FUNCTIONAL GENOMICS OF COHESIN ACETYLTRANSFERASES IN HUMAN CELLS

by

Sadia Rahman

A Dissertation

Presented to the Faculty of the Louis V. Gerstner, Jr.

Graduate School of Biomedical Sciences,

Memorial Sloan Kettering Cancer Center

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

New York, NY

March, 2014

_____ _____

Prasad V. Jallepalli, MD, PhD
Dissertation Mentor

Date

ABSTRACT

Accurate chromosome segregation during cell division requires that sister chromatids be physically linked from the time of their replication until their separation at anaphase. The cohesin complex, consisting of SMC1, SMC3, RAD21 and SCC3 arranges to form a ring-shaped structure that holds sister chromatids together. Acetylation of the cohesin SMC3 subunit by acetyltransferases ESCO1 and ESCO2 is essential for cohesion establishment. In addition to cohesion, cohesin also has roles in gene expression through its regulation of chromatin architecture.

Acetylation of cohesin by ESCO1/2 is regulated temporally and spatially. In human cells, it begins in G1 phase, rises in S-phase and persists until mitosis. The reaction occurs only on DNA-bound cohesin and SMC3 is quickly deacetylated after cohesin is removed from DNA. In this study, we map genome-wide ESCO1/2 and AcSMC3 sites by ChIP-Seq, study their regulation, and contribution to cohesion and gene expression functions.

Genome-wide mapping of ESCO1/2 reveals that they differ in their distribution: ESCO1 has many discrete binding sites that largely overlap with cohesin/CTCF sites, whereas ESCO2 has few sites of enrichment. A monoclonal antibody against the acetylated form of cohesin was also generated in this study to map cohesin acetylation, and this shows that cohesin is already acetylated in G1 at the majority of its sites and that this depends on ESCO1.

Identification of ESCO1 binding sites allowed study of the timing and recruitment to its sites by *cis* and *trans* elements. ESCO1 is targeted to cohesin/CTCF sites in G1, before the requirement for cohesion, in a cohesin-dependent manner. In addition, ESCO1 is found to target via several vertebrate-specific regions in the otherwise divergent N-

terminus. Deletion of several of these regions affects cohesion, indicating that its enrichment at these sites is important for ESCO1 function in cohesion.

In addition, *in vitro* binding experiments reveal that the N-terminus of ESCO1 interacts directly with SMC1A, PDS5B and CTCF. The targeting of ESCO1 in G1 and its colocalization and interaction with CTCF raised the question of whether it may function in gene expression. Microarray analysis reveals that many genes indeed have altered expression after ESCO1 loss. Furthermore, it is revealed that nearly all ESCO2 enrichment sites contain the REST/NRSF (RE1 silencing transcription factor/neural-restrictive silencing factor) binding motif and most nearby genes have functions in neuronal processes. ESCO2 as well as ESCO1 loss activates the expression of these genes. Therefore, a role in gene expression is demonstrated for both ESCO1 and ESCO2.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3C - chromosome conformation capture

ACT - acetyltransferase

APC - anaphase promoting complex

CdLS - Cornelia de Lange syndrome

ChIP - chromatin immunoprecipitation

ChIP-Seq - ChIP-sequencing

co-IP - co-immunoprecipitation

DAPI - 4′,6-diamidine-2-phenylindole dihydrochloride

DSB - double-strand break

DTT - dithiothreitol

ES cells - embryonic stem cells

FACS - fluorescence-activated cell sorting

FDR - false discovery rate

FRAP - fluorescence recovery after photobleaching

IF - immunofluorescence

IVT - *in vitro* transcription & translation

kb - kilobase

KD - knockdown

LOF - loss of function

mb - megabase

MEFS - mouse embryonic fibroblasts

PCH - pericentric heterochromatin

PCR - polymerase chain reaction

PIP - PCNA-interacting peptide

PMSF - phenylmethylsulfonyl fluoride

qPCR - quantitative polymerase chain reaction

RBS - Roberts–SC phocomelia

RT-qPCR - reverse transcription - quantitative polymerase chain reaction

WB - Western blot

WCE - whole cell extract

ZF - zinc finger

CHAPTER ONE

Introduction

Sister chromatid cohesion physically links pairs of sister chromatids from the time of their replication in S-phase until their separation at anaphase. Cohesion is required to ensure the accurate division of genetic content to daughter cells during cell division. Sister chromatids must be recognized as pairs when they are formed and tied together until a cell is ready to separate them at the metaphase-to-anaphase transition.

Central to this process is the cohesin complex which brings together four subunits that arrange to form a large ring-shaped structure proposed to encircle sister chromatids within the ring. Several accessory factors are also required for cohesion, including the acetyltransferases ESCO1 and ESCO2 in humans. These proteins belong to the Eco1-family of acetyltransferases that target two lysine residues on the cohesin subunit SMC3, an essential reaction for the process of cohesion establishment. Cohesion must then be maintained through G2 and into mitosis, and this requires the function of several other factors. Cohesin is then removed in vertebrate cells in two steps, first in prophase and then during anaphase. Cohesion dissolution at anaphase allows sisters to be pulled apart by opposing spindle forces.

Cohesin is not only involved in cohesion between sister chromatids, but also functions in DNA repair and gene expression. Defects in cohesin function lead to aneuploidy and cell death, and mutations in cohesin and related proteins including its acetyltransferase ESCO2 have been identified in human diseases.

The cohesin complex

Cohesin is a multi-subunit complex that associates with chromatin. It also associates with several regulatory proteins needed for its proper function throughout the cell cycle.

Cohesin structure

Cohesin is composed of four core subunits: SMC1A/B, SMC3, RAD21 (also known as SCC1 or MCD1), and SA1/SA2 (SCC3) (Figure 1.1) (Nasmyth and Haering 2005, Peters 2012). SMC1 and SMC3 are members of the SMC-family of proteins that are conserved from bacteria to humans and share structural and functional similarities (Losada et al., 1998). SMC-family proteins fold in half to form a hinge domain at one end that tightly dimerize to bring SMC1 and SMC3 together (Haering et al., 2002). The N- and C- termini of each protein join to create an ATPase head at the other end. The two head domains are joined by a third subunit called RAD21/SCC1, a kleisin-family protein (Schleiffer et al., 2003). The N-terminus of RAD21 binds to SMC3 and the C-terminus binds to SMC1. A fourth subunit consisting of a series of HEAT repeats, named SA1/2 in vertebrate cells and SCC3 in yeast, binds to RAD21 (Sumara et al., 2000, Haering et al., 2002).

SMC1 and SMC3 join to form a V-shaped structure as observed by electron microscopy, and their hinge and head domains are separated by a 45-nm long flexible coiled-coil domain (Anderson et al., 2002, Haering et al., 2002). SMC1, SMC3 and RAD21 subunits arrange to form a large tripartite ring of 35-nm diameter. Studies suggest that cohesin topologically embraces DNA inside this ring (Gruber et al., 2003, Ivanov and Nasmyth 2005).

Several models exist for how cohesin binds to DNA to tether sister chromatids together. In the simplest “embrace” model, a single complex topologically encircles and entraps both sisters, which are each 10-nm wide fibers, within the ring (Gruber et al., 2003,

Haering et al., 2008). This model is supported by several studies, for example using circular minichromosomes that release cohesin once linearized presumably because cohesin slides off the linear DNA. Also, evidence from experiments using engineered TEV sites at various surfaces that release cohesin from minichromosomes once cleaved, as well as covalent cross-linking experiments that show that the ring must be broken to allow at least one opening for cohesin dissociation, support the embrace model (Ivanov and Nasmyth 2005, Haering et al., 2008). An alternative oligomerization model, in which one cohesin complex encircles each sister chromatid and they interact with each other to bring sisters together, also remains a possible mode of cohesin-DNA interaction to generate cohesion (Huang et al., 2005).

DNA loading and unloading

Cohesin binds to DNA dynamically until a fraction becomes stably bound in S-phase. Fluorescence recovery after photobleaching (FRAP) assays on cohesin show that most cohesin is reversibly bound to DNA with a mean residence time of a half hour, until after S phase when a pool of cohesin becomes more stably bound with a residence time of several hours (Gerlich et al., 2006). This latter pool is thought to be the cohesin that participates in cohesion.

Loading of cohesin occurs on unreplicated DNA and therefore before the requirement for cohesion. Cohesin's association with DNA begins in late G1 phase in yeast, and even earlier, in telophase of the previous cell cycle, in vertebrates (Ciosk et al., 2000, Gillespie and Hirano 2004). Loading requires the SCC2/4 (yeast) or NIPBL/Mau2 (humans) complex (Ciosk et al., 2000, Seitan et al., 2006, Watrin et al., 2006). In *Xenopus* (but not in yeast), SCC2/4 association with DNA depends on formation of pre-replication complexes (pre-RCs) (Takahashi et al., 2004).

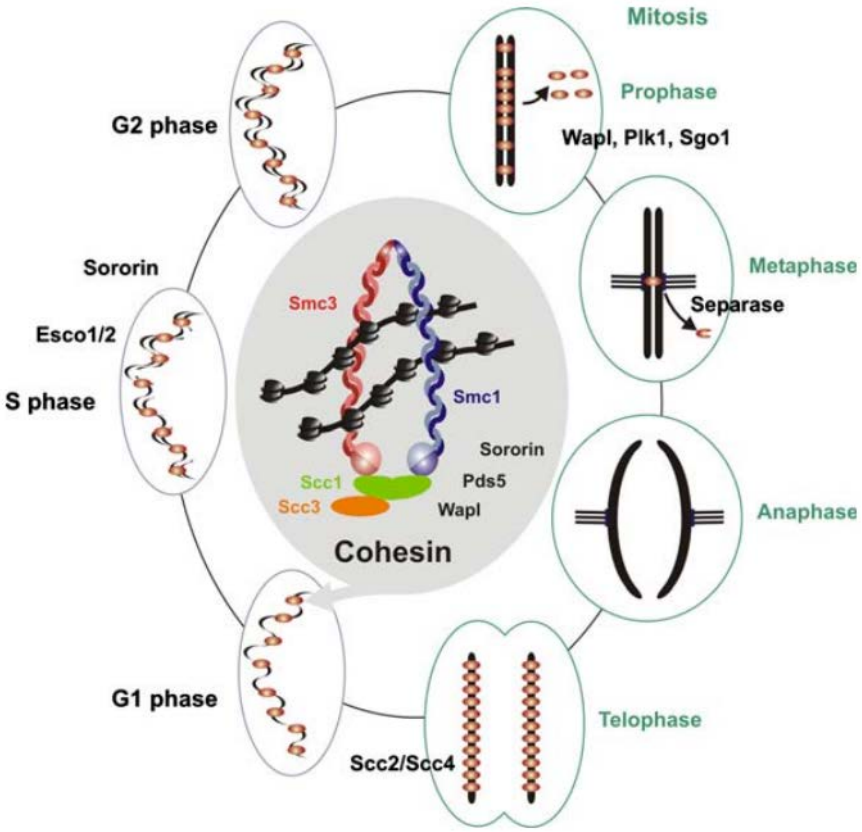


Figure 1.1. The cohesin cycle. The cohesin complex consists of four subunits, SMC1, SMC3, SCC1 (RAD21), and SCC3, which arrange to form a ring-shaped structure that has been proposed to topologically embrace both sister chromatids to link them together. Cohesin begins to associate with DNA in telophase when it is loaded by SCC2/4. Cohesion establishment is mediated by ESCO1/2 in S-phase and cohesion maintenance is regulated by vertebrate-specific Sororin. Cohesin removal occurs in two steps in human cells and involves WAPL and PLK1 in the prophase pathway to remove arm cohesin, while centromeric cohesin is protected by SGO1 and PP2A, and the remaining cohesin is removed by Separase through cleavage at anaphase onset. Cohesion dissolution at anaphase allows spindle forces to pull apart sister chromatids to opposite poles. Springer and Chromosome Research, 17(2), 2009, 201-14, How cohesin and CTCF cooperate in regulating gene expression., Wendt KS, Peters JM., Figure 1, with kind permission from Springer Science and Business Media.

DNA binding requires both the SMC1/3 hinge as well as ATPase activity at the head domains. The hinge and head domains can interact with each other in budding yeast (Mc Intyre et al., 2007). Mutation of SMC1/3 ATPase residues prevent cohesin binding to DNA (Arumugam et al., 2003). Opening of the hinge dimerization interface is also necessary for cohesin to load onto DNA, as experiments artificially linking the hinge domains by rapamycin-induced Frb-FKBP12 binding block DNA loading (Gruber et al., 2006, Nasmyth 2011). DNA exits is via a separate opening at the SMC3/RAD21 gate (Arumugam et al., 2003, Chan et al., 2012). Thus, DNA loading and unloading occurs through separate cohesin openings (Figure 1.2).

In yeast, cohesin is destroyed at anaphase when the RAD21 subunit is cleaved by separase/Esp1 at anaphase. The inhibitory chaperone of separase, securin/Pds1, is active for most of the cell cycle until all chromatid pairs have aligned at the metaphase plate, at which point the spindle checkpoint is turned off thereby activating the anaphase promoting complex (APC). The APC is a ubiquitin ligase that targets securin for destruction, activating the cysteine protease separase to cleave RAD21, which triggers anaphase by allowing the pulling apart of sister chromatids by opposing mitotic spindle forces (Ciosk et al., 1998, Uhlmann et al., 2000).

In contrast to yeast, cohesin in vertebrate mitotic cells is removed in two steps (Waizenegger et al., 2000). In the “prophase pathway”, the bulk of cohesin is removed from chromosome arms, without RAD21 cleavage. This removal depends on PLK1-mediated phosphorylation of SA2, and on WAPL (Sumara et al., 2002, Kueng et al., 2006). The prophase pathway allows removal of cohesin in the absence of its destruction to ensure that a large pool of cohesin is intact and ready to bind to DNA at the end of mitosis, thus allowing its quick re-loading in telophase in vertebrates.

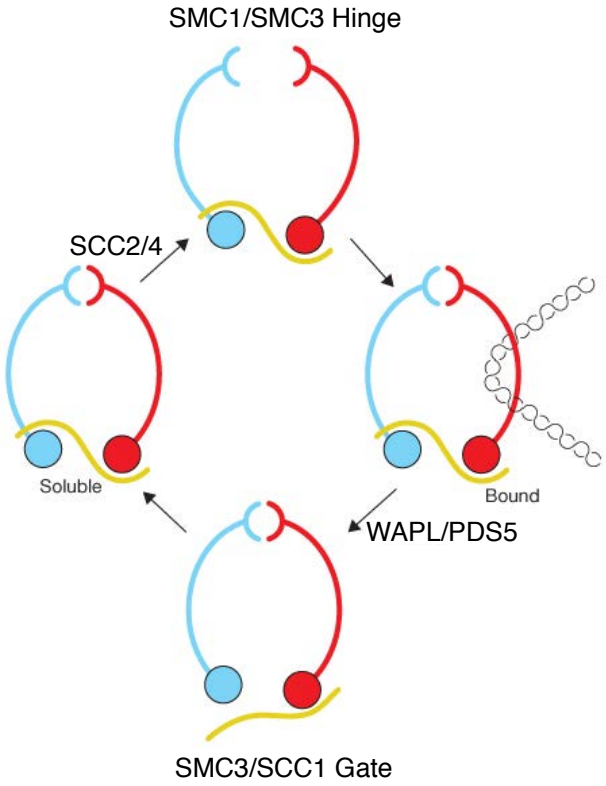


Figure 1.2. Cohesin has separate DNA entry and exit gates. Cohesin is loaded onto DNA by SCC2/SCC4 (or NIPBL/Mau2 in humans) through the opening of the SMC1/SMC3 hinge interface. Cohesin unloading is facilitated by WAPL/PDS5 and requires the opening of the SMC3/RAD21 interface, regulated by acetylation of the SMC3 head domain by ESCO1/2. Reprinted by permission from Macmillan Publishers Ltd: NATURE CELL BIOLOGY 13(10):1170-7, copyright 2011. http://www.nature.com/nc-b/journal/v13/n10/full/ncb2349.html

Centromeric cohesion is protected from the prophase pathway by the Shugoshin-PP2A complex, which dephosphorylates cohesin at centromeres (Kitajima et al., 2006). The remaining cohesin is then cleaved after APC activation by separase-mediated RAD21 cleavage as in yeast.

Cohesin-associated factors

In addition to the core cohesin subunits, several other proteins physically and functionally associate with cohesin. These include PDS5A/B, WAPL, and the vertebrate-specific protein Sororin.

PDS5 is a well-conserved HEAT-repeat containing protein (a motif that commonly interacts with other proteins and is also found in SCC2 and SCC3), that interacts genetically and physically with cohesin (Panizza et al., 2000, Losada et al., 2005). It appears to have complicated and context- as well as species- specific roles, contributing as both a positive and negative regulator of cohesion. PDS5 in *Saccharomyces cerevisiae* (budding yeast) is an essential protein required for cohesion (Panizza et al., 2000). In contrast, it is not essential in *Saccharomyces pombe* (fission yeast) and acts as an anti-establishment factor until counteracted by Eco1 homologue Eso1, but it stabilizes cohesion once established (Tanaka et al., 2001).

There are two homologues of PDS5 in vertebrate cells, PDS5A and PDS5B. These homologues are similar in their N-termini and consist of numerous HEAT repeats. The C-termini differ in sequence and PDS5B has two AT-hook domains, a domain that can bind AT-rich regions in the minor groove of DNA. These are lacking in PDS5A where there is one degenerate AT-hook domain (Zhang et al., 2009). An AT-hook domain is also found in another cohesin subunit, SA1, where it is required for SA1 association with DNA and function in telomeric cohesion (Bisht et al., 2013).

In *Xenopus*, depletion of PDS5A and PDS5B leads to partial defects in centromeric cohesion but an increase in cohesin stability on chromatin (Losada et al., 2005). PDS5A/B binding to DNA is dependent on cohesin in *Xenopus* and humans.

Mouse knockouts of PDS5A and PDS5B die before birth with defects reminiscent of cohesin-pathway human disease Cornelia de Lange Syndrome (CdLS, discussed below), including growth retardation, skeletal abnormalities, cleft palate and cardiac malformations (Zhang et al., 2007, Zhang et al., 2009). PDS5A-/- or PDS5B-/- MEFS do not show obvious cohesion defects. While both PDS5A and PDS5B are nuclear, PDS5B shows stronger nucleolus staining and PDS5A is more enriched in the rest of the nucleus. Deletion of both leads to very early embryonic lethality, and one allele of PDS5A/B leads to mid-gestational lethality in mice between E11.5 and E12.5 with no obvious cohesion defects. It was therefore hypothesized that one allele is sufficient for cohesion but leads to defects in development through gene expression abnormalities.

However, another mouse knockout study revealed centromeric cohesion defects in PDS5B-/-, and a defect in Sororin recruitment to pericentric heterochromatin (PCH) (Carretero et al., 2013). PDS5A and PDS5B double depletion by siRNA or deletion in conditional knockout MEFS result in an increase in cohesin bound to DNA. Both PDS5A and PDS5B contributed to telomere and arm cohesion by increase in telomere fragility and fragile site assays where there is a significant increase in fragility only when both homologues are depleted, but only PDS5B was shown to contribute to centromeric cohesion. A loss of characteristic ESCO2 and Sororin (cohesion establishment and maintenance factors; discussed below) staining at PCH is seen in PDS5B-/- MEFS, suggesting that PDS5B may contribute to centromeric cohesion by facilitating ESCO2-mediated acetylation of cohesin and subsequent Sororin binding at PCH.

Other factors associated with cohesin are WAPL (or Rad61 in yeast) and Sororin. WAPL is an anti-establishment factor that promotes cohesin removal from DNA . In yeast, Rad61 co-deletion can rescue lethality caused by cohesion establishment factor Eco1 deletion (Rolef Ben-Shahar et al., 2008, Sutani et al., 2009) indicating that Rad61 and Eco1 have counteracting effects. WAPL participates in the prophase pathway of cohesin removal from chromosome arms in mammalian cells (Gandhi et al., 2006). WAPL knockout mice are lethal, indicating that it is required during development (Tedeschi et al., 2013). Depletion of WAPL from HeLa cells leads to unresolved sister chromatids, or tighter binding along their lengths, at prophase, and increased cohesin binding to chromosomes. Loss of WAPL also increases cohesin residence time on chromatin by FRAP analysis (Tedeschi et al., 2013).

Sororin is a cohesion establishment and/or maintenance factor that is found specifically in vertebrates. It is required for cohesion but not for cohesin association with chromatin (Rankin et al., 2005, Schmitz et al., 2007), though Sororin binding to DNA does depend on cohesin binding, as well as on cohesin acetylation and DNA replication (Nishiyama et al., 2010). In *Xenopus*, Sororin is dispensable for cohesion in the absence of WAPL, as depletion of both mimics a Wapl depletion phenotype, i.e. unresolved sister chromatids in prophase.

WAPL and Sororin compete for binding to PDS5 in vertebrate cells (Nishiyama et al., 2010). Both contain FGF motifs that have been shown by mutational analyses to interact with PDS5 (Shintomi and Hirano 2009, Nishiyama et al., 2010). WAPL has three FGF motifs in its vertebrate-specific N-terminus and Sororin has one conserved motif. Mutation of the Sororin motif makes it incapable of promoting cohesion. WT Sororin, but not the FGF motif mutant, is capable of displacing WAPL from PDS5A when added to

Xenopus extracts, suggesting that these proteins compete for PDS5 binding via the FGF motifs. It has also been shown that Sororin binding to PDS5 is cell-cycle regulated by phosphorylation, as mitotic, phosphorylated Sororin is unable to displace WAPL from PDS5 (Nishiyama et al., 2010).

Cohesin localization

Cohesin associates with centromeres, telomeres, as well as along chromosome arms. It associates differently in different species. ChIP studies in yeast show that the binding pattern of cohesin differs from that of its loader, SCC2 (Lengronne et al., 2004). It has been proposed that cohesin may slide to its enrichment sites after its loading by the action of RNA polymerase II machinery, to sites of convergent transcription. In mammalian cells, however, cohesin is enriched at intergenic regions as well as within genes (Parelho et al., 2008, Wendt et al., 2008). A large number of these sites overlap with CCCTC-binding factor, or CTCF, a chromatin insulator protein. Cohesin interacts functionally with CTCF and other factors to regulate gene transcription, as discussed below.

ESCO1 and ESCO2

Cohesin loading onto DNA is not sufficient for cohesion between sister chromatids. Additional factors are required for cohesion. ESCO1 and ESCO2 are the human homologues of the yeast Eco1/Ctf7 acetyltransferase, and this family of proteins is essential for cohesion establishment, through acetylation of the cohesin SMC3 subunit in S-phase.

Chromatin localization and cell-cycle regulation

Eco1 is chromatin-associated throughout the cell cycle in budding yeast (Toth et al., 1999). Eco1 physically interacts with PCNA in *S. cerevisiae*, and this interaction has

been suggested to be required for its binding to chromatin (Moldovan et al., 2006). Mutation of the PCNA-interaction PIP motif in yeast leads to defects in cohesion. ESCO2 also tested positive for interaction with PCNA in vitro by yeast two-hybrid assay, suggesting that PCNA recruitment may be a conserved feature for chromatin localization of Eco1 homologues. However, the functional importance of the PCNA-interaction motif has not been shown aside from in yeast Eco1.

ESCO1 and ESCO2 in humans have been shown to be regulated differently throughout the cell cycle (Hou and Zou 2005). While ESCO1 levels are constant throughout, ESCO1 is phosphorylated and ESCO2 is degraded in mitosis, the latter reappearing in late G1/ or S-phase of the next cell cycle (Hou and Zou 2005). In yeast, Eco1 is a substrate of Cdk1, and its levels decrease when degraded in late S/G2 into mitosis, and this degradation is dependent on ubiquitination by SCF (Brands and Skibbens 2008, Lyons and Morgan 2011).

SMC3 acetylation

ESCO1/2 depletion does not affect overall cohesin levels bound to chromatin by fractionation (Hou and Zou 2005). Therefore, a defect in cohesion is not due to lower overall binding of cohesin to chromatin. Instead, it is the acetyltransferase activity that modifies the SMC3 subunit of cohesin to promote cohesion, a function of Eco1-family members that is conserved from yeast to humans.

Temperature-sensitive yeast Eco1 mutants arrested in different cell cycle stages revealed that Eco1 is essential only during S-phase, and not in G1, G2 or mitosis. Therefore, Eco1 is necessary for establishment of cohesion but not for cohesin loading onto DNA or for cohesion maintenance after its establishment (Skibbens et al., 1999, Toth et al., 1999).

The acetyltransferase activity of Eco1-family members is essential for their function in cohesion establishment (Rolef Ben-Shahar et al., 2008, Unal et al., 2008, Zhang et al., 2008). The critical substrate of these proteins for S-phase cohesion is SMC3. Two adjacent lysine residues are acetylated by Eco1 and ESCO1/2: K112,K113 in yeast and K105,K106 in humans. These residues are on the ATPase head domain of SMC3.

Lysine-to-arginine substitutions of these acetylated sites mimic Eco1 loss in yeast, causing a loss in viability and preventing cohesion (although there is normal cohesin association with DNA), and this is rescued by Wapl deletion (Rolef Ben-Shahar et al., 2008, Unal et al., 2008). A K113 lysine-to-asparagine mutation which is believed to mimic acetylation makes Eco1 dispensable. In humans, expression of unacetylatable mutants of SMC3 (K105A,K106A) in 293T cells leads to a loss of cohesion (Zhang et al., 2008). The consequence of Eco1 family proteins and of AcSMC3 is to destabilize the WAPL-cohesin interaction and therefore counteract WAPL's removal activity on cohesin (Sutani et al., 2009, Terret et al., 2009). Experiments in which yeast Smc3 is fused to Rad21, therefore blocking cohesin unloading from DNA, allows cohesion establishment in cells lacking Eco1 (Chan et al., 2012), demonstrating that the function of Eco1 acetylation is to stabilize the cohesin ring on DNA by preventing DNA exit.

ESCO1/2 activity and SMC3 acetylation are regulated such that only chromatin-associated SMC3 is acetylated. It was observed that ESCO1 could not acetylate SMC3 when incubated with purified recombinant cohesin, or unless chromatin was added in Xenopus extracts (Nishiyama et al., 2010).

Acetylated cohesin is only observed on chromatin and not in the soluble fraction following fractionation. This is because SMC3 is quickly deacetylated when removed

from chromatin, by Hos1 in yeast and its homologue HDAC8 in humans (Beckouet et al., 2010, Borges et al., 2010, Deardorff et al., 2012). In yeast, Hos1 deletion increases levels of acetylated Smc3 in the chromatin fraction and also leads to an appearance of acetylated Smc3 in the soluble fraction. Cell cycle synchronization experiments show that this increase in the *hos1*Δ background occurs after S-phase and that Smc3 co-immunoprecipitates with Hos1 after S-phase, with the interaction rising as cells progress to anaphase (Xiong et al., 2010).

Hos1 and Eco1 levels do not regulate the acetylation dynamics of Smc3 (Borges et al., 2010). Hos1 levels are constant and it does not appear to be post-translationally modified throughout the cell cycle. The regulation of Smc3 acetylation in yeast is also independent of Eco1 levels or its acetyltransferase activity. Eco1 levels rise in G1/S when acetylation begins but decline before binucleate formation when Smc3 acetylation is reduced, and overexpression of Eco1 throughout the cell cycle does not change the kinetics of acetylation and deacetylation. Acetyltransferase activity of Eco1 also does not coincide with Smc3 acetylation levels as suggested by its auto-acetylation activity after purification from synchronized cell populations. However, the localization of Smc3, whether it's on or off chromatin, does affect its acetylation status. Uncleavable Scc1/ Rad21 that does not allow for cohesin dissociation from DNA causes a persistence of acetylated Smc3. Conversely, early cleavage of Scc1/Rad21 in metaphase, with an engineered Scc1/Rad21 TEV cleavage site and expression of TEV protease, leads to early reduction in acetylation.

Loss of unacetylatable Smc3 for use in the next cell cycle is what causes the cohesion defect observed in *hos1*Δ cells, as Smc3 must be newly acetylated in S-phase for it to establish cohesion (Borges et al., 2010). In human cells, HDAC8 is found exclusively in the soluble fraction and undetectable in the chromatin fraction, and loss of HDAC8

activity leads to increased AcSMC3 throughout the cell cycle, suggesting that it also targets SMC3 after removal from chromatin and is active throughout the cell cycle, as is true for Hos1 in yeast (Deardorff et al., 2012). WAPL depletion also causes an increase in AcSMC3, presumably because of the prevention of cohesin removal from DNA by blocking the prophase pathway. Loss of HDAC8 activity appears to cause an increase (albeit a slight one) in RAD21 cleavage fragments associated with the cohesin complex on chromatin. It is also reported to lead to a change in the distribution of cohesin peaks on DNA (Deardorff et al., 2012). Therefore, as in yeast, the SMC3 deacetylase targets only the soluble pool of cohesin after its removal from DNA and is required to recycling a fresh pool of cohesin for reloading onto DNA for the next cell cycle.

Cohesin-associated factors affect the level of its acetylation. Sororin depletion decreases SMC3 acetylation and WAPL depletion greatly increases SMC3 acetylation in vertebrates (Nishiyama et al., 2010). This may be because cohesin is removed from DNA and therefore targeted by its deacetylase or remains on DNA and protected from the deacetylase, respectively, in these situations. Interestingly, Pds5, which has complicated roles in cohesion but has been thought of generally as an anti-establishment factor, is needed for normal levels of cohesin acetylation in both *S. cerevisiae* and *S. pombe* (Vaur et al., 2012, Chan et al., 2013). There is a decrease in acetylation in Pds5 mutants in both species, and this is independent of Wapl activity. Pds5 binds to Scc1/Rad21 through conserved residues adjacent to Scc1/Rad21's Smc3-interacting region and there is a decrease in acetylation when this interaction is abolished or when Pds5 is mutated (Chan et al., 2013). It was further shown that Pds5 in yeast is involved in both promoting *de novo* acetylation of cohesin in S-phase as well as preventing its deacetylation by Hos1 to maintain cohesion. Relevant to these findings, a yeast two-hybrid interaction was reported between *S. pombe* Eso1 and Pds5 (Tanaka et al., 2001), but this interaction has not been investigated further or *in vivo*.

The promotion of acetylation by PDS5 is also seen in mouse. Loss of both PDS5 homologues leads to a strong reduction in cohesin acetylation in mouse MEFS, as well as a decrease in Sororin binding to chromatin (Carretero et al., 2013).

Link to replication

Several findings point to a link between DNA replication and cohesion establishment. First, sister DNA molecules are immediately linked from the time of their replication, suggesting that they are linked as they emerge from the fork perhaps by cohesin molecules that are already bound to DNA, in addition to concatenations that arise due to replication. In addition, Eco1 functions during S-phase and cannot establish cohesion outside of this stage (Toth et al., 1999). Eco1 was shown to be enriched, though poorly, in a ChIP study to replication forks by colocalization with BrDU (Lengronne et al., 2006), although this enrichment was very weak .

Moreover, Eco1 interacts with replication factors, and several replication proteins when mutated in yeast or depleted in human cells, lead to cohesion defects. Yeast Eco1 genetically interacts with several replication factors. Among these are PCNA (Pol30 gene in yeast), over-expression of which rescues Eco1 temperature-sensitivity (Skibbens et al., 1999). Other factors include Ctf4, Ctf18 and several other RFC complex subunits, polymerase Pol2, checkpoint proteins Tof1 and Csm3, polymerase kappa encoded by TRF4 and DNA helicase Chl1 (Skibbens et al., 1999, Hanna et al., 2001, Edwards et al., 2003, Kenna and Skibbens 2003, Skibbens 2004). Interestingly, Ctf4 and Chl1 yeast mutants have more severe cohesion phenotypes than other replication proteins, and they appear to function independently of and in parallel to the Eco1-pathway, as deletion of Wapl does not relieve their cohesion defect (Borges et al., 2013). In contrast, Wapl removal does relieve the cohesion defects in other replication protein mutants.

The fact that Eco1 has a S-phase specific function and physically and genetically interacts with all of these replication proteins suggests that it is recruited to the replisome where its activity is required for cohesion between the emerging sister chromatids. Therefore, models of Eco1 function at the replication fork have been proposed in yeast, in which Eco1 travels with the fork to establish cohesion by either affecting fork geometry to allow passage through cohesin rings or to enable transient dissociation and reassociation of cohesin during fork passage (Figure 1.3) (Lengronne et al., 2006).

In human cells, knockdown or loss of some of the cohesion-pathway replication factors also leads to cohesion defects (Farina et al., 2008, Terret et al., 2009). Furthermore, deletion of an RFC-CTF18 complex subunit, DCC1, leads to a defect in fork progression, as measured by slower replication rates by DNA fiber analysis, which is also observed in Roberts Syndrome patient cells. This phenotype is recapitulated with ESCO1 or ESCO2 depletion in HeLa cells, where it is rescued by WAPL or PDS5 co-depletion (Figure 1.4) (Terret et al., 2009). This demonstrates that SMC3 acetylation and the same network of proteins are involved in both cohesion and normal fork progression, supporting the link between cohesion and replication and a role for ESCO1/2 at the fork. This impediment of cohesin to replication fork progression is not observed in yeast (Lopez-Serra et al., 2013), suggesting that it is a feature of cohesin pathway proteins that has evolved in higher organisms.

Rad21 acetylation and DNA repair

Repair of double-strand breaks (DSBs) in mitotic cells requires cohesion in order to use the sister chromatid as a template (Sjogren and Nasmyth 2001). This process involves the recruitment of cohesin to sites of DSB and establishment of cohesion in G2 cells, therefore independently of replication (Strom et al., 2004). Eco1 is also required for this

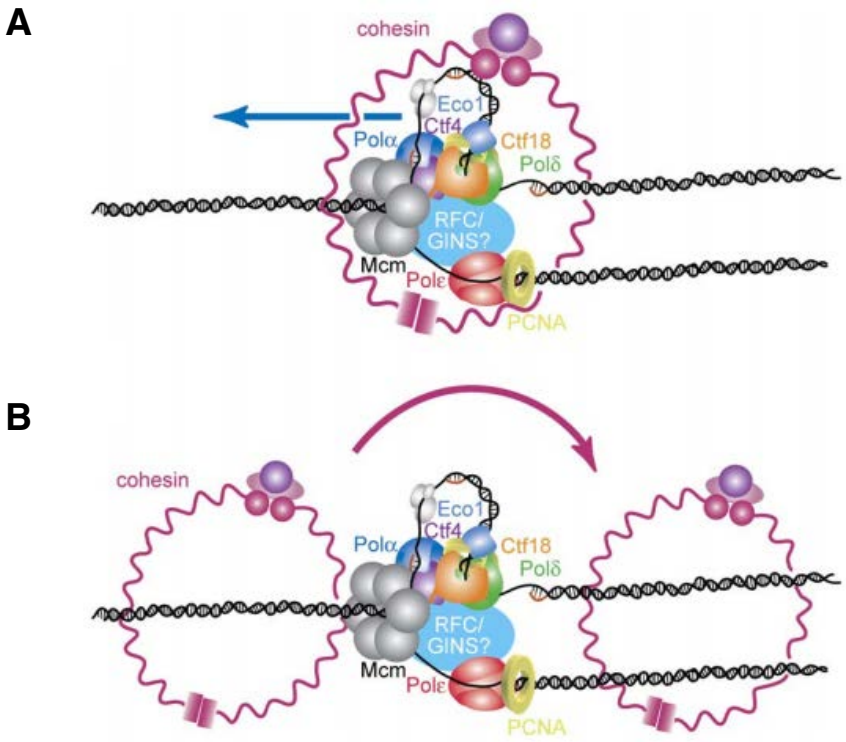


Figure 1.3. Models of Eco1 function at replication forks. Eco1 has been proposed to localize to and travel with replication forks, by interaction with PCNA and/or other replisome components that are important for cohesion. The replisome may slide through cohesin rings (A), in which case the function of Eco1 may be a role in altering replisome conformation to allow it to pass through cohesin. Alternatively, cohesin may briefly dissociate from DNA to allow fork passage and reassociate behind the fork (B). Eco1 and other establishment factors may be involved in this transient dissociation and reassociation of cohesin. Reprinted from Molecular Cell, 23(6), Lengronne A, McIntyre J, Katou Y, et al., Establishment of sister chromatid cohesion at the S. cerevisiae replication fork., 787-99, Copyright 2006, with permission from Elsevier.

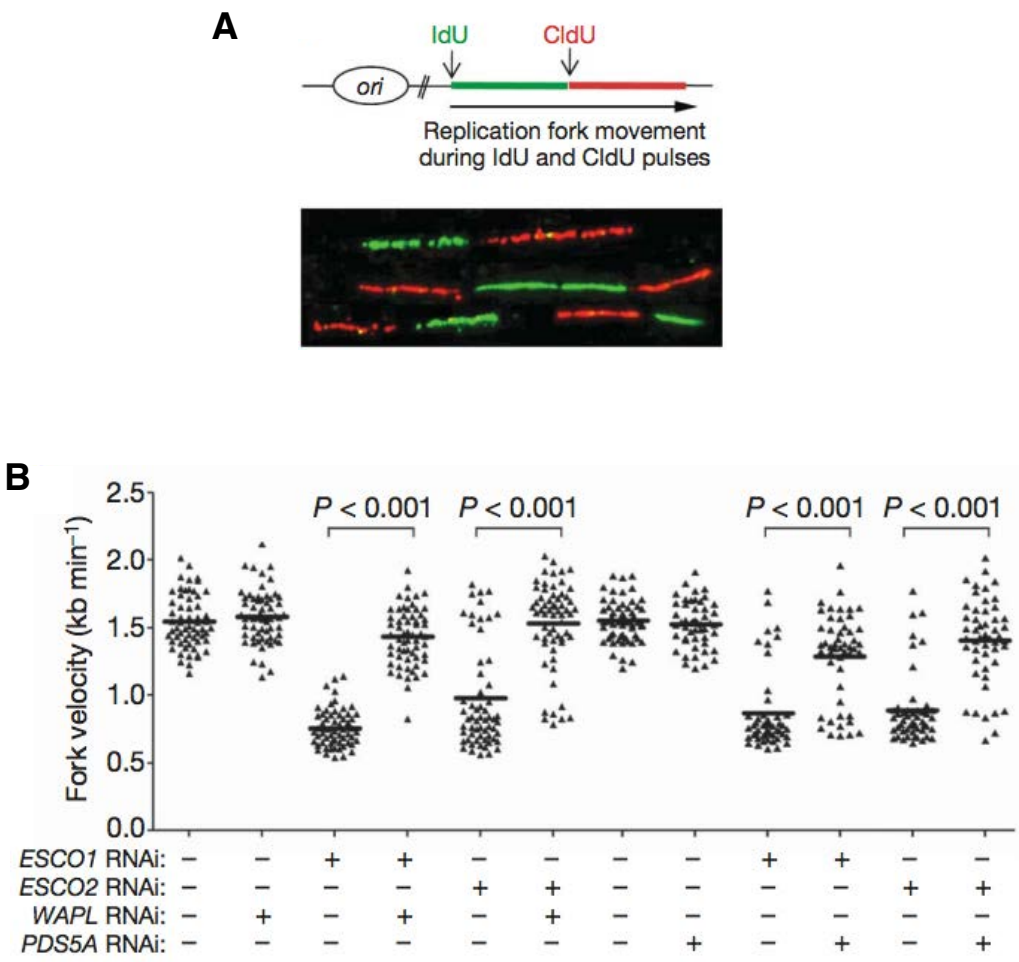


Figure 1.4. Loss of cohesin acetyltransferases in human cells causes a replication fork progression phenotype that is rescued by WAPL/PDS5 loss. (A) Replication fork dynamics are measured by labeling cells with pulses of two thymidine analogues, IdU and CldU, detected with specific antibodies to measure their track lengths. (B) ESCO1- and ESCO2- depleted cells have slower replication fork rates. Co-depletion of WAPL or PDS5A reverses this phenotype. Reprinted from Terret ME, Sherwood R, Rahman S, Qin J, Jallepalli PV., Cohesin acetylation speeds the replication fork., Nature. 2009 Nov 12;462(7270):231-4. http://www.nature.com/nature/journal/v462/n7270/full/nature08550.html

process in yeast, where it mediates genome-wide cohesion in G2 cells exposed to DSBs (Strom et al., 2007, Unal et al., 2007). This requires acetylation of the Scc1/Rad21 subunit of cohesin rather than Smc3.

DSBs induce checkpoint kinases to phosphorylate Scc1/Rad21 in yeast which in turn leads to its acetylation by Eco1, thereby counteracting Wapl activity to establish cohesion in G2/M. Mutation of Scc1/Rad21 acetylation sites cannot establish cohesion in response to DNA damage in G2/M, despite the presence of WT Smc3, and loss of Smc3 acetylation sites with WT Scc1/Rad21 cannot establish S-phase cohesion (Heidinger-Pauli et al., 2009). Therefore, at least in yeast, Eco1 can establish cohesion both in S-phase and in response to DSBs in G2/M, by targeting distinct cohesin subunits.

ESCO1/2 domains

Yeast Eco1 contains a PIP-box PCNA binding motif, C2H2 zinc finger, and acetyltransferase domains (Ivanov et al., 2002, Moldovan et al., 2006).

The two human homologues of Eco1, ESCO1 and ESCO2, both share the PIP box motif, zinc finger domains, and acetyltransferase domains of their yeast counterpart (Hou and Zou 2005, Moldovan et al., 2006). These domains are also highly conserved in all Eco1 homologues including *S. pombe* Eso1, *Xenopus* XEco1 and XEco2, Drosophila deco, and mouse ESCO1 and ESCO2 (Tanaka et al., 2000, Williams et al., 2003, Takagi et al., 2008). The fission yeast homologue Eso1 contains an extended N-terminus that is a fusion of Rad30 translesion DNA polymerase eta and is dispensable for cohesion (Tanaka et al., 2000). ESCO1 and ESCO2 are 77% similar and 59% identical in their C-terminal Eco1 domains (Hou and Zou 2005) and there is no similarity in their N-termini. This domain organization of Eco1 homologues is illustrated in Figure 1.5.

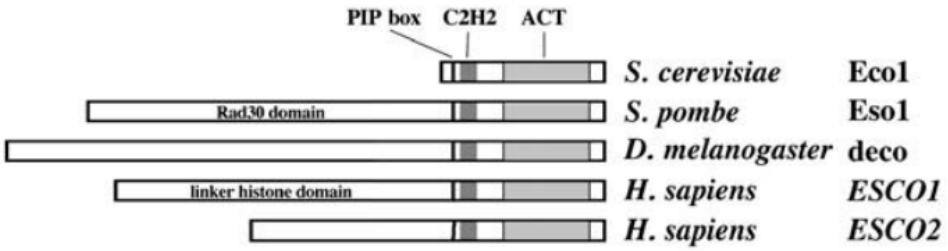


Figure 1.5. Conservation of domains in Eco1-family proteins. Diagram shows the conservation of C-terminal PIP box, C2H2 Zinc Finger and acetyltransferase (ACT) domains. All Eco1 homologues have an extended N-terminus that is highly divergent between homologues and is lacking in *S. cerevisiae* Eco1. The *S. pombe* Eso1 N-terminus is a translesion polymerase and is completely dispensable for its cohesion establishment function. Humans have two orthologues, ESCO1 and ESCO2, that also share conservation in the C-terminus but have divergent N-termini. Reprinted from Molecular Cell, 23(5), Moldovan GL, Pfander B, Jentsch S., PCNA controls establishment of sister chromatid cohesion during S phase., 723-32, Copyright 2006, with permission from Elsevier.

A study on yeast zinc finger mutation showed that it may be required for normal levels of acetylation and for cohesion and cell viability *(Onn et al., 2009)*. In human cells expressing a mutant ESCO1 with a substitution of a zinc finger residue, there was no affect on DNA localization *(Hou and Zou 2005)*. The importance of the zinc finger in human cells has not been studied further.

The human homologues are partially but not fully redundant in establishing cohesion; depletion of either in HeLa cells leads to a loss of cohesion and co-depletion of both causes a more severe phenotype (Hou and Zou 2005). Loss of the acetyltransferases does not however affect cohesin association with chromatin. About 70% of each protein associates with chromatin by fractionation. Interestingly, both ESCO1 and ESCO2 were shown to bind to chromosomes via their divergent N-terminal domains: the conserved C-terminus alone was not sufficient to bind chromatin whereas the N-terminus was necessary and sufficient for binding.

Cohesin Functions

Cohesin has essential functions in cohesion between sister chromatids and DNA repair. In addition to these roles, cohesin is important in regulating gene expression by controlling chromatin architecture.

Cohesion

Cohesin is key to the cohesion between sister chromatids from their generation in S-phase until their separation in anaphase. Cohesion is generated in S-phase by Eco1-family acetyltransferases and this cohesion must be maintained through G2/M until all chromosomes are bioriented on the metaphase plate. Eco1 is the only essential establishment factor required for cohesion and its acetylation of the SMC3 subunit of cohesin is necessary for S-phase cohesion, as discussed above. Loss of cohesin and

cohesion establishment factors leads to activation of the checkpoint and cell cycle delays, premature separation of sister chromatids, aneuploidy and apoptosis.

DNA repair

Cohesin and Eco1 roles in cohesion are also important for DNA repair, in order to provide a template for the damaged strands in the undamaged sister chromatid, for repair by homologous recombination (Strom et al., 2007, Unal et al., 2007). If Eco1 is inactivated in S-phase, its activity is normally limiting in G2/M so that it is unable to generate cohesion at this stage. Overexpression of Eco1 is required to generate cohesion in G2/M in the absence of DNA damage. However, damage can also activate endogenous levels of Eco1 to establish genome-wide cohesion in G2/M. Cohesin can be loaded *de novo* in G2/M in response to DSBs and Eco1 acetylates Scc1/Rad21 to mediate G2/M damage-induced cohesion (Heidinger-Pauli et al., 2008).

Gene expression

Cohesin is expressed in post-mitotic neurons, which suggests non-cohesion functions of cohesin. Indeed, several studies in zebrafish, *Drosophila* and mammalian cells have revealed a role for cohesin in regulating gene expression.

In yeast, cohesin is localized at AT-rich regions both in the centromeres and pericentric regions, and along arms at transcriptional termination sites (Glynn et al., 2004).Evidence for cohesin regulation of gene expression in budding yeast comes from a study showing yeast Scc2 involvement in the activation of Rec8 (meiotic kleisin) expression in meiosis through cohesin loading (Lin et al., 2011). Cohesin has also been shown to block read-through transcription in *S. pombe* (Gullerova and Proudfoot 2008).

In contrast to the localization of cohesin outside of transcribed regions in yeast, *Drosophila* ChIP-chip studies have revealed that cohesin co-localizes with its loader Nipped-B (SCC2/NIPBL) and is localized within active genes, where tissue-specific transcriptional changes are observed between different cell types in genes with varying levels of cohesin at nearby sites (Misulovin et al., 2008). Mutations of Nipped-B and PDS5 or decreased SCC3 or SMC1 in *Drosophila* affects expression of the *cut* gene, which has a cohesin binding site between remote promoter and enhancer regions located tens of kilo-bases away (Rollins et al., 1999, Rollins et al., 2004, Dorsett et al., 2005). In addition, in post-mitotic *Drosophila* neurons, loss of cohesin through SMC1, RAD21 or SCC3 mutations or inactivation is lethal and causes defects in axon pruning by gene expression changes in EcR receptor, a key regulator of pruning. Ectopic expression of EcR rescues the pruning phenotype (Pauli et al., 2008, Schuldiner et al., 2008). These studies in post-mitotic neurons, where cohesion between sister chromatids is not a factor, show that the gene regulation function of cohesin is independent of its role in sister chromatid cohesion.

Cohesin sites have also been mapped by ChIP-chip studies in mouse and human cells (Parelho et al., 2008, Rubio et al., 2008, Wendt et al., 2008). This has revealed cohesin enrichment at thousands of sites that co-localize with the transcriptional insulator protein CTCF.

CTCF is a vertebrate-specific highly conserved zinc finger protein (Filippova et al., 1996) that binds to insulators and prevents enhancer-promoter interactions, for example at the Igf2/H19 imprinted locus where it binds allele-specifically in a methylation-sensitive manner (Figure 1.6) (Bell and Felsenfeld 2000). CTCF also acts as a barrier element, binding to regions between active and repressive chromatin marks and preventing the spread of heterochromatin (Cuddapah et al., 2009). Genome-wide mapping of CTCF

binding has revealed tens of thousands of sites, 75% of which have the consensus motif CCGCGNGGNGGCAG. It functions as both an activator and repressor, possibly by different binding partners at different sites and/or because of post-translational modifications (Lee and Iyer 2012).

In one study, 89% of cohesin sites overlap with CTCF sites, and about half of CTCF sites are bound by cohesin (Wendt et al., 2008). The association of cohesin with chromatin is unaffected when CTCF is depleted, but enrichment at its discrete sites is reduced. Wendt et al. also reported a 50% decrease in CTCF occupancy after cohesin depletion, demonstrating that their enrichment at these sites is inter-dependent.

Chromosome conformation capture (3C) assays have shown cohesin functionally interacts with CTCF to regulate gene expression by forming chromatin loops. This has been demonstrated at several loci, including the IGF2/H19 imprinted locus (Figure 1.6), as well as at the β-globin and protocadherin loci (Nativio et al., 2009, Chien et al., 2011, Guo et al., 2012). In addition, cohesin as well as its loader NIPBL, also interact with the mediator complex, a transcriptional coactivator, at thousands of sites in mouse ES cells. DNA looping is observed by 3C between promoters and enhancers at Nanog, Oct4 and other genes where they bind and this interaction is cell-type specific, observed in ES cells but not MEFS (Kagey et al., 2010).

Recent studies point to a role in gene expression for other cohesion pathway proteins, namely HDAC8 and WAPL (Deardorff et al., 2012, Tedeschi et al., 2013). Mutant HDAC8-mutant patient CdLS lines exhibit gene expression changes that classify with CdLS cells containing mutations in NIPBL (Deardorff et al., 2007). Also, microarray analysis in WAPL-KO MEFS show altered gene expression profiles compared to WT

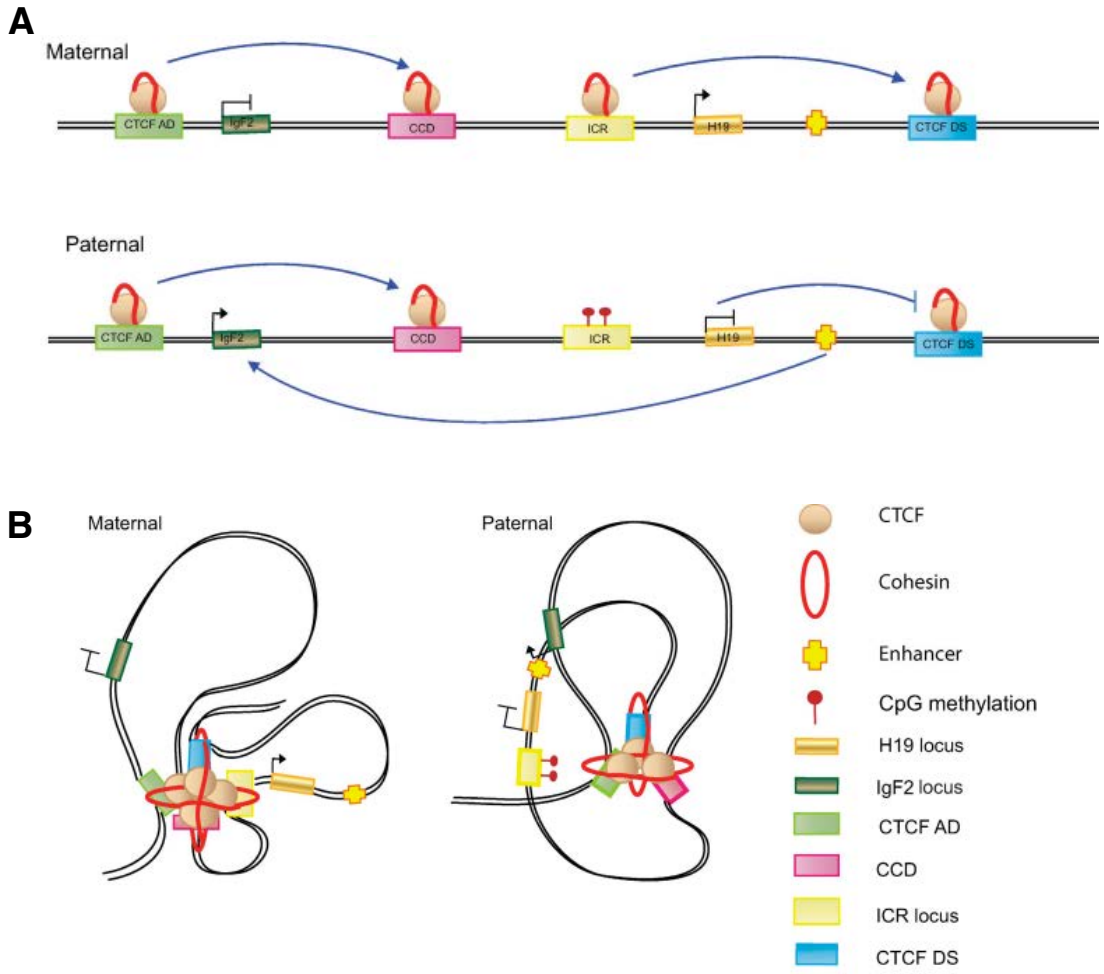


Figure 1.6. Cohesin and CTCF functionally cooperate in DNA looping events to control gene expression at the imprinted Igf2/H19 locus. (A) Organization of Igf2/H19 maternal and paternal loci with interactions between regions depicted by blue arrows. The imprinting control region (ICR) is differentially methylated in the two alleles. Lack of methylation on the maternal allele allows CTCF-cohesin binding and interaction with a downstream CTCF (CTCF DS) site to allow H19 transcription. On the paternal allele, ICR methylation blocks CTCF-cohesin binding and instead an enhancer interacts with Igf2 to allow its transcription. CTCF AD (CTCF-binding site) and CCD (centrally conserved DNase I hypersensitive site) interactions occur on both alleles. (B) 3C assays have revealed different looping events at maternal and paternal alleles. ICR methylation on the paternal allele causes CTCF DS to remain out of the loop, leading to enhancer activation of Igf2. The unmethylated ICR on the maternal allele leads to an interaction between ICR and CTCF DS, which prevents Igf2 activation by the enhancer. Originally published in J CELL BIOL. doi: 10.1083/jcb.200912129.

cells (Tedeschi et al., 2013). These changes appear to be due to a difference in cohesin occupancy at its binding sites near genes.

Cohesion-pathway mutations in human diseases

In humans, cohesin and related genes are mutated in diseases known as cohesinopathies. These include Cornelia de Lange Syndrome (CdLS), caused by mutations in several cohesion-pathway components, and Roberts Syndrome/SC phocomelia (RBS), caused by ESCO2 mutations.

Developmental defects including craniofacial abnormalities, limb defects, microcephaly, and mental retardation are characteristic CdLS phenotypes (Dorsett and Krantz 2009). About half of CdLS patients have heterozygous mutations in NIPBL, ~5% have hemizygous or heterozygous mutations in the X-linked SMC1A, and one patient was found with a mutation in SMC3. Additionally, there have been several patients with HDAC8 mutations and one patient with PDS5B missense mutation in an AT-hook domain in the C-terminus (Zhang et al., 2009, Deardorff et al., 2012).

RBS is a separate disease that is characterized by some CdLS-distinct features, as well as some CdLS-overlapping features such as slow growth, limb defects, microcephaly and mental retardation (Dorsett 2007). Unlike CdLS, it is a recessive disorder where ESCO2 is biallelically mutated, and one study reported eight RBS mutations in ESCO2 including one missense mutation (W539G) in a highly conserved residue of the ACT domain, a nonsense mutation in the N-terminus and six frameshift mutations, all of which truncate the protein upstream of the ACT domain. Therefore, the acetyltransferase activity is abrogated in all cases. RBS patient cells are also sensitive to some DNA-damaging agents (van der Lelij et al., 2009).

Although cohesion pathway proteins are mutated in CdLS, patient cells surprisingly do not display obvious cohesion defects or cell cycle delays. Instead, altered gene expression is believed to lead to CdLS phenotypes.

Heterochromatic repulsion is characteristic of RBS cells (Schule et al., 2005, Vega et al., 2005). Cell cycle delays, increased micronuclei formation, and apoptosis have also been reported in RBS cells and/or animal models, (Monnich et al., 2011, Morita et al., 2012), and these are not observed in CdLS. Thus, it has been proposed that RBS causes phenotypes through its heterochromatic repulsion defect and apoptosis, and not through gene expression alterations as in CdLS.

Models of RBS have been generated in zebrafish and mouse. Zebrafish depletion of ESCO2 results in some RBS-mirrored phenotypes such as fin reduction (analogous to limb reduction) and craniofacial abnormalities (Monnich et al., 2011). There is also high apoptosis in the brain and peripheral nervous system. ESCO2 morphants do not show strong gene misregulation and it was concluded that their phenotypes are due to some mis-regulated genes that are involved in cell cycle and apoptosis, rather than altered gene expression (Monnich et al., 2011).

A mouse ESCO2 knockout (Whelan et al., 2012) also provides some insight into RBS phenotypes. ESCO2-KO is lethal in the pre-implantation stages prior to the eight-cell stage with loss of centromeric cohesion, and lagging chromosomes were observed in the second mitosis of embryos and in KO MEFS. Deletion of ESCO2 conditionally in the cerebral cortex using Emx1-induced Cre recombinase leads to microcephaly in which most of the cortex fails to develop (Figure 1.7), which is relevant to findings in this study (Chapter Five).

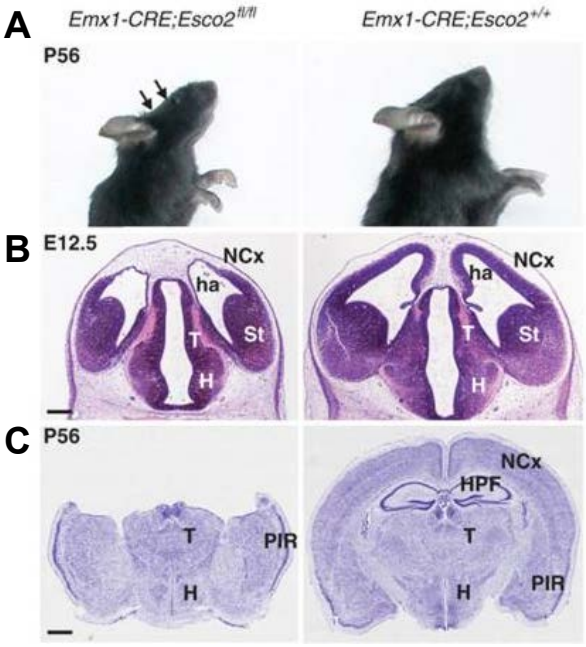


Figure 1.7. ESCO2-deficiency in cortical progenitors leads to brain developmental defects. (A) Flattened forehead (arrows) due to severe microcephaly is observed in Emx1-CRE;Esco2fl/fl mouse. Emx1 is a homeobox transcription factor expressed in the developing cerebrum. (B) Developmental failure of hippocampal and neocortical primordia is observed in ESCO2-deficient forebrain. (C) Section from adult mouse shows lack of hippocampus (HPF) and neocortex (NCx). Whelan G, Kreidl E, Wutz G, Egner A, Peters JM, Eichele G., Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin., EMBO J. 2012 Jan 4;31(1):71-82, with permission from John Wiley & Sons, Inc.

ESCO2 expression in mouse starts out low in G1, rises sharply in mid- to late-S, and declines in G2, consistent with the timing of its foci formation and intensity by IF. This timing reflects pericentric replication timing, and ESCO2 foci colocalize with PCH as marked by characteristic HP1α and PCNA staining and confirmed by ChIP-qPCR. In KO MEFS, acetylation and Sororin association with chromatin is reduced. By ChIP-qPCR, Sororin is reduced at PCH. Therefore, ESCO2 is localized to PCH in mouse cells and is involved in sororin recruitment to this region. ESCO2 function at PCH and its importance for centromeric constriction helps to explain the 'railroad track' morphology of RBS chromosomes.

There is also evidence that ESCO2 is involved in ribosome biogenesis. Ectopically-expressed ESCO2 localizes to the nucleus and is highly enriched in the nucleolus, where cohesin also localizes and has roles in ribosome biogenesis (van der Lelij et al., 2009, Gerton 2012). Mutations that affect ribosome biogenesis lead to developmental defects including craniofacial anomalies (Gerton 2012), which are also seen in RBS patients. RBS patient cells have been reported to have reduced global translation as measured by ^{35}S-methionine incorporation as well as fewer active ribosomes, and this is also observed in yeast cells expressing an RBS-analogous missense mutation in Eco1 (Bose et al., 2012). This suggests that these cells have ribosome assembly defects leading to altered translation that may contribute to RBS phenotypes.

Thesis Aims

The goal of the research in this study is to understand the mechanism, regulation, and functions of ESCO1/2 and their substrate, acetylated cohesin. To achieve this, the genome-wide binding sites of ESCO1/2 are first identified by ChIP-Sequencing, to reveal thousands of genome-wide ESCO1 binding sites but few ESCO2 sites of enrichment.

Using this data, I ask how ESCO1 binds to its enrichment sites, investigating the timing and requirement of *cis* and *trans* elements required for its targeting. This leads to the findings that ESCO1 is targeted to its discrete sites in G1 in a cohesin-dependent manner. The divergent N-terminal portion of the protein, through several vertebrate-specific regions, is shown to target ESCO1 to its sites. Deletion of several of these regions leads to mislocalization of ESCO1 from its enrichment sites and some of these regions are also required for ESCO1's function in cohesion. Additionally, the N-terminus is shown to interact directly *in vitro* with cohesin subunit SMC1A, and with PDS5B and CTCF. Depletion of these factors in human cells leads to reduced ESCO1 targeting, indicating that they also interact and recruit ESCO1 *in vivo*.

In addition, mapping of AcSMC3 sites is performed by employing a novel monoclonal antibody generated for this study that is specific for the modified, acetylated form of cohesin. It is discovered that cohesin is acetylated at the majority of its enrichment sites and this acetylation occurs on unreplicated DNA, i.e. before the requirement for cohesion. The acetylation sites largely overlap with ESCO1 sites, and acetylation at these discrete sites is dependent on ESCO1. Furthermore, AcSMC3 mapping at different cell-cycle time-points reveals that acetylation does not change significantly at these sites after G1.

The findings that cohesin acetylation at its enrichment sites and that ESCO1 is targeted to cohesin/CTCF sites in G1 and interacts directly with CTCF led to investigation of a role for the cohesin acetyltransferases in transcription. This reveals that ESCO1 loss causes altered transcription of many genes. Its vertebrate-specific N-terminal regions are also shown to play a role in its transcriptional function. Furthermore, I look at the binding sites of ESCO2 and find that it is enriched at REST/NRSF (RE1 silencing transcription factor/neural-restrictive silencing factor) motifs where it functions to repress target genes.

CHAPTER TWO

Materials and Methods

Cell culture and synchronizations

HeLa cells were grown in Dulbecco's modified essential medium (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin. RPE cells were grown in DMEM:F-12 with 10% FBS and 1% penicillin–streptomycin. HCT116 cells were grown in McCoy's medium with 10% FBS and 1% penicillin–streptomycin. 0.4mg/ml active G418 was used for selection and 0.2-0.4mg/ml active G418 was used to maintain stable resistant cell lines.

For AcSMC3 time-point ChIP, S1, S2 and G2 populations were arrested by double thymidine block (2.5mM thymidine for 16-18hours, 8 hour release, followed by 2.5mM thymidine for 16 hours) and 0, 3 and 6 hour release, respectively. AcSMC3 G1 cells were harvested after a thymidine block (20 hours) and release for 3 hours, 100ng/ml nocodazole block for 10-12 hours and shake-off with 3-5 hour release. For all other synchronizations, cells were arrested in G1 or G2 by double thymidine block and release for 13 hours (G1) and 5-6 hours (G2). AcSMC3 ChIP-Seq was done in two biological replicates for S1, S2 and G2 time-points. Sequenced and mapped reads were combined for these replicates for peak-calling.

Cell lines

ChIP-Seq on ESCO1 and ESCO2 was performed using FLAG-tagged lentivirally-transduced cell lines. FLAG-ESCO1 or FLAG-ESCO2 was cloned into pLVX (Clontech) and virus was generated with Lenti-X 293T cell line using the Lentiphos HT System according to the manufacturer's instructions. All other stable cell lines were made by

retroviral transduction. Constructs were cloned into pQCXIN (Clontech) and co-transfected with pVSV-G into Phoenix cells. Viral supernatants were applied to HeLa cells with 0.01 mg/ml polybrene. Transductants were selected with 0.4 mg/ml G418 after 48 hours and cloned by limiting dilution to make clonal cell lines where indicated.

ESCO1-knockout (RPE) and HDAC8-knockout (HCT116) cell lines were made in the lab by Rebecca Sherwood and Mathew Jones, respectively, by adenoviral targeting (Berdougo et al., 2009). A conditional ESCO1-knockout was generated by targeting exon 6. HDAC8 is X-linked with one copy in HCT116 cells. Exon 4 of HDAC8 was targeted and a cell line containing the targeted floxed allele and a neo cassette, before Cre excision of the floxed exon 4 (HDAC8-flox-neo) was used for ChIP-Seq in this study. Although the exon is not deleted in these cells, protein levels are severely affectedFigure 3.9), likely because of gene splicing disruption by the neo cassette. There is also a 13-fold increase in acetylation by Western on WCE in the HDAC8-flox-neo cells. Cells after Cre excision were also used for ChIP-qPCR, where ChIP-Seq results were validated.

RNA interference

siRNA was used at final concentrations of 50-100nM for all except 5-30nM for PDS5A and PDS5B, and transfected with Oligofectamine (Invitrogen). Target sequences used were as follows: ESCO1, GGACAAAGCTACATGATAGTT; ESCO2, TAAGTCCACTGTCTATCCATT; RAD21, GGTGAAAATGGCATTACGG ; WAPL, CGGACTACCCTTAGCACAA; RAD21, TGGAAGATCTCCTAACTAA; PDS5A, TTCTTCCTCAGGAACCCCATT; PDS5B, ; CTCF, GGACGATACCCAGATTATAAC; GL2, CGTACGCGGAATACTTCGATT;

Antibodies and immunoblotting

AcSMC3 antibody was generated in this study with help from the MSKCC Monoclonal Antibody Core Facility. Screening was performed with hybridoma supernatants by ELISA using mono- K105 or K106 acetylated, di-K105,K106-acetylated and non-acetylated peptides and selected clones were further validated by Western blot by testing specificity on extracts from SMC3-WT and SMC3-AA (K105A,K106A mutant) cells. Hybridoma clone 21A7 was selected and CELLine (Integra) bioreactor supernatant was purified by the Monoclonal facility.

Additional antibodies used in this study are as follows: FLAG (Stratagene 200472), ESCO1 (Terret et al., 2009), CTCF (Millipore 07-729), Histone H3 (Abcam ab1791), PDS5A (Bethyl A300-088A), PDS5B (Bethyl A300-538A-1), SMC3 (Bethyl A300-060A), tubulin (Santa Cruz sc-5286), topo II (Millipore MAB4197), RAD21 for Western (Bethyl A300-080A), RAD21 for ChIP (Abcam ab992), and HDAC8 (Santa Cruz sc-11405).

Cell pellets were either extracted with NETN buffer (20 mM Tris (pH 8.0), 100 mM NaCl, 1 mM EDTA, and 0.5% NP-40) to isolate chromatin where indicated or lysed directly in sample buffer for WCE.

Chromatin immunoprecipitation (ChIP)

ChIP was performed as described (Kagey et al., 2010) with protocol and buffers as detailed at http://younglab.wi.mit.edu/hES_PRC/ChIP.html with slight modifications. Briefly, cells were cross-linked with 1% formaldehyde for 10 min and quenched with 125 mM glycine and washed twice with cold PBS prior to harvesting with a silicon scraper. Cells were lysed and sonicated (Bioruptor) to yield fragments about 500bp in size. (DNA was re-sheared with a Covartis sonicator to produce shorter (~100-150bp) fragments by the MSKCC Genomics Core Lab prior to library preparation.) Extracts were incubated

with Dynal protein G magnetic beads and antibodies overnight. Antibody/bead complexes were incubated with 50-fold molar excess of peptides for 1 hour where indicated. Beads were washed with RIPA buffer four times and once with TE containing 50mM NaCl. Bound protein-DNA complexes were eluted at 65°C for 20 minutes with vortexing. Cross-links were reversed by incubation at 65°C for 6 hours to overnight. RNA and protein were digested with sequential RNase A and proteinase K treatments and DNA was extracted with phenol:chloroform. WCE input DNA was also purified. ChIP-qPCR primers are listed in Table 3.2.

Sequencing and mapping

Quality control with Agilent Bioanalyzer and sequencing on a SOLiD5500 platform were performed by the MSKCC Genomics Core Lab. ChIP and WCE input samples were sequenced to generate 50-bp reads. Reads were aligned by the MSKCC Bioinformatics Facility using Bioscope to UCSC hg19 genome.

ChIP-Seq analysis

Mapped reads were filtered to keep only uniquely aligned hits. Peaks were called with MACS2.0 using a q-value of 0.05, or p-value cutoff of 0.1 which was used only for IDR analysis. Duplicate reads were removed for peak-calling. Default MACS2.0 parameters were used except for shift-size of 250. Overlaps between sets were called with BedTools. Mapped reads were visualized using merged BAM files of two replicates where replicates were available. Normalization of read numbers in the time-point acetylation was performed by dividing by the mapped tags in the time-point sample with the lowest number of mapped tags. HOMER was used for *de novo* motif analysis and to plot histograms of ChIP-Seq read and motif densities.

Chromosome spreads

0.2 ug/ml nocodazole was added to cells for 2 hours prior to harvest. Cells were incubated with 0.075M NaCl at 37°C for 15 minutes and fixed with 3:1 volume methanol/ glacial acetic acid and incubated at -20^{0}C overnight or longer before preparation of slides. DNA was stained with Hoechst dye 33258.

***In vitro* binding assays**

ESCO1 N-terminal fragments were cloned into pET28a for expression of 6xHis-tagged protein in Rosetta (DE3) cells with induction by Isopropyl β-D-1-thiogalactopyranoside addition for 4 hours. Protein was purified with Ni-NTA beads (Novagen) and dialyzed for 24 hours to remove imidazole. Prey proteins were transcribed/translated from T7-promoter plasmids with TNT Quick Coupled Transcription/Translation System (Promega) with ^{35}S-Methionine (Perkin-Elmer). Binding assays were performed with His-Tag Isolation Dynabeads (Invitrogen) in 50mM Tris-HCl pH8.0, 120mM NaCl,1% NP-40 buffer and washed three times in same buffer after binding of bait and again after prey pulldown. Bound complexes were eluted by boiling beads in sample buffer.

Microarray and RT-PCR

Total RNA was isolated with RNeasy Mini Kit (Qiagen). cDNA was prepared with Superscript III First-Strand Synthesis (Invitrogen). Three independent knockdown samples were used for microarray, with siRNA omitted from control samples. Three knockdowns were used for RT-qPCR validation of microarray results with siRNA targeting GL2 used as control. Microarrays were performed by the MSKCC Genomics Core Lab on Human HT-12 Illumina arrays. Partek software was used for analysis.

RT-qPCR primers are listed in Table 5.1. RT-qPCR was done in technical replicates using indicated number of biological replicates, and analyzed using the delta delta Ct

method with GAPDH as the internal reference gene. Fold-change was normalized to GAPDH levels.

CHAPTER THREE

Genome-wide Mapping of ESCO1/2 and Dynamics of Cohesin Acetylation

Introduction

Genome-wide mapping of cohesin in mammalian cells by ChIP-chip has shed light on its function as a transcriptional regulator in cooperation with the transcription factor CTCF, with which it shares a large number of binding sites (Parelho et al., 2008, Wendt et al., 2008), as well as with the mediator complex, with which it also co-localizes at thousands of non-CTCF sites in mouse ES cells (Kagey et al., 2010). CTCF is required for cohesin to be enriched at its sites, and cohesin is also required for maximal CTCF binding at these sites (Wendt et al., 2008) . Cohesin cooperates with both CTCF and mediator functionally to form DNA loops, to bring together promoters and enhancers in some cases and sequester enhancers from target genes in other cases, and thereby regulate gene expression.

We sought to identify cohesin acetyltransferase binding sites to understand where and how they interact with DNA and other proteins. ESCO1/2 are chromatin-bound and act only on DNA-bound cohesin while soluble cohesin cannot be acetylated (Nishiyama et al., 2010). Yeast Eco1 that loses the ability to efficiently interact with chromatin by mutation of its PCNA-interacting motif (PIP box) is defective for cohesion (Moldovan et al., 2006). Therefore, their DNA localization is essential for function. In addition, ESCO1 and ESCO2 have non-redundant functions in acetylation, cohesion, and replication fork speed (Hou and Zou 2005, Terret et al., 2009). One possibility for the non-redundancy is that they localize to different sites and acetylate separate pools of cohesin. It has been shown that the divergent N-terminus, and not the highly conserved C-terminus, targets these proteins to chromatin (Hou and Zou 2005). These different N-termini may therefore

target ESCO1 and ESCO2 to distinct sites. ChIP-Seq provides a non-biased approach to identify binding sites and test this.

We also wanted to study specifically the acetylated form of cohesin to understand its cell-cycle dynamics and regulation. Therefore, a monoclonal antibody that recognizes the acetylated form of SMC3 was generated and used to map acetylation sites genome-wide. Cohesin acetylation was further mapped in different cell cycle stages as well as under different contexts, of ESCO1 and HDAC8 loss, to understand its dynamics and regulation.

Results

ESCO1 and ESCO2 Localization

In order to map ESCO1 and ESCO2 binding to DNA, clonal FLAG-tagged lines in HeLa cells were created for each protein and anti-FLAG antibody was used to immunoprecipitate the proteins. Two separate biological replicates were performed for each protein and the consistency between the replicates as measured by irreproducibility discovery rate (IDR) analysis and overlap between replicates indicate good reproducibility (Figure 3.6). The overlap between replicates is called ‘consensus sites’ and used for downstream analysis. ChIP-Seq statistics are listed in Table 3.1.

ChIP-Seq results show that ESCO1 is enriched at many sites throughout the genome (11,728 replicate-consensus sites) and ESCO2 is enriched at very few sites (23 sites). However, the enrichment of ESCO2 at these sites is high, indicating that they are true enrichment sites (Figure 3.1). ChIP-Seq enrichment at these sites were validated by ChIP-qPCR (Figure 3.2; all ChIP-qPCR primers are listed in Table 3.2).

Cohesin sites have been previously mapped in HeLa cells (ENCODE) and were also mapped in this study with RAD21 antibody (Figure 3.5). Visual inspection of the ESCO1 and RAD21 ChIP-Seq reads mapped to the genome indicates a high degree of overlap between ESCO1 and cohesin. This is confirmed by analyzing the overlap of peak calls in these sets, which shows 70% of ESCO1 sites co-localize with RAD21, and by plotting ESCO1 peak intensities around RAD21 sites (Figure 3.3). It is further established by *de novo* motif analysis of ESCO1 binding site sequences, which shows enrichment of the CTCF motif (Figure 3.3C,D).

Together, these results show that ESCO1 is enriched throughout the genome at cohesin/ CTCF binding sites, whereas ESCO2 is not similarly enriched at discrete sites throughout the genome, though it is highly enriched at a few sites. These ESCO2 sites are analyzed in Chapter Five.

Generation of AcSMC3 monoclonal antibody and AcSMC3 ChIP-Seq

To specifically study the modified, acetylated form of cohesin, a mouse monoclonal antibody was generated in collaboration with the MSKCC Monoclonal Antibody Facility, against acetylated SMC3. A synthetic peptide surrounding lysines K105 and K106, both acetylated, was used to immunize mice and generate hybridomas. The selected antibody was specific for AcSMC3 on Western blot with no detectable background and behaved as predicted under ESCO1/2 single and double knockdown conditions, with a decrease in signal after single depletion and a further decrease after double depletion (Figure 3.4A). Titration ELISA using varying amount of peptides that are mono-acetylated on either K105 or K106, or di-acetylated or unacetylated on both residues, shows that the antibody is highly specific for the mono-K105 and di-K105,K106-acetylated forms, but not the mono-K106 and unacetylated forms (Figure 3.4).

Table 3.1 ChIP-Seq Statistics

Sample	Cell Line	Antibody	Total reads	Mapped reads	Peaks
ESCO1 replicate 1	HeLa FLAG-ESCO1 clone 1	FLAG	33,365,016	15,967,094	14,152
ESCO1 replicate 2	HeLa FLAG-ESCO1 clone 1	FLAG	20,216,297	6,854,234	16,296
ESCO2 replicate 1	HeLa FLAG-ESCO1 clone 32	FLAG	36,844,226	14,614,906	38
ESCO2 replicate 2	HeLa FLAG-ESCO1 clone 32	FLAG	19,094,748	7,022,089	67
AcSMC3 replicate 1 (G2)	HeLa	AcSMC3	43,123,384	26,206,213	26,531
AcSMC3 replicate 2 (G1)	HeLa	AcSMC3	28,448,979	22,460,470	27,110
AcSMC3 (S1)	HeLa	AcSMC3	41,608,724	25,850,047	26,013
AcSMC3 (S2)	HeLa	AcSMC3	60,382,204	42,432,281	21,023
RAD21 replicate 1 (G2)	HeLa	RAD21	22,178,483	13,036,456	17,117
RAD21 replicate 2 (G1)	HeLa	RAD21	21,767,001	15,710,342	34,520
WT	WT RPE (G1)	AcSMC3	3,282,867	2,812,004	5,941
ESCO1-KO	ESCO1-KO RPE (G1)	AcSMC3	3,048,730	2,384,597	1,128
WT	WT HCT116	AcSMC3	21,226,501	18,625,183	21,193
HDAC8-neo	HDAC8-neo HCT116	AcSMC3	28,325,840	21,042,230	27,896
WT	WT HCT116	RAD21	3,032,492	2,134,673	7,606

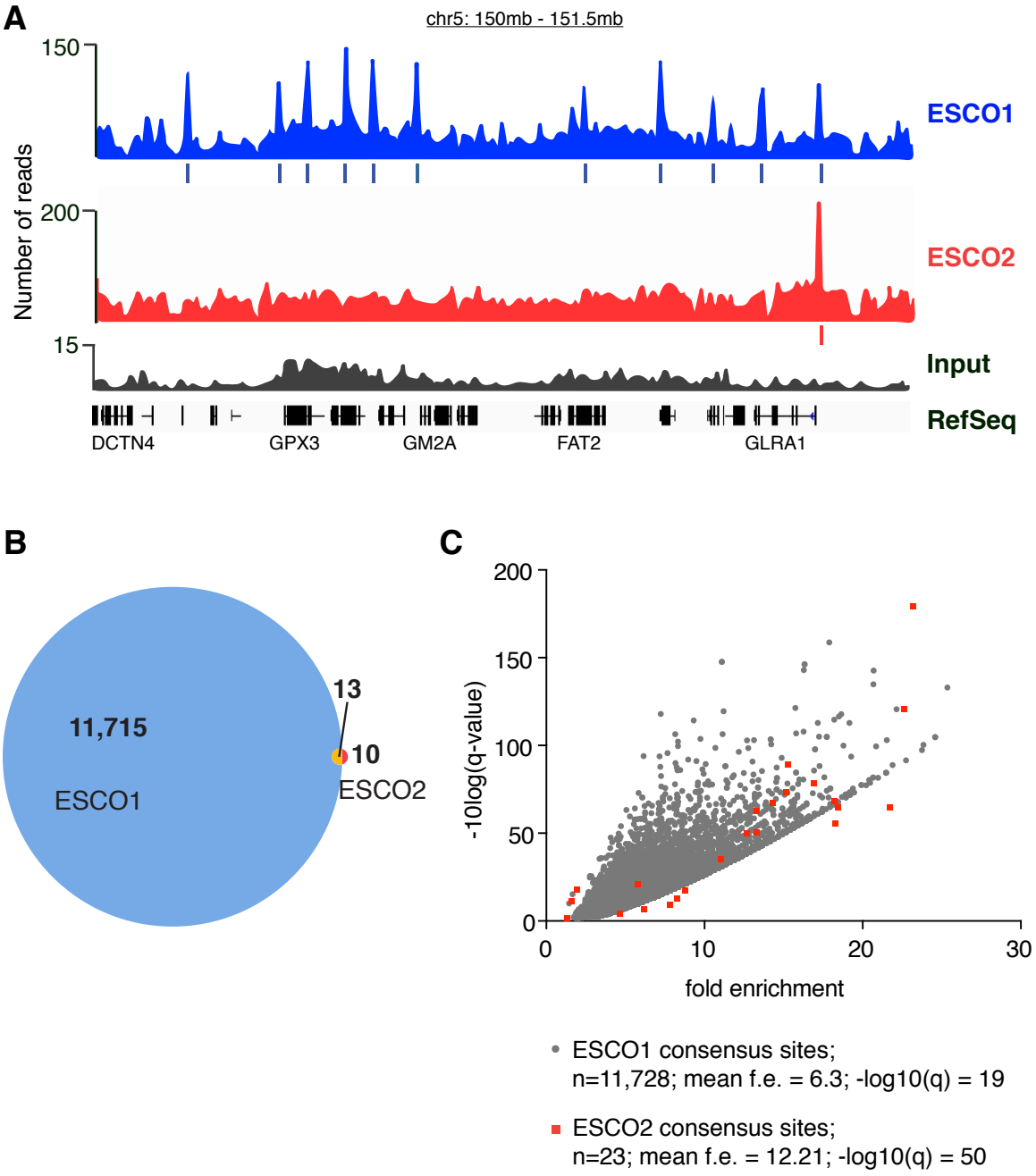


Figure 3.1. ESCO1 and ESCO2 occupancy in HeLa cells. (A) Sequence coverage of ESCO1 and ESCO2 over a 1.5mb region of chromosome 5. Bars below coverage tracks indicate replicate-consensus peaks. (B) Comparison of number of consensus peaks called for ESCO1 and ESCO2 and overlap between sets. (C) Scatter plot comparing fold-enrichment over input and q-values of ESCO1 and ESCO2 peaks. Mean fold enrichment (f.e.) and -log10(q) values are listed for each set of consensus peaks.

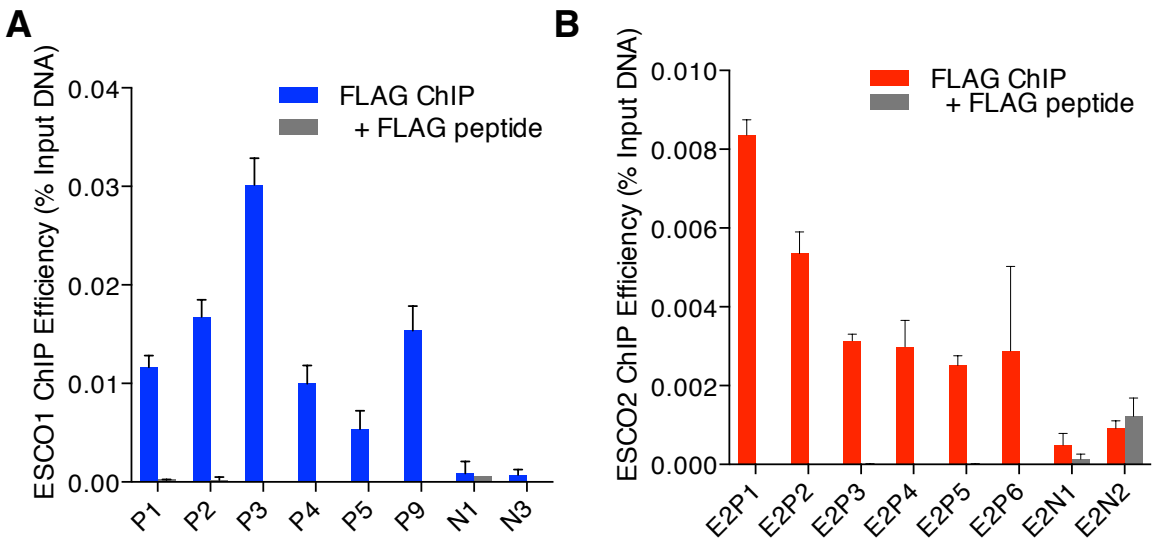


Figure 3.2. ChIP-qPCR validation of ESCO1 and ESCO2 peaks. (A) FLAG ChIP was performed using FLAG-ESCO1 tagged HeLa cell line followed by qPCR assaying six positive sites (P1-P9) called for ESCO1 ChIP-Seq, and two negative sites (N1 and N3), and (B) using FLAG-ESCO2 line assaying six positive (E2P1-E2P6) and two negative (E2N1 and E2N2) sites for ESCO2 ChIP-Seq, with and without FLAG peptide competition. Error bars represent standard deviation in triplicate qPCR measurements.

Table 3.2 List of ChIP-qPCR primers

primer	hg19 coordinates	Forward	Reverse	Description
P1	chr5: 31,270,891-31,270,978	CTTCCTGGGCTCAAGATGTC	TTTTCAGGGATAGCCTGGTG	CDH6 intron
P2	chr1: 35,208,594-35,208,675	AAGCAAAGTCCTCCGTTCAG	TATGGCAGATGCCTAACAGG	intergenic
P3	chr11: 62,980,339-62,980,488	GTTTCAGCTTGCCATCAACC	TGCACCAGCCTTCACTTATC	SLC22A25 intron
P4	chr15: 44,196,128-44,196,200	ACAGGCCCGTGTTTATTCTG	CATGGCTCCTCAAGTCACTC	FRMD5 intron
P5	chr16: 115,793-115,933	GGGCACAAATGACTTTCCAG	CATGTCATGGTAGAGCCAAGC	RHBDF1 intron
P6	chr17: 46,882,330-46,882,399	AGGCTCTGGGGAAATACCAG	TGCAGGTGGATCTCTGTAACTG	TTLL6 intron
P7	chr17: 46,891,895-46,892,028	TGAGATCTTCCAAGCCAGAG	CCAAATAACCCGCAATGG	TTLL6 intron
P8	chr17: 47,633,900-47,633,976	CAGGGCCTCTGATGGAAAC	GCTGACTCCGGGCAGAAC	NR_103773 intron
P9	chr18: 19,154,432-19,154,599	GCCTTTGTGATCTCCGGTTA	GGTCATCAAAGGCAGCATCT	ESCO1 exon
P10	chr18: 61,515,478-61,515,699	AGGACAGGAGTTGCTTGTTACC	CAGCACTTGTCTGACTTCACG	intergenic
N1	chr7: 27,211,083-27,211,200	CTTTGCACCATTGACCTCAG	AAATGGCCCCTGTCTTCG	NR_037940 exon
N2	chr17: 47,699,872-47,700,013	AGATTCCCCCTTTCAACCTC	TGGGCTTCTTGCTTTCTCTG	SPOP intron
N3	chr10: 93,064,214-93,064,306	TCGGCCTAATAAGGGAACTG	AAAACTCCCCAACTCTGGTG	intergenic
N4	chr18: 19,137,537-19,137,659	CTTGAAGGAAAGGGCATAGC	TGGCTGAGAAAAACTCAATCC	ESCO1 intron
E2P1	chr1: 26,735,119-26,735,245	ACTTGGAGGTGGGAACTCAA	GGCCTGCTTCTCTGTCCA	LIN28A upstream
E2P2	chr5: 151,304,268-151,304,367	GAGCGAGGGGGTCGTAGATA	AATACTCTTCGCGCTCCTGA	GLRA1 exon
E2P3	chr19: 50,490,502-50,490,585	TGCTGAACTCATACCCACAGC	ACAAAAGGGAAGTGCTAGGG	VRK3 intron
E2P4	chr20: 35,012,534-35,012,603	ACCTTTGAGCTGTGCTGTTTG	ACCCCTGACATCAGCACTAAG	DLGAP4 intron
E2P5	chr2: 163,697,286-163,697,376	ACAAGAGTTTCTCTCCCAGGTG	GTTGTAGGCAGAGAAGAGTTGC	intergenic; KCNH7 downstream
E2P6	chr11: 70,496,525-70,496,658	AATCCTCCCAGCTGTCTAGTTC	TGCACAGAGCAAGAGCATTG	SHANK2 intron
E2N1	chr10: 88,176,634-88,176,714	ACCTGCTGCTTCCTAAGTCTTG	GTAACTCAAACCGCGCAAAG	intergenic
E2N2	chr4: 39,839,540-39,839,656	CTTTGGCATTACCTGCTTCC	AATCTCAGGGCAATGCTACC	PDS5A exon

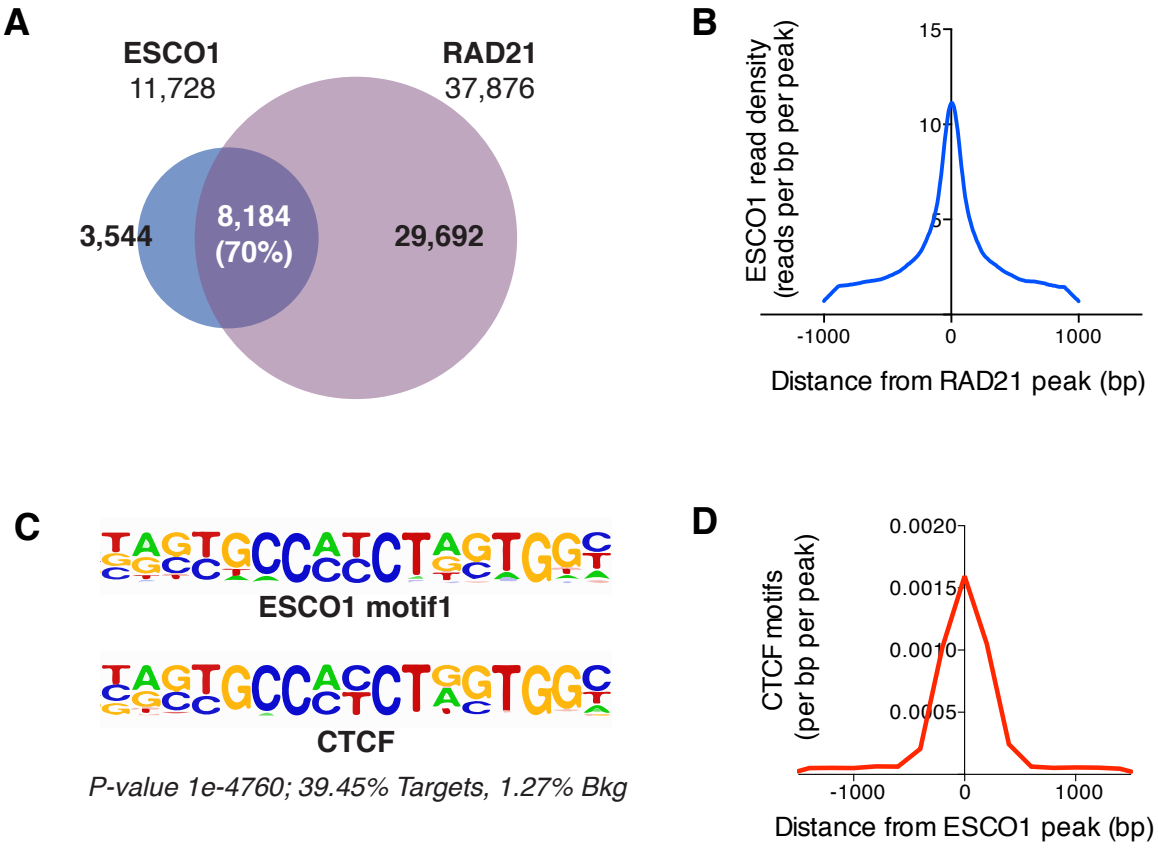
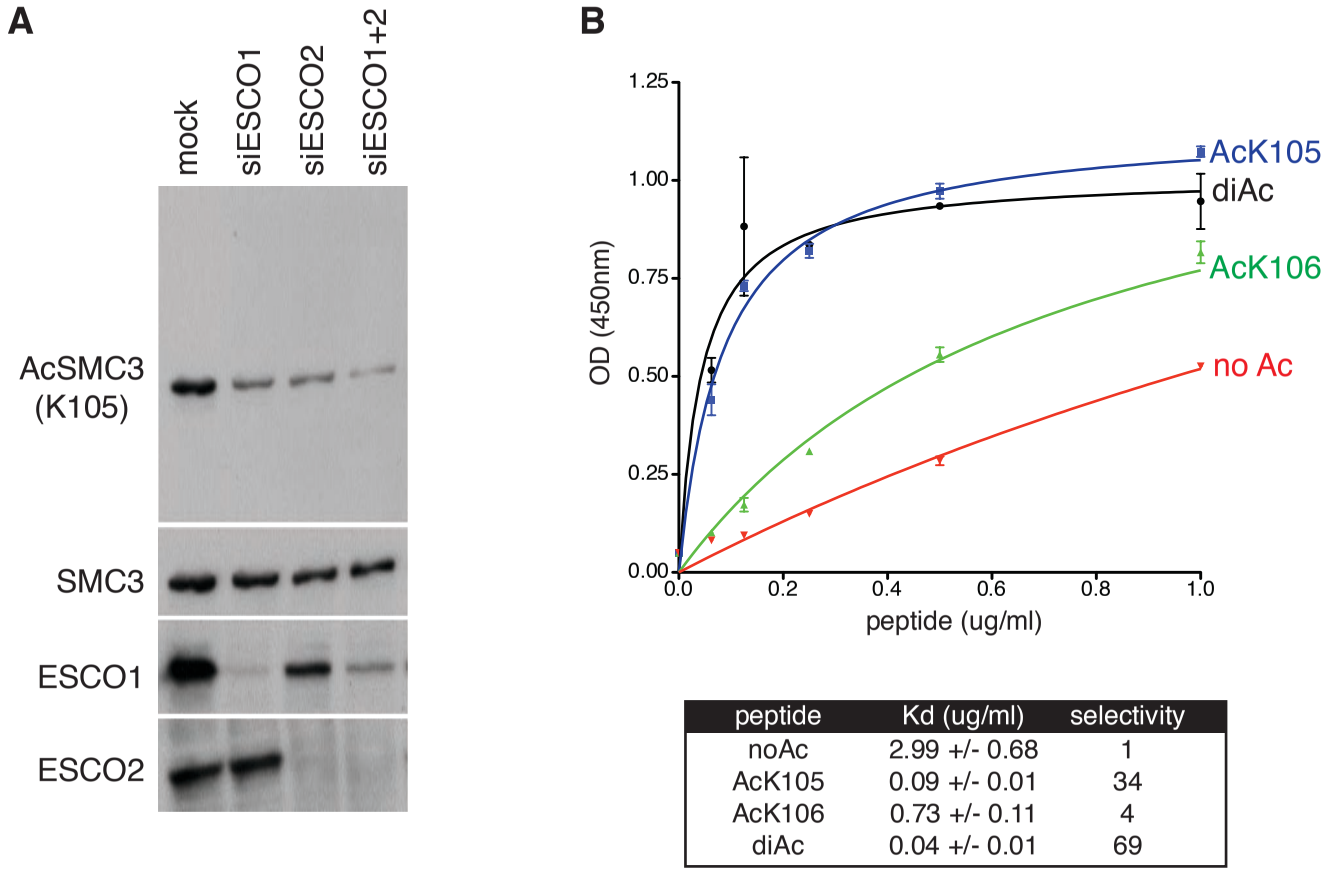


Figure 3.3. ESCO1 overlaps with cohesin/CTCF binding sites. (A) Venn diagram depicts overlap between consensus ESCO1 sites and union, in both replicates, of RAD21 ChIP-Seq sites. (B) Histogram plots ESCO1 ChIP-Seq read depth around RAD21 peaks. (C) *De novo* motif analysis using ESCO1 binding sites shows an enrichment of the CTCF motif. (D) Histogram shows enrichment of CTCF motif density at ESCO1 binding sites.



peptide	Kd (ug/ml)	selectivity
noAc	2.99 +/- 0.68	1
AcK105	0.09 +/- 0.01	34
AcK106	0.73 +/- 0.11	4
diAc	0.04 +/- 0.01	69

Figure 3.4. Characterization AcSMC3 monoclonal antibody. (A) Western blot demonstrates specificity of antibody and expected decrease of signal under ESCO1 and ESCO2 knockdown conditions. (B) ELISA of antibody binding to different concentrations of K105/K106 mono-acetylated, di-acetylated, and non-acetylated peptides (CRRVIGAKKDQYFLD) reveals a strong preference for K105 mono-acetylated and K105,K106 di-acetylated peptides.

ChIP-Seq of AcSMC3 in HeLa cells using the monoclonal antibody was performed in two biological replicates, with high consistency between the replicates and 17,185 replicate-consensus sites were identified (Figure 3.6). Overlap of acetylation sites and ESCO1 or total cohesin sites shows a high overlap between ESCO1 and acetylated cohesin and shows that most cohesin sites are acetylated (Figure 3.5B,C). These acetylation sites were also validated by ChIP-qpCR with peptide competition samples, and the signal is competed specifically by the acetylated peptide but not by the non-acetylated peptide (Figure 3.9E).

Effect of ESCO1 or HADC8 loss on cohesin acetylation

As shown above, many ESCO1 enrichment sites overlap with AcSMC3 sites. To test if these sites are acetylated in the absence of ESCO1, I performed ChIP-Seq with an ESCO1-KO Retinal Pigment Epithelial (RPE) cell line (KO line generated by R. Sherwood). Note that RPE cells behave differently from HeLa cells, in that there is no cohesion phenotype with loss of either ESCO1 or ESCO2, but a cohesion defect is seen with double-depletion (data not shown). However, there is a defect in acetylation levels with loss of only ESCO1 (Figure 3.7A), making it possible to study acetylation dynamics in this background.

Unlike WT RPE cells in which acetylation begins in G1 and then rises in S-phase, ESCO1-KO cells have no detectable acetylation by Western when synchronized in G1 by nocodazole block, and acetylation only appears in S-phase (data not shown). As shown in Figure 3.7A, the level of overall acetylation remains lower in these cells.

To analyze how acetylation is affected at discrete sites in the absence of ESCO1, and to dissect the contribution of ESCO1 and ESCO2 acetylation after cells are acetylated, cells were arrested in G2 to perform ChIP-Seq with ACSMC3 antibody. The results show

that the overall number of acetylation peaks, as well as the intensities of those peaks that are called in ESCO1-KO cells, are reduced relative to WT RPE cells (Figure 3.7). These data indicate that acetylation at these discrete sites is mediated by ESCO1 even into G2 when ESCO2 is active.

Next, to address if a loss in acetylation at specific sites affects cohesin and/or CTCF stability at these sites, I performed AcSMC3 ChIP-qPCR in WT and ESCO1-KO cells arrested in G1 by double-thymidine block and release. G1 synchronization was confirmed by FACS with >88% of cells in G1 for both WT and ESCO1-KO (data not shown). Acetylation at specific sites assayed by qPCR is greatly reduced in KO cells compared to WT. However, cohesin and CTCF stability are not affected (Figure 3.8). This shows that acetylation is not required to maintain cohesin or CTCF levels at enriched sites.

An HDAC8-deficient cell line was also generated in the lab (by M. Jones) in HCT116 background. The one copy of X-linked HDAC8 in these cells was targeted by a flox-neo cassette to flox exon 4. Cells prior to Cre-mediated exon 4 excision were used for AcSMC3 ChIP-Seq. These cells have depleted HDAC8 protein levels (Figure 3.9A) likely due to disruption in gene splicing by insertion of the neo cassette.

In HDAC8-deficient cells, there is an overall increase in acetylation signal by Western (Figure 3.9A). ChIP-Seq results show an increase in AcSMC3 peak number and fold-enrichment of the peaks that are called in HDAC8-deficient cells compared to those called in WT (Figure 3.9). ChIP-qPCR confirmed these results in cells after Cre-excision of exon 4 (HDAC8-KO; Figure 3.9E). Therefore, HDAC8 loss results in increased acetylation levels at discrete cohesin binding sites on chromatin.

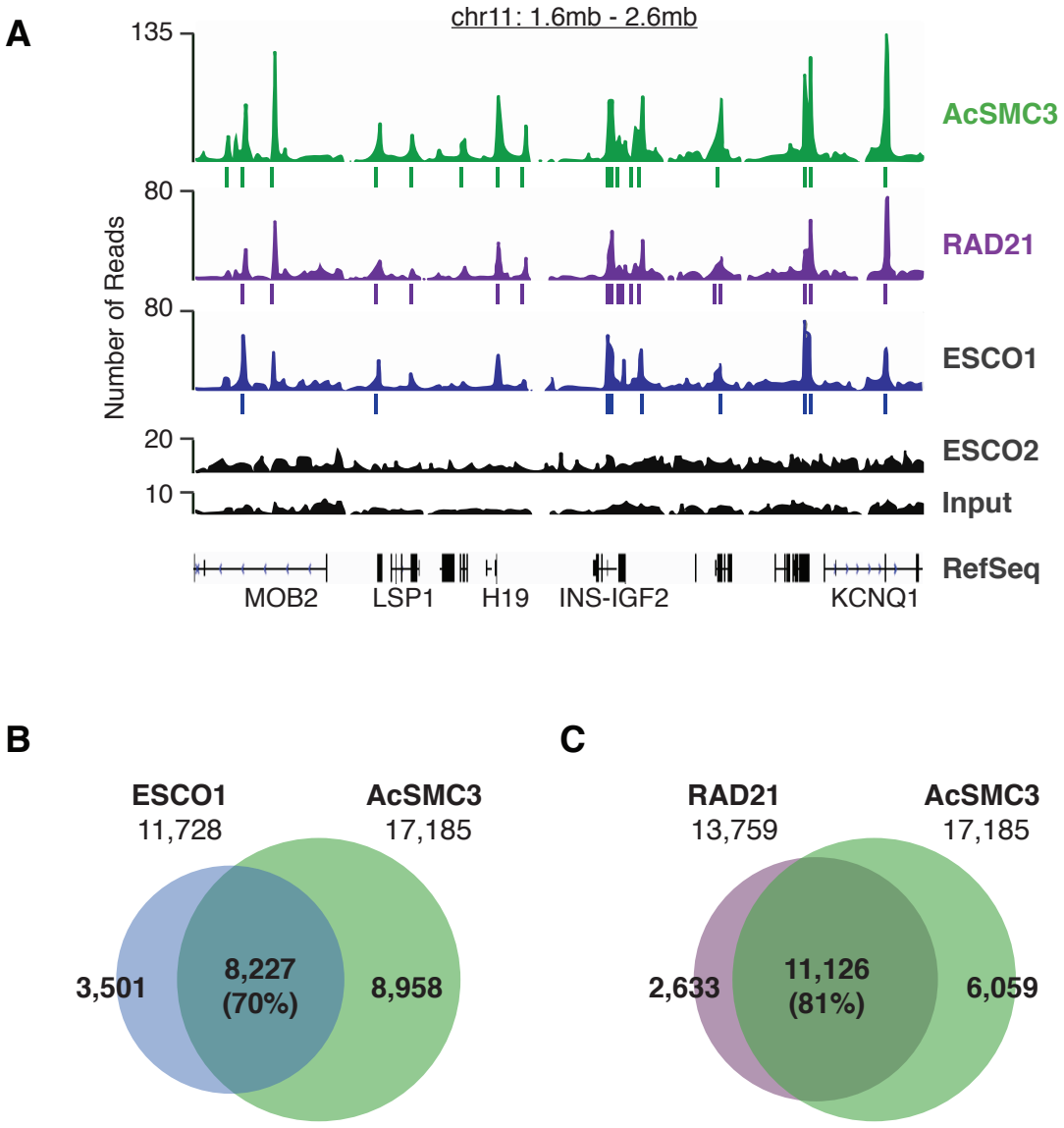


Figure 3.5. AcSMC3 occupancy in HeLa cells. (A) AcSMC3, RAD21, ESCO1, ESCO2 ChIP-Seq and input reads are visualized over 1mb region of chromosome 11. (B) Venn diagram of consensus ESCO1 sites and acetylated cohesin sites shows large overlap in these sets. (C) Venn diagram of consensus sites of total and acetylated cohesin shows that the majority of cohesin sites are acetylated.

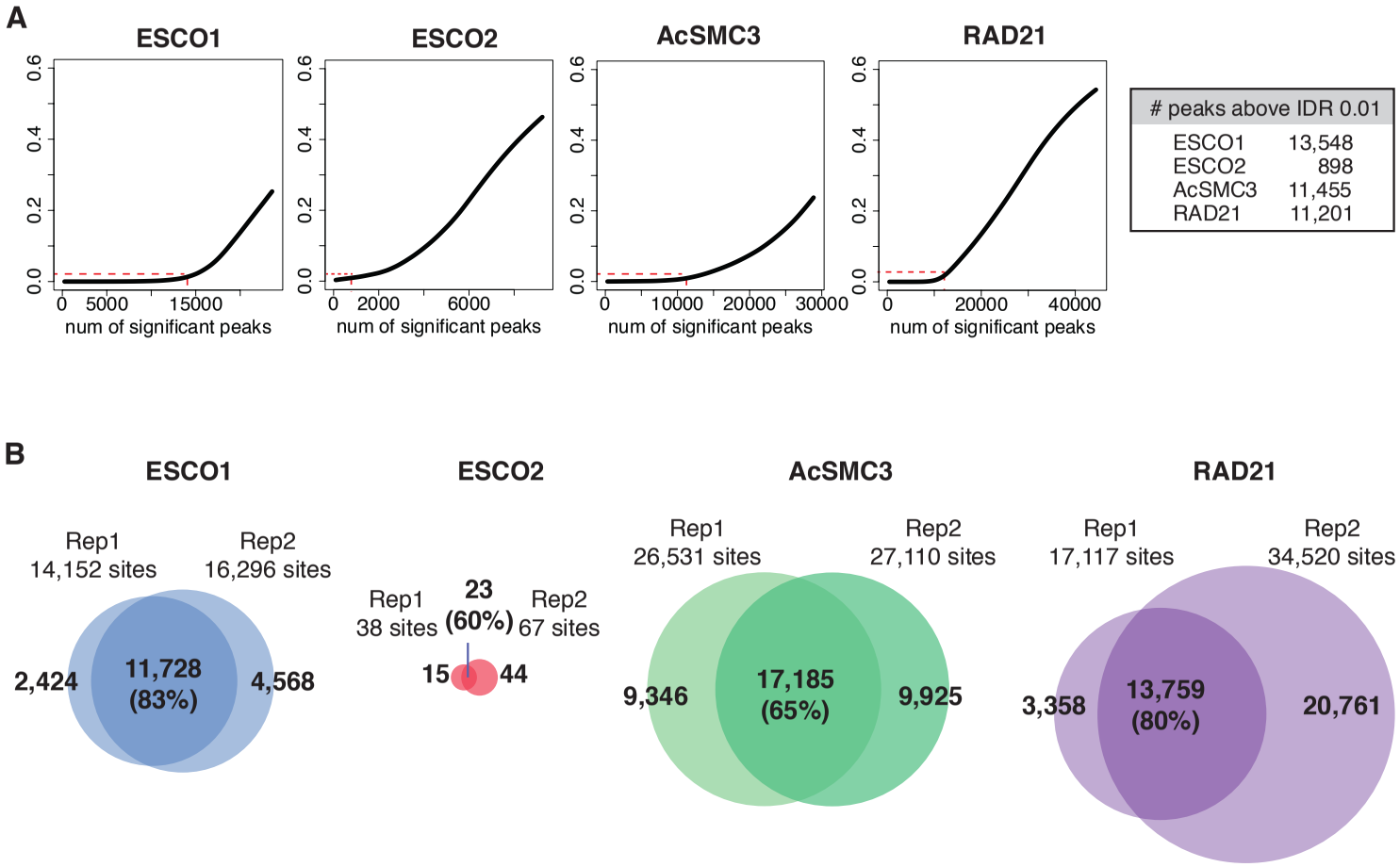

A
ESCO1
ESCO2
AcSMC3
RAD21
num of significant peaks
peaks above IDR 0.01
ESCO1 13,548
ESCO2 898
AcSMC3 11,455
RAD21 11,201
B
ESCO1
Rep1
14,152 sites
Rep2
16,296 sites
2,424
11,728
(83%)
4,568
ESCO2
Rep1
38 sites
23
(60%)
Rep2
67 sites
15
44
AcSMC3
Rep1
26,531 sites
Rep2
27,110 sites
9,346
17,185
(65%)
9,925
RAD21
Rep1
17,117 sites
Rep2
34,520 sites
3,358
13,759
(80%)
20,761

Figure 3.6. Irreproducible Discovery Rate (IDR) analysis of ChIP-Seq replicates and replicate overlap. (A) IDR analysis was performed on HeLa ChIP-Seq replicates with peaks called using MACS2.0 p-value cut-off of 0.1. Red dashed lines mark IDR cutoff of 0.01 and number of peaks above cutoff are listed to the right. (B) Overlap of number of peaks for ChIP-Seq replicates called with q-value cut-off of 0.05, termed consensus sites and used for further downstream analysis. Venn diagrams are drawn to scale except ESCO2 is 3x enlarged.

Cohesin acetylation throughout the cell cycle

Cohesin acetylation begins in G1 in human cells, therefore on unreplicated DNA before cohesion is established in S-phase, and discrete sites are also already acetylated in G1 as shown in Figure 3.9. Overall acetylation levels by Western blot rise when cells enter S-phase. To ask if acetylation at discrete sites also increases in S-phase, and look at how acetylation changes as cells progress through the cell cycle, AcSMC3 ChIP was performed in cells arrested in various cell cycle stages: in G1 by nocodazole block and release, and at three time-points after double thymidine block, early-S (S1), mid-S (S2) and late-S/G2 (G2) (Figure 3.10A).

ChIP-Seq mapping of acetylation at these time-points does not show significant differences in the number of peaks at the various time-points (Figure 3.10B, 3.11A). Acetylation levels also are not increased as cells go from G1 through S-phase at discrete sites (Figure 3.11A,B), in contrast to overall acetylation measured by Western blot. This is confirmed by ChIP-qPCR assaying specific sites (Figure 3.11C). Instead, it appears to be slightly higher in the G1 population, as seen by a shift in the fold-enrichment histogram and as confirmed by qPCR (Figure 3.11A,C).

Discussion

Difference in ESCO1 and ESCO2 distribution

The results presented here show that ESCO1 and ESCO2 bind differently to DNA. ESCO1 is enriched at tens of thousands of discrete sites throughout the genome, whereas ESCO2 has very few sites of enrichment. ESCO2 is still detected and highly enriched at some sites, as confirmed by qPCR with and without peptide competition, so this is not due to a problem ESCO2 to ChIP in this assay.

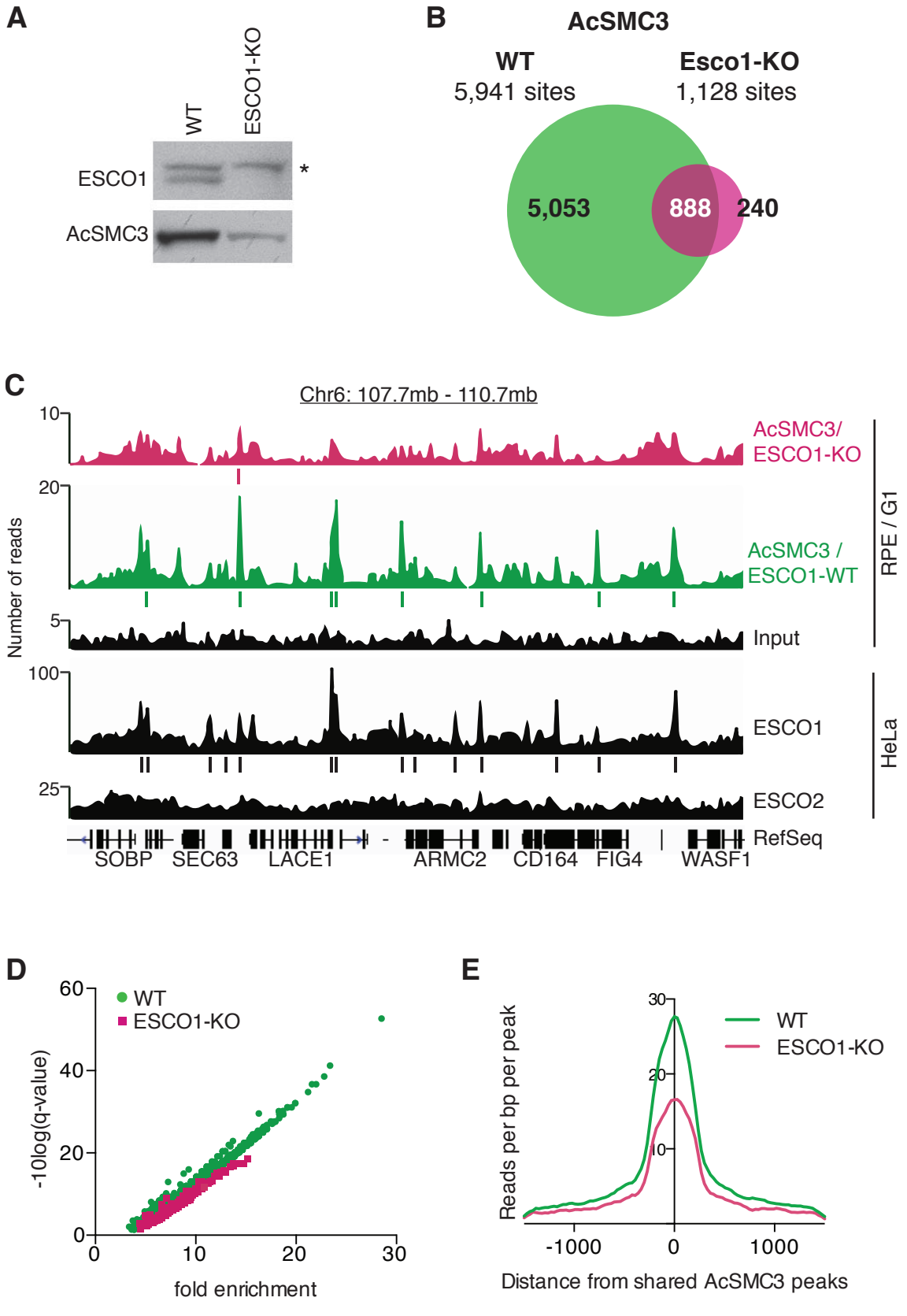

A
WT
ESCO1-KO
ESCO1
AcSMC3
*
B
AcSMC3
WT
5,941 sites
Esco1-KO
1,128 sites
5,053
888
240
C
Chr6: 107.7mb - 110.7mb
Number of reads
AcSMC3/
ESCO1-KO
AcSMC3 /
ESCO1-WT
Input
RPE / G1
ESCO1
ESCO2
HeLa
RefSeq
SOBP
SEC63
LACE1
ARMC2
CD164
FIG4
WASF1
D
WT
ESCO1-KO
-10log(q-value)
fold enrichment
E
WT
ESCO1-KO
Reads per bp per peak
Distance from shared AcSMC3 peaks

Figure 3.7. Acetylation is reduced in G2-arrested ESCO1-KO cells. (A) Western blot of RPE WT and ESCO1-KO asynchronous cells showing loss of ESCO1 signal and reduced overall acetylation in KO cells. Asterisk marks non-specific band. (B) Venn diagram comparing number of G2 acetylation peaks called for WT and ESCO1-KO cells. (C) ChIP-Seq mapped reads of AcSMC3 in ESCO1-KO and WT cells over a 3mb region of chromosome 6. ESCO1 and ESCO2 read data from HeLa cells is included for comparison. (D) Graph plotting fold enrichment and q-values for WT and ESCO1-KO cells and (E) histogram comparing read depth around overlapping peaks show a decrease in AcSMC3 peak enrichment in ESCO1-KO cells compared to WT cells.

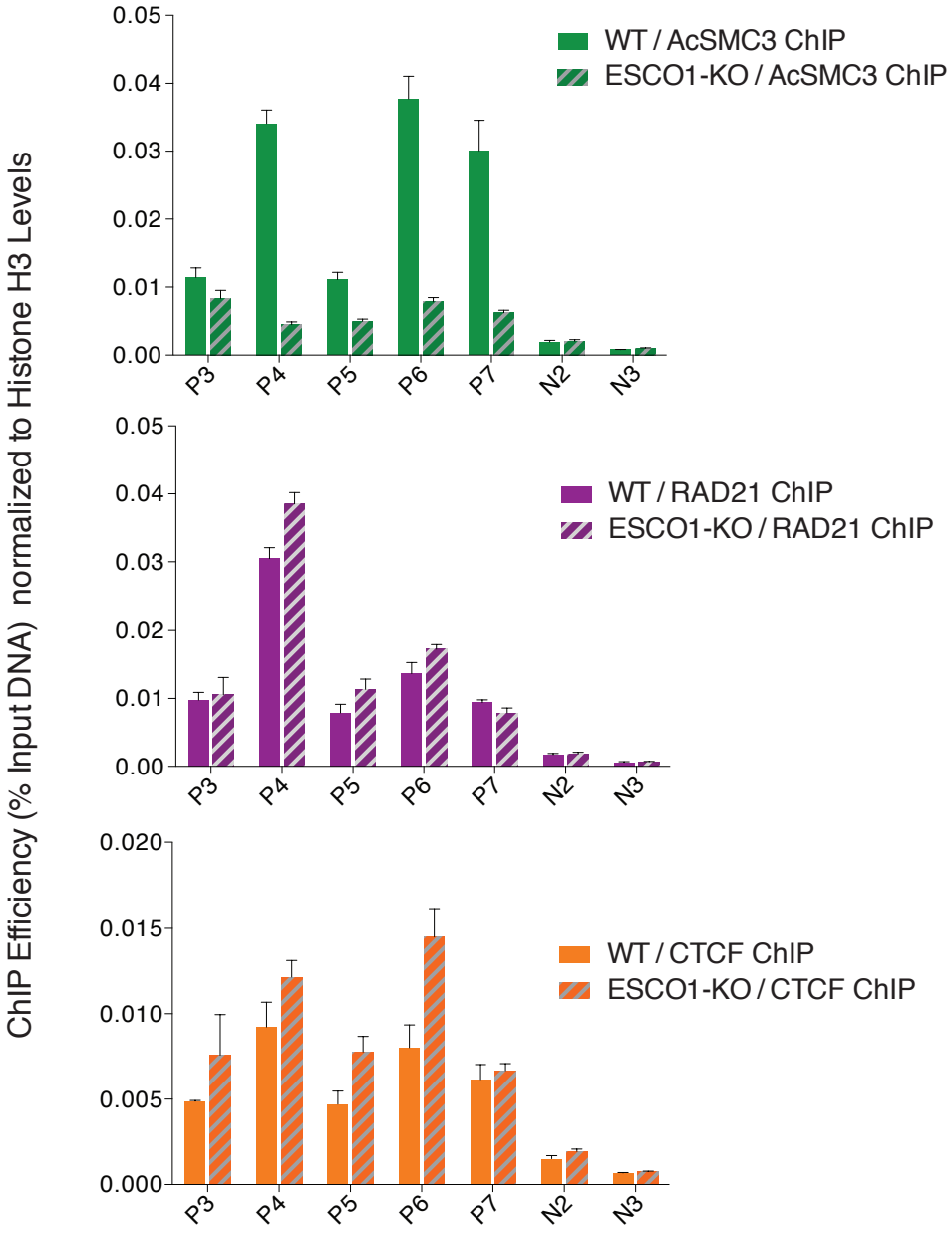


Figure 3.8. Cohesin/CTCF stability is not affected at enrichment sites when acetylation is reduced. ChIP-qPCR using AcSMC3, RAD21 or CTCF antibodies in G1-arrested WT and ESCO1-KO cells shows decreased acetylation at several AcSMC3 sites without negatively affecting RAD21 or CTCF levels at these sites. Signals were normalized to Histone H3 ChIP levels. Error bars represent standard deviation in triplicate qPCR measurements.

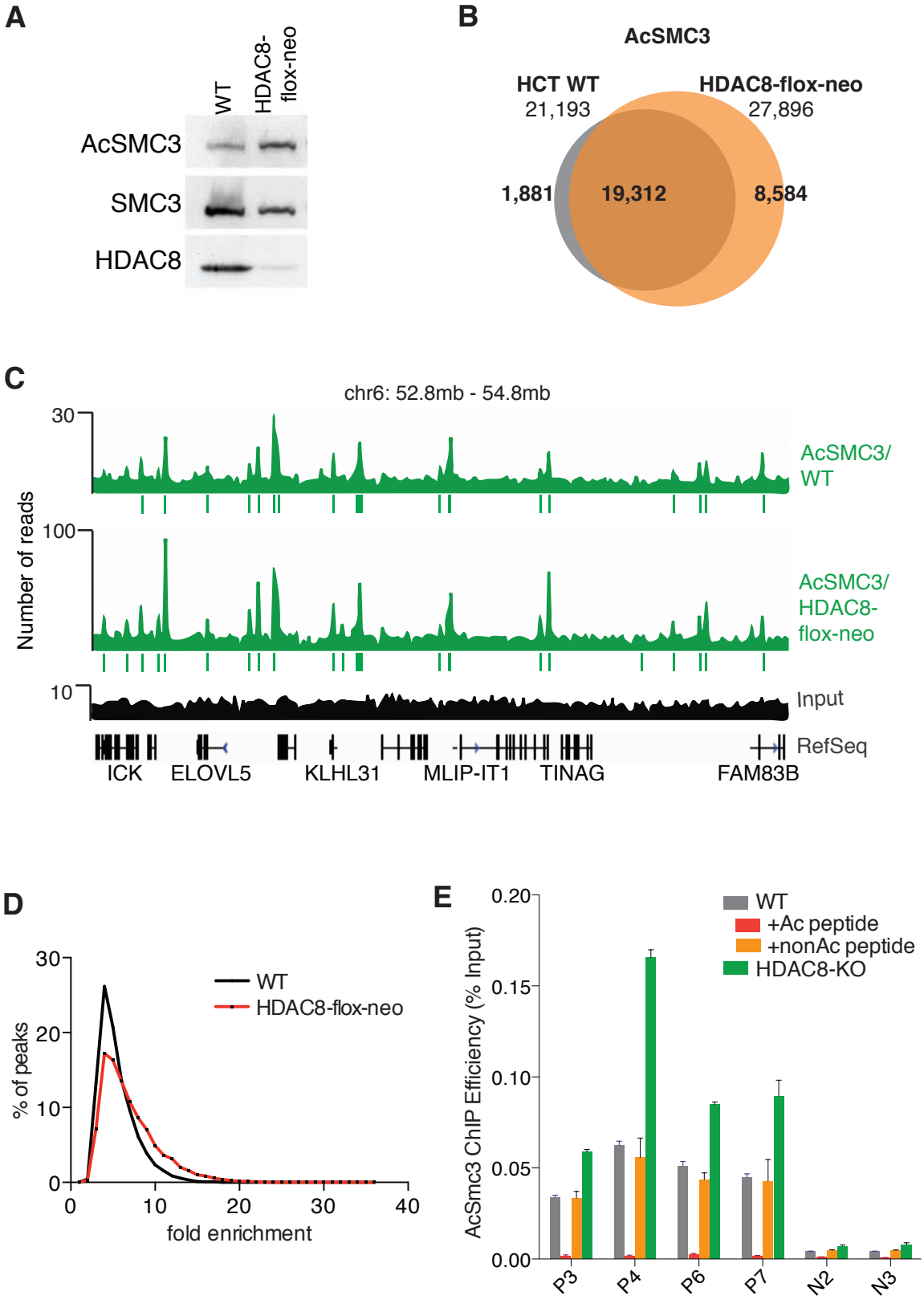

A
WT
HDAC8-flox-neo
AcSMC3
SMC3
HDAC8
B
AcSMC3
HCT WT
21,193
HDAC8-flox-neo
27,896
1,881
19,312
8,584
C
chr6: 52.8mb - 54.8mb
30
100
10
Number of reads
AcSMC3/ WT
AcSMC3/ HDAC8-flox-neo
Input
RefSeq
ICK
ELOVL5
KLHL31
MLIP-IT1
TINAG
FAM83B
D
% of peaks
fold enrichment
WT
HDAC8-flox-neo
E
AcSmc3 ChIP Efficiency (% Input)
WT
+Ac peptide
+nonAc peptide
HDAC8-KO
P3
P4
P6
P7
N2
N3

Figure 3.9. Acetylation peak number and intensity increase in HDAC8-KO cells. (A) Western blot (by M. Jones) showing loss of HDAC8 signal and rise in overall acetylation in HDAC8-deficient cells. (B) Venn diagram shows an increase in number of acetylation peaks called in HDAC8-deficient cells. (C) AcSMC3 ChIP-Seq reads in WT and HDAC8-deficient cells are visualized mapped to a 2mb region of chromosome 6. (D) Histogram shows increase in fold enrichment over input of AcSMC3 peaks in HDAC8-deficient cells compared to WT. (E) AcSMC3 ChIP-qPCR on cells after neo excision (HDAC8-KO) confirms ChIP-Seq data. Acetylated or non-acetylated peptides compete or do not compete the AcSMC3 signal as expected in WT samples. An increase in acetylation levels is observed after HDAC8 loss. P3-P7 are cohesin-positive sites, and N2 & N3 are negative.

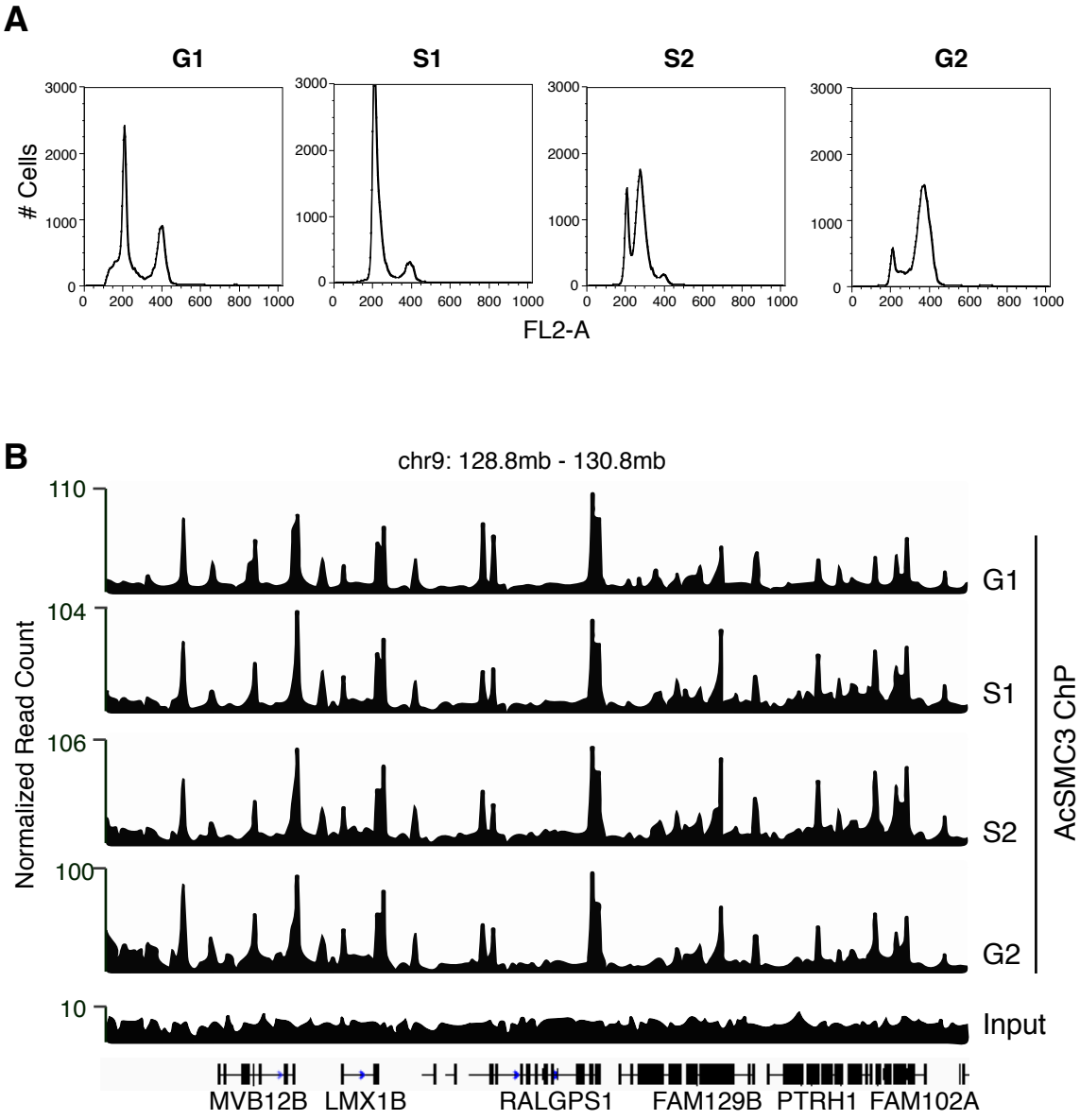


Figure 3.10. Cohesin peaks are acetylated in G1 and remain acetylated at similar levels as cells progress through S-phase. (A) FACS analysis with propidium iodide staining (FL2-A) of samples used in ChIP. Cells were arrested in G1, early S-phase (S1), mid-S (S2) and late-S/G2 (G2). (B) ChIP-Seq reads visualized over 2mb of chromosome 9. Number of reads are normalized to sample with lowest read depth, to compare samples.

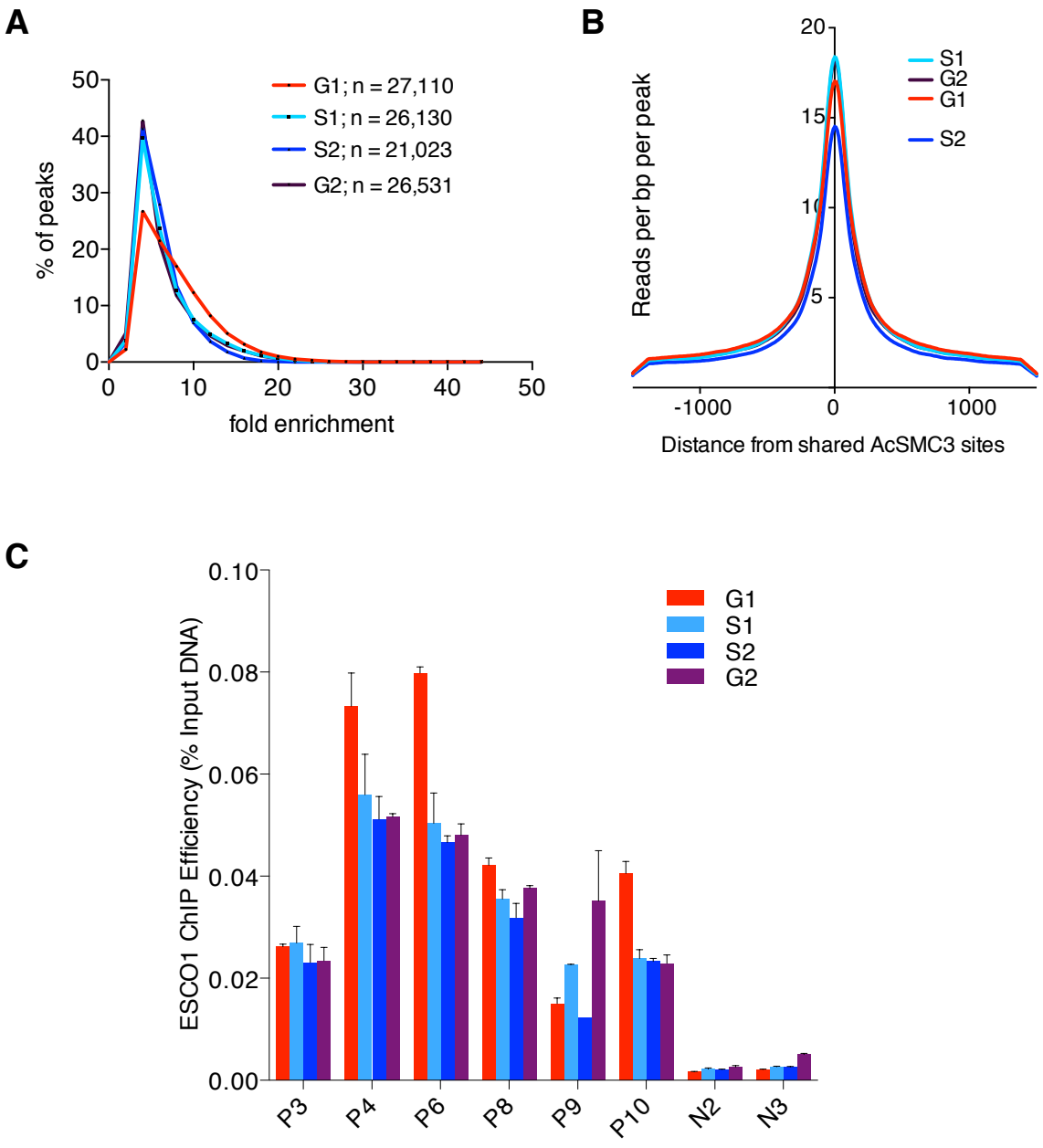


Figure 3.11. Acetylation levels are similar from G1 to G2. Histograms of (A) fold enrichment (over WCE input) and (B) read densities plotted around overlapping peaks in all four time-points. (C) AcSMC3 ChIP-qPCR on time-point samples comparing levels at 6 cohesin sites (P3-P10) and two negative sites (N2, N3). Error bars represent standard deviation of qPCR triplicates.

The difference in localization may be due to targeting by the divergent N-terminal domains, which have been shown to be required for chromatin binding (Hou and Zou 2005). While the C-termini are 77% similar and conserved features found in yeast Eco1 and all other Eco1 homologues, the N-termini are not similar. The role of the N-terminus in ESCO1 targeting is analyzed in Chapter Four.

ESCO1 co-occupies cohesin/CTCF binding sites, as shown by peak overlap and motif analysis (Figure 3.3). It is not surprising that it localizes with its substrate, but ESCO2 does not show the same binding pattern. It is interesting that the two orthologues bind to DNA differently, despite both proteins having similar defects in depletion experiments in acetylation, cohesion, and replication fork speed assays (Hou and Zou 2005, Terret et al., 2009).

ESCO2 is not enriched at discrete sites and may instead localize to repeat regions of the genome that are not mappable by ChIP-Seq. It has been shown recently that mouse ESCO2 localizes by IF and ChIP-qPCR to pericentric heterochromatin (Whelan et al., 2012). It is possible that human ESCO2 also binds to PCH. ESCO2 has also been reported to be enriched in overexpression studies by immunofluorescence in nucleoli (van der Lelij et al., 2009). ChIP-Seq would not show enrichment at these regions containing repetitive sequences.

Alternatively, ESCO2 may bind to DNA dynamically. It may be recruited to and track with replication forks as has been proposed for Eco1 (Lengronne et al., 2006). ESCO2 interacts with PCNA *in vitro* by yeast two-hybrid assay (Moldovan et al., 2006), though whether this is relevant *in vivo* remains unknown. The fact that its expression is regulated so that ESCO2 expression increases and it acetylates cohesin in S-phase may indicate that it has an S-phase specific role, perhaps at the replication fork.

Cohesin acetylation and dynamics

A new monoclonal antibody was generated for this study and acetylation was mapped genome-wide using this highly specific K105-acetyl antibody. This antibody only recognizes modified, acetylated SMC3. This is clear from the antibody's behavior in ELISA and Western blot detection after ESCO1/2 depletion (Figure 3.4). Also, the acetylated peptide competes for the ChIP signal while non-acetylated peptide has no effect (Figure 3.9E). Furthermore, the ChIP signal is diminished in ESCO1-KO cells and increases in HDAC8-KO cells (Figures 3.7 & 3.9).

AcSMC3 mapping reveals that most cohesin sites are acetylated. This is contrary to a recent study that reports only a small subset of cohesin sites as being acetylated by their antibody in the same cell line (Deardorff et al., 2012). The antibody used there preferentially recognizes K106 acetylation whereas the antibody used in this study recognizes K105 rather than K106 mono-acetylation. However, both recognize the double-K105,K106-acetylated form of SMC3 as well. Calling peaks using their raw data and the peak-calling method used here also identifies a small set of peaks, fewer than 4000 despite greater read depth, in their results (data not shown), so this is not due to a difference in data analysis. Whether the difference is due to the difference in antibody recognition or in ChIP conditions is not known.

Cohesin is acetylated in G1 by Western blot, though acetylation levels rise as cells enter and progress through S-phase. I therefore analyzed the dynamics of cohesin acetylation at its ChIP binding sites at various time-points to ask how levels change from G1 through S-phase and into G2. Unexpectedly, these sites are already acetylated in G1 and at similar levels as in S/G2 cells. Acetylation does not increase at these sites as cells go

through S-phase as it does overall in WCE or chromatin fraction by Western blot. This rise may be regulated by ESCO2 (and/or ESCO1) at other locations.

AcSMC3 mapping in ESCO1-KO cells shows that the absence of ESCO1 diminishes acetylation at its binding sites even in G2 cells when there is acetylation by Western blot in these cells. This suggests that ESCO2 cannot completely fulfill the role of ESCO1 in acetylating cohesin at these sites. However, there is some acetylation at these sites in the absence of ESCO1. Perhaps ESCO2 is bound at low levels throughout the genome including at these sites that is not seen as enrichment by ChIP. It can therefore acetylate cohesin to some degree by G2. Alternatively, ESCO2 may travel through these sites, for example riding the replication fork, to acetylate cohesin as it passes through. However, ESCO2 cannot fully take over for the function of ESCO1 to acetylate cohesin at these sites, nor in cohesion establishment function, as knockdown of ESCO1 causes a cohesion defect in HeLa cells.

Conversely, there is an increase in acetylation in HDAC8-deficient cells at cohesin ChIP sites. HDAC8 and its yeast homolog Hos1 have been reported to be important for recycling cohesin for use in the next cell cycle by deacetylating the soluble pool of SMC3 after cohesin's removal from DNA (Borges et al., 2010, Xiong et al., 2010). Although HDAC8 functions on soluble cohesin, loss of HDAC8 has been shown previously to increase chromatin levels of AcSMC3 (Deardorff et al., 2012), and as shown here, increases acetylated cohesin levels at its ChIP sites. This indicates that removal of acetylated cohesin and its reloading is possible when it is still acetylated.

CHAPTER FOUR

Regulation of ESCO1 Binding to DNA

Introduction

ESCO1 and ESCO2 share a conserved C-terminal region with all Eco1-family members, consisting of a PCNA-interacting motif, zinc finger domain, followed by an acetyltransferase domain at the extreme C-terminus (Ivanov et al., 2002, Moldovan et al., 2006). The latter is required for acetylation and cohesion establishment. The PIP-box has been suggested to recruit Eco1 to replication forks in yeast (Lengronne et al., 2006) but its functional significance in the human homologues is not understood.

Eco1 is bound to DNA throughout its expression in the cell cycle (Toth et al., 1999). Both ESCO1 and ESCO2 are associated with chromatin and this association depends on the N-termini in both proteins (Hou and Zou 2005). How these proteins are targeted to DNA and regulated by their cis-elements and by other proteins is not known.

With knowledge of ESCO1 target sites from the ChIP-Seq data, we asked which *cis* domains are required for ESCO1 to bind these sites, and whether this targeting is important for cohesion. Interacting proteins were also studied by *in vitro* binding assays, followed by depletion of positive *in vitro* hits in cells to assay targeting by ESCO1 ChIP-qPCR, to look at their effect on ESCO1 recruitment *in vivo*.

Results

ESCO1 is targeted in G1

Acetylation of cohesin begins in G1 phase by Western, and at discrete DNA binding sites by ChIP (Figure 3.10). Therefore, I asked if ESCO1 is bound to these sites in G1. FLAG-ESCO1 cells were synchronized in G1 or G2 phase by double thymidine block and release and FLAG ChIP-qPCR was performed to assay ESCO1-enriched sites. ESCO1 is observed at similar levels in both G1 and G2 arrested populations (Figure 4.1). Therefore, it is already targeted to its sites on unreplicated DNA before the need for cohesion establishment, in G1.

ESCO1 targeting is cohesin-dependent

As shown in Figure 3.3, ESCO1 and cohesin largely overlap. To determine if ESCO1 is targeted downstream or independently of cohesin, RAD21 was knocked down by siRNA, and GL2 siRNA was used as a control. There was a significant reduction in ESCO1 binding to several assayed binding sites after cohesin depletion (Figure 4.2). Therefore, cohesin is required for ESCO1 enrichment at these sites and either directly or indirectly recruits ESCO1.

N-terminus targets ESCO1 to its binding sites

To test if the conserved domains in the C-terminus are required for targeting to ESCO1 enrichment sites, or if the N-terminus, which is required for chromatin-association (Hou and Zou 2005) is also required for enrichment at these sites, FLAG-tagged N- and C-terminal truncations were expressed in HeLa cells and assayed for binding by ChIP-qPCR. Surprisingly, the C-terminus is dispensable for binding to ESCO1 sites and the N-terminus is necessary and sufficient for binding (data not shown). Therefore, the

domains conserved in the C-terminus from *S. cerevisiae* Eco1 are dispensable for ESCO1 targeting to these sites.

Similarity in vertebrate ESCO1 N-termini

The N-terminus of ESCO1 has no similarity to ESCO2 N-terminus and has been reported to not have similarity with its homologues. However, because this region is required for targeting to ESCO1 sites in human cells, we asked if there are regions that share similarity in vertebrate homologues. A closer inspection of the N-terminus by comparison to mouse and *Xenopus laevis* Eco1 revealed that there are several small regions of similarity in the N-termini (Figure 4.3). These regions are called A (comprised of 25 amino acids), B (5aa), G (5aa) and D (20aa). We wanted to study their contribution to ESCO1 function by deletion analysis.

N-terminal regions are required for ESCO1 targeting

The vertebrate-specific N-terminal regions were deleted from ESCO1 constructs and expressed in HeLa cells to assay binding by fractionation and ChIP-qPCR. Chromatin-fractionation shows that all deletion mutants still bind to chromatin (Figure 4.4A). Fragments of the N-terminus and C-terminus were also expressed and used as controls, and behaved as expected with the former targeted to the chromatin fraction and the latter targeted exclusively to the soluble fraction.

To then determine whether these regions are required for ESCO1 enrichment at its ChIP binding sites, deletion mutants were assayed by ChIP-qPCR. Three of these regions, A, G and D, are required for ESCO1 to target to enrichment sites, while deletion of B did not affect targeting and bound ESCO1 sites at similar levels to the wild-type protein.

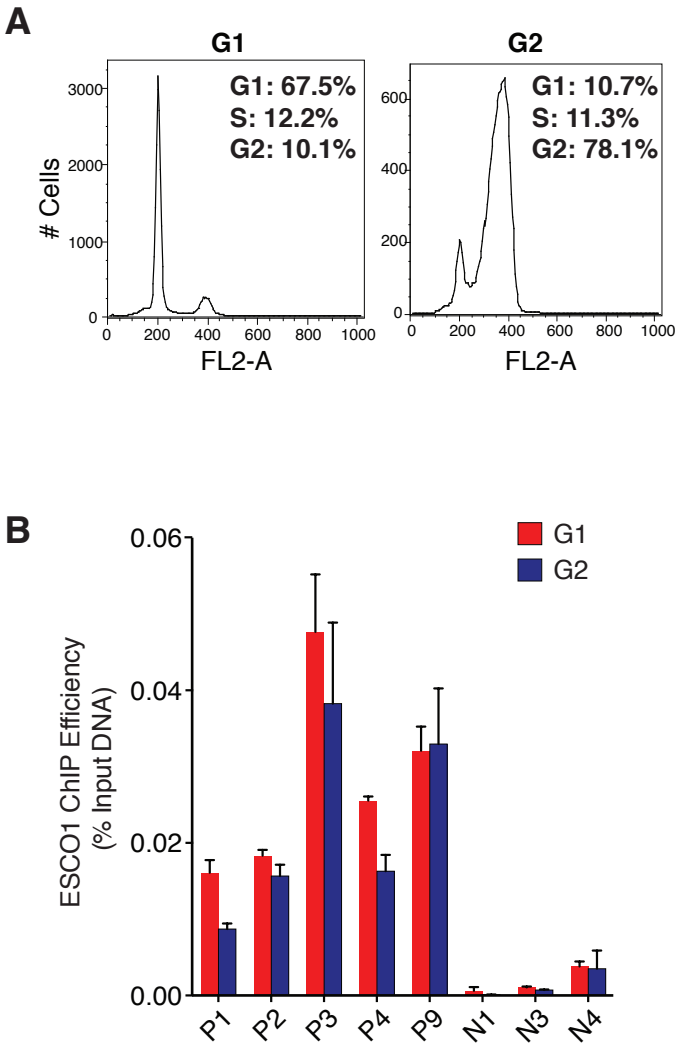


Figure 4.1. ESCO1 is targeted to enriched sites in G1. (A) FLAG-ESCO1 HeLa cells were arrested in G1 or G2, following a double thymidine block and release of 13 hours or 5 hours, respectively. FACS shows propidium iodide staining (FL2) and enrichment of G1, S and G2 populations quantified by Dean-Jett-Fox method. (B) FLAG ChIP-qPCR on ESCO1-enriched sites shows no significant differences in ESCO1 levels at enriched sites in G1 and G2 populations. Five positive sites (P1-P9) and three negative sites (N1-N3) were assayed. Error bars represent standard deviation of triplicate qPCR measurements.

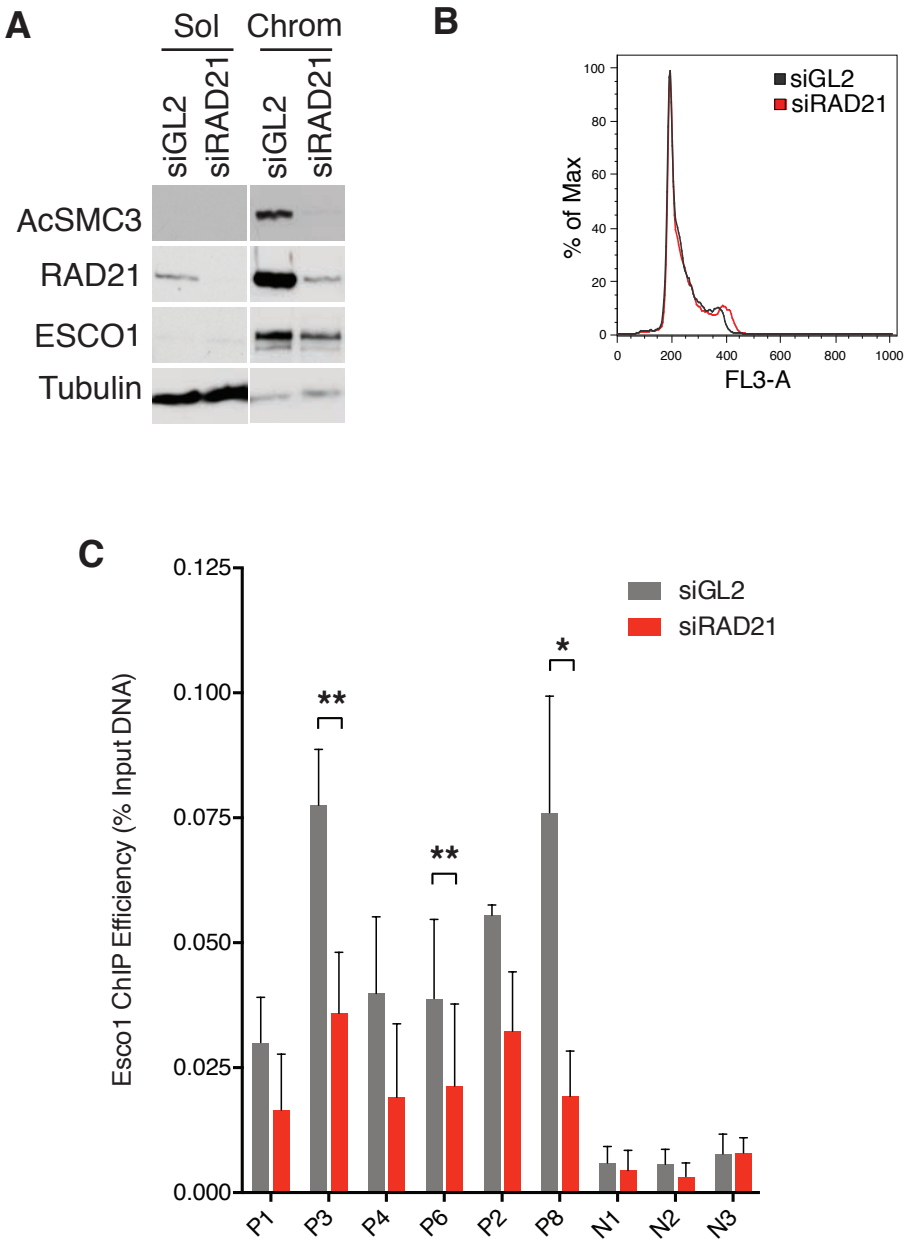


Figure 4.2. ESCO1 is targeted to enriched sites in a cohesin-dependent manner. (A) Western blots on NETN-lysed soluble and chromatin fractions following 56-hour depletion of RAD21 with 16-hour thymidine block before harvest. (B) FACS profiles of RAD21-KD vs control GL2-KD showing similar cell cycle profiles. (C) ESCO1 ChIP-qPCR on five positive and three negative ESCO1-binding sites shows decrease in binding after depletion of RAD21. Error bars represent standard deviations from three independent experiments. Significance calculated using two-tailed Student's t-test; *P<0.05, **P<0.005.

A

```
X.laevis     MAQKRKLSQAESSAKHQKLDKDHKHPSVATKNANGKSKTSTANHASNKKINALKPHGKANTPTKKAIKLSPQKILQKTRNELKSAKTSKEGPTSSRTTQGPLYSVQASKASKRGKADVGK   120
M.musculus   ------------------------------------------------MSIQEKSKENSS--IVTKESEDENLEEEVESSQNSPTKKSGSKEAVKT--P---VRFSNKSKTNESEFGM    63
H.sapiens    ------------------------------------------------MMSIQEKSKENSS--KVTKKSDDKNSETEIQDSQKNLAKKSGPKETIKS--Q---AKSSSESKINQPELET    64

X.laevis     KASTRANVQKSKNISNVGKNLVKNTSSARLKVSASGKNEESKTGTKRSSRLLELKHGKAQSKEPPKNSKKDIHSSCTKTEPLKGRSRLVRKTAQNKGSNSEHRGNADRKSIQQIPSSK-K   239
M.musculus   RMSTRSASCSADKT--ATNSFNKN------TVTLKGQSQESSKTKKLCQEK--LSL----------------------GILKGNEQLHRRS---------------QRLQQLTECTTR   134
H.sapiens    RMSTRSSKAASNDK--ATKSINKN------TVTVRGYSQESTK-KKLSQKK--LVH----------------------ENPKANEQLNRRS---------------QRLQQLTEVSRR   134

X.laevis     SAVIGKKQAQIRTVTHAQPKESSSKKHSKQLQATKKLSQAKTANGRQSKRARPLSPVNRKNKGTQGHQGPGIKTAIKRSDLRTGVKGPNDEKKLKAIQNDGMMEIPLHPNTDVTENEVDT   359
M.musculus   SLRSREIHGQIQTVKQNQQSARREQC----------------NSTQSKCN--------KVKVNQK--------HVKRKVLEIKSDCKED----RHSVTNEVINSPKGKKRKVQHQTTST   217
H.sapiens    SLRSREIQGQVQAVKQSLPPTKKEQC----------------SSTQSKSN--------K--TSQK--------HVKRKVLEVKSDSKED----ENLVINEVINSPKGKKRKVEHQTACA   215

                                                                                                                          A
X.laevis     CSHSSPSLKDSGASFS-------KKPLE--KSQRPKQKRSSIVASAKKKSPSAKPPHKKLTV--KRQKTGHSGTAKNTSAPTLPDVPTAKRISILDLCNEIAGEIESDTVEVKKDMPSSQ   468
M.musculus   CSSQC--NQGSEKCLQKTSRKEEIKPVPVTADIRKLKAATSVVSK---KNELRKSAHTQVSTSTKRPQIPLPLVPEHSDDQELEQAGKSKRGSILQLCEEIAGEIESDTVEVKKESSCVE   332
H.sapiens    CSSQC--TQGSEKCPQKTTRRDETKPVPVTSEVKRSKMATSVVPK---KNEMKKSVHTQVNTNTTLPKSPQPSVPEQSD-NELEQAGKSKRGSILQLCEEIAGEIESDNVEVKKESSQME   329

                  B                 G                 D
X.laevis     DIQKEEHPIVENKESAQPEAVNVCHTGEKEETQSKRFFSSKTVVRLKCKLDRKNSPAPKNSKWNKIKLKKANNLSGMNVPKIHAVLPNLDVIKARSK--LTQTGHSVTGT-------SDK   579
M.musculus   -SVKEEKPAEVKLQGTDAERQILHHKEANQDVRSNRFFPSRKTKPVKCVLNGINSSTKKNSNWTKIKLSKFNSVQQHKLDS--QVSPKLNLLQTGLSTSVLEMPHPVSQSTFLEMKAH-G   448
H.sapiens    -SVKEEKPTEIKLEETSVERQILHQKETNQDVQCNRFFPSRKTKPVKCILNGINSSAKKNSNWTKIKLSKFNSVQHNKLDS--QVSPKLGLLRTSFSPPALEMHHPVTQSTFLGTKLHDR   446

X.laevis     P--LDKKLPAGINSKSTLPDLKKSKSAAIIQKTSQEKAENGLLDKHINHDLELTFDE-----------GFKLHLDSSPESSPLKKPPMKPQQLEPEHKKPDVIASKACASKQLFNSSLR   685
M.musculus   NVTCQRDKMKGIKSE----EVKINNIAIEINKATKRDPGNCNLDNHIKPSPDSSLDNQMKLSCESAPDQNFSICSASEVETNPLENTAAASTLLSQAKIDED-RTFPGSAP-NQQHSVLS   562
H.sapiens    NITCQQEKMKEINSE----EVKINDITVEINKTTERAPENCHLANEIKPS-DPPLDNQMKHSFDSASNKNFSQCLESKLENSPVENVTAASTLLSQAKIDTGENKFPGSAP--QQHSILS   559

                                  PIP                 ZF
X.laevis     DQMAPSDEGRNTVTSHVSSAKSTCVLSEVNIQKEVKKLKEAEKDSNKQSIIDAGQKRFGAISCNVCGMLYTASNPEDEAQHLLFHNQFISAVKYVGWKKERIVSEYPDGRIIMVLPDDPK   805
M.musculus   DEASINRKNRDVPPNHSQL-K-HDSHLEITIP-KSLKLKDSEKVDEKQLVIDAGHKRFGAVSCNICGMLYTASNPEDETQHLLFHNQFISAVKYVGWKKERILAEYPDGRIIMVLPEDPK   679
H.sapiens    NQTSKSSDNRETPRNHSLP-K-CNSHLEITIP-KDLKLKEAEKTDEKQLIIDAGQKRFGAVSCNVCGMLYTASNPEDETQHLLFHNQFISAVKYVGWKKERILAEYPDGRIIMVLPEDPK   676

                                                                                                                        ACT
X.laevis     YALKKVDEIREMVDNDLGFQQVPLRLHSRTKTLLFISSDKKVAGCLIAEHIQWGYRVIDDLIPEGTSEKEKALSERVKAWCCSTSPEPAICGVSRIWVFSMMRRKKIASRMLECLRNNFI   925
M.musculus   YALKKVDEIREMVDNDLGFQQAPLMCYSRTKTLLFISNDKKVVGCLIAEHIQWGYRVIEEKLPVIRSEEEKVRFERQKAWCCSTLPEPAICGISRIWVFSMMRRKKIASRMIECLRSNFI   799
H.sapiens    YALKKVDEIREMVDNDLGFQQAPLMCYSRTKTLLFISNDKKVVGCLIAEHIQWGYRVIEEKLPVIRSEEEKVRFERQKAWCCSTLPEPAICGISRIWVFSMMRRKKIASRMIECLRSNFI   796

X.laevis     YGSFLNKDEIAFSDPTPDGKLFATRYCGTSQFLVYNFISGHS--    967
M.musculus   YGSYLSKEEIAFSDPTPDGKLFATQYCGTGQFLVYNFINGQNTT    843
H.sapiens    YGSYLSKEEIAFSDPTPDGKLFATQYCGTGQFLVYNFINGQNST    840
```

B

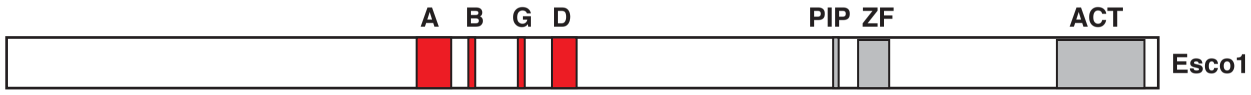


Figure 4.3. Similar N-terminal regions in vertebrate ESCO1 homologs. (A) Xenopus X.Eco1, mouse and human ESCO1 alignment highlighting C-terminal conserved PIP box, zinc finger (ZF), and acetyltransferase (ACT) domains in gray. Regions of similarity in N-terminus (named A, B, G & D) are highlighted in yellow. (B) Schematic shows C-terminal domains and the shared vertebrate-specific N-terminal regions.

Some targeting mutants are defective in cohesion

Next, the functional importance of these regions for cohesion establishment was assayed by testing the ability of siRNA-resistant ESCO1 mutants to rescue endogenous ESCO1 depletion, which causes a defect in cohesion. Chromosome spreads were prepared and scored as having partial (sister chromatids that are still aligned but have lost centromeric constriction) or severe (scattered single chromatids unaligned to sisters) for cohesion defect. An empty vector line was used as a control.

WT-ESCO1, and ΔB and ΔG deletions rescued the defect observed in the empty-vector cell line, while ΔA and ΔD did not rescue the cohesion defect. Therefore, some targeting mutants are unable to establish cohesion, indicating that enrichment at ESCO1 binding sites may be important for its function in cohesion.

ESCO1 interacts with CTCF and cohesion-pathway proteins

ESCO1 may directly interact with DNA, or the interaction may be mediated by other proteins. As cohesin is required for targeting to its binding sites, we wanted to determine if it interacts with cohesin or related proteins in the cohesion pathway. *In vitro* binding experiments were performed with bacterially purified N-terminus of ESCO1 (fragment N1) and several ^{35}S-methionine labeled *in vitro transcribed/translated* candidate proteins at cohesin/CTCF sites.

Some of these prey proteins had background binding to beads and it was therefore inconclusive whether they interact with ESCO1-N1 (data not shown). Among the others, SMC3, SA1 and WAPL did not interact (data not shown), while SMC1A, PDS5B and CTCF were positive for interaction with the ESCO1-N terminus (Figure 4.5). The interaction between PDS5B was further narrowed down to its C-terminal region, beginning at the end of its HEAT repeat region, about the final quarter of the protein (331

amino acids). This region contains two AT-hooks and lacks similarity with PDS5A, unlike the N-termini of these proteins, which are similar and contain a series of HEAT repeats (Zhang et al., 2009). The C-terminus of PDS5A was then tested and failed to interact with ESCO1-N1 in this assay (data not shown).

The vertebrate-specific N-terminal regions of ESCO1 were next tested for binding to these proteins. Deletions of these regions were expressed in ESCO1-N1 and purified from bacteria. As shown in Figure 4.5, the various deletions did not have a large effect on binding to SMC1A. Binding to PDS5B (C-terminus) appeared to be reduced partially with several deletions (Figure 4.5). CTCF interaction was similar to WT for all deletions except there was a decrease with the ΔB mutation. Therefore, region B may be important for direct interaction with CTCF.

Combined deletions of these regions may have a stronger defect than the individual deletions. To test this, two fragments of ESCO1-N1 were purified, N1N and N1C , the latter containing all of these vertebrate-specific regions (Figure 4.6A). PDS5B is able to bind both fragments, though with slightly lower affinity, while binding of SMC1A and CTCF is relatively lower than WT with N1N but not N1C (Figure 4.6B). The defect of CTCF interaction with N1N is especially strong. Therefore, it appears that the stretch containing the vertebrate-conserved regions interacts with SMC1A and CTCF, while PDS5B may interact with the 118 amino acid overlapping region in these fragments.

Effect of depleting ESCO1-interacting proteins on ESCO1 targeting

In order to test if these *in vitro* interactions are relevant *in vivo* for ESCO1 targeting, knockdown of these proteins in HeLa cells was followed by ESCO1 ChIP-qPCR. First, both orthologues of PDS5 were knocked down. RAD21 depletion was used as a positive

control and parallel ChIP of Histone H3 was performed for each sample to normalize

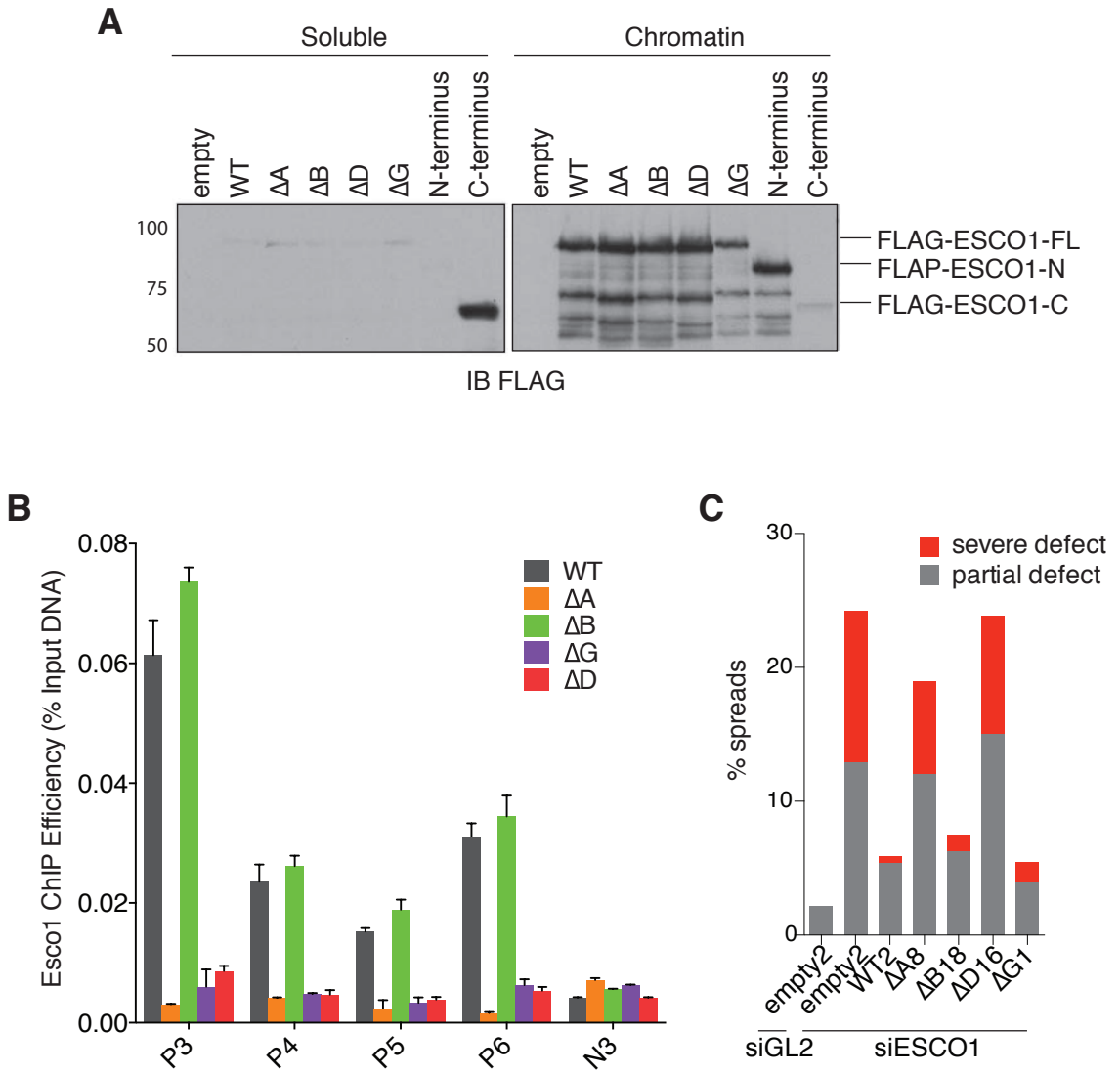


Figure 4.4. N-terminal regions A, D and G are required for ESCO1 enrichment at discrete sites, and regions A and D are required for cohesion. (A) ESCO1 constructs were expressed in HeLa cells and lysed to isolate soluble and chromatin fractions and immunoblotted with FLAG antibody. The N-terminal fragment and small N-terminal deletions all localize to chromatin, while the C-terminal fragment alone is unable to bind chromatin. (B) ChIP-qPCR on enriched sites shows that ΔB deletion targets to enriched sites but ΔA, ΔG, and ΔD are not enriched at these sites. Four positive sites and one negative site (N3) are shown. (C) Complementation assay testing ability of siRNA-resistant deletion mutants to rescue cohesion phenotype after depletion of endogenous ESCO1 shows rescue by WT, ΔB and ΔG, but ΔA and ΔD. Spreads were scored as partially defective for cohesion if sister chromatids were aligned but separated with no centromeric constriction, and severely defective if they were scattered as single chromatids.

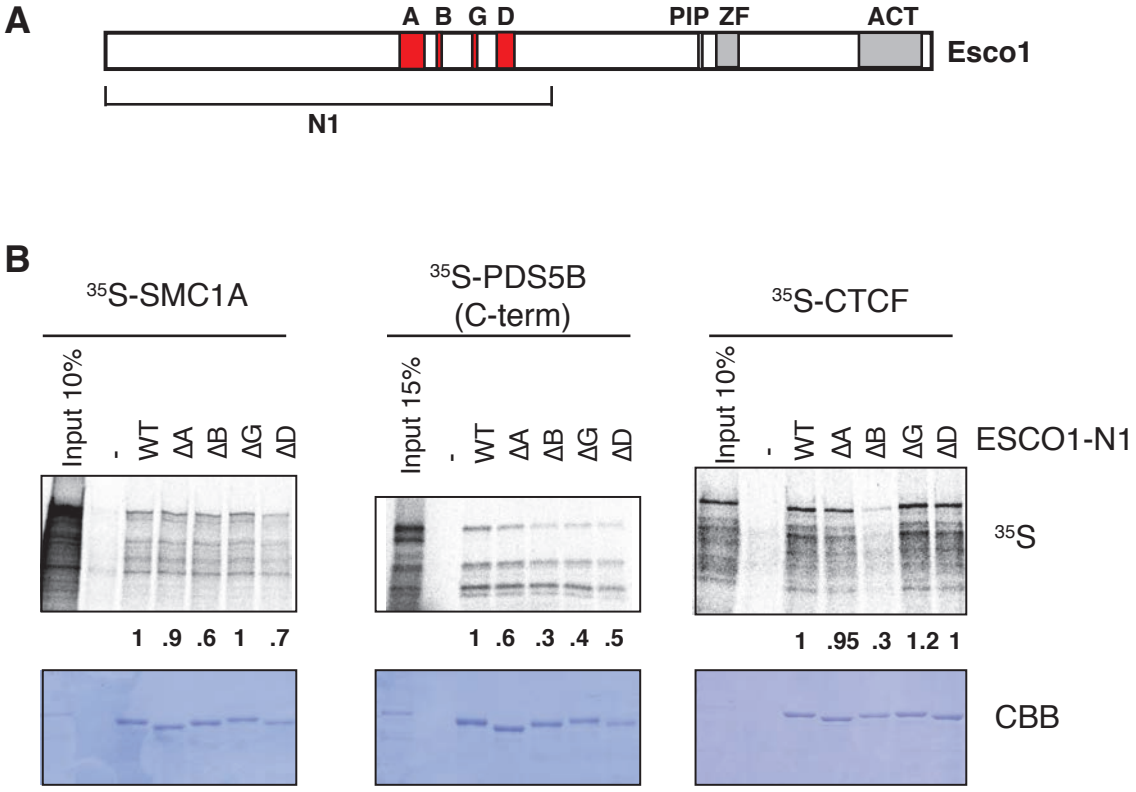


Figure 4.5. Interactions of ESCO1 N-terminal domains with SMC1A, PDS5B and CTCF. (A) Schematic of ESCO1-N1 fragment and deletion regions used for binding assays. (B) Binding was performed with bacterially purified ESCO1-N1 bait and ^{35}S-labeled SMC1A, the C-terminus of PDS5B, and CTCF. Binding is quantified for each lane, normalized to WT, and adjusted to bait protein levels in coomassie brilliant blue (CBB) staining.

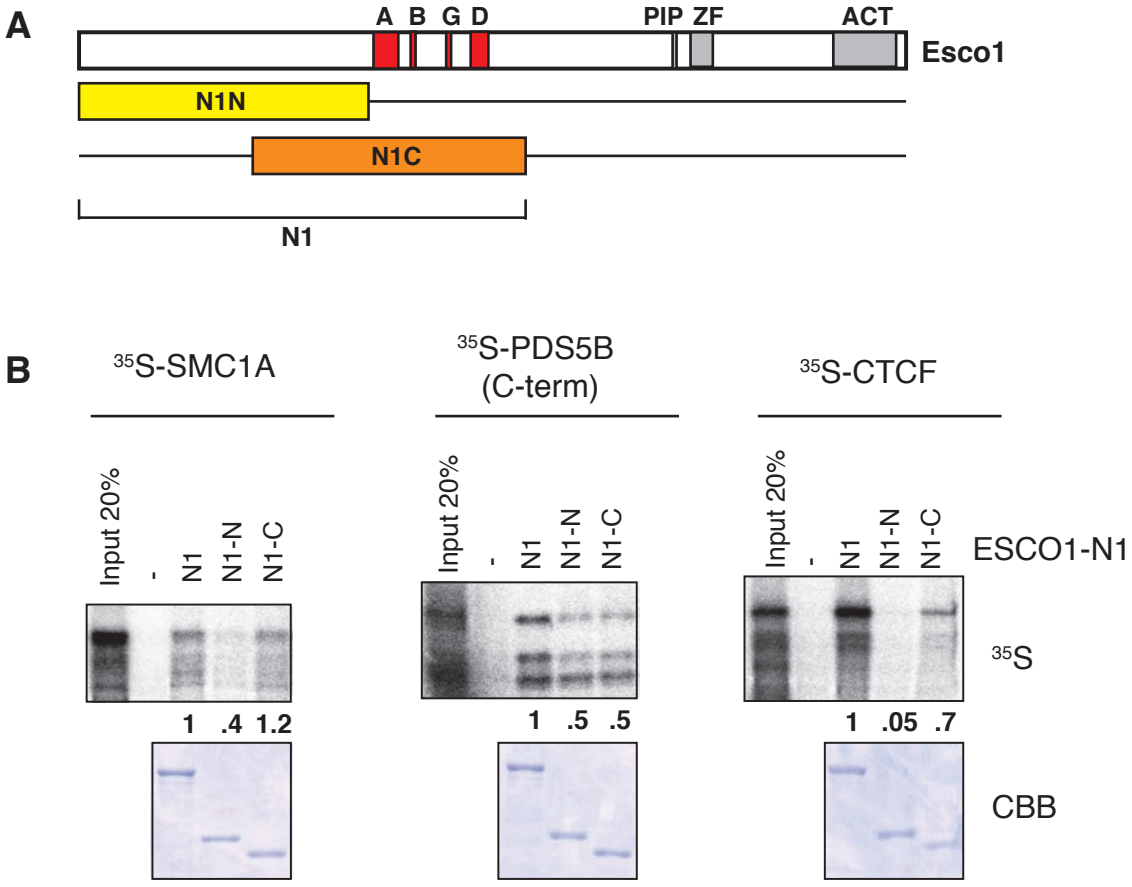


Figure 4.6. Interactions of ESCO1 N-terminal fragments. (A) Schematic of ESCO1 N-terminal fragments used for binding assays. N1N and N1C fragments share 118 amino acid overlapping region. (B) Binding is partially reduced with N1N relative to entire N-terminus (N1) with SMC1A and severely affected with CTCF. PDS5B binding is partially reduced with both fragments. Quantifications for each lane are shown below ^{35}S image and are adjusted to bait protein levels.

levels to H3 ChIP. Note that depletion of PDS5A/B results in decreased overall acetylation of cohesin as observed by FACS analysis of AcSMC3-antibody staining (Figure 4.7B). qPCR shows a decrease in ESCO1 binding to its enrichment sites after PDS5A/B knockdown (Figure 4.7C). However, the decrease is relatively small, but this may be explained by inefficiency of PDS5B depletion as seen in the immunoblot. The incomplete depletion likely masks a more severe effect of PDS5 loss on ESCO1 recruitment.

Knockdown of CTCF resulted in a decrease in ESCO1 recruitment as well, similar in degree to RAD21/SMC3 depletion (Figure 4.8). This could either indicate direct ESCO1 recruitment by cohesin, which also depends on CTCF for its enrichment at these sites (Wendt et al., 2008), or by CTCF, or both of these factors.

WAPL depletion was also performed in this experiment (Figure 4.8), and results in an increase in ESCO1 targeting. An increase in cohesin stability after WAPL depletion (Tedeschi et al., 2013) at these sites could account for this increase. This would further support a role for cohesin (or CTCF) in recruiting ESCO1 to these sites.

Discussion

G1 targeting and dependence on cohesin

Here, I show that ESCO1 is targeted to its binding sites already in G1, when it binds at similar levels as in G2 cells. This G1 targeting is likely responsible for acetylation of cohesin at these sites in G1 (Figure 3.10). Because ESCO1 binds these sites and acetylates cohesin on unreplicated DNA prior to a need for cohesion, there may be another function of this targeting and/or acetylation. A role for ESCO1 in transcription is explored in Chapter Five.

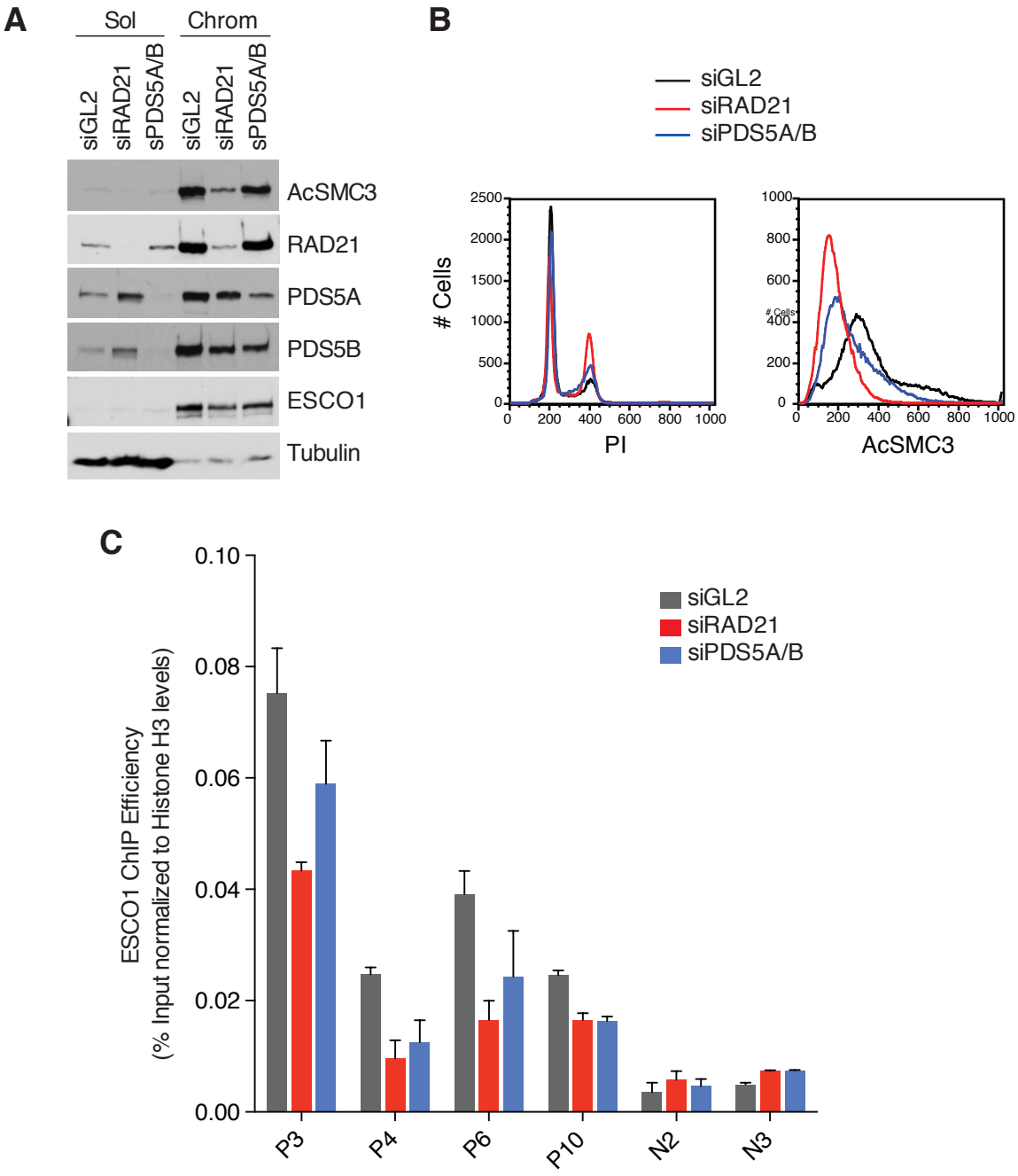


Figure 4.7. Effect of PDS5 depletion on ESCO1 targeting. (A) GL2, RAD21 and PDS5A&B were depleted by siRNA in FLAG-ESCO1 HeLa cells and harvested 48 hours after transfection. Cells were NETN-lysed to separate soluble (Sol) and chromatin (Chrom) fractions for immunoblotting. (B) FACS on KD samples shows similar PI profiles and changes in AcSMC3 staining; RAD21 depletion reduces AcSMC3 staining intensity as does PDS5A&B depletion. (C) ChIP-qPCR was performed in KD samples using FLAG antibody as well as Histone H3 antibody to normalize levels across samples. Four positive sites and two negative sites for ESCO1 binding were assayed. Error bars represent standard deviation in triplicate qPCR measurements.

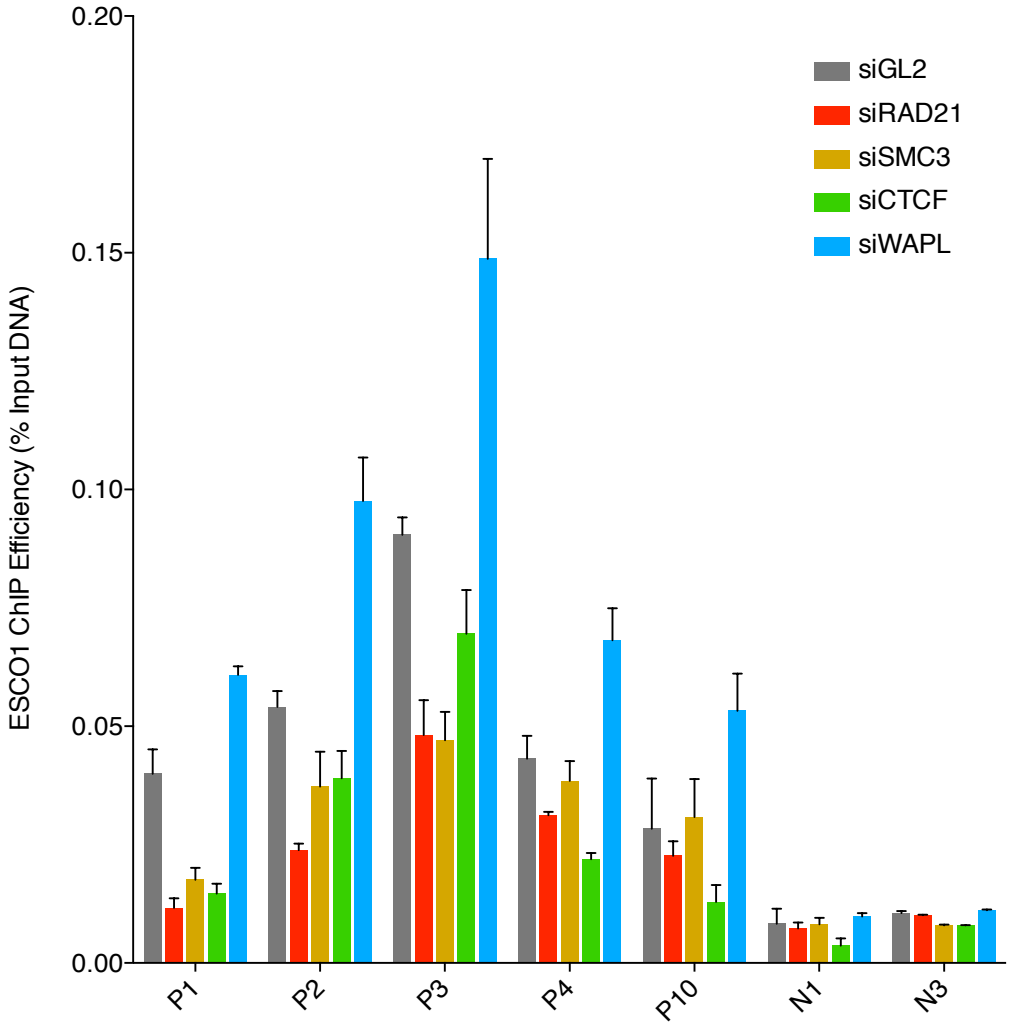


Figure 4.8. Effect of cohesin, CTCF and WAPL depletion on ESCO1 targeting. Cohesin subunits RAD21 and SMC3, CTCF, and WAPL were depleted in FLAG-ESCO1 expressing HeLa cells and ChIP was performed with anti-FLAG antibody. Five ESCO1-positive sites (P1-P10) and two negative sites (N1, N3) were assayed by qPCR, showing a decrease with cohesin and CTCF knockdown, and an increase with WAPL knockdown, of ESCO1 targeting to its binding sites.

The targeting of ESCO1 depends on cohesin, as ESCO1 levels at cohesin sites are attenuated when cohesin is depleted (Figure 4.2). ESCO1 could be recruited directly by cohesin, or it could be through factors on which cohesin also depends for its recruitment, such as CTCF. It has been shown that cohesin depletion also affects CTCF positioning, as depletion of cohesin results in a ~50% decrease in CTCF binding at these sites (Wendt et al., 2008). Therefore, this cohesin KD - ESCO1 ChIP assay does not distinguish between recruitment by cohesin or recruitment more directly by CTCF. It remains possible that ESCO1 is recruited directly by CTCF with which it interacts *in vitro* (Figure 4.5), or it may be recruited by both CTCF and cohesin.

Role of vertebrate-specific N-terminal domains

Although most of the N-terminus is not similar between ESCO1 homologues, there is similarity in several regions between vertebrate N-termini (Figure 4.3). When some of these regions (A, G, D) are deleted, they no longer target to normal ESCO1 enrichment sites although they all bind chromatin at levels comparable to WT (Figure 4.4). Therefore, all deletions bind chromatin but regions A, G and D contribute to ESCO1 enrichment at cohesin/CTCF sites. Deletions of these regions result in a displacement of ESCO1 from its normal enriched sites.

Two out of these three targeting-defective regions (A and D) are required for proper cohesion establishment (Figure 4.4C). This may be due to the defect in targeting in these deletion mutants. However, ESCO1-ΔG should also have a cohesion phenotype in this case. If targeting is important for cohesion function, the lack of a defect in the ΔG deletion mutant may be due to either a problem with its ability to ChIP, because of altered protein conformation or a difference in cross-linking efficiency. Alternatively, targeting to these sites may not be important for ESCO1 role in cohesion. Regions A and D may have a cohesion defect because of a functional role in this process.

ESCO1 N-terminal interactions with SMC1A, CTCF, and PDS5B

The *in vitro* binding assays show that the N-terminus of ESCO1 interacts with at least three factors at its binding sites: SMC1A, PDS5B, and CTCF. *In vivo* depletion followed by ESCO1 ChIP shows that PDS5 has a slight effect on ESCO1 recruitment (although its knockdown efficiency is poor), while both cohesin and CTCF depletion have a stronger effect. ESCO1 may have multiple modes of recruitment to these sites, with PDS5 as well as cohesin and/or CTCF.

Recruitment of cohesin acetyltransferases by PDS5 would explain a defect of yeast Pds5 mutants to acetylate cohesin, as has been observed in several studies (Vaur et al., 2012, Chan et al., 2013), as well as a defect in acetylation in mouse PDS5-KO MEFS (Carretero et al., 2013). A defect in acetylation is also shown here in human cells after PDS5A/B depletion (Figure 4.7B). However, a mechanism for how PDS5 affects cohesin acetylation is not clear. The biochemical evidence for direct interaction of PDS5 with cohesin's acetyltransferase, and therefore a role of PDS5 in ESCO1 recruitment to its substrate presented in this study, would explain this acetylation defect.

CHAPTER FIVE

ESCO1 and ESCO2 Regulate Gene Expression

Introduction

With the discovery that cohesin co-localizes with CTCF in mammalian cells (Parelho et al., 2008, Wendt et al., 2008), there have been many recent studies on transcriptional regulation by cohesin and cohesin now has a well-established role in transcription. Cohesin mutations in CdLS do not cause cohesion defects and instead cause developmental phenotypes due to altered gene expression.

Other cohesion-pathway proteins have also been shown to be involved in transcriptional regulation. WAPL-deficient MEFS have altered gene expression profiles by microarray in comparison to WT MEFS (Tedeschi et al., 2013). Loss of the cohesin deacetylase HDAC8 also appears to change gene expression, as CdLS patient cells with HDAC8 mutations group with NIPBL-mutant patient cell lines based on gene expression profiling of a CdLS classifier gene set (Deardorff et al., 2012). This is suggested to be due to altered cohesin targeting in the HDAC8-mutant cells.

Roberts Syndrome (RBS) is instead believed to be caused by cohesion and cell cycle defects, according to studies of RBS models of zebrafish and mouse that have mild cohesion defects and increased apoptosis (Monnich et al., 2011, Whelan et al., 2012). RBS patient cells also have a mild cohesion defects at centromeric regions as observed by heterochromatic repulsion (Vega et al., 2005). A transcriptional defect in these cells has therefore been discounted.

However, the findings in this study that acetylation of cohesin and ESCO1 targeting to its discrete binding sites occurs on unreplicated DNA, as well as the direct interaction between CTCF and ESCO1, prompted me to investigate a role for the cohesin acetyltransferases in transcription. Here, I analyze ESCO2 binding sequences and gene expression patterns after ESCO1 loss by microarray.

Results

ESCO2-enriched sites contain REST motif

Although there are few sites of ESCO2 enrichment, these sites have very high levels of ESCO2 binding (Figure 3.1). I analyzed these ESCO2-bound sequences using HOMER *de novo* motif analysis. The RE1 silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF) motif, RE1/neuron-restrictive silencer element (NRSE), is enriched in these sequences, found in about 80% of ESCO2 consensus sites (Figure 5.4A,B). REST is a zinc finger transcription factor that blocks transcription of its target genes by binding to RE1/NRSE sequences found in their regulatory regions, and it generally functions to repress neuronal genes in non-neuronal cell types (Coulson 2005, Lunyak and Rosenfeld 2005).

Analysis of the genes near ESCO2 binding sites shows that nearly all of the genes have processes in neuronal development or function, as would be expected with REST-bound genes. Table 5.2 lists the genes within 5 kb of ESCO2 binding sites and the presence of the REST motif near these genes as annotated in a transcription factor binding database, showing that nearly all genes contain the REST motif.

ESCO1/2 are required for repression of REST target genes

Visualization of ESCO2 reads mapped to the genome shows that ESCO1 peaks co-occupy most of these ESCO2 sites. Some of these sites are shown in Figure 5.4C. To test if these genes are regulated by ESCO1/2, these factors were depleted in HeLa cells and GL2 was depleted as a negative control, followed by RT-qPCR to analyze levels of ESCO2 bound/nearby genes. ESCO1 and ESCO2 knockdown resulted in upregulation of all of these genes (Figure 5.4D). This indicates that ESCO1 and ESCO2 repress expression of these genes.

ESCO1-KD microarray reveals altered gene expression patterns

ESCO1 has many binding sites genome-wide, as shown in Chapter Three. To determine if it regulates genome-wide transcription, ESCO1 knockdown-microarray was performed in HeLa cells, with mock-transfected cells used as negative control and RAD21 knockdown as a positive control. Triplicate knockdowns were performed for each and cells were harvested with a thymidine block and release to minimize cell cycle differences. The protein levels as measured in a parallel knockdown for each siRNA show efficient depletion and FACS shows similar cell cycle profiles across different knockdowns (Figure 5.1A,B).

The microarray analyzed by unsupervised clustering of samples with principal component analysis (PCA) mapping separates control, RAD21 and ESCO1 knockdown samples on one axis (PC#3) (Figure 5.1C), indicating that there are differences in gene expression and these are consistent across biological replicates. Hierarchichal clustering was performed with genes that have a change in expression with FDR ≤ 0.05 in ESCO1 depletion samples compared to control cells. 548 genes meet this criteria and most are upregulated after ESCO1 depletion, with some changing in the same direction and some in the opposite direction after RAD21 depletion (Figure 5.2A). As shown in Figure 5.2B of

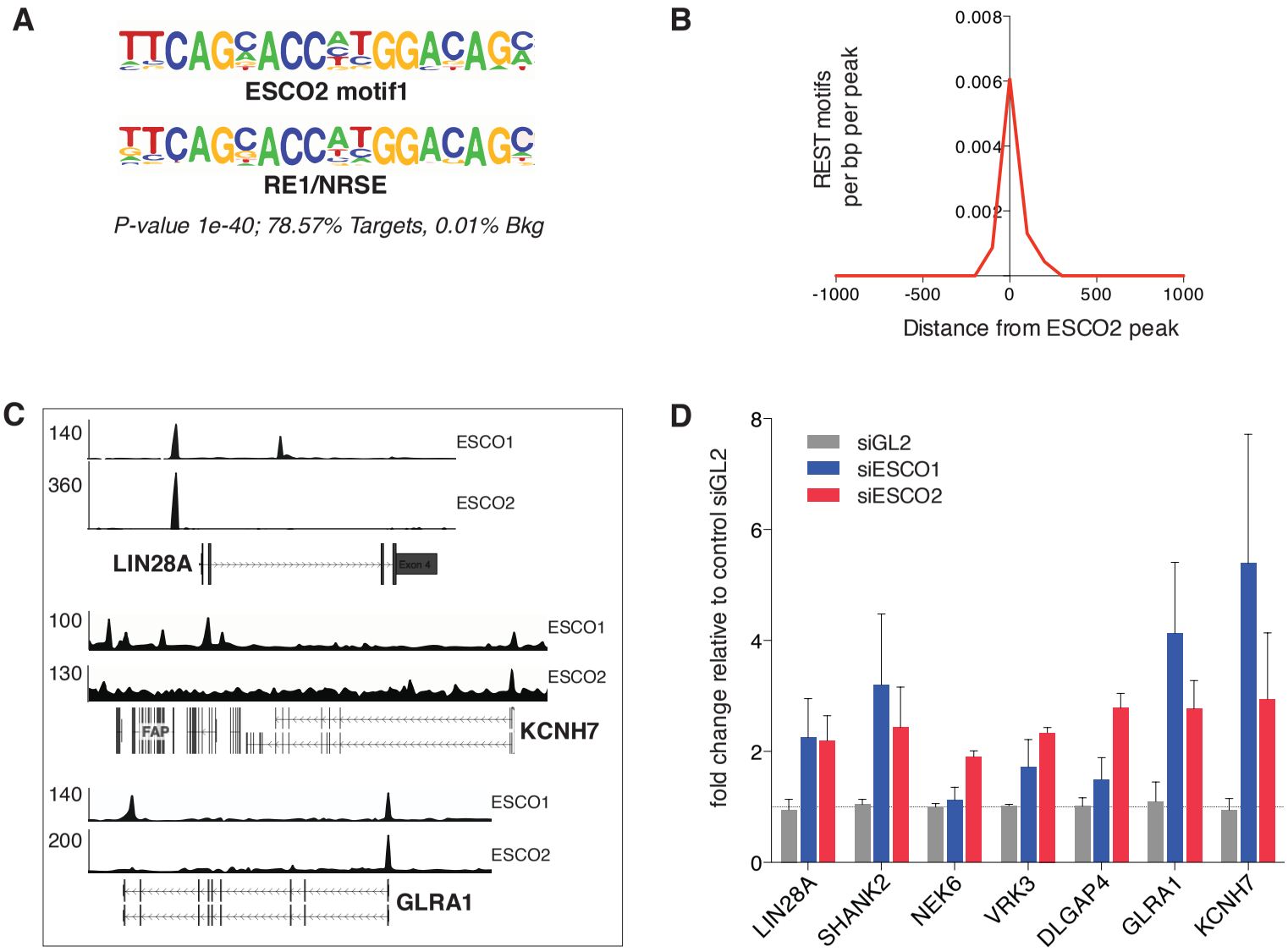

A
ESCO2 motif1
RE1/NRSE
P-value 1e-40; 78.57% Targets, 0.01% Bkg
B
REST motifs per bp per peak
0.008
0.006
0.004
0.002
-1000
-500
0
500
1000
Distance from ESCO2 peak
C
140
ESCO1
360
ESCO2
LIN28A
Exon 4
100
ESCO1
130
ESCO2
FAP
KCNH7
140
ESCO1
200
ESCO2
GLRA1
D
fold change relative to control siGL2
siGL2
siESCO1
siESCO2
0
2
4
6
8
LIN28A
SHANK2
NEK6
VRK3
DLGAP4
GLRA1
KCNH7

Figure 5.1. ESCO2 is targeted mainly to REST motifs and ESCO1/2 regulate expression of REST target genes. (A) *De novo* motif analysis of ESCO2 binding sequences reveals that the majority of sites contain REST consensus motif (RE1/NRSE). (B) Histogram of REST motifs around ESCO2 peaks shows ESCO2 binding is centered around the REST motif. (C) Enrichment of ESCO1 and ESCO2 ChIP-Seq reads visualized around a few of the ESCO2-bound REST target genes: LIN28A, KCNH7 and GLRA1. ESCO1 overlaps with most of the ESCO2 binding sites. Number of reads are plotted on Y axis. (D) Depletion of ESCO1 or ESCO2 results in upregulation of REST target genes bound by ESCO1/2. Values are normalized to internal reference GAPDH and are relative to control GL2 knockdown. Error bars represent standard deviation of three knockdowns for each siRNA.

Table 5.1 List of RT-PCR primers

Gene	Forward	Reverse
GAPDH	GAAGGTGAAGGTCGGAGTCAAC	CAGAGTTAAAAGCAGCCCTGGT
C4BPB	TGCAAAAGTAGGGACTGTGACC	TGCACGCCCACTAAGTAGTAC
LIN28A	ACCAGCAGTTTGCAGGTG	CAGCAGCTGAGGCTCGTC
DLGAP4	GAGGAAGGAAGAGAAGAAACCAC	GGCGTCTGAGGCCTTGTC
SHANK2	GAAGAAGGATAAACCCGAGGAG	TTCCACAGCCATGTTCTCAG
GLRA1	TTAAAGGTCCCCCAGTGAAC	CATGGTTGTCTCAGCAATGG
KCNH7	CGCAAAGGCTGATCTCCTAC	TTCACTTTCAAATGACAATTTCC
NEK6	GATGAATCTCTTCTCCCTGTGC	GACCAGTTCTCGTAACTTCTCG
VRK6	ACTCAAACTGGATGCCAAGG	CCACTTGTTGACTTGCAGAGG
CLDN11_NM1	CTCATCCTGCCGGGCTAC	CAGTCAGCAGCAGTAAAATGGC
CLDN11_NM2	TCATTCTGCTGGCTCTCTGC	TGCATACAGGGAGTAGCCAAAG
KCNK1	TGCTCTCCACCACAGGTTATG	TGAAGGGAATGCCAATGACG
NDFIP2	ATCTGCGGATTTGGCCTTTC	AAGAAAAGGAGCAGGCCAAG
MAP7	TTGAGGAGGACAAAGAACGC	ATCTGCACTGTGGATGCTAGG
TAF12	ACAGGCTCCGGAGAGTTAATTG	CATAATCTGCCGAGCTTTGGAC
TNNC1	GCATGGATGACATCTACAAGGC	GGCTGCCTTGAACTCATTTTTC
ACOT13	ACAATGGCTCTGCTATGCAC	TGCAGGTGACATGTACGTTATG
DNER	CTGCCAGCTTGTTGCAGATC	AAGTGCCTGTTCACAGTTGG

Table 5.2 Genes near ESCO2 binding sites

Genes within 5kb of ESCO2 consensus sites:

Gene	Description[1]	REST site[2]
BC127192	-	
CELF4	RNA-binding protein implicated in the regulation of pre-mRNA alternative splicing	✓
DLGAP4	guanylate kinase found at the postsynaptic density in neuronal cells	✓
FAM24A		✓
GLRA1	subunit of a inhibitory glycine receptor that mediates postsynaptic inhibition	✓
KCNH7	voltage-gated potassium (Kv) channels, involved in neurotransmitter release	✓
LIN28A	microRNA-binding protein	✓
MIR7-3HG	RNA gene affiliated with the lncRNA class	
NPAS4	neuronal PAS domain protein 4; transcriptional regulator, can activate the CNS midline enhancer	✓
PNPLA6	function in neurite outgrowth and process elongation during neuronal differentiation	✓
RTBDN	preferentially expressed in the retina and may play a role in binding retinoids	✓
SHANK2	synaptic protein with function in the postsynaptic density	✓
UNC5A	mediates axon guidance	
VRK3	member of the vaccinia-related kinase family of serine/threonine protein kinases	✓

Additional genes within 5kb of all ESCO2 binding sites (in both replicates) that contain REST motif:

BRSK2	serine/threonine-protein kinase that plays a key role in polarization of neurons and axonogenesis	✓
CADM3	involved in the cell-cell adhesion	✓
INSM2	may function as a growth suppressor in liver cells and in certain neurons	✓
L1CAM	axonal glycoprotein involved in neuron-neuron adhesion, neurite fasciculation, outgrowth of neurites, etc	✓
SYT7	member of the synaptotagmin gene family that mediate calcium-dependent regulation of membrane trafficking in synaptic transmission	✓
XKR7	XK, Kell blood group complex subunit-related family, member 7	✓

[1]From GeneCards.
[2]Presence of REST motif(s) according to http://oreganno.org/tfview/cgi-bin/sitelist.pl?speciesname=Homo%20sapiens&tf=REST.

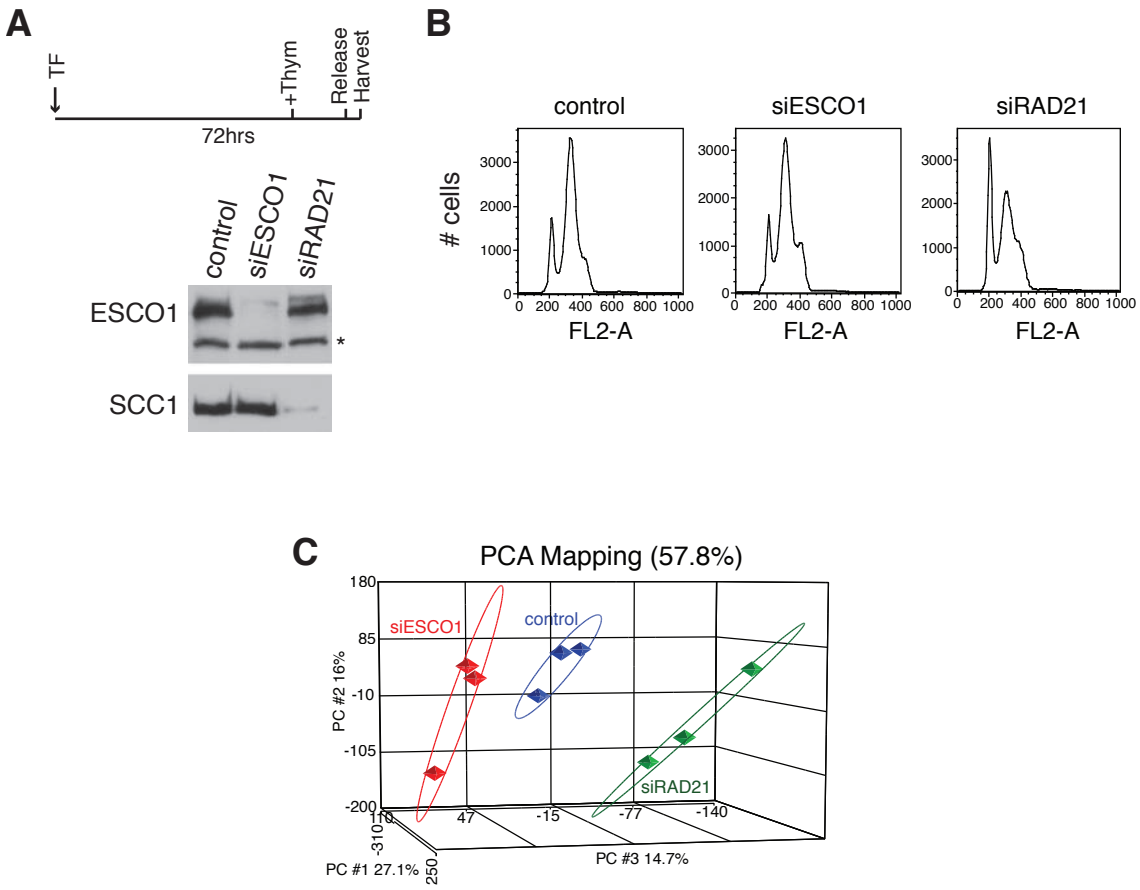


Figure 5.2. ESCO1-knockdown microarray. (A) Schematic of knockdown and Western blot of WCE samples. Cells were harvested 72 hours after transfection with a 16 hour thymidine block and 4 hour release. (B) FACS profiles of propidium iodide staining (FL2-A) for each sample show similar cell cycle profiles. (C) PCA mapping of triplicate knockdown samples shows clustering of RAD21 and ESCO1 knockdown samples separately from control cells on PC3 axis.

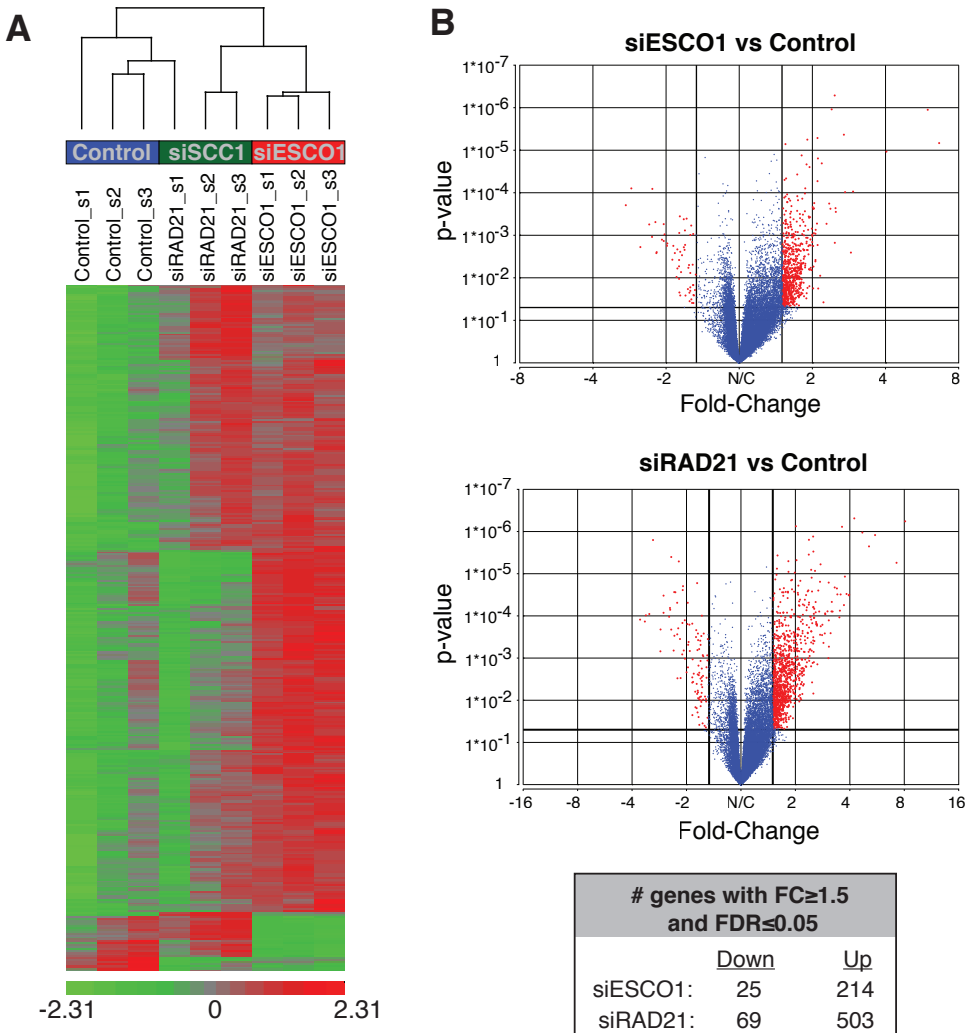


Figure 5.3. Gene expression changes are observed after ESCO1 depletion by microarray. (A) Hierarchichal clustering using genes that change in expression with FDR ≤ 0.05 in siESCO1 vs control (n=548). Standardized probe intensity is shown on scale. (E) Volcano plots of siESCO1 or siRAD21 samples vs control, with fold changes plotted against p-values. Genes with FC≥1.5 and p-value≤0.05 are colored in red. Table lists number of genes upregulated or downregulated for each knockdown vs control, and overlapping upregulated or downregulated genes in both ESCO1 and RAD21 knock-downs.

volcano plots of ESCO1 or RAD21 versus control knockdown expression profiles, many genes are upregulated or downregulated with a fold-change cutoff of 1.5 and p-value cutoff 0.05 after either ESCO1 or RAD21 depletion. The number of genes that change with FDR≤0.05 are listed. Therefore, the microarray data clearly demonstrate that the expression of many genes is changed after ESCO1 depletion.

RT-qPCR confirms altered gene expression after ESCO1 loss

In order to confirm the changes observed by microarray profiling, I assayed several genes by RT-qPCR in independent knockdown samples. Instead of mock (no siRNA) transfection as control as used in the microarray, cells that were transfected with GL2 siRNA were used as negative control for RT-qPCR to rule out siRNA artifacts, and triplicate knockdowns were performed for each GL2 and ESCO1.

Genes chosen for validation have nearby ESCO1-binding sites by ChIP-Seq and some of these are shown in Figure 5.4A. RT-qPCR confirmed the microarray expression changes, with downregulated genes also downregulated by RT-qPCR, and upregulated genes also up-regulated (Figure 5.4B). Therefore, ESCO1 loss clearly results in genome-wide gene expression changes.

Vertebrate-specific N-terminal domains play a role in regulation of gene expression by ESCO1

Next, I analyzed the effect of expressing the vertebrate-specific N-terminal deletions on levels of several of RT-qPCR validated genes. Expressing some ESCO1 deletion mutants alters the levels of these genes. These deletions have varying effects on the various genes (Figure 5.5A). However, deletions of A, B and D generally have similar effects on most genes, and WT and deletion of G have similar effects. Also, they appear to act dominant negatively, as depletion of endogenous ESCO1 is not required for their

effects, though it does cause more severe changes (Figure 5.5B). The effects of each deletion mutants on each gene are consistent in both experiments, with and without ESCO1 depletion.

Discussion

ESCO1/2 regulate transcription of REST target genes

Analysis of ESCO2 binding sequences by motif analysis reveals the presence of the REST motif at nearly all of these sites. RT-qPCR analyzing the expression of bound RST target genes after ESCO1/2 depletion demonstrates that they may function with REST to repress transcription of these genes.

In one study, ESCO2 was reported to interact with coREST and several histone deacetylases and methyltransferases (Kim et al., 2008). CoREST is the co-repressor and an interactor of REST, with which it functions for long-term repression of neuronal genes to maintain cell identity in non-neuronal cell types (Andres et al., 1999). This interaction was found by immunoprecipitation of ESCO2 from HeLa cells followed by mass spectrometry of co-immunoprecipitated proteins, which identified coREST as well as other REST complex members (histone deacetylases BRAF35, PHF21A, HDAC1/2 and LSD1). The study went on to show that ESCO2 fused to Gal4-DNA binding domain represses transcription in a luciferase reporter construct. The same group showed in another study that ESCO1 interacts with histone demethylase LSD1 and HDAC1/2, and ESCO1 fusion to Gal4-DBD also represses transcription in a luciferase assay (Choi et al., 2010). There have not been follow-up studies or other reports of ESCO1 or ESCO2 involvement in transcription, but these interactions with coREST complex members are worth revisiting in light of the findings presented here.

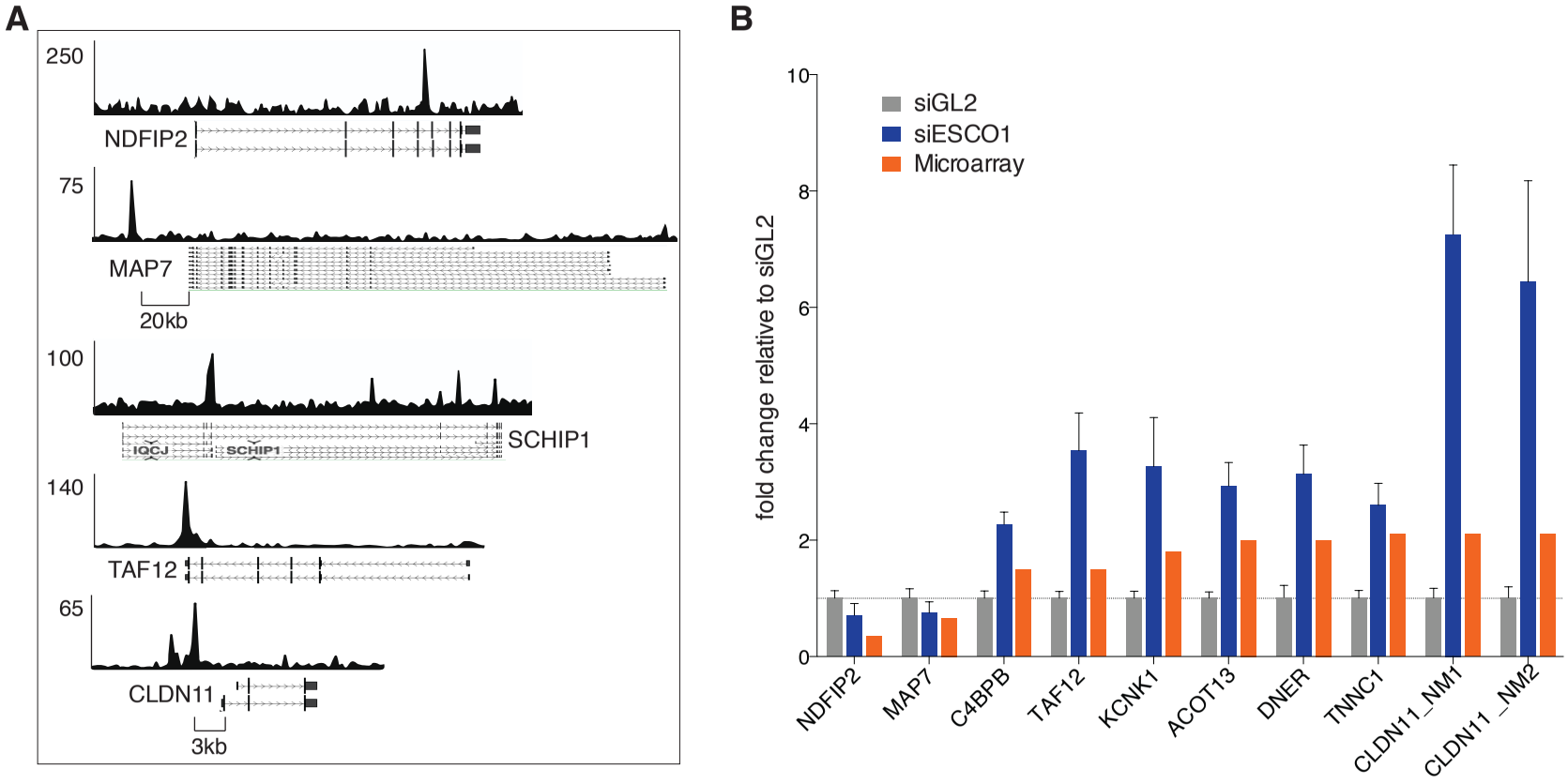


Figure 5.4. Validation of microarray results after ESCO1 depletion. (A) ESCO1 ChIP-Seq reads mapped surrounding a few genes validated by RT-qPCR. (B) Genes that are downregulated in microarray (NDFIP2 and MAP7) are downregulated in RT-qPCR. Microarray-upregulated genes are also upregulated by RT-qPCR. Genes are arranged by fold-change in ESCO1-KD versus control in microarray: NDFIP2: 0.34; MAP7: 0.66; upregulated genes range from 1.5-fold to 2.1-fold upregulated by microarray. Graph shows fold change relative to GL2 control knockdown by RT-qPCR, using GAPDH as internal reference and to normalize levels. CLDN11 is assayed by two primer pairs, NM1 and NM2. NM1 is specific for one transcript and NM2 recognizes both CLDN11 transcripts.

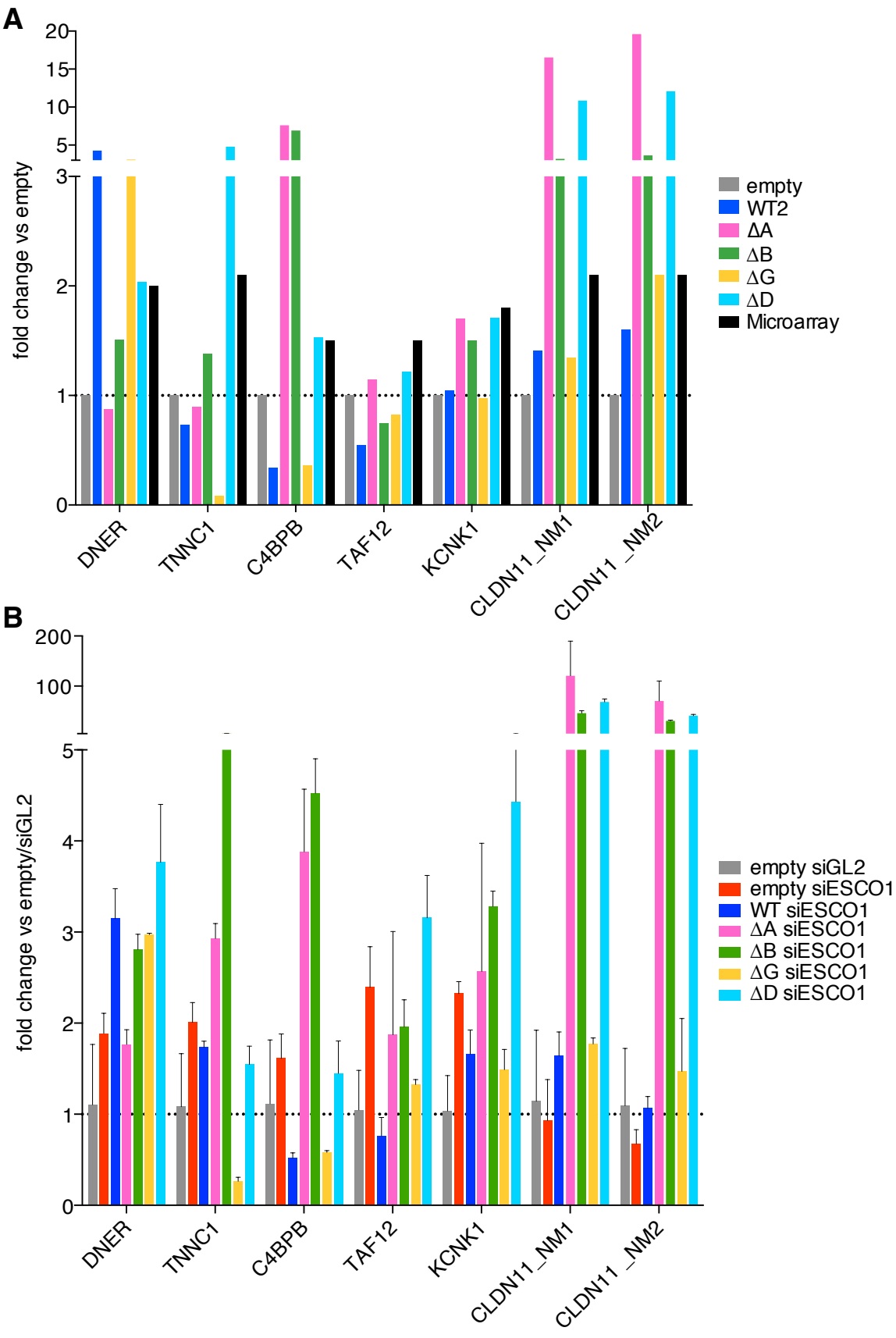

A
fold change vs empty
20
15
10
5
3
2
1
0
DNER
TNNC1
C4BPB
TAF12
KCNK1
CLDN11_NM1
CLDN11 _NM2
empty
WT2
ΔA
ΔB
ΔG
ΔD
Microarray
B
fold change vs empty/siGL2
200
100
5
4
3
2
1
0
DNER
TNNC1
C4BPB
TAF12
KCNK1
CLDN11_NM1
CLDN11 _NM2
empty siGL2
empty siESCO1
WT siESCO1
ΔA siESCO1
ΔB siESCO1
ΔG siESCO1
ΔD siESCO1

Figure 5.5. Effect of ESCO1 N-terminal deletions on gene expression. (A) Expression of N-terminal deletion mutants alone, without ESCO1 depletion, causes gene expression changes. Clonal HeLa cell lines expressing empty vector, WT-ESCO1, or indicated deletions in full length ESCO1 are used. (B) Effect of ESCO1 depletion in N-terminal deletion cell lines. Error bars represent the average of two independent knockdown samples. Note that there are variations in how the deletion mutants affect various genes but the changes are generally consistent in both experiments, under both no KD (A) and KD (B) conditions, for the various deletions.

There have also been findings that have predicted a role for ESCO2 in neuronal/brain development. One study showed that ESCO2 cooperates in the Notch pathway to promote neuronal differentiation (Leem et al., 2011). ESCO2 was shown to co-immunoprecipitate with Notch and disrupt Notch intracellular domain (NICD) interaction with its nuclear binding partner, transcription factor CBF1, on a downstream gene promoter. Interestingly, ESCO2 overexpression caused differentiation of embryonic carcinoma cells and neural progenitor cells, while its depletion blocked differentiation. It will be interesting to analyze if ESCO2 contribution in repression of REST target genes, in addition to its involvement in the Notch pathway, leads to these effects.

It is also interesting to note that mouse ESCO2-deficient cells have severe defects in brain development (Figure 1.7, (Whelan et al., 2012)). ESCO2 knocked out in cortical progenitor cells causes microcephaly. Though the study suggested that a slight defect in pericentric cohesion, where ESCO2 localizes in mouse cells, and eventually apoptosis is the cause of these defects, a role of ESCO2 in regulating proper neuronal development may also contribute. The fact that ESCO2 loss in the neuronal progenitor cells affects proper brain development supports a role for ESCO2 in neuron differentiation and development.

Finally, although findings that RBS patient cells have mild heterochromatic repulsion at centromeres and increased apoptosis (Vega et al., 2005) have suggested that patient phenotypes are are due to these defects instead of gene expression changes, one study reported that two RBS patient lines clustered with CdLS patient samples by expression profiling, instead of with control samples, appearing to have an intermediate gene expression profile to controls and CdLS (Liu et al., 2009). This indicates that RBS cells may in fact have altered transcriptional profiles and is worth further study.

Genome-wide transcriptional changes in ESCO1 depleted cells and role of vertebrate-specific N-terminal domains

The findings here clearly demonstrate a role for ESCO1 in regulating the expression of many genes. Furthermore, the vertebrate-specific N-terminal regions are shown to be important in this function. Unlike in cohesion, where a ESCO1-ΔB does not have a defect, there is a defect in transcription when region B is deleted. The *in vitro* data shows that this region interacts with CTCF. This may explain its involvement in transcription, and why ESCO1-ΔB deletion acts as a separation of function mutant for cohesion and transcription.

RT-qPCR analysis of gene expression after ESCO1 depletion in the deletion mutants raises the question of why there is such a strong effect on transcription with some of these mutants, much stronger than the ESCO1-depletion phenotype. This may be explained by incomplete depletion of ESCO1 by siRNA. The remaining protein may fulfill some of its function in transcription. Expression of the mutants may cause them to act as dominant negatives. Alternatively, the mutants may lose interaction with proteins that ESCO1 negatively regulates, or sequester interacting proteins and not allowing them to carry out their normal transcriptional function.

Conclusions and Future Directions

In summary, the findings in this study reveal that cohesin acetyltransferases ESCO1 and ESCO2 localize differently to DNA. ESCO1 associates with many thousands of sites genome-wide, and is targeted to unreplicated DNA in a cohesin/CTCF-dependent manner. In addition, I show that cohesin is acetylated at high levels at its ChIP-enriched sites already in G1 and acetylation at these sites does not increase as cells pass through S-phase. ESCO1 is responsible for the acetylation at these discrete sites, as acetylation levels are greatly reduced in the absence of ESCO1.

In addition, analysis of vertebrate ESCO1 homologs reveal regions of similarity in the otherwise divergent N-termini. The N-terminus is also shown to interact with cohesin subunit SMC1A, CTCF and PDS5B, and ESCO1 targets to its binding sites through interaction with these proteins. Some of the vertebrate-specific N-terminal regions of ESCO1 are important for its role in cohesion.

Furthermore, I show that these acetyltransferases are not only important for cohesion establishment but also for the regulation of gene expression. ESCO1 regulates the expression of many genes while ESCO2 is required at least for repression of REST target genes. The N-terminal regions of ESCO1 play a role in transcriptional regulation of genes and one region (B) is shown to both interact with CTCF and have a role in transcription, while lacking a function in cohesion.

Future studies of ESCO2 targeting are required to determine where and how it binds to DNA. It is possible that ESCO2 binds to repeat regions of the genome such as pericentric heterochromatin or nucleoli, or it may bind dynamically, for example by tracking with replication forks. It also remains to be seen if the ESCO2 N-terminus has a similar role in its targeting as the ESCO1 N-terminus.

Further analysis of the vertebrate-specific N-terminal regions is needed to clarify the mechanism of their function in ESCO1 targeting. Also, investigation of the different dependencies of these regions in cohesion and transcriptional function of ESCO1 will reveal their contributions these processes.

Whether or not the transcriptional function of ESCO1/2 are mediated through their acetylation of SMC3 or another substrate, or if the acetylation activity is required at all, also requires further study. Yeast Eco1 is also able to acetylate Rad21, which is important for DNA repair, as well as Scc3 and Pds5 *in vitro* (Ivanov et al., 2002), though *in vivo* data for these substrates has not been reported. There might exist other unidentified substrates as well. Alternatively, the functions of ESCO1/2 in transcription may be independent of their acetyltransferase activity. This seems unlikely however, given that the cohesin deacetylase and other cohesion-pathway proteins, namely WAPL and PDS5, also appear to have transcriptional functions. It is likely that ESCO1/2 function in transcription through their regulation of cohesin, but this remains open to future studies on the transcriptional roles of these factors.

REFERENCES

Anderson, D. E., A. Losada, H. P. Erickson and T. Hirano (2002). "Condensin and cohesin display different arm conformations with characteristic hinge angles." J Cell Biol **156**(3): 419-424.

Andres, M. E., C. Burger, M. J. Peral-Rubio, E. Battaglioli, M. E. Anderson, J. Grimes, J. Dallman, N. Ballas and G. Mandel (1999). "CoREST: a functional corepressor required for regulation of neural-specific gene expression." Proc Natl Acad Sci U S A **96**(17): 9873-9878.

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

Beckouet, F., B. Hu, M. B. Roig, T. Sutani, M. Komata, P. Uluocak, V. L. Katis, K. Shirahige and K. Nasmyth (2010). "An Smc3 acetylation cycle is essential for establishment of sister chromatid cohesion." Mol Cell **39**(5): 689-699.

Bell, A. C. and G. Felsenfeld (2000). "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." Nature **405**(6785): 482-485.

Berdougo, E., M. E. Terret and P. V. Jallepalli (2009). "Functional dissection of mitotic regulators through gene targeting in human somatic cells." Methods Mol Biol **545**: 21-37.

Bisht, K. K., Z. Daniloski and S. Smith (2013). "SA1 binds directly to DNA through its unique AT-hook to promote sister chromatid cohesion at telomeres." J Cell Sci **126**(Pt 15): 3493-3503.

Borges, V., C. Lehane, L. Lopez-Serra, H. Flynn, M. Skehel, T. Rolef Ben-Shahar and F. Uhlmann (2010). "Hos1 deacetylates Smc3 to close the cohesin acetylation cycle." Mol Cell **39**(5): 677-688.

Borges, V., D. J. Smith, I. Whitehouse and F. Uhlmann (2013). "An Eco1-independent sister chromatid cohesion establishment pathway in S. cerevisiae." Chromosoma **122**(1-2): 121-134.

Brands, A. and R. V. Skibbens (2008). "Sister chromatid cohesion role for CDC28-CDK in Saccharomyces cerevisiae." Genetics **180**(1): 7-16.

Carretero, M., M. Ruiz-Torres, M. Rodriguez-Corsino, I. Barthelemy and A. Losada (2013). "Pds5B is required for cohesion establishment and Aurora B accumulation at centromeres." EMBO J **32**(22): 2938-2949.

Chan, K. L., T. Gligoris, W. Upcher, Y. Kato, K. Shirahige, K. Nasmyth and F. Beckouet (2013). "Pds5 promotes and protects cohesin acetylation." Proc Natl Acad Sci U S A **110**(32): 13020-13025.

Chan, K. L., M. B. Roig, B. Hu, F. Beckouet, J. Metson and K. Nasmyth (2012). "Cohesin's DNA exit gate is distinct from its entrance gate and is regulated by acetylation." Cell **150**(5): 961-974.

Chien, R., W. Zeng, S. Kawauchi, M. A. Bender, R. Santos, H. C. Gregson, J. A. Schmiesing, D. A. Newkirk, X. Kong, A. R. Ball, Jr., A. L. Calof, A. D. Lander, M. T. Groudine and K. Yokomori (2011). "Cohesin mediates chromatin interactions that regulate mammalian beta-globin expression." J Biol Chem **286**(20): 17870-17878.

Choi, H. K., B. J. Kim, J. H. Seo, J. S. Kang, H. Cho and S. T. Kim (2010). "Cohesion establishment factor, Eco1 represses transcription via association with histone demethylase, LSD1." Biochem Biophys Res Commun **394**(4): 1063-1068.

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

Ciosk, R., W. Zachariae, C. Michaelis, A. Shevchenko, M. Mann and K. Nasmyth (1998). "An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast." Cell **93**(6): 1067-1076.

Coulson, J. M. (2005). "Transcriptional regulation: cancer, neurons and the REST." Curr Biol **15**(17): R665-668.

Cuddapah, S., R. Jothi, D. E. Schones, T. Y. Roh, K. Cui and K. Zhao (2009). "Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains." Genome Res **19**(1): 24-32.

Deardorff, M. A., M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K. E. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L. J. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin, M. Kaur, T. Kiyono, P. M. Lombardi, L. Magnaghi-Jaulin, G. R. Mortier, N. Nozaki, M. B. Petersen, H. Seimiya, V. M. Siu, Y. Suzuki, K. Takagaki, J. J. Wilde, P. J. Willems, C. Prigent, G. Gillessen-Kaesbach, D. W. Christianson, F. J. Kaiser, L. G. Jackson, T. Hirota, I. D. Krantz and K. Shirahige (2012). "HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle." Nature **489**(7415): 313-317.

Deardorff, M. A., M. Kaur, D. Yaeger, A. Rampuria, S. Korolev, J. Pie, C. Gil-Rodriguez, M. Arnedo, B. Loeys, A. D. Kline, M. Wilson, K. Lillquist, V. Siu, F. J. Ramos, A. Musio, L. S. Jackson, D. Dorsett and I. D. Krantz (2007). "Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation." Am J Hum Genet **80**(3): 485-494.

Deardorff, M. A., J. J. Wilde, M. Albrecht, E. Dickinson, S. Tennstedt, D. Braunholz, M. Monnich, Y. Yan, W. Xu, M. C. Gil-Rodriguez, D. Clark, H. Hakonarson, S. Halbach, L. D. Michelis, A. Rampuria, E. Rossier, S. Spranger, L. Van Maldergem, S. A. Lynch, G. Gillessen-Kaesbach, H. J. Ludecke, R. G. Ramsay,

M. J. McKay, I. D. Krantz, H. Xu, J. A. Horsfield and F. J. Kaiser (2012). "RAD21 mutations cause a human cohesinopathy." Am J Hum Genet **90**(6): 1014-1027.

Dorsett, D. (2007). "Roles of the sister chromatid cohesion apparatus in gene expression, development, and human syndromes." Chromosoma **116**(1): 1-13.

Dorsett, D., J. C. Eissenberg, Z. Misulovin, A. Martens, B. Redding and K. McKim (2005). "Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila." Development **132**(21): 4743-4753.

Dorsett, D. and I. D. Krantz (2009). "On the molecular etiology of Cornelia de Lange syndrome." Ann N Y Acad Sci **1151**: 22-37.

Edwards, S., C. M. Li, D. L. Levy, J. Brown, P. M. Snow and J. L. Campbell (2003). "Saccharomyces cerevisiae DNA polymerase epsilon and polymerase sigma interact physically and functionally, suggesting a role for polymerase epsilon in sister chromatid cohesion." Mol Cell Biol **23**(8): 2733-2748.

Farina, A., J. H. Shin, D. H. Kim, V. P. Bermudez, Z. Kelman, Y. S. Seo and J. Hurwitz (2008). "Studies with the human cohesin establishment factor, ChlR1. Association of ChlR1 with Ctf18-RFC and Fen1." J Biol Chem **283**(30): 20925-20936.

Filippova, G. N., S. Fagerlie, E. M. Klenova, C. Myers, Y. Dehner, G. Goodwin, P. E. Neiman, S. J. Collins and V. V. Lobanenkov (1996). "An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes." Mol Cell Biol **16**(6): 2802-2813.

Gandhi, R., P. J. Gillespie and T. Hirano (2006). "Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase." Curr Biol **16**(24): 2406-2417.

Gerlich, D., B. Koch, F. Dupeux, J. M. Peters and J. Ellenberg (2006). "Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication." Curr Biol **16**(15): 1571-1578.

Gillespie, P. J. and T. Hirano (2004). "Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts." Curr Biol **14**(17): 1598-1603.

Glynn, E. F., P. C. Megee, H. G. Yu, C. Mistrot, E. Unal, D. E. Koshland, J. L. DeRisi and J. L. Gerton (2004). "Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae." PLoS Biol **2**(9): E259.

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

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

Gullerova, M. and N. J. Proudfoot (2008). "Cohesin complex promotes transcriptional termination between convergent genes in S. pombe." Cell **132**(6): 983-995.

Guo, Y., K. Monahan, H. Wu, J. Gertz, K. E. Varley, W. Li, R. M. Myers, T. Maniatis and Q. Wu (2012). "CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice." Proc Natl Acad Sci U S A **109**(51): 21081-21086.

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

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

Hanna, J. S., E. S. Kroll, V. Lundblad and F. A. Spencer (2001). "Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion." Mol Cell Biol **21**(9): 3144-3158.

Heidinger-Pauli, J. M., E. Unal, V. Guacci and D. Koshland (2008). "The kleisin subunit of cohesin dictates damage-induced cohesion." Mol Cell **31**(1): 47-56.

Heidinger-Pauli, J. M., E. Unal and D. Koshland (2009). "Distinct targets of the Eco1 acetyltransferase modulate cohesion in S phase and in response to DNA damage." Mol Cell **34**(3): 311-321.

Hou, F. and H. Zou (2005). "Two human orthologues of Eco1/Ctf7 acetyltransferases are both required for proper sister-chromatid cohesion." Mol Biol Cell **16**(8): 3908-3918.

Huang, C. E., M. Milutinovich and D. Koshland (2005). "Rings, bracelet or snaps: fashionable alternatives for Smc complexes." Philos Trans R Soc Lond B Biol Sci **360**(1455): 537-542.

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

Ivanov, D., A. Schleiffer, F. Eisenhaber, K. Mechtler, C. H. Haering and K. Nasmyth (2002). "Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion." Curr Biol **12**(4): 323-328.

Kagey, M. H., J. J. Newman, S. Bilodeau, Y. Zhan, D. A. Orlando, N. L. van Berkum, C. C. Ebmeier, J. Goossens, P. B. Rahl, S. S. Levine, D. J. Taatjes, J. Dekker and R. A. Young (2010). "Mediator and cohesin connect gene expression and chromatin architecture." Nature **467**(7314): 430-435.

Kenna, M. A. and R. V. Skibbens (2003). "Mechanical link between cohesion establishment and DNA replication: Ctf7p/Eco1p, a cohesion establishment factor, associates with three different replication factor C complexes." Mol Cell Biol **23**(8): 2999-3007.

Kim, B. J., K. M. Kang, S. Y. Jung, H. K. Choi, J. H. Seo, J. H. Chae, E. J. Cho, H. D. Youn, J. Qin and S. T. Kim (2008). "Esco2 is a novel corepressor that associates with various chromatin modifying enzymes." Biochem Biophys Res Commun **372**(2): 298-304.

Kitajima, T. S., T. Sakuno, K. Ishiguro, S. Iemura, T. Natsume, S. A. Kawashima and Y. Watanabe (2006). "Shugoshin collaborates with protein phosphatase 2A to protect cohesin." Nature **441**(7089): 46-52.

Kueng, S., B. Hegemann, B. H. Peters, J. J. Lipp, A. Schleiffer, K. Mechtler and J. M. Peters (2006). "Wapl controls the dynamic association of cohesin with chromatin." Cell **127**(5): 955-967.

Lee, B. K. and V. R. Iyer (2012). "Genome-wide studies of CCCTC-binding factor (CTCF) and cohesin provide insight into chromatin structure and regulation." J Biol Chem **287**(37): 30906-30913.

Leem, Y. E., H. K. Choi, S. Y. Jung, B. J. Kim, K. Y. Lee, K. Yoon, J. Qin, J. S. Kang and S. T. Kim (2011). "Esco2 promotes neuronal differentiation by repressing Notch signaling." Cell Signal **23**(11): 1876-1884.

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

Lengronne, A., J. McIntyre, Y. Katou, Y. Kanoh, K. P. Hopfner, K. Shirahige and F. Uhlmann (2006). "Establishment of sister chromatid cohesion at the S. cerevisiae replication fork." Mol Cell **23**(6): 787-799.

Lin, W., H. Jin, X. Liu, K. Hampton and H. G. Yu (2011). "Scc2 regulates gene expression by recruiting cohesin to the chromosome as a transcriptional activator during yeast meiosis." Mol Biol Cell **22**(12): 1985-1996.

Liu, J., Z. Zhang, M. Bando, T. Itoh, M. A. Deardorff, D. Clark, M. Kaur, S. Tandy, T. Kondoh, E. Rappaport, N. B. Spinner, H. Vega, L. G. Jackson, K. Shirahige and I. D. Krantz (2009). "Transcriptional dysregulation in NIPBL and cohesin mutant human cells." PLoS Biol **7**(5): e1000119.

Lopez-Serra, L., A. Lengronne, V. Borges, G. Kelly and F. Uhlmann (2013). "Budding yeast Wapl controls sister chromatid cohesion maintenance and chromosome condensation." Curr Biol **23**(1): 64-69.

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

Losada, A., T. Yokochi and T. Hirano (2005). "Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts." J Cell Sci **118**(Pt 10): 2133-2141.

Lunyak, V. V. and M. G. Rosenfeld (2005). "No rest for REST: REST/NRSF regulation of neurogenesis." Cell **121**(4): 499-501.

Lyons, N. A. and D. O. Morgan (2011). "Cdk1-dependent destruction of Eco1 prevents cohesion establishment after S phase." Mol Cell **42**(3): 378-389.

Mc Intyre, J., E. G. Muller, S. Weitzer, B. E. Snydsman, T. N. Davis and F. Uhlmann (2007). "In vivo analysis of cohesin architecture using FRET in the budding yeast Saccharomyces cerevisiae." EMBO J **26**(16): 3783-3793.

Misulovin, Z., Y. B. Schwartz, X. Y. Li, T. G. Kahn, M. Gause, S. MacArthur, J. C. Fay, M. B. Eisen, V. Pirrotta, M. D. Biggin and D. Dorsett (2008). "Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome." Chromosoma **117**(1): 89-102.

Moldovan, G. L., B. Pfander and S. Jentsch (2006). "PCNA controls establishment of sister chromatid cohesion during S phase." Mol Cell **23**(5): 723-732.

Monnich, M., Z. Kuriger, C. G. Print and J. A. Horsfield (2011). "A zebrafish model of Roberts syndrome reveals that Esco2 depletion interferes with development by disrupting the cell cycle." PLoS One **6**(5): e20051.

Morita, A., K. Nakahira, T. Hasegawa, K. Uchida, Y. Taniguchi, S. Takeda, A. Toyoda, Y. Sakaki, A. Shimada, H. Takeda and I. Yanagihara (2012). "Establishment and characterization of Roberts syndrome and SC phocomelia model medaka (Oryzias latipes)." Dev Growth Differ **54**(5): 588-604.

Nasmyth, K. (2011). "Cohesin: a catenase with separate entry and exit gates?" Nat Cell Biol **13**(10): 1170-1177.

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

Nativio, R., K. S. Wendt, Y. Ito, J. E. Huddleston, S. Uribe-Lewis, K. Woodfine, C. Krueger, W. Reik, J. M. Peters and A. Murrell (2009). "Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus." PLoS Genet **5**(11): e1000739.

Nishiyama, T., R. Ladurner, J. Schmitz, E. Kreidl, A. Schleiffer, V. Bhaskara, M. Bando, K. Shirahige, A. A. Hyman, K. Mechtler and J. M. Peters (2010). "Sororin mediates sister chromatid cohesion by antagonizing Wapl." Cell **143**(5): 737-749.

Onn, I., V. Guacci and D. E. Koshland (2009). "The zinc finger of Eco1 enhances its acetyltransferase activity during sister chromatid cohesion." Nucleic Acids Res **37**(18): 6126-6134.

Panizza, S., T. Tanaka, A. Hochwagen, F. Eisenhaber and K. Nasmyth (2000). "Pds5 cooperates with cohesin in maintaining sister chromatid cohesion." Curr Biol **10**(24): 1557-1564.

Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H. C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, B. S. Cobb, K. Yokomori, N. Dillon, L. Aragon, A. G. Fisher and M. Merkenschlager (2008). "Cohesins functionally associate with CTCF on mammalian chromosome arms." Cell **132**(3): 422-433.

Pauli, A., F. Althoff, R. A. Oliveira, S. Heidmann, O. Schuldiner, C. F. Lehner, B. J. Dickson and K. Nasmyth (2008). "Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons." Dev Cell **14**(2): 239-251.

Peters, J. M. (2012). "The many functions of cohesin--different rings to rule them all?" EMBO J **31**(9): 2061-2063.

Rankin, S., N. G. Ayad and M. W. Kirschner (2005). "Sororin, a substrate of the anaphase-promoting complex, is required for sister chromatid cohesion in vertebrates." Mol Cell **18**(2): 185-200.

Rolef Ben-Shahar, T., S. Heeger, C. Lehane, P. East, H. Flynn, M. Skehel and F. Uhlmann (2008). "Eco1-dependent cohesin acetylation during establishment of sister chromatid cohesion." Science **321**(5888): 563-566.

Rollins, R. A., M. Korom, N. Aulner, A. Martens and D. Dorsett (2004). "Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/ Scc3 cohesion factor to facilitate long-range activation of the cut gene." Mol Cell Biol **24**(8): 3100-3111.

Rollins, R. A., P. Morcillo and D. Dorsett (1999). "Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes." Genetics **152**(2): 577-593.

Rubio, E. D., D. J. Reiss, P. L. Welcsh, C. M. Disteche, G. N. Filippova, N. S. Baliga, R. Aebersold, J. A. Ranish and A. Krumm (2008). "CTCF physically links cohesin to chromatin." Proc Natl Acad Sci U S A **105**(24): 8309-8314.

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

Schmitz, J., E. Watrin, P. Lenart, K. Mechtler and J. M. Peters (2007). "Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase." Curr Biol **17**(7): 630-636.

Schuldiner, O., D. Berdnik, J. M. Levy, J. S. Wu, D. Luginbuhl, A. C. Gontang and L. Luo (2008). "piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning." Dev Cell **14**(2): 227-238.

Schule, B., A. Oviedo, K. Johnston, S. Pai and U. Francke (2005). "Inactivating mutations in ESCO2 cause SC phocomelia and Roberts syndrome: no phenotype-genotype correlation." Am J Hum Genet **77**(6): 1117-1128.

Seitan, V. C., P. Banks, S. Laval, N. A. Majid, D. Dorsett, A. Rana, J. Smith, A. Bateman, S. Krpic, A. Hostert, R. A. Rollins, H. Erdjument-Bromage, P. Tempst, C. Y. Benard, S. Hekimi, S. F. Newbury and T. Strachan (2006). "Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance." PLoS Biol **4**(8): e242.

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

Sjogren, C. and K. Nasmyth (2001). "Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae." Curr Biol **11**(12): 991-995.

Skibbens, R. V. (2004). "Chl1p, a DNA helicase-like protein in budding yeast, functions in sister-chromatid cohesion." Genetics **166**(1): 33-42.

Skibbens, R. V., L. B. Corson, D. Koshland and P. Hieter (1999). "Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery." Genes Dev **13**(3): 307-319.

Strom, L., C. Karlsson, H. B. Lindroos, S. Wedahl, Y. Katou, K. Shirahige and C. Sjogren (2007). "Postreplicative formation of cohesion is required for repair and induced by a single DNA break." Science **317**(5835): 242-245.

Strom, L., H. B. Lindroos, K. Shirahige and C. Sjogren (2004). "Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair." Mol Cell **16**(6): 1003-1015.

Sumara, I., E. Vorlaufer, C. Gieffers, B. H. Peters and J. M. Peters (2000). "Characterization of vertebrate cohesin complexes and their regulation in prophase." J Cell Biol **151**(4): 749-762.

Sumara, I., E. Vorlaufer, P. T. Stukenberg, O. Kelm, N. Redemann, E. A. Nigg and J. M. Peters (2002). "The dissociation of cohesin from chromosomes in prophase is regulated by Polo-like kinase." Mol Cell **9**(3): 515-525.

Sutani, T., T. Kawaguchi, R. Kanno, T. Itoh and K. Shirahige (2009). "Budding yeast Wpl1(Rad61)-Pds5 complex counteracts sister chromatid cohesion-establishing reaction." Curr Biol **19**(6): 492-497.

Takagi, M., K. Bunai, K. Yanagi and N. Imamoto (2008). "Cloning of Xenopus orthologs of Ctf7/Eco1 acetyltransferase and initial characterization of XEco2." FEBS J **275**(24): 6109-6122.

Takahashi, T. S., P. Yiu, M. F. Chou, S. Gygi and J. C. Walter (2004). "Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex." Nat Cell Biol **6**(10): 991-996.

Tanaka, K., Z. Hao, M. Kai and H. Okayama (2001). "Establishment and maintenance of sister chromatid cohesion in fission yeast by a unique mechanism." EMBO J **20**(20): 5779-5790.

Tanaka, K., T. Yonekawa, Y. Kawasaki, M. Kai, K. Furuya, M. Iwasaki, H. Murakami, M. Yanagida and H. Okayama (2000). "Fission yeast Eso1p is required for establishing sister chromatid cohesion during S phase." Mol Cell Biol **20**(10): 3459-3469.

Tedeschi, A., G. Wutz, S. Huet, M. Jaritz, A. Wuensche, E. Schirghuber, I. F. Davidson, W. Tang, D. A. Cisneros, V. Bhaskara, T. Nishiyama, A. Vaziri, A. Wutz, J. Ellenberg and J. M. Peters (2013). "Wapl is an essential regulator of chromatin structure and chromosome segregation." Nature **501**(7468): 564-568.

Terret, M. E., R. Sherwood, S. Rahman, J. Qin and P. V. Jallepalli (2009). "Cohesin acetylation speeds the replication fork." Nature **462**(7270): 231-234.

Toth, A., R. Ciosk, F. Uhlmann, M. Galova, A. Schleiffer and K. Nasmyth (1999). "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication." Genes Dev **13**(3): 320-333.

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

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

Unal, E., J. M. Heidinger-Pauli and D. Koshland (2007). "DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7)." Science **317**(5835): 245-248.

van der Lelij, P., B. C. Godthelp, W. van Zon, D. van Gosliga, A. B. Oostra, J. Steltenpool, J. de Groot, R. J. Scheper, R. M. Wolthuis, Q. Waisfisz, F. Darroudi, H. Joenje and J. P. de Winter (2009). "The cellular phenotype of Roberts syndrome fibroblasts as revealed by ectopic expression of ESCO2." PLoS One **4**(9): e6936.

Vaur, S., A. Feytout, S. Vazquez and J. P. Javerzat (2012). "Pds5 promotes cohesin acetylation and stable cohesin-chromosome interaction." EMBO Rep **13**(7): 645-652.

Vega, H., Q. Waisfisz, M. Gordillo, N. Sakai, I. Yanagihara, M. Yamada, D. van Gosliga, H. Kayserili, C. Xu, K. Ozono, E. W. Jabs, K. Inui and H. Joenje (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**(5): 468-470.

Waizenegger, I. C., S. Hauf, A. Meinke and J. M. Peters (2000). "Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase." Cell **103**(3): 399-410.

Watrin, E., A. Schleiffer, K. Tanaka, F. Eisenhaber, K. Nasmyth and J. M. Peters (2006). "Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression." Curr Biol **16**(9): 863-874.

Wendt, K. S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, K. Yahata, F. Imamoto, H. Aburatani, M. Nakao, N. Imamoto, K. Maeshima, K. Shirahige and J. M. Peters (2008). "Cohesin mediates transcriptional insulation by CCCTC-binding factor." Nature **451**(7180): 796-801.

Whelan, G., E. Kreidl, G. Wutz, A. Egner, J. M. Peters and G. Eichele (2012). "Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin." EMBO J **31**(1): 71-82.

Williams, B. C., C. M. Garrett-Engele, Z. Li, E. V. Williams, E. D. Rosenman and M. L. Goldberg (2003). "Two putative acetyltransferases, san and deco, are required for establishing sister chromatid cohesion in Drosophila." Curr Biol **13**(23): 2025-2036.

Xiong, B., S. Lu and J. L. Gerton (2010). "Hos1 is a lysine deacetylase for the Smc3 subunit of cohesin." Curr Biol **20**(18): 1660-1665.

Zhang, B., J. Chang, M. Fu, J. Huang, R. Kashyap, E. Salavaggione, S. Jain, S. Kulkarni, M. A. Deardorff, M. L. Uzielli, D. Dorsett, D. C. Beebe, P. Y. Jay, R. O. Heuckeroth, I. Krantz and J. Milbrandt (2009). "Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies." PLoS One **4**(5): e5232.

Zhang, B., S. Jain, H. Song, M. Fu, R. O. Heuckeroth, J. M. Erlich, P. Y. Jay and J. Milbrandt (2007). "Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome." Development **134**(17): 3191-3201.

Zhang, J., X. Shi, Y. Li, B. J. Kim, J. Jia, Z. Huang, T. Yang, X. Fu, S. Y. Jung, Y. Wang, P. Zhang, S. T. Kim, X. Pan and J. Qin (2008). "Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast." Mol Cell **31**(1): 143-151.