and the centrosome migrates to opposite poles of the cell. During this stage, each replicated chromosome consists of two sister chromatids that are held together by a structure called centromere. In metaphase, the chromosomes align along the equatorial plate. As the cell enters into early anaphase, chromatids initiate to separate to the opposite poles through kinetochore attachment to the spindle microtubules. Once sister chromatids have separated at the end of anaphase, the final steps of telophase and cytokinesis yield two identical daughter cells. The accuracy of the chromosome segregation depends on the fact that replicated chromatids are held together from the time they form in S-phase until their separation in anaphase. This process, which is known as sister chromatid cohesion, is a requirement for the bipolar attachment of sister chromatids to the spindle fibers during mitosis. Absence of this cohesion could result in early separation of sister chromatids from each other, before the chromatids attach to both poles of the spindle. Early separation of sister chromatids in turn could lead to an unequal distribution of sister chromatids to the daughter cells. Cohesion is present along the sister chromatid arms and at centromeres. However, be-

cause of the importance of accurate microtubule attachment to the kinetochores, cohesion is especially crucial at centromeres.

Two general mechanisms for cohesion have been proposed: First, a persistent catenation of sister DNA molecules, which physically interlocks (catenates) DNA across the sister chromatids (Murray and Szostak, 1985). Second, a protein connection of DNA molecules *via* the cohesin complex, which physically tether the sister chromatids (Michaelis et al., 1997; Nasmyth, 2009).

In support of the second mechanism, genetic screens of a variety of mutants, defective for sister chromatid cohesion, in various species identified a number of proteins involved in cohesion (Birkenbihl and Subramani, 1992; Davis, 1971; Guacci et al., 1997; Holt and May, 1996; Kerrebrock et al., 1992; Michaelis et al., 1997). These proteins were found to be subunits of the cohesin complex or regulators of this complex and are highly conserved among different eukaryotes.

The cohesin complex is loaded onto chromatin in G1 (yeast) or telophase (vertebrates) by Nipbl-Mau2 heterodimer (Scc2-Scc4 in yeast). This process is highly reversible due to presence of the unloading factors (Wapl-Pds5). During S-phase, acetylation of cohesin by acetyl transferases Esco1 and Esco2 (Eco1 in yeast) establishes cohesion between two sister chromatids. Cohesin remains chromatin associated through G2-phase. During mitosis, cohesins which are located between the arms of sister chromatids are released in prophase and the remaining cohesins at centromeres are dissociated in anaphase. This allows proper sister chromatid resolution and efficient segregation.

1.1.2   Cohesin complex architecture

The cohesin complex is evolutionarily conserved throughout eukaryotes (Table 1) and consists of four core subunits (Anderson et al., 2002; Losada et al., 1998; Peters et al., 2008; Sumara et al., 2000). Two cohesin core subunits, Smc1 and Smc3, belong to the Smc family of chromosomal ATPases (ATP-binding cassette [ABC] family) (Figure 1). Smc proteins have a rod-like shape with a hinge domain on one end and an ATPase head domain at the other end. A long antiparallel coiled-coil separates the hinge domain from the head domain. Interaction of Hinge domains Smc1 and Smc3 leads to the formation of a Smc heterodimer. The

third main cohesin subunit, Rad21 (Scc1 in yeast), belongs to the kleisin protein family. The C-terminal of the Rad21 subunit binds to the Smc1 head domain and the N-terminal of this subunit binds to the Smc3 head domain. In this way, Rad21 creates a triple ring. The forth subunit, SA1 or SA2 (Scc3 in yeast), binds to the middle region of Rad21 (Gligoris et al., 2014; Haering et al., 2008; Haering et al., 2002; Hara et al., 2014; Hirano and Hirano, 2002; Melby et al., 1998). The ring-like structure of the cohesin complex with a diameter of ~50 nm has been shown by electron microscopic images of this complex, either purified from cell extracts or reconstituted using recombinant proteins (Anderson et al., 2002; in 't Veld et al., 2014). This ring is big enough to topologically trap the two sister chromatids and thereby keep them together (Haering et al., 2008; Haering et al., 2002; Ivanov and Nasmyth, 2005; Murayama and Uhlmann, 2014).
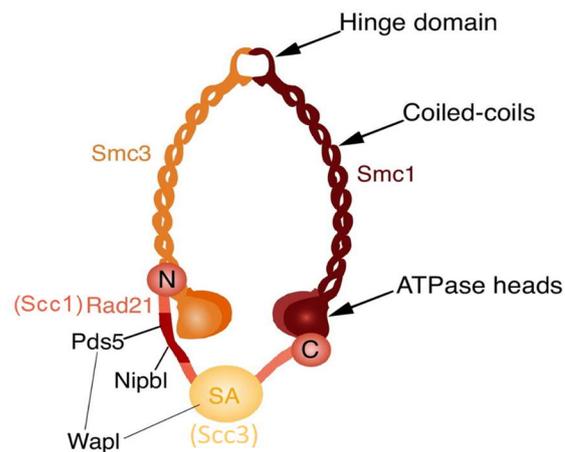
Figure 1: Architecture of the cohesin complex.
In vertebrates, the cohesin core complex consists of Smc1, Smc3, Rad21, and either SA1 or SA2. The names of the orthologs in yeast are shown in parentheses. Smc proteins fold back on themselves with the hinge domain at one end and the ATPase domain at the other end. Smc1 and Smc3 hinges interact directly together while their head domains connect *via* the kleisin subunit Rad21. The SA subunit associates with the middle region of the Rad21. Known interaction sites for cohesin regulators Pds5, Nipbl, and Wapl are shown. Nipbl and Pds5 bind the same part on Rad21 (darkertone). The illustration is modified from Morales and Losada (2018).

Table 1: Cohesin subunits and regulatory protein homologs.

	Budding yeast	Fission yeast	Mammals	Xenopus	C. elegans	Drosophila
Cohesin Subunits	Smc1	Psm1	SMC1a	XSMC1	Him1	DCAP
	Smc3	Psm3	SMC3	XSMC3	Smc3	Smc3
	Scc1/Mcd1	Rad21	RAD21	XRAD21	Coh2/Scc1	DRAD21
	Scc3/Irr1	Psc3	SA1/STAG1 SA2/STAG2	XSA1, XSA2	Scc3	DSA1
Loading	Scc2	Mis4	NIPBL	SCC2	Pqn-85	Nipped-B
	Scc4	Ssl3	Mau2/Scc4	XSCC4	Mau2	Mau2
Establishment	Ctf7/Eco1	Eso1	ESCO1, ESCO2	XECO1, XECO2		Deco/San
Maintenance	Pds5	Pds5	PDS5A, PDS5B	PDS5A, PDS5B	Pds5/Evl14	Pds5
	Rad61	Wpl1	WAPL			
Dissolution	Pds1	Cut2	Securin	Securin		PIM
	Esp1	Cut1	Separase	Separin		SSE/THR
	Cdc5	Plo1	PLK1	PLX1		POLO
	Sgo1	Sgo2	Shugoshin/ SGOL1, SGOL2	Shugoshin-like1 (xSGO1)		MEI-S332

1.1.3 Cohesin functions

As noted above, there are several lines of evidence that the cohesin ring traps the two sister chromatids and functions as molecular glue ensuring equal segregation of sister chromatids during mitosis (Ciosk et al., 2000; Losada et al., 1998; Tanaka et al., 2000) and meiosis (Klein et al., 1999).

Beyond cohesion, different reports have shown that cohesin mutants are defective in DNA damage repair. Following DNA damage, cohesin normally accumulates on the DNA double-strand breaks (DSB) sites and allows postreplicative homologous recombination repair of DNA DSBs (Kim et al., 2002; Sjogren and Nasmyth, 2001).

In addition, cohesin has been shown to be important for CTCF-dependent gene regulation, which is known to be mediated by chromatin looping. Since cohesin is able to physically connect DNA strands, it has been proposed that it could form or stabilize chromatin loops (Wendt and Peters, 2009; Wendt et al., 2008). This hypothesis has been supported by chromatin con-

formation capture (3C) experiments which revealed that long-range chromosomal interactions are cohesin dependent (Hadjur et al., 2009; Kagey et al., 2010; Nativio et al., 2009).

1.1.4 Cohesin loading and unloading

Cohesin rings, *via* entrapping DNA, regulate chromosome segregation, transcription, and DNA repair. All these functions are dependent on the proper regulation of cohesin association and disassociation on chromatin (Haarhuis et al., 2014a).

The cohesin ring has two outer gates: (1) the interface between the head domain of Smc3 and the N-terminal of Rad21 (exit gate) and (2) the interface between the hinge domains of Smc1 and Smc3 (entry gate) (Figure 2) (Gruber et al., 2006; Haarhuis et al., 2014b; in 't Veld et al., 2014). In addition, cohesin has a third 'inner gate' between the head domains of Smc1 and Smc3, which is regulated by binding and hydrolysis of ATP (Figure 2). Binding of two ATP molecules to the head domains of Smc1 and Smc3 closes the inner gate. Hydrolysis of these ATP molecules on the other hand drives the head domains apart and opens this gate (Beckouet et al., 2016; Murayama and Uhlmann, 2015).

Cohesin is loaded onto chromatin as vertebrate cells enter G1-phase by a mechanism that depends on the heterodimer Nipbl-Mau2 (Scc2-Scc4 in budding yeast), whereas it is released from the chromatin by the Wapl and Pds5 regulators (Chao et al., 2017a; Ciosk et al., 2000; Tedeschi et al., 2013; Weitzer et al., 2003).

It is believed that the heterodimer, Nipbl-Mau2, promotes loading of cohesin onto chromatin by stimulating ATP hydrolysis of the Smc heads (Chao et al., 2015; Chao et al., 2017a; Weitzer et al., 2003). However, in the absence of this cohesin loader *in vitro*, cohesin can still bind topologically to DNA to some extent (Murayama and Uhlmann, 2014). It has been proposed that cohesin entraps DNA through an entry gate between the Smc1 and Smc3 hinge domains (Figure 2) (Gruber et al., 2006). The coiled-coils of Smc might transmit conformational changes induced by ATP hydrolysis from the head to the hinge domains and promote entry gate opening (Nasmyth, 2011). The heterodimer Nipbl-Mau2 interacts with SA1/2. Hence, SA1/2 subunit connects the cohesin loader to the Smc head domains and increases both the ATPase activity of cohesin and its loading onto chromatin. However, the SA1/2 still stimulates the cohesin loading and ATPase activity in the absence of cohesin loader to some

extent. This suggests that SA1/2 also affects the loading process independent of Nipbl-Mau2 (Murayama and Uhlmann, 2014).

It has been proposed that after cohesin loading onto DNA, ATP rebinding closes the inner gate of cohesin and locks DNA between the Smc1 and Smc3 coiled-coils. The entrapped DNA in turn stimulates ATPase activity of Smc3 head domain, likely by direct interaction with a basic patch on this domain, harboring the two conserved lysines, K105 and K106 (Camdere et al., 2015; Murayama and Uhlmann, 2015; Yu, 2016). Subsequently, ATP hydrolysis and nucleotide release result in dissociation of the Smc3 and Smc1 head domains and allow passage of the DNA to the inner gate (Beckouet et al., 2016; Camdere et al., 2015; Elbatsh et al., 2016). Next, new ATP molecules bind the Smc head domains and close the inner gate. After this step, Wapl-Pds5 dissociates the Smc3-Rad21 dimer and hence, DNA is released from the cohesin ring through the exit gate (Ouyang and Yu, 2017).

In vivo analyses the of cohesin binding pattern to chromosomes have shown that the main population of cohesin complexes are loaded around centromeres, in addition to gene promotors along chromosome arms (D'Ambrosio et al., 2008; Ocampo-Hafalla et al., 2007). In mammalian cells, many of the cohesin binding sites overlap with the CTCF and are different from their loading sites. Translocation of cohesin from promoters to CTCF-binding sites is not well understood (Parelho et al., 2008; Uhlmann, 2016; Wendt et al., 2008). Recent studies suggest that loop extrusion could play an important role in this process (Barrington et al., 2017).
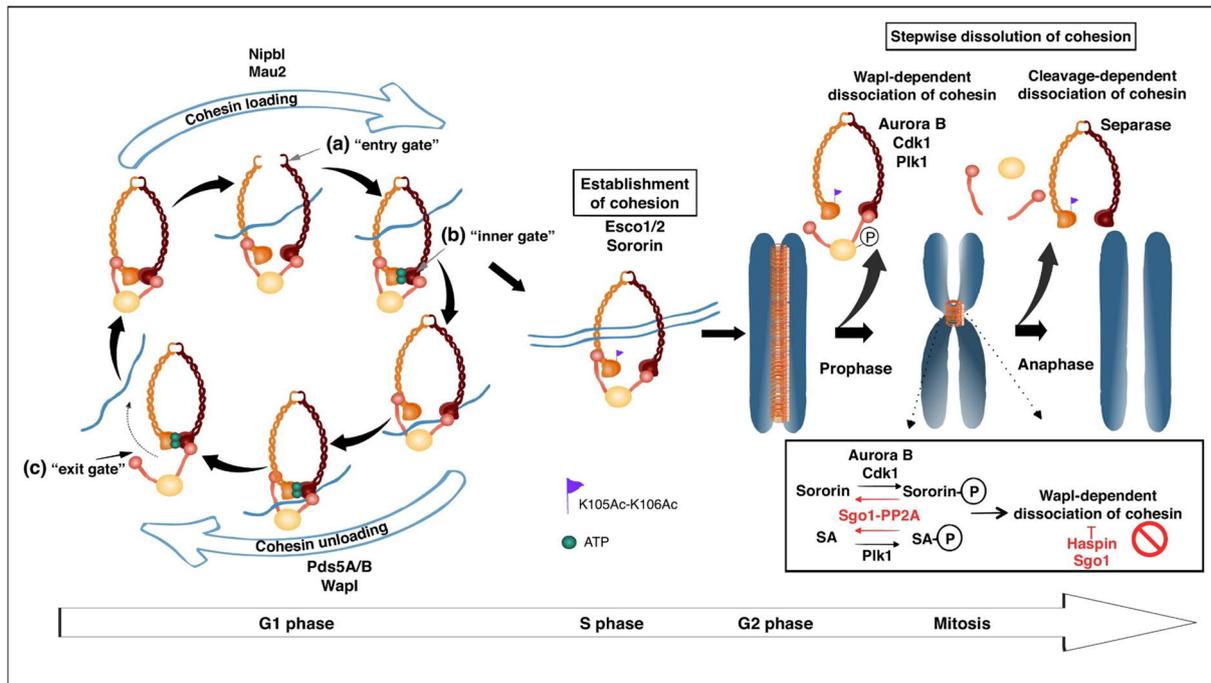
Figure 2: Cohesin loading and unloading onto chromatin during the cell cycle in vertebrate cells.
Left: in G1-phase, the Nipbl-Mau2 dimer is loading cohesin onto chromatin by stimulating the ATP hydrolysis of the Smc head domains. The entry gate is located between the hinge domains of the Smc1 and Smc3 proteins (a). The inner gate resides between the two Smc head domains. This gate closes and opens by ATP binding and hydrolysis, respectively (b). The unloading process depends on Wapl and Pds5. DNA is released *via* an exit gate located between the Smc3 head domain and Rad21 (c). During DNA replication in S-phase (middle), cohesion establishment occurs *via* Smc3 acetylation by Esco1 and Esco2. Subsequently, acetylated cohesin recruits Sororin, which antagonizes Wapl and mediates sister chromatid cohesion. **Right**: Dissociation of cohesin during mitosis proceeds in two steps. The majority of cohesin is released from the chromosome arms during prophase by Wapl. In this phase, cohesin and Sororin become phosphorylated by Cdk1 and Aurora B kinases. At centromeres, cohesin is protected from dissociation by the Sgo1–PP2A complex, which dephosphorylates cohesin and Sororin in this region. In addition, Sgo1 and Haspin compete with binding of Wapl to cohesin and Pds5, respectively. Finally, during anaphase, centromeric cohesin is released by Separase-mediated cleavage of Rad21 that allows the accurate separation of the sister chromatids. The illustration is adapted from Morales and Losada (2018).

1.1.5 Cohesion establishment

In S-phase, cohesion can be established once DNA has been replicated in a process that depends on the acetylation of the two conserved lysine residues (K105 and K106) on the basic patch of the Smc3 head domain. It has been proposed that acetylation of these two lysine residues neutralize the positive charge of the basic patch, thereby weakening the DNA binding and reducing ATPase activity. This stops the release of DNA through the cohesin exit gate, counteracts Wapl and Pds5 functions and leads to stable DNA entrapment inside the cohesin ring (Ben-Shahar et al., 2008; Murayama and Uhlmann, 2015; Unal et al., 2008; Yu, 2016). In

vertebrates, cohesion establishment involves an additional component, Sororin, which competes with Wapl for binding to Pds5 and in this way antagonizes the Wapl-Pds5 releasing activity (Carretero et al., 2013; Nishiyama et al., 2010). Smc3 acetylation is essential for Sororin recruitment to cohesin.

How cohesin holds two sister chromatids together is not well understood. Protein crosslinking experiments have provided evidence for a "one ring model", in which one cohesin ring entraps both sister chromatids. Alternatively, a "handcuff model" has been proposed. This model suggests that each of two sister chromatids is entrapped by one cohesin ring, which interconnect to each other. A few other models have also been proposed, including fusion of the two cohesin rings to make a larger ring that embraces sister chromatids (Haering et al., 2008; Zhang et al., 2008; Zhang and Pati, 2015). To achieve embracing of both sister chromatids by one cohesin ring (one ring model) one possibility could be that the replisome passes through DNA-bound cohesin rings. Because of the size restrictions, presumably the replisome cannot simply pass through the cohesin ring (Stigler et al., 2016). Instead, cohesin could capture both sister chromatids in the vicinity of the replication fork (Lengronne et al., 2006). This possibility, may link cohesin *de novo* loading and cohesin acetylation and therefore explain the requirement of cohesin's ATPase activity for Smc3 acetylation (Ladurner et al., 2014).

Acetylation of the two conserved lysine residues in the Smc3 head domain is mediated by the cohesin acetyltransferases Esco1 and Esco2 (Eco1 in yeast). Esco1 and Esco2 both consist of a divergent N-terminus, a C2H2 zinc finger and a conserved C-terminus acetyltransferase domain (Hou and Zou, 2005). The relative contributions of these paralogs to cohesin regulation are not entirely clear. Several studies have suggested that both Esco1 and Esco2 are engaged in sister chromatid cohesion, as depletion of both enzymes in cells results in sister chromatid cohesion defects which are more severe than either single depletion (Hou and Zou, 2005; Kawasumi et al., 2017; Minamino et al., 2015). In contrast, some evidence suggests that Esco2 is mainly responsible for the establishment of cohesion, possibly *via* recruitment of Sororin to cohesin and stabilizing cohesin around sister chromatids during S-phase. In contrast, Esco1 may contribute to gene regulation likely by stabilizing cohesin around single chromatids during interphase (chromatin looping) (Rahman et al., 2015). Esco1 and Esco2 show distinct patterns of expression during the cell cycle. Esco1 is constantly present throughout the cell cycle, while Esco2 is highly abundant during the S-phase. Esco2 is a sub-

strate of the anaphase promoting complex/cyclosome (APC/C), an E3 ubiquitin ligase that is activated at mitotic exit (Lafont et al., 2010; Song et al., 2012). Thus, Esco2 levels are low in mitosis and G1-phase, and only increase as APC activity reduces during S-phase. It has been shown that Esco1 directly interacts with cohesin *via* Pds5. Such a direct interaction with cohesin has not been observed for Esco2 (Minamino et al., 2015). Instead, Esco2 interacts with the replication proteins, proliferating cell nuclear antigen (PCNA) (Higashi et al., 2012; Song et al., 2012) and minichromosome maintenance protein complex (MCM) (Ivanov et al., 2018; Minamino et al., 2018).

Similar to cohesin and its regulatory subunits, Esco1 and Esco2 also are associated with human developmental disorders (Krantz, 2014; Liu and Krantz, 2008). Mutations in or overexpression of *Esco1* have been associated with endometrial and bladder cancer, respectively (Price et al., 2014; Zhang et al., 2016). Mutations in *Esco2* have been associated with Roberts syndrome (RBS), a childhood autosomal recessive disorder (Gordillo et al., 1993; Gordillo et al., 2008; Vega et al., 2005). RBS patients are characterized by various degree of mental retardation and a number of dysmorphologies. Most of *Esco2* mutations involve premature stop codons in the N-terminal part of the enzyme. These mutations lead to the expression of a truncated Esco2 protein with no enzymatic activity (Gordillo et al., 2008). Metaphase chromosomes from RBS patient show a loss of cohesion in the pericentric heterochromatin (PCH) while cohesion is maintained on the arms (Van Den Berg and Francke, 1993). These chromosomes show a parallel alignment of sister chromatids that in combination with repulsion in PCH, results in a 'railroad track' appearance of chromosomes (Maserati et al., 1991). These findings indicate that Esco2 function must be compensated by Esco1 to some extent at least in human cells, since its deficiency is compatible with life (Vega et al., 2005). By contrast, *Esco2*-deficient mouse embryos die early in development (Whelan et al., 2012). Similar to human RBS patient cells, *Esco2*-deficient mouse embryonic fibroblasts (MEFs) show severe centromeric cohesion defects while cohesion is sustained along the arms. Therefore, Esco2 could play the key role in the establishment of cohesion around centromeres while Esco1 might be involved in cohesion along arms in the absence of Esco2 (Whelan et al., 2012).

1.1.6 The prophase pathway of cohesin dissociation

In most eukaryotic cells, cohesin dissociation takes place during two phases of mitosis (Losada et al., 1998; Peters et al., 2008; Sumara et al., 2000; Waizenegger et al., 2000). The

first dissociation takes place during prophase and prometaphase, when most of the cohesin dissociates from the chromosome arms. The second phase occurs at the onset of anaphase, when the residual cohesin on chromosomes, mostly at centromeres, dissociates because of cleavage of the Rad21 subunit by Separase (Hauf et al., 2001; Kumada et al., 2006; Nakajima et al., 2007; Uhlmann et al., 1999; Uhlmann et al., 2000; Wirth et al., 2006). Several proteins and events are required for the prophase pathway of cohesin unloading. A number of studies indicate that polo-like kinase 1 (Plk1) contributes to cohesin dissociation by phosphorylation of the SA subunit (Hauf et al., 2005; Morales and Losada, 2018). Other mitotic kinases, e.g. Aurora B and cyclin-dependent kinase 1 (Cdk1), are also required for efficient dissociation of cohesin from chromosomes in prophase (Gimenez-Abian et al., 2004; Losada et al., 2002). These kinases phosphorylate Sororin, which results in the dissociation of this protein from Pds5. Subsequently, Wapl replaces Sororin, binds to Pds5 and opens the cohesin ring (Nishiyama et al., 2013). Together, these mitotic phosphorylation events are thought to activate the cohesin unloading process in prophase. Cohesin at the centromere is protected from this activity by dephosphorylation. The protein Shugoshin (Sgo1) is targeted to the centromere by Aurora B and budding uninhibited by benzimidazoles 1 (Bub1) kinase, recruiting the protein phosphatase 2A (PP2A) (Huang et al., 2007; Kitajima et al., 2004; Kitajima et al., 2006; Riedel et al., 2006). This causes localized dephosphorylation of both cohesin and Sororin, thus inhibiting cohesin unloading at the centromere (Liu et al., 2013; McGuinness et al., 2005). Another mitotic kinase, Haspin, is also recruited to centromere and outcompetes Wapl binding to Pds5. Haspin and Bub1 phosphorylate histones H3 and H2A, respectively, resulting in the recruitment of the chromosomal passenger complex (CPC), which is essential for centromeric cohesion (Hengeveld et al., 2017).

At the end of metaphase, when cells have bioriented all of their chromosomes on the mitotic spindle, the spindle checkpoint signaling is stopped and the Anaphase-promoting complex/cyclosome (APC/C) becomes active. This results in ubiquitylation and subsequent degradation of several APC/C substrates, including the Separase inhibitors: Securin and the activating subunit of Cyclin-dependent kinase 1 (Cdk1). These reactions lead to activation of separase (Peters, 2002). Before APC/C is active, Separase is inhibited in interphase and early mitosis by interaction with Securin (Hornig et al., 2002; Waizenegger et al., 2002). Separase in vertebrates is additionally inhibited by Cdk1-mediated phosphorylation and by interaction with Cdk1's cyclin B subunit (Gorr et al., 2005; Huang et al., 2008; Huang et al., 2005;

Stemmann et al., 2001). In vertebrate cells, APC/C activation thus releases Separase from two inhibitory mechanisms *via* ubiquitylation of Securin and cyclin B. As soon as Separase turns into its active state it cleaves the Rad21 subunit, which results in opening of the cohesin ring, dissociation of cohesin from chromosomes and separation of sister chromatids (Uhlmann et al., 1999; Uhlmann et al., 2000).

1.2 Smc3 acetylation by Esco1 and Esco2

1.2.1 Esco1 and Esco2 belong to GCN5 family

Acetylation reactions, catalyzed by several groups of enzymes, play a key role in various biological processes including intracellular localization, enzyme activity, protein-protein interactions, protein stability and transcriptional regulation. The best characterized acetyltransferase enzymes are histone acetyltransferases (HATs), which catalyze histone acetylation and play a role in many epigenetic processes (Eberharter and Becker, 2002; Yang and Seto, 2007). HATs are classified into different subfamilies including Histone acetyltransferase 1 (HAT1), Gcn5/PCAF, MYST, CBP/p300, and Rtt109, based on sequence and substrate acetylation profiles. All HATs share a structurally similar acetyl coenzyme A (AcCoA) binding site (Tanner et al., 2000a; Tanner et al., 2000b; Wang et al., 2008). It has been shown that some of the HATs, such as members of the CBP/p300 and MYST families, also acetylate non-histone proteins. Esco1 and Esco2 belong to the GCN5-related N-acetyltransferase (GNAT) family. Despite the highly divergent N-terminal parts of Esco1 and Esco2, which presumably are the cause of their functional differences, the C-terminal GNAT acetyltransferase domains of these two enzymes are considerably conserved (Hou and Zou, 2005; Ivanov et al., 2002; Neuwald and Landsman, 1997; Roth et al., 2001). Esco1 and Esco2 contain a zinc-finger (ZnF) domain that is similar to those found in other HAT family members. The ZnF typically mediates both DNA binding and protein interactions in HATs (Akhtar and Becker, 2001; Koehler et al., 2014; Toleman et al., 2006). However, it has also been shown that the ZnF is important for acetyltransferase activity in Esco1 and Esco2 (Onn et al., 2009).

1.2.2 GNAT family structure, acetyl-CoA binding and substrate recognition

Members of the GNAT superfamily have been identified in various organisms. These enzymes catalyze the transfer of an acetyl group from AcCoA to the primary amine moiety of a wide range of substrates such as glucosamine 6-phosphate, aminoglycoside antibiotics, spermidine, spermine, dopamine, histones and non-histone proteins (Dyda et al., 2000; Majorek et al., 2013; Vetting et al., 2005). Protein acetyltransferases from *Sulfolobus solfataricus* (SsPAT) (Brent et al., 2009), human α-tubulin acetyltransferase 1 (αTAT1) (Friedmann et al., 2012), human Naa50p (Liszczak et al., 2011) and *M. tuberculosis* AcCoA synthetase N-acetyltransferase (Rv0998) (Lee et al., 2012) are representative members of this family that acetylate histone and non-histone proteins (Salah Ud-Din et al., 2016). All of these enzymes acetylate the amino group of a lysine in the protein substrate, except for Naa50p, which transfers the acetyl group to the α-amino group of N-terminal methionine in protein substrates. The structures of these enzymes have been determined: They consist of a β-sheet-helix core region which is structurally conserved among all HATs (colored blue in Figure 3), irrespective of the sequence conservation (Salah Ud-Din et al., 2016). This core region is flanked by variable N- and C-terminal structural motifs (colored green in Figure 3), which are likely responsible for substrate recognition. The AcCoA cofactor is wedged in a groove between structurally conserved α-helices (Figure 3). Residues that contribute to the protein-AcCoA interaction are typically not conserved. Similar to other HATs, these acetyltransferases use CoA both as an acetyl donor in enzymatic reactions and as a molecule to stabilize the overall folding of the acetyltransferase domain (Friedmann and Marmorstein, 2013; Trievel et al., 1999; Yuan et al., 2012).
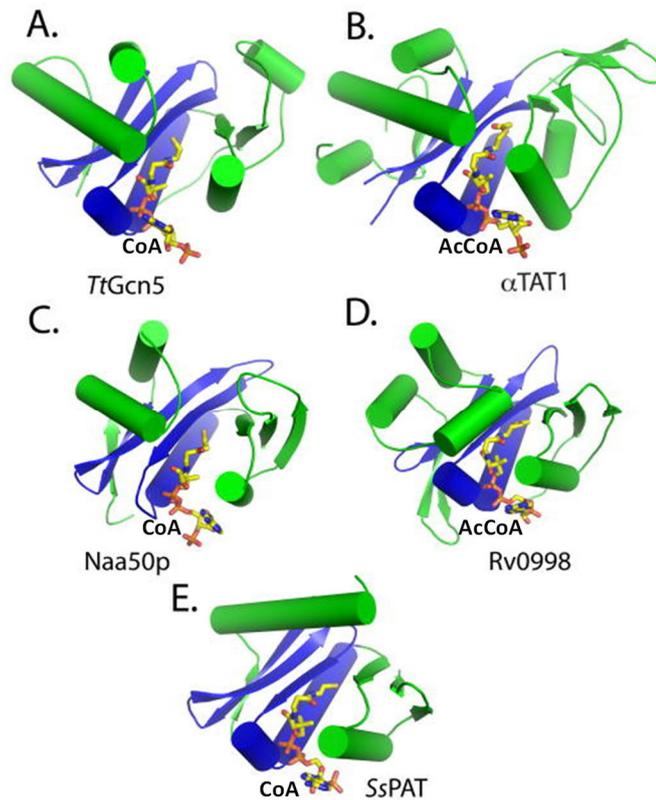
Figure 3. Structure of non-histone acetyltransferases.
Ribbon representation of acetyltransferase domains of **(A)** *Tetrahymena GCN5 (*TtGcn5), **(B)** αTAT1, **(C)**, Naa50p, **(D)** Rv0998, and **(E)** *Ss*PAT. The conserved acetyltransferase core region is colored blue, and the variable flanking fragments are colored green. AcCoA and CoA are shown as stick and colored according to element: carbon, yellow; nitrogen, blue; oxygen, red. Figure adapted from Friedmann and Marmorstein (2013).

1.2.3   GNAT family catalysis mechanism

AcCoA dependent acetyltransferases utilize one of two catalytic mechanisms, the ping-pong or the sequential mechanism (Figure 4). The Ping-pong mechanism involves the formation of an acetylated enzyme intermediate after binding and reaction with AcCoA (Figure 4, upper panel). The product (CoASH) is released; protein substrate binds, and the ϵ-amino group of lysine reacts with the intermediate to generate the final acetylated protein product. In a sequential mechanism (Figure 4, lower panel), both AcCoA and the substrate bind to the enzyme and form a ternary complex allowing the lysine to directly attack the bound AcCoA, without the formation of a covalent enzyme intermediate.

Ping-Pong Reaction

Sequential Reaction
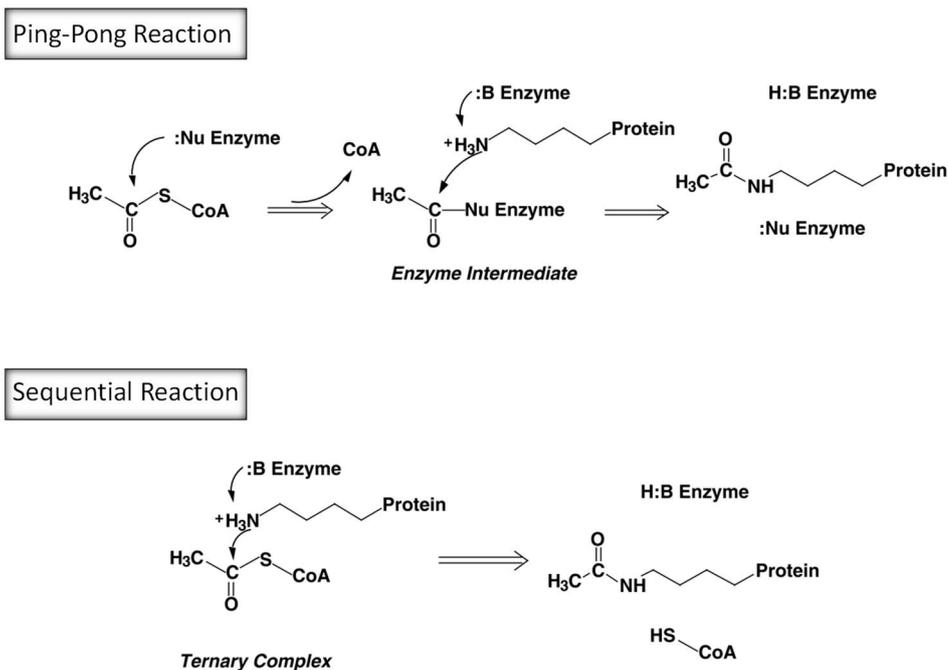
Figure 4: Catalytic mechanisms for acetyltransferase enzymes.
The reaction at the top shows the ping-pong catalytic mechanism in which an acetyl-enzyme intermediate is formed. The reaction at the bottom shows the sequential catalytic mechanism. In this mechanism, both AcCoA and protein bind to the enzyme forming a ternary complex, prior to any chemical step. The figure was adapted from Berndsen and Denu (2005).

Several lines of evidence show that the GNAT family uses a sequential mechanism (Tanner et al., 2000a; Tanner et al., 2000b; Tanner et al., 1999; Yan et al., 2002), using a conserved glutamate residue as a general base (:B in Figure 4) (Clements et al., 1999; Trievel et al., 1999). Gcn5 is one of the GNAT enzymes that uses a glutamate as a general base for catalysis and has a very efficient turnover number (k_{cat} = 210 min^{-1}) (Poux et al., 2002). Structural and kinetic analyses have shown that Rv0998 uses Glu235, which is positioned in similar way as the glutamate of Gcn5 (Lee et al., 2012). Interestingly, some of the Gcn5 family members do not use this particular glutamate for deprotonation as is the case in SsPAT, Naa50p and αTAT1. SsPAT has a glutamate residue (E76) at the equivalent position in its structure, but SsPAT does not solely use this residue as a general base for deprotonation. Instead, this enzyme relies on a number of additional residues (Y38, E42, E43, D53, H72, E76) which function as a "proton wire" to deprotonate the substrate lysine (Brent et al., 2009). The rate of reaction by SsPAT (k_{cat} = 2 min^{-1}) is much slower than that of Gcn5. In addition, studies on Naa50p sug-

gest that catalysis by this enzyme does not rely on one particular residue. Naa50p utilize a tyrosine and histidine (Y73 and H112) to deprotonate the amino group of the substrate *via* a water molecule (Liszczak et al., 2011). This reaction appears to occur at $k_{cat} = 7$ min^{-1}. Different reports have proposed an acetyl transfer mechanism involving several catalytic residues also for αTAT1 that uses D157, C120 and Q58 as general bases (Friedmann et al., 2012; Taschner et al., 2012). Kinetic experiments revealed that αTAT1 has a very inefficient catalytic rate ($k_{cat} = 12$ min^{-1}) (Friedmann et al., 2012; Taschner et al., 2012). In addition to these non-histone protein acetyltransferase enzymes, some of the GNAT members, which acetylate small molecules, have been shown to use more than one catalytic residue. Dopamine N-acetyltransferase from *D. melanogaster* (Dat) is an example for a group of enzymes which catalyze the acetylation reaction using E47 and S182 (Cheng et al., 2012).

1.3 Aims of the study

In mammals, Esco1 and Esco2 acetylate Smc3, which subsequently counteracts Wapl and Pds5 function and thereby stabilizes cohesin on chromatin. In this way, Esco1 and Esco2 are engaged in various cellular functions of cohesin, such as sister chromatid cohesion, regulation of gene expression and DNA repair. Esco1 and Esco2 belong to the GNAT family of histone acetyltransferases. The C-terminal acetyltransferase domains of these paralogs show a high degree of homology with 60% sequence identity (Hou and Zou, 2005; Ivanov et al., 2002; Neuwald and Landsman, 1997).

Thorough investigation of the catalytic mechanism of the acetyltransferases Esco1 and Esco2 is essential to understand the role of cohesin acetylation in different cellular functions. Before we started to work on this project, no structural and comprehensive functional study of these enzymes was available. Thus, the main goal of this thesis was to combine the Esco1 or/and Esco2 structures, *in vitro* biochemistry and cell-based studies in order to identify the catalytic residues of these enzymes and to gain insights into their functions in catalysis. In the course of completing this thesis, three studies were published about structures, along with limited bio-chemical characterization of *Homo sapiens* ESCO1 (HsESCO1) and its ortholog in *Xenopus* Eco2 (xEco2) (Chao et al., 2017b; Kouznetsova et al., 2016; Rivera-Colon et al., 2016). One study proposed a substrate-assisted catalytic mechanism for HsESCO1 (Kouznetsova et al., 2016). The authors argued that HsESCO1 lacks a carboxylic acid side chain within the active site that might function as a general base. The authors propose that HsESCO1 uses an aspartate (D107) of the substrate Smc3 for deprotonation of the neighboring lysine residues K105 and K106. However, Smc3 D107 is located on the opposite side of the ϵ-amino group of the targeted lysines in the xEco2/Smc3 peptide complex structure and interacts with two conserved residues of xEco2. This implies that D107 plays a role in enzyme binding rather than catalysis (Chao et al., 2017). Another group combined a structure-guided study of HsESCO1 with site-directed mutagenesis. They studied a number of potential catalytic residues using *in vitro* acetylation assay with a Smc3 peptide as substrate and reported that D810 in HsESCO1 could play a role as a general base (Rivera-Colon et al., 2016). However, this residue is not conserved in ScEco1 (Figure 7), which suggests that in yeast some other residues carry out substrate lysine deprotonation.

In our study, we aimed to overcome the limitations of the previous studies by using *in vivo* and *in vitro* assays of Esco1/2 catalysis along with structure that we determined independently. To achieve this goal we did the following:

1) Determine the structure of the murine Esco2 acetyltransferases because our experimental, cell-based analysis are conducted with the murine model.

2) From the structure, determine the architecture of the active site and compare this structure with other acetyltransferases and with published Esco structures from other organisms. This includes identification of the general bases that capable of abstracting the proton from the ϵ-amino group of Smc3 substrate lysines.

3) *In vitro* Smc3 acetylation using recombinant wild type and mutants Esco1 or/and Esco2 protein. These *in vitro* studies are based on use of the entire cohesin ring as substrate.

4) *In vivo* Smc3 acetylation experiments with wild type murine Esco1 and with a series of single and double mutants of residues that may be involved in catalysis.

5) *In vivo* Smc3 acetylation along with complementation assays, sister chromatid cohesion and Aurora B localization, with wild type and mutants MmEsco2.

2 Material and Methods

2.1 Antibodies

Rabbit antibody against MmEsco1 was generated in the laboratory of Gregor Eichele using a haemocyanin-conjugated peptide comprising amino acids 521 to 606 of mouse Esco1 (1:1000). The following previously described custom-made antibodies were used: anti-Esco2 (Whelan et al., 2012) (1:1000), mouse anti-acetyl-Smc3 (a gift from K. Shirahige) (Nishiyama et al., 2010) (1:1000). The following commercial antibodies were used: rabbit anti-Smc3 (Cell Signaling D47B5, 1:3000), conjugated mouse anti-TBP (Abcam 197874, 1:5000), conjugated mouse anti-His tag (Novus 31055H, 1:1000), anti-Aurora B (Sigma-Aldrich A5102, 1:100, for immunofluorescence).

2.2 Protein expression in insect cells

2.2.1 Cloning and mutagenesis

Truncated mouse $Esco2^{368-592}$ with a C-terminal His-tag was cloned into the pFL vector. Full-length *HsEsco1* and *HsEsco2* were cloned into the pFastbac-HTC vector (Invitrogen) with an N-terminal His-tag. Standard restriction-ligation methods were used to introduce these cDNAs into the plasmids. Briefly, cDNAs were amplified using PCR and purified using ethanol precipitation. Vectors and purified PCR products were digested for 4 h at 37 °C using restriction endonucleases (New England Biolabs, NEB). Digestion of the $MmEsco2^{368-592}$ PCR product and pFL vector were performed using EcoRI/HindIII restriction enzymes, digestion of *HsEsco1* and *HsEsco2* PCR products and pFastbac-HTC vector were done using BamHI/HindIII and SalI/BamHI, respectively. Digested vectors were subsequently dephosphorylated by addition of antarctic phosphatase (NEB). DNA fragments were separated by agarose gel electrophoresis (0.7% w/v agarose) and extracted using the QIAquick gel extraction kit (Qiagen). PCR product and linearized vector (1:5 molar ratio of insert to vector) were ligated overnight at 16 °C using T4 DNA ligase (NEB). After ligation, DNA was transformed into DH10B Cells (NEB) using electroporation. Afterwards, cells were plated on Lysogeny broth (LB) agar plates containing the corresponding antibiotics for selection of transformed cells and incubated overnight at 37 °C. A single colony was used to inoculate 4 ml of Lysogeny broth medium (LB-medium) containing the corresponding antibiotic. The culture

was grown overnight at 37 °C and plasmids were isolated using the spin miniprep kit (Qiagen).

HsScc1 was cloned into a 438-C vector, containing an N-terminal His-tag followed by a maltose binding protein (MBP) tag and a tobacco etch virus (TEV) protease cleavage site. Cloning into this vector was performed using ligation independent cloning (LIC) method (Figure 5). The principle of LIC cloning is that linear DNA fragments can be assembled into a defined chimeric plasmid by making DNA ends single stranded and complementary to one another. Complementary DNA ends are achieved by the addition of nucleotide tags to the 5' ends of the oligos used to PCR the targets. Complementary DNA ends are then made single stranded using an exonuclease. Annealing of the DNAs results in a nicked-chimeric plasmid that is then transformed directly into *E. coli*. Based on this method, *HsScc1* was amplified by PCR using primers carrying tags at their 5' end:

Forward tag:

5´TAC TTC CAA TCC AAT GCA xxxxxxxxxxxxxx3´
    Y   F   Q   S   N   A   target PCR oligo

Reverse tag:

5´TTATCCACTTCCAAT G TTA TTA xxxxxxxxxxxxxxxx3´
                         Stop   Stop   target PCR oligo

The PCR product was purified using ethanol precipitation. The 438-C vector was linearized with SspI restriction enzyme. PCR product and linearized vector were run on an agarose gel and extracted. Purified vector and PCR product were treated with T4 DNA polymerase in the presence of dGTP and dCTP, respectively. The protocol is described below:

10 μl gel purified vector or PCR product (50-150 ng), 2 μl dCTP or dGTP (25 mM stock), 2 μl T4 DNA pol 10x Buffer, 1 μl 100 mM DTT, 0.4 μl EMD/Novagen T4 DNA pol and 4.6 μl H2O.

The reactions were incubated in a thermocycler at 22 °C for 30 min followed by 75 °C for 20 minutes. To anneal, 2 μl LICed PCR and 2 μl LICed vector were incubated in 10 μl of total volume for 10 min at RT. The annealed DNAs were transformed into DH10B cells (NEB) using electroporation. Isolation of plasmids was done as described above for other plasmids.
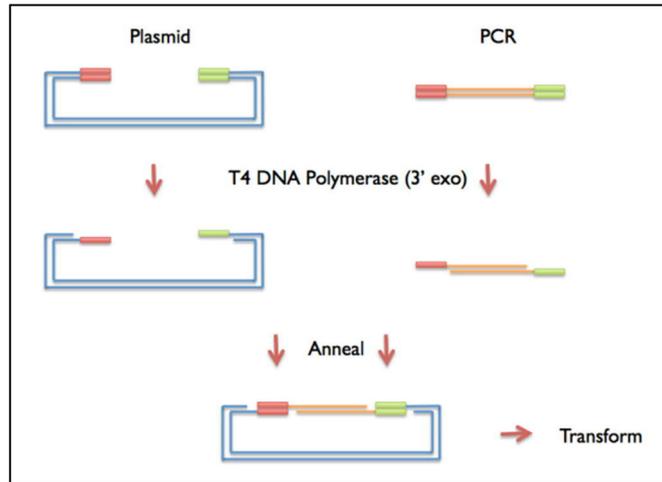
Figure 5: Sschematic illustration of the ligation independent cloning (LIC) method used for cloning. *HsSmc3-FLAG*, *HsSmc1-His* in pFastbac and combined *Smc1*, *Smc3-FLAG*, *Scc1* and *His-SA1* in a pFL multibac vector were provided by Jan-Michael Peters lab (Research Institute of Molecular Pathology, Vienna). Point mutations in *HsEsco1* were introduced with the QuikChange II XL site-directed mutagenesis kit (Agilent Technologies) according to the manufacturer's manual and were verified by DNA sequencing.

HsSmc3-FLAG and *HsSmc1-His* in pFastbac vector were provided by peters lab (Ladurner et al., 2014). To produce tetramer complex, combined *Smc1*, *Smc3-FLAG*, *Scc1* and *His-SA1* in a pFL multibac vector was also provided by peters lab (Ladurner et al., 2014).

2.2.1 Baculovirus generation

The recombinant plasmids were transformed into DH10BAC competent cells, which contain a modified viral bacmid Bmon14272 (Invitrogen) and a helper plasmid (Pmon7124), using electroporation. The bacmid has a yellow fluorescent protein (YFP) expression cassette to monitor protein expression and a mini-attTn7 target site and lacZα gene. The helper plasmid is tetracycline resistant and encodes a transposase for Tn7 transposition. The mini-Tn7 element on the recombinant plasmid can transpose to the mini-attTn7 target site on the bacmid in the presence of transposase provided by the helper plasmid. Colonies containing recombinant bacmids were identified by antibiotic selection (100 µg ml^{-1} ampicillin, 10 µg ml^{-1} tetracycline, 50 µg ml^{-1} kanamycin, 7 µg ml^{-1} gentamicin) and blue/white screening (X-gal and inducer Isopropyl β-D-1-thiogalactopyranoside [IPTG]). White clones were chosen as positive for transposition, since the transposition results in disruption of the LacZα gene and prevents colonies from turning blue in the presence of X-gal along IPTG. Positive clones were inoculated in LB containing the corresponding antibiotic and incubated at 37 °C overnight. Recom-

binant bacmid was isolated with a partially modified protocol using the miniprep kit (Qiagen). Briefly, cell pellets were resuspended in buffer P1 (Qiagen) followed by adding buffer P2 and N3 (Qiagen). The lysed cells were cleared by centrifugation (13000 rpm, 15 min). The supernatant containing DNA was processed for DNA extraction using isopropanol precipitation. The DNA was washed with 70% ethanol and resuspended in water.

To generate V_0 baculoviruses, about 1 million Sf9 cells were seeded in 1 well of a 6-well plate using Sf900 II SFM (Gibco) medium. DNAs (bacmids) were transfected to the adherent Sf9 cells using X-tremeGENE 9 DNA Transfection Reagent (Sigma-Aldrich). Transfection efficiency was monitored by YFP signal using a fluorescent microscope. V_0 viruses were harvested 72 h after transfection and applied to produce V_1 viruses using infection of 50 ml of Sf9 suspension culture. V_1 viruses were harvested 48 h after proliferation arrest and stored at 4 °C in dark.

2.2.2 Protein expression in insect cells

For protein expression, 600 ml (1×10^6 cells ml^{-1}) of Sf9 cells (grown in Sf900 II SFM [Gibco] medium) or Hi5 cells (grown in ESF921 medium [Expression Technologies]) were infected with 500-1000 μl of V_1 virus. To assess the protein expression level in cells, YFP expression was measured. Briefly, 500 μl of cells were collected at different time points and the intensity of YFP signals measured using a Victor X3 multi-label plate reader (PerkinElmer). After proliferation arrest, YFP signal was increased and then dropped (48-72 h after proliferation arrest). Cells were harvested by centrifugation (4000 rpm, 4 °C, 30 min), washed with 1x PBS, snap frozen, and stored at -80 ºC.

The expression levels and stability of different proteins were evaluated in Sf9 and Hi5 cells using SDS page and Coomassie blue staining. Based on the results, Hi5 cells were chosen for expression of truncated Esco2 and Sf9 for expression of other proteins in this study.

2.3 Protein purification

2.3.1 Purification of mouse Esco2$^{368-592}$

MmEsco2$^{368-592}$ recombinant protein was purified using anion exchange chromatography followed by affinity chromatography and size-exclusion chromatography. Briefly, the cell pellet was thawed from -80 °C and resuspended in lysis buffer (50 mM HEPES pH 7.2, 200 mM NaCl, 10% glycerol, 2 mM DTT and complete EDTA-free protease inhibitors cocktail [Roche]) and lysed 3 times using a Microfluidizer (Micro fluidics). The lysate was cleared by centrifugation (10000 rcf, 4 °C, 30 min) and the supernatant was applied onto a 50 ml anion-exchange Q-Sepharose column (GE Healthcare) equilibrated with lysis buffer. The column was washed with 10 column volumes of lysis buffer. The bound proteins were eluted with a linear gradient from 100 mM to 1 M NaCl. Different fractions were analyzed by SDS-PAGE and Coomassie blue staining. The peak fractions containing Esco2 were pooled and applied onto a 1 ml nickel- nitrilotriacetic acid (Ni-NTA) Superflow column (Qiagen), equilibrated with 20 mM HEPES pH 7.2, 500 mM NaCl, 10% glycerol, 2 mM DTT and 10 mM imidazole. The column was washed with 10 column volumes of buffer containing 10 mM imidazole. The nickel bound proteins were eluted with a linear imidazole gradient of 15-250 mM. The peak fractions were analyzed by SDS-PAGE and Coomassie blue staining. The corresponding fractions were pooled, concentrated to the volume of 2 ml using centrifugal filters (10-kDa concentrator; Amicon Ultra, Millipore) and applied onto a S75 16/600 pg size exclusion column (GE Healthcare), equilibrated with 10 mM HEPES pH 7.2, 150 mM NaCl, 5% glycerol and 2 mM DTT. Peak fractions were analyzed by SDS-PAGE and Coomassie blue staining, concentrated and flash frozen, and stored at −80 °C. Single point Esco2 mutants were purified using the same strategy described above. A list of buffer solutions applied for this purification is shown in Table 2.

Table 2: List of buffer solutions applied in MmEsco2$^{368\text{-}592}$ purification.

A		B		C
Lysis buffer	**Elution buffer**	**Lysis buffer**	**Elution buffer**	**buffer**
50mM Hepes 7.2	50mM Hepes 7.2	20mM Hepes 7.2	20mM Hepes 7.2	10mM Hepes 7.2
200mM NaCl	1 M NaCl	500mM NaCl	1 M NaCl	150mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	5% glycerol
2mM DTT	2mM DTT	10mM imidazol	250mM imidazol	2mM DTT
		2mM DTT	2mM DTT	

(A) Anion Exchange Chromatography buffers. **(B)** Affinity chromatography buffers. **(C)** Gel filtration buffer.

2.3.2 Purification of full-length human ESCO1

500 µl of Ni-NTA agarose beads (Qiagen) were washed 3 times and equilibrated with lysis buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 30 mM imidazole, 1 mM TCEP and complete EDTA-free protease inhibitors cocktail [Roche]) and used for 5 ml of cell pellet.

Frozen cell pellets were thawed and resuspended in lysis buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 30 mM imidazole, 1 mM TCEP and complete EDTA-free protease inhibitors cocktail [Roche]) and lysed by sonication (60 s "on" 30 % amplitude and 60 s "off", repeated 3 times) (Branson Ultrasonics™ Sonifier Modell 250 CE). The lysate was cleared by centrifugation (7500 rcf, 4 °C, 30 min [after 15 min spinning was stopped, the supernatant transferred to a fresh tube for an extra 15 min of centrifugation]). Subsequently, the supernatant containing HsESCO1 was filtered using 0.8 µm filters (Millipore) and incubated with equilibrated Ni-NTA beads for 2 h at 4 °C. Ni-NTA beads were washed with 10 bead volumes (BV) of lysis buffer, followed by 10 BV of high salt buffer (20 mM HEPES pH 7.5, 1 M NaCl, 10% glycerol, 30 mM imidazole and 1 mM TCEP), lysis buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 30 mM imidazole and 1 mM TCEP) and finally low salt buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 30 mM imidazole and 1 mM TCEP). Bound proteins to Ni-NTA beads were eluted with lysis buffer containing 150 mM NaCl, 500 mM imidazole and dialyzed for 16 h against dialysis buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 10% glycerol, 1 mM DTT). The dialyzed samples were aliquoted and snap-frozen. HsESCO1 mutants were expressed and purified as described for wild type HsESCO1.

2.3.3   Purification of trimeric and tetrameric cohesin complex

Trimeric cohesin complexes were expressed in Sf9 cells using coinfection with *Smc1-His* and *Smc3-FLAG* and *Scc1-MBP* viruses. Cells were lysed in lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 1 mM DTT and complete EDTA-free protease inhibitors cocktail [Roche]) supplemented with 0.02% NP40 and 1 mM PMSF. After sonication and clarification (centrifugation at 8500 rcf, 4 °C, 60 min), the lysate was applied onto a 5 ml amylose column (GE Healthcare) and equilibrated with lysis buffer. The bound proteins were eluted with a linear gradient of 10-100 mM maltose. The peak fractions were pooled, concentrated and applied onto a S200 16/600 pg size exclusion column (GE Healthcare) and equilibrated with 10 mM HEPES pH 7.2, 150 mM KCl, 5% glycerol and 2 mM DTT. Peak fractions were concentrated, flash frozen and stored at −80 °C. The Tetrameric cohesin complexes were expressed in Hi5 cells using coinfection with Smc1, Smc3-FLAG, Scc1, His-SA1 viruses. Cells were lysed in lysis buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 2 mM DTT, 30 mM imidazole and complete EDTA-free protease inhibitors cocktail [Roche]) supplemented with 1 mM TCEP, 1 mM Pefabloc and 0.05% Tween-20. After sonication (60 s "on" 30% amplitude and 60 s "off", repeated 3 times) and clarification (centrifugation at 8500 rcf, 4 °C, 60 min), the supernatant containing the tetramer was filtered using 0.8 μm filters (Millipore). Subsequently, the lysate was incubated with 1 ml of Ni-NTA beads for 2 h at 4 °C. Ni-NTA beads were washed with 10 bead volumes (BV) of lysis buffer, followed by 10 BV of high salt buffer (50 mM HEPES pH 7.5, 1 M NaCl, 10% glycerol, 30 mM imidazole and 0.01% Tween-20), lysis buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 30 mM imidazole and 0.01% Tween-20) and finally low salt buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 30 mM imidazole and 0.01% Tween-20). Bound proteins to Ni-NTA beads were eluted with lysis buffer containing 150 mM NaCl and 250 mM imidazole. Eluates were incubated with 200 μl of anti-FLAGM2 agarose beads (Sigma) for 2 h at 4 °C. Beads were washed with washing buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT and 0.01% Tween-20). The complex was eluted in elution buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT and 0.5 mg ml^{-1} FLAG peptide).

2.4 Characterization of purified proteins

2.4.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated by SDS-PAGE according to their size using 4-15% Mini-PROTEAN® TGX™ precast protein gels and the Mini-PROTEAN® tetra cell systems chambers (Biorad). Gel electrophoresis was carried out with 1x Tris/glycine/SDS running buffer at 250 volts for 30 to 45 min. Protein subunit size was estimated using a prestained protein marker (ThermoFisher). 0.8 and 3 µl of this marker were used for Coomassie and silver staining, respectively.

2.4.2 Protein staining

Protein gels were either stained using Coomassie or silver staining. To stain with Coomassie, staining solution (0.1% Coomassie brilliant blue R-250, 40% methanol and 10% glacial acetic acid) was added to the gel and put into a microwave at 900 watts for 40 seconds to 1 minute (until the Coomassie stain solution boiled). The gel was then incubated in the Coomassie solution for 10 minutes to overnight on a rocking table. Fresh destaining solution (40% methanol and 10% glacial acetic acid) was added to the gel and put into the microwave at 900 watts for 40 seconds to 1 minute (until the destaining solution boiled). The gel was incubated for 10 minutes in the destaining solution on a rocking table. Destaining was repeated until a sufficient level of destaining was reached.

To stain with silver, the gel was incubated for 60 minutes to overnight in the fixing solution (50% methanol and 10% glacial acetic acid). After fixing, silver staining was done according to the manufacturer's manual (ProteoSilver silver stain kit, Sigma).

2.4.3 Protein concentration measurement

The concentrations of protein solutions were determined by NanoDrop 2000 (Thermo Fisher Scientific) using the protein specific extinction coefficient and molecular weight. A dilution series of bovine serum albumin (BSA) (67 kDa, NEB) was used as standard to estimate the concentration of low quantities of purified proteins.

2.5   Crystallization and structure determination

MmEsco2[368-592] was crystallized using the sitting-drop vapor-diffusion method at 20 °C, using a Cartesian liquid dispensing robotic device. Different commercial crystallization reagents were screened to identify initial conditions. Crystals were obtained from droplets consisting of 100 nl of MmEsco2[368-592] (in buffer with 10 mM HEPES pH 7.2, 150 mM NaCl, 5% glycerol and 2 mM DTT) and 100 nl reservoir solution. The crystal growth was monitored by imaging with RockImager. After harvesting, crystals were cryoprotected in 15–20% ethylene glycol and flash frozen in liquid nitrogen.

Diffraction data were collected at beamline PXII of SLS (Paul Scherrer Institute, Villigen, Switzerland), processed and scaled using XDS (Kabsch, 2010). By making use of the natively bound zinc ion, the crystal structure was determined by single-wavelength anomalous dispersion (SAD) from a dataset collected at the zinc peak wavelength.

The final model was built manually using COOT32 (Emsley et al., 2010) and structure refinement was performed with Phenix33 (Adams et al., 2010).

2.6   *In vitro* acetylation assay

Acetylation assays were performed by preincubation of 100-500 nM of trimer (dependent on the experiment) or 100 nM of tetramer with 240 µM ATP, 10 µM AcCoA, 3.3 nM pcDNA3.1 plasmid, 25 mM HEPES pH 7.5, 25 mM NaCl, 1 mM $MgCl_2$ and 0.05 mg ml^{-1} BSA at 32 °C. After 1 h, 50 nM HsESCO1 and additional NaCl to the final concentration of 100 mM were added for further incubation at 37 °C. The reactions were stopped at the different time points by adding an equal volume of 2X SDS loading buffer, and denatured at 100 °C for 5 min. SMC3 acetylation was detected using immunoblotting and AcSmc3-specific antibody. Quantification was performed using ImageQuant software (GE Healthcare).

In vitro acetylation assays using MmEsco2[368-592] and different lengths of the Smc3 peptides (14 residues; RVIGAKKDQYFLDK, 24 residues: EEVSLRRVIGAKKDQYFLDKKMTK and 29 residues: DNRLPIDKEEVSLRRVIGAKKDQYFLDKKMTKNDVMNLLES) were performed with incubation of 5µM of enzyme and 50 µM of peptide in buffer (50 mM Tris-

HCl, pH 7.9, 0.1 mM EDTA, 1 mM DTT, 50 µg/ml acetylated BSA and 20 µM [^{14}C] AcCoA) for 2 h at 30°C.

2.7 Cell culture, transfection and synchronization

Wild type *MmEsco1* was cloned into a pEF6/Myc-His B vector using standard restriction-ligation methods and KpnI/EcoRV restriction enzymes (NEB). Point mutations in *MmEsco1-myc/His* were introduced with the QuikChange II XL site-directed mutagenesis kit (Agilent Technologies). Immortalized MEFs$^{Esco1-/-}$ in standard medium (DMEM, supplemented with 10% fetal bovine serum [FBS], 100 U/ml penicillin and 100 µg/ml streptomycin [all Thermo Fisher Scientific]) were transiently transfected with the wild type and mutant versions of *MmEsco1* constructs using Lipofectamine LTX Reagent (Thermo Fisher) according to the manufacturer's manual with minor changes. In summary, 0.7×10^5 cells were seeded in one well of a 6-well plate. After 24 h, cells were transfected with 3 µg of plasmid DNA. To synchronize cells in G1, 36 h after transfection, the medium was changed to DMEM medium supplemented with 10% FBS and 25 µM lovastatin. Cells were harvested after 24 h (G1) and synchronization was assessed by flow cytometry. Subsequently, the transfection efficiency of the cells was monitored using western blotting with Esco1-specific antibody that can detect endogenous levels of the protein. Cells expressing MmEsco1 close to endogenous levels of MmEsco1 were taken for further analysis.

Wild type *MmEsco2*-myc/his and *H2B-mCherry* were cloned into the pVITRO2-hygro-mcs vector in two steps. First, full-length *MmEsco2* was cloned into the pcDNA3.1/myc-His vector. Subsequently, *MmEsco2-myc/his* and *H2B-mCherry* were amplified from the vectors pcDNA3.1/myc-His and pcDNA3 (Addgene), respectively, and cloned into the pVITRO2-hygro-mcs vector. Point mutations in *MmEsco2-myc/his* were introduced with the Quik-Change II XL site-directed mutagenesis kit (Agilent Technologies). Primary MEFs$^{Esco2fl/fl}$ were isolated from E12.5 embryos following standard procedures and immortalized by serial passages. To delete exon 2 and 3 of *Esco2*, immortalized MEFs$^{Esco2fl/fl}$ were grown to confluence in standard medium (DMEM, supplemented with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin [all Thermo Fisher Scientific]) at 37 °C in a humidified incubator with 5% CO_2. After reaching confluency, cells were transduced with Ad-Cre-GFP adenoviruses

(SignaGen) in low serum medium containing 3% FBS (Figure 6). After two days, the medium was changed to fresh low-serum medium and cells were cultured for another 48 h. Immortalized MEFs$^{Esco2-/-}$ were stably transfected with wild type and mutant versions of *MmEsco2* using Hygromycin selection. Clones that stably expressed the mutants close to the endogenous level were selected using western blotting and MmEsco2-specific antibody sensitive enough to detect endogenous protein levels. For synchronization, cells were treated twice with 2 mM thymidine for 14 h with an intermittent release of 9 h. Cells were harvested 2 h after the second thymidine release and further processed for subsequent analyses.
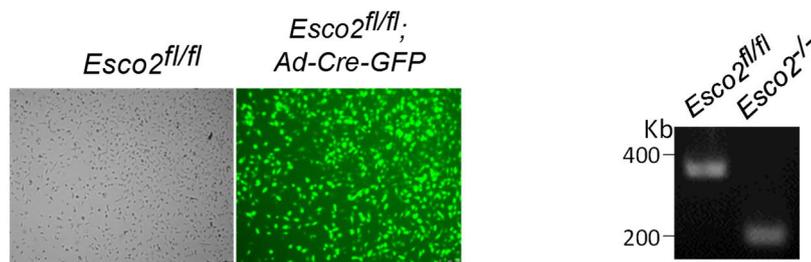
Figure 6. Generation of MEFs$^{Esco2-/-}$ from MEFs$^{Esco2\,fl/fl}$.
MEFs$^{Esco2fl/fl}$ were transduced with Ad-Cre-GFP to delete exon 2 and 3 of *Esco2*. Efficiency of transduction was assessed by monitoring GFP fluorescence using a fluorescent microscopy (left). Wild type and knock out cells were PCR-genotyped using published primers (Whelan et al., 2012) to confirm their genotype (right).

2.8   Flow cytometry

To confirm cell cycle synchronization , an aliquot of collected cells (5×10^5) was analyzed for DNA content by flow cytometry. Cells were fixed for at least 30 min in 70% ethanol at -20 °C. Subsequently, cells were washed with cold 1x phosphate-buffered saline (PBS) supplemented with 5% FBS (resuspension and centrifugation at 1000 g, 4 °C) and stained with 500 µl of FxCycle™ PI/RNase Staining Solution (Thermo Fisher Scientific) for 30 min at room temperature. Samples were analyzed using the Accuri C6 Flow Cytometer (BD).

2.9   Cell extract preparation and immunoblotting

For whole-cell extracts, cells were collected, washed in cold 1x PBS, resuspended in 2X SDS loading buffer and sonicated in an ultrasonic bath for 15 min. Chromatin fractionation was

performed according to the protocol described by Mendez and Stillman (2000) with minor modifications. Briefly, cells (1.2×10^6 cells) were washed twice with cold 1x PBS and lysed with 200 µl buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 1 mM DTT and complete EDTA-free protease inhibitors cocktail [Roche]). Subsequently, Triton X-100 (0.1%) was added, and the cells were incubated for 8 min on ice. Nuclei were collected in pellet 1 (P1) by low-speed centrifugation (5 min, 1300 g, 4 °C). The supernatant (S1) was further clarified by high-speed centrifugation (5 min, 20000 g, 4 °C) to remove cell debris. Pellet 1 (nuclei) was washed once in 200 µl buffer A, and then lysed in 100 µl buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT and complete EDTA-free protease inhibitors cocktail [Roche]). Cells in buffer B were incubated for 30 min at 4 °C on a rotating wheel. The insoluble chromatin was collected by centrifugation (5 min, 1700 g, 4 °C), washed once in 100µl buffer B, and centrifuged again under the same conditions. The final chromatin pellet (P3) was denatured in 2X SDS loading buffer and sonicated in an ultrasonic bath for 15 min. Cytoplasmic (S1), nucleoplasmic (S2) and chromatin fractions (P3) were analyzed by western blotting. Signal intensities were detected by ImageQuant LAS4010 imager (General Electric) and quantified using ImageQuant software.

2.10 Prometaphase chromosome spreads, cytospin and immunofluorescence analysis

MEFs were grown in T-75 flasks with standard culture medium (DMEM, supplemented with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin [all Thermo Fisher Scientific]). Cells with 60 % confluency were arrested using nocodazole (400 ng ml^{-1}) for 4 h. Mitotic cells were harvested by shaking off and incubated with 1 ml of 75 mM KCl for 20 min at 37 °C. Prometaphase chromosomes were fixed by adding 1 ml of the fixing solution (methanol: acetic acid [3:1]), washed 6 times with fixing solution and dropped onto humidified positively charged microscope slides (VWR). Prometaphase chromosome spreads were stained and mounted using Vectashield mounting medium with DAPI (Vector Laboratories) and visualized using a Leica fluorescence microscope. For immunofluorescence on prometaphase chromosomes, mitotic cells were spun onto poly-L-lysine coated coverslips at 1000 rpm for 10 min using a Cytospin-4 centrifuge (Shandon) after incubation with 75 mM KCl. Cells were

pre-extracted with 0.5% Triton X-100 in 1x PBS for 5 min at 4 °C and fixed in 4% paraform-aldehyde (PFA) in 1x PBS for 20 min at RT. Fixed cells were incubated further in blocking reagent (5% BSA, 0.1% Tween-20 in 1x PBS) followed by incubation with Aurora B antibody and a fluorescent secondary antibody. Nuclei were stained with DAPI.

2.11 Statistical analysis

Statistical analysis was conducted using GraphPad Prism software (version 5.04). All data are presented as means ± SEM.

3 Results

3.1 Expression, purification and crystallization of the MmEsco2 protein

To gain insights into the structure and catalytic mechanism of an enzyme, crystallization and X-ray structure determination are essential. This requires sufficient amounts of highly pure recombinant protein. Protein expression in *Escherichia coli* is a fast, cost-efficient, and well-established method. To produce soluble Esco2 protein, full-length N-terminal or C-terminal His-tagged mouse or human Esco2 (592 and 601 amino acids long, respectively) were expressed in *E. coli*. Protein Expression was analyzed by SDS-PAGE following Coomassie blue staining and Western blotting. The results showed that all these constructs were poorly expressed in *E. coli* (data not shown). This could be explained by the lack of specific chaperon systems or the absence of post-translational modifications required for protein stability (Trowitzsch et al., 2010). Likewise full-length proteins (Figure 7A, top) were poorly expressed in Sf9 and Hi5 insect cells (not shown) which prompted us to produce several fragments which included the acetyltransferase (ACT) domain and the N-terminally located C2H2 zinc finger (ZnF) (Figure 7A). These MmEsco2 fragments were expressed in Sf9 and Hi5 insect cells. A high yield was obtained with the C-terminal His-tagged MmEsco2$^{368-592}$ construct in Hi5 cells (Figure 7B). This fragment was also soluble (Figure 7C) and migrated in SDS-PAGE at the predicted molecular mass of approximately 28 kDa (Figures 7B and 7C). Hence large-scale of MmEsco2$^{368-592}$ expression was performed in 1 L Hi5 cells. Cell extracts were sequentially purified starting with anion exchange chromatography followed by a Ni-NTA affinity column and size-exclusion chromatography (Figures 8A - 8C). This resulted in a pure 28 kDa MmEsco2$^{368-592}$ recombinant protein (Figure 8C) as judged by SDS-PAGE stained with Coomassie blue and on a Western blot with a His-antibody (Figure 8C). Note that MmEsco2$^{368-592}$ is positively charged and does not bind to the anion exchange resin. This allows the removal of the majority of nucleic acid contaminants and leads to partial purification of MmEsco2$^{368-592}$ from crude extract (Figure 8B). Ni-affinity chromatography further separated MmEsco2$^{368-592}$ from contaminant proteins and MmEsco2$^{368-592}$ eluted sharply from the Ni-NTA column with increasing imidazole concentration (Figure 8B). Size exclusion chromatography on Superdex 75 yielded approximately 0.4 ml of a pure fraction of MmEsco2$^{368-592}$ at a concentration in a range of 10 mg/ml. This efficient 3-step purification and a high purity of MmEsco2$^{368-592}$ provided a good resource for subsequent crystallization experiment for MmEsco2$^{368-592}$.
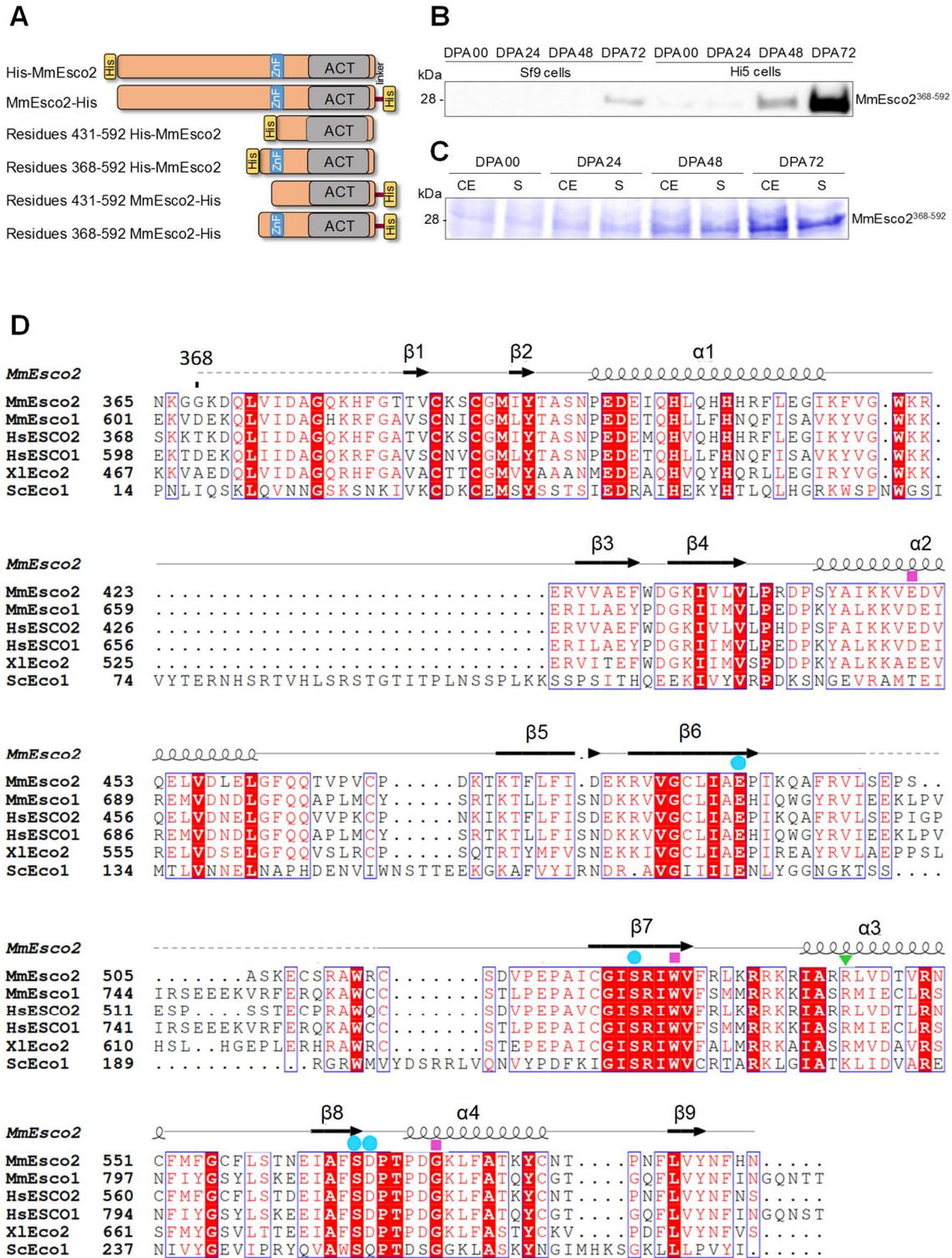
Figure 7: Expression of MmEsco2 in insect cells and sequence alignment of ACT domains from various species.
(A) Scheme of MmEsco2 indicating the zinc-finger (ZnF) and the acetyltransferase (ACT) domain. Full-length and different truncated variants of MmEsco2 with N- or C-terminal His-tag were designed for expression in insect cells. (B) Comparison between the expression of MmEsco2$^{368-592}$ in Sf9 and Hi5 insect cells at different time points after virus infection. Expression of protein was analyzed by Western blotting -

using His-antibody. **(C)** Monitoring expression and solubility of MmEsco2$^{368-592}$ in Hi5 insect cells. Expression of protein was analyzed by SDS-PAGE followed by Coomassie blue staining. CE: cell extraction; S: soluble fraction; DPA: days post anthesis. **(D)** Sequence alignment of Esco2 orthologs. Sequences shown are *Homo sapiens* ESCO1 (Hs), HsESCO2, *Mus musculus* Esco1(Mm), MmEsco2, *Xenopus laevis* Eco2 (Xl) and *S. cerevisiae* Eco1 (Sc). Strictly conserved residues are highlighted in red, and highly conserved residues are boxed. Numbering and secondary structural elements above the sequence alignment are shown for MmEsco2$^{368-592}$. Dashed lines show the disordered regions. Blue circles highlight residues presumed to be important for catalysis. Pink squares highlight residues mutated in Roberts syndrome. Sequence alignments were illustrated using ESPript (Gouet et al., 1999).
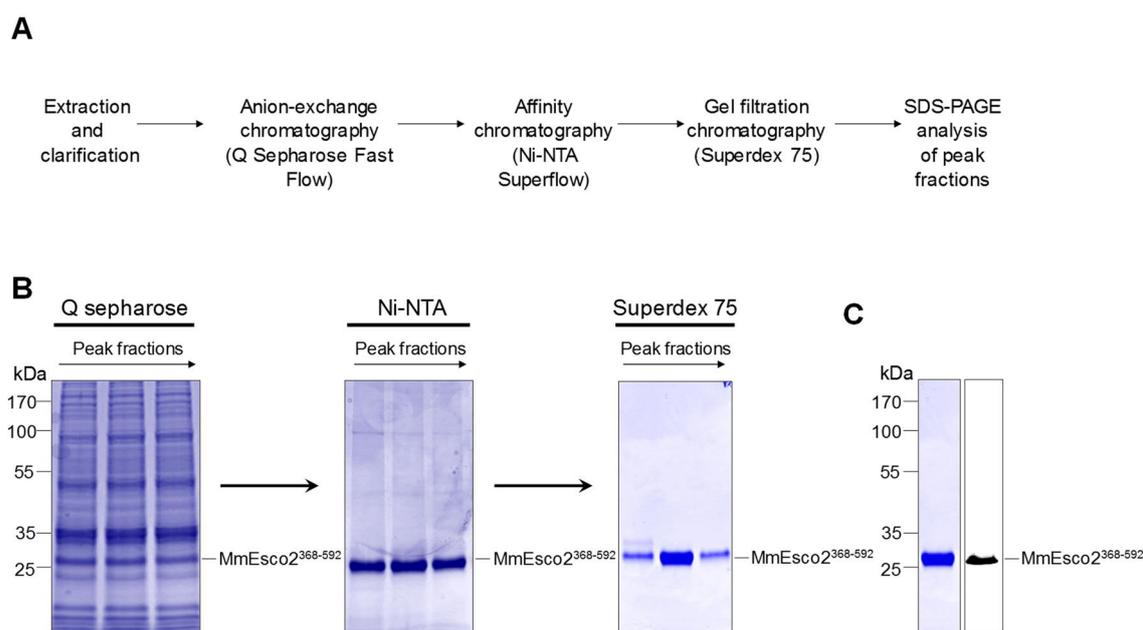
Figure 8: Purification of recombinant MmEsco2$^{368-592}$.
(A) Scheme of the MmEsco2$^{368-592}$ purification steps. **(B)** Purification of MmEsco2$^{368-592}$. The protein was purified in three subsequent steps using anion-exchange (Q-Sepharose), affinity (Ni-NTA) and size exclusion (Superdex 75) chromatography. The peak fractions after each step were analyzed by SDS-PAGE followed by Coomassie blue staining. The band corresponding to MmEsco2$^{368-592}$ is indicated on the right side. The masses of the molecular weight markers (kDa) are indicated on the left side. **(C)** The peak fractions from size exclusion (Superdex 75) chromatography were analyzed by SDS-PAGE followed by Coomassie blue staining (left) or Western blotting using anti-His-antibody (right).

Next, initial crystallization conditions were determined. Several commercially available kits were used, each of which consists of 96 different crystallization conditions (variables are salt type, precipitant, additive, pH etc.). Kits used were Classics™, Classics Lite™, PEGs™, NH$_4$SO$_4$, Anions™, Cations™, ComPAS™, pH Clear I™, pH Clear II™, Index screen™ and Wizard 1™ and Wizard 2™. Crystallization experiments for MmEsco2$^{368-592}$ were performed at 20 °C with the sitting-drop vapor-diffusion method using a Cartesian liquid dispensing ro-

botic device. Crystals appeared after 3 days from droplets in pH Clear I screen plates. 100 nl of 10 mg/ml MmEsco2$^{368\text{-}592}$ in buffer (10 mM HEPES pH 7.2, 150 mM NaCl, 5% glycerol and 2 mM DTT) was mixed with an equal volume of pH Clear I solution in the reservoir. Crystals grew optimally in the presence of 100 mM Tris, 20% (v/v) 2-Methyl-2,4-pentanediol (MPD) at pH 8. They had a conical appearance and a length of up to 580 µm (Figure 9). Grown crystals were harvested and cryoprotected in 15–20% ethylene glycol and flash frozen in liquid nitrogen. Diffraction data were collected at beamline PXII. Crystals diffracted up to 1.8 Å resolution. Initial molecular replacement attempts using published Gcn5-related N-acetyltransferases (GNAT) protein structures (PDB ID code 1QST) as model were unsuccessful to provide electron density maps suitable for model building. By making use of the natively bound zinc ion, the crystal structure was determined by single-wavelength anomalous dispersion (SAD) from a dataset collected at the zinc peak wavelength. Crystal structure data are summarized in a table (Table 3).
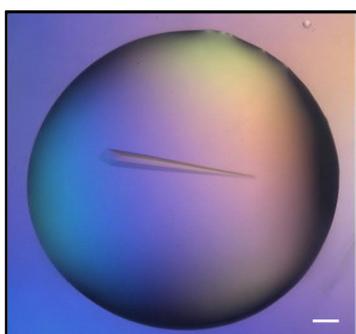
Figure 9: MmEsco2$^{368\text{-}592}$ crystal.
Image of a typical MmEsco2$^{368\text{-}592}$ crystal obtained in a pH Clear I screen plate, consisting of equal volumes of MmEsco2 (10 mg/ml) and reservoir solution (100 mM Tris, 20% (v/v) MPD, pH 8). Scale bar: 100 µm.

Table 3: Data collection, phasing and refinement statistics.

	Native I - Zn K-Edge	Native II
Data collection		
Space group	P4$_3$	P4$_3$
Cell dimensions		
a, b, c (Å)	52.6, 52.6, 106.1	52.7, 52.7, 107.5
α, β, γ (°)	90, 90, 90	90, 90, 90
Wavelength	1.28	1.0
Resolution (Å)	50.0-2.3 (2.3)	50-1.8 (1.8)
R_{sym}	2.1 (50.8)	4.7 (29.9)
I / σI	90.8 (4.1)	31.7 (4.8)
Completeness (%)	98.7 (99.1)	99.2 (98.0)
Redundancy	7.0 (7.1)	3.4 (3.5)
CC(1/2)	100 (93.6)	99.5 (95.7)
Refinement		
Resolution (Å)		47.4-1.8
No. reflections		26908
R_{work} / R_{free}		18.5/21.1
No. atoms		
Protein		1581
Ligand		49
Water		21
B-factors		
Protein		36.4
Ligand		30.1
Water		34.5
R.m.s deviations		
Bond lengths (Å)		0.019
Bond angles (°)		1.9
Ramachandran		
favored (%)		98.5
allowed (%)		1.0
outliers (%)		0.5

One crystal was used for each dataset. Values in parentheses are for highest-resolution shell.

3.2 Overall structure of the MmEsco2$^{368-592}$ in complex with Coenzyme A (CoA)

The refined MmEsco2$^{368-592}$ /CoA structure revealed continuous electron density from amino acids 368 to 592, except for two short unresolved regions (residues 368-383 and 501-514), likely because these regions are structurally disordered (Figure 7D). Residues 423-592 of MmEsco2$^{368-592}$ adopt an overall fold that shares many common features with other acetyl-transferases of the GNAT family (Salah Ud-Din et al., 2016) (Figure 10A). This fold includes a structurally conserved core region, consisting of β5, β6, β7, α3, and α4, which is flanked by structurally variable regions, likely responsible for substrate binding (Salah Ud-Din et al., 2016) (Figure 10A). The CoASH cofactor is natively present in a complex with MmEsco2 in a groove formed by β7 and β8 strands and α3 and α4 helices (Figures 10A and 10C), similar to the position of the CoA/Acetyl-CoA cofactor bound to other GNAT family members (Salah Ud-Din et al., 2016) (Figure 3). The zinc finger domain (residues 385-416), located N-terminally of the catalytic domain, consists of two β-strands and one α-helix and folds around the zinc ion, which is coordinately bound to the protein through four residues (C386, C389, H404 and H407) (Figure 10B).
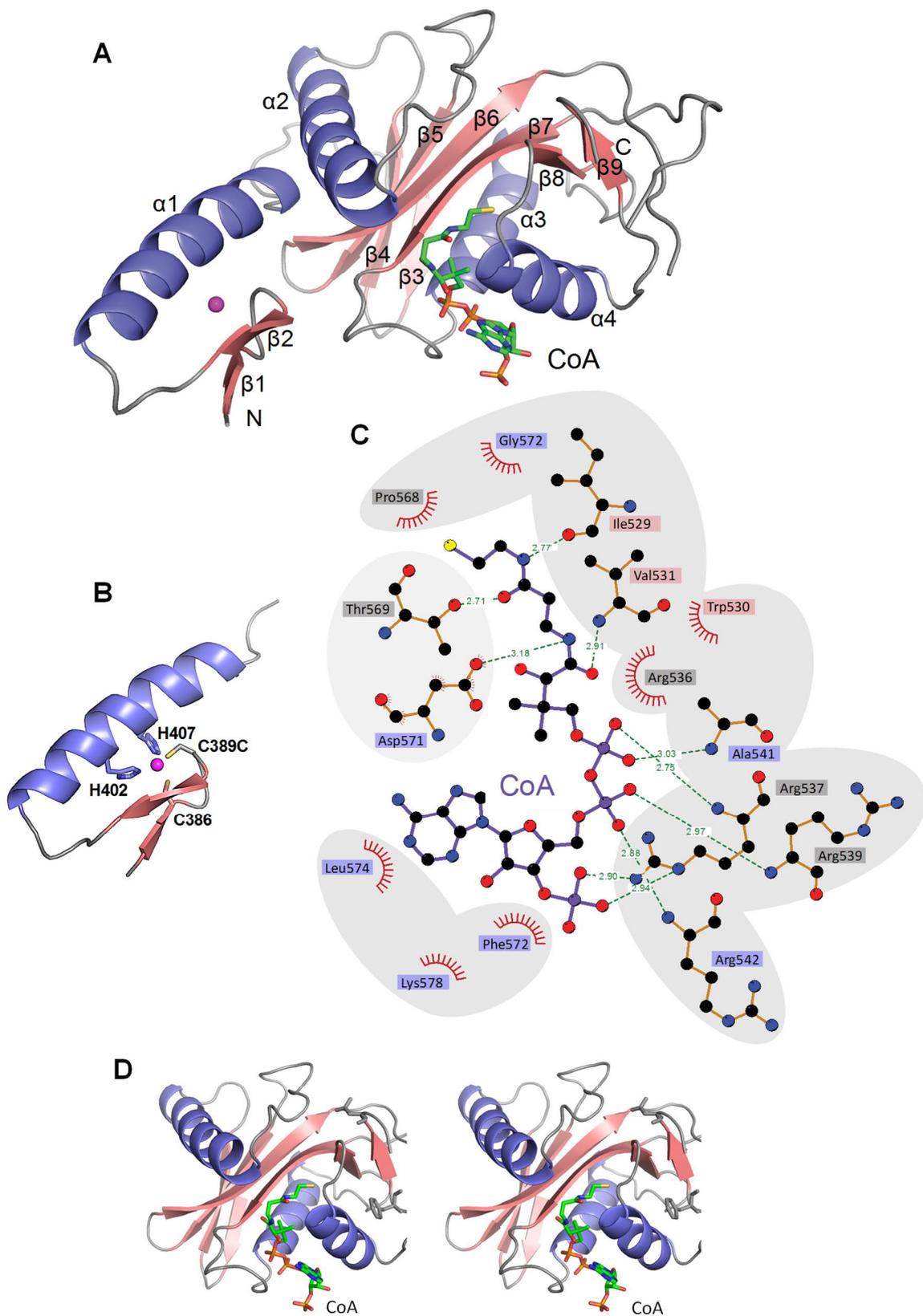
Figure 10: Structure of the MmEsco2[368-592]/CoA Complex.
(**A**) Ribbon representation of the MmEsco2[368-592]/CoA complex. α-Helices are shown in blue, β-strands in raspberry, and loop regions in grey. CoA is represented as sticks and colored according to elements: carbon, green,

nitrogen, blue; sulfur, yellow; oxygen, red and the zinc ion shown as a magenta sphere. **(B)** Zinc finger residues involved in zinc ion ligation are labeled. **(C)** CoA interactions with its binding residues in MmEsco2. Residues located on β-sheets are highlighted in raspberry, residues located on α-helices are highlighted in blue, residues located on loops are highlighted in grey. Green dashed lines indicate hydrogen bonds. The hydrophobic contacts are represented by arcs with spokes radiating towards the CoA. The diagram was generated using LigPlot (Wallace et al., 1995). **(D)** Stereo view of the MmEsco2368-592/CoA complex structure. The CoA cofactor is located in a groove formed by β7 and β8 strands and α3 and α4 helices.

3.3 Active site architecture and implications for catalysis

To begin to characterize the active site of MmEsco2, we chose a number of potential catalytic residues which could serve as general base/acid and proton shuttle (see Introduction section 1.2.3). This selection was based on the MmEsco2$^{368-592}$/CoA complex structure and took into consideration side chain orientations and proximities of the candidate residues relative to the acetyl group of AcCoA (Figures 11A and 11B). The most obvious candidate residues are serine 566 and aspartate 567, which are located at the C-terminus of the β8 strand and in a loop between β8 strand and α3 helix, respectively (Figures 11A). The β-hydroxyl group of S566 and γ-carboxyl group of D567 are ~ 6 and ~ 7.6 angstrom (Å) away from the carbonyl carbon of AcCoA, respectively (Figures 11A and 11B). In addition, we considered serine 527, which is in the middle of β7 (Figure 11A). The β-hydroxyl group of this serine is ~ 7.9 Å away from carbonyl carbon of AcCoA (Figures 11A and 11B). Finally, we considered E491, which resides at the end of the β6 strand (Figure 11A). The δ-carboxyl group of E491 is ~ 9.6 Å away from carbonyl carbon of AcCoA (Figures 11A and 11B). Also, note a water molecule located between these four residues and CoA, which might be involved in proton transfer from the ε-amino groups of substrate lysines to the catalytic residues (Figure 11C). Additional support for the engagement of these candidate residues in catalysis is coming from superposition of the structures MmEsco2/CoA with HsESCO1/AcCoA, xEco2/K105-CoA and xEco2/K106-CoA (Figures 12 and 13). Positioning of putative catalytic residues of MmEsco2 toward the AcCoA in HsESCO1/AcCoA and toward the K105 and K106 of the Smc3 peptide model substrate in xEco2/K105-CoA and xEco2/K106-CoA allows a more educated depiction of the candidates in the context of their model substrate. Further evidence for the importance of these four residues we chose for enzyme activity is coming from their strict and high conservation among Esco2 orthologs. E491, S527 and S566 residues are strictly conserved among Esco homologs. D567 residue is also highly conserved except in yeast Eco1 (Figure 7D).

Engagement of the D567 residue in catalysis has been also shown for its structurally equivalent residue (D157) in HsαTAT1 acetyltransferase (Szyk et al., 2014), a member of the GNAT family that shows a high structural similarity to MmEsco2 (Figure 14). However, superposition of MmEsco2/CoA and XEco2/K105-CoA suggests that D567 is also well positioned to play a role in substrate recognition and binding (Chao et al., 2017b) (Figure 12A).

In summary, our structure together with the comparison to other structures of related acetyltransferases identifies four putative catalysis-relevant residues in MmEsco2. Given the precedent of multiple side chains participating in acetyl transfer (See introduction section 1.2.3) we cannot *a priori* rule out a similar mechanism for MmEsco2.

To investigate the functional importance of these catalytic candidate residues, we performed site directed mutagenesis.
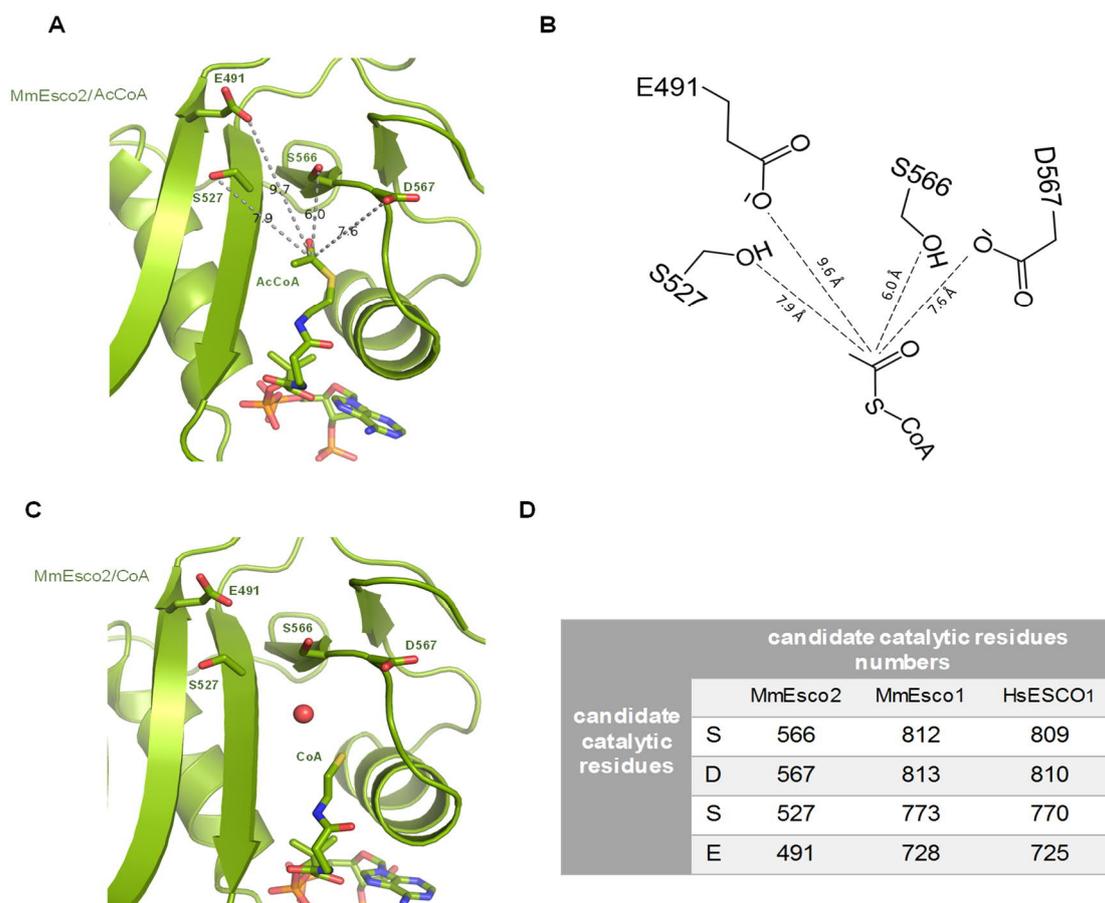
Figure 11: Active site of MmEsco2.
(A) Close-up view of the active site of the MmEsco2$^{368-592}$/AcCoA model (the acetyl moiety was modeled based on superimposition of MmEsco2$^{368-592}$/CoA and HsαTAT1/AcCoA [PDB ID code 4GS4]). Putative catalytic residues are labeled. Dashed lines indicate the distance in Ångström (Å) of putative catalytic residues to the acetyl group of AcCoA. **(B)** 2D schematic diagram showing the distances of S566, S527, D567 and E491 from AcCoA. **(C)** Close-up view of the active site of MmEsco2$^{368-592}$/CoA showing a water molecule located between the catalytic candidate residues and CoA. The water molecule is well positioned to allow proton transfer. **(D)** Numbering of equivalent putative catalytic residues of MmEsco2 in MmEsco1 and HsESCO1.
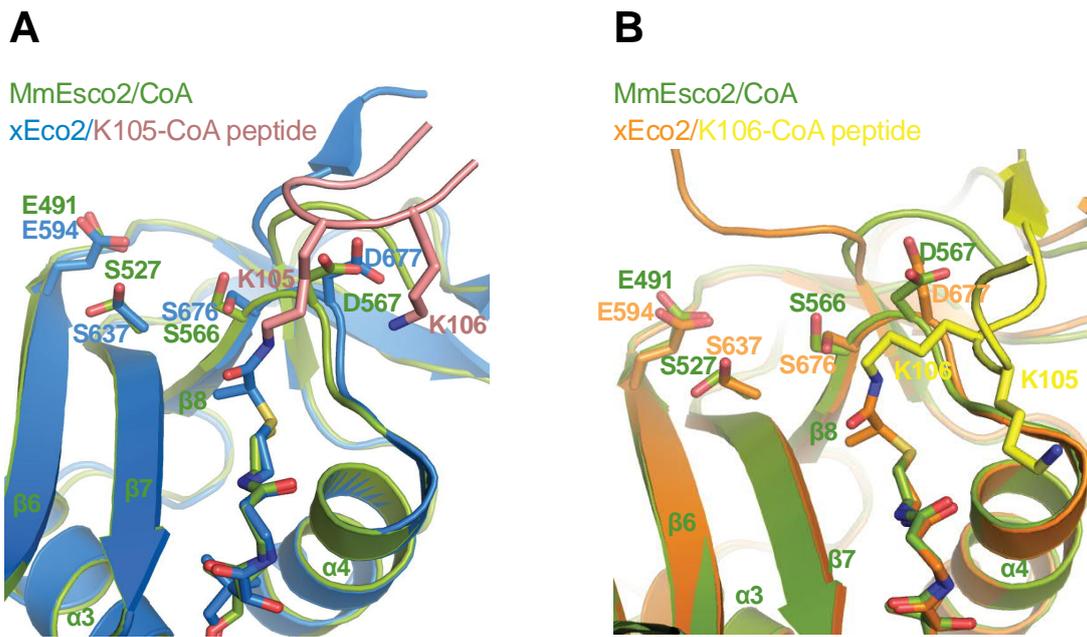
Figure 12: Close-up view and comparison of the active sites of MmEsco2/CoA, xEco2/K105-CoA and xEco2/K106-CoA.
(A) Putative catalytic residues of MmEsco2 (green) and xEco2/K105-CoA (blue; PDB ID code 5N1W) are shown as sticks. K105 in the Smc3 peptide model substrate is shown in raspberry. **(B)** Putative catalytic residues of MmEsco2 (green) and xEco2/K106-CoA (orange; PDB ID code 5N22) are shown as sticks. K106 of the Smc3 peptide is shown in yellow. Active site architecture and positioning of putative catalytic residues of MmEsco2 show high structural homology to xEco2.
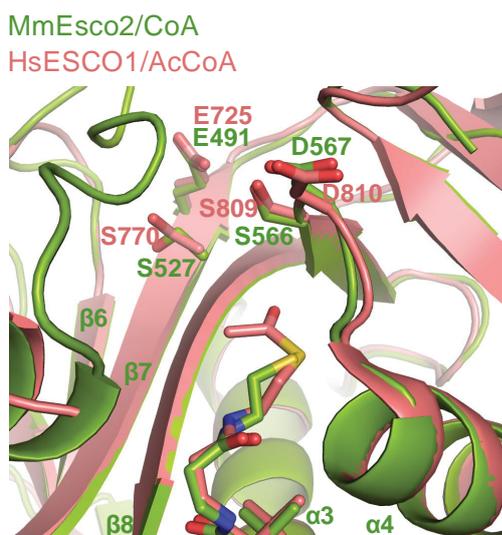
Figure 13: Close-up view of the active sites of MmEsco2/CoA and HsESCO1/AcCoA.
Putative catalytic residues of MmEsco2 (green) and HsESCO1 (raspberry; PDB ID code 4MXE) show high structural conservation.
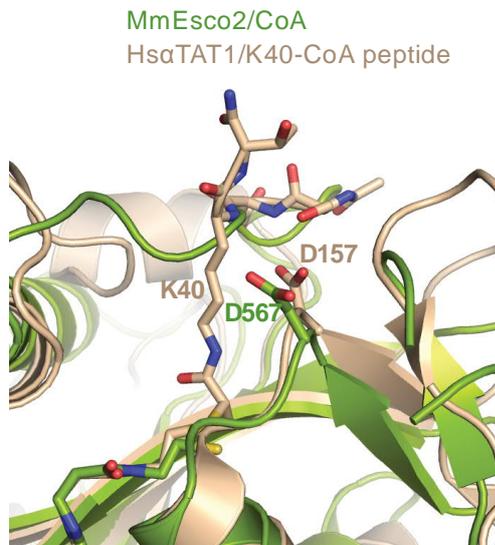
Figure 14: Close-up view of the active sites of MmEsco2/CoA and HsαTAT1/K40-CoA.
Close-up view of the active sites of MmEsco2/CoA (green) superimposed on structure of HsαTAT1/K40-CoA α-tubulin peptide (sand; PDB ID code 4GS4). D567 of MmEsco2 and D157 of HsαTAT1 show similar localization in the active site.

3.4 Effect of site-directed mutagenesis of the active site residues on acetyltransferase activity of Esco1 and Esco2

3.4.1 Expression and purification of HsESCO1 and HsESCO2

Based on the structure of murine Esco2 and on previously published GNAT acetyltransferase studies (Friedmann et al., 2012; Rivera-Colon et al., 2016) we focused on mutations in serine 566 and the neighboring residue aspartic acid 567 (Figure 11A). All proteins were synthesized in insect cells. While MmEsco2$^{368-592}$ was capable of acetylating itself at several lysine residues, this reaction was slow with a half-time in the range of hours (see Appendix, Figures 29B and 29C). The Smc3 peptide could not be acetylated by MmEsco2$^{368-592}$ although a human ESCO1 fragment (residues 590 – 840) was reported to acetylate this peptide albeit at a slow rate (Rivera-Colon et al., 2016). In addition, Ladurner et al. (2014) showed that full-length HsESCO1 is catalytically active and acetylates recombinant cohesin under *in vitro* conditions.

Full-length, N-terminally His-tagged versions of HsESCO1, MmEsco1, HsESCO2, MmEsco2, and of full-length His-MBP tagged HsESCO2 were expressed in Sf9 cells. They all gave very low yields with the exception of His-tagged HsESCO1 and MBP-tagged

HsESCO2. Purification of these two recombinant proteins from a 500 ml culture was done using affinity chromatography (Ni-NTA or amylose resin) (Figures 15A – 15C) and resulted in µg amounts of protein (HsESCO1: 20 µg and HsESCO2: 2 µg). SDS-PAGE of HsESCO1 showed a strong band with a molecular mass of about 120 kDa (molecular weight of HsESCO1 including His-tag and TEV recognition site was predicted as 100 kDa) which also reacted with an Esco1 polyclonal antibody on a Western blot (Figure 15C). MBP-tagged HsESCO2 was somewhat less pure and the yield was lower (Figure 15B, left lane). Therefore, all experiments reported here were carried out with HsESCO1 whose active site is very similar to that of MmEsco2 (Figure 13).

Next, we generated mutants of the HsESCO1 enzyme to evaluate their catalytic activity. Mutants S809A, S809C, D810A, and D810N (see Figure 11D for equivalent residues of MmEsco1 and MmEco2) were purified using the purification procedure described above for wild type HsESCO1 (Figures 15A and 15D).
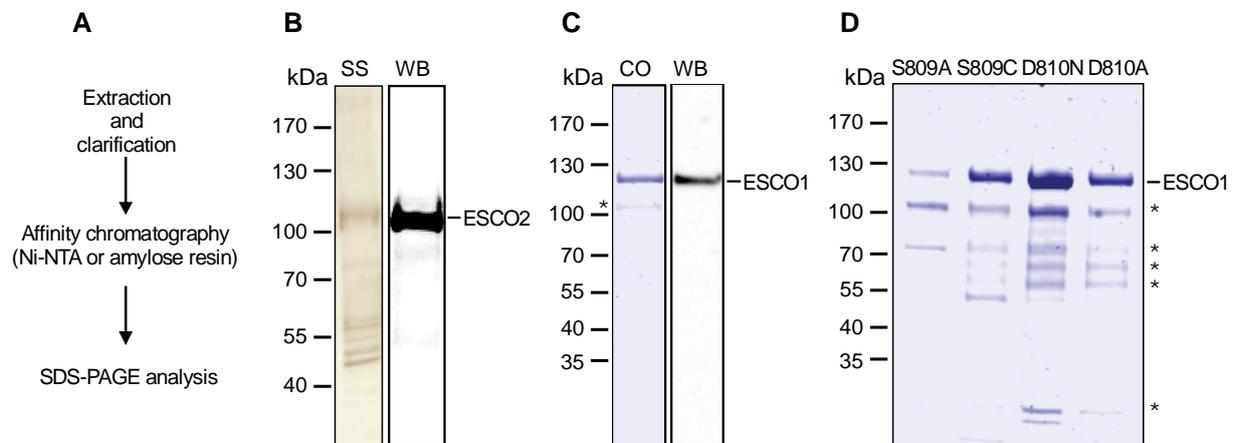
Figure 15: Purification of recombinant HsESCO1 and HsESCO2.
(A) Schematic presentation of the purification procedure for HsESCO1 and HsESCO2. **(B)** Purified HsESCO2 analyzed by SDS-PAGE followed by silver staining (SS) or Western blotting (WB) using MBP-specific antibody. The protein was detected as a protein band with a molecular mass of about 115 kDa (predicted: 110 kDa). **(C)** Purified HsESCO1 analyzed by SDS-PAGE followed by Coomassie blue staining (CO) or Western blotting (WB) using Esco1 antibody. The protein was detected as a protein band with a molecular mass of about 120 kDa. **(D)** Purified HsESCO1 mutants analyzed by SDS-PAGE followed by Coomassie blue staining. Asterisks indicate contaminants in Coomassie blue stainings in C and D.

3.4.2 Expression and purification of trimeric and tetrameric cohesin complexes

As noted above, autoacetylation of MmEsco2 and acetylation of the Smc3 peptide occurs at low rates and it might therefore problematic to use these substrates in assessing changes in activity evoked by site-directed mutagenesis of candidate catalytic residues. Moreover, it has been shown that the ATPase activity residing in the cohesin subunit SMC3 is essential for acetylation of this subunit by HsESCO1 (Ladurner et al., 2014). In addition, binding of DNA to cohesin stimulates the ATPase activity of Smc3 (Camdere et al., 2015). It has also been reported that the yeast ortholog of the subunit SA1/2 stimulates loading of cohesin onto DNA and increases DNA-dependent ATP hydrolysis by cohesin (Murayama and Uhlmann, 2014). Altogether, a more reliable assessment of Esco1-mediated catalysis is to use trimeric (SMC1, SMC3, SCC1) or tetrameric (SMC1, SMC3, SCC1, SA1 or SA2) cohesin in the presence of ATP and DNA.

To generate the physiological substrate of HsESCO1, trimeric and tetrameric human cohesin complexes were reconstituted and purified. The trimeric human cohesin complex was expressed in Sf9 cells using coinfection of SMC1-His, SMC3-FLAG and SCC1-MBP viruses. A two-step purification (amylose affinity chromatography and gel filtration) of the modified trimer resulted in a stoichiometric and highly concentrated complex at an amount 900 µg. All three subunits were detected using SDS-PAGE and Coomassie blue staining (Figures 16A – 16C).
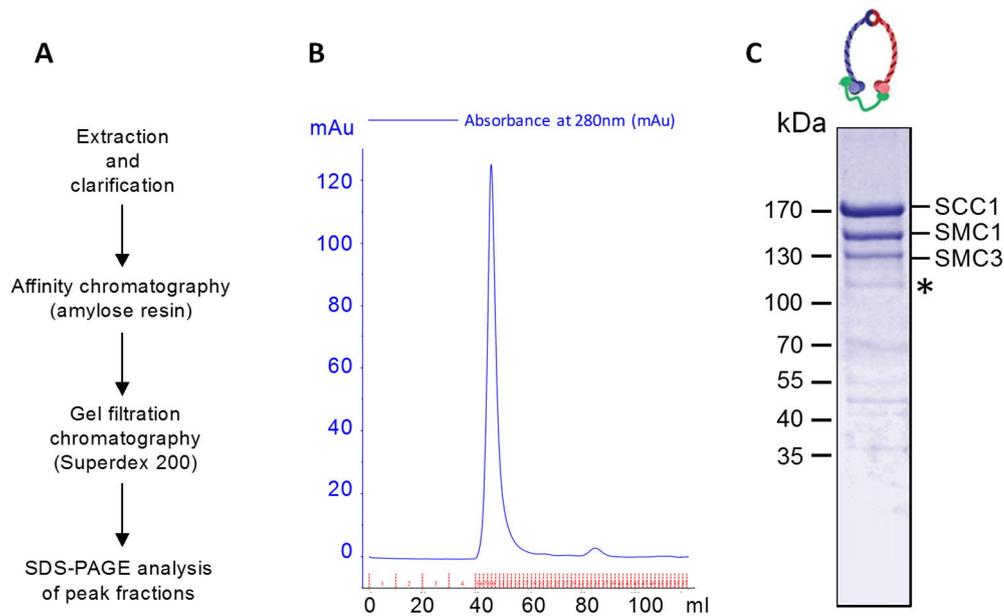
Figure 16: Purification of the trimeric human cohesin complex.
(A) Scheme of the purification procedure for trimeric human cohesin. **(B)** Trimer complex migrates as a sharp single peak purified on a HiLoad 16/600 Superdex 200 size exclusion column. mAu; milli absorbance units. **(C)** Purified human trimeric human cohesin analyzed by SDS-PAGE followed by Coomassie staining. The bands corresponding to SCC1, SMC1 and SMC3 are indicated on the right side. The masses of the molecular weight marker (kDa) are indicated on the left side. The asterisk indicates a minor contaminant. In the scheme of the subunit composition of trimeric cohesin SMC3, SMC1 and SCC1 are shown in blue, red and green respectively.

Tetrameric human cohesin complex was expressed in Hi5 cells using infection with a multi-bac virus harboring a SMC1, SMC3-FLAG, SCC1 and His-SA1 expression cassette. The complex was purified by two steps of affinity chromatography using Ni-NTA and anti-FLAG M2 agarose beads. All four subunits of tetramer complex were detected using SDS-PAGE followed by silver staining (Figures 17A and 17B). A final yield of about 20 µg was obtained.
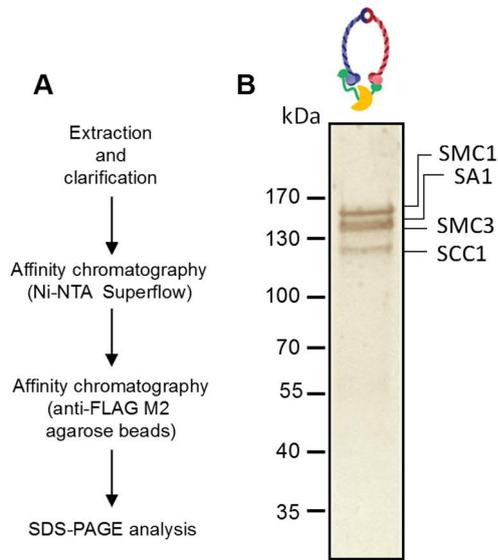
Figure 17: Purification of the tetrameric human cohesin complex.
(A) Scheme of the tetrameric human cohesin purification procedure. (B) Purified human tetrameric cohesin analyzed by SDS-PAGE followed by silver staining. The bands corresponding to SCC1, SMC1, SMC3 and SA1 are indicated on the right side. The masses of the molecular weight marker (kDa) are indicated on the left. In the scheme of the subunit composition of tetrameric cohesin SMC3, SMC1, SCC1 and SA1 are shown in blue, red, green and yellow respectively.

3.4.3   Establishment of *in vitro* cohesin acetylation by HsESCO1

Recombinant HsESCO1 was incubated with trimeric or tetrameric cohesin in the presence or absence of ATP and of topologically different DNAs for 1 hr. SMC3 acetylation was assessed by Western blotting using an AcSmc3 specific antibody (Nishiyama et al., 2010). These experiments revealed that in the presence of ATP and DNA, SMC3 acetylation in both trimeric and tetrameric cohesin is highly efficient (Figures 18A - 18C). ATP or DNA alone was not sufficient to promote acetylation of SMC3 neither in trimeric nor tetrameric cohesin (Figures 18A and 18B). However, addition of DNA and ATP together enabled HsESCO1 to catalyze SMC3 acetylation. Consistent with previous work (Ladurner et al., 2014), blocking ATPase activity with ATP analogs abolished acetylation.
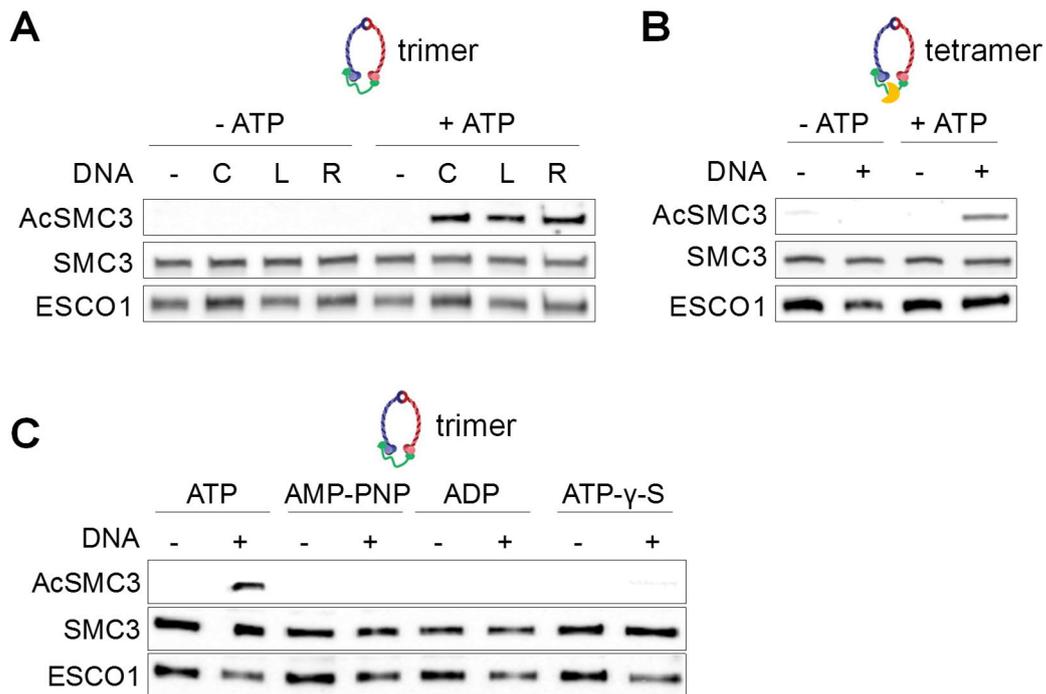
Figure 18: ATP and DNA stimulate acetylation of cohesin by HsESCO1.
(A) The purified trimeric cohesin complex was incubated with HsESCO1 and AcCoA in the presence or absence of ATP and different topologies of the pcDNA3 plasmid including linear DNA [L], relaxed circular DNA [R] or covalently closed circular DNA [C]. Different topologies of the DNA show a similar effect on SMC3 acetylation. (B) The purified tetrameric cohesin complex was incubated with HsESCO1 and AcCoA in the presence or absence of ATP and covalently closed circular DNA. (C) The purified trimeric cohesin complex was incubated with HsESCO1 and AcCoA in the presence of ATP, adenylyl-imidodiphosphate (AMP-PNP), Adenosine diphosphate (ADP) or Adenosine 5′-[γ-thio] triphosphate (ATP-γ-S). In the cohesin complex models, Smc1, Smc3, Scc1, and SA1 subunits are shown in red, blue, green, and yellow, respectively.

Next, we determined the time course of HsESCO1-mediated SMC3 acetylation for both trimeric and tetrameric cohesin. The reaction was carried out in the presence of covalently closed circular DNA, ATP and AcCoA and a 1:2 molar ratio of HsESCO1 to oligomeric cohesin. The results showed that the addition of the SA1 subunit strongly stimulates the rate of SMC3 acetylation (Figures 19A and 19B).

Together, our *in vitro* acetylation experiments show that a combination of tetrameric cohesin, DNA, and ATP provides a robust assay on HsESCO1 catalytic activity that can readily be extended to enzyme assays of site-directed HsESCO1 mutants.
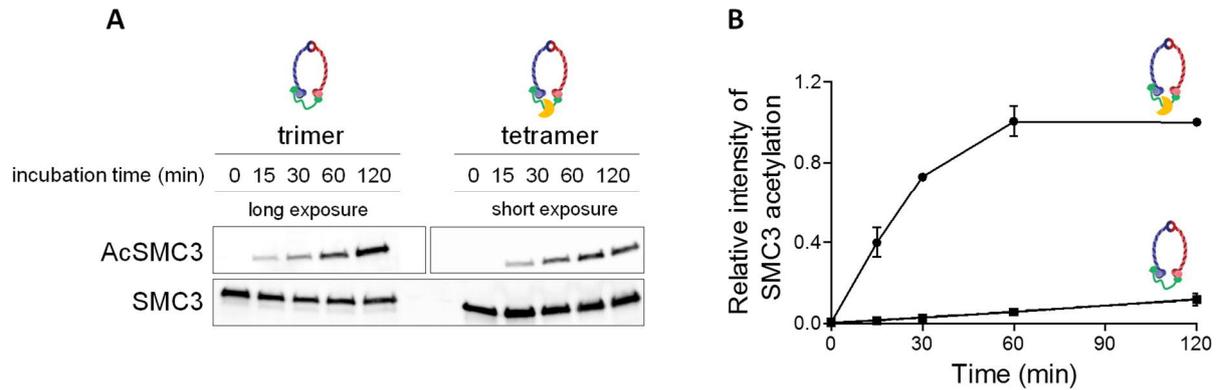
Figure 19: SA1 subunit boosts cohesin acetylation.
(A) and **(B)** Time course quantification of SMC3 acetylation after incubation of purified trimeric or tetrameric cohesin complexes with HsESCO1 in the presence of ATP, DNA and AcCoA. The level of SMC3 acetylation level was analyzed by quantitative Western blotting using an AcScm3-specific antibody. Data were normalized to the maximal signal seen for the tetramer. Data are shown as mean $\pm$ SEM (n = 2). Half-times for the tetramer is ~ 20 min however, acetylation of the trimer acetylation is considerably slower.

3.4.4 *In vitro* analysis of active site mutants of HsESCO1

We next quantified the rate of SMC3 acetylation by site-directed mutants of HsESCO1. We focused on the S809 and D810 residues and incubated HsESCO1[S809A], HsESCO1[S809C], HsESCO1[D810A] and HsESCO1[D810N] mutants (for residue equivalence see Figure 11D) with cohesin complexes. Consistent with the above findings, wild type HsESCO1 efficiently acetylated trimeric and tetrameric substrates. In mutants, acetylation of SMC3 could neither be detected with trimeric nor tetrameric cohesin substrates (Figures 20A - 20D). With the exception of the HsESCO1[S809A] mutant, HsESCO1 protein levels were similar (Figure 15D) to those obtained with wild type HsESCO1. These four mutants also had previously been shown to exhibit similar thermal stability to wild type protein arguing against the possibility that mutant proteins were unfolded (Rivera-Colon et al., 2016). Importantly, wild type HsESCO1 was able to acetylate SMC3 in the presence of HsESCO1 mutants such as HsESCO1[D810N] and HsESCO1[D810A] (Figure 20E). Hence, we can rule out an inhibitory effect of impurities present in mutant protein purifications. Furthermore, this experiment suggests that HsESCO1 mutants did not act in a dominant negative fashion.

Together, these results indicate that both S809 and D810 are indispensable for the catalytic activity of the HsESCO1 enzyme under *in vitro* condition.
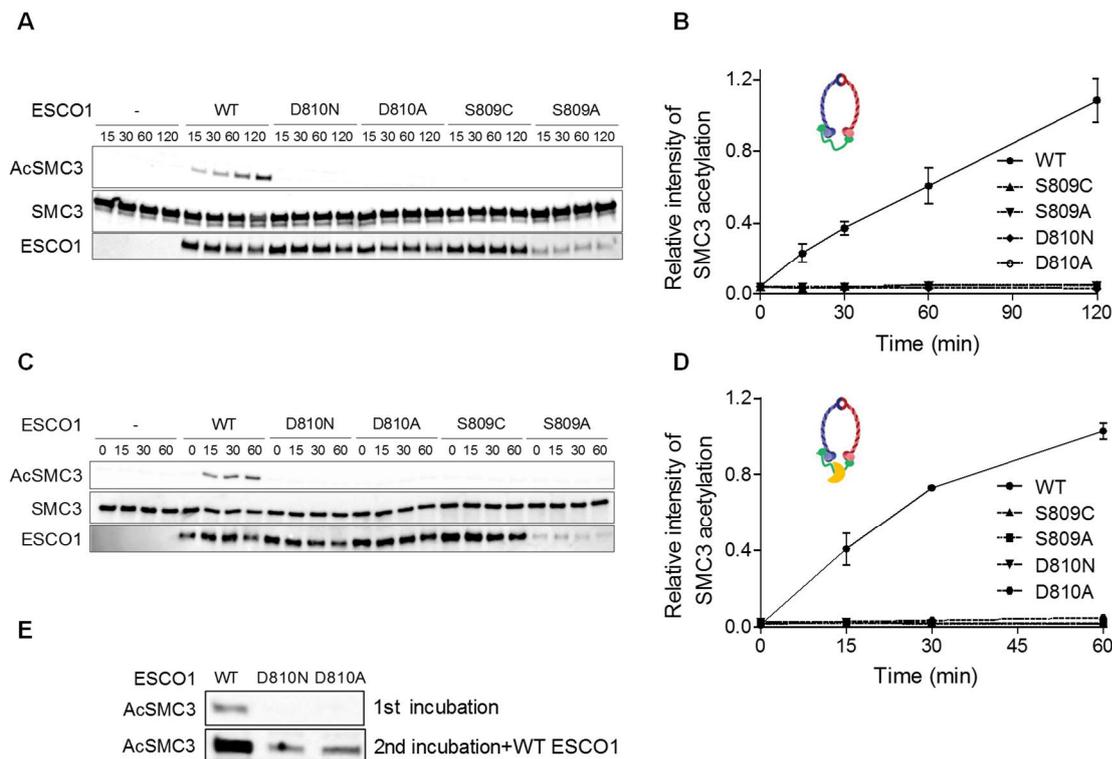
Figure 20: S809 and D810 are crucial for HsESCO1 activity under *in vitro* condition.
Time course of SMC3 acetylation after incubation of purified trimeric cohesin (**A**) and (**B**) and tetrameric (**C**) and (**D**) complexes with wild type or mutants HsESCO1. The molar ratio of HsESCO1 to trimer and tetramer was 1:10 and 1:2 respectively. Relative intensity of SMC3 acetylation level was analyzed by quantitative Western blotting using a specific AcSmc3 antibody. Data were normalized to maximal signal and are shown as mean ± SEM (n=2). (**E**) Test of whether contaminations of mutant protein preparation are inhibiting. The first 30 min of incubation were with trimeric cohesin complex and wild type or mutant HsESCO1. In the subsequent second incubation (30 min), wild type HsESCO1 was added to all reactions. Wild type enzyme produced acetylated SMC3 although mutant enzyme was present.

3.5 Effect of catalytic site mutants of Esco1 and Esco2 on *in vivo* acetyltransferase activity

In vitro assay with tetrameric cohesin shows efficient HsESCO1-mediated acetylation of the SMC3 cohesin subunit but all mutations of candidate catalytic residues failed to acetylate SMC3. This suggests that S809 and D810 are both indispensable for SMC3 acetylation. This poses the question of whether that also holds true for catalytic activity *in vivo* where additional cohesin components such as Wapl, Pds5A, Pds5B, SA1, SA2, Sororin and Nipbl are present (Carretero et al., 2013; Nishiyama et al., 2010; Remeseiro et al., 2012; Remeseiro et al., 2013; Tedeschi et al., 2013). Therefore, we used *in vivo* assays that assess Smc3 acetylation

and sister chromatid cohesion in various mutants of MmEsco1/2 and MmEsco2 mouse embryonic fibroblast, respectively.

3.5.1   Effect of catalytic site mutants of MmEsco1 on Smc3 acetylation *in vivo*

To investigate the activity of MmEsco1 mutants, we used a complementation assay in which the capability of mutants in Smc3 acetylation were assessed in *Esco1*-deficient MEFs.

Esco1-deficient MEFs (MEFs$^{Esco1-/-}$) were obtained from *Esco1$^{-/-}$* mouse embryos. MEFs$^{Esco1-/-}$ were transiently transfected with C-terminally myc-tagged *MmEsco1*, and the following myc-tagged mutant constructs: *MmEsco1^{S812A}*, *MmEsco1^{S812C}*, *MmEsco1^{D813A}*, and *MmEsco1^{D813N}*. Synchronization of cells in G1-phase showed that the mutations did not affect cell cycle progression (Figure 21A). Cells deficient for *Esco1* showed a marked reduction in Smc3 acetylation compared to controls (Figures 21C and 21D). Residual Smc3 acetylation is likely due to MmEsco2 that is present at low-amounts during the G1-phase (Lafont et al., 2010; Whelan et al., 2012). Transfection with *MmEsco1-myc* resulted in strong Smc3 acetyl ation.

Transfection of MEFs$^{Esco1-/-}$ with different MmEsco1 single mutants showed that only MmEsco1^{D813N} could not restore Smc3 acetylation above background (Figures 21D and 21E). In fact, MmEsco1^{D813A}, MmEsco1^{S812A}, and MmEsco1^{S812C} showed 30-50% residual acetyltransferase activity, relative to wild type (Figures 21D and 21E). None of these substitutions is able to deprotonate the substrate. However, if either of these two residues S812 and D813 is retained, then either the hydroxyl group of serine or the carboxyl group of aspartate could act as a general base. Hence, we transfected the double mutant MmEsco1$^{S812A;D813A}$ into MEFs$^{Esco1-/-}$. This double mutant resulted in base-line level activity of MmEsco1 (Figures 21D and 21E) implying that both of S812 and D813 can perform deprotonation of the substrate lysines.

The residual activity of the MmEsco1^{D813A} mutant argues that another general base could be involved in the deprotonation of the substrate in case of D813 being incapacitated. A good candidate is E728. However, the carboxyl group of this residue is nearly 10 Å away from the carbonyl group of the acetylated CoA model structure (Figures 11A and 11B). Thus, direct substrate lysine deprotonation by E728 is less likely and would require either water molecules or polar residues as proton shuttle. S812 and S773 are well positioned to transfer the proton from the substrate lysine ε-amino group to the E728 (Figure 11B). The MmEsco1^{E728Q} mutant

showed 40% residual activity relative to wild type suggesting that E728 could indeed play the role of a general base if D813 is mutated. A double mutant of D813 and E728 (MmEsco1[D813A;E728Q]) was enzymatically inactive supporting the assumption that one of the general bases, D813 or E728, can deprotonate the ϵ-amino group of substrate lysines. MmEsco1[S812A] and MmEsco1[S773A] mutants still reached 30-40% of activity in acetylating Smc3, relative to wild type. However, the double mutant MmEsco1[S812A;S773A] was completely inactive. These findings suggest that either of these two serines can act as a proton shuttle.

Altogether, the most cautious interpretations of our data suggest that S812, D813, S773 and E728 residues can redundantly cooperate in deprotonation. In case of a mutation they can replace each other in an *in vivo* complementation assay. *In vitro*, however, such substitution is not seen.

Importantly, the catalytically inactive MmEsco1 mutants bound to the chromatin in the same manner as reported for the endogenous MmEsco1 (Hou and Zou, 2005) (Figure 21B). This would thus argue against an effect on protein mislocalization.
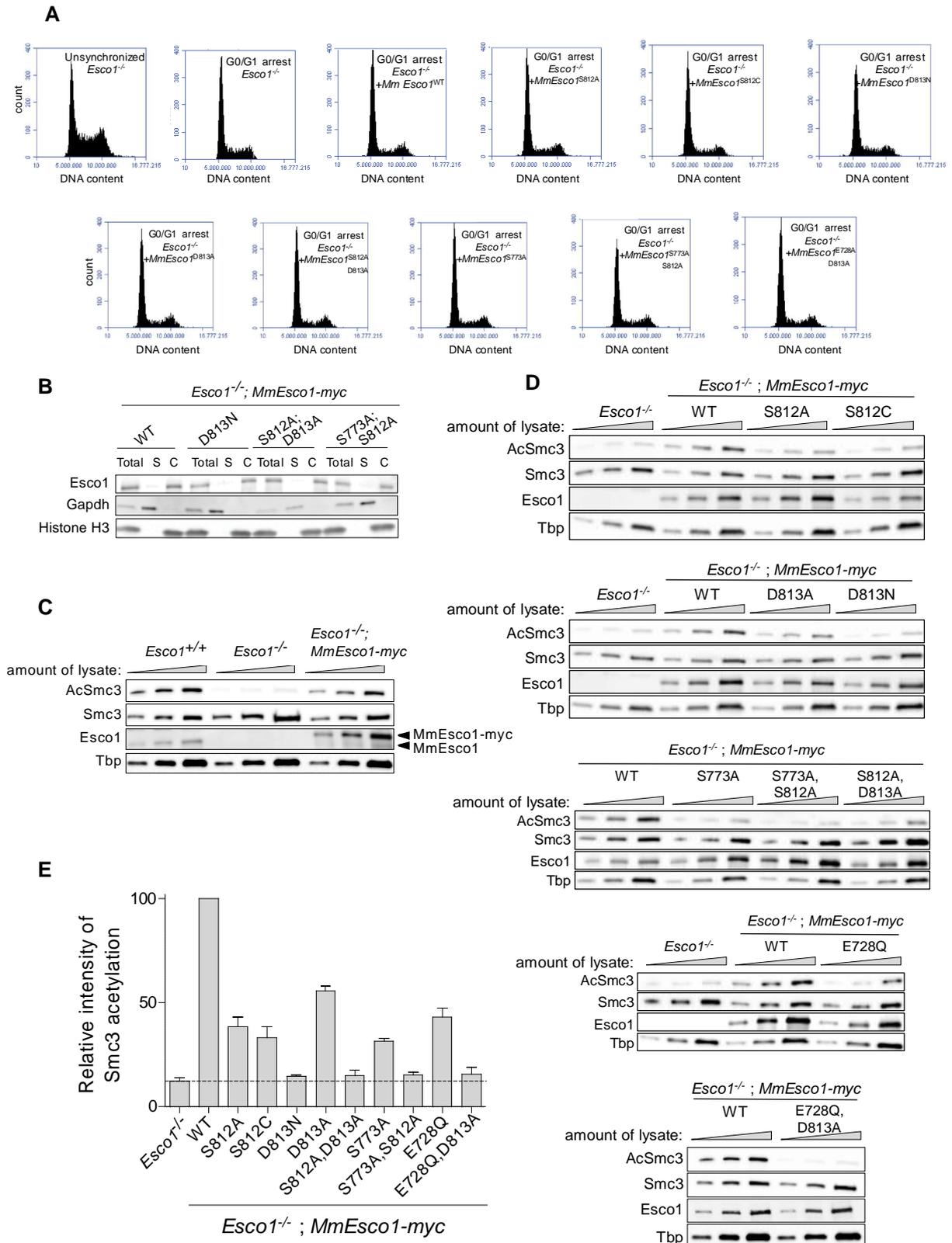
Figure 21: S812, D813, S773, and E728 function cooperatively in the catalysis of MmEsco1.
(**A**) Flow cytometry profiles of G1-phase arrested MEFs$^{Esco1-/-}$ expressing wild type or mutant MmEsco1. Unsynchronized MEFs$^{Esco1-/-}$ were used as controls. (**B**) Immunoblot analysis of MEFs$^{Esco1-/-}$ expressing wild type and catalytically dead mutants of MmEsco1-myc. Wild type and mutants localize to chromatin. S: soluble fraction, C: chromatin. (**C**) Establishment of an *in vivo* acetylation assay in MEFs$^{Esco1-/-}$. The

Smc3 acetylation level of chromatin-bound fractions was analyzed by Western blotting using an AcSmc3-specific antibody. The lysate for each sample was loaded in three dilutions: 1X, 2X, and 4X. Of note, transiently transfected MmEsco1 was expressed at a level about 4 fold that of endogenous MmEsco1 (compare lanes 3 and 7). **(D)** Representative immunoblot of MEFs$^{Esco1-/-}$ transiently expressing wild type or mutant versions of MmEsco1-myc, arrested in G0/G1-phase. The ability of the MmEsco1 mutants to acetylate Smc3 was assessed by quantitative Western blotting using an AcSmc3-specific antibody. Chromatin-bound fractions were used. Note that MEFs$^{Esco1-/-}$ expressed comparable levels of different MmEsco1-myc variants. The lysate for each sample was loaded in three dilutions: 1X, 2X, and 4X. **(E)** Quantification of the data shown in (D). The dotted line indicates the Smc3 acetylation value for MEFs$^{Esco1-/-}$. Data were normalized to maximal signal and are shown as mean ± SEM (n=3).

3.5.2 Effect of catalytic site mutants of MmEsco2 on Smc3 acetylation

The above *in vivo* complementation results were examined using mutant MmEsco2 versions, transfected into *Esco2*-deficient MEFs.

Esco2 deficient MEFs (MEFs$^{Esco2-/-}$) were generated by transduction of MEFs$^{Esco2fl/fl}$ with adenoviral Cre (AdCre). MEFs$^{Esco2-/-}$ were stably transfected with C-terminally myc-tagged *MmEsco2 and* myc-tagged mutants *MmEsco2^{S566A}*, *MmEsco2^{D567A}*, *MmEsco2^{D567N}*, *MmEsco2^{S527A}*, *MmEsco2^{E491Q}*, *MmEsco2$^{S566A;D567A}$*, and *MmEsco2$^{S527A;S566A}$* (Figure 22B). Synchronization of cells in S-phase showed that the mutant versions did not affect cell cycle progression (Figure 22A). Cells deficient for *Esco2* showed reduced Smc3 acetylation (Figure 22C). Residual Smc3 acetylation is most likely due to MmEsco1 that is expressed throughout the cell cycle (Lafont et al., 2010; Whelan et al., 2012). Transfection of *Esco2*-deficient MEFs with *MmEsco2* but not with a *MmEsco2^{D567N}* mutant construct and the two double mutant constructs, *MmEsco2$^{S566A;D567A}$* and *MmEsco2$^{S527A;S566A}$*, restored Smc3 acetylation (Figures 22D and 22E). These results are in close agreement with the effect of the same mutants on MmEsco1 activity. However, the effect of single mutations (S566A, D567A, S527A and E491Q) on the degree of Smc3 acetylation was less pronounced than that observed for the MmEsco1 isoform. For instance, MmEsco2^{S566A} and MmEsco2^{S527A} showed about 60-80% residual acetyltransferase activity relative to wild type. In addition, MmEsco2^{D567A} and MmEsco2^{E492Q} mutants acetylated Smc3 to the same extent as its wild type control (Figures 22D and 22E). This difference between Esco1 and Esco2 could be due to the fact that *in vivo* Esco1 and Esco2 act in a different molecular context. For example, only Esco2 acetylates Smc3 in the context of DNA replication (Hou and Zou, 2005; Ivanov et al., 2018).

In summary, our *in vivo* mutant data propose that MmEsco2 and MmEsco1 function in different contexts but show similarity in their catalytic mechanism. In agreement with MmEsco1, MmEsco2 can engage alternative catalytic residues S566, D567, S527, and E491 that redundantly act in deprotonation of substrate lysines.

Importantly, all MmEsco2 variants bound to the chromatin in the same manner as reported for the endogenous MmEsco2 (Hou and Zou, 2005) (Figure 22B). This would thus argue against an effect on protein mislocalization.
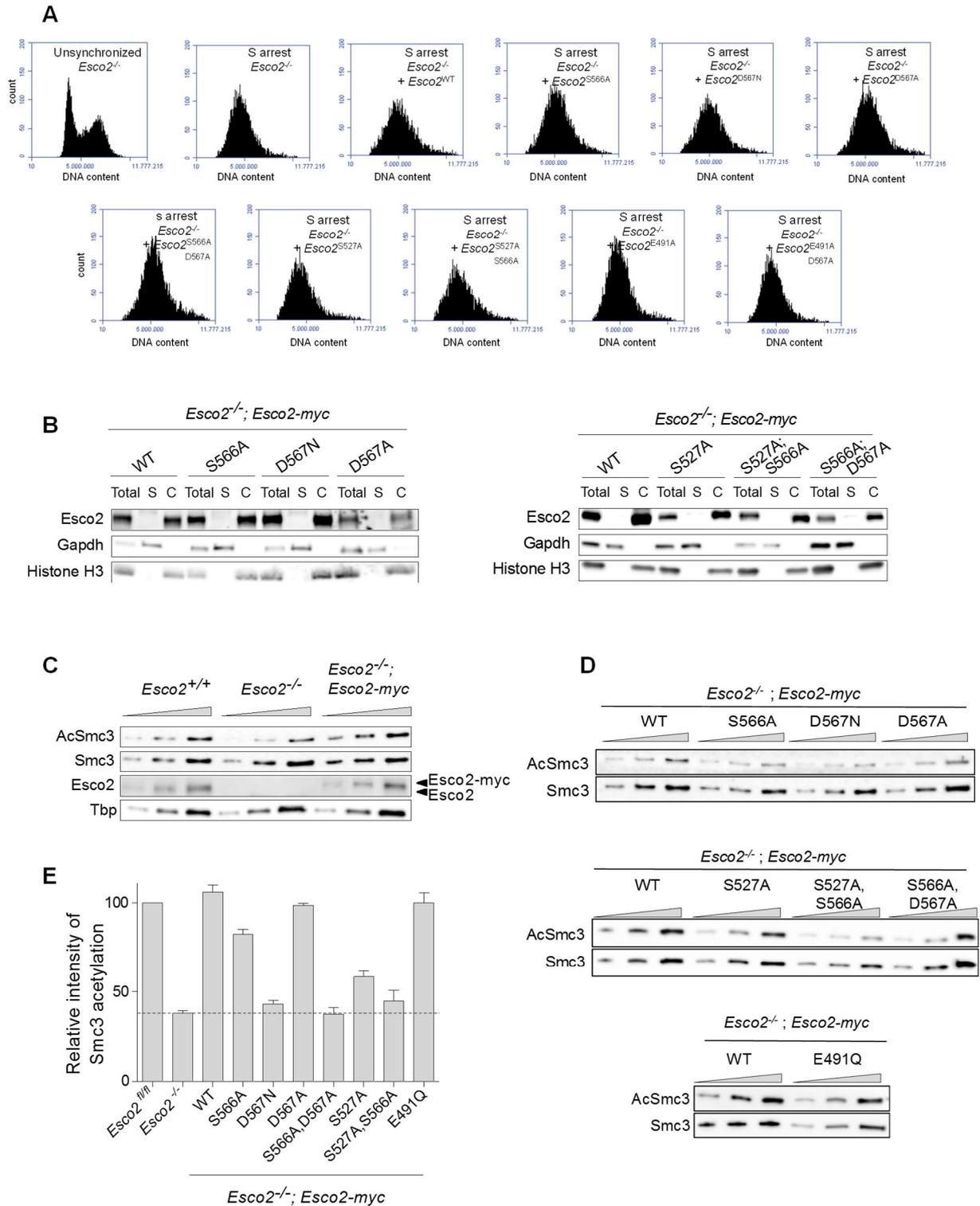
Figure 22: S566, D567, S527, and E491 function cooperatively in the catalysis of MmEsco2.
(**A**) Flow cytometry profiles of S-phase arrested MEFs$^{Esco2-/-}$ expressing wild type or mutant MmEsco2. Unsynchronized MEFs$^{Esco2-/-}$ were used as controls. (**B**) Immunoblot analysis of MEFs$^{Esco2-/-}$ expressing equivalent amounts of ectopic MmEsco2-myc variants. Wild type and mutant MmEsco2-myc localize to chromatin. S: soluble fraction, C: chromatin. Representative mutants are shown. (**C**) Establishment of an *in vivo* acetylation assay in MEFs$^{Esco2-/-}$. The Smc3 acetylation level of chromatin-bound fractions was analyzed by Western blotting using an AcSmc3-specific antibody. The lysate for each sample was loaded in three dilutions: 1X, 2X, and 4X.

(D) Representative immunoblot of MEFs$^{Esco2-/-}$ expressing wild type or mutant versions of MmEsco2-myc, arrested in S-phase. The ability of the MmEsco2 mutants to acetylate Smc3 was assessed by quantitative Western blotting using an AcSmc3-specific antibody. Chromatin-bound fractions were used. The lysate for each sample was loaded in three dilutions: 1X, 2X, and 4X. **(E)** Quantification of the data shown in (D). The dotted line indicates the Smc3 acetylation value for MEF$^{Esco2-/-}$. Data were normalized to signal of MEFs$^{Esco2fl/fl}$ and are shown as mean ± SEM (n = 3).

3.6 Effect of catalytic site mutants of MmEsco2 on sister chromatid cohesion

A number of naturally occurring mutations in *HsEsco2* cause Roberts syndrome (RBS), a developmental disease in which prometaphase chromosomes have a railroad track appearance (with separate centromeres and cohered arms). Most of these mutations lead to nonsense-mediated decay or affect the catalytic domain of the enzyme (Ivanov et al., 2002; Rivera-Colon et al., 2016). Therefore, loss of HsESCO2 acetyltransferase activity is believed to cause loss of cohesion in RBS patients. *Esco2*-deficiency in MEFs recapitulates the defect seen in RBS in mouse prometaphase chromosomes (Figure 23A) (Whelan et al., 2012). Therefore, to confirm the functional importance of the catalytic mutants of MmEsco2 and to complement the *in vivo* Smc3 acetylation analysis above (see 3.5), we investigated the effect of these mutants on sister chromatid cohesion.

MEFs$^{Esco2-/-}$ were transfected with wild type or different mutant versions of *MmEsco2*. Cells were synchronized in prometaphase, isolated by mitotic shake-off, and used for chromosome spread preparation. Consistent with the results from the Smc3 acetylation rescue experiments (Figures 22D and 22E), chromosome morphology analysis revealed that MmEsco2^{D567N}, MmEsco2$^{S566A;D567A}$ and MmEsco2$^{S527A;S566A}$ do not rescue the sister chromatid cohesion defect (Figure 23B). However, single mutants MmEsco2^{S566A}, MmEsco2^{D567A}, MmEsco2^{S527A}, and MmEsco2^{E491Q} that still showed intrinsic acetyltransferase activity (Figures 22D and 22E) also restored wild type appearance of cohesion at least to some extent (Figure 23B). These observations are in close agreement with the effect of MmEsco2 catalytic mutants on Smc3 acetylation and hence support the notion of a redundant function of S566, D567, S527, and E491 in the catalytic activity of MmEsco2.

Another phenotype that is related to sister chromatid cohesion is Aurora B localization on the chromosomes (Whelan et al., 2012; Carretero et al., 2013). Localization of Aurora B on centromeres is essential for centromeric cohesion through recruitment of Shugoshin 1 (Sgo1) (Kelly et al., 2010, Wang et al., 2010, Wang et al., 2011, De Antoni et al., 2012). It has been shown that *Esco2*-deficiency leads to relocalization of Aurora B from centromeres to the arms of prometaphase chromosomes (Figure 24A) (Whelan et al., 2012). To further confirm the functional importance of the catalytic MmEsco2 mutants, we studied the effect of two catalytic mutations of MmEsco2, D567A and D567N, on localization of Aurora B on mitotic chromosomes. MmEsco2^{D567A} is representative of mutants with no effect on enzyme activity and MmEsco2^{D567N} is one of the catalytic mutants that abolish the enzyme activity. Aurora B immunostaining of prometaphase chromosomes was done in MEFs$^{Esco2-/-}$ transfected with wild type or the two different mutant versions. The results showed a relocalization of Aurora B immunoreactivity from centromeres to the arms for the MmEsco2^{D567N} mutant but not the MmEsco2^{D567A} mutant (Figures 24A and 24B). Thus, the acetyltransferase-deficient MmEsco2^{D567N} mutant cannot restore the proper localization of Aurora B in *Esco2*-deficient cells. These observations confirm the involvement of the employed catalytic candidate residue, D567, in the catalytic activity of MmEsco2.

In summary, these findings in addition to other *in vivo* assays suggest that MmEsco2 uses at least four residues in the active site including S566, D567, S527, and E491 that cooperatively involve in the catalytic activity of the enzyme.
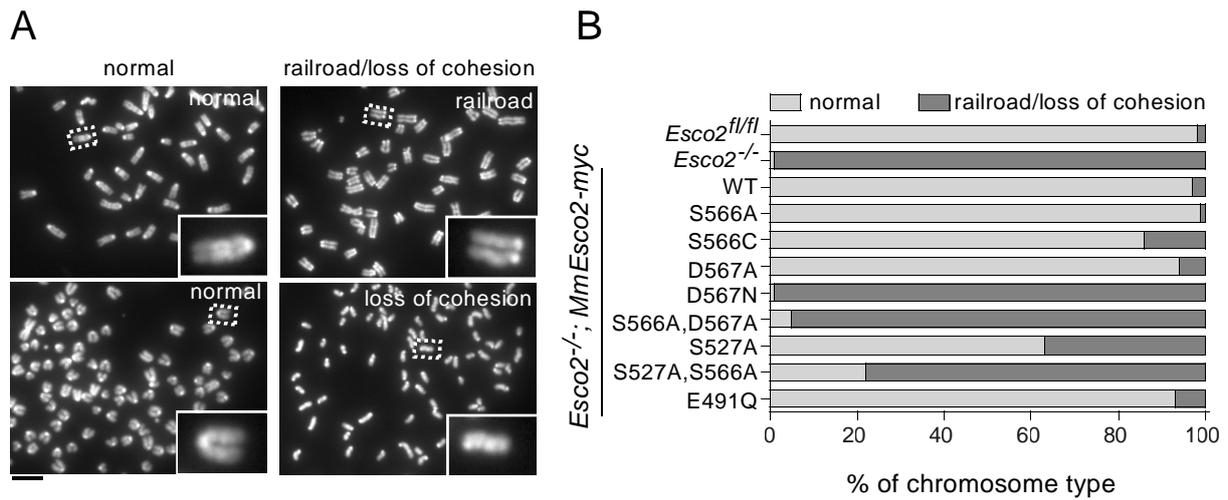
Figure 23: Sister chromatid cohesion is dependent on the acetyltransferase activity of Esco2.
(A) Representative prometaphase chromosomes spreads including normal, railroad and single chromatids. Scale bar: 10 µm **(B)** Frequency of chromosome types were assessed in 1000 prometaphase chromosomes per *MmEsco2* mutant.
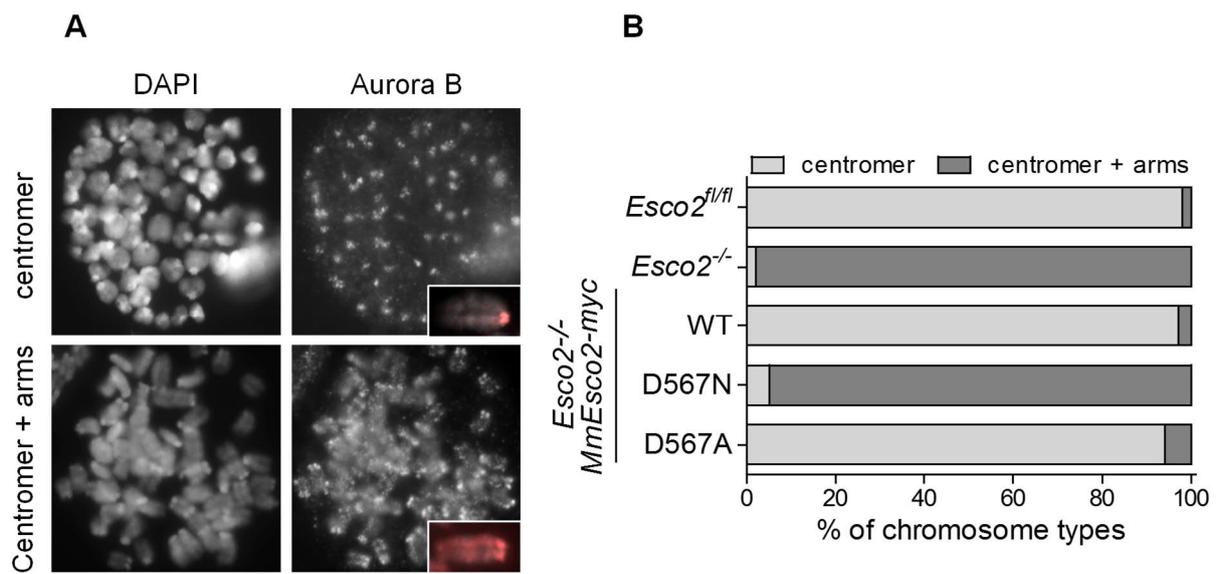
Figure 24: Aurora B localization is dependent on the acetyltransferase activity of MmEsco2.
(A) Representative immunoreactivity staining of Aurora B in prometaphase chromosomes spreads showing centromere localization or centromere-arms localization. Nuclei were counterstained with DAPI. Scale bar: 10 µm **(B)** Frequency of prometaphase chromosomes with Aurora B enrichment at the centromeres or centromeres and arms for D567 Esco2 mutants. 1000 prometaphase chromosomes per *MmEsco2* mutant were assessed.

4 Discussion

4.1 The architecture of MmEsco2 in complex with CoA

We determined the structure of the MmEsco2$^{368-592}$ acetyltransferase domain natively complexed with CoA. It contains a mixed α-helix and β-sheet structure forming a globular fold that is structurally conserved among all GNAT family members (Figure 10A). CoA is wedged into a groove consisting of two central β-strands (β7 and 8) and two α-helices (α3 and 4) (Figure 10A). The presence of CoA is likely required to stabilize the overall folding of the MmEsco2 acetyltransferase domain, as has previously been suggested for other acetyltransferases (Friedmann and Marmorstein, 2013). Our structure is from recombinant protein containing CoA but not AcCoA. Attempts to substitute CoA by AcCoA in the crystal were not successful. However, other Esco acetyltransferase domain crystal structures contain AcCoA and thus the position of the acetyl moiety can be modeled into the MmEsco2$^{368-592}$ acetyltransferase domain.

A structural similarity search against available structures in the Protein Data Bank (PDB) using DALI (Holm and Rosenstrom, 2010) or SSM (Krissinel and Henrick, 2004) revealed that acetyltransferase domain of MmEsco2 shows structural similarities to other GNAT family members (xEco2, HsESCO1, HsαTAT1, HsGCN5) as well as to the acetyltransferase domain of the MYST family member ScEsa1 (Table 4). Table 4 shows that the xEco2 folds have the highest Z-scores and lowest RMSD compared with other Esco orthologs. A conserved C-terminal extension and an extended β-hairpin between two β-strands in xEco2-K105-CoA mediate specific binding of the substrate (Chao et al., 2017). The corresponding regions in MmEsco2 are structurally conserved suggesting the same regions would be involved in substrate binding (Figure 25). It should be noted that in MmEsco2 and xEco2, part of the extended β-hairpin (S501-C515 of MmEsco2) (L609-W623 of xEco2) is unresolved suggesting structural flexibility that might enable interaction with subunits of the cohesin ring or cohesin regulators.

Table 4: Structural similarity search using SSM.

Proteins(accession numbers)	RMSD* (Å)	length of alignment	Identity (%)	Z-score
xEco2 (5N1W)	0.78	152	68	15.7
xEco2 (5N1U)	0.92	154	68	14.6
HsESCO1 (4mxe)	1.13	155	58	12.9
HsESCO1 (5T53)	1.58	152	58	10.7
HsαTAT1 (4PK2)	2.23	125	10	8.2
ScEsa1 (1MJB)	2.46	126	10	6.9
HsGcn5 (1Z4R)	2.65	120	11	7.9

* RMSD: root-mean-square deviation
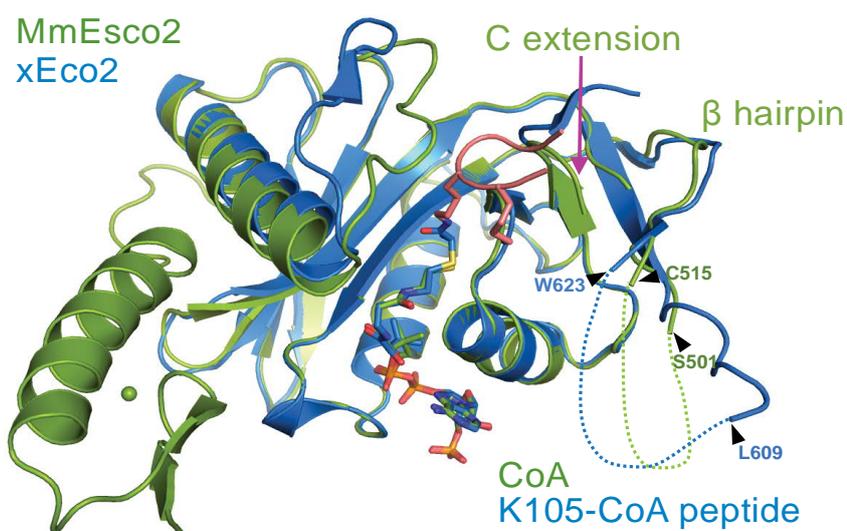
Figure 25: Superposition of MmEsco2/CoA with xEco2/K105-CoA.
Alignment of MmEsco2/CoA (green) and the xEco2/K105-CoA peptide (blue; PDB ID code 5N1W). The Smc3 peptide K105 is shown in raspberry. The functionally important β hairpin and C-terminal extension of MmEsco2 are labeled. Note that one part of the β hairpin is unresolved (shown with dashed lines) in the acetyltransferase domain of MmEsco2 and xEco2.

4.2 Characterization of the active site of Esco1 and Esco2

The active site of MmEsco2[368-592] (Figure 11A) is readily identified by the location of the co-factor coenzyme A. Based on the interatomic distances, four potential catalytic residues could be identified (S566, D567, S527 and E491) (Figures 11A and 11B). To explore whether these residues contribute to catalysis, we developed *in vitro* acetyltransferase assays as well as cell-based assays that detect cohesin acetylation *in vivo*. For *in vitro* assays, we used recombinant HsESCO1, since we failed to produce sufficient amounts of either mouse or human recombinant Esco2.

4.2.1 *In vitro* cohesin acetylation requires ATP hydrolysis and DNA

So far, only autoacetylation of Esco1 and Esco2 and Smc3 peptide acetylation by Esco1 have been shown. These are slow reactions and thus may not properly recapitulate *in vivo* activity of Esco enzymes. The sole exception to this was Peters's lab work, that has used trimeric cohesin without nucleic acid (Ladurner et al., 2014). To better model the true substrate, we used tetrameric cohesin complex. We found that full-length HsESCO1 can acetylate SMC3 *in vitro* in the presence of both ATP and DNA. The requirement for DNA is consistent with the fact that *in vivo* only chromatin bound cohesin is acetylated (Morales and Losada, 2018). Moreover, requirement for ATP is consistent with the finding that cohesin loading onto chromatin and subsequent cohesin acetylation depends on ATP hydrolysis mediated by the SMC3 head domain (Camdere et al., 2015; Ladurner et al., 2014; Murayama and Uhlmann, 2014).

Of note, we found that addition of DNA and/or ATP did not enhance HsESCO1 autoacetylation. Therefore, the physiological function of nucleic acids and ATP has to be viewed in the context of Esco proteins acetylating the cohesin ring.

4.2.2 *In vitro* analysis of active site mutants of HsESCO1

The active site model shown in Figure 11A and 11B implicates S566 and D567 as potential catalytic residues in MmEsco2[368-592]. Accordingly, equivalent serine and aspartic acid residues in HsESCO1 (see Figure 11D) were mutated as follows: S809A, S809C, D810A, and D810N. All mutants were catalytically inactive in the *in vitro* assay. In agreement with these results, an *in vitro* mutagenesis study of recombinant HsESCO1[590–840] with a 21 amino acid long SMC3 peptide as substrate also showed that S809A and D810N mutations are catalyti-

cally inactive (Rivera-Colon et al., 2016). This suggests that both S809 and D810 (S566 and D567 of MmEsco2) are involved in catalysis. The position of D810 in the structure of xEco2 complexed with a K105-CoA-Smc3 peptide is such that it could act as a general base to deprotonate the ε-group of the K105 (Figure 12A). Chao (2017) also noted the presence of a salt bridge between D810 and K106 of the Smc3 peptide. Hence, it is possible that D810 is involved in an enzyme-substrate interaction and additionally abstracts a proton from K105. We propose that there could be a proton transfer from K105 to D810 that involves S809 as a proton shuttle. Once K105 is acetylated, CoA dissociates and a new cofactor is bound to Apo Esco. As a next step, the holoenzyme now acetylates K106. Because D810 is no longer properly positioned for salt bridge formation, it could act either directly as the general base or S809 again acts as a proton shuttle. This model would explain why S809A, S809C, D810A, and D810N mutations are all catalytically inactive.

4.2.3 *In vivo* analysis of active site mutants of MmEsco1

Are the above active site mutants catalytically inactive *in vivo*? Transfecting different *MmEsco1* mutants into *Esco1*-deficient MEFs and assessing Smc3 acetylation showed that all mutants except for D813N had surprisingly still partial activity. All mutants used in our work have been shown to be properly folded based on their gel filtration elution profiles (Rivera-Colon et al., 2016). The full loss of activity of the D813N mutant despite the residual activity of the D813A mutant enzyme support the view that D813 is not only involved in catalysis but also in substrate binding (see above). Engagement of an active site residue in both catalysis and substrate binding has also been found for the HsαTAT1 acetyltransferase (Szyk et al., 2014). D157 of HsαTAT1 acts as a general base in catalysis and interacts with the substrate lysine (Szyk et al., 2014). D157 is located in the equivalent position as D813 MmEsco1 and D567 in MmEsco2 (Figure 14). Reminiscent of MmEsco1, a D157N mutation in HsαTAT1 resulted in a complete loss of acetyltransferase activity (Taschner et al., 2012).

Why are S812A/C and D813A partially active *in vivo*? As pointed out above, β-strands 6 and 7 project side chains of E728 and S773 into the active site. These residues could act as a general base and proton shuttle, respectively. Single mutations of residues S773A and E728Q did not abrogate the catalytic activity of MmEsco1. However, mutations of both general bases (E728Q;D813A) led to a complete loss of enzymatic activity. This suggests that either D813 or E728 can deprotonate substrate lysines. We have argued above that S812 acts as a proton

shuttle and S773 could do the same. This possibility is supported by the absence of any enzymatic activity in S812A;S773A MmEsco1 double mutant. The role of serine as proton shuttle in substrate deprotonation was also proposed for the Dat acetyltransferase, which employs a glutamate residue to deprotonate the amino group of the substrate through a serine residue (Cheng et al., 2012).

It is possible that S773, S812, D813 and E728 play a redundant role in deprotonation of the substrate even in the native enzyme. However, it is also possible that the catalytic activity of E728 and S773 comes to the fore only in the absence of S812 and D813. An equivalence of the two pairs would require similar rates of catalysis mediated by either pair of residues. Currently these rates are unknown. However, in our *in vivo* context, we observed that extend of Smc3 acetylation is very similar in either E728 and S773 or S812 and D813 mutations.

In vitro, human ESCO1 mutations S809A, S809C, D810A and D810N are inactive. Interestingly, we found that the mouse ortholog residues mutations (S812A, S812C and D813A) showed up to 50% of wild type activity *in vivo*. The other single MmEsco1 mutants, S773A and E728Q, with partial catalytic activity in our *in vivo* assays are also catalytically inactive *in vitro* (Rivera-Colon et al., 2016). Why is that? It is known that cohesin regulators Pds5A, Pds5B, MCM and chromatin are required for Smc3 acetylation in cells (Carretero et al., 2013; Ivanov et al., 2018; Minamino et al., 2015). Therefore, we speculate that presence of one or several of these factors enhance Esco activity. This limitation makes the point that an *in vivo* assessment of various Esco mutant activity is crucial for understanding catalysis.

4.2.1 *In vivo* analysis of active site mutants of MmEsco2

The results from the *in vivo* mutational studies for MmEsco2 using both Smc3 acetylation and sister chromatid cohesion as readouts are consistent. Thus, mutations impairing Smc3 acetylation also cause sister chromatid cohesion defects. Furthermore, identical mutations in either MmEsco1 or MmEsco2 evoke Smc3 acetylation deficiencies to a similar qualitative extent (Figures 21E and 22E). Nonetheless, there are some quantitative differences. An example for such a quantitative difference is the effect of the single mutation D813A in MmEsco1 (corresponding to D567A in MmEsco2). While the extent of Smc3 acetylation is reduced by 50% in the MmEsco1 mutant, wild type levels are retained in the MmEsco2 mutant. This difference between Esco1 and Esco2 could be due to the fact that they act in a very different molecular

context *in vivo*. Esco1 directly acetylates Smc3 while bound to chromatin. In contrast, Esco2 acetylates cohesin only in the context of a DNA replication fork (Hou and Zou, 2005; Ivanov et al., 2018).

Taken together, we proposed that the general base in either Esco1 or Esco2 abstracts the substrate proton *via* a serine residue (Figure 26). This hypothetical mechanism is based on the crystallographic evidence along with site-directed mutagenesis performed in this study as well as in recent reports from other laboratories (Chao et al., 2017b; Rivera-Colon et al., 2016). In the case of mutagenesis of D567 or S566, active site residues E491 and S527 can compensate. In a way, the active sites of Esco1 and Esco2 are mirror symmetrical. Esco1 and Esco2 sequentially acetylate two distinct lysine residues located next to each other on Smc3. It is possible that the dual active reflects this task.
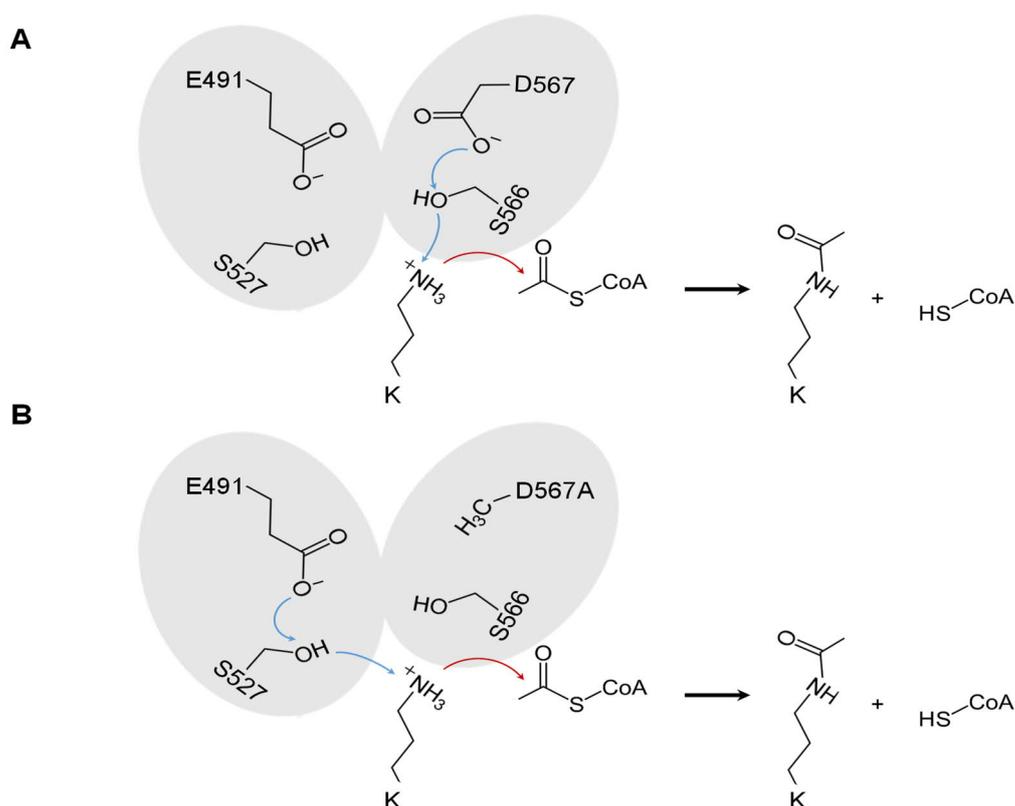
Figure 26: Proposed catalytic mechanism for Esco1 and Esco2.
(A) The proposed mechanism involves D566, acting as a general base, to initiate the reaction by abstracting a proton from the hydroxyl group of S566. Subsequently, the hydroxylate of S566 can then act as base catalyst to deprotonate the amino group of K105 and /or K106. This is followed by the nucleophilic attack of the amine on the carbonyl carbon of AcCoA. Blue arrows indicate deprotonation and red arrows the nucleophilic attack. The numbering of putative catalytic residues is based on the MmEsco2 sequence (Figure 11D). **(B)** In the case of mutagenesis of the D567 active site residue, E491 using S527 as a proton shuttle compensate.

4.3 Structural interpretation of Roberts syndrome mutations

The structure of MmEsco2$^{368-592}$ provides an interpretation of the Roberts-syndrome-associated mutations (Vega et al., 2010). In Roberts syndrome most mutations lead to stop codons in the amino terminal domain that are not present in the MmEsco2$^{368-592}$ structure. However, W539G, G581R, and deletion of E453 are in the catalytic domain. Human W539 (W530 in MmEsco2), which resides on the β7-strand, is part of a hydrophobic core (Figure 27) that might stabilize the AcCoA binding fold and α2-helix (Figure 27). Our xEco2-Smc3 docking study suggests that the α2 helix could represent an important domain for substrate binding (Figure 28). Therefore, mutation of W539 might indirectly interfere with the substrate binding. In addition, it could also disturb the AcCoA binding and hence, indirectly reduce the enzyme stability. This idea is supported by the finding that an equivalent mutation in HsESCO1 (W773G) leads to a dramatic reduction in thermal stability of the mutant protein and possibly substrate binding (Rivera-Colon et al., 2016). E453 is located in the middle of the α2-helix (Figure 27), a potential region for substrate binding (see above). Removing E453 from the helix could results in a ~120° reorientation of the side chains of adjacent C-terminally positioned residues. Thus, there is a rotation that puts hydrophobic residues of α2 towards the protein surface and hydrophilic residues towards the inside of the catalytic domain, which subsequently could impair substrate binding. G581 (G572 in MmEsco2) is located in the α4-helix and is in close proximity to CoA (Figure 27). This implies that a G to R mutation may sterically obstruct the entry of AcCoA into the active site.
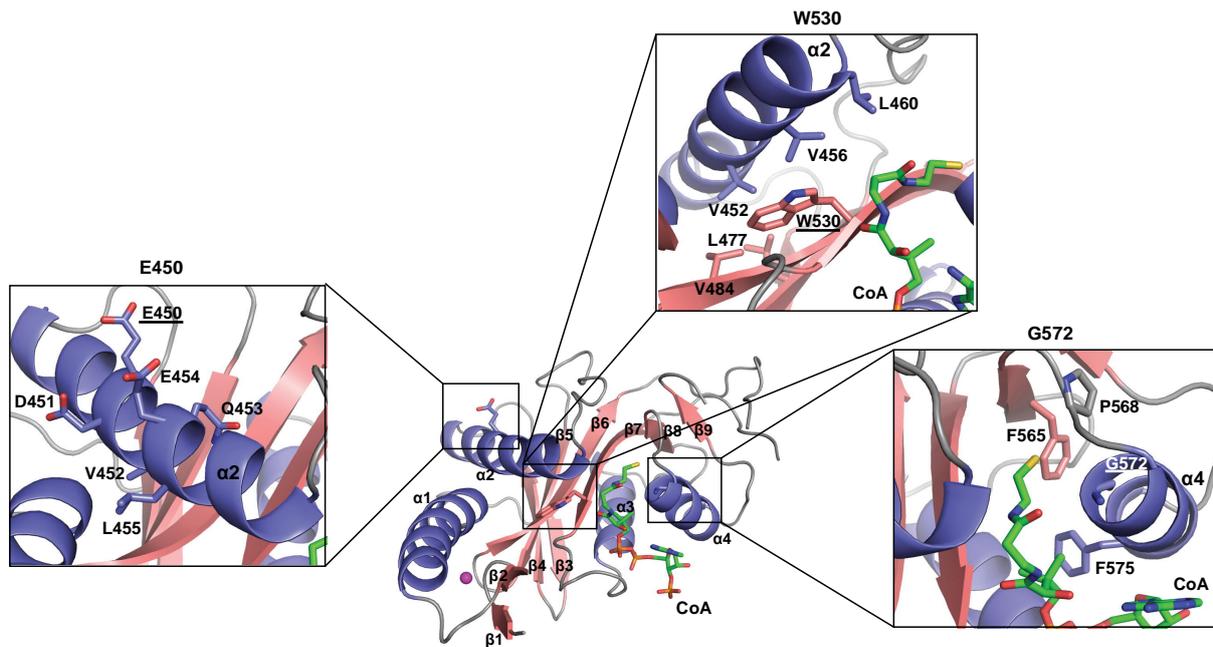
Figure 27: Disease-associated *Esco2* mutations.
Esco2 residues that are mutated in Roberts Syndrome are underlined. Adjacent interacting residues are labeled. The numbering of putative catalytic residues is based on the MmEsco2 sequence.
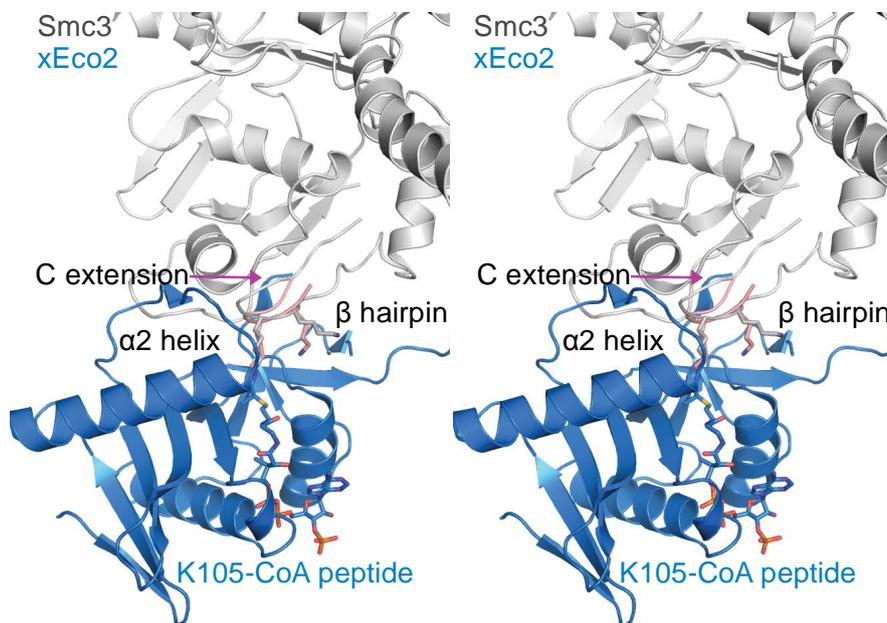
Figure 28: Substrate binding sites of Esco acetyltransferases.
Stereo view of the docking of the xEco2-K105-CoA structure (xEco2 in blue and Smc3 peptide in salmon) onto the *S. cerevisiae* Smc3 (gray; PDB ID 4UX3) structure based on the positions of the K105 and K106 (*S. cerevisiae* K112 and K113) residues. Potential substrate binding regions, α2 helix, β hairpin and C extension are labeled.

5 Additional Data

5.1 MmEsco2$^{368-592}$ shows inefficient autoacetyltransferase activity

To investigate the contribution of the potential catalytic residues, a thorough characterization of the corresponding mutants is essential. To do this, we first tried to establish an *in vitro* acetylation assay using MmEsco2$^{368-592}$ and different lengths of the Smc3 peptides (14, 24 and 29 residues). The results showed that MmEsco2$^{368-592}$ could not acetylate the Smc3 peptides, at least under our *in vitro* conditions (Figure 29A). However similar to previous reports, the enzyme showed autoacetylation with a very slow rate (Gordillo et al., 2008) (Figures 29A and 29C). After establishment of the *in vitro* autoacetylation assay, the acetyltransferase activity of different MmEsco2$^{368-592}$ mutants (S566A, S566C, D567A and D567Q) was assessed and compared to wild type MmEsco2$^{368-592}$. Unexpectedly, all mutants showed some degree of autoacetylation activity (Figure 29D). This was in contrast to the strong structural evidence for engagement of the two adjacent residues, S566 and D567, in catalysis. Since, MmEsco2$^{368-592}$ autoacetylation was highly inefficient (saturation took more than 12 hours) (Figure 29C), we assume that some part of autoacetylation activity may result from non-enzymatic activity.
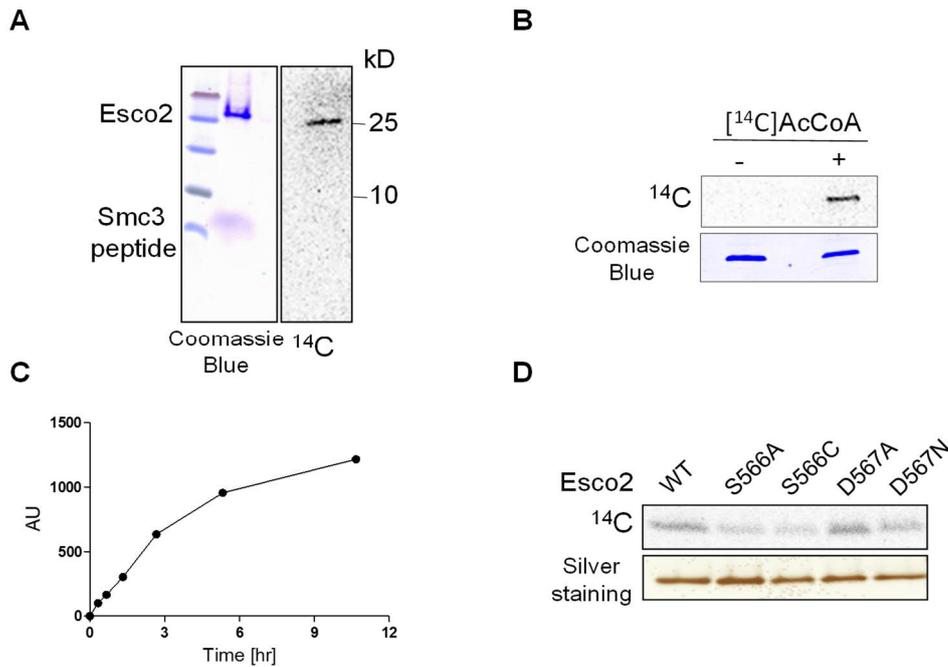
Figure 29: MmEsco2$^{368-592}$ shows inefficient autoacetyltransferase activity.
(A) MmEsco2$^{368-592}$ does not acetylate the Smc3 peptide. The Smc3 peptide (29 residues; 4.8 kDa) was incubated with purified MmEsco2$^{368-592}$ (28kD) and [^{14}C] AcCoA. Acetylation was analyzed by SDS-PAGE followed by Coomassie blue staining and phosphorimaging. **(B)** MmEsco2$^{368-592}$ shows autoacetylation activity. MmEsco2$^{368-592}$ (28kD) was incubated with or without [^{14}C] AcCoA and acetylation activity was assessed as in (A). **(C)** Autoacetylation activity of MmEsco2$^{368-592}$ is inefficient. Purified MmEsco2$^{368-592}$ was incubated with [^{14}C] AcCoA for different amounts of time. Acetylation activity was assessed as in (A) and normalized to time point t=0 hr. **(D)** The effect of MmEsco2$^{368-592}$ mutants on autoacetylation activity was assessed by SDS-PAGE followed by silver staining and phosphorimaging.

5.2   HsESCO2 does not acetylate the trimer under *in vitro* conditions

We could not purify sufficient amount of recombinant HsESCO2 for all *in vitro* assays. However, it was enough to assess the activity of this enzyme using a limited number of acetylation assays. MBP-tagged HsESCO2 was incubated with trimeric cohesin in the presence of ATP, DNA and AcCoA. The results showed that HsESCO2 does not acetylate the trimer, although an equal amount of HsESCO1 acetylates the trimer under the same *in vitro* conditions (Figure 30). One explanation for these results could be that HsESCO2 requires the presence of additional factors for cohesin acetylation. However, we do not rule out the possibility that the lack of enzymatic activity of HsESCO2 might be due to MBP fusion. Limitations in production of recombinant HsESCO2 did not allow us to study this issue further.
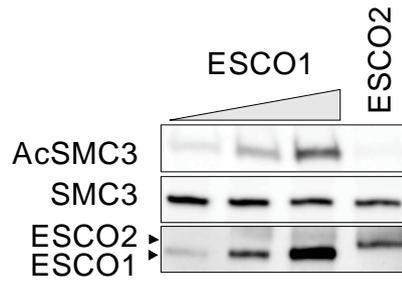
Figure 30: HsESCO2 does not acetylate the trimer under *in vitro* conditions.
Purified trimeric cohesin complex was incubated with increasing concentrations of HsESCO2 in the presence of AcCoA, ATP and DNA. HsESCO1 was used as a positive control. Note that the structures of the catalytic domains of ESCO1 and ESCO2 are highly conserved (Figure 13). SMC3 acetylation levels were analyzed by Western blotting using a AcSmc3-specific antibody.

5.3 *In vitro* cohesin acetylation shows salt sensitivity

Our results implicated the need of topological loading of cohesin on DNA for SMC3 acetylation. It has been shown that cohesin topological loading onto DNA requires low salt concentrations under *in vitro* condition. Therefore, we speculated that the rate of SMC3 acetylation might increase by reducing the salt concentration in our *in vitro* conditions to allow more efficient loading of cohesin onto DNA. To test this assumption, first trimeric cohesin complex, ATP and DNA were incubated in the presence of low and high concentrations of NaCl for 1 hour (first incubation). To initiate the acetylation process, HsESCO1 was added to the reactions and concentrations of NaCl were adjusted (second incubation). Interestingly, the results revealed that efficient SMC3 acetylation by HsESCO1 required an initial incubation of cohesin together with ATP and DNA in low salt conditions, probably to allow cohesin loading onto DNA (Figure 31). However, the acetylation process by HsESCO1 required a higher concentration of NaCl to be more efficient. These findings support our previous observation, that topological loading of cohesin onto DNA is required for cohesin acetylation.
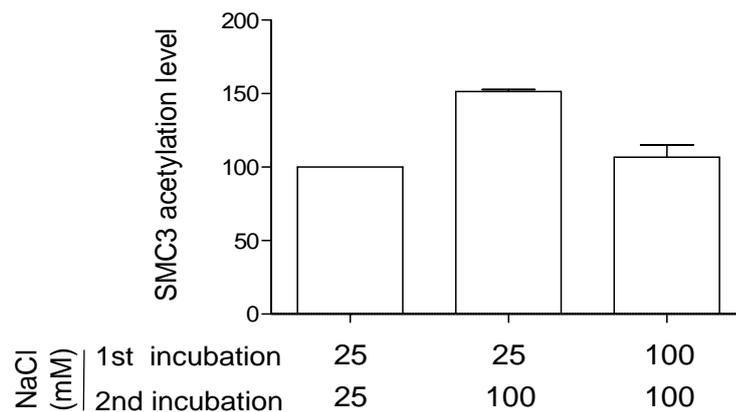
Figure 31: Salt sensitivity of cohesin acetylation *in vitro*.
The first incubation containing trimeric cohesin complex, ATP, AcCoA and DNA was carried out in the presence of the indicated concentrations of NaCl for 1 hour (first incubation). To initiate the acetylation process, HsESCO1 was added to the reactions in the presence of the indicated concentrations of NaCl (second incubation). SMC3 acetylation was analyzed by Western blotting using an AcSmc3 specific antibody.

5.4 S809 and D810 are crucial for the autoacetylation activity of HsESCO1 under *in vitro* conditions

HsESCO1 showed autoacetyltransferase activity independent of DNA and ATP under *in vitro* conditions (Figure 32A). Therefore, we quantified the autoacetylation of wild type and mutant versions of ESCO1 by Western blotting using an pan-acetyl antibody. In regard to the *in vitro* SMC3 acetylation results, autoacetylation of all tested mutants was markedly reduced (Figures 32B and 32C).

These results show that both S809 and D810 are required for the catalytic activity of the enzyme. However, superposition of MmEsco2/CoA and xEco2/K105-CoA suggests that S809 and D810 are also well positioned to play a role in substrate recognition and binding (Figure 12A).

A

C

B
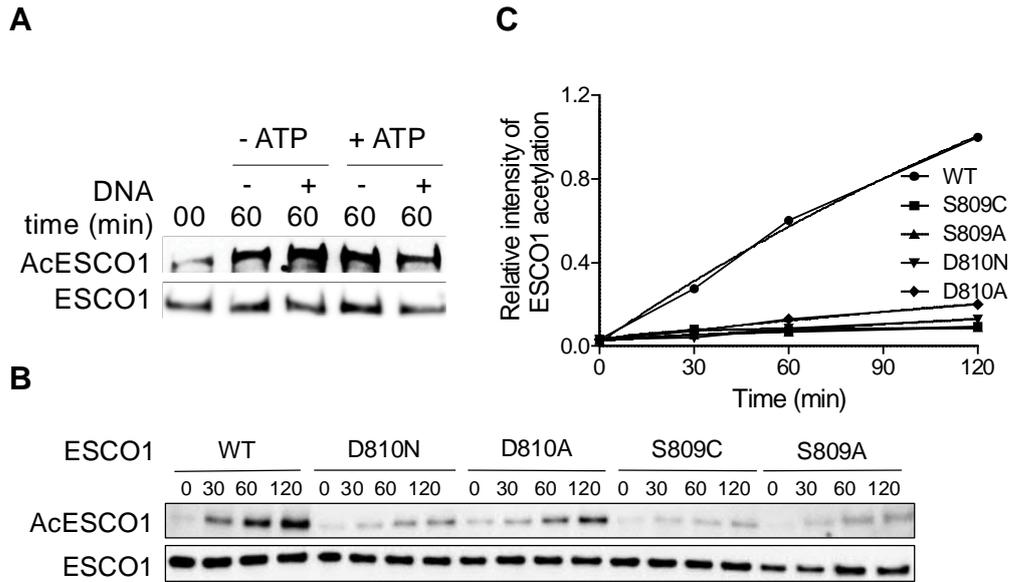
Figure 32: S809 and D810 are crucial for the autoacetylation activity of HsESCO1 under *in vitro* condition.

(A) ATP and DNA do not affect the autoacetylation activity of HsESCO1. Purified HsESCO1 was incubated with AcCoA in the presence or absence of ATP and DNA. The level of HsESCO1 autoacetylation was analyzed by Western blotting using an anti-pan-Aacetyl antibody. **(B)** and **(C)** Time course quantification of the outoacetylation activity of wild type and putative active site mutants of HsESCO1. Relative intensity of HsESCO1 autoacetylation was quantified as in (A). Data were normalized to the maximal signal and are shown as mean ± SEM (n=2).

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221.

Akhtar, A., and Becker, P.B. (2001). The histone H4 acetyltransferase MOF uses a C2HC zinc finger for substrate recognition. EMBO Rep 2, 113-118.

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

Barrington, C., Finn, R., and Hadjur, S. (2017). Cohesin biology meets the loop extrusion model. Chromosome Res 25, 51-60.

Beckouet, F., Srinivasan, M., Roig, M.B., Chan, K.L., Scheinost, J.C., Batty, P., Hu, B., Petela, N., Gligoris, T., Smith, A.C., et al. (2016). Releasing Activity Disengages Cohesin's Smc3/Scc1 Interface in a Process Blocked by Acetylation. Molecular Cell 61, 563-574.

Ben-Shahar, T.R., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M., and Uhlmann, F. (2008). Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion. Science 321, 563-566.

Berndsen, C.E., and Denu, J.M. (2005). Assays for mechanistic investigations of protein/histone acetyltransferases. Methods 36, 321-331.

Birkenbihl, R.P., and Subramani, S. (1992). Cloning and characterization of rad21 an essential gene of Schizosaccharomyces pombe involved in DNA double-strand-break repair. Nucleic Acids Res 20, 6605-6611.

Brent, M.M., Iwata, A., Carten, J., Zhao, K., and Marmorstein, R. (2009). Structure and biochemical characterization of protein acetyltransferase from Sulfolobus solfataricus. J Biol Chem 284, 19412-19419.

Camdere, G., Guacci, V., Stricklin, J., and Koshland, D. (2015). The ATPases of cohesin interface with regulators to modulate cohesin-mediated DNA tethering. Elife 4.

Carretero, M., Ruiz-Torres, M., Rodriguez-Corsino, M., Barthelemy, I., and Losada, A. (2013). Pds5B is required for cohesion establishment and Aurora B accumulation at centromeres. Embo J 32, 2938-2949.

Chao, W.C., Murayama, Y., Munoz, S., Costa, A., Uhlmann, F., and Singleton, M.R. (2015). Structural Studies Reveal the Functional Modularity of the Scc2-Scc4 Cohesin Loader. Cell Rep 12, 719-725.

Chao, W.C., Murayama, Y., Munoz, S., Jones, A.W., Wade, B.O., Purkiss, A.G., Hu, X.W., Borg, A., Snijders, A.P., Uhlmann, F., et al. (2017a). Structure of the cohesin loader Scc2. Nat Commun 8, 13952.

Chao, W.C., Wade, B.O., Bouchoux, C., Jones, A.W., Purkiss, A.G., Federico, S., O'Reilly, N., Snijders, A.P., Uhlmann, F., and Singleton, M.R. (2017b). Structural Basis of Eco1-Mediated Cohesin Acetylation. Sci Rep 7, 44313.

Cheng, K.C., Liao, J.N., and Lyu, P.C. (2012). Crystal structure of the dopamine N-acetyltransferase-acetyl-CoA complex provides insights into the catalytic mechanism. Biochem J 446, 395-404.

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

Clements, A., Rojas, J.R., Trievel, R.C., Wang, L., Berger, S.L., and Marmorstein, R. (1999). Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. Embo J 18, 3521-3532.

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

Davis, B.K. (1971). Genetic analysis of a meiotic mutant resulting in precocious sister-centromere separation in Drosophila melanogaster. Mol Gen Genet 113, 251-272.

Dyda, F., Klein, D.C., and Hickman, A.B. (2000). GCN5-related N-acetyltransferases: a structural overview. Annu Rev Biophys Biomol Struct 29, 81-103.

Eberharter, A., and Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep 3, 224-229.

Elbatsh, A.M.O., Haarhuis, J.H.I., Petela, N., Chapard, C., Fish, A., Celie, P.H., Stadnik, M., Ristic, D., Wyman, C., Medema, R.H., et al. (2016). Cohesin Releases DNA through Asymmetric ATPase-Driven Ring Opening. Mol Cell 61, 575-588.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486-501.

Friedmann, D.R., Aguilar, A., Fan, J., Nachury, M.V., and Marmorstein, R. (2012). Structure of the alpha-tubulin acetyltransferase, alphaTAT1, and implications for tubulin-specific acetylation. Proc Natl Acad Sci U S A 109, 19655-19660.

Friedmann, D.R., and Marmorstein, R. (2013). Structure and mechanism of non-histone protein acetyltransferase enzymes. FEBS J 280, 5570-5581.

Gimenez-Abian, J.F., Sumara, I., Hirota, T., Hauf, S., Gerlich, D., de la Torre, C., Ellenberg, J., and Peters, J.M. (2004). Regulation of sister chromatid cohesion between chromosome arms. Curr Biol 14, 1187-1193.

Gligoris, T.G., Scheinost, J.C., Burmann, F., Petela, N., Chan, K.L., Uluocak, P., Beckouet, F., Gruber, S., Nasmyth, K., and Lowe, J. (2014). Closing the cohesin ring: structure and function of its Smc3-kleisin interface. Science 346, 963-967.

Gordillo, M., Vega, H., and Jabs, E.W. (1993). Roberts Syndrome. In GeneReviews((R)), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA)).

Gordillo, M., Vega, H., Trainer, A.H., Hou, F., Sakai, N., Luque, R., Kayserili, H., Basaran, S., Skovby, F., Hennekam, R.C., et al. (2008). The molecular mechanism underlying Roberts syndrome involves loss of ESCO2 acetyltransferase activity. Hum Mol Genet 17, 2172-2180.

Gorr, I.H., Boos, D., and Stemmann, O. (2005). Mutual inhibition of separase and Cdk1 by two-step complex formation. Mol Cell 19, 135-141.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305-308.

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

Guacci, V., Koshland, D., and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. Cell 91, 47-57.

Haarhuis, J.H., Elbatsh, A.M., and Rowland, B.D. (2014a). Cohesin and its regulation: on the logic of X-shaped chromosomes. Dev Cell 31, 7-18.

Haarhuis, J.H.I., Elbatsh, A.M.O., and Rowland, B.D. (2014b). Cohesin and Its Regulation: On the Logic of X-Shaped Chromosomes. Dev Cell 31, 7-18.

Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G., and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature 460, 410-U130.

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

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

Hara, K., Zheng, G., Qu, Q., Liu, H., Ouyang, Z., Chen, Z., Tomchick, D.R., and Yu, H. (2014). Structure of cohesin subcomplex pinpoints direct shugoshin-Wapl antagonism in centromeric cohesion. Nat Struct Mol Biol 21, 864-870.

Hartwell, L.H., and Weinert, T.A. (1989). Checkpoints - Controls That Ensure the Order of Cell-Cycle Events. Science 246, 629-634.

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

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

Hengeveld, R.C.C., Vromans, M.J.M., Vleugel, M., Hadders, M.A., and Lens, S.M.A. (2017). Inner centromere localization of the CPC maintains centromere cohesion and allows mitotic checkpoint silencing. Nat Commun 8.

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

Holt, C.L., and May, G.S. (1996). An extragenic suppressor of the mitosis-defective bimD6 mutation of Aspergillus nidulans codes for a chromosome scaffold protein. Genetics 142, 777-787.

Hornig, N.C., Knowles, P.P., McDonald, N.Q., and Uhlmann, F. (2002). The dual mechanism of separase regulation by securin. Curr Biol 12, 973-982.

Hou, F.J., and Zou, H. (2005). Two human orthologues of Eco1/Ctf7 acetyltransferases are both required for proper sister-chromatid cohesion. Mol Biol Cell 16, 3908-3918.

Huang, H., Feng, J., Famulski, J., Rattner, J.B., Liu, S.T., Kao, G.D., Muschel, R., Chan, G.K., and Yen, T.J. (2007). Tripin/hSgo2 recruits MCAK to the inner centromere to correct defective kinetochore attachments. J Cell Biol 177, 413-424.

Huang, X., Andreu-Vieyra, C.V., York, J.P., Hatcher, R., Lu, T., Matzuk, M.M., and Zhang, P. (2008). Inhibitory phosphorylation of separase is essential for genome stability and viability of murine embryonic germ cells. Plos Biol 6, e15.

Huang, X., Hatcher, R., York, J.P., and Zhang, P. (2005). Securin and separase phosphorylation act redundantly to maintain sister chromatid cohesion in mammalian cells. Mol Biol Cell 16, 4725-4732.

in 't Veld, P.J.H., Herzog, F., Ladurner, R., Davidson, I.F., Piric, S., Kreidl, E., Bhaskara, V., Aebersold, R., and Peters, J.M. (2014). Characterization of a DNA exit gate in the human cohesin ring. Science 346, 968-972.

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

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

Ivanov, M.P., Ladurner, R., Poser, I., Beveridge, R., Rampler, E., Hudecz, O., Novatchkova, M., Heriche, J.K., Wutz, G., van der Lelij, P., et al. (2018). The replicative helicase MCM recruits cohesin acetyltransferase ESCO2 to mediate centromeric sister chromatid cohesion. Embo J 37.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr 66, 125-132.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature 467, 430-435.

Kawasumi, R., Abe, T., Arakawa, H., Garre, M., Hirota, K., and Branzei, D. (2017). ESCO1/2's roles in chromosome structure and interphase chromatin organization. Genes Dev 31, 2136-2150.

Kerrebrock, A.W., Miyazaki, W.Y., Birnby, D., and Orr-Weaver, T.L. (1992). The Drosophila mei-S332 gene promotes sister-chromatid cohesion in meiosis following kinetochore differentiation. Genetics 130, 827-841.

Kim, S.T., Xu, B., and Kastan, M.B. (2002). Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. Genes Dev 16, 560-570.

Kitajima, T.S., Kawashima, S.A., and Watanabe, Y. (2004). The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. Nature 427, 510-517.

Kitajima, T.S., Sakuno, T., Ishiguro, K., Iemura, S., Natsume, T., Kawashima, S.A., and Watanabe, Y. (2006). Shugoshin collaborates with protein phosphatase 2A to protect cohesin. Nature 441, 46-52.

Klein, F., Mahr, P., Galova, M., Buonomo, S.B., Michaelis, C., Nairz, K., and Nasmyth, K. (1999). A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. Cell 98, 91-103.

Koehler, C., Bonnet, J., Stierle, M., Romier, C., Devys, D., and Kieffer, B. (2014). DNA binding by Sgf11 protein affects histone H2B deubiquitination by Spt-Ada-Gcn5-acetyltransferase (SAGA). J Biol Chem 289, 8989-8999.

Kouznetsova, E., Kanno, T., Karlberg, T., Thorsell, A.G., Wisniewska, M., Kursula, P., Sjogren, C., and Schuler, H. (2016). Sister Chromatid Cohesion Establishment Factor ESCO1 Operates by Substrate-Assisted Catalysis. Structure 24, 789-796.

Krantz, I.D. (2014). Cohesin embraces new phenotypes. Nat Genet 46, 1157-1158.

Kumada, K., Yao, R., Kawaguchi, T., Karasawa, M., Hoshikawa, Y., Ichikawa, K., Sugitani, Y., Imoto, I., Inazawa, J., Sugawara, M., et al. (2006). The selective continued linkage of centromeres from mitosis to interphase in the absence of mammalian separase. J Cell Biol 172, 835-846.

Ladurner, R., Bhaskara, V., Huis in 't Veld, P.J., Davidson, I.F., Kreidl, E., Petzold, G., and Peters, J.M. (2014). Cohesin's ATPase activity couples cohesin loading onto DNA with Smc3 acetylation. Curr Biol 24, 2228-2237.

Lafont, A.L., Song, J., and Rankin, S. (2010). Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. Proc Natl Acad Sci U S A 107, 20364-20369.

Lee, H.J., Lang, P.T., Fortune, S.M., Sassetti, C.M., and Alber, T. (2012). Cyclic AMP regulation of protein lysine acetylation in Mycobacterium tuberculosis. Nat Struct Mol Biol 19, 811-818.

Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfner, K.P., Shirahige, K., and Uhlmann, F. (2006). Establishment of sister chromatid cohesion at the S. cerevisiae replication fork. Mol Cell 23, 787-799.

Liszczak, G., Arnesen, T., and Marmorstein, R. (2011). Structure of a ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. J Biol Chem 286, 37002-37010.

Liu, H., Rankin, S., and Yu, H. (2013). Phosphorylation-enabled binding of SGO1-PP2A to cohesin protects sororin and centromeric cohesion during mitosis. Nat Cell Biol 15, 40-49.

Liu, J., and Krantz, I.D. (2008). Cohesin and human disease. Annu Rev Genomics Hum Genet 9, 303-320.

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

Losada, A., Hirano, M., and Hirano, T. (2002). Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. Genes Dev 16, 3004-3016.

Majorek, K.A., Kuhn, M.L., Chruszcz, M., Anderson, W.F., and Minor, W. (2013). Structural, functional, and inhibition studies of a Gcn5-related N-acetyltransferase (GNAT) superfamily protein PA4794: a new C-terminal lysine protein acetyltransferase from pseudomonas aeruginosa. J Biol Chem 288, 30223-30235.

Maserati, E., Pasquali, F., Zuffardi, O., Buttitta, P., Cuoco, C., Defant, G., Gimelli, G., and Fraccaro, M. (1991). Roberts syndrome: phenotypic variation, cytogenetic definition and heterozygote detection. Ann Genet 34, 239-246.

McGuinness, B.E., Hirota, T., Kudo, N.R., Peters, J.M., and Nasmyth, K. (2005). Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. Plos Biol 3, 433-449.

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

Mendez, J., and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol Cell Biol 20, 8602-8612.

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

Minamino, M., Ishibashi, M., Nakato, R., Akiyama, K., Tanaka, H., Kato, Y., Negishi, L., Hirota, T., Sutani, T., Bando, M., et al. (2015). Esco1 Acetylates Cohesin via a Mechanism Different from That of Esco2. Current Biology 25, 1694-1706.

Minamino, M., Tei, S., Negishi, L., Kanemaki, M.T., Yoshimura, A., Sutani, T., Bando, M., and Shirahige, K. (2018). Temporal Regulation of ESCO2 Degradation by the MCM Complex, the CUL4-DDB1-VPRBP Complex, and the Anaphase-Promoting Complex. Curr Biol 28, 2665-2672 e2665.

Morales, C., and Losada, A. (2018). Establishing and dissolving cohesion during the vertebrate cell cycle. Curr Opin Cell Biol 52, 51-57.

Murayama, Y., and Uhlmann, F. (2014). Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature 505, 367-371.

Murayama, Y., and Uhlmann, F. (2015). DNA Entry into and Exit out of the Cohesin Ring by an Interlocking Gate Mechanism. Cell 163, 1628-1640.

Murray, A.W., and Szostak, J.W. (1985). Chromosome Segregation in Mitosis and Meiosis. Annu Rev Cell Biol 1, 289-315.

Nakajima, M., Kumada, K., Hatakeyama, K., Noda, T., Peters, J.M., and Hirota, T. (2007). The complete removal of cohesin from chromosome arms depends on separase. J Cell Sci 120, 4188-4196.

Nasmyth, K. (2009). How does cohesin hold sister DNAs together? Comp Biochem Phys A 153a, S41-S41.

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

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

Neuwald, A.F., and Landsman, D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. Trends Biochem Sci 22, 154-155.

Nishiyama, T., Ladurner, R., Schmitz, J., Kreidl, E., Schleiffer, A., Bhaskara, V., Bando, M., Shirahige, K., Hyman, A.A., Mechtler, K., et al. (2010). Sororin Mediates Sister Chromatid Cohesion by Antagonizing Wapl. Cell 143, 737-749.

Nishiyama, T., Sykora, M.M., Huis in 't Veld, P.J., Mechtler, K., and Peters, J.M. (2013). Aurora B and Cdk1 mediate Wapl activation and release of acetylated cohesin from chromosomes by phosphorylating Sororin. Proc Natl Acad Sci U S A 110, 13404-13409.

Ocampo-Hafalla, M.T., Katou, Y., Shirahige, K., and Uhlmann, F. (2007). Displacement and re-accumulation of centromeric cohesin during transient pre-anaphase centromere splitting. Chromosoma 116, 531-544.

Onn, I., Guacci, V., and Koshland, D.E. (2009). The zinc finger of Eco1 enhances its acetyltransferase activity during sister chromatid cohesion. Nucleic Acids Res 37, 6126-6134.

Ouyang, Z., and Yu, H. (2017). Releasing the cohesin ring: A rigid scaffold model for opening the DNA exit gate by Pds5 and Wapl. Bioessays 39.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., et al. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell 132, 422-433.

Peters, J.M. (2002). The anaphase-promoting complex: proteolysis in mitosis and beyond. Mol Cell 9, 931-943.

Peters, J.M., Tedeschi, A., and Schmitz, J. (2008). The cohesin complex and its roles in chromosome biology. Gene Dev 22, 3089-3114.

Poux, A.N., Cebrat, M., Kim, C.M., Cole, P.A., and Marmorstein, R. (2002). Structure of the GCN5 histone acetyltransferase bound to a bisubstrate inhibitor. Proc Natl Acad Sci U S A 99, 14065-14070.

Price, J.C., Pollock, L.M., Rudd, M.L., Fogoros, S.K., Mohamed, H., Hanigan, C.L., Le Gallo, M., Program, N.I.H.I.S.C.C.S., Zhang, S., Cruz, P., et al. (2014). Sequencing of candidate chromosome instability genes in endometrial cancers reveals somatic mutations in ESCO1, CHTF18, and MRE11A. PLoS One 8, e63313.

Rahman, S., Jones, M.J., and Jallepalli, P.V. (2015). Cohesin recruits the Esco1 acetyltransferase genome wide to repress transcription and promote cohesion in somatic cells. Proc Natl Acad Sci U S A 112, 11270-11275.

Remeseiro, S., Cuadrado, A., Gomez-Lopez, G., Pisano, D.G., and Losada, A. (2012). A unique role of cohesin-SA1 in gene regulation and development. Embo J 31, 2090-2102.

Remeseiro, S., Cuadrado, A., Kawauchi, S., Calof, A.L., Lander, A.D., and Losada, A. (2013). Reduction of Nipbl impairs cohesin loading locally and affects transcription but not cohesion-dependent functions in a mouse model of Cornelia de Lange Syndrome. Biochim Biophys Acta 1832, 2097-2102.

Riedel, C.G., Katis, V.L., Katou, Y., Mori, S., Itoh, T., Helmhart, W., Galova, M., Petronczki, M., Gregan, J., Cetin, B., et al. (2006). Protein phosphatase 2A protects centromeric sister chromatid cohesion during meiosis I. Nature 441, 53-61.

Rivera-Colon, Y., Maguire, A., Liszczak, G.P., Olia, A.S., and Marmorstein, R. (2016). Molecular Basis for Cohesin Acetylation by Establishment of Sister Chromatid Cohesion N-Acetyltransferase ESCO1. J Biol Chem 291, 26468-26477.

Roth, S.Y., Denu, J.M., and Allis, C.D. (2001). Histone acetyltransferases. Annu Rev Biochem 70, 81-120.

Salah Ud-Din, A.I., Tikhomirova, A., and Roujeinikova, A. (2016). Structure and Functional Diversity of GCN5-Related N-Acetyltransferases (GNAT). Int J Mol Sci 17.

Sjogren, C., and Nasmyth, K. (2001). Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. Curr Biol 11, 991-995.

Song, J., Lafont, A., Chen, J., Wu, F.M., Shirahige, K., and Rankin, S. (2012). Cohesin acetylation promotes sister chromatid cohesion only in association with the replication machinery. J Biol Chem 287, 34325-34336.

Stemmann, O., Zou, H., Gerber, S.A., Gygi, S.P., and Kirschner, M.W. (2001). Dual inhibition of sister chromatid separation at metaphase. Cell 107, 715-726.

Stigler, J., Camdere, G.O., Koshland, D.E., and Greene, E.C. (2016). Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. Cell Rep 15, 988-998.

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

Szyk, A., Deaconescu, A.M., Spector, J., Goodman, B., Valenstein, M.L., Ziolkowska, N.E., Kormendi, V., Grigorieff, N., and Roll-Mecak, A. (2014). Molecular basis for age-dependent microtubule acetylation by tubulin acetyltransferase. Cell 157, 1405-1415.

Tanaka, T., Fuchs, J., Loidl, J., and Nasmyth, K. (2000). Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. Nat Cell Biol 2, 492-499.

Tanner, K.G., Langer, M.R., and Denu, J.M. (2000a). Kinetic mechanism of human histone acetyltransferase P/CAF. Biochemistry 39, 15652.

Tanner, K.G., Langer, M.R., Kim, Y., and Denu, J.M. (2000b). Kinetic mechanism of the histone acetyltransferase GCN5 from yeast. J Biol Chem 275, 22048-22055.

Tanner, K.G., Trievel, R.C., Kuo, M.H., Howard, R.M., Berger, S.L., Allis, C.D., Marmorstein, R., and Denu, J.M. (1999). Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. J Biol Chem 274, 18157-18160.

Taschner, M., Vetter, M., and Lorentzen, E. (2012). Atomic resolution structure of human alpha-tubulin acetyltransferase bound to acetyl-CoA. Proc Natl Acad Sci U S A 109, 19649-19654.

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V., et al. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature 501, 564-+.

Toleman, C.A., Paterson, A.J., and Kudlow, J.E. (2006). The histone acetyltransferase NCOAT contains a zinc finger-like motif involved in substrate recognition. J Biol Chem 281, 3918-3925.

Trievel, R.C., Rojas, J.R., Sterner, D.E., Venkataramani, R.N., Wang, L., Zhou, J., Allis, C.D., Berger, S.L., and Marmorstein, R. (1999). Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. Proc Natl Acad Sci U S A 96, 8931-8936.

Trowitzsch, S., Bieniossek, C., Nie, Y., Garzoni, F., and Berger, I. (2010). New baculovirus expression tools for recombinant protein complex production. J Struct Biol 172, 45-54.

Uhlmann, F. (2016). SMC complexes: from DNA to chromosomes. Nat Rev Mol Cell Bio 17, 399-412.

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

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

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

Van Den Berg, D.J., and Francke, U. (1993). Roberts syndrome: a review of 100 cases and a new rating system for severity. Am J Med Genet 47, 1104-1123.

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

Vetting, M.W., LP, S.d.C., Yu, M., Hegde, S.S., Magnet, S., Roderick, S.L., and Blanchard, J.S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433, 212-226.

Waizenegger, I., Gimenez-Abian, J.F., Wernic, D., and Peters, J.M. (2002). Regulation of human separase by securin binding and autocleavage. Curr Biol 12, 1368-1378.

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

Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8, 127-134.

Wang, L., Tang, Y., Cole, P.A., and Marmorstein, R. (2008). Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. Curr Opin Struct Biol 18, 741-747.

Weitzer, S., Lehane, C., and Uhlmann, F. (2003). A model for ATP hydrolysis-dependent binding of cohesin to DNA. Curr Biol 13, 1930-1940.

Wendt, K.S., and Peters, J.M. (2009). How cohesin and CTCF cooperate in regulating gene expression. Chromosome Res 17, 201-214.

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

Whelan, G., Kreidl, E., Wutz, G., Egner, A., Peters, J.M., and Eichele, G. (2012). Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. Embo J 31, 71-82.

Wirth, K.G., Wutz, G., Kudo, N.R., Desdouets, C., Zetterberg, A., Taghybeeglu, S., Seznec, J., Ducos, G.M., Ricci, R., Firnberg, N., et al. (2006). Separase: a universal trigger for sister chromatid disjunction but not chromosome cycle progression. J Cell Biol 172, 847-860.

Yan, Y., Harper, S., Speicher, D.W., and Marmorstein, R. (2002). The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. Nat Struct Biol 9, 862-869.

Yang, X.J., and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene 26, 5310-5318.

Yu, H. (2016). Magic Acts with the Cohesin Ring. Mol Cell 61, 489-491.

Yuan, H., Rossetto, D., Mellert, H., Dang, W., Srinivasan, M., Johnson, J., Hodawadekar, S., Ding, E.C., Speicher, K., Abshiru, N., et al. (2012). MYST protein acetyltransferase activity requires active site lysine autoacetylation. Embo J 31, 58-70.

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

Zhang, N., and Pati, D. (2015). C-terminus of Sororin interacts with SA2 and regulates sister chromatid cohesion. Cell Cycle 14, 820-826.

Zhang, S., Li, J., Zhou, G., Mu, D., Yan, J., Xing, J., Yao, Z., Sheng, H., Li, D., Lv, C., et al. (2016). Increased expression of ESCO1 is correlated with poor patient survival and its role in human bladder cancer. Tumour Biol 37, 5165-5170.

List of Figures

List of Tables

Acknowledgements

Finally, I would like to thank my father who is not in this world anymore, but I feel his support and presence in every single moment of my life. I also thank my amazing mom for her eternal love and my family for their unconditional support.