Structural Maintenance of Chromosomes Protein C-terminal Domains Bind Preferentially to DNA with Secondary Structure*

(Received for publication, May 6, 1998)

Alexandre T. Akhmedov, Christian Frei‡, Monika Tsai-Pflugfelder‡, Börries Kemper§, Susan M. Gasser‡,¶ and Rolf Jessberger†

From the Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland; ‡ISREC, Chemin des Boveresses 155, CH-1066 Epalinges/Lausanne, Switzerland; and the §Institute for Genetics, University of Cologne, Zülpicher Strasse 47, D-50674 Cologne, Germany

Structural maintenance of chromosomes (SMC) proteins interact with DNA in chromosome condensation, sister chromatid cohesion, DNA recombination, and gene dosage compensation. How individual SMC proteins and their functional domains bind DNA has not been described. We demonstrate the ability of the C-terminal domains of Saccharomyces cerevisiae SMC1 and SMC2 proteins, representing two major subfamilies with different functions, to bind DNA in an ATP-independent manner. Three levels of DNA binding specificity were observed: 1) a >100-fold preference for double-stranded versus single-stranded DNA; 2) a high affinity for DNA fragments able to form secondary structures and for synthetic cruciform DNA molecules; and 3) a strong preference for AT-rich DNA fragments of particular types. These include fragments from the scaffold-associated regions, and an alternating poly(dA-dT)-poly(dT-dA) synthetic polymer, as opposed to a variety of other polymers. Reannealing of complementary DNA strands is also promoted primarily by the C-terminal domains. Consistent with their in vitro DNA binding activity, we show that overexpression of the SMC C termini increases plasmid loss without altering viability or cell cycle progression.

The structural maintenance of chromosomes (SMC)1 protein family, with members from lower and higher eukaryotes, may be divided into four subfamilies (SMC1 to SMC4) and two SMC-like protein subfamilies (SMC5 and SMC6) (for reviews see Refs. 1–7). Members of this family appear to function primarily as heterodimers and are implicated in a large range of activities that modulate chromosome structure and organization. Studies of yeast strains deficient in SMC1 and SMC2 show defects in the segregation of mitotic chromosomes (8–10). A role in chromosome condensation was demonstrated in a cell-free chromosome condensation assay based on Xenopus laevis oocyte extracts (11), in which SMC2 and SMC4 subtypes were identified as essential components of the condensin protein complex (12). More recently, the SMC1p and SMC3p of Saccharomyces cerevisiae were shown to be essential for sister chromatid cohesion (13, 14), and in Caenorhabditis elegans, two SMC protein homologs, MIX-1 and DPY-27, are involved in gene dosage compensation (15, 16). Finally, evidence for a role of SMC proteins in DNA recombination and repair has been supported by both genetic and biochemical studies. The Schizosaccharomyces pombe Rad18 gene and its S. cerevisiae homolog RHC18 are SMC-like and were found to act in an unusual postreplicative recombinational repair pathway (17), whereas mammalian SMC1 and SMC3 proteins were described as essential subunits of a recombinational repair protein complex (RC-1), isolated from calf thymus (18, 19).

SMC proteins display a very characteristic structure: two coiled-coil domains separate evolutionarily conserved head and tail domains, which contain an NTP binding motif (Walker A box) in the N terminus, and a DA box (Walker B) in the C terminus (20). The similarity of SMC proteins to motor proteins such as kinesin and myosin has been noted, and both chromatin clamp and DNA motor functions have been proposed for SMC proteins (1–3). SMC proteins generally tend to form either SMC1/SMC3 or SMC2/SMC4 heterodimers (7). These heterodimers then form subcomponents of larger multiprotein complexes, which serve specific biological functions. Examples are the X. laevis 13 S condensin, in which XCAP-C and -E (SMC4 and 2, respectively) are complexed with three non-SMC proteins, required for chromosome condensation (12); the bovine recombination protein complex RC-1, in which SMC1 and SMC3 molecules are complexed with DNA ligase III and DNA polymerase ε (19); or the C. elegans dosage compensation complex consisting of the two SMC proteins DPY-26 and MIX-1, associated with other proteins (16).

Thus, it has become evident that SMC proteins in their various heterodimeric combinations and multiprotein complexes act in a range of reactions involving DNA and chromosome dynamics. To date, however, knowledge about the enzymatic activities and the molecular characteristics that allow the SMC proteins or their individual domains to fulfill these various functions has remained very limited. Central to these reactions is the interaction of SMC proteins with DNA. Therefore, we have analyzed DNA binding activities in vitro and in vivo of domains of the S. cerevisiae SMC1 and SMC2 proteins, which represent the two major classes of SMC heterodimers and complexes.

EXPERIMENTAL PROCEDURES

Purification of S. cerevisiae SMC Protein Domains—N- and C-terminal domains of SMC1 and SMC2 were expressed in E. coli as GST fusion proteins. Domains were cloned by polymerase chain reaction from plasmid pAS130 (SMC1) or pAS405 (SMC2) (8, 10) using following primers: SMC1-N, 5'-CCGGAATTCCGCAATGGGACGTTT AGTTGGC and 5'-GGCGAATTCCGCAATGGGACGTTT AGTTGGC; SMC2-N, 5'-GGCGAATTTAATGCCAATGAGCAGGTTT TTGTTTTTGCTATGG; SMC2-C, 5'-C CGGAATTTAATGCCAATGAGCAGGTTT TTGTTTTTGCTATGG; SMC2-N, 5'-C CGGAATTTAATGCCAATGAGCAGGTTT TTGTTTTTGCTATGG; and SMC2-C, 5'-C CGGAATTTAATGCCAATGAGCAGGTTT TTGTTTTTGCTATGG.

* The Basel Institute for Immunology was founded and is supported by Hoffman-LaRoche Inc., Basel, Switzerland. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡ Supported by ISREC, the Swiss National Science Foundation, the Human Frontiers Science Program, and the Swiss Cancer League.
¶ To whom correspondence should be addressed. Tel.: 41-61-605-1284; Fax: 41-61-605-1364; E-mail: jessberger@bii.ch.
1 The abbreviations used are: SMC, structural maintenance of chromosomes; SAR, scaffold-associated region; bp, base pair(s); CEN, centromere; GST, glutathione S-transferase; ss, single-stranded; ds, double-stranded.
SMC Proteins Bind to DNA with Secondary Structure

SMC Proteins Bind to DNA with Secondary Structure

TCTTAG AAGATTGTTGCCCTCGTGCGG, SMC1-C, 5'-GGGCAATTC-TTATGGCAGCACATTGACCC and 5'-GGCGTTCTTTTGCGG-GGTATACAGGAGTTTGG; SMC2-C, 5'-GGCGAATTCAGAAGACATTGCAAAACATGGGG and 5'-GGGGCTTTTTTACAATTACTAAGGG. Polymerase chain reactions were performed with a Pfu polymerase (Boehringer Mannheim) and fused in frame into the pGEX-3X plasmid (Amersham Pharmacia Biotech): SMC1-N, pCF310; SMC2-N, pCF301; SMC1-C, pCF312; and SMC2-C, pCF314. All domains were sequenced to detect potential mutations. See “Results” for exact boundaries of the domains.

For protein purification, the fusion proteins were induced in transformants harboring pTS1 for 16 h with 1 mM IPTG at 22 °C. The following steps were all performed at 4 °C. The cells were washed, resuspended in phosphate-buffered saline and 1 mM phenylmethylsulfonyl fluoride, 1 mM MgCl2, 1 mM dithiothreitol, 1 mM ATP, 0.13 M NaCl, 1 mM diethiothreitol, 1 mM ATP, 1 mM glycerol, and 10% glycerol. The proteins were stored in phosphate-buffered saline and 1 mM phenylmethylsulfonyl fluoride, 1 mM MgCl2, 1 mM dithiothreitol, 1 mM ATP, 0.13 M NaCl, 1 mM diethiothreitol, 1 mM ATP, 1 mM glycerol, and 10% glycerol. The proteins were stored in aliquots at −80 °C.

**Gel Retardation Assays**—These assays were performed in 10-μl reaction mixtures containing 0.5–1 ng (3000–6000 cpm) of 32P-labeled DNA in 20 mM HEPES (pH 7.5), 0.5 mM EDTA. All gels were fixed (60 min in 10% acetic acid, 10% ethanol), dried, and exposed for autoradiography. The following DNA substrates were used: a 210-bp EcoRI-Kod M13mp18 DNA fragment (positions 6001–6231), a 232-bp AvaII fragment from pUC19 (positions 1837–2059), the polymers poly(dA-dT), poly(dC-dG), poly(dA-dT)-poly(dT-dA), poly(dG-dC)-poly(dC-dG), poly(dA-dC)-poly(dG-dT) (average length, 8000 bp; synthesized enzymatically; purchased from two sources, Sigma and Amersham Pharmacia Biotech), the gel-purified, structure-tested cruciform DNA substrates CF01 (29 nucleotides/oligonucleotide; Ref. 21) or CFM13 (29 nucleotides/oligonucleotide; Ref. 22), a 600-bp centromeric region DNA from CEN3 of *S. cerevisiae* (23); and a 189-bp bent fragment from the Drosophila *ftz* regulatory region DNA. DNA-protein complexes were resolved by electrophoresis at 4 °C in non-denaturing polyacrylamide gels in 20 mM HEPES (pH 7.5), 0.5 mM EDTA. All gels were fixed (60 min in 10% acetic acid, 10% ethanol), stained with propidium iodide, and analyzed by FACsScan analysis.

**Fluorescence-activated Cell Sorting Analysis**—To follow the cell cycle progression of yeast cells overexpressing SMC proteins, we monitored the genomic DNA content by fluorescence-activated cell sorting analysis. The diploid yeast strain, RS453, was transformed with the SMC-containing expression plasmids or pBSD.04-HA as control. Transformed yeast cells were grown for 12 h in synthetic drop-out medium lacking uracil and supplemented with 2% (w/v) glucose, 3% (v/v) glycerol, and 2% (w/v) lactate. The cultures were diluted (1:100) into the same medium, lacking glucose, and grown for 24 h to deplete glucose from the medium and to allow maximal induction of the hybrid GALURA/CYC promoter by galactose. Measurements were taken every 2 h after the addition of 2% (w/v) galactose, and cells were fixed by washing in 70% ethanol, stained with propidium iodide, and analyzed by FACsScan analysis. Control cultures were treated identically, but growth was continued on glucose throughout.

**RESULTS**

**Purification of SMC Protein Terminal Domains**—In order to analyze the biochemical properties of the individual SMC proteins, the N- and C-terminal globular domains of the *S. cerevisiae* SMC1 and SMC2 proteins were expressed in bacteria as GST-fusion proteins and were purified in their native forms. All purified proteins were >95% homogeneous (Fig. 1), except the SMC2-C preparation, which contained one minor contaminant of roughly 68 kDa. The SMC2-C domain appears to be somewhat less stable because 29-kDa GST fragment is visible as a breakdown product (see comparison with purified GST, Fig. 1). Both types of the fusion polypeptides were designed to contain the entire N- and C-terminal globular domains, as follows: SMC1-N, amino acids 1–184; SMC2-N, amino acids 1–174; SMC1-C, amino acids 1070–1225; SMC2-C, amino acids 1026–1170 (see Fig. 1).
stranded DNA—The binding of the SMC terminal domains to a variety of DNA substrates was monitored by gel retardation assays. Increasing amounts (0–80 ng) of N- or C-terminal domain proteins of SMC1 and SMC2 were incubated with a 230-bp fragment of M13 RF DNA (Fig. 2A). A specific signal was obtained with both C-terminal domains with as little as 10 ng of protein (0.23 pmol; 70-fold molar excess over the probe), but not with either N-terminal domain, even when 80 ng, or a 500-fold molar excess over the probe, was added. Mixing equal amounts of both domains to a total of 80 ng of protein did not affect the gel shift (data not shown). The GST domain alone does not bind DNA at 40 or 80 ng of protein. Thus we conclude that the C-terminal domain, which includes the DA-box and almost no coiled-coil region, is sufficient for DNA binding.

For dsDNA is therefore about 130-fold. Similar results were obtained for the SMC1 C-terminal domain and an overexpressed C-terminal domain of the bovine SMC1 protein (not shown).

The use of either full-length ds or ss M13 DNA revealed similar results: the ds form competes at least 100-fold better than the ss form (data not shown). It was also observed, however, that when intact full-length plasmids were used as competitor, the ds M13 DNA was nearly 100-fold more efficient than either the ds φX174 or pBR322 DNA at competing for the interaction with the 230-bp M13 fragment (Fig. 2C). As discussed below, this suggested to us a preference of the SMC domain for structure or sequence features in the M13 DNA. These features are examined in more detail below.

**DNA Secondary Structures Are Preferred Binding Substrates for SMC C-terminal Domains**—Computer-aided analysis (GCG stem-loop program, Madison, Wisconsin; based on Ref. 29) of the 230-bp M13 fragment (bp 6001–6231), which was used as a labeled binding substrate in the experiments described above, showed a very high potential of this fragment to form secondary structures, i.e. stem-loops. The highest score stem-loop in this fragment (bp 6170–6206) contains 14 bp in the stem, 9 in the loop and a quality score of 30, the second highest of the entire M13 genome.

The 13 S condensin protein complex from *X. laevis*, which contains SMC proteins of the SMC2 and 4 subfamilies, has been shown earlier to bind to synthetic cruciform DNA (30). It was not clear, however, which of the five polypeptide subunits of that complex possessed the binding activity, or whether the complex needed to be intact to bind these substrates. Moreover, it was unknown whether individual SMC proteins or their DNA binding domains could do so. Using the purified SMC1 and SMC2 C-terminal domains, we found the C-terminal domains bind preferentially to cruciform DNA (substrate CF01), as compared with a linear dsDNA substrate bearing a part of the same DNA sequence (Fig. 3A). In the gel retardation assay,
SMC Proteins Bind to DNA with Secondary Structure

four bands could be detected, two of which were also present as very faint bands in the negative control and in lanes containing the N-terminal domains. Several new retarded species were generated by interactions with the C-terminal domains of SMC1 and SMC2. The nature of the weak background bands is not clear, but they may reflect differently folded forms of cruciform DNA. In any case, the intensities of these bands are substantially enhanced in the presence of SMC proteins as well.

Almost no binding was detected with the linear substrate, whereas binding to the cruciform was readily detectable in both the SMC1 and SMC2 C-terminal proteins. As before, the N-terminal domains and the GST protein control did not interact significantly with either DNA. The apparently lower binding capacity of the SMC1-C compared with SMC2-C in Fig. 3A is not significant because in other experiments SMC1-C gave an even stronger signal than SMC2-C (Table I, Experiment 2). This variability is in part due to a tendency for the SMC domains to aggregate or misfold upon repeated freeze-thaw cycles (not shown). In no case, however, was the linear DNA substrate bound with an efficiency comparable to that of the cruciform substrate. Quantitation of the autoradiograph showed in average a 20-fold higher affinity for the cruciform over the linear substrate (Table I). Similar results were obtained using another cruciform and using a C-terminal domain cloned from the bovine SMC1 protein with the same cruciform DNA substrate (not shown).

**Preference for Alternating dA-dT Polymers**—SMC proteins have been proposed to bind to the AT-rich SARs, which are implicated in the compaction of loops during chromosome condensation (31). These regions are naturally bent, are generally >70% A+T, contain oligo(dA-dT) stretches, and often have potential to form stem-loop structures (24, 32, 33). We therefore tested the efficiency of a series of synthetic DNA polymers as competitors for the complex formed with the labeled 230-bp M13 fragment. The polymers included poly(dA)-poly(dT), poly(dG)-poly(dC), and the alternating polymers poly(dA-dT)-poly(dT-dA), poly(dG-dC)-poly(dC-dG), and poly(dA-dC)-poly(dT-dG) with an average length of 8000 bp. Increasing amounts of these competitor DNAs were added to the binding reaction with the 230-bp M13 dsDNA fragment as labeled binding substrate (Fig. 4A). Both homopolymers, poly(dA-dT) and poly(dG-dC), compete with low efficiency, as did the alternating poly(dG-dC)-poly(dC-dG) and poly(dA-dC)-poly(dT-dG) polymers. The alternating poly(dA-dT)-poly(dT-dA) polymer, however, was over 200-fold more efficient than all other synthetic polymer competitors, and it matches the efficiency of competition by the homologous M13 sequence (see Fig. 4B).

Because the synthetic competitor DNA used reflects the structure of A/T (or G/C)-rich sequences in natural DNA regions to only a limited extent, we tested DNA fragments derived from Drosophila SAR and yeast centromere (CEN) regions as competitors. Like Drosophila SAR DNA, the yeast centromeric DNA contains both a highly stable stem-loop structure and appropriately spaced oligo-da stretches that confer both a narrow minor groove and bent character. Both fragments also bind the yeast nuclear scaffold (23). In a gel retardation assay using the SMC1 C terminus and the labeled 230-bp M13 fragment, we have used as competitors a 189-bp ftz SAR DNA fragment (80% A/T; Ref. 24), a 600-bp centromeric DNA fragment derived from *S. cerevisiae* CEN3 (80% A/T; Ref. 23), and the 230-bp M13 control DNA (Fig. 4B). No significant difference in competition by either the SAR or the CEN DNA substrates could be observed, and both compete only slightly less well than the M13 DNA fragment, which is, serendipitously, one of the best competitors identified. We conclude that the SAR and CEN fragments are strongly bound substrates for SMC proteins.

Because these results again suggest a preference of SMC protein binding to DNA secondary structures, we also selected as a binding substrate a DNA fragment of the same length (232 bp) from the plasmid pUC19 (bp 1837–2059), which has no or a very low potential to form stem-loops. The overall G/C content of the M13 and the pUC19 fragments is very similar (51.7 and 49.8% respectively), yet this DNA fragment was very poorly bound by the SMC1-C protein (Fig. 4C), confirming the earlier observation that SMC proteins have significantly higher affinity for DNA with a propensity to form secondary structures.

**DNA Reannealing Is Promoted by the C-terminal Domain**—From the above studies, we concluded that C-terminal domains of SMC proteins not only have the capacity to bind DNA but also have very distinct substrate preferences. We previously demonstrated the ability of the bovine SMC1/3 heterodimer to promote reannealing of complementary strands of DNA (19). This reaction relates to a possible pairing function of SMC proteins in DNA recombination. The *S. pombe* cut3/cut14 heterodimer was also shown to promote such a reaction (34). To determine whether the reaction is inherent in the DNA binding capacity of the C-terminal domains of SMC proteins, we compared the efficiency of SMC-C, SMC-N, and *Escherichia coli* RecA proteins in this assay.

Increasing amounts of either SMC1 or SMC2 N- or C-terminal domains were incubated with 32P-labeled 422-nucleotide complementary single-stranded DNA substrates, and the products were analyzed by gel electrophoresis (Fig. 5). Although both terminal domains promoted the reannealing reaction to some extent, about 10-fold lower amounts of the C termini of both SMC proteins (between 0.2 and 0.8 pmol) were necessary to obtain a comparable signal as revealed by densitometric quantification of the reannealing products (Table II). No signal above background was obtained at these low amounts with the N-terminal domains (not shown). Thus, the C termini are the primary mediators in this reaction. The reaction with both N- and C-terminal domains does not depend on, nor is it affected by, the presence of ATP (not shown). The efficiency of this reaction was standardized to the reannealing catalyzed by the RecA protein. On a molar basis, the SMC C termini are about 2-fold more efficient than RecA protein at promoting this reaction (Table II).

---

**Table I**

| DNA substrate | Protein | Protein amount | Bound |
|---------------|---------|----------------|-------|
| **Experiment 1** | | | |
| Cruciform (CF01) | No protein | 0 | <0.1 |
| Cruciform (CF01) | GST | 20 | 2.1 |
| Cruciform (CF01) | SMC1-N | 10 | 1.1 |
| Cruciform (CF01) | SMC2-N | 10 | 2.1 |
| Cruciform (CF01) | SMC1-C | 10 | 8.5 |
| Cruciform (CF01) | SMC1-C | 10 | 19.4 |
| Cruciform (CF01) | SMC2-C | 10 | 44.8 |
| Cruciform (CF01) | SMC2-C | 20 | 54.2 |
| Linear | No protein | 0 | 0.9 |
| Linear | SMC1-C | 10 | 1.6 |
| Linear | SMC2-C | 20 | 2.8 |
| **Experiment 2** | | | |
| Cruciform (CF01) | No protein | 0 | <0.1 |
| Cruciform (CF01) | GST | 30 | 1.1 |
| Cruciform (CF01) | SMC1-N | 30 | 1.9 |
| Cruciform (CF01) | SMC1-C | 10 | 23.6 |
| Cruciform (CF01) | SMC2-C | 20 | 43.1 |
| Cruciform (CF01) | SMC2-C | 10 | 12.3 |
| Cruciform (CF01) | SMC2-C | 20 | 14.6 |
Protein. ss substrates and ds product are indicated. The indicated amounts of either the SMC N-terminal or C-terminal protein were added to the reannealing reaction, which contained heat-denatured ssDNA. The indicated amounts of either the SMC N-terminal or C-terminal protein were added to the reannealing reaction, which contained heat-denatured ssDNA. The indicated amounts of either the SMC N-terminal or C-terminal protein were added to the reannealing reaction, which contained heat-denatured ssDNA.

The indicated amounts of either the SMC N-terminal or C-terminal protein were added to the reannealing reaction, which contained heat-denatured ssDNA. The indicated amounts of either the SMC N-terminal or C-terminal protein were added to the reannealing reaction, which contained heat-denatured ssDNA.

The N-terminal SMC Protein Domains Bind 8-Azido-ATP—Although the C-terminal domains possess specific DNA binding and reannealing activities, they also contain a DA box, which is thought to be involved in NTP binding and/or hydrolysis. The major NTP binding element, a Walker A motif (20), however, is located in the N-terminal domain. Using the photoflavin ATP-analogue 8-azido-[γ-32P]ATP (25), we investigated binding of ATP to the individual domains and found that only the N-terminal domains bind the ATP analogue in vitro (Fig. 6). In a related experiment, we observed 8-azido-ATP binding to the N-terminal domain of the human SMC1 protein SB1.8 (35), which was efficiently competed by addition of ATP (not shown). Addition of the C-terminal to the N termini or of DNA did not affect the reaction (not shown). Thus, the C termini can neither bind ATP on their own, nor are they required for binding of ATP by the N-terminal domains.

Overexpression of the SMC C Termini Increases CEN Plasmid Loss in Vivo—We next asked whether the observed DNA binding specificities are important in vivo. It is well established that SMC1 and SMC2 genes are essential for proper chromo-

some segregation in yeast (8, 10) and that SMC1 is a key component of the cohesin complex that mediates sister chromatid cohesion (12, 13). By analogy to XCAP-E, SMC2 is thought to be part of condensin, which helps mediate chromosome condensation (10, 12). If the observed interaction of the SMC C termini with DNA is of physiological relevance, we might expect that overexpression of this domain would enhance plasmid loss, by competing for the binding of the larger cohesin or condensin complex. To test this, we have overexpressed the C-terminal subdomains of SMC1 and SMC2, and monitored the rate of loss of an ADE2-containing centromere plasmid (26). As shown in Table III, we observed a consistent increase in plasmid loss rates from 2.4 to 4.0 or 3.7% when either SMC1-C or SMC2-C is overexpressed, as compared with the host strain grown under identical conditions carrying a vector that expresses only a short epitope tag. The rather large standard deviation in these elevated loss rates correlates with fluctuations in expression levels of these proteins, as determined by Western blot analysis (data not shown).

Overexpression of full-length SMC1 and SMC2 from identical single copy plasmids also results in a slight increase in plasmid loss (Table III). The levels of induction for the C-terminal domains and the full length proteins are approximately equal as judged by Western blotting, yet the overexpression of full-length SMC proteins provokes an accumulation of cells in late G2 and a delay for transition through mitosis, which was not observed when cells were grown on glucose or when the C-terminal domains were overexpressed (data not shown and Table III). In no case did we observe a loss in cell viability. The plasmid loss due to full-length SMC overexpression may reflect an imbalance in the assembly of the condensin and cohesin complexes. It is noteworthy that the small SMC1C-terminal domain is even more effective at provoking plasmid loss than the full-length SMC1, although it does not alter cell cycle progression and is unable to homo- or heterodimerize. Its effect may therefore reflect interference at the level of interac-

2 C. Frei, unpublished results.
tions between the sister chromatid cohesion complex (12, 13) and plasmid DNA.

DISCUSSION

To date, almost all studies of SMC proteins have used either heterodimers formed by two SMC proteins or large multisubunit protein complexes, which contain a particular SMC heterodimer. Therefore, our knowledge of biochemical functions of SMC proteins themselves, and of their highly conserved N- and C-terminal domains, remained very limited. Cloning and overexpression of domains of S. cerevisiae SMC1 and SMC2 proteins has allowed us to study molecular properties of these proteins in detail and to assign specific functions to these domains. The C-terminal DA box domain of both SMC1 and SMC2 was found to be sufficient for DNA binding (Fig. 2). We cannot exclude the possibility that the binding specificity might be altered by dimerization nor the possibility that the coiled-coil regions might also bind DNA; but neither mixing the N- and C-terminal domains together in one reaction nor using full-length SMC1 or SMC2 proteins yielded stronger signals in the gel shift experiments (not shown). The C-terminal domains of both SMC proteins preferred dsDNA over ssDNA (Fig. 2). As revealed by titration of competitor DNA, the affinity for dsDNA is more than 100-fold higher than that for ssDNA of the same origin. Both linear and supercoiled circular DNA served efficiently as competitors.

Two further levels of DNA binding specificity have been revealed by our studies. We observed (i) a strong preference for DNA fragments able to form DNA secondary structures, and (ii) an equally strong preference for AT-rich DNA fragments of a particular type. The secondary structure preference was observed with both synthetic cruciform DNA and with phage DNA fragments having a high potential to form such structures. Plasmid-derived DNA fragments carrying no secondary structures were bound poorly or not at all (Fig. 4C), as were ds oligonucleotides from 25 to 75 bp in length, which also lack secondary structures (Fig. 3 and data not shown). Little difference was observed between short DNA fragments and large circular DNA competitor molecules, as long as both contained secondary structures (Figs. 2C and 4). The localization of SMC proteins on mitotic chromosomes (8, 10, 36) and their role in chromosome condensation has led to the idea that SMC proteins may have a preferred association with SARs (2, 37). These regions are characterized by a high AT content with oligo(dA-dT) stretches, low \( T_m \), and a narrow minor groove (24, 38). A strict consensus sequence, however, has not been found, prompting the idea that DNA structure is the defining feature for these regions. SAR sequences have been implicated in chromosome condensation in vitro (31), and also in chromosome segregation in vivo, because the yeast centromere has SAR activity (23). SMC proteins might preferentially recognize these structures and thereby mediate loop formation, chromosome condensation, and/or sister chromatid cohesion.

Among synthetic polymers, only the poly(dA-dT)-poly(dT-dA) was bound with an efficiency comparable to that of the 230-bp M13 DNA fragment or SAR and CEN fragments. The alternating poly(dA-dT)-poly(dT-dA) polymer may have features partially resembling those of SARs, for in contrast to the homopolymer poly(dA-dT), which is very rigid, it readily forms secondary structures. Although the alternating poly(dG-dC)-poly(dC-dG) polymer can also theoretically form stem-loop structures, it does not act as a good competitor. The difference between these two polymers is that poly(dA-dT)-poly(dT-dA) has a low \( T_m \), perhaps allowing the duplex to melt and readily form the secondary structures, whereas the poly(dG-dC)-poly(dC-dG) polymer has a very high \( T_m \). In this scenario, the affinity of SMC proteins for the alternating poly(dA-dT)-poly(dT-dA) substrate would be attributable to a preference for secondary structure, in agreement with the preference for cruciform DNA, and for DNA fragments like the 230-bp M13 DNA fragment with a high potential to form secondary structures as described above.

This secondary structure preference may have important biological ramifications. In DNA replication, secondary structures and topologically strained DNA can be generated if two replication bubbles approach and encounter each other completing replication of a certain chromosomal region (for review, see Ref. 39). At this stage, the partially entangled sister chromatids need to be resolved by topoisomerase II, and at about the same time, sister chromatid cohesion may start, requiring specific associations between the two DNA duplexes. Cruciform-like structures or similar DNA distortions may serve as signals for SMC proteins to bind and co-mediate sister chromatid cohesion. Importantly, both SARs and yeast centromeres have the potential to form stable stem-loop structures, indicating that the interactions observed in vitro may have implications in chromosome stability in vivo.

Consistent with this, we observed enhanced plasmid loss in S. cerevisiae upon expression of the SMC C-terminal domains from single copy plasmids, although the expression has no effect on either cell cycle progression and viability. These results support our proposal that essential DNA-protein interactions are mediated by these domains. These assays are performed under fairly low levels of overexpression, (the induced proteins are not visible by Coomassie Blue staining of total protein extracts), because strong galactose-induced expression from 2-\( \mu \)m plasmids leads not only to rapid loss of the ADIE2-marked reporter but to loss of the expression plasmid itself and to inviability (data not shown). Thus, whereas high levels of induction are counterelective and do not allow accurate quantitation of plasmid loss, low level overexpression results in a small but reproducible increase in plasmid loss (Table III).

The 13 S condensin multiprotein complex has been shown to bind preferentially to synthetic cruciform DNA (30). It seems likely that this property of the complex is contributed by its SMC protein subunits, in particular their C-terminal domains. The complex acts in chromosome condensation, and it was hypothesized that local DNA distortions, such as bent DNA or other secondary structures, may be preferentially recognized by the complex, which is then proposed to aid condensation by introducing superhelical tension (30).

In DNA recombination, cruciform DNA is the paradigm of Holliday junction intermediates of homologous recombination (40). Earlier, we showed that the bovine SMC1/3 heterodimer catalyzes the formation of joint molecules,\(^3\) which represent early intermediates in DNA recombination. In this context, the preferential binding of SMC proteins to cruciform DNA may be important. During strand invasion, crossed DNA structures

\(^3\) R. Jessberger, submitted for publication.
may be generated, which may be recognized and stabilized by SMC proteins, thereby driving the reaction toward the product. In a standard homologous recombination reaction, Holliday junctions exist after strand exchange and may be stabilized by SMC proteins as well. Intriguingly, stem-loop structures have been shown to coincide with in vivo recombination sites in some cases (41).

Several proteins that are known to bind DNA secondary structures or distortions in the DNA structure with high affinity, such as the HMG proteins or SRY, not only bind but also bend the DNA (42–44). Similar to what we demonstrate for SMC proteins, only a particular domain of the HMG proteins, the HMG box, is needed for its binding specificity, and also the bending of DNA (43, 45). Other chromatin remodeling factors, such as SWI/SNF, also bind to duplex DNA with preferences for distorted or secondary structures (46). The preference for cruciform DNA may also reflect a high affinity for DNA at the base of the nucleosome where two helices are in close proximity. It remains to be shown whether SMC proteins bind the nucleosome fiber.

The presence of an N-terminal NTP binding motif, the Walker A box, is one of the characteristics of the SMC protein family (3–7, 37). We show here, however, binding of 8-azido-ATP to the N-terminal but not to the C-terminal domain of the S. cerevisiae SMC1 and SMC2 proteins (Fig. 6). So far, however, all attempts to detect ATP hydrolysis by isolated SMC proteins or subdomains failed. We were also unable to detect ATP hydrolysis by the isolated bovine SMC1/3 heterodimer under various experimental conditions (19). The complex RC-1, however, possesses an ATPase activity, which is partially stimulated by the presence of DNA (19). Likewise, the complete X. laevis 13 S condensin complex, which contains SMC2 and SMC4 proteins, possesses ATPase activity (30). Thus, it seems likely that ATP hydrolysis by SMC proteins requires interaction(s) with other proteins. ATP is also not required for promotion of DNA reannealing by a mammalian SMC heterodimer (19), nor is it required by the C termini of either SMC1 or SMC2. This again underscores the autonomy of the SMC C termini for these reactions with DNA. In a more complex, recombination-specific D-loop assay for joint molecule formation, however, the C termini alone did not work (not shown), whereas a bovine SMC heterodimer was active. Thus, although the C termini are sufficient for DNA recognition, more complex interactions will no doubt require the SMC heterodimer.

Acknowledgments—We thank Drs. Fraser McBlane and Heinz Jacobs for critical reading of the manuscript.

REFERENCES

1. Peterson, C. L. (1994) Cell 79, 389–392
2. Gasser, S. M. (1995) Curr. Biol. 5, 357–360
3. Saitoh, N., Goldberg, I., and Earshaw, W. C. (1995) BioEssays 17, 759–766
4. Hohloch, T. (1995) Trends Biochem. Sci. 20, 357–361
5. Hirano, T., Mitchison, T. J., and Swedlow, J. R. (1995) Curr. Biol. 5, 329–336
6. Kosland, D., and Strunnikov, A. (1996) Annu. Rev. Cell. Dev. Biol. 12, 305–333
7. Jessberger, R., Frei, C., and Gasser, S. M. (1998) Curr. Opin. Genet. Dev. 8, 254–259
8. Strunnikov, A., Lariozon, V. L., and Kosland, D. (1993) J. Cell Biol. 123, 1635–1648
9. Saka, Y., Sutani, T., Yamashita, Y., Saitoh, S., Takeuchi, M., Nakaseko, Y., and Yanagida, M. (1994) EMBO J. 13, 4938–4952
10. Strunnikov, A. V., Hogan, E., and Koshland, D. (1995) EMBO J. 14, 985–991
11. Hirano, T., and Mitchison, T. J. (1994) Cell 79, 449–458
12. Hirano, T., Kobayashi, R., and Hirano, M. (1997) Cell 89, 511–521
13. Gasser, V., Kosland, D., and Strunnikov, A. (1997) Cell 89, 47–58
14. Michaels, C., Ciosk, R., and Nasmyth, K. (1997) Cell 91, 35–46
15. Chuan, P. T., Lieb, J. D., and Meyer, B. (1996) Science 274, 1736–1738
16. Lieb, J. D., Albrecht, M. R., Chuan, P.-T., and Meyer, B. J. (1998) Cell 92, 33–32
17. Lehmann, A. R., Walicka, M., Griffiths, D. J., Murray, J. M., Watts, F. Z., McCreasy, S., and Carr, A. M. (1995) Mol. Cell. Biol. 15, 7067–7080
18. Jessberger, R., Pedust, V., Hülshaber, U., and Berg, P. (1995) J. Biol. Chem. 268, 15070–15079
19. Jessberger, R., Riwar, B., Baechtold, H., and Akhmedov, A. T. (1996) EMBO J. 15, 4061–4068
20. Walker, J. E., Sarasate, M., Runswick, M. J., and Gay, N. J. (1982) EMBO J. 1, 945–951
21. Kupfer, C., and Kemper, B. (1996) Eur. J. Biochem. 238, 77–87
22. Gola, S., Christoph, A., Birkenshamp-Dentro-Dreer, R., and Kemper, B. (1997) Eur. J. Biochem. 245, 573–580
23. Amati, B., and Gasser, S. M. (1988) Cell 54, 967–978
24. Amati, B., Pick, L., Larocche, T., and Gasser, S. M. (1990) EMBO J. 9, 4007–4016
25. Knight, K. L., and McEntee, K. (1985) J. Biol. Chem. 260, 867–872
26. Hieter, P., Mann, C., Snyder, M., and Davis, R. W. (1985) Cell 40, 381–392
27. Doty, V., Wepf, R., and Hart, E. C. (1994) EMBO J. 13, 6062–6075
28. Stotz, A., and Linder, P. (1990) Gene 95, 91–98
29. Tinoco, I., Borer, P. N., Dinger, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1992) Nature 360, 40–41
30. Kimura, K., and Hirano, T. (1997) Cell 90, 625–634
31. Strick, R., and Laemmli, U. K. (1995) Cell 83, 1137–1148
32. Gasser, S. M., and Laemmli, U. K. (1986) Cell 46, 521–530
33. Homberger, H. P. (1989) Chromosoma 98, 99–104
34. Sutani, T., and Yanagida, M. (1997) Nature 388, 788–801
35. Rocques, P. J., Clark, J., Ball, S., Crew, J., Gill, S., Christodoulou, Z., Borte, R., Louis, E. J., Davies, K. E., and Cooper, C. S. (1995) Hum. Mol. Genet. 4, 243–249
36. Saitoh, N., Goldberg, I. G., Wood, E. R., and Earshaw, W. C. (1994) J. Cell Biol. 127, 303–318
37. Heck, M. M. S. (1997) Cell 91, 5–8
38. Laemmli, U. K., Kas, E., Poljak, L., and Adachi, Y. (1992) Curr. Opin. Genet. Dev. 2, 275–285
39. Bastia, D., and Mohanty, B. K. (1996) in DNA Replication (de Pamphilis, M., ed.) Monograph, Vol. 31, pp. 177–216, CSH Press
40. Holliday, R. (1964) Genet. Res. 5, 282–304
41. Musmann, R., Court, M., Schwager, J., and Du Pasquier, L. (1997) Eur. J. Immunol. 27, 2610–2619
42. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992) EMBO J. 11, 4497–4506
43. Pil, P. M., and Lippard, S. J. (1992) Science 256, 234–237
44. Po¨hler, J. R. G., Norma, D. G., Bramham, J., Bianchi, M. E., and Lilley, D. M. J. (1998) EMBO J. 17, 817–826
45. Quinn, J., Fyrberg, A. M., Gauster, R. W., Schmidt, M. C., and Peterson, C. L. (1996) Nature 379, 844–847