The Coiled Coils of Cohesin Are Conserved in Animals, but Not In Yeast

Glenn E. White1, Harold P. Erickson2*

1 Department of Biological and Environmental Sciences, Longwood University, Farmville, Virginia, United Kingdom, 2 Department of Cell Biology, Duke University Medical Center, Durham, North Carolina, United States of America

Abstract

**Background:** The SMC proteins are involved in DNA repair, chromosome condensation, and sister chromatid cohesion throughout Eukaryota. Long, anti-parallel coiled coils are a prominent feature of SMC proteins, and are thought to serve as spacer rods to provide an elongated structure and to separate domains. We reported recently that the coiled coils of mammalian condensin (SMC2/4) showed moderate sequence divergence (≈10–15%) consistent with their functioning as spacer rods. The coiled coils of mammalian cohesins (SMC1/3), however, were very highly constrained, with amino acid sequence divergence typically <0.5%. These coiled coils are among the most highly conserved mammalian proteins, suggesting that they make extensive contacts over their entire surface.

**Methodology/Principal Findings:** Here, we broaden our initial analysis of condensin and cohesin to include additional vertebrate and invertebrate organisms and multiple species of yeast. We found that the coiled coils of SMC1/3 are highly constrained in *Drosophila* and other insects, and more generally across all animal species. However, in yeast they are no more constrained than the coils of SMC2/4 and Ndc80/Nuf2p, suggesting that they are serving primarily as spacer rods.

**Conclusions/Significance:** SMC1/3 functions for sister chromatid cohesion in all species. Since its coiled coils apparently serve only as spacer rods in yeast, it is likely that this is sufficient for sister chromatid cohesion in all species. This suggests an additional function in animals that constrains the sequence of the coiled coils. Several recent studies have demonstrated that cohesin has a role in gene expression in post-mitotic neurons of *Drosophila*, and other animal cells. Some variants of human Cornelia de Lange Syndrome involve mutations in human SMC1/3. We suggest that the role of cohesin in gene expression may involve intimate contact of the coiled coils of SMC1/3, and impose the constraint on sequence divergence.

Introduction

The structural maintenance of chromosome (SMC) proteins interact with DNA to carry out several critical functions within the cell including DNA repair, chromosome condensation, and sister chromatid cohesion during mitosis [1–8]. Three SMC protein complexes have been identified in eukaryotes: SMC1/3 (cohesin), SMC2/4 (condensin), and SMC5/6. Each SMC complex consists of two SMC protein subunits and a varying number of accessory proteins. The SMCs are structurally characterized by a long, anti-parallel coiled coil that exhibits the hydrophobic packing of Crick [9]. The SMC subunits fold at the central hinge, and the coiled coils bring the N- and C-terminal domains together to form the ATPase head domain (Fig. 1). The two SMC subunits of each pair form a dimer, with the hinges associated in the middle and the ATPase domains at the ends of the long coiled-coil rods (Fig. 1).

Research efforts on the function of cohesin (SMC1/3 and two associated proteins) have focused primarily on its well-documented role in holding sister chromatids together until they are ready for separation in anaphase [10–15], and recently to a role in mitotic spindle pole formation [16]. During mitosis, SMC1 complexes with SMC3 to form cohesin (we use SMC1 to designate SMC1p (yeast) and SMC1A or SMC1L1 (animals)). During meiosis in animals, a second isoform (SMC1β) complexes with SMC3 to form a second cohesin [17,18]. An attractive model for the mechanism of sister chromatid cohesion proposes that the ATPase heads come together and the long coiled coils form a ring that could trap the DNA of each sister chromatid [5,19], but see also [20]. In this model, and indeed in most thinking about the possible function of cohesin and condensin, the coiled coils are thought to serve a structural role as long spacer arms. However, we recently reported that the mammalian cohesins possess the most highly conserved coiled-coil domains in the genome [21], suggesting that they are not just spacer arms (discussed below).

Emerging evidence demonstrates that cohesin also has important functions in interphase cells, independent of its role in cohesion. Cohesin binds to specific chromosomal sites across the genome, and is involved in regulating gene expression in post-mitotic cells. In yeast, the cleavage of the Scc1/Mdc1p subunits of cohesin play a role in transcriptional silencing [22] and SMC1 and SMC3 may be a part of the mechanism defining the boundaries of the transcriptionally silenced HMR locus [23]. Cohesin mutations...
Cohesin's Coiled Coils

Figure 1. SMC protein structure. SMC proteins have a hinge region in the middle of the sequence where the protein folds to form long segments of anti-parallel coiled coils. This brings the N-terminus (Nt) and C-terminus (Ct) together to form the ATPase heads. The hinge region of each protein interacts to form the heterodimer. The head of SMC1 can bind the head of SMC3 to make a ring. The linked heads are stabilized by binding of accessory proteins (not shown). The coiled coils are drawn to scale.

doi:10.1371/journal.pone.0004674.g001

in the coiled coils of both the SMC1 and SMC3 subunits were discovered in patients with a mild variant of Cornelia de Lange Syndrome (CdLS) [24–27] – a human neurological developmental disorder associated with mental retardation. In Drosophila, studies with cohesin mutants demonstrated that cohesin is required for developmental axonal pruning [28] and abnormal cohesin complex cleavage alters wild type larval locomotion [29]. Finally, cohesin binding sites and those of the zinc-finger protein CCCTC-binding factor (CTCF), which acts as an insulator protein that blocks enhancer-promoter interactions, overlap significantly in mammalian chromosomes ([30–34]; see [33,34] for commentary). CTCF is found in Drosophila and vertebrates [35]. 9,000 cohesin-binding sites were found in the human genome, and 90% of them also bound CTCF [31–33]. Additionally, this post-mitotic function for cohesin appears not to be limited to neurons as immunoblots demonstrate that cohesin is differentially expressed in a variety of murine tissue extracts [32]. Overall, these studies demonstrate that cohesin functions in regulating gene expression in addition to its role in sister chromatid cohesion.

In a previous study, we noted that the coiled-coil segments of SMC1 and SMC3 were among the most highly conserved mammalian proteins, showing sequence divergence across different mammalian species of only 0–1% over ~700 amino acids [21]. To put this in context, we analyzed a variety of coiled-coil proteins. Some coiled coils (Ndc80, Nufl2p, giantin) showed sequence divergence of ~20% across different species of mammals, and we concluded that this was a typical sequence divergence of coiled coils serving primarily as inert spacer rods. Ndc80/Nuf2p appear to be found in all eukaryotes and form a heterodimeric coiled coiled that is involved in linking the kinetochore to microtubules [36–38]. The coiled coils of condensins (SMC2/4) showed 10–15% sequence divergence, suggesting that they are also serving primarily as rods. Proteins that are known to use their coiled coils for packing into filaments, such as skeletal muscle myosin II and intermediate filament proteins, showed a divergence of only 1–3%, reflecting the constraints to maintain the protein-protein contacts over their surface. The coiled coils of SMC1/3 were even more highly conserved than myosin II and intermediate filament proteins, which implied these coiled coils are not serving just as spacer rods. We concluded that the coiled coils of SMC1/3 probably have interactions over their entire length and circumference that constrain sequence divergence.

We also found that the coiled coils of SMC1/3 were much more highly conserved than those of SMC2/4 when we compared human sequences to avian, amphibian, and Drosophila orthologs [21]. This suggested that the mechanism imposing the constraint in SMC1/3 coiled coils was found in all vertebrates, and perhaps all animal species, and raised the question whether this mechanism is found in all eukaryotes. As genomic sequences have become increasingly available for a number of yeast, nematode, Drosophila, and other invertebrate species [39–43], we decided to extend our analysis to yeast and additional animal species. We determined that the coiled coils of SMC1/3 are more conserved than those of SMC2/4 and Ndc80/Nuf2p in vertebrates and several species of invertebrates, but are no more conserved than those of SMC2/4 and Ndc80/Nuf2p across multiple yeast species. In addition, the meiotic SMC1β coils are not constrained in animals.

Tropomyosin, an actin binding coiled-coil protein found throughout Eukaryota served as our control reference sequence of a coiled coil that is universally highly conserved [9,44,45]. Tropomyosin is an essential element of the thin filament of striated muscle while its non-muscle isoforms have been implicated in a number of functions including actin filament stability and cytokinesis [46,47]. Its sequence conservation is apparently due to its binding to actin and the tropinin proteins, involving most of the surface of the coiled coil. Our finding that SMC1/3 coils are not constrained in yeast suggests that there is a special mechanism involving the SMC1/3 coils that operates in metazoans but not in yeast nor in animal cells undergoing meiosis.

Methods

Protein Sequence Acquisition and Coiled-Coil Domain Determination

Seven coiled-coil proteins (SMC1/3, SMC2/4, Ndc80, Nufl2p, tropomyosin (Tm)) from Saccharomyces cerevisiae, Caenorhabditis elegans, Drosophila melanogaster, and Homo sapiens served as the reference sequences for this study. Protein sequences for each were obtained from GenBank with the majority annotated as Reference Sequences. The remaining sequences were obtained by using these sequences as the query in either a BLAST search of GenBank's protein database, or a Eukaryota Genomic BLAST (tBLASTn) search of the whole genome shotgun sequence databases (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism = euk). For those species which only had whole genome shotgun sequences available, orthogonal proteins were identified based upon the highest % identity with the query sequence for that organism, and that their coiled-coiled domains, as identified using the COILS prediction program (see below), closely approximated those of the query sequences. The gene sequences were then translated using the SIXFRAME tool at Biology Workbench at the San Diego Super Computer (http://seqtool.sdsc.edu/CGI/BW.cgi).

The coiled-coil domains of each protein were identified using the 28-residue window output from the COILS program (http://www.ch.embnet.org/software/COILS_form.html) and to maintain consistency with our previous study, only those regions of the sequence where the score was ≥0.5 were used for the analysis of the coils [21,48]. All SMC proteins in this study had disruptions of their coiled coils [49]. The boundaries of the individual coiled-coil segments were identified and the non-coil segments were removed. For a select subset of SMC proteins, coiled coil boundaries were determined using other coiled coil prediction software including
Cohesin's Coiled Coils

PairCoil2 ([50]; http://groups.csail.mit.edu/ch/paircoil2/) and PCOILS ([51]; http://toolkit.tuebingen.mpg.de/pcolls). All three programs predicted slightly different boundaries for the coiled-coil segments, and using our analysis criteria, resulted in slightly different percent divergences. However, the sequence divergences never changed more than 1 or 2 percent, so the trends we report using the COILS program (in this and our previous study), are not affected by the small differences in boundaries.

For ease of viewing, the SMC1/3, SMC2/4, Ndc80, NuF2p and tropomyosin sequences are grouped by organism. For those sequences that served as the reference sequences in this study, the following additional information is provided: 1) protein length (or noted as a partial sequence), and 2) the boundaries of the coiled-coil segments as determined by COILS. All remaining sequences are listed by their GenBank accession number only with translated ORFs and partial length sequences reported as such. In some instances, the protein sequence was reconstructed from multiple genomic segments, and that is noted.

As all sequences (except for those from Ciona intestinalis) were obtained from GenBank, their accession numbers are provided.

Yeast

S. cerevisiae SMC1 (= SMC1p) [NP_116647.1, 1225 amino acids, N and C Coil: G17 - T212, E328 - S323, Q320 - T337, Y395 - I441, S456 - L496, S707 - Y745, K795, N805 - N943, K807 - E1065], SMC3 [NP_012461.1, 1230 amino acids, N and C Coil: A172 - Y230, S259 - Q347, R358 - K387, I328 - S592, A777 - K1020], and Tm [AACI02000420.1, translated ORF], Ndc80 [AACI02000158.1, translated ORF], NuF2p [AACH10006641.1, translated ORF], and Tm [AACH10006641.1, translated ORF].

S. paradoxus SMC1 [AABY01000015.1, translated ORF], SMC3 [AABY01000267.1, translated ORF], SMC2 [AABY01000211, translated ORF], SMC4 [AABY01000051.1, translated ORF], Ndc80 [AABY01000004.1, translated ORF], NuF2p [AABY01000044.1, translated ORF], and Tm [AABY01000091.1, translated ORF, (partial)].

Abyba gyosypsi ATCC 10895 SMC1 [NP_986643.1], SMC3 [NP_982960.1], SMC2 [NP_986902], SMC4 [NP_986754], Ndc80 [Not available], NuF2p [NP_984846.1], and Tm [NC_005786.1 (partial)].

Candida glabrata XCP [NP_446063.1], XMP [NP_446916.1], XMC2 [NP_445638.1], XMC4 [NP_449298.1], XNdc80 [NP_445332.1], NuF2p [NP_448216.1], and Tm [NC_006033.1].

Klugeryonimyces lactis XCP1 [XP_453397.1], XMC3 [XP_451012.1], XMC2 [XP_453765], XMC4 [XP_454536.1], NuF2p [XP_454608.1], and Tm [NC_006041.1 (partial)].

Klugeryonimyces waristi XMC1 [AADM01000020, translated ORF], XMC3 [AADM01000235.1, translated ORF], XMC2 [AADM01000229.1, translated ORF], XMC4 [AADM01000282.1, translated ORF], Ndc80 [AADM01000214.1, translated ORF, (partial)], NuF2p [AADM01000270.1, translated ORF, (partial)], and Tm [AADM01000003.1, translated ORF (partial)].

Nematode

C. elegans SMC1 [NP_001040658.1, 1281 amino acids, N and C Coil: T164 - L235, H238 - Q408, L417 - A499, A673 - L760, R763 - F793, N810 - L933, N9143 - L1108], SMC3 [NP_4949532.1, 1261 amino acids, L180 - V334, K394 - F452, Q471 - A503, D675 - F750, E758 - L836, K842 - L947, R1023 - A1068], SMC2 [NP_496331.1, 1244 amino acids, N and C Coil: C244 - L303, Q306 - K353, I415 - E479, Q702 - E954, S1004 - E1044], XMC4 [NP_497935.1, 1549 amino acids, N and C Coil: M326 - T490, A492 - K604, K786 - S832, A844 - R1058, E1144 - R1213], XNdc80 [NP_501830.1, 590 amino acids, N and C Coil: C246 - N296, H298 - H342, K347 - E376, R442 - I471, R488 - D525, S527 - A555], NuF2p [Q21925.1, 490 amino acids, N and C Coil: A136 - V242, E247 - 1283, E303 - N412], and Tm [NP_001021698.1, 284 amino acids, Coil: M1 - V284 (all)].

C. briggsae SMC1 [CAE73105.1 (partial)], XMC3 [XP_001664944.1], XMC2 [XP_001665682.1], XMC4 [NP_001040658.1], Ndc80 [CAPE39778.1, (partial)], NuF2p [XP_001664731.1], and Tm [CAQ900701].

Brugia malayi SMC1 [EDP37102.1 (partial)], XMC3 [EDP31459.1], XMC2 [EDP3066.1], XMC4 [Not available], Ndc80 [XP_001894611.1], NuF2p [XP_001895760.1 (partial)], and Tm [AAQ01004217.1, NZ_AAQA01000177.1, NZ_AAQA01000167], reconstructed protein sequence from translated ORF, (partial).

Insect

Drosophila melanogaster SMC1 [CAB76376.1, 1238 amino acids, N and C Coil: L186 - Q247, H233 - R318, L320 - V526, H683 - S804, V811 - N951, N1037 - R1082], SMC3 [NP_523374, 1200 amino acids, N and C Coil: E173 - Q397, N397 - A301, N671 - D730, K756 - N334, A446 - A498, R498 - P989], SMC2 [NP_610995.1, 1179 amino acids, N and C Coil: Q173 - E201, K255 - M292, T312 - Q383, T389 - Q432, D441 - Q507, A580 - L724, M726 - K782, R785 - I939, V988 - H1017], SMC4 [NP_723996.1, 1409 amino acids, N and C Coil: L272 - F307, I328 - S592, A777 - K1020], Ndc80 [NP_572902.1, 675 amino acids, N and C Coil: Q173 - E201, K255 - M292, T312 - Q383, T389 - Q432, D441 - Q507, A580 - L724, M726 - K782, R785 - I939, V988 - H1017], SMC4 [NP_572902.1, 675 amino acids, N and C Coil: Q173 - E201, K255 - M292, T312 - Q383, T389 - Q432, D441 - Q507, A580 - L724, M726 - K782, R785 - I939, V988 - H1017], SMC4 [NP_572902.1, 675 amino acids, N and C Coil: Q173 - E201, K255 - M292, T312 - Q383, T389 - Q432, D441 - Q507, A580 - L724, M726 - K782, R785 - I939, V988 - H1017].
acids, N and C Coil: K302 - L429, D525 - D539, R565 - L662, Nuf2p [AAC56691.1, 502 amino acids, Coil: E205 - R303, T366 - T432, L449 - I406] and Tm [NP_732012.1, 204 amino acids, Coil: M1 - Y204].

Drosophila sechellia SMC1 [AAKO01000393.1, translated ORF], SMC3 [AAKO01002722.1 and AAK001000647.1, reconstructed protein sequence from translated ORF], SMC2 [AAKO01000647.1, reconstructed protein sequence from translated ORF], SMC4 [AAKO01001629.1, translated ORF, (partial)], Ndc80 [AAKO01002741.1, translated ORF, (partial)], Nuf2p [AAKO01000021.1, (partial)], and Tm [XP_002031091.1, 1064 amino acids, N and C Coil: L142 - V172, Q211 - D284, S288 - Q880, S883 - K934, D967 - E994, R1022 - R1049], Ndc80p [XP_217011.3, Ndc80p [XP_217489.3], Nuf2p [NP_001012023.1], and Tm [NP_001029240.1].

Gallus gallus SMC1 [NP_998947.1], SMC3 [NP_998948.1], SMC2 [NP_996561.1], SMC4 [NP_999049.1], SMC1β [XP_416467.2], Ndc80 [NP_998908.2], Nuf2p [Q76590], and Tm [P04268.2].

Xenopus laevis SMC1 [NP_001030490.1], SMC3 [CAD59446.1], SMC2 [PS0533], SMC4 [PS0532], SMC1β [Not available], Ndc80 [NP_001008237.1], Nuf2p [NP_001008237.1], and Tm [NP_001125846.1].

Danio rerio SMC1 [NP_688120], SMC3 [NP_999854.1], SMC2 [XP_00191036.1], SMC4 [NP_775360.1], SMC1β [XP_001334257.1, Ndc80 [NP_001003963.1], Nuf2p [NP_956604.1], and Tm [NP_571180.1].

Strongylocentrotus purpuratus SMC1 [XP_001177541.1], SMC3 [XP_001202053.1 and XP_001176817, reconstructed protein sequence, (partial)], SMC2 [Not available], SMC3 [XCAP-C] [XP_797583.2], Ndc80 [NP_001197919.1], and Tm [XP_001202479.1, (partial)].

Ciona intestinalis SMC1 [Ciona:290909], SMC3 [Ciona:236566], SMC2 [Ciona:269866], SMC4 [Ciona:259758, (partial), Nuf2p, Nuf2p [Not available], and Tm [Not available]]. Sequences were downloaded from the Ciona intestinalis v2.0 website: http://genome.jgi-psf.org/Cioin2/Cioin2.home.html.

Homo sapiens SMC1 [NP_001092481.1], and Tm [NP_001013608.1]. Sequences were available for the first four, which are the closest to S. cerevisiae. The analysis is most informative for the first four, which are the closest to S. cerevisiae (members of Saccharomyces sensu stricto [40], and which diverged approximately 10-20 MYA (5.26K generations/year) [52,53], but the conclusions are similar for the more divergent species. The coils of SMC1/3 are no more conserved than the coils of SMC2/4 and Ndc80/Nuf2p. Our reference coiled-coil protein tropomyosin, however, is highly constrained across these same species, providing a benchmark coiled coil sequence within Saccharomyces. The
fact that tropomyosin’s coils are highly constrained seems consistent with its recently demonstrated essential role in cytokinesis [47], and the involvement of much of its surface in binding to actin.

Table 2 shows the same comparisons for insects and nematodes. Based upon the Drosophila phylogeny recently generated from the analysis of whole genome shotgun sequences, three species of Drosophila were selected for comparison to D. melanogaster: D. sechellia (time to last common ancestor (TLCA)=1.2 MYA [10 generations/year]), D. pseudoobscura (TLCA=24 MYA), and D. grimshawi (TLCA=40 MYA) [43, 54, 55]. In several insect species, the Ndc80/Nuf2p sequences returned from the tBLASTn search of the whole genome shotgun sequences showed either exceptionally high or low sequence divergence compared to D. melanogaster. We included in the tables all sequences that seemed to have the correct domain structure, in particular coiled coils of approximately the right length and position. We note that Ndc80 shows a very high divergence and Nuf2p a very low divergence across Drosophila species, but have not explored this further. Curiously, when other animal Ndc80/Nuf2p sequences are compared to Ndc80/Nuf2p from humans, it is the Ndc80 coil that is constrained relative to the Nuf2p [21]. The functional significance of these results is not clear at this time. Finally, tropomyosin coils are constrained in these insects but at different levels in Drosophila and mosquitoes.

In contrast to yeast, insect species including Drosophila and mosquitoes (Drosophila/Anopheles TLCA≈250 MYA [56]) show a strong conservation of SMC1/3 coils relative to SMC2/4. This suggests that there is a mechanism constraining sequence divergence of cohesin’s coiled coils, and this mechanism is found across all insect species. The nematode comparisons are less definitive but provide similar results. Their SMC1/3 coils are more conserved than are those of SMC2/4 and Ndc80/Nuf2p (C. elegans/C. briggsae TLCA≈24 MYA [6 generations/year in soil] [55]), but the difference is not as striking as in Drosophila and mammals, perhaps reflecting their lack of CTCF [35].

Table 3 extends our previous analysis of mammals and vertebrates to include more distant members of the animal kingdom. The conservation of SMC1/3 coils is especially striking for mammals (0-0.6% divergence), but the SMC1/3 coils are also much more highly conserved than the SMC2/4 coils when comparing H. sapiens to other vertebrates. It is only when H. sapiens SMC1/3 are compared to invertebrates such as sea urchin, sea squirt, sea anemone, Drosophila, and C. elegans that high sequence divergences are observed. Nevertheless, these invertebrate SMC1/3 coils are still more conserved than SMC2/4 or Ndc80/Nuf2p coils. This constraint upon SMC1 coils in animals does not include SMC1B, whose coils show divergences consistent with a spacer rod function. This suggests that the special mechanism involving the coiled coils of SMC1/3 cohesin is found across the entire animal kingdom but is not associated with meiotic cohesin function.

Analysis of SMC1/SMC3 Coiled Coil Mutations in Cornelia de Lange Syndrome

Most CdLS mutations are not in the SMC1/3, but in the accessory protein NIPBL (Nipped-B-Like) [57]. This protein is not a part of the cohesin complex but is involved in loading the cohesin complex onto chromosomes. As noted in the Introduction, however, recent studies have demonstrated that mutations in the coiled-coil domains of both SMC1 and SMC3 can cause CdLS as well [24, 26]. Deardorff et al. [24] reported that the COILS program predicts that the mutations found in the coiled-coil of SMC1 change the probability of the formation of the coiled-coil domains over localized segments of the protein. We repeated this analysis in detail, and in Table 4 summarize the COILS prediction regarding the potential impact of each of these SMC1/3 mutations. The majority (6/7) reduce the probability of coiled-coil formation over relatively small segments. Two mutations: R711W and D831E-Q832Del are predicted to have the greatest impact. The R711W mutation reduces the probability of coiled-coil formation earlier in this segment, and COILS also predicts a shifting of the end of this segment of the coil by 15 amino acids towards the N-terminal of the protein. The D831E-Q832Del mutation would shift the helical groove over the remaining 104 amino acids of this segment of the coil, but produced only a small, localized weakening of the coils. One mutation: R790Q is
predicted to increase the probability of coiled-coil formation of this segment by reducing the size of a small upstream interruption. In contrast to the SMC1 mutations, for the only SMC3 coil mutation reported for CdLS patients (E488Del), COILS predicts it will not impact the coil negatively. Overall, it seems that the deleterious effects of the mutations may not involve significant alteration of the coiled-coil structure.

### Discussion

As noted in the Introduction, cohesin in yeast has been implicated in transcriptional silencing and in defining the boundaries of the silenced HMR locus \([22,23]\), but no role for SMC1/3 in regulating gene expression has been discovered yet. In animals, however, the CdLS mutations and *Drosophila* data are

---

#### Table 2. Nematode and Insect Tropomyosin, Ndc80, Nuf2p, and SMC Protein and Coil Divergences.

|           | % AA Divergence vs. *D. melanogaster* |
|-----------|--------------------------------------|
| **Insects** |                                      |
| D. sechellia | Tm 4.6 6.8 (8.3 p) Ndc80 1.3 (9.7) Nuf2p 3.8 (1.8) SMC2 3.4 (2.7) SMC4 SMC1 SMC3 0.3 (4.5) 0.2 (0.3) |
| D. pseudoobscura | 6.3 63.4 (54.9) 8.7 (11.9) 27.3 (19.4) 32.2 (19.4) 6.4 (6.6) 4.3 (3.3) |
| D. grimshawi | 6.0 64.0 (60.7p) 7.6 (23.5p) 31.3 (23.3) 37.9 (32.6) 8.2 (8.7) 6.8 (5.1) |
| Anopheles gambiae | 21.1 73.2 (77.7) 31.9 (51.9) 58.4 (46.2) 62.0 (48.6) 34.8 (30.9) 28.1 (21.2) |
| Aedes aegypti | 20.1 74.9 (78.6) NA 58.0 (44.7) 61.2 (45.6) 33.1 (30.1) 27.2 (19.3) |
| Apis mellifera | 17.3 NA NA 63.6 (55.5) 73.6 (55.5) 60.0 (39.1) 57.1 (32.2) |
| A. gambiae vs. A. aegypti | 4.6 67.8 (60.1) NA 29.2 (22.8) 28.9 (24.0) 16.2 (15.5) 16.5 (10.6) |

**Nematodes**

| % AA Divergence vs. *C. elegans* |
|----------------------------------|
| C. briggsae Tm 7.0 44.4 (42.4) Ndc80 45.5 (46.9) Nuf2p 59.8 (49.0) SMC2 47.5 (38.1) SMC4 32.4 (24.1) SMC1 SMC3 37.3 (32.0) |
| B. malayi 11.8p 85.1 (75.6) 80.4 (80.1) 73.6 (65.1) 74.0 (69.3) 63.0 (50.7) 58.8 (52.5) |

Data presented as percentage sequence divergence for the combined coiled-coils (and in parentheses the divergence for the whole protein). Tropomyosin sequences are all coiled-coil.

NA = sequence not available.

p = sequence is partial (see Methods).

doi:10.1371/journal.pone.0004674.t002

---

#### Table 3. Animal Tropomyosin, Ndc80, Nuf2p, and SMC Protein and Coil Divergences*.

|           | % AA Divergence vs. *H. sapiens* |
|-----------|----------------------------------|
| *M. musculus* | u-Tm 0.4 19.0 (16.2) Ndc80 27.4 (27.9) Nuf2p 10.5 (7.9) SMC2 13.3 (10.5) SMC4 0.5 (0.6) SMC1 0.0 (0.1) SMC3 18.9 (16.4) |
| *R. norvegicus* | 0.8p 16.1 (14.8) 21.5 (20.8) 9.6 (6.7) 13.6 (10.5) 0.3 (0.4) 0.6 (0.6) 25.7 (16.2) |
| *B. taurus* | 0.7 8.0 (11.1) 12.2 (8.5) NA NA 0.0 (0.2) 0.0 (1.5) 14.2 (10.8p) |
| *G. gallus* | 4.2 27.8 (28.7) 40.2 (46.1) 28.9 (32.9) 28.9 (23.0) 4.3 (5.4) 0.6 (0.7) 44.4 (36.8) |
| *K. laevis* | 6.0 44.8 (43.0) 55.9 (55.3) 32.7 (18.6) 32.7 (23.0) 5.9 (5.9) 4.8 (3.3) NA |
| *D. rerio* | 7.7 57.3 (54.2) 70.3 (65.3) 29.9 (25.0) 40.5 (31.3) 11.2 (10.8) 4.3 (4.8) 61.1 (46.3) |
| *S. purpuratus* | 70.4p 766.7 (72.6) NA NA 60.3 (50.6) 45.4 (37.2) 42.5 (39.2p) ND |
| *C. intestinalis* | NA NA NA 61.8 (47.3) 65.6 (53.5) 494.4 (43.9) 56.9 (44.7) ND |
| *N. vectensis* | 60.4 73.4 (68.2) NA NA 59.6 (45.7) 55.1 (46.0) 58.5 (50.7) ND |
| *D. melanogaster* | 51.6 74.8 (82.7) 77.1 (82.2) 77.7 (58.9) 76.9 (60.2) 57.5 (50.3) 60.1 (48.1) ND |
| *C. elegans* | 43.3 71.3 (78.9) 81.2 (80.5) 73.2 (67.1) 78.4 (68.4) 69.6 (58.3) 69.6 (61.1) ND |

Data presented as percentage sequence divergence for the combined coiled-coils (and in parentheses the divergence for the whole protein). Tropomyosin sequences are all coiled-coil.

NA = sequence not available.

p = sequence is partial (see Methods).

ND = not determined.

*Portions of data taken from White and Erickson, *Journal of Structural Biology* 154 (2006): 111–121.

doi:10.1371/journal.pone.0004674.t003

---
compelling evidence for a post-mitotic function for SMC1/3 cohesin in the nervous system. In Drosophila, SMC1 and Nipped-B are co-localized at multiple locations on the chromosomes, primarily at the promoter regions of active genes [58]. Curiously, reduced levels of Nipped-B reduced expression of the act gene (a homeobox protein important in morphogenesis), while reduced SMC1 increased its expression [25,59,60].

Furthermore, the role of cohesin in gene regulation may not be limited to neurons and may not always require an intact cohesin complex. Ghiselli and Iozzo found that overexpressing SMC3 alone approximately 3-fold in NIH and Balb/c 3T3 cells caused them to adopt a transformed phenotype. They also found that SMC3 (SMC1 was not examined) was elevated in 70% of human colon carcinoma samples [61]. Human 293 cells stably transformed by SMC3 overexpression upregulated the expression of at least 65 genes [62]. Though the mechanism by which SMC3 overexpression causes these changes remains unknown, the data suggest further that SMC3 can modulate gene expression without being part of the cohesin complex. It would be interesting to test this for SMC1, and for the isolated coiled-coil segments of the two subunits.

Cohesin’s role in regulating gene expression in post-mitotic cells has only recently been added to its list of biological functions. In neurons, the mechanism apparently involves the whole cohesin complex: cleavage of the Rad21 subunit in post-mitotic neurons, the mechanism apparently involves the whole cohesin complex. The coiled coil is involved in sister chromatid cohesion, which was the only known function for cohesin. A recent study has provided evidence for lateral interactions of yeast cohesins during this process, presumably along the length of their coiled coils [63]. This mechanism is apparently not one that requires extreme conservation of the coiled coil sequence, since we found that the coiled coils are not highly conserved in Saccharomyces. However, several recent studies have demonstrated that cohesin also has a separate function involving regulation of gene expression in post-mitotic neurons and other cells. We now suggest that the coiled coils of cohesin in metazoans are an important part of these mechanisms. It is interesting that five out of seven CdLS mutations in SMC1 change positively charged arginines to neutral residues. These arginines may be involved in binding to the negatively charged phosphates of DNA.

Conclusions

Our previous analysis showed that the coiled coils of cohesin are very highly conserved across vertebrates, which implied a function involving the entire length and circumference of the coiled coil. At that time we suggested that the surfaces of the coils would be involved in sister chromatid cohesion, which was the only known function for cohesin. A recent study has provided evidence for lateral interactions of yeast cohesins during this process, presumably along the length of their coiled coils [63]. This mechanism is apparently not one that requires extreme conservation of the coiled coil sequence, since we found that the coils are not highly conserved in Saccharomyces. However, several recent studies have demonstrated that cohesin also has a separate function involving regulation of gene expression in post-mitotic neurons and other cells. We now suggest that the coiled coils of cohesin may play a key role in this second mechanism. The high conservation of the coiled coils of cohesin in metazoans supports the hypothesis that the entire surface of the coils may be involved in binding interactions, perhaps to the DNA of the various genes they regulate.

Author Contributions

Conceived and designed the experiments: GEW HPE. Analyzed the data: GEW. Wrote the paper: GEW HPE.

References

1. Hirano T (2005) SMC proteins and chromosome mechanics: from bacteria to humans. Philos Trans R Soc Lond B Biol Sci 360: 507–514.
2. Cobbe N, Heck MM (2000) Review: SMG in the world of chromosome biology from prokaryotes to higher eukaryotes. J Struct Biol 129: 125–143.
3. Cobbe N, Heck MM (2004) The evolution of SMC proteins: phylogenetic analysis and structural implications. Mol Biol Evol 21: 332–347.
4. Hagtrom KA, Meyer BJ (2003) Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet 4: 520–534.
5. Haering CH, Lowe J, Hochwagen A, Nasmyth K (2002) Molecular architecture of SMC proteins and the yeast cohesin complex. Mol Cell 9: 773–780.
6. Haering CH, Nasmyth K (2003) Building and breaking bridges between sister chromatids. Bioessays 25: 1178–1191.
7. Morikawa H, Morishita T, Kawanow S, Iwasaki H, Carr AM, et al. (2004) Rad62 protein functionally and physically associates with the smc3/smc6 protein complex and is required for chromosome integrity and recombination repair in fission yeast. Mol Cell Biol 24: 9401–9413.
8. Jones S, Sgroiros J (2001) The cohesin complex: sequence homologies, interaction networks and shared motifs. Genome Biol 2: RESEARCH0009.
9. Crick FHC (1953) The packing of alpha-helices: simple coiled-coils. Acta Crystallograpica 6: 689–697.
10. Mihinovich M, Utal E, Ward C, Skibbens R, Koshland D (2007) A multi-step pathway for the establishment of sister chromatid cohesion. PLoS Genet 3: e12.
11. Gruber S, Arumugam P, Katou Y, Kuquitsh D, Heuhaert W, et al. (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell 127: 525–537.

12. Vollrath H, Anke W, Jäger R, Beumer J, Uhlmann F (2007) Sequential loading of cohesin subunits during the meiotic prophase of grasshopper spermatocytes. PLoS Genet 3: e28.

13. Yu HG, Koshland D (2005) Chromosome morphogenesis: condensin-dependent cohesin removal during meiosis. Cell 123: 397–407.

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

15. Zhang J, Shi X, Li Y, Kim BJ, Jia J, et al. (2008) Acetylation of Smc3 by Ecor1 is required for S phase sister chromatin cohesion in both human and yeast. Mol Cell 31: 145–151.

16. Wong RW, Blobel G (2008) Cohesin subunit SMC1 associates with mitotic microtubules at the spindle pole. Proc Natl Acad Sci U S A 105: 15441–15445.

17. Revenkova E, Eijpe M, Huygen F, Gross B, Jessberger R (2008) Novel metazoan-specific isoform of mammalian SMC1. Mol Cell Biol 28: 6984–6996.

18. Revenkova E, Jessberger R (2005) Keeping sister chromatin together: cohesins in meiosis. Reproduction 130: 783–790.

19. Huang CE, Milutinovich M, Koshland D (2005) Rings, bracelet or snaps: fashionable alternatives for Smc complexes. Philos Trans R Soc Lond B Biol Sci 360: 537–542.

20. Guacci V (2007) Sister chromatid cohesion: the cohesin cleavage model does not ring true. Genes Cells 12: 699–708.

21. White GE, Erickson HP (2006) Sequence divergence of coiled-coils–structural rods, myosin filament packing, and the extraordinary conservation of cohesins. J Struct Biol 154: 111–121.

22. Lau A, Blintzblau H, Bell SP (2002) Cell-cycle control of the establishment of mating-type silencing in S. cerevisiae. Genes Dev 16: 2935–2945.

23. Douzé D, Adams CR, Rine J, Kamakaka RT (2002) The boundaries of the Fab-8 insulator. EMBO Rep 3: 165–170.

24. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, et al. (2005) Cohesin and Nipped-B with transcriptionally active regions of the Drosophila phagocyte. Nature 435: 203–210.

25. Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, et al. (2008) Association of cohesin with transcriptionally active regions of the Drosophila genome. Nature 451: 51–66.

26. Misulovin Z, Schultz YB, Li X, Kahn TG, Gause M, et al. (2008) Cohesin connects cohesin in Drosophila to the actin cytoskeleton. EMBO Rep 9: 66–71.

27. Warner JR (1999) The economics of ribosome biosynthesis in yeast. Trends Biochem Sci 24: 382–389.

28. Perry SV (2003) Vertebrate tropomyosins: distribution, properties and function. J Muscle Res Cell Motil 24: 5–9.

29. Lin JJ, Warren KS, Wamboldt DD, Wang T, Lin JL (1997) Tropomyosin isoforms in nonmuscle cells. Int Rev Cytol 170: 1–38.

30. Thomas JA, Loch HM, Bomburg JR, Gunning PW, Weisberger RP (2008) A tropomyosin I induced defect in cytokinesis can be rescued by elevated expression of cofilin. Cell Motil Cytoskeleton 65: 979–990.

31. Dorsett D, Fallon P, Chemes M, de Lange T (2002) The means to an end. Curr Opin Cell Biol 14: 227–231.

32. Davis TN, Wordeman L (2007) Rings, bracelets, sleeves, and chevrons: new functional features in Saccharomyces genomes by phylogenetic footprinting. Science 310: 71–76.

33. Davis TN, Wordeman L (2007) Rings, bracelets, sleeves, and chevrons: new structures of kinetochore proteins. Trends Cell Biol 17: 377–382.

34. Tanaka TU, Desai A (2008) Kinetochore-microtubule interactions: the means to the end. Curr Opin Cell Biol 20: 53–63.

35. Clift P, Sadasivan P, Desai A, Fulton L, Fulton B, et al. (2003) Finding functional features in Saccharomyces genomes by phylogenetic footprinting. Science 301: 71–76.

36. Rainieri S, Zambonelli G, Kaneko Y (2003) Saccharomyces sensu strictio: systematic, genetic diversity and evolution. J Biotechnol 96: 1–9.

37. Stein LD, Bao Z, Blazier D, Blumenenthal T, Brent MR, et al. (2003) The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. PLoS Biol 1: e45.

38. Dorsett D, Sato Y, Campbell RK, Chapman J, Degnan B, et al. (2002) The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. Science 298: 2157–2167.

39. Clark AG, Eren MB, Smith DR, Bergman CM, Oliver B, et al. (2007) Evolution of yeast and genomes on the Drosophila phylogeny. Nature 450: 203–210.

40. Gunning P, O’Neill G, Hardeman F (2008) Tropomyosin-basal regulation of the actin cytoskeleton in time and space. Physiol Rev 88: 1–35.

41. Perry SV (2003) Vertebrate tropomyosins: distribution, properties and function. J Muscle Res Cell Motil 24: 5–9.

42. Guacci V (2007) Sister chromatid cohesion: the cohesin cleavage model does not ring true. Genes Cells 12: 699–708.

43. White GE, Erickson HP (2006) Sequence divergence of coiled-coils–structural rods, myosin filament packing, and the extraordinary conservation of cohesins. J Struct Biol 154: 111–121.

44. Dorsett D, Fallon P, Chemes M, de Lange T (2002) The means to an end. Curr Opin Cell Biol 14: 227–231.

45. Davis TN, Wordeman L (2007) Rings, bracelets, sleeves, and chevrons: new functional features in Saccharomyces genomes by phylogenetic footprinting. Science 310: 71–76.

46. Davis TN, Wordeman L (2007) Rings, bracelets, sleeves, and chevrons: new structures of kinetochore proteins. Trends Cell Biol 17: 377–382.

47. Davis TN, Wordeman L (2007) Rings, bracelets, sleeves, and chevrons: new structures of kinetochore proteins. Trends Cell Biol 17: 377–382.