Structure and DNA binding activity of the mouse condensin hinge domain highlight common and diverse features of SMC proteins

Julia J. Griese, Gregor Witte and Karl-Peter Hopfner*

Department of Biochemistry, Gene Center, Center for Integrated Protein Sciences and Munich Center for Advanced Photonics, Ludwig-Maximilians University Munich, Feodor-Lynen-Str. 25, D-81377 Munich, Germany

Received December 22, 2009; Revised and Accepted January 14, 2010

ABSTRACT

Structural Maintenance of Chromosomes (SMC) proteins are vital for a wide range of processes including chromosome structure and dynamics, gene regulation and DNA repair. Eukaryotes have three SMC complexes, consisting of heterodimeric pairs of six different SMC proteins along with several specific regulatory subunits. In addition to their other functions, all three SMC complexes play distinct roles in DNA repair. Cohesin (SMC1–SMC3) is involved in DNA double-strand break repair, condensin (SMC2–SMC4) participates in single-strand break (SSB) repair, and the SMC5–SMC6 complex functions in various DNA repair pathways. SMC proteins consist of N- and C-terminal domains that fold back onto each other to create an ATPase ‘head’ domain, connected to a central ‘hinge’ domain via long coiled-coils. The hinge domain mediates dimerization of SMC proteins and binds DNA, but it is not clear to what purpose this activity serves. We studied the structure and function of the condensin hinge domain from mouse. While the SMC hinge domain structure is largely conserved from prokaryotes to eukaryotes, its function seems to have diversified throughout the course of evolution. The condensin hinge domain preferentially binds single-stranded DNA. We propose that this activity plays a role in the SSB repair function of the condensin complex.

INTRODUCTION

Throughout all kingdoms of life, Structural Maintenance of Chromosomes (SMC) proteins are responsible for the faithful inheritance of genetic information. They are involved in a wide range of vital cellular processes from cell division to gene regulation and DNA repair, acting as global organizers and safeguards of the genome. Whereas prokaryotic genomes encode for only one SMC protein that exists as a homodimer, eukaryotes possess six different SMC proteins that form three distinct heterodimeric complexes, with the holocomplexes additionally containing several specific regulatory non-SMC subunits (1).

In prokaryotes, the SMC complex is required for chromosome condensation and segregation (2). In eukaryotes, the complex containing SMC1 and SMC3, named cohesin, is responsible for sister chromatid cohesion during mitosis and meiosis (3). The condensin complex with SMC2 and SMC4 at its core is required, but not solely responsible for proper chromosome condensation and segregation during cell division (4). It seems to organize and maintain the chromosome scaffold rather than actually establishing it (5,6), but how it accomplishes this function is still unresolved. The as yet unnamed SMC5–SMC6 complex is involved in several DNA repair pathways as well as homologous recombination in meiosis (7).

Both cohesin and condensin are also involved in gene regulation (8–13) and DNA repair (14–20). While their roles are partially overlapping, they seem to be involved in different DNA repair pathways. Cohesin on the one hand is specifically recruited to DNA double-strand breaks (DSBs) in postreplicative cells and promotes DNA repair from the sister chromatid (17,21,22). Induction of a single DSB indeed leads to genome-wide establishment of cohesion independently of DNA replication (23,24), thus cohesin acts like a safeguard of genome integrity.

The function of condensin in DNA repair on the other hand is less well-characterized. Studies in yeast have implied that condensin has an interphase-specific function in DNA repair, but they have not clarified which DNA repair pathway is affected (14,16). There are, however, indications that condensin is involved in DNA single-strand break (SSB) repair (25,26). There are two forms of condensin in vertebrates that differ in their...
non-SMC subunits, but have the same core consisting of SMC2 and SMC4 (27). Human condensin I was shown to interact in an interphase-specific manner with the DNA nick-sensor poly(ADP-ribose) polymerase 1 (PARP1), and this interaction increased upon SSB damage induction. While nuclear retention of condensin was enhanced in the presence of SSBs, this was not the case for cohesin. Depletion of condensin I compromised SSB, but not DSB repair.

The interaction between condensin and PARP1 was particularly strong in S phase, suggesting an involvement of condensin also in normal DNA replication (25). Indeed, condensin was found to accumulate at stalled replication forks in budding yeast (28), and was shown to be required for the replication checkpoint response after stalling replication by hydroxyurea treatment in fission yeast (14). Since it is not known whether condensin is also found at moving replication forks, its presence at stalled forks might either be due to its DNA repair function, or a function in undisturbed DNA replication.

SMC proteins have a striking domain architecture consisting of a long antiparallel coiled-coil region with globular domains at both ends (1,29). Their N- and C-terminus interact at one end of the coiled-coil to make head domains and vice versa. Structural changes along the coiled-coil region to the hinge domains are quite similar. Two hinge domain monomers interact with each other via two interfaces to create a doughnut-shaped homodimer with 2-fold symmetry. The coiled-coils are formed intramolecularly and emerge from the same face of the dimer (29,36,37).

Most investigations focusing on the hinge domain have so far been conducted with bacterial SMC proteins. To shed more light on the structure and function of eukaryotic SMC hinges, we focused our efforts on the condensin hinge domain from mouse. We solved its crystal structure and furthermore studied its DNA binding activity. Interestingly, the condensin hinge domain preferentially binds single-stranded DNA (ssDNA), while its interaction with double-stranded DNA (dsDNA) is non-specific. Taking into account the data placing condensin in SSB repair (25), we propose that the ssDNA binding activity of the hinge domain supports the DNA repair function of the condensin complex.

**MATERIALS AND METHODS**

**Cloning, expression and protein purification**

We designed two mouse condensin hinge domain constructs of different lengths. The longer construct, designated mSMC2h4h-l, spans residues 492–680 of SMC2, and residues 581–766 of SMC4, while the shorter one, mSMC2h4h-s, contains residues 506–666 of SMC2, and residues 595–752 of SMC4. DNA fragments encoding the desired hinge domain constructs were PCR-amplified from cDNA vectors (SMC2: image ID 30543190; SMC4: image ID 6841276, imaGenes) and cloned into a modified bicistronic pET-21b vector (Novagen) carrying a second ribosome binding site between the SalI and NotI sites. The smc2 hinge fragment was inserted between the NdeI/EcoRI sites, and the smc4 hinge fragment was cloned into the NceI/XhoI sites.

Expression was carried out in *E. coli* Rosetta (DE3) (Novagen). Cultures were grown at 37°C in LB medium supplemented with the appropriate antibiotics to an optical density (600 nm) of ~0.7, cooled down to 18°C, and induced with 0.5 mM IPTG. Cells were harvested by centrifugation 20 h after induction. Cell pellets were stored at −20°C until further use. To obtain selenomethionine-labelled protein, constructs were expressed in *E. coli* B834 (DE3) additionally containing the pRARE plasmid (Novagen) in LeMaster’s medium (39) supplemented with the appropriate antibiotics and selenomethionine.

Proteins were purified via nickel chelate affinity chromatography and gel filtration. All purification steps were carried out at 8°C. Cells from 2 l of culture were lysed by sonication in buffer A (25 mM Tris–HCl pH 8.0, 300 mM NaCl, 20 mM imidazole). The lysate was cleared by centrifugation and applied to a gravity flow column containing Ni-NTA agarose beads (Qiagen). Nickel chelate chromatography was performed using buffer A for washing and buffer B (25 mM Tris–HCl pH 8.0, 300 mM NaCl, 250 mM imidazole) for elution. The eluate was concentrated in centrifugal filter units (Amicon Ultra, 10,000 MWCO, Millipore) and applied to a Superdex 200 pg 26/60 gel filtration column (GE Healthcare) equilibrated in buffer C (5 mM Tris–HCl pH 8.0, 100 mM NaCl, 0.1 mM EDTA). Fractions containing only the heterodimeric SMC2–SMC4 hinge domain were pooled, and the protein was concentrated to 30–40 mg/ml. Protein concentration was determined using a calculated extinction coefficient at 280 nm (40). The purification process was monitored by SDS-PAGE (41). Concentrated protein was aliquoted, flash frozen in...
liquid nitrogen and stored at −80°C until further use. Selenomethionine-labelled protein was purified analogously with the addition of 1 mM DTT to all buffers. TCEP was added to the concentrated protein to a final concentration of 1 mM to prevent oxidation of the selenomethionine residues.

**Small-angle X-ray scattering of protein solutions**

To prepare samples suitable for small-angle X-ray scattering (SAXS) measurements, proteins were additionally purified via gel filtration on a Superdex 200 column (GE Healthcare), and concentrated to yield samples in concentration ranges from 2 to 20 mg/ml in buffer C. The flowthrough of the concentration step was used as buffer reference for SAXS measurements. SAXS data were collected at beamline X33 at EMBL/DESY, Hamburg. Scattering profiles of BSA and lysozyme were measured as reference for molecular mass determination. The ATSAS package (42) was used to process and analyze data. Theoretical scattering profiles from atomic resolution models were calculated and fitted to measured profiles with CRYSOL. *Ab initio* models of mSMC2h4h-l were reconstructed from the experimental data using the program GASBORp without imposing any symmetry or other restrictions on possible models. Ten independently reconstructed envelopes were aligned and averaged with SUPCOMB and DAMAVER. Envelope representations were calculated using the Situs package (43), which was also used to dock atomic resolution models into the envelope.

**Protein crystallization and structure determination**

The short mouse condensin hinge construct mSMC2h4h-s was crystallized by vapor diffusion in the hanging-drop setup at 20°C. Crystallization was optimized with selenomethionine-labelled protein. The refined crystallization condition contained 15% (w/v) PEG 4000, 5% (v/v) isopropanol, 20% (v/v) glycerol and 100 mM Tris–HCl pH 8.5. For data collection, crystals were flash-frozen in liquid nitrogen without additional cryoprotection. Data were collected at beamline PXI of the Swiss Light Source (SLS, Villigen, Switzerland). The crystals belong to space group P2 1 2 1 and contain one molecule each of the SMC2 and SMC4 subunits in the asymmetric unit (see Supplementary Table S1 for crystallographic data, phasing and refinement statistics). The structure was determined by single-wavelength anomalous dispersion (SAD) phasing from a peak wavelength dataset of the selenomethionine-labelled protein crystals. Data were indexed and integrated using the XDS package (44). Phases were calculated with AutoSHARP (45). The model was largely automatically built with ARP/wARP (46–48) and completed by manual model building in Coot (49). Initial refinement was carried out with CNS (50), followed by several rounds of refinement with phenix.refine (51) and rebuilding in Coot. Refinement included simulated annealing in initial cycles, individual atomic coordinate and anisotropic B factor refinement, and bulk solvent corrections. Solvent molecules were added with phenix.refine and manually. The $R_{free}$ factor was calculated from 10% of the data which were removed at random before the structure was refined. The structure was validated using MolProbity (52) and PROCHECK (53). The electrostatic surface potential was calculated with the Adaptive Poisson–Boltzmann Solver (APBS) (54). All figures were prepared with PyMOL (DeLano, W.L. The PyMOL Molecular Graphics System. (2008) DeLano Scientific LLC, Palo Alto, CA, USA. http://www.pymol.org). Coordinates and structure factors were deposited at the Protein Data Bank (PDB) with accession number 3L51.

**In vitro DNA-binding assays**

We used DNA oligonucleotide substrates carrying a 6-FAM fluorescence label to monitor their binding to the mouse condensin hinge domain (Table 1). HPLC-purified DNA oligonucleotides (Thermo Scientific) were dissolved in water. DNA concentration was determined using a calculated extinction coefficient at 260 nm (55). To anneal oligonucleotides, they were mixed with a 1.1-fold molar excess of the unlabelled oligonucleotide in 40 mM Tris–HCl pH 7.5, 100 mM NaCl, 10 mM MgCl 2, incubated in a thermocycler (Biometra T personal) for 5 min at 95°C, and then cooled down to 4°C at a cooling rate of 0.1°C/s.

Samples for electrophoretic mobility shift assays (EMSAs) contained 12.5 nM of a DNA substrate and protein in a 0-, 1-, 2-, 10-, 25-, 50-, 100-, 250-, 500- and 1000-fold molar excess over the DNA in 1 x PBS in a total volume of 20 µl. They were incubated at room temperature for 30 min before addition of 5 µl 50% (v/v) glycerol. The samples were then loaded onto an 0.5% (w/v) agarose gel in 1 x TB buffer and separated for 2 h at 4 V/cm and 8°C. Gels were scanned on a Typhoon 9400 fluorescence scanner (GE Healthcare).

Fluorescence quenching titrations were performed in a Horiba Jobin Yvon FluoroMax-P fluorimeter, using a 1.5 ml fluorescence cuvette with a stirring bar, at 20°C under constant stirring. The titration solution contained 25 nM 6-FAM-labelled DNA substrate (Table 1) in 1 x PBS in a starting volume of 800 µl. Protein was added successively from a concentrated stock solution. After each addition of protein, the mixture was allowed to reach equilibrium for 1 min before measuring fluorescence. 6-FAM fluorescence was excited at 495 nm, and measured at the experimentally determined emission maximum of the DNA substrate (between 415 and 418 nm). Excitation and emission slit width were adjusted so that the signal was in the linear range of the photon counting multiplier. All measurements were performed in triplicate. Data were normalized and fitted using a non-linear least squares fit algorithm with a single-site binding model. Dissociation constants are the result of global fits to triplicates, errors are the standard deviations of dissociation constants resulting from independent fits to the three measurements.
Table 1. Oligonucleotides used for DNA-binding assays

<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
<th>Oligonucleotide sequences</th>
</tr>
</thead>
</table>
| 30-mer ssDNA       |                            | 5′-6-FAM-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Figure 1. Overall structure of the mouse condensin hinge domain (construct mSMC2h4h-s, residues 506–666 of SMC2 and 595–752 of SMC4). (A) Stereo view of the SMC2–SMC4 hinge domain, as seen from the bottom face, i.e. the face on the opposite side of the coiled-coils, colored by subdomains. The SMC2 subunit is colored in shades of blue, the SMC4 subunit in shades of red, with the N-terminal subdomain shown in the lighter shade. The long loop connecting the subdomains is shown in dark and light grey for SMC2 and SMC4, respectively. The hexahistidine tag on the C-terminals of the SMC4 hinge is colored white. (B) The SMC2–SMC4 hinge domain heterodimer as seen from the top face, in the open and closed conformation. The top panel shows the open conformation found in the crystal structure, the bottom panel depicts the model of the closed conformation, generated by separately superposing the SMC2 and SMC4 subunits with the T. maritima SMC hinge domain dimer (pdb 1GXL) (29). The SMC2 subunit is shown in blue, the SMC4 subunit in red and orange for the open and closed conformation, respectively. (C) Superposition of the SMC2 and SMC4 subunit, shown in stereo. The color scheme is the same as in (A).

Figure 2. Alignment of the hinge domains of SMC1α, SMC2, SMC3 and SMC4 from mouse. Only residues present in the crystallized construct mSMC2h4h-s are shown for SMC2 and SMC4 (residues 506–666 of SMC2 and 595–752 of SMC4). Numbering of residues is for mSMC2. The secondary structure of the SMC2 hinge domain is shown above the alignment, that of the SMC4 hinge domain below. α-Helices are displayed as red rectangles, β-strands as blue arrows. The residues that probably form β8 of SMC2 are invisible in the electron density. The asterisks mark mSMC2-K506/mSMC4-K510 and mSMC2-K613/mSMC4-K698. In the consensus sequence, lower case letters are used for ≥50%, upper case letters for ≥90% conservation; exclamation mark denotes any one of IV, dollar symbol: any one of LM, hash: any one of NDQE. The alignment was generated with MultAlin (62).
missing to close the ring. The structural similarity to the TmaSMC hinge domain however strongly suggests that the subunit cores are correctly folded. The rmsd between the Cα traces of the SMC2 and SMC4 hinge domain and the TmaSMC hinge (pdb 1GXL) is 2.2 and 2.6 Å, respectively, and the rmsd between the SMC2 and the SMC4 hinge is only 1.8 Å. Most secondary-structure elements are conserved.

Notable differences in secondary structure between the TmaSMC and the mouse condensin hinge domain are found in the SMC4 subunit (Figures 1C and 2, Supplementary Figure S2). In the mSMC2 and TmaSMC hinge domain, the N-terminal β-sheet merges directly into the long loop connecting the subdomains which terminates in a helix on the outside of the C-terminal β-sheet (mSMC2h zF). This helix is followed by a strand of the C-terminal β-sheet (mSMC2h β4). The mSMC2 and mSMC4 hinges have the same number of helices, but their F helices do not correspond to each other and are in completely different positions, flanking the C-terminal β-sheet in mSMC2h, and the helical core on the outside surface in mSMC4h. In the mSMC4 hinge, helix F lies between the N-terminal β-sheet and the connecting loop which terminates in helix G of the helical core, so that the C-terminal β-sheet of mSMC4h consists of one less strand than that of the mSMC2 and TmaSMC hinge (provided that the predicted C-terminal β-strand of mSMC2h is indeed formed).

The SMC2–SMC4 hinge interface explains dimerization specificity of eukaryotic SMC proteins

Most intersubunit contacts are formed by hydrophobic interactions, supported by few hydrogen bonds (Figure 3). It has long been unclear how eukaryotic SMC proteins specifically assemble into defined heterodimers (e.g. SMC1–SMC3, SMC2–SMC4), while prokaryotic SMC proteins form homodimers. Taking a closer look at the interface, the reason for dimerization specificity is revealed. While most residues contributing to the interaction are conserved or replaced by similar residues within the subfamilies SMC2/SMC3 and SMC1β/SMC4 (56), the few non-conservative exchanges are apparently enough to make wrong pairing impossible. For SMC2/SMC3, good candidate residues for dimerization specificity within the N-terminal interface are K570R571R572 of SMC2 which are replaced by P582G583E584 in SMC3 (Figure 2). These residues are placed in an otherwise conserved region, hence it is very likely that they are in similar positions in the SMC2 and SMC3 tertiary structures. Looking at the dimerization interface, it is immediately obvious that the PGE sequence in SMC3 would not be able to form the same interactions with SMC4 as the KRR sequence in SMC2 (Figure 3). For example, SMC2-R572 forms a hydrogen bond with the backbone oxygen of SMC4-G740. The shorter glutamate side chain in SMC3 could not support this interaction. Between the SMC4 and SMC1β C-terminal interface residues there is only one non-conservative exchange, namely SMC4-T723 for SMC1β-R626 (Figure 2). Again, this residue is situated in an otherwise conserved region. Assuming that it therefore adopts a similar position in SMC1β as in SMC4, the arginine side chain would clash with K561 of SMC2, thus making it impossible for SMC1β to dimerize with SMC2. SMC3 on the other hand has a small threonine residue in place of SMC2-K561, so that the SMC3 interface can accommodate the large side chain of SMC1β-R626.

Solution scattering analysis shows that the condensin hinge domain can also adopt a closed conformation

To acquire structural information about the conformation of the condensin hinge in solution, we measured SAXS profiles of both the short and the long hinge domain construct (Figure 4A). The best data were obtained with protein concentrations of 8 and 20 mg/ml for the short and long construct, respectively. While the short construct showed a slight tendency to aggregate at high concentrations, the long construct did not show any such signs even at 20 mg/ml, thus yielding data of very high quality with little noise even at higher values of the scattering vector. The molecular mass determined from the scattering intensity extrapolated to zero angle confirms that both constructs exist as heterodimers in solution. We compared the observed scattering profiles with SAXS profiles calculated from the crystal structure and the model of the closed conformation (Figure 4A). The observed profile of mSMC2h4h-s matches perfectly with the profile calculated from the crystal structure. This means that the crystallized construct mSMC2h4h-s adopts the same open conformation in solution as in the crystal. Hence, this conformation is not produced by crystal contacts.

The long construct mSMC2h4h-l, on the other hand, clearly exhibits a different conformational state. The SAXS profiles suggest that this construct has the
The condensin hinge domain preferentially binds ssDNA

Previous work has demonstrated that the SMC hinge domain can bind DNA (30–33), but because of the purely qualitative nature of the assays performed in these studies, the specificity for different DNA substrates could not be unambiguously determined. To learn more about the potential role of the condensin complex in SSB repair, we quantitatively analyzed different DNA substrates for binding to the mouse condensin hinge domain. Sequences and structures of the DNA substrates used can be found in Table 1. The short construct mSMC2h4h-s binds DNA only very weakly and non-specifically (Supplementary Figure S4 and data not shown), probably due to its partially open conformation or the lack of the transition into the coiled-coil region which might harbor additional DNA binding sites. The long construct mSMC2h4h-l, however, which adopts a closed conformation and contains a short stretch of the coiled-coil regions, binds DNA quite efficiently and shows interesting differential affinity towards different DNA substrates.

We performed EMSAs with 6-FAM-labelled DNA substrates (Supplementary Figure S4). The results suggest a qualitatively different binding behavior towards ssDNA and dsDNA. ssDNA was shifted upwards of the well, that is, the protein–ssDNA complex migrated towards the cathode, suggesting that the protein is so positively charged that even the complex with a 30-mer ssDNA oligonucleotide still has a surplus of positive charge. In fact, the long mouse condensin hinge construct has a theoretical isoelectric point of ~9.3, and positively charged residues are distributed all over the surface of the protein (Figure 6). While a 30-mer dsDNA substrate was also bound, the EMSAs implied that it was bound much more weakly than the ssDNA substrate of the same length (Supplementary Figure S4). Also, the protein–dsDNA complex did not migrate towards the cathode, but remained in the wells. Thus, EMSAs are not the method of choice to quantitatively analyze DNA binding.

We therefore performed fluorescence quenching titrations to get a better picture of the DNA binding activity of the mouse condensin hinge, exploiting the fact that binding of 6-FAM-labelled DNA to the mouse condensin hinge leads to quenching of 6-FAM fluorescence. The resulting titration curves are not only a means to quantify the affinity towards different DNA substrates with great accuracy, but also contain information about the binding mode. All measurements were performed at physiological salt concentrations. Inclusion of Mg²⁺ in

Figure 4. Small-angle X-ray scattering analysis of the mouse condensin hinge domain. (A) SAXS profiles of the short and long construct in solution in comparison with profiles calculated from atomic resolution models. The scattering profile of the long construct mSMC2h4h-l is shown in black, the profile of the short construct mSMC2h4h-s in blue. The calculated scattering profiles of the atomic resolution models of mSMC2h4h-s in the open and closed conformation are shown in red and orange, respectively. (B) Solution envelope reconstruction of the long construct mSMC2h4h-l, superimposed with the atomic resolution models of mSMC2h4h-s in the open (top left) and closed conformation (bottom right). The SAXS envelope is shown as grey mesh, the atomic resolution models are depicted in cartoon representation with the SMC2 subunit colored blue, and the SMC4 subunit colored red in the open and orange in the closed conformation.
the assay buffer did not have any influence on DNA binding by the condensin hinge (data not shown). As can be seen in Figure 5A, the binding to dsDNA does not reach saturation even at a 1000-fold excess of protein and cannot be fitted using a simple binding model. This non-saturation behavior shows that the mouse condensin hinge binds dsDNA non-specifically. We therefore conclude that non-specific binding produces such large protein–DNA aggregates that they cannot enter the gel matrix and remain in the wells in EMSAs.

On the other hand, the mouse condensin hinge domain binds all DNA substrates tested that are at least partially single-stranded specifically and with high affinity (Figure 5A and Table 2). We used oligo(dT) as ssDNA substrate because it does not form intramolecular base pairing or stacking interactions and is a model substrate to study ssDNA binding specificity (57). All titration curves obtained with partially or completely single-stranded substrates could be fitted using a single-site binding model, meaning that one hinge heterodimer binds one DNA molecule. The 30-mer ssDNA substrate is bound with a dissociation constant of 0.45 ± 0.04 μM. A 15-mer ssDNA substrate is still bound specifically, albeit with a 10-fold higher dissociation constant ($K_d = 3.21 ± 0.14$ μM), suggesting that 15 nucleotides constitute the minimal binding length. We presumed that the condensin hinge might fall off the ends of this short oligonucleotide, while it would not so quickly dissociate from the twice as long 30-mer ssDNA substrate. To test this hypothesis, we designed 15-mer ssDNA substrates that are ‘capped’ on one or both ends by a 15-mer dsDNA stretch. Indeed, the 30-mer ds–ssDNA substrate, where one end is capped, is bound twice as strongly as the 15-mer ssDNA. Dissociation constants were found to be the same within the range of error, regardless of whether the ssDNA stretch was a 3’ or 5’ overhang ($K_d = 1.55 ± 0.05$ μM for 30-mer ds–ssDNA-3’, $1.77 ± 0.08$ μM for 30-mer ds–ssDNA-5’). This rules out specific recognition of a particular ssDNA–dsDNA transition by the condensin hinge. Since the fluorescence label is on the 5’-end of the short 15-mer strand (Table 1), it is close to the ssDNA stretch in the 30-mer ds–ssDNA-3’ substrate and far away from it in the 30-mer ds–ssDNA-5’ substrate. Titrations with these two substrates also showed that only the ssDNA stretch is bound, as the

---

**Figure 5.** Fluorescence quenching titrations. One representative titration curve of each triplicate measurement is shown. Solid lines represent fits according to a single-site binding model. (A) Titrations of different DNA substrates with the long mouse condensin hinge domain construct mSMC2h4h-l wt. DNA substrates are: filled circles, 30-mer ssDNA; open circles, 45-mer ds–ss–dsDNA; filled triangles, 30-mer ds–ssDNA-3’; open triangles, 15-mer ssDNA; filled diamonds, 30-mer dsDNA. The titration with the 30-mer dsDNA substrate could not be fitted with a simple binding model. For clarity, the titration with the 30-mer ds–ssDNA-5’ substrate is not shown. For structures and sequences of the DNA substrates see Table 1. (B) Titrations of the 30-mer ssDNA substrate with mSMC2h4h-l and three point mutants. Proteins are: filled circles, mSMC2h4h-l wt; filled triangles, mSMC2h4h-l K566E4h-K657E-l; open triangles, mSMC2h4h-l K613E4h-K698E-l. All titrations were performed with 25 nM of the 6-FAM labelled DNA substrate in 1× PBS at 20°C in a Horiba Jobin Yvon FluoroMax-P fluorimeter.

**Figure 6.** (A) Electrostatic surface potential of the mouse condensin hinge domain. Positively charged regions are colored blue, negatively charged regions red, and neutral regions white. The panels, from top left to bottom, show the top face, the view from the side onto the closed interface, and the bottom face. (B) Lysine residues in the mouse condensin hinge domain. The structure is depicted looking onto the closed interface. The backbone of the condensin hinge is shown in cartoon representation in white, lysines are shown as stick models in green. Lysine residues mutated to glutamate in the point mutants SMC2-K566E/SMC4-K657E and SMC2-K613E/SMC4-K698E of mSMC2h4h-l are colored blue and purple, respectively.
absolute change in fluorescence intensity during the titration was half as big when the label was far away from the ssDNA stretch as when it was close to it. Capping both ends of the 15-mer ssDNA results in a 5-fold tighter binding with respect to the uncapped 15-mer ssDNA ($K_d = 0.66 \pm 0.02 \mu M$), hence the condensin hinge binds the 45-mer ds–ss–dsDNA substrate almost as tightly as the 30-mer ssDNA substrate. This suggests that the hinge domain is held in place on the ssDNA stretch by the dsDNA caps, making it less likely to dissociate. Although at least two protein molecules should theoretically fit onto the 45-mer ds–ss–dsDNA substrate, data could be fitted very well with a single-site binding model, giving further proof that in partially double-stranded and partially single-stranded substrates only the ssDNA stretch is bound by the condensin hinge.

Having shown that the mouse condensin hinge preferentially binds ssDNA, we wanted to characterize this binding more closely. We therefore made three lysine-to-glutamate point mutants of the long construct mSMC2hK566E4h-l, and two double mutants, mSMC2hK566E4hK657E-l, and mSMC2hK613E4hK698E-l. The first combination of lysine residues (SMC2-K$_{566}$/SMC4-K$_{657}$) was chosen because this is the only lysine residue that is highly conserved among SMC proteins from all species (Figure 2). The second combination was chosen due to the position of these lysine residues in the structure. While SMC2-K$_{566}$/SMC4-K$_{657}$ are at the C-terminus of $\alpha$-helix E, the helix capping off the dimer interface, SMC2-K$_{613}$/SMC4-K$_{698}$ are $\sim 90^\circ$ removed from the dimer interface, in the short loop connecting helices G and H, on the outside surface of the hinge domain (Figures 2 and 6B). Additionally, they are positioned in a region of the protein where the electrostatic surface potential is almost neutral, whereas there is a cluster of positively charged residues around the dimer interface (Figure 6A). All point mutant constructs behaved like the wild-type construct mSMC2h4h-l during purification. To confirm that the mutations do not disturb the protein fold, we measured SAXS profiles of all three mutant proteins which matched the profile of wild-type mSMC2h4h-l, although the mutant proteins, especially mSMC2hK613E4hK698E-l, displayed a higher tendency to aggregate than the wild-type protein (Supplementary Figure S5).

EMSAs showed a reduction of the non-specific dsDNA binding for the single mutant and even more dramatically for both double mutants (Supplementary Figure S4), but results with the 30-mer ssDNA substrate were less clear. This is likely to be due to the fact that EMSAs are unsuitable to reveal subtle differences in binding strength (10-fold and less), as our data show. Fluorescence quenching titrations of the 30-mer ssDNA substrate with the mutant proteins clearly demonstrate a reduction in affinity as compared to wild-type (Figure 5B and Table 2). The effect of the mutations is additive, since the dissociation constant for the single mutant mSMC2hK$_{566}$E4h-l is half as big as that of the corresponding double mutant. Both double mutants have roughly the same affinity towards the 30-mer ssDNA, it is reduced 7- to 8-fold as compared to wild-type. The mutations also reduce specificity of binding, as the titrations show a contribution of non-specific interaction. Especially for the double mutants, the binding does not saturate completely, and data could only be fitted up to a 500-fold excess of protein over DNA.

### DISCUSSION

To learn more about the mechanism of eukaryotic SMC proteins, we determined the structure of a eukaryotic, heterodimeric SMC hinge domain from the mouse condensin complex, and quantitatively analyzed its DNA binding activity and specificity.

Our structure shows that the SMC hinge domain fold is conserved from prokaryotes to eukaryotes. Although the SMC2–SMC4 hinge heterodimer adopts an open conformation in our crystals, dimerizing via only one of the two expected interfaces, the subunit cores themselves resemble other SMC hinge domains already described (29,36,37). Interestingly, the SMC2 hinge is more similar to the prokaryotic SMC hinge (29) than the SMC4 hinge. The structural features of the SMC4 hinge that are different from its bacterial counterpart might be involved in specific functions of the condensin complex which the prokaryotic SMC complex does not have. Since these structural features are exposed on the outer surface of the SMC4 hinge, they could constitute a binding interface for an interaction partner like Cti1/C1D, a protein that was found to interact with the SMC4 hinge domain in fission yeast (16) and is implicated in DNA repair functions (58,59). Further research is clearly necessary to find out whether condensin also interacts with C1D in higher eukaryotes, and if so, whether it does this via the SMC4 hinge domain.

The open conformation that we observe in our crystals might be caused by the SMC2–SMC4 hinge domain construct being too short for the second dimerization interface to be stable, but it is also possible that it represents a functional intermediate during assembly of SMC complexes or their action on DNA. The SMC2–SMC4
hinge heterodimer is stable although one dimer interface is disrupted, indicating that one intact interface suffices for dimerization. This observation is in close agreement with mutational studies of the bovine cohesin hinge domain (30). The intact interface between the SMC2 and SMC4 subunit in our structure also offers an explanation as to how dimerization specificity of SMC proteins is created. Comparing the interacting residues with the corresponding residues in SMC1α and SMC3, we found that some of these residues would either clash with, or not be able to interact with the interface residues of the ‘wrong’ partners.

The interaction between the two subunits is strong enough to withstand some structural rearrangements, as the interface remains intact despite the hinge being bent open along the interface axis, leading to a wider angle between subunits than in the closed conformation. Our data therefore provide additional evidence for the structural flexibility of the hinge domain, a quality that is probably very important for the dynamic interactions of SMC proteins with DNA (32,60,61). The open conformation we observe would at first glance seem to suggest that the hinge domain could indeed open up to allow DNA to enter into the intra-coiled-coil space, as has been proposed (60). While we do not want to rule out this possibility, the space between the SMC2 and SMC4 subunits in our crystal structure is not big enough to accomodate a DNA double helix, and the charge distribution on the inner surface of the hinge domain would rather repulse than attract DNA (Figure 6A). In fact, the observed charge distribution with only one strongly basic patch argues for a preference for ssDNA over dsDNA.

It has been demonstrated that the transition into the coiled-coil region is necessary for DNA binding by the cohesin hinge, but not for its dimerization (30). Similar results were obtained in a previous study of the BsSMC protein (31), and we found this to be true for the mouse condensin hinge domain as well. In our experiments, the construct without coiled-coil regions bound to DNA only very weakly and non-specifically, whereas the construct carrying a short stretch of coiled-coil bound ssDNA strongly and specifically. Our structural data show that in the construct without coiled-coil regions only one dimer interface is intact, while the longer construct has both expected dimer interfaces. Taken together, these results suggest that the transition into the coiled-coil region does not directly participate in DNA binding, but rather confers structural stability to the hinge domain, especially to the basic patch at the dimer interface which is essential for DNA binding.

The effects of our lysine-to-glutamate mutants of the condensin hinge domain in DNA-binding assays imply that ssDNA wraps around the outer surface of the hinge domain, and all positively charged residues contribute to binding. In spite of prokaryotic SMC proteins also showing a preference for ssDNA over dsDNA, as has been observed previously with the B. subtilis SMC (BsSMC) hinge domain (31–33,35), their DNA binding surface is likely to be different. While the outer surface of the condensin hinge domain is neutral-to-positively charged (Figure 6A), the TmaSMC hinge has a neutral-to-negatively charged outer surface, and only its inner surface is positively charged with the basic patch at the interface also being less pronounced (29). The DNA binding activity of the TmaSMC hinge has never been studied, but due to the strong sequence conservation it is probably quite similar to that of the BsSMC hinge domain, or conversely, the surface charge distributions of both hinge domains are probably quite similar. In studies of the BsSMC protein (32), the authors found a complete disruption of dsDNA binding when the conserved lysine residue K 565 (corresponding to mSMC2-K 666/mSMC4-K 667) was mutated to glutamate, whereas ssDNA binding seemed to be only modestly affected. Only mutation of three consecutive lysine residues (K 666–K 668) to glutamate, that are part of the same basic patch at the dimer interface as K 565, resulted in a complete loss of dsDNA as well as ssDNA binding (32). Our own studies suggest that prokaryotic SMC hinge domains bind DNA with lower affinity than the condensin hinge domain, whilst always having a preference for ssDNA (unpublished data). In prokaryotic SMC proteins, the DNA is therefore probably bound only by the basic patch at the hinge domain dimer interface. Interestingly, studies of the BsSMC protein also suggest that the prokaryotic SMC hinge domain interacts with dsDNA and ssDNA in mechanistically distinct manners (32,33), as our results show the eukaryotic condensin hinge domain does.

What could be the function of the ssDNA binding activity of the condensin hinge domain? We propose that this activity might support the SSB repair function of the condensin complex. In higher eukaryotes, PARP1 recognizes single-strand breaks and facilitates base-excision repair (BER) (26). There are two BER pathways, short- and long-patch BER (26). Short-patch BER repairs single-nucleotide gaps, employing DNA polymerase β and DNA ligase IIIα. Long-patch BER on the other hand requires the action of FEN1, DNA polymerase δ/ε (or β) and DNA ligase I to repair gaps of up to 12 nucleotides length, and is additionally stimulated by PARP1 and PCNA (26). In addition to its interaction with PARP1, upon SSB damage induction condensin was found to interact with the BER scaffold protein XRCC1 as well as FEN1 and DNA polymerase δ/ε (25), suggesting it is involved in long-patch BER. The ssDNA binding activity of the condensin hinge domain might tether it to the damage site and might help to organize the DNA structure for repair. This function might also be important during normal DNA replication which transiently produces ssDNA stretches, as suggested by the replication checkpoint defect of condensin mutants (14), and the accumulation of condensin at stalled replication forks (28). Lower eukaryotes like yeast do not possess PARP1, therefore condensin must have a slightly different function in DNA repair in these organisms (14,16). However, it is likely that condensin participates in SSB repair in lower eukaryotes as well, and possibly even in prokaryotes, since prokaryotic SMC hinge domains also preferentially bind ssDNA (31–33). It is conceivable that ssDNA binding has been enhanced during the evolution from prokaryotic SMC proteins to condensin while genome size increased and DNA repair...
pathways consequently became more and more sophisticated.

**REFERENCES**


