<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Proteins: Structure, Function, and Bioinformatics*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Griese, J J., Hopfner, K P. (2011)

Structure and DNA-binding activity of the *Pyrococcus furiosus* SMC protein hinge domain

Proteins: Structure, Function, and Bioinformatics, 79(2): 558-568

<https://doi.org/10.1002/prot.22903>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-381650>

Full title:

Structure and DNA-Binding Activity of the *Pyrococcus furiosus* SMC Protein Hinge Domain

Short title: The *Pyrococcus furiosus* SMC hinge domain

Julia J. Griesse¹ and Karl-Peter Hopfner^{1*}

¹Gene Center, Center for Integrated Protein Sciences and Munich Center for Advanced Photonics, Department of Biochemistry, Ludwig-Maximilians University Munich, Feodor-Lynen-Str. 25, D-81377 Munich, Germany

*Corresponding author:

Karl-Peter Hopfner

phone: 0049-(0)89-218076953

fax: 0049-(0)89-218076999

email: hopfner@lmb.uni-muenchen.de

Key words: Structural Maintenance of Chromosomes (SMC), hinge domain, coiled-coil, single-stranded DNA, chromosome condensation, chromosome segregation

ABSTRACT

Structural Maintenance of Chromosomes (SMC) proteins are essential for a wide range of processes including chromosome structure and dynamics, gene regulation, and DNA repair. While bacteria and archaea have one SMC protein that forms a homodimer, eukaryotes possess three distinct SMC complexes, consisting of heterodimeric pairs of six different SMC proteins. SMC holocomplexes additionally contain several specific regulatory subunits. The bacterial SMC complex is required for chromosome condensation and segregation. In eukaryotes, this function is carried out by the condensin (SMC2-SMC4) complex. SMC proteins consist of N-terminal and C-terminal domains that fold back onto each other to create an ATPase “head” domain, connected to a central “hinge” domain via a long coiled-coil region. The hinge domain mediates dimerization of SMC proteins and binds DNA. This activity implicates a direct involvement of the hinge domain in the action of SMC proteins on DNA. We studied the SMC hinge domain from the thermophilic archaeon *Pyrococcus furiosus*. Its crystal structure shows that the SMC hinge domain fold is largely conserved between archaea and bacteria as well as eukarya. Like the eukaryotic condensin hinge domain, the *P. furiosus* SMC hinge domain preferentially binds single-stranded DNA (ssDNA), but its affinity for DNA is weaker than that of its eukaryotic counterpart, and point mutations reveal that its DNA-binding surface is more confined. The ssDNA-binding activity of its hinge domain might play a role in the DNA-loading process of the prokaryotic SMC complex during replication.

INTRODUCTION

Structural Maintenance of Chromosomes (SMC) proteins act as global organizers and safeguards of the genome in all domains of life. By directly and indirectly influencing chromosome structure and dynamics, they participate in a vast number of vital cellular processes ranging from cell division to gene regulation and DNA repair. Prokaryotic genomes contain only one *smc* gene whose product forms a homodimer, whereas in eukaryotes six different SMC proteins are found that form three distinct heterodimeric complexes. The SMC holocomplexes additionally contain several specific regulatory subunits¹.

The bacterial SMC complex is required for chromosome condensation and segregation². In eukaryotes, this function is carried out by the condensin complex with SMC2 and SMC4 at its core, the closest homologue of the prokaryotic SMC complex³. The complex containing SMC1 and SMC3, named cohesin, is responsible for sister chromatid cohesion during mitosis and meiosis⁴. The most divergent of the eukaryotic SMC complexes, consisting of SMC5 and SMC6, remains unnamed, as its precise function is still enigmatic. It is involved in several DNA repair pathways as well as homologous recombination in meiosis⁵.

SMC proteins have a striking domain architecture consisting of a long antiparallel coiled-coil region with globular domains at both ends^{1,6}. Their N and C terminus interact at one end of the coiled-coil to make up an ATP-binding cassette (ABC)-type ATPase “head” domain, with the Walker A motif being contributed by the N-terminal half, while the Walker B and the ABC signature motifs are located in the C-terminal half. The “hinge” domain at the other end of the coiled-coil mediates dimerization of SMC proteins. Upon ATP binding, the head domains engage to form two shared ATPase active sites with the ATP sandwiched between the Walker A and B motifs from one monomer and the signature motif from the other^{7,8}. Most of the non-SMC subunits bind to the head domains and regulate their ATPase activity

and consequently their engagement-disengagement cycle^{9,10}. It seems obvious therefore to assume that the head domains constitute the active part of the SMC complex, while the coiled-coil and hinge domain only play a structural role. However, the function of the hinge domain exceeds that of a simple dimerization domain, as it has been shown to also bind DNA^{9,11-14} as well as proteins^{15,16}. In case of the *Bacillus subtilis* SMC protein, ATP binding to the head domains stimulates DNA binding to the hinge domains⁹, and this in turn stimulates ATP hydrolysis by the heads^{12,13,17,18}. This indicates that the hinge domain is capable of transmitting conformational changes along the coiled-coil region to the head domains and vice versa, and hence that the hinge domain is closely involved in the actions of SMC proteins on DNA.

To date, the structures of a eukaryotic and two bacterial SMC hinge domains have been determined, these being the hinge domains of the *Thermotoga maritima* SMC (TmaSMC)⁶, *Escherichia coli* MukB^{19,20}, and mouse condensin (SMC2-SMC4)¹⁴. MukB is a divergent SMC protein which has closer homologues only in other γ -proteobacteria, and its hinge domain is substantially smaller than that of other bacterial and eukaryotic SMC proteins. Nonetheless the MukB hinge domain is structurally quite similar to the TmaSMC hinge. The similarity between the structures of the TmaSMC and the mouse condensin hinge domain is even more striking. Although eukaryotic SMC hinge domains are heterodimers, whereas the prokaryotic ones are homodimeric, the hinge domain fold is largely conserved between bacteria and eukaryotes.

Here, we investigated the structure and function of the SMC hinge domain from the thermophilic archaeon *Pyrococcus furiosus*. Its crystal structure more closely resembles the mouse condensin hinge domain than the TmaSMC hinge. Like bacterial^{9,12,13} and eukaryotic¹⁴ condensin hinge domains, but unlike the cohesin hinge domain¹¹, the *P. furiosus* SMC (PfuSMC) hinge displays a marked preference for single-stranded DNA (ssDNA) over

double-stranded DNA (dsDNA). However, its affinity for DNA is decidedly lower than that of the eukaryotic condensin hinge, and point mutations reveal that its DNA-binding surface is also much smaller.

MATERIALS AND METHODS

Cloning, expression and protein purification

The DNA fragment encoding the desired hinge domain construct spanning residues 488-667 of PfuSMC was PCR-amplified from a pET-21b vector (Novagen) containing the full-length *P. furiosus smc* gene and inserted into the NdeI/XhoI sites of pET-21b, to be expressed with the vector-encoded C-terminal hexahistidine tag. Point mutations were introduced into the vector by site-directed mutagenesis using the QuikChange method (Stratagene). All constructs were verified by sequencing.

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, and induced with 0.5 mM IPTG. Cells were harvested by centrifugation 5 h after induction. Cell pellets were stored at -20°C until further use. To obtain selenomethionine-labelled protein, the construct was expressed in *E. coli* B834 (DE3) additionally containing the pRARE plasmid (Novagen) in LeMaster's medium²¹ supplemented with the appropriate antibiotics and selenomethionine.

Proteins were purified via heat denaturation of contaminating proteins, nickel chelate affinity chromatography and gel filtration. 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 incubated at 70°C for 10 min, then 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 homodimeric PfuSMC hinge domain were pooled, and the protein was concentrated to ~50 mg/ml. Protein concentration was determined using a calculated extinction coefficient at 280 nm²². The purification process was monitored by SDS-PAGE²³. 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.

Protein crystallization and structure determination

The PfuSMC hinge domain 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 1.8 M ammonium sulphate, 0.2 M potassium-sodium tartrate, and 0.1 M trisodium citrate pH 6.0. Cryoprotection was achieved by briefly soaking crystals in mother liquor supplemented with 15% (v/v) D(-)-2,3-butane diol. Data were collected at beamline PXI of the Swiss Light Source (SLS, Villigen, Switzerland). The crystals belong to space group C222₁ and contain two molecules in the asymmetric unit (see Table I for crystallographic data, phasing and refinement statistics). The structure was determined by MAD phasing. Data were indexed and integrated using the XDS package²⁴. Phases were calculated with AutoSHARP²⁵. One chain was built manually in Coot²⁶, and the second chain was then generated by molecular replacement using Phaser^{27,28}. The model was refined against the high remote wavelength dataset of the selenomethionine derivative. NCS

restraints were only used in the first stages of refinement because the two chains present in the asymmetric unit are not completely identical. Initial refinement was carried out with CNS²⁹, followed by several rounds of refinement with phenix.refine³⁰ and rebuilding in Coot. Refinement included simulated annealing in initial cycles, individual atomic coordinate and isotropic *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 MolProbity³¹, RAMPAGE³², and PROCHECK³³. The electrostatic surface potential was calculated with the Adaptive Poisson-Boltzmann Solver (APBS)³⁴. All figures were prepared with PyMOL³⁵. Coordinates and structure factors have been deposited at the Protein Data Bank (PDB) with accession number **3NWC**.

Small-angle X-ray scattering of protein solutions

To prepare samples suitable for SAXS measurements, the PfuSMC hinge domain was 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 package³⁶ was used to process and analyze data. Theoretical scattering profiles from atomic resolution models were calculated and fitted to measured profiles with CRY SOL³⁷. *Ab initio* models of the PfuSMC hinge domain were reconstructed from the experimental data using the program GASBORi³⁸, initially without imposing any symmetry or other restrictions on possible models. Since all models were clearly two-fold symmetric, further models were calculated imposing two-fold symmetry. Eight independently reconstructed envelopes were

aligned and averaged with SUPCOMB³⁹ and DAMAVER⁴⁰. An envelope representation was calculated using the Situs package⁴¹, which was also used to dock atomic resolution models into the envelope.

***In vitro* DNA-binding assays**

DNA oligonucleotide substrates carrying a 6-FAM fluorescence label were used to monitor their binding to the PfuSMC hinge domain (Table II). HPLC-purified DNA oligonucleotides (Thermo Scientific) were dissolved in water. DNA concentration was determined using a calculated extinction coefficient at 260 nm⁴². 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 MgCl₂, 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 EMSAs contained 12.5 nM of a DNA substrate and protein in a 0-, 10-, 25-, 50-, 100-, 250-, 500-, 1000-, 2500-, and 5000-fold and 0-, 100-, 500-, 1000-, 2500-, 5000-, 10 000-, 25 000-, 50 000-, and 100 000-fold molar excess over ssDNA and dsDNA, respectively, in 1× 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× 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).

RESULTS

Protein purification, crystallization and structure determination

To facilitate crystallization, we designed an expression construct of the PfuSMC protein that spans the entire hinge domain including short stretches of the coiled-coil region (residues 488-667), with a hexahistidine tag at the C terminus. The construct was expressed in *E. coli*,

and the protein was purified via heat denaturation of contaminating proteins, nickel chelate affinity chromatography, and gel filtration. We obtained crystals in space group $C222_1$ with two molecules in the asymmetric unit. Selenomethionine derivative crystals diffracted to 1.7 Å resolution and allowed phase determination by multiple-wavelength anomalous dispersion (MAD). The resulting electron density map was easily interpretable. After several cycles of manual model building and refinement, the final R factors were 19.4% for R_{work} and 23.5% for R_{free} . In the final model, one chain spans residues 488-667, while the other chain encompasses residues 488-662. The geometry of both chains is well within acceptable range. Crystallographic data, phasing and refinement statistics are shown in Table I. An example of the initial and refined electron density can be found in Fig. 1A.

Crystal structure of the *P. furiosus* SMC hinge domain

The PfuSMC hinge domain consists of two pseudo-two-fold-symmetric subdomains that are linked by a long ordered loop passing along the bottom face of the hinge domain, i.e. the face opposite the coiled-coils. Each subdomain contains a mixed β -sheet and several α -helices on both sides of the sheet. The two subdomains interact to form a half-ring structure with an α -helical core that is bordered by a β -sheet on each side. The β -sheets are then flanked on the outside by one or two α -helices (Fig. 1B; see also the sequence alignment in Fig. 4A).

The tertiary structure of the archaeal PfuSMC hinge domain is similar to that of the bacterial TmaSMC hinge⁶ (Fig. 1C); the rms deviation between the $C\alpha$ traces of both proteins is 2.0 Å. Unexpectedly, however, the predicted coiled-coil segments at the N and C terminus of the PfuSMC hinge do not join up into a coiled-coil. Instead, the N-terminal coiled-coil segment folds into a helix that is lying on top of the hinge domain core, while the C-terminal segment forms a long loop (Fig. 1B and C). This conformation appears to be stabilized by crystal contacts: the N-terminal helices of both chains in the asymmetric unit contact each other, and

the C-terminal coiled-coil segment of one chain loops around the other chain (Fig. 1D and Fig. 2A). Since the coiled-coil segments are short – consisting of only four or five helical turns – the interaction between them is probably not very strong and therefore easily disrupted by crystal packing forces.

In crystals, the TmaSMC⁶ and the *E. coli* MukB^{19,20} hinge domains form two-fold-symmetric doughnut-shaped homodimers. The two chains present in the asymmetric unit of the PfuSMC hinge domain crystals assemble in a different manner, interacting via the predicted coiled-coil segments (Fig. 1D), but the doughnut-shaped dimer is generated by symmetry operators (Fig. 1E). In this symmetry-generated dimer, the two symmetrical interfaces between the monomers are formed largely by β -strand 3 of one monomer interacting with $\beta 6$ and $\beta 7$ of the other monomer to form two continuous mixed seven-stranded β -sheets ($\beta 1-3 + \beta 4-7$).

Additional dimer interactions are contributed by helices αE and αH which flank the β -sheets on the outside (Fig. 1E, Fig. 2B). As can be seen in Fig. 2, there are far more contacts between the symmetry-related chains than between the two chains in the asymmetric unit (e.g. 28 vs. 7 hydrogen bonds). The residues making up the interface between the symmetry-related chains are also among the most conserved residues in the hinge domains of SMC proteins from all three phylogenetic domains, suggesting that this dimeric assembly is the physiologically relevant oligomeric state of the SMC hinge domain (Fig. 4A).

Solution scattering analysis

To unambiguously determine whether the doughnut-shaped dimer of the SMC hinge domain is the assembly present in solution, we performed small-angle X-ray scattering (SAXS) experiments with the PfuSMC hinge domain. The SAXS data are of very high quality, as the protein did not show any signs of aggregation even at concentrations of up to 20 mg/ml (Fig. 3A). The molecular mass determined from the scattering intensity extrapolated to zero angle

confirms that the PfuSMC hinge domain exists as a homodimer in solution. The shape of the scattering profile indicates that this dimer has a globular conformation. To compare the solution to the crystal structure, theoretical scattering curves were calculated of both possible dimeric assemblies observed in the crystals (Fig. 3A). The scattering profile of the dimer in the asymmetric unit does not fit the observed scattering profile at all; this dimer is more elongated than the dimer present in solution. The theoretical scattering curve of the symmetry-generated dimer, however, has the same shape as the observed one, although the two profiles do not overlap completely. This is probably due to the conformation of the coiled-coil segments that is stabilized in the crystals, but likely to be flexible in solution, as would also be suggested by the solution envelope (see below, Fig. 3B). Initially, *ab initio* envelope reconstructions were calculated without imposing any symmetry. Since the resulting models were clearly two-fold symmetric, two-fold symmetry was then imposed to generate models of higher quality. Eight independent models were averaged to generate the final solution envelope model of the PfuSMC hinge domain into which both possible dimeric assemblies were docked (Fig. 3B). The dimer present in the asymmetric unit is too narrow to fill the solution envelope completely and too long to completely fit into it. The symmetry-generated dimer however is docked into the SAXS model so that the hole in the ring nicely coincides with the dent in the envelope. Only the C-terminal coiled-coil segment does not fit into the envelope, which is not surprising, since it would be expected to be disordered in solution. The flexibility of the coiled-coil segments probably also accounts for the fact that the envelope is slightly bigger than the symmetry-generated dimer in the crystal structure. We therefore conclude that the biological assembly of the SMC hinge domain is indeed the doughnut-shaped dimer.

The *P. furiosus* SMC hinge domain binds single-stranded, but not double-stranded DNA

Since other SMC hinge domains have been shown to bind DNA^{9,11-14}, we tested different DNA substrates for binding to the PfuSMC hinge domain (DNA sequences are listed in Table II). Electrophoretic mobility shift assays (EMSAs) were performed to assess DNA binding, using a 5'-6-FAM label to visualize the DNA substrates. In these assays, the PfuSMC hinge bound a 30-mer ssDNA substrate relatively well, whereas a dsDNA substrate of the same length was not bound (Fig. 5A and B). ~50% of the 30-mer ssDNA were bound at a 50 – 100-fold excess of protein over DNA, and at a 250 – 500-fold excess of protein, the ssDNA substrate was completely shifted. The protein concentration at which half the ssDNA bound was 0.6 – 1.3 μ M, therefore the dissociation constant of the complex should be in this range. The protein-ssDNA complex migrated as a defined band, indicating that a specific complex was formed. With the 30-mer dsDNA, only a smeared shift could be obtained at a very high (25 000 – 100 000-fold) excess of protein (Fig. 5B). This suggests that the shift was simply caused by the high protein concentration, not by specific binding. We also tested longer DNA substrates which essentially yielded the same results, while shorter ssDNA substrates were not bound efficiently (data not shown).

In order to more accurately define the DNA-binding interface of the PfuSMC hinge domain, we constructed two lysine-to-glutamate point mutants. We chose the lysines to be mutated in analogy to the mutations made in the mouse condensin hinge domain in our previous study¹⁴, but based our choice on a tertiary structure rather than a sequence alignment (Fig. 4).

While SMC hinge domains generally contain many lysines, there is only one lysine residue that is highly conserved among SMC proteins from all species (PfuSMC-K⁵⁶⁸, mSMC2-K⁵⁶⁶/mSMC4-K⁶⁵⁷) (Fig. 4A). We mutated this conserved lysine to glutamate in the mouse SMC2-SMC4 hinge domain in our previous study¹⁴. This residue is located at the C terminus

of the α -helix capping the dimer interface in all SMC hinge domain structures solved so far. In the mSMC2 hinge domain, K⁵⁶⁶ points outwards from the protein surface, whereas the corresponding residue in the PfuSMC hinge domain, K⁵⁶⁸, is involved in dimer interactions, forming a hydrogen bond with E⁶⁴¹ of the other chain (Fig. 4B). Since it has previously been shown that dimerization is necessary for DNA binding by SMC hinge domains^{11,12,14}, we did not want to disrupt this interaction. PfuSMC-K⁵⁶⁵, however, is in a similar position in the tertiary structure as mSMC2-K⁵⁶⁶ and was therefore mutated to glutamate instead of the conserved residue PfuSMC-K⁵⁶⁸. The second residue mutated to glutamate was PfuSMC-K⁶⁰⁵, located in the same region of the hinge domain as mSMC2-K⁶¹³/mSMC4-K⁶⁹⁸ which we mutated previously¹⁴ (Fig. 4A). This residue was chosen because, in contrast to PfuSMC-K⁵⁶⁵, it is distant from the dimer interface and located in an overall negatively charged region (Fig. 4B and Fig. 6).

Interestingly, while both mutants displayed reduced affinity towards ssDNA as compared to wild-type, the effect of the K⁵⁶⁵E mutation was stronger than that of the K⁶⁰⁵E mutation (Fig. 5C and D). The K⁵⁶⁵E mutant protein only started shifting the 30-mer ssDNA at a concentration of 3.1 μ M (250-fold excess), whereas the K⁶⁰⁵E mutant started shifting the DNA substrate at 0.6 μ M (50-fold excess), indicating that the affinity of the K⁵⁶⁵E mutant protein for ssDNA is approximately 5-fold lower than that of the K⁶⁰⁵E mutant. However, both mutant proteins failed to form a defined protein-DNA complex, instead producing only a smeared shift, and a complete shift was not obtained with either of the two mutant proteins even at high protein concentrations. This suggests that while both the basic patch as well as positively charged residues outside of the basic patch contribute to binding specificity, the binding strength is achieved mostly by the basic patch.

DISCUSSION

To close the gap between bacterial and eukaryotic SMC hinge domains, we solved the crystal structure of the archaeal SMC hinge domain from *P. furiosus*. The primary sequence of SMC hinge domains is strongly conserved: between SMC hinge domains from all three phylogenetic domains, as exemplified by those whose structures have been solved, i.e. the TmaSMC⁶, the PfuSMC, and the mouse condensin hinge domain¹⁴, the sequences are ~30% identical and 50 – 60% similar. Likewise, the tertiary structures of these SMC hinge domains are quite similar, with rms deviations all below 3 Å (Table III, Fig. 1C, Fig. 4B). However, while the overall fold of the SMC hinge domain is conserved from prokaryotes to eukaryotes, functionally relevant structural details have changed. Most importantly, the surface charge of the hinge domain has been almost reversed throughout evolution (Fig. 6). Whereas bacterial and archaeal SMC hinge domains have a quite acidic outer surface, the outer surface of eukaryotic SMC hinges, and especially of the condensin hinge, is mostly basic. As discussed below, this has implications in the DNA-binding activity and specificity of prokaryotic and eukaryotic SMC hinge domains.

The PfuSMC hinge domain binds DNA and displays a marked preference for ssDNA over dsDNA – more so than the mouse condensin hinge domain which binds dsDNA, albeit non-specifically¹⁴. The differential effects of the two lysine-to-glutamate point mutations on DNA binding by the PfuSMC hinge domain suggest that its DNA-binding activity resides mostly within the basic patch at the dimer interface. In contrast, in the mouse condensin hinge domain the analogous mutations both had the same quantitative and qualitative effect, reducing affinity as well as specificity, which implies that the ssDNA wraps around the outer surface of the hinge domain¹⁴. The affinity of the PfuSMC hinge domain for ssDNA is obviously lower than that of the mouse condensin hinge domain. In EMSAs, under the same assay conditions, the same 30-mer ssDNA substrate was shifted completely by the mouse

condensin hinge domain at a 100-fold excess of protein over DNA¹⁴, whereas with a 100-fold excess of the PfuSMC hinge domain only ~50% of the DNA are shifted (Fig. 5A). Comparing the surface charge distribution of both proteins, this is not surprising. While the PfuSMC hinge domain – like the TmaSMC hinge⁶ – has a mostly negatively charged surface with only one basic patch at the dimer interface, this basic patch is much more pronounced in the mouse condensin hinge, and the entire surface of the latter is mostly positively charged and contributes to DNA binding¹⁴ (Fig. 6). The mouse condensin hinge domain has a higher affinity for DNA than the PfuSMC hinge at the cost of a lower specificity for the structure of the DNA bound.

In comparison, our results indicate that the DNA-binding surface of the PfuSMC hinge domain is similar, but significantly smaller than that of the mouse condensin hinge. A previous mutational study of the *B. subtilis* SMC protein¹³ supports this conclusion, although residues outside of the basic patch were not mutated in this study. However, the authors found that mutation of three consecutive lysine residues in this basic patch of the *B. subtilis* SMC hinge led to a complete loss of DNA binding.

The *E. coli* MukB protein on the other hand has an extensive DNA-binding site on the head domain surface between the coiled-coils⁴³. No DNA-binding activity was observed with the MukB hinge domain which is even more acidic than genuine prokaryotic SMC hinge domains^{19,20}. This suggests that the DNA-binding activity is a feature acquired by SMC and MukB proteins after γ -proteobacteria diverged from other bacteria.

The fact that the ssDNA-binding activity appears to have been enhanced in condensin as compared to the prokaryotic SMC, while the cohesin hinge domain preferentially binds dsDNA¹¹, argues that this activity plays a role in a function acquired specifically by the condensin complex. Since condensin has been shown to be involved in single-strand break repair⁴⁴, we proposed previously that ssDNA-binding by its hinge domain might help tether

the condensin complex to a break site and/or organize the DNA structure for repair¹⁴. Prokaryotes employ different mechanisms to repair single-strand breaks. Hence the prokaryotic SMC complex must have a different function in DNA repair, if at all. But while condensin is loaded onto chromosomes in prophase of mitosis^{45,46}, the prokaryotic SMC complex appears to be loaded during replication, at replication origins⁴⁷⁻⁴⁹. The ssDNA-binding activity of its hinge domain might therefore play a role in the DNA-loading process of the prokaryotic SMC complex.

With more and more data on structure and activity of SMC hinge domains being available, it is becoming increasingly clear that while the SMC hinge domain fold has been strongly conserved throughout evolution, its function has diversified.

FUNDING

This work was supported by grants from the Deutsche Forschungsgemeinschaft (SFBs 646 and 684 and HO2489/3); the German Excellence Initiative (Center for Integrated Protein Science Munich and Munich Center for Advanced Photonics); and the Boehringer Ingelheim Fonds to J.J.G.

ACKNOWLEDGEMENTS

We thank Alexandra Schele for general assistance in the laboratory, the staff of beamlines X06SA/Swiss Light Source (SLS) and EMBL/DESY X33 for help and support during data collection, Katja Lammens and Gregor Witte for help with crystallographic and SAXS data, Matthew Bennett for critical reading of the manuscript, and the Hopfner group for valuable discussions. J.J.G. thanks the Boehringer Ingelheim Fonds for financial and personal support.

REFERENCES

1. Hirano T. At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* 2006;7(5):311-322.
2. Graumann PL, Knust T. Dynamics of the bacterial SMC complex and SMC-like proteins involved in DNA repair. *Chromosome Res* 2009;17(2):265-275.
3. Hudson DF, Marshall KM, Earnshaw WC. Condensin: architect of mitotic chromosomes. *Chromosome Res* 2009;17(2):131-144.
4. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. *Annu Rev Genet* 2009;43:525-558.
5. De Piccoli G, Torres-Rosell J, Aragon L. The unnamed complex: what do we know about Smc5-Smc6? *Chromosome Res* 2009;17(2):251-263.
6. Haering CH, Lowe J, Hochwagen A, Nasmyth K. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell* 2002;9(4):773-788.
7. Haering CH, Schoffnegger D, Nishino T, Helmhart W, Nasmyth K, Lowe J. Structure and stability of cohesin's Smc1-kleisin interaction. *Mol Cell* 2004;15(6):951-964.
8. Lammens A, Schele A, Hopfner KP. Structural biochemistry of ATP-driven dimerization and DNA-stimulated activation of SMC ATPases. *Curr Biol* 2004;14(19):1778-1782.
9. Hirano M, Hirano T. Positive and negative regulation of SMC-DNA interactions by ATP and accessory proteins. *EMBO J* 2004;23(13):2664-2673.
10. Arumugam P, Nishino T, Haering CH, Gruber S, Nasmyth K. Cohesin's ATPase activity is stimulated by the C-terminal Winged-Helix domain of its kleisin subunit. *Curr Biol* 2006;16(20):1998-2008.
11. Chiu A, Revenkova E, Jessberger R. DNA interaction and dimerization of eukaryotic SMC hinge domains. *J Biol Chem* 2004;279(25):26233-26242.
12. Hirano M, Hirano T. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *EMBO J* 2002;21(21):5733-5744.
13. Hirano M, Hirano T. Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions. *Mol Cell* 2006;21(2):175-186.
14. Griesse JJ, Witte G, Hopfner KP. Structure and DNA binding activity of the mouse condensin hinge domain highlight common and diverse features of SMC proteins. *Nucleic Acids Res* 2010;38(10):3454-3465.
15. Chen ES, Sutani T, Yanagida M. Ctf1/C1D interacts with condensin SMC hinge and supports the DNA repair function of condensin. *Proc Natl Acad Sci USA* 2004;101(21):8078-8083.
16. Duan X, Yang Y, Chen YH, Arenz J, Rangi GK, Zhao X, Ye H. Architecture of the Smc5/6 Complex of *Saccharomyces cerevisiae* Reveals a Unique Interaction between the Nse5-6 Subcomplex and the Hinge Regions of Smc5 and Smc6. *J Biol Chem* 2009;284(13):8507-8515.
17. Hirano M, Anderson DE, Erickson HP, Hirano T. Bimodal activation of SMC ATPase by intra- and inter-molecular interactions. *EMBO J* 2001;20(12):3238-3250.
18. Hirano M, Hirano T. ATP-dependent aggregation of single-stranded DNA by a bacterial SMC homodimer. *EMBO J* 1998;17(23):7139-7148.
19. Li Y, Schoeffler AJ, Berger JM, Oakley MG. The crystal structure of the hinge domain of the *Escherichia coli* structural maintenance of chromosomes protein MukB. *J Mol Biol* 2010;395(1):11-19.
20. Ku B, Lim JH, Shin HC, Shin SY, Oh BH. Crystal structure of the MukB hinge domain with coiled-coil stretches and its functional implications. *Proteins: Struct Funct Bioinform* 2010;78(6):1483-1490.

21. LeMaster DM, Richards FM. ^1H - ^{15}N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry* 1985;24(25):7263-7268.
22. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. *The Proteomics Protocols Handbook*: Humana Press; 2005. p 571-607.
23. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227(5259):680-685.
24. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 1993;26:795-800.
25. Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Methods Mol Biol* 2007;364:215-230.
26. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Cryst D* 2010;66(Pt 4):486-501.
27. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. *Acta Cryst D* 1994;50(Pt 5):760-763.
28. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. *Acta Cryst D* 2005;61(Pt 4):458-464.
29. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Cryst D* 1998;54(Pt 5):905-921.
30. Afonine PV, Grosse-Kunstleve, R.W., Adams, P.D. The Phenix refinement framework. *CCP4 Newsletter* 2005;42:contribution 8.
31. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 2007;35(Web Server issue):W375-383.
32. Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by C alpha geometry: phi, psi and C beta deviation. *Proteins: Struct Funct Genet* 2003;50(3):437-450.
33. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck - a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26:283-291.
34. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 2001;98(18):10037-10041.
35. DeLano WL, DeLano Scientific LLC, Palo Alto, CA, USA. The PyMOL Molecular Graphics System; 2008.
36. Konarev PV, Petoukhov MV, Volkov VV, Svergun DI. ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Crystallogr* 2006;39:277-286.
37. Svergun D, Barberato C, Koch MHJ. CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 1995;28:768-773.
38. Svergun DI, Petoukhov MV, Koch MH. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 2001;80(6):2946-2953.
39. Kozin MB, Svergun DI. Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 2001;34:33-41.
40. Volkov VV, Svergun DI. Uniqueness of *ab initio* shape determination in small-angle scattering. *J Appl Crystallogr* 2003;36:860-864.

41. Wriggers W, Chacon P. Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. *J Appl Crystallogr* 2001;34:773-776.
42. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* 2007;35(Web Server issue):W43-46.
43. Woo JS, Lim JH, Shin HC, Suh MK, Ku B, Lee KH, Joo K, Robinson H, Lee J, Park SY, Ha NC, Oh BH. Structural studies of a bacterial condensin complex reveal ATP-dependent disruption of intersubunit interactions. *Cell* 2009;136(1):85-96.
44. Heale JT, Ball AR, Jr., Schmiesing JA, Kim JS, Kong X, Zhou S, Hudson DF, Earnshaw WC, Yokomori K. Condensin I interacts with the PARP-1-XRCC1 complex and functions in DNA single-strand break repair. *Mol Cell* 2006;21(6):837-848.
45. Hirota T, Gerlich D, Koch B, Ellenberg J, Peters JM. Distinct functions of condensin I and II in mitotic chromosome assembly. *J Cell Sci* 2004;117(Pt 26):6435-6445.
46. Oliveira RA, Heidmann S, Sunkel CE. Condensin I binds chromatin early in prophase and displays a highly dynamic association with *Drosophila* mitotic chromosomes. *Chromosoma* 2007;116(3):259-274.
47. Mascarenhas J, Volkov AV, Rinn C, Schiener J, Guckenberger R, Graumann PL. Dynamic assembly, localization and proteolysis of the *Bacillus subtilis* SMC complex. *BMC Cell Biol* 2005;6:28.
48. Gruber S, Errington J. Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* 2009;137(4):685-696.
49. Sullivan NL, Marquis KA, Rudner DZ. Recruitment of SMC by ParB-parS organizes the origin region and promotes efficient chromosome segregation. *Cell* 2009;137(4):697-707.
50. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988;16(22):10881-10890.

FIGURE LEGENDS

Fig. 1. (A) Electron density around residues H⁶⁵³-F⁶⁵⁴-R⁶⁵⁵ of PfuSMC in the crystal structure of the PfuSMC hinge domain (residues 488-667). Left panel, experimental electron density contoured at 1.5 σ ; right panel, final 2F_o-F_c electron density contoured at 1.5 σ . Residues are shown as stick models with carbon atoms colored green, nitrogen: blue, and oxygen: red. (B) Stereo view of the PfuSMC hinge domain monomer in cartoon representation, colored by subdomains. The N-terminal subdomain is colored blue, the C-terminal subdomain violet. The N- and C-terminal segments predicted to belong to the coiled-coil region and the loop connecting the subdomains are colored grey. (C) Superposition of the PfuSMC hinge domain monomer with the TmaSMC hinge domain (pdb 1GXL)⁶. The PfuSMC hinge is colored as in (B), the TmaSMC hinge is shown in white. (D) The dimer present in the asymmetric unit of the PfuSMC hinge domain crystals. Chain A is shown in blue, chain B in green. (E) The symmetry-generated dimer of the PfuSMC hinge domain. Chain A is shown in blue, chain A' in pink.

Fig. 2. (A) Stereo view of the interface between the two chains in the asymmetric unit of the PfuSMC hinge domain crystals. Residues contributing to the interaction are shown as stick models with carbon atoms colored blue and green for chain A and B, respectively, nitrogen: dark blue, and oxygen: light red. The protein backbone is depicted as ribbon model in light blue and light green for chain A and B, respectively. Hydrogen bonds are represented by dashed blue lines. (B) Stereo view of the dimer interface between the symmetry-related chains in the PfuSMC hinge domain crystal structure. Residues contributing to the subunit interaction are shown as stick models with carbon atoms colored blue and pink for chain A and chain A', respectively, nitrogen: dark blue, and oxygen: light red. The protein backbone

is depicted as ribbon model in light blue and violet for chain A and chain A', respectively. Hydrogen bonds are represented by dashed blue lines.

Fig. 3. (A) Small-angle X-ray scattering profiles of the PfuSMC hinge domain. The scattering profile of the protein in solution is shown in black, the theoretical scattering profiles of the dimer present in the asymmetric unit of the crystals and the symmetry-generated dimer are shown in green and pink, respectively. (B) Solution envelope reconstruction of the PfuSMC hinge domain (grey mesh); left panel, superimposed with the dimer present in the asymmetric unit of the PfuSMC hinge domain crystals (chain A colored blue, chain B green); right panel, superimposed with the symmetry-generated dimer (chain A colored blue, chain A' pink).

Fig. 4. (A) Sequence alignment of the hinge domains of bacterial, archaeal and eukaryotic condensin SMC proteins. Pfu, *P. furiosus*; Tma, *T. maritima*; Bsu, *B. subtilis*; Sc, *S. cerevisiae*; m, *Mus musculus*. Numbering of residues is for PfuSMC. The secondary structure of the PfuSMC hinge domain is shown above the alignment, that of the TmaSMC hinge domain (pdb 1GXL)⁶ below, with α -helices displayed as purple rectangles and β -strands as blue arrows. The lysine residues mutated in PfuSMC are indicated by blue asterisks, those mutated in mSMC2 and mSMC4¹⁴ by orange asterisks. In the consensus sequence, lower case letters are used for $\geq 50\%$, upper case letters for $\geq 90\%$ conservation; ! is any one of I/V, and # is any one of N/D/Q/E. The alignment was generated with MultAlin⁵⁰. (B) Superposition of the PfuSMC hinge domain (chain A colored light blue, chain A' light pink) with the mouse SMC2 hinge domain (orange; pdb 3L51)¹⁴. The lysine residues mutated in the PfuSMC hinge domain are shown as stick models in blue, corresponding lysine residues in the mSMC2 hinge in orange. Note that the conserved residue K⁵⁶⁶ of mSMC2 corresponds to PfuSMC-K⁵⁶⁸ (shown as stick model in black) in a primary sequence alignment, but superimposes with

PfuSMC-K⁵⁶⁵ in the tertiary structure. PfuSMC-K⁵⁶⁸ is involved in dimer interactions, forming a hydrogen bond with E⁶⁴¹ of the other chain (shown as stick model in pink).

Fig. 5. Electrophoretic mobility shift assays. (A) Titration of a 30-mer ssDNA substrate with the wild-type PfuSMC hinge domain (PfuSMCh^{wt}). (B) Titration of a 30-mer dsDNA substrate with PfuSMCh^{wt}. (C) Titration of the 30-mer ssDNA substrate with the point mutant PfuSMCh^{K565E}. (D) Titration of the 30-mer ssDNA substrate with the point mutant PfuSMCh^{K605E}. The 6-FAM-labelled DNA substrates at a fixed concentration of 12.5 nM were incubated with increasing concentrations of protein (in a molar excess over the DNA as detailed underneath each lane) in 1× PBS for 30 min at room temperature. After addition of 10% glycerol, samples were separated on 0.5% agarose gels in 1× TB buffer. Asterisks indicate free DNA, the arrow indicates a defined protein-DNA complex.

Fig. 6. (A) Surface charge distribution of the PfuSMC hinge domain, looking onto the dimer interface and side surface; and (B) the mouse SMC2-SMC4 hinge domain (pdb 3L51)¹⁴, looking onto the face from which the coiled-coils would emerge. Negatively charged regions are colored red, positively charged regions blue, and neutral regions white. Asterisks mark PfuSMC-K⁵⁶⁵ and the corresponding residues in the mouse SMC2-SMC4 hinge domain mSMC2-K⁵⁶⁶/mSMC4-K⁶⁵⁷, arrows mark PfuSMC-K⁶⁰⁵ and the corresponding residues mSMC2-K⁶¹³/mSMC4-K⁶⁹⁸.

Table I. Crystallographic data, phasing and refinement statistics. Data were collected at beamline PXI of the Swiss Light Source (SLS). The values given in parentheses are for the highest resolution shell. Friedel mates were treated as independent reflections. The high remote wavelength dataset was used for refinement. The R_{free} factor was calculated from 10% of the data which were removed at random before the structure was refined. Ramachandran statistics were calculated with RAMPAGE³⁹.

Dataset	peak	inflection	high remote
wavelength (Å)	0.9788	0.9794	0.9778
space group	C222 ₁	C222 ₁	C222 ₁
unit cell dimensions (Å, a/b/c)	69.79 / 118.88 / 82.83	69.87 / 119.06 / 83.00	69.71 / 118.65 / 82.64
unit cell angles (°, α/β/γ)	90.0 / 90.0 / 90.0	90.0 / 90.0 / 90.0	90.0 / 90.0 / 90.0
redundancy	6.8 (6.6)	6.8 (6.6)	6.7 (6.5)
R_{sym} (%)	4.7 (34.9)	5.2 (44.1)	4.6 (33.4)
resolution (Å)	35.0-1.8 (1.9-1.8)	35.0-1.8 (1.9-1.8)	35.0-1.7 (1.8-1.7)
completeness (%)	99.1 (95.5)	99.0 (94.9)	99.3 (96.8)
$I/\sigma I$	24.18 (4.77)	21.73 (3.91)	24.09 (4.73)
phasing power (acentric-anomalous)	2.161	0.832	1.776
R_{Cullis} (acentric-anomalous)	0.60	0.88	0.65
Figure of merit (acentric)	0.42		

Refinement

Dataset	high remote
number of reflections used in refinement	38,167
resolution (Å)	34.0-1.7
Model $R_{\text{work}}/R_{\text{free}}$ (%)	19.4 / 23.5
protein nonhydrogen atoms	2840
water molecules	251
overall B factor (Å ²)	
all atoms	29.5
protein main and side chains	29.0
water	36.1
rms deviation from ideal geometry	
bonds (Å)	0.006
angles (°)	0.991
Ramachandran statistics (%)	
most favored/allowed/disallowed regions	97.7 / 2.3 / 0.0

Table II. Oligonucleotides used for DNA-binding assays.

Name	Oligonucleotide sequences
30-mer ssDNA	5'-6-FAM-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
30-mer dsDNA	strand 1: 5'-6-FAM-CCGGAAAGCATCTAGCATCCTGTCAGCTGC strand 2: 5'-GCAGCTGACAGGATGCTAGATGCTTTCCGG

Table III. Rms deviations between the C α traces of the SMC hinge domains from *P. furiosus* (PfuSMCh), *T. maritima* (TmaSMCh; pdb 1GXL)⁶, and mouse condensin (mSMC2h, mSMC4h; pdb 3L51)¹⁴.

	TmaSMCh	mSMC2h	mSMC4h
PfuSMCh	2.0 Å	1.4 Å	1.9 Å
TmaSMCh		2.2 Å	2.6 Å
mSMC2h			1.8 Å

Accepted Article

Figure 1
159x108mm (300 x 300 DPI)

Figure 2
160x61mm (300 x 300 DPI)

Accepted Pre-proof

Figure 3
160x49mm (600 x 600 DPI)

John Wiley & Sons, Inc.

Accepted Article

Figure 5
160x117mm (600 x 600 DPI)

Figure 6
77x116mm (300 x 300 DPI)