OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

Exploratory Structural Analysis of SMC3 Acetylation Mimics with SMC1A as a Comparative Model in Roberts Syndrome

Hannah Vásquez Liberman
hannah_vl@icloud.com

ABSTRACT

Roberts Syndrome is a rare cohesinopathy that is caused by pathogenic variants in ESCO2, an acetyltransferase required for proper cohesin regulation to take place. This study used AlphaFold Server (AlphaFold 3) to explore whether lysine to glutamine (K to Q) substitutions, which are used as approximate acetylation mimics, were associated with predicted structural differences at SMC3 K105/K106 and, as an exploratory comparison, SMC1A K579/K584. Wild-type and K to Q models were compared using Cα structural alignment, local residue-level derivations, DSSP secondary structure assignment, and AlphaFold confidence scores. For SMC3, the local Cα RMSD across residues 95-115 was 0.044 Å, DSSP assigned the same secondary structure states at residues 105 and 106 in both models, and local pLDDT values were similar. For SMC1A, the local Cα RMSD across residues 569-594 was 0.116 Å, with unchanged DSSP assignments at residues 579 and 584 and nearly identical local pLDDT values. Whole protein RMSDs were larger (5.37 Å for SMC3 and 10.28 Å for SMC1A), but because the structures were independently predicted with different model seeds, these global differences cannot be directly attributed to the K to Q substitutions. Overall, the models do not provide evidence that these substitutions cause a β-sheet-to-turn transition, protein weakening, or altered cohesin mechanics. These findings are exploratory and highlight the need for molecular dynamics or experimental validation.

INTRODUCTION

Roberts Syndrome is a rare genetic disorder that is classified as cohesinopathy and is caused by pathogenic variants in ESCO2 which is a gene that encodes an acetyltransferase required for proper establishment of sister chromatid cohesion. This is a group of diseases caused by defects in the proteins that are responsible for maintaining proper chromosome structure (GeneReviews, 2006; Vega et al., 2005; Mfarej et al., 2020). Manifestations of Roberts Syndrome include severe developmental abnormalities which include limb malformations, craniofacial defects, and growth retardation (GeneReviews, 2006; Vega et al., 2005). Everything starts at the cellular level. When chromosome cohesion and DNA organization disrupt normal cell division and gene regulation(Mfarej et al., 2020; Nishiyama, 2019).

Cohesion proteins play a crucial role in holding sister chromatids together during cell division and in organizing chromatin within the nucleus of the cell (Gruber et al., 2003; Nishiyama, 2019; Horsfield et al., 2023). For development to proceed as it should, these proteins need to keep their mechanical properties intact. These include binding to DNA, staying structurally stable, and handling the physical forces inside the cell without exception because otherwise, everything falls apart (Gligoris et al., 2014; Song et al., 2012; Villa-Hernandez and Bermejo, 2018).

ESCO2 dysfunction is known to disrupt normal cohesin regulation, but the structural consequences or altered acetylation at the level of individual cohesin subunits remain incompletely understood (Mfarej et al., 2020; GeneReviews, 2006). Recent advancements in protein structure prediction with artificial intelligence, like AlphaFold, have made it possible to model proteins whose experimental structures are not known or difficult to obtain (Jumper et al., 2021; Tunyasuvunakool et al., 2021). While AlphaFold has demonstrated strong performance in predicting wild-type protein structure, its ability to evaluate subtle effects of mutations and post-translational modification is still quite limited (Jumper et al., 2021; Tunyasuvunakool et al., 2021).Therefore, the purpose of this study is to evaluate whether K to Q acetylation-mimic substitutions are associated with detectable structural differences in predicted SMC protein models and to identify questions that may require dynamic or experimental investigation.

This study focuses on an exploratory structural comparison of SMC3 and SMC1A. SMC3 was the main focus because ESCO2-mediated acetylation of SMC3 is established in cohesin regulation. SMC1A was only included as an exploratory comparison because it forms the core SMC heterodimer with SMC3. The SMC1A sites examined are not presented as established Roberts Syndrome targets. The study does not model ESCO2 disease-causing variants directly. Lysine to glutamine substitutions were used as approximate acetylation mimics instead, and independently predicted structures were compared to identify local differences that could motivate the future experimental or molecular dynamics work.

BACKGROUND

Chromosome cohesion and the cohesin complex

For cells to divide correctly, they must copy their DNA and then separate it so that each new cell receives one complete set of chromosomes. After DNA is replicated, each chromosome consists of two identical sister chromatids. These chromatids must stay attached until the correct moment during cell division. This attachment is maintained by a group of proteins known as the cohesin complex, which forms a structure that is ring-like and can physically hold DNA strands together (Gruber et al., 2003). The cohesin complex, its ring-like architecture, and protein composition are illustrated in Figure 1.

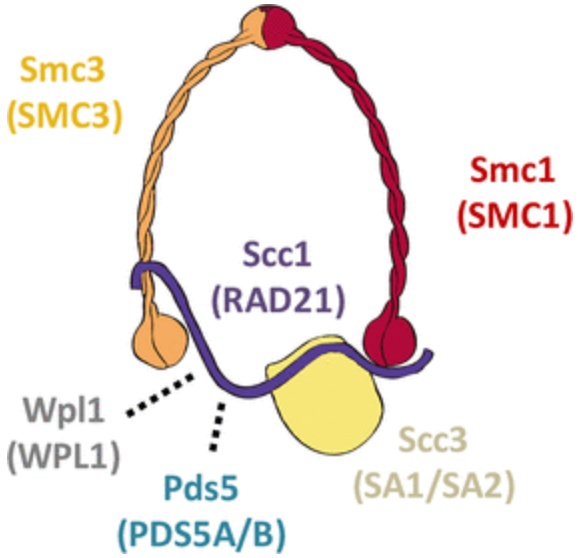

Figure 1. Structure of the SMC cohesin complex. Schematic representation of the core cohesin complex and its interacting factors, including SMC1A, SMC3, RAD21, and STAG proteins. Reproduced from Villa-Hernández and Bermejo (2018), *Cohesin dynamic association to chromatin and interfacing with replication forks in genome integrity maintenance*, *Current Genetics*, 64(5), 1005–1013. https://doi.org/10.1007/s00294-018-0824-x. © Springer-Verlag GmbH Germany, part of Springer Nature.

The core of the cohesin complex is made up of the proteins SMC1A and SMC3, as can be seen in Figure 1. A third protein called RAD21 and a regulatory subunit known as STAG connect these proteins. This protein complex allows cohesin to bind to DNA and maintain chromosome structure during the cell cycle (Gligoris et al., 2014; Horsfield et al., 2023). Chromosomes may separate too early or fail to organize correctly without a properly functioning cohesin complex.

Cohesin has other roles that do not include cell division. It helps organize DNA inside the nucleus by forming chromatin loops that bring distant regions of DNA close together. These loops help control gene expression by determining which genes can interact with regulatory elements. Many times, cohesin works with a protein called CTCF to create these organized regions of the genome that are known as topologically associated domains (TADs) (Gabriele et al., 2022). Even small disruptions in cohesin function can have serious effects since development depends on tightly regulated gene expression (Nishiyama, 2019).

Cohesinopathies and sensitivity during development

Diseases caused by mutations in cohesin proteins or in proteins that regulate cohesin are known as cohesinopathies. Two cohesinopathies that are well known are Cornelia de Lange Syndrome and Roberts Syndrome. While these disorders manifest themselves in different ways and with different levels of severity, they are both linked to problems in the function of cohesin and organization of genomes (Kline et al., 2018).

Studies about Cornelia de Lange Syndrome show that diseases that are related to cohesin are not always a result of complete chromosome separation failure. Instead, many of the symptoms come from changes seen in chromatin structure and gene regulation. This comes to show that cohesin is not only important for holding chromosomes together. It is also important for controlling how genes are turned on and off during development (Kline et al., 2018). A consequence of this is that even small changes in cohesin behavior can lead to major developmental consequences.

Roberts Syndrome and the Role of ESCO2

Roberts Syndrome is a severe and rare cohesinopathy. It is caused by mutations in the ESCO2 gene. ESCO2 produces an enzyme that is necessary for establishing sister chromatid cohesion during DNA replication (Vega et al., 2005). The manifestations of Roberts Syndrome stem from characteristic chromosome defects that can be seen under the microscope such as premature separation at the centromeres (Vega et al., 2005).

At the molecular level, ESCO2 functions by acetylating specific lysine residues on SMC3, one of the main cohesin proteins. This acetylation helps stabilize cohesin on DNA and allows proper cohesion to form during the S phase of the cell cycle (Song et al., 2012). The S phase, or synthesis phase, is the stage of the cell cycle during which DNA replication happens. During this phase, each chromosome is copied to produce two identical sister chromatids that must remain properly connected until cell division (Nishiyama, 2019; Gruber et al, 2003). When this modification does not happen correctly, cohesin becomes less stable and cannot maintain proper attachment between sister chromatids Song et al., 2012 ; GeneReviews, 2006).

The related enzyme ESCO1 can also acetylate SMC3, though research shows that ESCO2 has a unique role during DNA replication that ESCO1 cannot fully replace.
This is why mutations in ESCO2 specifically lead to Roberts Syndrome instead of milder conditions (Alomer et al., 2017). Additionally, ESCO2 mutations have been associated with other cellular problems including abnormalities in nuclear structure and production of proteins. These findings suggest that defects in cohesin acetylation may affect many cellular processes which can contribute to the severity of Roberts Syndrome (Mfarej et al., 2020). SMC proteins have a very structurally distinctive configuration that is essential for their function. Both SMC1A and SMC3 are long, flexible proteins composed of extended coiled coil arms with a hinge domain in the middle and ATPase “head” domains at their ends. The hinge regions of SMC1A and SMC3 interact with each other forming a V-shaped structure. The head domains come together with RAD21 to complete the cohesin ring (Gligoris et al., 2014.; Gruber et al., 2003). This configuration, which is ring-like, allows cohesin to physically encircle DNA rather than binding it to a single point (see Figure. 1).

The acetylation of SMC3 happens near the head domain. This is a region that plays a key role in regulating cohesin’s interactions with DNA and other regulatory proteins. Due to the head domains being involved in ATP binding and conformational changes, modifications in this region can influence how

cohesin opens, closes, and remains bound to chromatin (Song et al., 2012; Nishiyama, 2019). Due to the SMC3 head region participating in ATP-dependent conformational changes and interactions that regulate cohesin, altered acetylation in this region could influence cohesin behavior. However, the direction and magnitude of any mechanical effect cannot be determined from static AlphaFold predictions alone.

Protein and DNA mechanics

Cohesin is a dynamic protein complex that responds to physical forces inside the cell. During cell division, cohesin must resist the pulling forces that are generated by the mitotic spindle while still maintaining its hold on DNA. Structural studies show that cohesin's shape is important for allowing it to act as a structure that is force-bearing.

Cohesin is a dynamic protein complex that responds to physical forces within the cell (Gligoris et al., 2014; Gruber et al., 2003).

DNA is also a mechanical molecule that bends, twists, and loops inside the nucleus. Cohesin plays a key role in maintaining these DNA structures by stabilizing chromatin loops and organizing the genome. If impaired cohesin regulation changes the stability, dynamics, or chromatin interactions of the complex, DNA organization and chromosome segregation could be affected. These consequences are linked to ESCO2 dysfunction in Roberts syndrome. This increases the chance of irregular gene regulation and errors in chromosome separation (Song et al., 2012; Nishiyama, 2019).

From a mechanical perspective, studying Roberts Syndrome helps demonstrate how molecular-level changes can lead to a wide range of developmental defects. The developmental manifestations of Roberts syndrome come from cellular consequences of ESCO2 dysfunction, which include abnormal cohesion and altered genome regulation. Protein flexibility and mechanics may contribute to these processes, but they are not established here as a single direct cause of the clinical phenotype.

METHODS

Protein selection and sequence retrieval

The primary protein of interest in the study was SMC3 because ESCO2-mediated acetylation of SMC3 is an established step in cohesion regulation. SMC1A was included as an exploratory structural comparison because SMC1A and SMC3 form the core SMC heterodimer of cohesin. The SMC1A residues examined in this study were not treated as established ESCO2 targets or as validated Roberts syndrome sites. Full-length human protein sequences were obtained from UniProt: SMC3 (UniProtKB Q9UQE7, 1217 amino acids) and SMC1A (UniProtKB Q14683, 1233 amino acids), and were used as single-chain inputs for structure prediction.

Wildtype structure prediction using AlphaFold

To establish a baseline for comparison, the first structures that were modeled were the wildtype structures of SMC1A and SMC3. All four structure predictions were generated with AlphaFold server using

AlphaFold3 output/format version 3. Each job contained one protein chain and had structure-template use enabled. The model_0 CIF output from each job was used for the quantitative comparison, and the corresponding summary-confidence file reported 10 recycles for every model. The recorded model seeds were 1196632961 for SMC3 wild-type, 662440371 for the SMC3 K105Q/K106Q mimic, 1170897056 for SMC1A wild-type, and 1004364299 for the SMC1A K579Q/K584Q mimic. Because the wild-type and variant structures were generated in separate runs with different model seeds, model to model variability was considered when interpreting the structural differences.

Modeling acetylated proteins as acetylation mimics

To create an approximate acetylation mimic, selected lysine (K) residues were replaced with glutamine (Q). This substitution can approximate the neutralization of lysine's positive charge, but it does not chemically reproduce Nε-acetyllysine and should not be interpreted as a direct model of the post-translational modification. Accordingly, differences between the independently predicted wild-type and K to Q structures were interpreted as exploratory effects of the sequence substitution within the modeling framework, not as proof of the structural consequences of true acetylation. For protein SMC3, the lysine residues K105 and K106 were mutated to glutamine. These were the residues chosen because they are known targets of acetylation during cohesion and they play a key role in stabilizing cohesion on DNA. For protein SMC1A, lysine residues K579 and K584 were mutated to glutamine. These were examined only as an exploratory comparison. Their inclusion does not imply that these residues are established ESCO2 targets or that their acetylation has a demonstrated role in Roberts syndrome.

After introducing the mutations, the modified sequences were resubmitted to the AlphaFold server using the same parameters as the wildtype models. Using the same modeling conditions reduced one source of procedural variation, but it did not eliminate model uncertainty or run-to-run prediction variability. Therefore, observed differences were not automatically attributed to acetylation or the K to Q substitution without additional structural and confidence assessment.

Structural comparison and analysis

Structural comparisons were performed on the AlphaFold model_0 CIF files. The visual ribbon-model comparisons were retained as a qualitative part of the analysis, and the quantitative measurements were used to determine whether the qualitative observations corresponded to measurable local structural differences. Corresponding Cα atoms were superposed with a least-squares Kabsch alignment and Cα root-mean-square deviation (RMSD) was calculated for the full protein and for local windows surrounding the modified sites (SMC3 residues 95-115; SMC1A residues 569-594). Residue-level Cα deviations at the substituted positions were calculated after local superposition. Secondary structure was assigned computationally using the DSSP method implemented in MDTraj (version 1.11.1.post1), rather than being determined solely through visual inspection of the AlphaFold models. AlphaFold per-residue confidence (pLDDT) was read from the model CIF files, and pTM values were taken from the AlphaFold summary-confidence outputs. Since each wild-type and K to Q structure came from an independent prediction with a different model seed, large-scale differences were interpreted cautiously and were not automatically attributed to the substitutions.

Measure	SMC3 WT vs K105Q/K106Q	SMC1A WT vs K579Q/K584Q
Sequence length	1217 aa in both models	1233 aa in both models
Whole-protein Cα RMSD	5.370 Å	10.284 Å
Local window	Residues 95-115	Residues 569-594
Local Cα RMSD	0.044 Å	0.116 Å
Target-residue Cα deviations	105: 0.106 Å; 106: 0.074 Å	579: 0.099 Å; 584: 0.098 Å
DSSP at modified sites	Same in WT and mimic: 105=bend/coil; 106=coil	Same in WT and mimic 579=coil; 584=3_{10} helix
pLDDT at modified sites	105: 83.69 -> 85.06; 106: 85.78 -> 86.81	579: 86.11 -> 86.11; 584: 87.88 -> 87.93
Mean local pLDDT	90.64 -> 91.16	87.91 -> 87.98
pTM	0.49 -> 0.48	0.47 -> 0.46
Interpretation	No comparable local transition is visually apparent. DSSP/local Cα analysis does not independently confirm a formal secondary-structure conversion at the exact sites.	No comparable local transition is visually apparent. DSSP/local Cα analysis likewise shows no confirmed local secondary-structure conversion.

Table 1. Comparison of structural features between wild-type and acetylation-mimic SMC3 and SMC1A models. The table summarizes global and local RMSD values, residue-specific deviations, DSSP assignments, pLDDT scores, pTM scores, and observed structural differences between the wild-type and K to Q mimic models.

Interpretation of structural predictions

The predicted wild-type and acetylation mimic structures were compared to evaluate whether structural differences were found at the regions that surrounded the modified residues. These structural comparisons

were interpreted in the context of cohesin's biological function in chromosome cohesion. However, because AlphaFold generates static structural predictions, the models were not used to draw conclusions about changes in protein flexibility, stability, force resistance, or mechanical behavior. Any functional or mechanical effects of acetylation would require additional methods, such as molecular dynamics simulations or experimental validation.

RESULTS

Overview of structural comparisons

Predicted three dimensional structures were generated for wild- type and acetylation mimic variants of the cohesin subunits SMC3 and SMC1A using AlphaFold. Acetylation mimic mutations were introduced at residues K105 and K106 in SMC3 and K579 and K584 in SMC1A. Each modified structure was compared directly to its corresponding wild-type model. Structural comparisons focused on changes in secondary structure elements, including β sheets, and turns. Predicted Aligned Error (PAE) maps were also examined to compare confidence in relative residue and domain positioning between the models.

Visual inspection and quantitative analysis gave complementary information. In the SMC3 images, the region around positions 105 and 106 shows no clear change between wild-type and the K to Q mimic. The SMC1A images do not show a comparable local secondary-structure transition at positions 579 or 584. Quantitatively, SMC3 residues 95-115 had a local Cα RMSD of 0.044 Å, while SMC1A residues 569-594 had a local Cα RMSD of 0.116 Å. DSSP did not assign a different secondary structure state at the exact modified residues in either protein.

Structural comparison of SMC3 at residue K105

To investigate how acetylation affects local protein structure, a structural comparison was performed between the wild-type SMC3 protein and an acetylation mimic at residue K105.

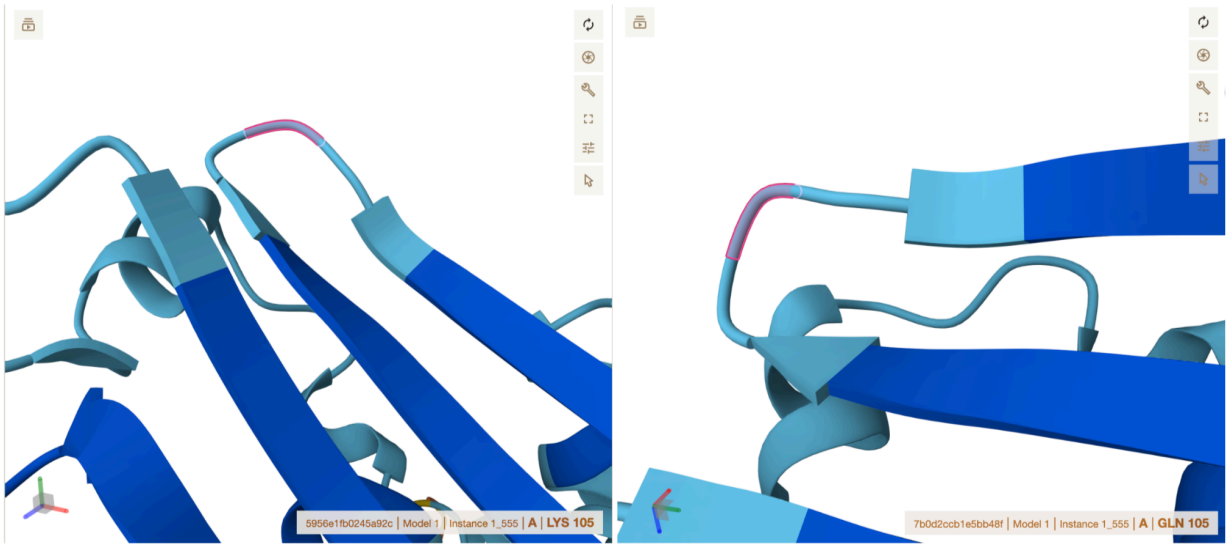

Figure 2. AlphaFold-predicted SMC3 wild-type and K105Q/K106Q acetylation-mimic models at residue 105. Structures were visualized in the AlphaFold Server structure viewer using ribbon representation, with ribbon colors indicating prediction confidence (pLDDT). Residue 105 is highlighted in pink/red. Left: wild-type (K105). Right: acetylation mimic (Q105).

In Figure 2 at the SMC3 105 region, the models show similar local geometry. DSSP gave the same assignment at residue 105 in both paired CIF models (bend/coil), and after local superposition of residues 95-115 the Cα displacement at residue 105 was 0.106 Å. The pLDDT value was 83.69 in the wild-type and 85.06 in the K to Q model. Together, these results indicate that the K105Q/K106Q substitution did not produce a detectable local secondary structure change at residue 105 in this static AlphaFold comparison.

Structural comparison of SMC3 at residue K106

Figure 3 presents a similar comparison for SMC3 residue 106. Again, the wild-type is on the left and the acetylated structure is on the right.

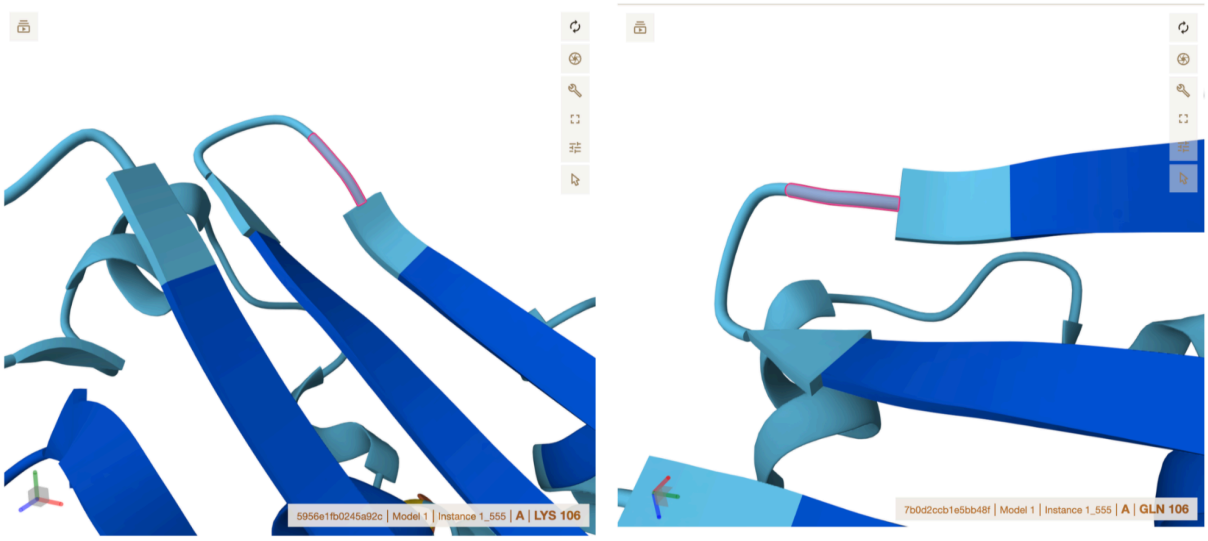

Figure 3. AlphaFold-predicted SMC3 wild-type and K105Q/K106Q acetylation-mimic models at residue 106. Structures were visualized in the AlphaFold Server structure viewer using ribbon representation, with ribbon colors indicating prediction confidence (pLDDT). Residue 106 is highlighted in pink/red. Left: wild-type (K106). Right: acetylation mimic (Q106).

At SMC3 residue 106, the matched models likewise show no obvious local secondary structure transition. DSSP assigned residue 106 as coil/unassigned in both pair models, its Cα displacement after local alignment was 0.074 Å, and pLDDT values were similar (85.78 in the wild-type and 86.81 in the K to Q model). Together with the 0.044 Å local RMSD across residues 95-115, these measurements support the conclusion that no detectable local structural change was predicted around K105/K106.

Comparison of SMC3 K105 and K106 effects

When residues 105 and 106 are considered together, neither the ribbon-model comparison nor the quantitative analysis identifies a reproducible local secondary structure conversion. The local Cα RMSD across residues 95–115 was only 0.044 Å, the target-residue deviations were small, and DSSP assignments at both sites were unchanged between the wild-type and the acetylation mimic model. Therefore, the most defensible result is that the K105Q/K106Q substitutions did not produce a detectable local structural change in the static AlphaFold predictions.

Structural comparison of SMC1A at residue K579

Figure 4 shows the structural comparison of SMC1A at residue 579, with the wild-type structure on the left and the acetylation mimic structure on the right. In the wild-type model, residue 579 is located within a region composed of alpha-helices and coiled coil structures. These are consistent with the known structure of SMC proteins.

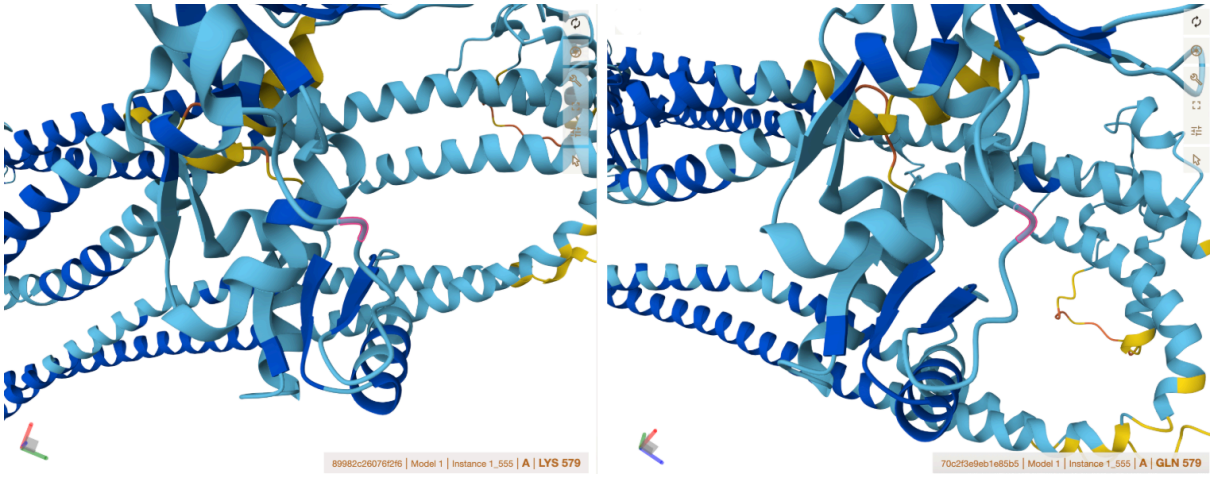

Figure 4. AlphaFold-predicted SMC1A wild-type and K579Q/K584Q acetylation-mimic models at residue 579. Structures were visualized in the AlphaFold Server structure viewer using ribbon representation, with ribbon colors indicating prediction confidence (pLDDT). Residue 579 is highlighted in pink/red. Left: wild-type (K579). Right: acetylation mimic (Q579).

In the acetylation mimic model, the same region maintains the same arrangement of alpha-helices and coiled coils. No visible transition from structured elements to turns is observed. The overall shape, orientation, and secondary structure composition of the region surrounding residue 579 appears to be the same between the two models.

Structural comparison of SMC1A at residue K584

Figure 5 presents the structural comparison of SMC1A at residue 584 with the wild-type structure shown on the left and the acetylation mimic structure on the right. Just like residue 579, residue 584 is located within a structured region dominated by alpha-helices and coiled coils.

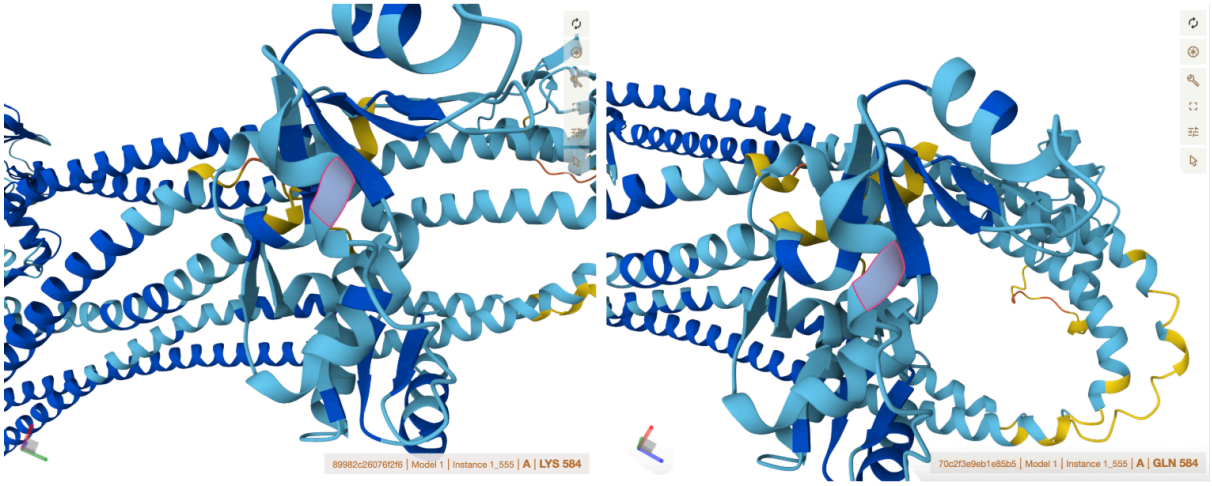

Figure 5. AlphaFold-predicted SMC1A wild-type and K579Q/K584Q acetylation-mimic models at residue 584. Structures were visualized in the AlphaFold Server structure viewer using ribbon representation, with ribbon colors indicating prediction confidence (pLDDT). Residue 584 is highlighted in pink/red. Left: wild-type (K584). Right: acetylation mimic (Q584).

It can be seen that the wild-type and the acetylation mimic models have similar secondary structure composition in this region. No clear changes can be observed in regards to the loss of β sheets or helices and no increase in turn-like regions is observed. The local fold is the same in both models.

The SMC1A wild-type and acetylation mimic (K to Q) ribbon models also show no obvious local geometry or secondary structure change at residues 579 and 584. Quantitatively, residues 569-594 had a local Cα RMSD of 0.116 Å. Residue 579 had a Cα deviation of 0.099 Å and remained coil/unassigned by DSSP in both models. pLDDT was essentially unchanged at the two sites (K579: 86.11 to 86.11; K584: 87.88 to 87.93). These results support describing SMC1A as showing no detectable local structural change in this static modeling setup.

Comparison of acetylation effects between SMC3 and SMC1A

Neither SMC3 nor SMC1A showed a clear local structural change in the correctly matched ribbon models or in the quantitative analysis. Both proteins had very small local Cα RMSDs and unchanged DSSP assignments at the exact sites that were examined. Whole protein Cα RMSDs were larger (5.37 Å for SMC3 and 10.28 Å for SMC1A), but the structures were generated independently with different model seeds, so these global differences should not be attributed to the acetylation mimic substitutions. The models therefore do not support the claim that SMC3 is structurally more sensitive than SMC1A. SMC3 is still the biologically relevant focus because its acetylation by ESCO2 is established, not because the AlphaFold comparison detected a larger structural response.

Predicted aligned error map comparison analysis for SMC3

Figure 6 shows Predicted Aligned Error (PAE) maps for the SMC3 wild-type and K to Q models. The wild-type model is on the left and the acetylation mimic is on the right. PAE estimates the expected positional error for one residue when the predicted structure is aligned on another residue. It measures confidence in relative residue or domain positioning rather than residue-residue correlation, interaction, or biological behavior.

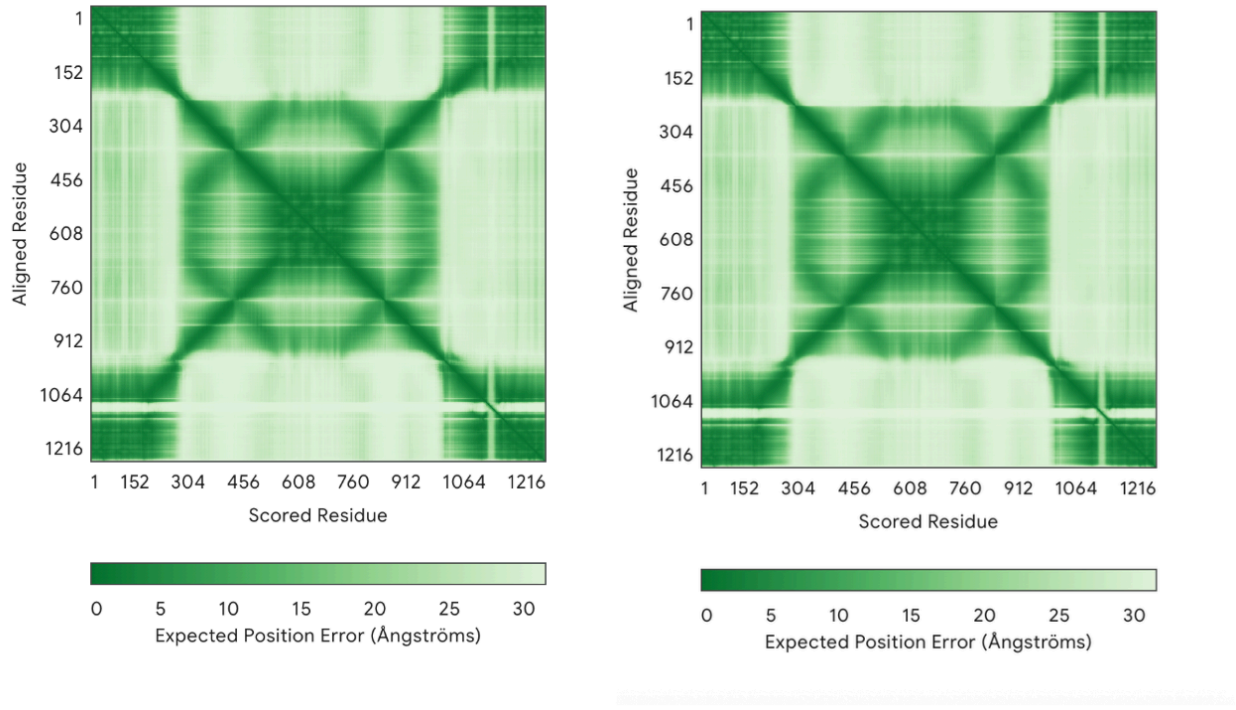

Figure 6. Predicted Aligned Error (PAE) maps for the SMC3 wild-type and K105Q/K106Q acetylation-mimic models generated by AlphaFold Server. Darker green indicates lower expected positional error and greater confidence in relative residue positioning, while lighter green indicates higher expected positional error. Left: wild-type. Right: acetylation mimic.

In the residue-residue comparison maps above, each axis represents the amino acid positions in the SMC3 protein sequence. The color scale signifies the predicted aligned error between pairs of residues. Darker green regions represent lower expected positional error (the program has higher confidence in their relative positioning) and lighter regions suggest greater uncertainty. Blocks of low error (dark green) typically belong to domains that are well defined structurally while higher error regions (light green) often indicate flexible linkers or regions with lower structural confidence.

The broad similarity of the maps can be described qualitatively, but small local differences should not be treated as structural evidence on their own. The summary-confidence outputs also show very similar

overall pTM scores (0.49 for wild-type and 0.48 for the K to Q model), which are consistent with broadly comparable model-level confidence.

Predicted aligned error map comparison analysis for SMC1A

Figure 7 shows the PAE maps for SMC1A with wild-type and acetylation mimics shown side-by-side. The overall PAE patterns for the two models look nearly identical across the protein.

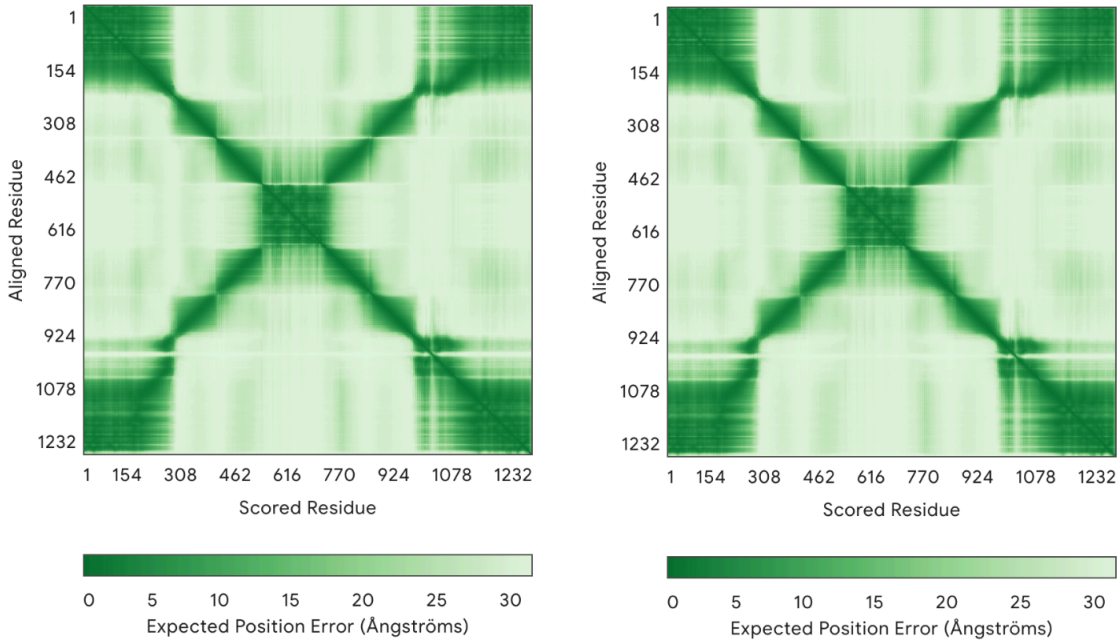

Figure 7. Predicted Aligned Error (PAE) maps for the SMC1A wild-type and K579Q/K584Q acetylation-mimic models generated by AlphaFold Server. Darker green indicates lower expected positional error and greater confidence in relative residue positioning, while lighter green indicates higher expected positional error. Left: wild-type. Right: acetylation mimic.

The SMC1A wild-type and K to Q PAE maps should likewise be described as model-confidence maps. The summary pTM scores were 0.47 for wild-type and 0.46 for the K to Q model, indicating little change in this global confidence metric. PAE and pTM do not establish that residue interactions are identical or that the acetylation has no functional effect.

DISCUSSION

This work aimed to identify structural changes that could help explain why defects in cohesin acetylation are associated with disorders like Roberts Syndrome by comparing the wildtype and acetylation mimic

models of proteins SMC1A and SMC3. Results show that the tested acetylation-mimic substitutions did not produce detectable local structural changes in the static SMC3 or SMC1A AlphaFold predictions. Quantitative re-analysis added an important second layer of interpretation, however. The SMC3 local RMSD was 0.044 Å and DSSP assigned the same secondary-structure states at residues 105 and 106 in both paired CIF models. SMC1A similarly showed a local RMSD of 0.116 Å with unchanged assignments at residues 579 and 584. These results indicate that the tested K to Q substitutions did not produce detectable local structural changes in the static AlphaFold predictions. This finding should not be used as evidence that acetylation has no biological consequence. Instead, it should be interpreted as a limitation of what was detected by the modeling approach.

Quantitative structural comparison of the SMC3 K to Q models

The main SMC3 finding is the absence of a local structural difference between the wild-type and K105Q/K106Q models. After local alignment of residues 95-115, the Cα RMSD was 0.044 Å, with residue level deviations of 0.106 Å at position 105 and 0.074 Å at position 106. DSSP assigned the same state at the exact positions in both paired models, while pLDDT values were similar. Therefore, this analysis does not provide evidence for a detectable static local conformational change at K105/K106. The biological importance of this region still comes from the established role of SMC3 acetylation in cohesin regulation, and effects on dynamics, interactions, or function may occur without appearing as a large static structural change in an isolated prediction.

These results should not be treated as contradicting experimental evidence that SMC3 acetylation supports stable cohesin function. The analysis did not detect a robust local structural change, and it does not measure protein stability, chromatin association, ATPase behavior, or mechanical strength. A lack of detectable change in these static predictions is compatible with acetylation acting through dynamics, intermolecular interactions, or chemical effects that a K to Q substitution and isolated AlphaFold model do not capture.

Limited structural effects of acetylation in SMC1A

The acetylation mimic mutations in protein SMC1A at residues K579 and K584 did not produce any visible changes in secondary structure. The alpha-helices and coiled coil regions that surround these residues were the same in the wild-type and the acetylated models.

SMC1A was included as an exploratory comparator and is not presented as a validated Roberts syndrome acetylation target. Its K579Q/K584Q local comparison was also highly similar to the wild-type (local Cα RMSD 0.116 Å), with unchanged DSSP assignments and essentially unchanged pLDDT at the modified sites.This result should be reported simply as an absence of a c;ear local structural difference in this modeling setup, not as any proof that SMC1A is biologically insensitive to acetylation.

Interpretation of residue-residue comparison maps

The PAE maps provide information about confidence in relative residue and domain positioning, not residue interactions or structural correlation. Their overall similarity, together with the nearly unchanged

pTM scores (SMC3 0.49 to 0.48; SMC1A 0.47 to 0.46), suggests that the K to Q jobs did not produce a major change in these model level confidence metrics. However, PAE and pTM cannot determine whether a post-translational modification changes real protein dynamics since this is all being done in a static model. These should be interpreted together with local alignment, DSSP assignment, and pLDDT rather than as independent evidence of an acetylation effect.

Implications for cohesin mechanics and function

Because the quantitative analysis did not identify a robust local secondary structure change at the sites that were tested, no conclusion about altered flexibility, force resistance, ATPase behavior, DNA binding, or chromatin association can be drawn from these models. Instead, the results define a limit of this static, K to Q, and AlphaFold approach. Molecular dynamics simulations using chemically appropriate acetyl-lysine parameters, modeling of the larger cohesin complex, and experiments would be needed to test whether acetylation changes SMC3 dynamics or functions in ways that are not visible in these static and isolated models.

Limitations and future directions

There are several limitations that should be considered when interpreting these results. First, AlphaFold only generates static structural predictions and does not model protein dynamics or interactions with DNA and other cohesin subunits. Hence, changes in flexibility of binding behavior may not be fully captured. Second, lysine to glutamine (k to Q) substitutions serve as an approximation of acetylation rather than a direct chemical modification which could also cause the protein to react differently than it did in this study. Third, each wild-type and acetylation mimic structure was generated in a separate AlphaFold Server run with a different model seed, so whole protein differences may include run to run prediction variability. Fourth, the models are isolated, static protein predictions and do not include the dull cohesin complex, DNA, ATP-dependent conformational cycling, or protein dynamics. The qualitative and quantitative analyses are consistent. Neither the correctly matched ribbon models nor DSSP/local Cα alignment identified a reproducible local structural transition at the tested residues. The absence of a predicted structural change should not be interpreted as evidence that true acetylation has no effect, because acetylation may alter dynamics, interactions, electrostatics, or complex-level behavior that are not captured here.

Future works could address these limitations by incorporating molecular dynamics simulations to explore how acetylation affects flexibility in proteins and their mechanical response over time. In addition, modeling larger portions of the cohesin complex including DNA could provide a more complete picture of how acetylation influences cohesin mechanics and therefore function.

Comparison with previous literature

These findings suggest that acetylation does not necessarily trigger localized structural destabilizations as well as flexibility in SMC3. Previous studies have come to the conclusion that acetylation stabilizes cohesion by promoting chromatin association and preventing premature release (Song et al., 2012; Nishiyama, 2019). In this study, however, the K105Q’K106Q acetylation mimic model did not show a

detectable local structural change compared with the wild-type SMC3 model. This finding does not contradict previous experimental work because the substitution being made (K to Q) is only an approximation of acetylation and AlphaFold does not model the full chemical modification, protein dynamics, or interactions within the cohesin complex as a whole.

Structural studies have shown that cohesin functions as a dynamic ring-like complex that is formed by SMC1A, SMC3, RAD21, and associated regulatory proteins (Gligoris et al., 2014; Gruber et al., 2003). Due to this, the functional effects of SMC3 acetylation may occur through changes in dynamics, intermolecular relationships, ATP-dependent conformational behavior, or chromatin association without producing a large detectable local structural change in a static model that is also isolated. Therefore, the results of this study emphasize the need for molecular-dynamics simulations and experimental approaches to further investigate acetylation-dependent effects that may not be captured by AlphaFold predictions**.**

CONCLUSION

Through computational modeling, this study investigated the influence of acetylation-related modifications on the cohesin subunits SMC1A and SMC3. This study aimed to identify structural changes that could explain how defects in cohesin acetylation contribute to disorders such as Roberts Syndrome by modeling wild-type and acetylation mimics of each protein using AlphaFold Server. A focus on specific lysine residues at the established SMC3 acetylation sites K105/K106 and at SMC1A K579/K584 as an exploratory comparison was done.

No local geometry or secondary structure change was observed at the tested regions in either protein. Quantitative analysis supported this result. The SMC3 the SMC3 local region (residues 95–115) showed a Cα RMSD of 0.044 Å, unchanged DSSP assignments at residues 105 and 106, and similar pLDDT values between wild-type and the K to Q model. Likewise, SMC1A showed a small local RMSD of 0.116 Å across residues 569–594, unchanged secondary-structure assignments at residues 579 and 584, and nearly identical local confidence values. Therefore, within the limitations of this acetylation mimic and AlphaFold Server approach, the tested substitutions did not produce detectable local structural changes. This result does not imply that true acetylation has no functional effect, due to AlphaFold not modeling full chemistry, dynamics, or complex-level interactions involved in cohesin regulation. SMC3 remains the biologically relevant focus because ESCO2 mediated SMC3 acetylation is established, while SMC1A remains an exploratory comparison. Future molecular dynamics and experimental studies are needed to test whether acetylation produces changes in SMC3 dynamics, interactions, or functions that were not captured by the static models.

Data Availability

The protein sequences that were analyzed in this study were obtained from publicly available databases which include UniProt. All protein structures were generated using the AlphaFold protein structure prediction server (last accessed February 2026) based on these sequences.

Acknowledgements

I would like to deeply thank my research mentor, Ms. Brenda Rubenstein, for guidance throughout this project and for their support throughout this whole process. I am also grateful to the Lumiere Research Program for providing me with the opportunity and resources to conduct this research.

REFERENCES

Alomer, Ramon M., et al. "Esco1 and Esco2 Regulate Distinct Cohesin Functions during Cell Cycle Progression." *Proceedings of the National Academy of Sciences*, vol. 114, no. 37, 2017, pp. 9906–9911, https://doi.org/10.1073/pnas.1708291114.

Gabriele, Michele, et al. "Dynamics of CTCF- and Cohesin-Mediated Chromatin Looping." *Science*, vol. 376, no. 6592, 2022, pp. 496–501, https://doi.org/10.1126/science.abn6583.

GeneReviews. "ESCO2 Spectrum Disorder." *GeneReviews®*, edited by Adam, Margaret P., et al., University of Washington, Seattle, updated 2023, National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/books/NBK1153/.

Gligoris, Thomas G., et al. "Closing the Cohesin Ring: Structure and Function of Its Smc3-Kleisin Interface." *Science*, vol. 346, no. 6212, 21 Nov. 2014, pp. 963–967, https://pubmed.ncbi.nlm.nih.gov/25414305/.

Gruber, Stephan, et al. "Chromosomal Cohesin Forms a Ring." *Cell*, vol. 112, no. 6, 2003, pp. 765–777, https://doi.org/10.1016/S0092-8674(03)00162-4.

Horsfield, Janine A., et al. "A Brief History of Cohesin and Gene Regulation." *FEBS Journal*, vol. 290, no. 2, 2023, pp. 337–356, https://doi.org/10.1111/febs.16271.

Jumper, John, et al. "Highly Accurate Protein Structure Prediction with AlphaFold." *Nature*, vol. 596, 2021, pp. 583–589, https://doi.org/10.1038/s41586-021-03819-2.

Kline, A. D., et al. "Cornelia de Lange Syndrome." *Nature Reviews Genetics*, vol. 19, 2018, pp. 607–622, https://doi.org/10.1038/s41576-018-0031-0.

Mfarej, Mohamad G., et al. "An Ever-Changing Landscape in Roberts Syndrome Biology." *PLOS Genetics*, vol. 16, no. 6, 2020, e1008866, https://doi.org/10.1371/journal.pgen.1008866.

Nishiyama, Tomohiro. "Cohesin and Chromatin Organization." *Current Opinion in Cell Biology*, vol. 58, 2019, pp. 8–14, https://doi.org/10.1016/j.ceb.2019.02.004.

Song, Jin, et al. "SMC3 Acetylation Is Essential for Sister Chromatid Cohesion." *Journal of Biological Chemistry*, vol. 287, no. 46, 2012, pp. 38505–38514, https://doi.org/10.1074/jbc.M112.395731. Vega, H., et al. "Roberts Syndrome Is Caused by Mutations in ESCO2, a Human Homolog of Yeast Eco1." *Nature Genetics*, vol. 37, 2005, pp. 468–470, https://doi.org/10.1038/ng1548.

Villa-Hernandez, Sara, and Rodrigo Bermejo. "Cohesin Dynamic Association to Chromatin and Interfacing with Replication Forks in Genome Integrity Maintenance." *Current Genetics*, vol. 64, 2018, https://doi.org/10.1007/s00294-018-0824-x.

Tunyasuvunakool, Kathryn, et al. "Highly Accurate Protein Structure Prediction for the Human Proteome." *Nature*, vol. 596, 2021, pp. 590–596, https://doi.org/10.1038/s41586-021-03828-1.