ScII: An Abundant Chromosome Scaffold Protein Is a Member of a Family of Putative ATPases with an Unusual Predicted Tertiary Structure

Noriko Saitoh, Ilya G. Goldberg, Edgar R. Wood, and William C. Earnshaw

Department of Cell Biology and Anatomy, Johns Hopkins School of Medicine, Baltimore, Maryland 21205

Abstract. Here, we describe the cloning and characterization of ScII, the second most abundant protein after topoisomerase II, of the chromosome scaffold fraction to be identified. ScII is structurally related to a protein, Smclp, previously found to be required for accurate chromosome segregation in *Saccharomyces cerevisiae*. ScII and the other members of the emerging family of *SMCI*-like proteins are likely to be novel ATPases, with NTP-binding A and B sites separated by two lengthy regions predicted to form an α-helical coiled-coil. Analysis of the ScII B site predicted that ScII might use ATP by a mechanism similar to the bacterial recN DNA repair and recombination enzyme. ScII is a mitosis-specific scaffold protein that colocalizes with topoisomerase II in mitotic chromosomes. However, ScII appears not to be associated with the interphase nuclear matrix. ScII might thus play a role in mitotic processes such as chromosome condensation or sister chromatid disjunction, both of which have been previously shown to involve topoisomerase II.

In interphase human cells, chromosomal DNA molecules totaling ~2 m long are packaged into nuclei that are only ~10 μm in diameter. At mitosis, the chromosomes become further condensed by about fourfold. This overall 10,000-fold compaction of the DNA is accomplished by a hierarchy of DNA and chromatin packaging (Earnshaw, 1991). At the lowest level, the DNA fiber is compacted six- to sevenfold by winding around the histone core of the nucleosome, generating fibers of ~10 nm diameter (Kornberg, 1974). At the second level, association of histone H1 with the 10-nm fiber causes the fiber to shorten and thicken to ~30 nm in diameter, bringing the overall compaction of the DNA to ~40-fold (Finch and Klug, 1976; Thoma et al., 1979; Horowitz et al., 1994). How the remaining 250-fold compaction of the 30-nm fiber is accomplished remains a matter of active investigation and considerable controversy (Earnshaw, 1991).

At present, the most widely accepted model for higher order chromosome structure proposes that the 30-nm fiber is gathered into loops, each containing ~50–100 kb of DNA, and tethered to nonhistone proteins of the nuclear scaffold or matrix (Laemmli et al., 1978). This model proposes that at the onset of mitosis, the scaffold proteins at the base of the loops associate with one another, thus pulling the chromosomal loop domains closer together. The aggregates of chromosomal scaffolding with their associated loops are thought to form either rosettes that coil along a helical path (Comings and Okada, 1971; Rattner and Lin, 1985; Boy de la Tour and Laemmli, 1988) or stack above one another to form minibands (Pienta and Coffey, 1984). Recent microscopy analysis using DNA fluorochromes under conditions where they bind preferentially to AT-rich or GC-rich DNA has suggested that mitotic chromosome arms consist of a more or less tightly coiled axial region of AT-rich DNA with loops of GC-rich DNA protruding from it (Saitoh and Laemmli, 1994).

The loop models of chromosome organization all suggest, however, that the chromatin fiber is packed into the final chromosome structure, and special molecules must exist that bind to the chromatin and define the base of each loop domain. At present, both the DNA sequences and polypeptide components that comprise this putative loop-fastener complex are unknown, although candidates for both have been suggested. The polypeptide components have been suggested to be components of the mitotic chromosome scaffold (or nuclear matrix) Adolph et al., 1977*a*, 1977*b*; Izaurralde et al., 1989; Zhao et al., 1993). The DNA sequences are known variously as MARs or SARs (matrix or scaffold attachment regions) (Mirkovitch et al., 1984; Gasser et al., 1989).

Chromosome scaffold proteins comprise the 5–10% of nonhistone chromosomal proteins that remain insoluble after treatment of isolated metaphase chromosomes with nucleases and subsequent extraction under a variety of conditions, including high salt (2 M NaCl), low ionic strength (dextran sulfate/heparin), or chaotropes (lithium diiodosalicylate) (Lewis and Laemmli, 1982). Although the chromo-

Address all correspondence to William C. Earnshaw, Department of Cell Biology and Anatomy, Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, MD 21205. Phone: (410) 955-2591. The present address for Edgar R. Wood is Division of Cell Biology, Burroughs Wellcome Co., 3030 Cornwallis Road, Research Triangle Park, NC 27709.

© The Rockefeller University Press, 0021-9525/94/10/303/16 $2.00

some scaffold is in reality a biochemical fraction, the term has been widely interpreted as describing a structural network within mitotic chromosomes. This, in part, results from the observation that isolated chromosome scaffolds retain the overall chromosomal morphology, with paired sister chromatids and condensed centromeres (Adolph et al., 1977*b*; Earnshaw and Laemmli, 1983). However, the role, if any, played by chromosome scaffold proteins in chromosome structure and function remains an important unsolved question.

The first chromosome scaffold protein to be conclusively identified was DNA topoisomerase II (Earnshaw et al., 1985; Berrios et al., 1985; Gasser et al., 1986) (initially termed ScI [Lewis and Laemmli, 1982]). This protein turns out to be the major component of the chromosome scaffold fraction (Gasser et al., 1986; Heck and Earnshaw, 1986). Several independent mapping techniques revealed that topoisomerase II is concentrated in the axial region of expanded mitotic chromosomes, and that it is largely absent from the expanded chromosomal loop domains (Earnshaw and Heck, 1985; Gasser et al., 1986).

Functional studies support the notion that topoisomerase II plays an essential role in mitotic chromosome structure and function both early and late in mitosis. The protein is required for normal chromosome condensation in fission yeast (Uemura et al., 1987), and also for chromosome condensation in vitro when interphase nuclei or naked DNA are added to mitotic extracts prepared from *Xenopus* eggs (Adachi et al., 1991; Hirano and Mitchison, 1993). The role of topoisomerase II during chromosome condensation is not known. On the one hand, it is possible that the enzyme is simply required to sort out DNA entanglements that impede orderly chromosome condensation. For example, the mitotic condensation process may serve as a rectification mechanism, whereby neighboring chromosomes are untangled from one another so that they can assort independently during mitosis (Holm, 1994). On the other hand, topoisomerase II may actually make a structural contribution to the condensed chromosome (Earnshaw et al., 1985; Gasser et al., 1986; Adachi et al., 1991). This could occur through interactions of the protein with the SAR/MAR sequences that have been proposed to form the base of chromosomal loop domains (Adachi et al., 1989). The notion that topoisomerase II plays a structural role in chromosomes is controversial, even when results obtained with the same experimental system are compared (Adachi et al., 1989; Hirano and Mitchison, 1993).

Genetic analysis in the yeasts has revealed that topoisomerase II is required for disjunction of sister chromatids at anaphase (DiNardo et al., 1984; Holm et al., 1985; Uemura and Yanagida, 1986). This function is also conserved in vertebrates, as shown both by drug treatments of cultured cells (Downes et al., 1991; Clarke et al., 1993) and by analysis of sister chromatid disjunction in *Xenopus* cell cycle extracts (Shamu and Murray, 1992). It has been speculated that assembly of topoisomerase II into the chromosomal structure might be important for regulation of its action during disjunction of sister chromatids (Earnshaw et al., 1985).

These studies of topoisomerase II provided the first concrete evidence that members of the chromosome scaffold fraction actually do play an important role in mitotic chromosome structure and function. However, with the exception of CENP-B (Earnshaw et al., 1984; Earnshaw and Rothfield, 1985), CENP-C (Earnshaw et al., 1984; Earnshaw and Rothfield, 1985), CENP-E (Yen et al., 1991), and the INCENPs (Cooke et al., 1987), all of which are concentrated in and around the centromere, topoisomerase II has remained the only member of this fraction to be characterized. Other abundant members of the fraction, including ScII (135 kD) (Lewis and Laemmli, 1982) and ScIII (140 kD), have remained unstudied.

In this paper, we provide the first characterization of ScII. We prepared antibodies to chicken ScII and used them to obtain cDNA clones encoding the chicken polypeptide. Like topoisomerase II, ScII is concentrated in the axial region of swollen chromosomes throughout the entire length of the chromosome arms. Biochemical fractionation confirms that ScII is a prominent component of the mitotic chromosome scaffold fraction. However, the protein associates only very loosely with interphase nuclei, with ≥95% leaking out into the cytoplasm during Dounce homogenization. Thus, ScII is not a component of the nuclear matrix. DNA sequence analysis reveals that ScII is a member of an emerging family of proteins with two internal regions of coiled-coil and highly conserved NTP-binding motifs at the amino and carboxy termini. The best characterized member of this family, Smclp, is required for accurate chromosome segregation in the budding yeast *Saccharomyces cerevisiae* (Strunnikov et al., 1993). Analysis of the deduced polypeptide sequence, together with previous results, suggests that ScII may be a chromosomal enzyme that may function in a complex with topoisomerase II.

Materials and Methods

Isolation of Chicken Chromosome Scaffold Proteins and Production of Guinea Pig Antibody

Mitotic chromosomes were isolated from chicken lymphoblastoid cell line MSB-1 as previously described (Earnshaw et al., 1985). Scaffolds were prepared by subjecting chromosomes to nuclease digestion and 2 M NaCl extraction (Adolph et al., 1977*b*). The pelleted scaffold fraction was applied onto preparative SDS-PAGE, the gels were stained with Coomassie blue, and a 135-kD band was excised and used to immunize a guinea pig to produce antiserum 5B2 (Earnshaw et al., 1985).

cDNA cloning

Library Screening. Antiserum 5B2 was used to screen an oligo-dT-primed λgt11 library prepared from chick embryo fibroblast mRNA (gift of B. Vennstrom, Karolinska Institute, Stockholm, Sweden). Expression screening (Ausubel et al., 1991) yielded three positive clones. These clones were used to screen a λZAP library made from MSB-1 mRNA (Mackay et al., 1993). The eight clones recovered by high stringency nucleic acid hybridization screening were characterized by restriction enzyme mapping and by partial sequencing. The largest clone, pBSc20 (insert size = 3.8 kb), was used in further experiments.

Rapid Amplification of cDNA Ends (RACE). To recover cDNA sequences 5′ to the beginning of the clone pBSc20, RACE-PCR (Frohman et al., 1988) was used. First, ScII-specific antisense primers (5′CCAAAGAGATCCTGTACACGGTTGTTGGAAGCATTCTA3′ [nt 545-507 in Fig. 2 *A*] or 5′TTGCTTAGTGGAATGATGGTGTAT3′ [nt 1889-1866 in Fig. 2 *A*]) were annealed to 1 μg of poly $(A)^+$ RNA isolated from logarithmically growing MSB-1 cells with the Fast Track mRNA Isolation Kit (Invitrogen, San Diego, CA). Next, the first strand of cDNA was synthesized with 10 U of AMV reverse transcriptase (Life Technologies, Inc., Bethesda, MD). A poly$(A)^+$ tail was attached to the 3′ end of the first strand with 25 U of terminal deoxynucleotidyl transferase (Boehringer-Mannheim Biochemicals Corp., Indianapolis, IN). The second strand was synthesized with 2.5 U of Amplitaq enzyme (Perkin-Elmer Corp., Norwalk, CT) and RACE-1 primer

(5'GATGGATCCTGCAGAAGCTTT$_{13}$3'), and both strands were subsequently amplified with the ScII specific primer (5'-CCAAAGAGATCCTGTACACGGTTGTTGGAAGCATTCTA-3' [545-507 nt in Fig. 2 *A*]) and RACE-2 primer (5'GATGGATCCTGCAGAAGC3'). Other PCR reaction conditions were as previously described (Frohman et al., 1988). The PCR products were treated with T4 DNA polymerase (Life Technologies) to generate blunt ends, and were ligated into the EcoRV site of the Bluescript II KS$^-$ vector (Stratagene, La Jolla, CA).

Sequencing. Phage inserts were transferred to Bluescript plasmids by recloning after EcoRI digestion (λgt11) or through in vivo excision (λZAP). Nested deletions were generated with the Erase-A-Base kit (Promega Biotech, Inc., Madison, WI) (Henikoff, 1984), and the resulting double-stranded templates were sequenced with Sequenase 2.0 according to the manufacturer's instructions (U.S. Biochemical Corp., Cleveland, OH).

Sequence information was entered into a Macintosh computer with MacVector 3.5 (IBI, New Haven, CT) hardware and software. A contig representing the ScII cDNA and including a single long open reading frame (ORF)[1] was assembled. MacVector 3.5 provided preliminary sequence information on the ORF, while an estimate of the probability of coiled-coil formation was generated by the COILCOIL program (Lupas et al., 1991). The ORF was compared to the set of nonredundant databases resident on the National Center for Biotechnology Information Internet server on June 15, 1994, by the TFASTA (Pearson, 1990) and BLASTP (Altschul et al., 1990) search network service.

Fluorescence Microscopy

Paraformaldehyde Fixation. MSB-1 cells were centrifuged at 1,200 *g* and resuspended in 1× RSB buffer (10 mM Tris/HCl, pH 7.4, 10 mM NaCl, 5 mM $MgCl_2$) for 5 min. Cells were placed onto glow-discharged coverslips (Earnshaw and Heck, 1985), and were centrifuged in a centrifuge (TJ6-R; Beckman) for 2.5 min at 2,800 rpm. The coverslips were immersed in D-PBS (8.06 mM Na_2HPO_4, 1.47 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 0.68 mM $CaCl_2$, and 0.492 mM $MgCl_2$) for 5 min, and were then fixed for 5 min in 3% paraformaldehyde in D-PBS at room temperature. Then, coverslips were washed with KB$^+$ (10 mM Tris/HCl, pH 7.7, 150 mM NaCl, 0.1% BSA, and 0.1% Triton X-100) three times, and were incubated with anti-ScII antibodies (1:200 dilution) in KB$^-$ (10 mM Tris/HCl, pH 7.7, 150 mM NaCl, and 0.1% BSA) at 37° for 1 h. After washing with KB$^-$, cells were incubated with biotinylated goat anti-rabbit IgG (1:500 dilution; Vector Laboratories, Burlingame, CA) at 37° for 30 min, followed by washing with KB$^-$. To visualize the secondary antibody, cells were incubated with streptavidin-Texas red (1:800 dilution; GIBCO-BRL, Gaithersburg, MD) at 37° for 30 min, followed by a wash with KB$^-$. To stain DNA, 4',6' diamino phenylindole (DAPI) (Calbiochem, San Diego, CA) was added to 1 μg/ml to the final wash (Williamson and Fennell, 1975).

Methanol-Acetic Acid Fixation. Cells were grown on coverslips, then swollen in 0.8% sodium citrate for 10 min before fixation with 50% methanol-acetic acid (3:1 vol/vol) for 2 min. This was followed by immersion of coverslips in 100% methanol-acetic acid (3:1 vol/vol) for 5 min at room temperature. Chromosomes were spread and dried using an aquarium air pump (Earnshaw et al., 1989) (Hagen 800), and swollen by immersion in three changes (2 min each) of TEEN (1 mM triethanolamine/HCl, pH 8.5, 0.2 mM NaEDTA, and 25 mM NaCl), to which 0.1% BSA and 0.5% Triton X-100 had been added (Earnshaw et al., 1989). Next, cells were incubated with primary antibody at 1:200 dilution in the same buffer for 1 h at 37°. Incubations with secondary and tertiary detection reagents were performed in KB$^+$ buffer as described above. For topoisomerase II staining, human autoimmune serum to topo II (Hoffmann et al., 1989) was added to fixed cells at 1:100 dilution, and followed by incubation with fluorescein-conjugated goat anti-human IgG at 1:100 dilution (Cappel Laboratories, Durham, NC).

Bacterial Expression and Antibody Production

DNA fragments corresponding to peptides A (aa 120-970), N (aa 120-567), M (aa 569-970), and C (aa 969-1189) were ligated in-frame into the pT7-7 vector (Tabor and Richardson, 1985), and transformed into *Escherichia coli* strain BL21(DE3) (Novagen Inc., Madison, WI). Transformants were cultured in L broth containing 50 μg/ml ampicillin at 37°. Expression was induced by making cultures 0.1 mM in IPTG, and allowing them 3–4 h of continued growth. To check protein expression, 1 ml of culture was centrifuged (1 min at 12,000 *g*), and the cell pellet was dissolved in 100 μl of SDS-PAGE sample buffer, briefly sonicated, and loaded onto an SDS-polyacrylamide gel. For large-scale preparation of antigens, partial purifications were performed as follows. Cells from 300 ml of induced culture were collected by centrifugation (5 min at 3,000 *g*), resuspended in 10-ml of solution A (10 mM Tris/HCl, pH 7.5, 25% sucrose, and 1 mM EDTA), then incubated on ice for 1 h in the presence of 1 mg/ml lysozyme (Sigma Immunochemicals, St. Louis, MO). Bacterial DNA was sheared by sonication on ice, then Nonidet NP-40 and NaCl were added to final concentrations of 2% and 0.1 M, respectively. Each of the expressed polypeptides (A, N, M, and C) was found to be concentrated in the pellet fraction after centrifugation for 10 min at 10,000 g. After two washes with solution B (50 mM Tris/HCl, pH 7.5, 0.1 M NaCl, 1 mM EDTA, and 0.1% NP-40), the pellet was dissolved in 1 ml sample buffer and subjected to SDS-PAGE. Each major Coomassie blue–stained band of the appropriate size was excised from the gel, ground to a fine powder in liquid nitrogen, resuspended in PBS, and injected into rabbits (Hazleton Research Products Inc., Denver, PA).

DNA and RNA Blots

For each lane of genomic Southern blots, 10 μg of MSB-1 chromosomal DNA was digested with a restriction enzyme. DNA was separated on a 0.8% agarose gel and transferred to BA83 nitrocellulose (Schleicher & Schuell, Inc., Keene, NH) according to the manufacturer's instructions. For each Northern blot lane, 1 μg of poly(A)$^+$ RNA was run on a formaldehyde–1% agarose gel, then transferred to nitrocellulose (Ausubel et al., 1991).

Probe sequences were isolated from agarose gels and labeled with [^{32}P]dCTP (Amersham Corp., Arlington Heights, IL) by random priming (Feinberg and Vogelstein, 1983, 1984). Probes were hybridized to filters in 5× SSC, 0.5% SDS, and 100 μg/ml salmon sperm DNA at 65° overnight. The filters were then washed twice for 10 min with 2× SSC and 0.1% SDS, followed by a 1-h wash with 0.3× SSC and 0.1% SDS at 65° (Ausubel et al., 1991).

Subcellular Fractionation

Dounce Homogenization. Exponentially growing cells ("interphase cells") or cells grown in the presence of 0.1 μg/ml colcemid ("mitotic cells") were collected by centrifugation at room temperature. Cells were resuspended and swollen in 1 × RSB (Lewis and Laemmli, 1982) at room temperature for 5 min, and centrifuged for 3 min at 800 g. Subsequent steps were performed on ice. The pellet was vigorously resuspended in 1× buffer A (15 mM Tris/HCl, pH 7.4, 80 mM KCl, 2 mM KEDTA, pH 7.4, 0.75 mM spermidine, 0.3 mM spermine, and 0.1% digitonin [Sigma Immunochemicals] [Lewis and Laemmli, 1982]), and cells were lysed with 10 strokes of a Dounce homogenizer. The lysate was centrifuged at 3,200 *g* for 20 min to separate the supernatant ("cytoplasm") from the pellet ("nuclei" or "chromosomes"). To reduce cross-contamination, the supernatant was centrifuged again at 3,200 *g* for 20 min, and the pellet was discarded. A small portion of the supernatant was observed under the microscope to confirm that nuclei and chromosomes were absent. The pellet was washed with solution 3 (5 mM Tris:HCl, pH 7.4, 2 mM KCl, 2 mM KEDTA, pH 7.4, 0.375 mM spermidine). Further fractionation of chromosomes to scaffold and nonscaffold fractions was performed as described previously (Cooke et al., 1987; Earnshaw et al., 1985). Total (lysate without centrifugation), cytoplasm, and nuclei or chromosome fractions were then mixed with SDS-PAGE sample buffer, boiled, sonicated, and applied to SDS-PAGE. Electrophoresis and immunoblotting were performed as described previously (Earnshaw et al., 1984).

Mass Enucleation. Preparation of cytoplasts and karyoplasts was performed as described in (Prescott and Kirkpatrick, 1973) with slight modifications. Briefly, cells were grown on round plastic coverslips (PGC Scientific, Gaithersburg, MD), and were pre-centrifuged in warm RPMI 1640 (Gibco Laboratories, Grand Island, NY) cell side down at 37° for 10 min at 5,000 *g* to remove dead cells. The coverslips were then transferred to centrifuged tubes containing warm RPMI plus 10 μg/ml cytochalasin B (Sigma Immunochemicals), and centrifuged for 10 min at 17,000 *g* at 37°. Karyoplasts in the pellet and cytoplasts still attached to the coverslip were dissolved in SDS-PAGE sample buffer and used for gel analysis. The cytoplasts were also observed in the microscope after fixation with 3% paraformaldehyde and staining with DAPI. More than 95% of the cells remaining attached to the coverslips had been enucleated.

Phylogenetic Analysis

Assembling the Walker B Profile. The region at the COOH terminus of ScII

1. *Abbreviations used in this paper*: DAPI, 4',6' diamino phenylindole; ORF, open reading frame; RACE, rapid amplification of cDNA ends.

that is most conserved between ScII, Smc2, Smc1, and P115 was used to construct a position-specific scoring table (profile) using the PILEUP program from the GCG package of sequence alignment programs (Devereux et al., 1984). This profile was used to search the entire coding region of representatives from the ABC transporter family and from the RNA helicase family using PROFILEGAP. In addition, using PROFILESEARCH, we searched the entire SwissProt data base to find B motifs more closely related to the ScII group. The highest scoring match for a known ATPase was for UvrA, with the recombination gene RecN being the most closely related protein overall. In each case, the best match to the profile was a conserved region homologous to the Walker B motif, which is well characterized for ABC transporters (Higgins et al., 1986) and for RNA helicases (Linder et al., 1989). Representative sequences from each group were added to the ScII profile using PILEUP to generate a profile with more divergent B motifs, the Walker B profile. To gain some understanding of the possible role of the Walker B motif in ATP binding or hydrolysis, the Walker B profile was aligned to the sequence for rabbit muscle adenylate kinase, an ATPase for which detailed structural information is available. The best match to the Walker B profile contained the entire so-called segment 3 (Fry et al., 1986), which is conserved in other ATPases and contains what has been implicated as the catalytic residue, Asp-119 (Asp-1120 in ScII).

Constructing the Phylogenetic Tree. Distances between sequences in the Walker B profile were computed using the DISTANCES program from the GCG package. The resultant matrix of distances was used to construct a phylogenetic tree using the Fitch-Margoliash algorithm as implemented in the program FITCH from the PHYLIP phylogeny inference package (version 3.52c) (Felsenstein, 1989). The order in which sequences were added to the tree was randomized five times, and the best tree of 7,028 examined is reported. The tree as presented was rooted at its centroid. The sequences of the Walker B sites used in the phylogenetic tree are given in Table I, together with the appropriate references.

Results

Generation of Antibody Probes for ScII

ScII was first described by Lewis and Laemmli as a 135-kD polypeptide component of mitotic chromosome scaffolds from HeLa cells (Lewis and Laemmli, 1982). This polypeptide was not characterized further. In a subsequent study of chicken mitotic chromosome scaffold proteins (Earnshaw et al., 1985), we isolated a prominent 135-kD band by preparative SDS-PAGE and injected it into a guinea pig. The resultant polyclonal serum (5B2) specifically reacted with a 135-kD polypeptide in immunoblots of isolated chicken and human mitotic chromosomes (not shown). Thus, adopting the nomenclature established for HeLa scaffold proteins, we designated this antigen chicken ScII. We note, however, that in the absence of specific probes for the HeLa protein seen originally by Lewis and Laemmli, we cannot be certain that the polypeptide we now designate ScII exactly corresponds to the original ScII scaffold protein. To begin the functional analysis of ScII, we have cloned a cDNA encoding the polypeptide, and we have determined the distribution of the protein in cells and mitotic chromosomes.

cDNA Cloning of Chicken ScII

The process of obtaining cDNA clones encoding the entire

Table I. Sequences Examined as NTP-binding B Sites and Used to Establish Phylogenetic Tree in Fig. 13

Protein	Amino acid sequences including B site	Notes	Reference
ScII	LSGGQRSLAA LSLILAILLF KPAPIYILDE VDAALD	Members of	This study
Smc2p	LSGGQRSLIA LSLIMALLQF RPAPMYILDE VDAALD	SMC1 family	Strunnikov (unpublished)
Smc1p	LSGGEKTVAA LALLFAINSY QPSPFFVLDE VDAALD		Strunnikov et al. (1993)
P115	FSGGEKAIIA ISLLFAILKA RPIPLCILDE VEAALD		Notarnicola et al. (1991)
Bs-RecN	ASGGELSRVM LAIKSIFSSQ QDVTSIIFDE VDTGVS	DNA repair-	Van Hoy and Hoch (1990)
Ec-RecN	ASGGELSRIA LAIQVITARK METPALIFDE VDVGIS	recombination	Rostas et al. (1987)
UvrA B2	LSGGEAQRIR LASQIGA-GL VGV-MYVLDE PSIGLH	proteins	Husain et al. (1986)
UvrA B3	LSGGEAQRVK LARELSKRGT GQT-LYILDE PTTGLH		Husain et al. (1986)
NodI	LSGGMKRRLT LAGALIND-- -PQ-LLILDE PTTGLD	ABC transporters	Evans and Downie (1986)
OppD	FSGGMRQRVM IAMALLCR-- -PK-LLIADE PTTALD		Higgins et al. (1985)
HlyB	LSGGQRQRIA IARALVNN-- -PK-ILIFDE ATSALD		Felmlee et al. (1985)
OppF	FSGGQCQRIG IARALILE-- -PK-LIICDD AVSALD		Hiles et al. (1987)
HisP	LSGGQQQRVS IARALAME-- -PD-VLLFDE PTSALD		Higgins et al. (1982)
PstB	LSGGQQQRLC IARGIAIR-- -PE-VLLLDE PCSALD		Surin et al. (1985)
FtsE	LSGGEQQRVG IARAVVNK-- -PA-VLLADE PTGNLD		Gill et al. (1986)
RbsA	LSIGDQQMVE IAKVLSF--- -ESKVIIMDE PTDALT		Bell et al. (1986)
MalK	LSGGQRQRVA IGRTLVAE-- -PS-VFLLDE PLSNLD		Gilson et al. (1982)
eIF-4AI	IVGTPGRVFD MLNR-RYLSP KYIKMFVLDE ADEMLS	RNA helicases	Nielsen et al. (1985)
eIF-4AII	VVGTPGRVFD MLNR-RYLSP KWIKMFVLDE ADEMLS		Nielsen and Trachsel (1988)
Tif1	VVGTPGRVFD NIQR-RRFRT DKIKMFILDE ADEMLS		Linder and Slonimski (1989)
p68	CIATPGRLID FLEC-GKTNL RRTTYLVLDE ADRMLD		Ford et al. (1988)
PL10	LVATPGRLVD MMER-GKIGL DFCKYLVLDE ADRMLD		Leroy et al. (1989)
vasa	VIATPGRLLD FVDR-TFITF EDTRFVVLDE ADRMLD		Lasko and Ashburner (1988)
MSS116	VIATPGRLID VLEKYSNKFF RFVDYKVLDE ADRLLE		Seraphin et al. (1989)
SrmB	VVATTGRLLQ YIKE-ENFDC RAVETLILDE ADRMLD		Nishi et al. (1988)
Adn kin Sg3	LIDGYPRQVQ QGEEFERRIA QPTLLLYVDA GPETMQ	Kinase	Kuby et al. (1984)

open reading frame of ScII was carried out in three stages. First, a chicken cDNA λgt11 expression library (Young and Davis, 1983) was screened with guinea pig anti-ScII. Several positive clones were characterized, and were found to cover a portion of the ScII cDNA. In the second round of screening, the inserts from the these phage were used for further DNA screening of a chicken MSB-1 λZAP library (Mackay et al., 1993) by plaque hybridization. Eight positive clones were obtained, the longest of which was designated pBSc20 (insert size = 3.8 kb) (Fig. 1 *A*). Primer extension analysis (data not shown) indicated that pBSc20 was lacking ~400 bp from its 5′ end relative to the 5′ end of the ScII mRNA. Therefore, in a third round of cloning, these missing sequences were recovered by the RACE-PCR method (Frohman et al., 1988). Four clones from two independent RACE reactions with ScII mRNA were sequenced and turned out to be identical to each other and contiguous with the 5′ end of pBSc20. These were used to complete the molecular characterization of the 5′ end of the ScII cDNA. pBSc20 was fully sequenced on both strands, subcloned for bacterial expression, and used for a probe in Southern (DNA) and Northern (RNA) hybridization analysis.

The 4,134-bp ScII cDNA encodes a single undisrupted ORF encoding a predicted polypeptide of 1,189 amino acids, with a calculated molecular mass of 134,900 D (Fig. 2 *A*). This corresponds to the apparent molecular mass of ScII previously determined from SDS-PAGE (135 kD) (Lewis and Laemmli, 1982). The ScII polypeptide is predicted to be slightly basic (calculated pI = 8.6) and highly hydrophilic (Fig. 2 *B*). At the 3′ end of the cDNA, a polyadenylation signal (AATAAA) was found 18 bp upstream from the poly A sequence.

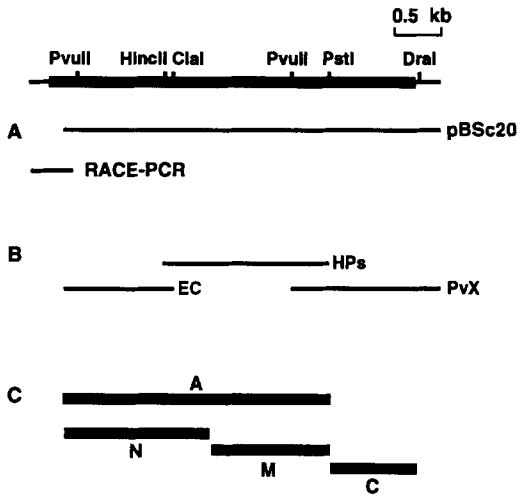


Figure 1. cDNA cloning and bacterial expression of ScII. The uppermost line shows the structure of the ScII cDNA with several restriction enzyme sites. The thick line indicates a continuous ORF. (*A*) Clones that were fully sequenced and used for further experiments. pBSc20 is the longest clone of eight final positive clones that were obtained by immuno and DNA screening of chicken cDNA libraries. RACE-PCR indicates four clones derived from two independent RACE-PCR reactions. (*B*) Probes used for a genomic Southern blot and RNA blots. HPs (HincII-PstI DNA fragment) was used for the genomic Southern blot in Fig. 4*A*. EC (EcoRI-ClaI DNA fragment) and PvX (PvuII-XhoI DNA fragment), respectively, were used as 5′ or 3′ probes for the RNA blots in Fig 4 *B*. The EcoRI and XhoI sites of EC and PvX, respectively, are derived from linker sequences of pBsc20. (*C*) Subregions that were expressed in *E. coli*. The peptides A (aa 120-970), N (aa 120-567), M (aa 568-969), and C (aa 968-1,189) were expressed as fusion proteins to pT7-7 vector–derived sequences.

Validation of the Clones

We obtained evidence that pBSc20 encodes bona fide ScII by demonstrating that the cloned and chromosomal polypeptides share at least three independent epitopes. Several subcloned regions of pBSc20 (shown in Fig. 1 *C*) were expressed in *E. coli* using a T7 RNA polymerase–based expression system (Tabor and Richardson, 1985). Peptides N, M, and C were designed to cover the $\underline{NH_2}$-terminal, middle and COOH-terminal regions, respectively. These polypeptides do not overlap one another. Between them, these molecules cover the entire ScII molecule, except for 120 amino acids at the NH_2 terminus. One further peptide, A, covers a region equivalent to the sum of N and M.

Bacterially expressed peptides N, M, and C produced prominent bands in SDS-PAGE of induced bacterial extracts, all of which migrated with approximately the expected molecular weights. Peptide A was somewhat degraded in vivo. Peptides A, N, and M reacted with the original guinea pig anti-ScII antibody in immunoblots; however, peptide C did not (data not shown). All four expressed proteins were isolated by SDS-PAGE and injected into rabbits for production of polyclonal antisera.

Despite the fact that antibodies N, M, and C are directed against epitopes on nonoverlapping peptides, all of them recognized a 135-kD band in the chromosome protein fraction by immunoblot. This immunoreactive polypeptide comigrated with ScII as detected with the original guinea pig antibody, 5B2 (Fig. 3). Thus, our cDNA clones encode multiple independent epitopes present on ScII, suggesting that pBSc20 is a bona fide cDNA clone for ScII.

ScII is Encoded by a Single Gene and Transcribed into a 4.2 kb mRNA

When chicken genomic DNA was digested with restriction endonucleases, electrophoresed and transferred to nitrocellulose, and hybridized with various subclones derived from pBSc20, only single or double bands were observed. This suggests that ScII is encoded by a single gene (Fig. 4 *A*).

Northern (RNA) blots of poly(A)$^+$ mRNA revealed a single 4.2-kb transcript with probes derived from either the 5′ or 3′ portion of pBSc20 (Fig. 4 *B*). This mRNA appears to be reasonably abundant, at least in the chicken MSB-1 lymphoblastoid cell line. Interestingly, lower levels of transcript were detected in a parallel analysis of mRNA from chick embryo fibroblasts (Fig. 4 *C*).

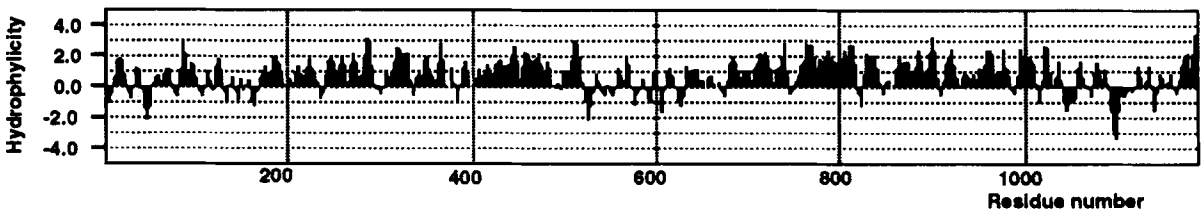
ScII is a Member of an Emerging Family of Putative Nucleotide-binding Proteins Defined by Yeast Protein Smclp

A computer database search with the ScII cDNA sequence revealed limited sequence similarities to several other proteins, including Smclp. This yeast protein (stability of minichromosomes) was recently shown to be required for proper chromosomal segregation in budding yeast. The *SMC1* gene product is the founder member of a group of proteins that

A

```
                                                                                                                    AAACGTTAGTTGTGACGAGCATGCGAGAGGG -118
GTTCGGTGGCTGCTGGCGGTTCACTGAGGGCTCCCTTCGAGCTCGAAGAGTTACAGTCCAGTCCTTTCCTTGTACTGCTGGGGGAAGCACCAACGTGGCCCCCTGACAGATTAAAGAAG
ATG TAC ATC AAA TCA ATT GTA CTT GAA GGA TTT AAA TCC TAT GCT CAG AGG ACA GAA ATC CGT GAT TTT GAT CCG TTA TTC AAT GCC ATT   90
 M  Y  I  K  S  I  V  L  E  G  F  K  S  Y  A  Q  R  T  E  I  R  D  F  D  P  L  F  N  A  I   30
ACC GGC TTG AAT GGT AGT GGC AAG TCC AAC ATC TTG GAC TCC ATC TGC TTC CTG TTG GGC ATC TCC AAT CTA TCT CAG GTG CGA GCT TCA  180
 T  G  L  N  G  S  G  K  S  N  I  L  D  S  I  C  F  L  L  G  I  S  N  L  S  Q  V  R  A  S   60
AGC CTC CAA GAT TTA GTT TAT AAA AAT GGG CAA GCT GGA GTG AAT AAA GCA ACT GTG TCT ATC ACG TTT GAT AAT TCC GAC AAG AAA AAT  270
 S  L  Q  D  L  V  Y  K  N  G  Q  A  G  V  N  K  A  T  V  S  I  T  F  D  N  S  D  K  K  N   90
AGT CCA CTG GGA TTT GAA AAT AAT GAT GAG ATC ACC ATC ACC AGG CAG GTT ATT GTT GGA GGT AGA AAT AAG TAT CTT ATC AAC GGT ATG  360
 S  P  L  G  F  E  N  N  D  E  I  T  I  T  R  Q  V  I  V  G  G  R  N  K  Y  L  I  N  G  M  120
AAT GCT TCC AAC AAC CGT GTA CAG GAT CTC TTT GGT TCT GTT GGA CTC AAT GTC AAT AAC CCT CAC TTC CTT ATT ATG CAG GGG CAA ATT  450
 N  A  S  N  N  R  V  Q  D  L  F  G  S  V  G  L  N  V  N  N  P  H  F  L  I  M  Q  G  Q  I  150
ACT AAA GTT CTA AAT ATG AAG CCG ACA GAG ATT CTA GCT ATG ATT GAA GAA GCA GCT GGT ACT AGG ATG TAT GAA TGC AAG AAA ATA ACT  540
 T  K  V  L  N  M  K  P  T  E  I  L  A  M  I  E  E  A  A  G  T  R  M  Y  E  C  K  K  I  T  180
GCC CAT AAG ACC ATA GAA AAG AAG GAA AGC AAA CTG GAT GAA ATT CGA AGG ATC ATA ACT GAA GAG ATC AGT CCA ACT CTA GAG AAA CTG  630
 A  H  K  T  I  E  K  K  E  S  K  L  D  E  I  R  R  I  I  T  E  E  I  S  P  T  L  E  K  L  210
AAA GAG GCA CGT GCA TCG TAT CTG GAG TAT CAA AAA ATG ACA CGA GAA GTG GAG AAT TTA AGG CGC ATC TAT GTG GCT TTT CAG TAT GTT  720
 K  E  A  R  A  S  Y  L  E  Y  Q  K  M  T  R  E  V  E  N  L  R  R  I  Y  V  A  F  Q  Y  V  240
CGT GCT GAA GAG ATA AAA GAT CGC TCT ACT AAC GCT TTA AAG GAA GCA CAG GCA AAT AAG AAG AAG ATT TTT GAA TCA ATG GCT GAG AAT  810
 R  A  E  E  I  K  D  R  S  T  N  A  L  K  E  A  Q  A  N  K  K  K  I  F  E  S  M  A  E  N  270
GAG AAG AAG GTA AAG GAA CTT GCG CAA CAA ATA GAA GAA ACA GAA AAG AAA AAC AAT GAG GAA TTT GGT GCT AAG TTG CAT TCT CTG GAA  900
 E  K  K  V  K  E  L  A  Q  Q  I  E  E  T  E  K  K  N  N  E  E  F  G  A  K  L  H  S  L  E  300
GCT GCT TTT TCA GAA CTC CAA AGA GTT GAT GCT AAA GTC CGC AGT GAT CTT GAT CAT AGG AAG CAA AAC TTA AAC AGT GAA GAA AAT AGG  990
 A  A  F  S  E  L  Q  R  V  D  A  K  V  R  S  D  L  D  H  R  K  Q  N  L  N  S  E  E  N  R  330
CTT AAA GAA TTG ATT AAA ATC ATG CAG GAG GAA TTT AAA GCA TTC ACA TCA AAG GAA AAG GAA ATA AAG AAG ATA AAG GAA GGA CTA AAT 1080
 L  K  E  L  I  K  I  M  Q  E  E  F  K  A  F  T  S  K  E  K  E  I  K  K  I  K  E  G  L  N  360
GGT TTA CAG GAG GAA AGT AAA AAA GAT GCA GAA GCT TTA GCC TCC GCT CAA CAG CAT TTT AAT GCT GTA TCT GCT GGC CTG TCT AGC AAT 1170
 G  L  Q  E  E  S  K  K  D  A  E  A  L  A  S  A  Q  Q  H  F  N  A  V  S  A  G  L  S  S  N  390
GAC AGT GGT CAA GGA ACT TCT CTT GCT GAC CAG ATG ATG ACC TGC AAA AAT GAG ATT AGC AAA GCA GCA ACA GAG GCT AAA CAG GCC CAA 1260
 D  S  G  Q  G  T  S  L  A  D  Q  M  M  T  C  K  N  E  I  S  K  A  A  T  E  A  K  Q  A  Q  420
ATG AAG TTG AAA TAT GCA CAG CAG GAG CTA AAA ACA AAA CAA GCT GAA GTT AAA AAG ATG GAT GGA AGC TAC AAG GAA GAC CAA GAG GCA 1350
 M  K  L  K  Y  A  Q  Q  S  L  K  T  K  Q  A  E  V  K  K  M  D  G  S  Y  K  E  D  Q  E  A  450
TTT GAA GCT ATC AGA AAA ACG AAA GAA AAA CTA CAG GAT GAA ATG AAA AAG CTG AAA TAT GAA GAA GCA GAA CAA GAA GCC CAT CTA GCA 1440
 F  E  A  I  R  K  T  K  E  K  L  Q  D  E  M  K  K  L  K  Y  E  E  A  E  Q  E  A  H  L  A  480
AAA AAG AAG CAA TTG AGT TCT GAG ATC AGT AGT CTG AGA GAA CTG TGT GAA AGT ATA GAA GCA AAG CAT CCT TAT CTA CGG TTT GAA TAC 1530
 K  K  K  Q  L  S  S  E  I  S  S  L  R  E  L  C  E  S  I  E  A  K  H  P  Y  L  R  F  E  Y  510
AAA AAT CCA GAA AAA AAT TGG AAC CCC AAC TGT GTG AAA GGC CTT GTT GTA ACT CTT ATC ACC GTG AAA GAC ATA TCC ACA TCA AAA GCT 1620
 K  N  P  E  K  N  W  N  P  N  C  V  K  G  L  V  V  T  L  I  T  V  K  D  I  S  T  S  K  A  540
CTA GAA GCA GTG GCT GGG GGA AAA CTC TAT AAC ATT GTT GTG GAC ACA GAG GCT ACT GGC AAA AAG ATT TTA GAA AAG GGT CAA CTA AAG 1710
 L  E  A  V  A  G  G  K  L  Y  N  I  V  V  D  T  E  A  T  G  K  K  I  L  E  K  G  Q  L  K  570
CAT CGA TAC ACC ATC ATT CCA CTA AGC AAA ATT TCA GCC AAC TCT ATT GGA CAT GAA ATT ATC TCA CTG GCC AAA AAC CTG ATT GGT CAT 1800
 H  R  Y  T  I  I  P  L  S  K  I  S  A  N  S  I  G  H  E  I  I  S  L  A  K  N  L  I  G  H  600
CGT GAA GTG CAT ATA GCC ATT TCT CTT ATT GAC TAT AAT TCT GAA CTA CAG AAG GCA ATG GAA TAT GTC TTT GGG ACA ACA CTG GTC TGT 1890
 R  E  V  H  I  A  I  S  L  I  D  Y  N  S  E  L  Q  K  A  M  E  Y  V  F  G  T  T  L  V  C  630
AGC AGC ATG GAT AAT GCT AAG AAA GTG ACC TTT GAC AAA AGA ATA ATG AGA AAA ACT GTT ACA CTT CAA GGA GAT ATA TTT GAC CCC CAG 1980
 S  S  M  D  N  A  K  K  V  T  F  D  K  R  I  M  R  K  T  V  T  L  Q  G  D  I  F  D  P  Q  660
GGA ACC CTT AGT GGT GGT GCA AGC TCA CAT GTT ACA CCT ATA TTG TCT AAA CTT AAA ACA ATG AGA GAT GCT GAA GAT GAA CTC AAA ATA 2070
 G  T  L  S  G  G  A  S  S  H  V  T  P  I  L  S  K  L  K  T  M  R  D  A  E  D  E  L  K  I  690
AAG ACG TCT CAG CTT GAA GCT ACA GAG AAA GAG CTA GCA AAC TTG AAG AAT ATG GCT GAA AAG TAC CAG CAT CTG AAG CAG CAG TGG GAG 2160
 K  T  S  Q  L  E  A  T  E  K  E  L  A  N  L  K  N  M  A  E  K  Y  Q  H  L  K  Q  Q  W  E  720
ATG AAG TCT GAA GAG GCA GAG CTA CTA CAA ACT AAG ATT CAG CAA AGC GCT TAT CAC AAA CAA CAG GAA GAC CTG CTT GCT CTG AAG AAA 2250
 M  K  S  E  E  A  E  L  L  Q  T  K  I  Q  Q  S  A  Y  H  K  Q  Q  E  D  L  L  A  L  K  K  750
ACT ATT GCG GAG TGT GAA GAG ACA TTG AAG AAA ACT GAA GAG AGC CAA CGG AAA GCT GAA GAA GAA TAC AAG GCA TTG GAG AAT AAA ATG 2340
 T  I  A  E  C  E  E  T  L  K  K  T  E  E  S  Q  R  K  A  E  E  E  Y  K  A  L  E  N  K  M  780
AAA AAT GCA GAA GCA GAA CGT GGG AAG GAA ATA AAA AAT GCC CAG CAG AAA CTA AAT AGC GCC AAG AAG AAA GCA GAC GAT TCA AGC AGA 2430
 K  N  A  E  A  E  R  G  K  E  I  K  N  A  Q  Q  K  L  N  S  A  K  K  K  A  D  D  S  S  R  810
AAA ATG AAA GAA AAG CAG CAA GAA GTT GAA GCA TTA GTT CTG GAA CTT GAA CAG CTG AAA CAA GAA CAA GCC TCA TAT AAG CAA CAG AGT 2520
 K  M  K  E  K  Q  Q  E  V  E  A  L  V  L  E  L  E  Q  L  K  Q  E  Q  A  S  Y  K  Q  Q  S  840
GAG GCA GCA CAA CAA GCA ATC GCA TCC TTA AAG GAG CAA GTC AGT GCT TTG GAA GCT GAG GCA GTG AAG ACA AGG GAA TCT CTA AAG AAC 2610
 E  A  A  Q  Q  A  I  A  S  L  K  E  Q  V  S  A  L  E  A  E  A  V  K  T  R  E  S  L  K  N  870
GCA GAG AAT GAG CTG TCC AGT GAA AAG GGA TTA ATG GAA GAA CGA ACT AAA GAT ATT AAA GCC AAA TCT GCA AAG ATA GAG AAA TAC AGA 2700
 A  E  N  E  L  S  S  E  K  G  L  M  E  E  R  T  K  D  I  K  A  K  S  A  K  I  E  K  Y  R  900
GAG CAA AAT AAT GAA TTG CAG CTC AGC ATT AAT GCA TTA GAA CAT GAC ATC AAC AAG TAT CAA CAA GAG ACT GCT GAT GCT TCT TCC ACG 2790
 E  Q  N  N  E  L  Q  L  S  I  N  A  L  E  H  D  I  N  K  Y  Q  Q  E  T  A  D  A  S  S  T  930
TTG GAC AAA CTG CTG AAA GAG TAT AAG TGG ATA GCT TCA GAA AAA GAG CTT TTT GGC CAG GCA GAC ACA ACT TAC GAC TTT GAA GCC AAC 2880
 L  D  K  L  L  K  E  Y  K  W  I  A  S  E  K  E  L  F  G  Q  A  D  T  T  Y  D  F  E  A  N  960
AAT CCT AAG GAA ACA GGT CAG AAG CTG CAG AAA CTG CTG ACA AAA AAA GAG AAG CTG GAG AAG TCC TTG AAC ATG AGA GCT ATG AAT TTG 2970
 N  P  K  E  T  G  Q  K  L  Q  K  L  L  T  K  K  E  K  L  E  K  S  L  N  M  R  A  M  N  L  990
CTT TCT GAG GCA GAG GAG AGG TAT AAC GAC TTA ATG AAG AAA AAG AGA ATG GTT GAG AAT GAC AAG ATA AAA ATT CTT GCA ACT ATT GAA 3060
 L  S  E  A  E  E  R  Y  N  D  L  M  K  K  K  R  M  V  E  N  D  K  I  K  I  L  A  T  I  E  1020
GAG CTC GAC CGA AAG AAA AAT AAA GCT TTA CAC ATT GCT TGG GAA AAG GTG AAC AAA GAC TTT GGT TCC ATC TTC TCG ATG CTT CTA CCA 3150
 E  L  D  R  K  K  N  K  A  L  H  I  A  W  E  K  V  N  K  D  F  G  S  I  F  S  M  L  L  P  1050
GGA GCC AAA GCT ATG TTG GTT CCA TCT AAA AAG CAA AAT ATC CTG GAT GGT CTG GAG TTC AGG GTT GGC TTA GGA GAC ATC TGG AAG GAG 3240
 G  A  K  A  M  L  V  P  S  K  K  Q  N  I  L  D  G  L  E  F  R  V  G  L  G  D  I  W  K  E  1080
AAC TTA ACA GAA CTT AGT GGA GGT CAA AGA TCA TTG GCG GCC TTG TCC TTG ATC TTA GCC ATC CTT CTC TTC AAA CCT GCT CCA ATT AAC TAC 3330
 N  L  T  E  L  S  G  G  Q  R  S  L  A  A  L  S  L  I  L  A  I  L  L  F  K  P  A  P  I  Y  1110
ATC TTG GAT GAA GTG GAT GCC GCT CTT GAC CTT TCT CAT ACC CAA AAT ATT GGA CAG ATG CTT CAT GCT CAT TTC AAA CAA TCC CAG TTC 3420
 I  L  D  E  V  D  A  A  L  D  L  S  H  T  Q  N  I  G  Q  M  L  H  A  H  F  K  Q  S  Q  F  1140
CTT GTG GTG TCC CTG AAG GAC GGA ATG TTT AAC AAC GCC AAT GTT CTC TAC AGG ACC AAG TTT GTT GAT GGC ATT TCC ACA GTT TCA AGA 3510
 L  V  V  S  L  K  D  G  M  F  N  N  A  N  V  L  Y  R  T  K  F  V  D  G  I  S  T  V  S  R  1170
CAC TGT CAG TTA AAA AAA AAG CAG CCA CTG AGT GAA GCC AGC AAC AAG GAT GAG TAA CAAACGATGCATAATGATGACTCAAACATTCTGTGCCCC 3609
 H  C  Q  L  K  K  K  Q  P  L  S  E  A  S  N  N  K  D  E  *                                                    1189
TTGATCAAAGGAAGCCCCAGAATCACCTTTTAAAAAAAAAATATATATATATACATATATATATTTTATCTTGATTCCTTTAAGCCATAGTTATAGTTTTTATCTTTTTTAATAA 3728
CTTAGTTCATGGCAGCCCACCGGAGCCATGCAGTGTGATGAAGCTTGCAGCATTTATGGCTTTCAGAAAACCTGCTGCAAACTAAGCTACTGCGTTCAGCTTTTGTTGGCAGTTAAAGT 3847
GATGTATGAACATGCAAGTTAGTGTAAGCTTTTATCTGTGTAACCTACTTATTTCTGAGATGTACAGTGGGCAGCAGTTTAACTGCCAATTTTTAGTTGCTTGGTAATGGCTTTAATAA 3966
ATATTAGACAGAAAACTTTA(A)n
```

B



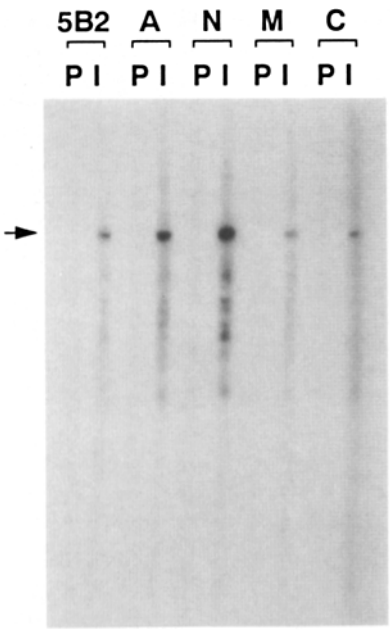


Figure 3. cDNA clones encode multiple independent epitopes present on ScII. Chromosomal proteins from MSB-1 cells were subjected to SDS-PAGE, blotted to nitrocellulose, cut into strips, and probed with antisera. *P*, preimmune serum; *I*, immune serum. *5B2*, Guinea pig antisera raised against the 135-kD band of the chicken chromosomal scaffold fraction. *A*, *N*, *M*, and *C*, Rabbit antisera raised against bacterially expressed peptides described in Fig. 1 *C*.

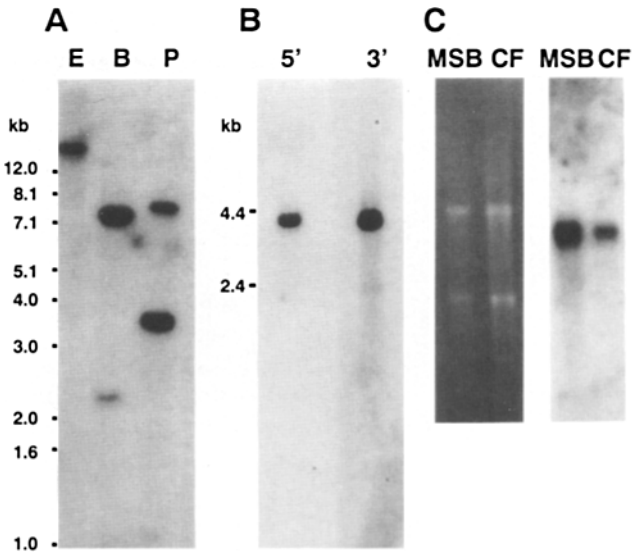


Figure 4. Chicken ScII is encoded by single gene and transcribed into a 4.2-kb mRNA. (*A*) Genomic southern blot. Total DNA from MSB-1 cells was digested with EcoRI (*E*), BamHI (*B*), and PstI (*P*), separated on an agarose gel, transferred to nitrocellulose, and probed with ^{32}P-labeled HPs fragment as described in Fig. 1 *B*. (*B*) RNA blot. 1 μg of poly(A)$^+$ RNA form MSB-1 cells was loaded in each lane. The blot was probed with either the 5′ (*EC*) or 3′ probe (*PvX*) as described in Fig. 1 *B*. (*C*) 1 μg of poly(A)$^+$ RNA from MSB-1 cells (*MSB-1*) or chicken embryo fibroblasts (*CF*) was separated on a denaturing agarose gel and stained with ethidium bromide (*left panel*), or blotted and probed with the 3′ probe (*right panel*).

share common amino acid sequence motifs at their amino and carboxy termini. ScII most closely resembles one of these family members, Smc2p (required for chromosome condensation, Strunnikov, A. V., and D. Koshland, personal communication), with which it shares 36% overall amino acid sequence identity.

As shown in Fig. 5 *A*, ScII and Smc2p share 65% identity over a 210-amino acid stretch near the NH_2 terminus, and 66% identity over a 136-amino acid stretch near the COOH terminus. As shown in Fig. 6, the amino-terminal sequence of all *SMC1* family members contains a consensus NTP-binding motif (GXXXXGKS, the so-called Walker A site [Walker et al., 1982]). The carboxy-terminal sequence of *SMC1* was suggested to be a novel motif, and was termed the DA box (Strunnikov et al., 1993). As described below, our analysis suggests that this is one version of the NTP-binding B site (the so-called Walker B site [Walker et al., 1982]). The B site is implicated in ATP hydrolysis, where structural evidence is available (Serpersu et al., 1986), or thought to be required in the coupling of ATP hydrolysis to ATPase function (Pause and Sonenberg, 1992). A third characteristic motif found in GTP-binding proteins is not present in ScII.

All *SMC1* family members have two internal regions that are predicted to form α-helical coiled coil. These regions (of 256 and 352 amino acids, respectively, in ScII) show no sequence similarity between ScII and Smc2, but the region between coils is 47% identical over 122 amino acids (Fig. 5, *B* and *C*).

ScII Localizes Predominantly in Interphase Cell Nuclei, but Leaks Out Readily during Subcellular Fractionation

To determine the intracellular localization of ScII in interphase cells, we subjected cultured cells to a standard procedure used for cell fractionation. This method uses Dounce homogenization of hypotonically swollen cells (Materials and Methods), followed by centrifugation to separate the nuclei from the cytoplasm.

When chicken or human cells were fractionated according to this method, ScII appeared nearly quantitatively in the cytoplasmic fraction (Fig. 7 *A*, left panel, data for DU249 and HeLa cells not shown). This was a surprise, since we would have expected a chromosome scaffold protein to be located

Figure 2. (*A*) Nucleotide sequence of the ScII cDNA reveals a deduced ORF of 1,189 amino acids. The probable NTP-binding A and B sites as discussed in the text are underlined with thick lines. Two putative coiled-coil regions are underlined with dotted lines. The putative polyadenylation signal at 3,962-3,967 nt is underlined with a thin line. These sequence data are available from EMBL/GenBank/DDBJ under accession number X80792. (*B*) The chicken ScII protein is predicted to be hydrophilic along its entire length. This Kyte-Doolittle hydrophilicity plot (window = 7) was generated by MacVector software (IBI).

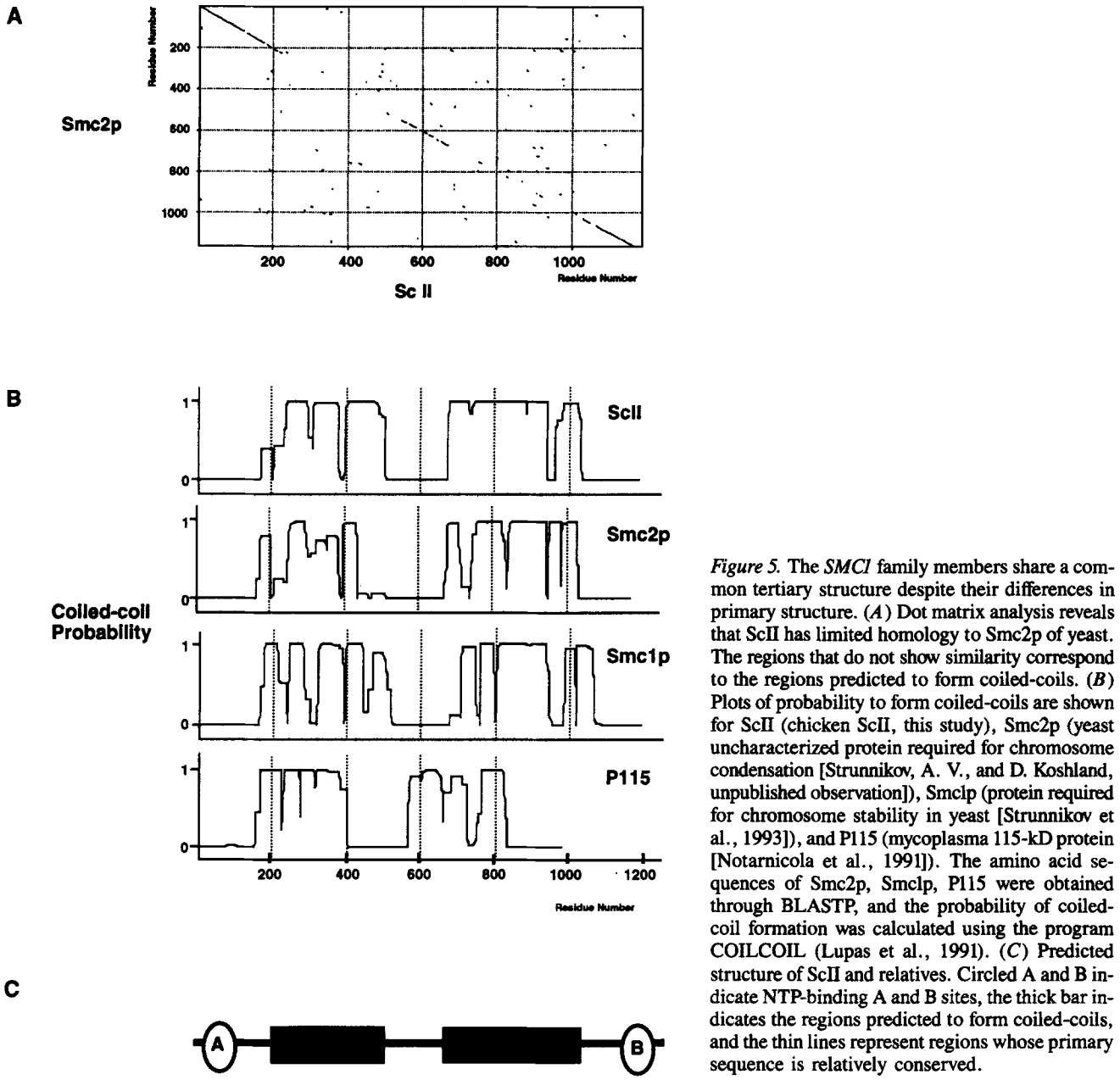


Figure 5. The *SMC1* family members share a common tertiary structure despite their differences in primary structure. (*A*) Dot matrix analysis reveals that ScII has limited homology to Smc2p of yeast. The regions that do not show similarity correspond to the regions predicted to form coiled-coils. (*B*) Plots of probability to form coiled-coils are shown for ScII (chicken ScII, this study), Smc2p (yeast uncharacterized protein required for chromosome condensation [Strunnikov, A. V., and D. Koshland, unpublished observation]), Smclp (protein required for chromosome stability in yeast [Strunnikov et al., 1993]), and P115 (mycoplasma 115-kD protein [Notarnicola et al., 1991]). The amino acid sequences of Smc2p, Smclp, P115 were obtained through BLASTP, and the probability of coiled-coil formation was calculated using the program COILCOIL (Lupas et al., 1991). (*C*) Predicted structure of ScII and relatives. Circled A and B indicate NTP-binding A and B sites, the thick bar indicates the regions predicted to form coiled-coils, and the thin lines represent regions whose primary sequence is relatively conserved.

in the nuclei, possibly as a component of the insoluble nuclear matrix. The result could not be explained by disruption of the nuclei during fractionation, since when the same preparations were probed with a control antibody recognizing topoisomerase II, this molecule was detected solely in nuclei (Fig. 7 *A*, right panel). A similar analysis of mitotic cells revealed that ~50% of the ScII is associated with the chromosomes (Fig. 8 *A*). Of this, ~70% remains as a component of th chromosome scaffolds after nuclease digestion and extraction of the chromosomes with 2 M NaCl (Fig. 8 *B*).

We considered two possible explanations for this result. (*a*) ScII might be a cytoplasmic protein in interphase cells that associates with chromosomes after nuclear envelope breakdown in mitosis. In this case, the different localizations of ScII during interphase and mitosis might reflect different molecular associations and possibly different functional roles. At one extreme, ScII could be a cytoplasmic protein that sticks to chromosomes at mitosis as a result of nonspecific interactions. (*b*) A second possibility was that ScII is actually a nuclear protein, but readily leaks out from the nuclei during Dounce homogenization. A number of well-characterized nuclear proteins such as DNA polymerase α and RNA polymerase I have been found to leak out from nuclei during mechanical cell lysis (Herrick et al., 1976).

To decide between these possiblities, we used another method of cell fractionation that eliminates leakage of proteins from the nucleus (Krek et al., 1992). This involves mass enucleation of adherent cells after exposure to cytochalasin B. In this method, coverslips with adherent cell monolayers were centrifuged cell side down at 37°C in medium containing cytochalasin B. Under these conditions, the cell cortex becomes much less robust, so that karyoplasts (nuclei surrounded by plasma membrane containing a thin

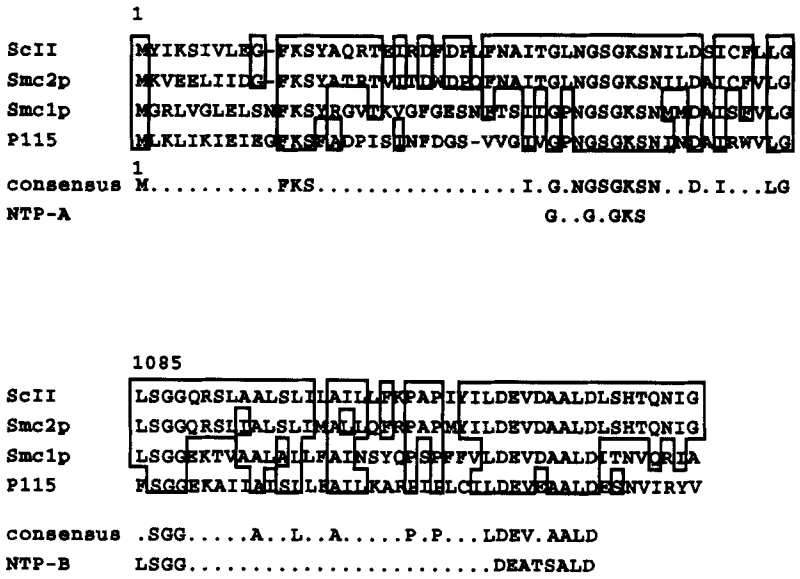


Figure 6. Probable NTP binding A and B sites were found in conserved regions of ScII near the NH_2 and COOH termini, respectively. Amino acid sequence alignments are shown for *SMCI* family members ScII, Smc2p, Smclp, and P115. (*Upper part*) Sequence alignment for the NH_2 terminus, including the NTP binding A site. (*Lower part*) Sequence alignment for the COOH terminus, including the putative NTP-binding B site. Numbers above the sequence correspond to amino acid numbering of ScII (Fig. 2 *A*). Identical amino acids are boxed. The consensus sequence derived from the *SMCI* family ("consensus") is shown together with consensus sequence for NTP-binding A and B sites of the ABC transporter family (*NTP-A* and *NTP-B*) (Fath and Kolter, 1993).

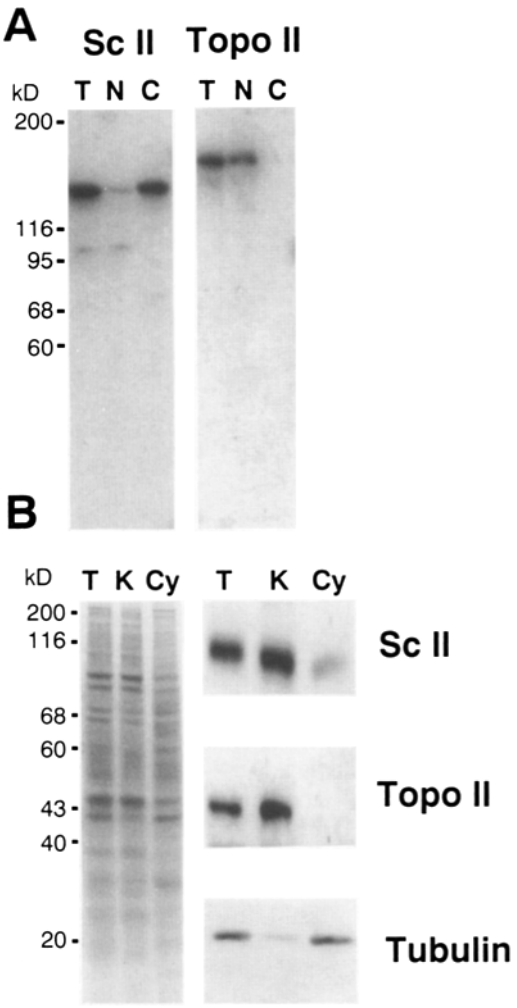


Figure 7. Determination of subcellular localization of ScII in interphase cells. (A) The majority of ScII is found in the cytoplasmic fraction when cells are fractionated by Dounce homogenization followed by centrifugation. Total (*T*), nuclear (*N*), and cytoplasmic (*C*) proteins of chicken cells were separated by SDS-PAGE, transferred to nitrocellulose, and probed with anti-ScII or anti-topoisomerase II antibodies. The amount of protein in each lane was normalized for the number of starting cells ($\sim 10^4$). Anti-topoisomerase II was used as a nuclear protein marker. (*B*) The majority of ScII is found in nuclei (karyoplasts) when cells are fractionated by mass enucleation. Total (*T*), karyoplast (*K*), and cytoplast (*Cy*) proteins of chicken cells were prepared as described in Materials and Methods. They were separated on SDS-PAGE, stained with Coomassie blue (*left panel*) or transferred to nitrocellulose, and probed with anti-ScII, anti-topoisomerase II, or anti-tubulin antibodies (a cytoplasmic protein marker).

layer of cytoplasm) sediment through it to the bottom of the tube. Cytoplasts (cell bodies lacking nuclei, but retaining all other cellular organelles) remain adherent to the coverslips (Prescott and Kirkpatrick, 1973). Since this process does not involve the generalized disruption of cellular membranes, leakage of nuclear components into the cytoplasts is minimized.

In contrast to the results obtained with Dounce homogenization, immunoblot analysis revealed that most of the ScII cosediments with karyoplasts. A similar result was obtained with topoisomerase II. In contrast, tubulin, used as a marker for cytoplasm, was enriched in the cytoplast fraction. These results imply that ScII is a nuclear protein that is not retained in nuclei during mechanical lysis and centrifugation (Fig. 7 *B*).

Although topoisomerase II and ScII were initially found to be in the chromosome scaffold fraction, these two proteins apparently associate with nuclear components in a different manner. This is surprising, since nuclear matrix and chromosome scaffold proteins are isolated by similar protocols, except for the starting material (i.e., the former use isolated nuclei, and the latter use isolated mitotic chromosomes). As a result, nuclear matrix and chromosomal scaffold proteins have traditionally been assumed to be equivalent. This generalization was supported by the fact that molecules such as topoisomerase II are found in both chromosome scaffolds and nuclear matrix (Berrios et al., 1985; Earnshaw et al.,

1985; Gasser et al., 1986). However, this notion seems not to apply to ScII, which is clearly not a component of the nuclear matrix, since it is not even a component of nuclei isolated by Dounce homogenization.

ScII Is Distributed along the Arms of Mitotic Chromosomes

ScII was localized in mitotic chromosomes using two different fixation and spreading protocols.

In the first, cells were hypotonically swollen, placed on coverslips whose surface had been rendered hydrophilic by glow discharging (Earnshaw and Migeon, 1985), and centrifuged briefly to burst them open. The coverslips were then transferred immediately to a buffer at physiological ionic strength, and were fixed with 3% paraformaldehyde. Chromosomes prepared in this way are frequently expanded, presumably as a result of shear forces during lysis of the cells. However, it is important to note that they have not been subjected to any chemical extraction. It is also important to note that these preparations were made with cells that had not been previously blocked with colcemid, thus minimizing the possibility of artifactual redistribution of the antigens in vivo (Compton et al., 1991).

Under these conditions, ScII was found to be distributed along the axis of the chromosome arms over their entire length, including the centromere (Fig. 9). The DNA was often considerably more diffuse in aspect, suggesting that chromatin loops were extended outward from the axial region containing ScII. These images look virtually identical to earlier images obtained using this technique in conjunction with antibodies to topoisomerase II (Earnshaw and Heck, 1985).

Condensed staining of centromere regions by anti-ScII was reproducibly observed by this technique. This may reflect a concentration of ScII in the heterochromatin of the centromere in vivo. Alternatively, it could simply reflect a differential tendency of the centromeric heterochromatin and chromosome arms to expand during centrifugation.

These images should not be interpreted as showing that ScII forms an axial core up the center of the chromosome arms, as in some versions of the chromosome scaffold model (Earnshaw, 1991). Instead, we suggest that ScII is associated with an insoluble network that is distributed throughout the chromatid arms. When the chromatin loops splay outwards during cell lysis, this network more or less retains its original shape. Thus it is unlikely that ScII is tightly associated with the bulk chromatin of the radial loops. Instead, is it apparently associated either with the bases of these loops, or with some other nonchromatin component of the chromosomes.

Our second procedure for visualizing ScII in mitotic chromosomes involves hypotonic swelling of cells, followed by fixation with methanol–acetic acid and spreading induced by blowing with a stream of air from an aquarium pump. This is a standard procedure for obtaining cytological spreads of chromosomes, which we have previously modified to preserve the reactivity of antigens with anticentromere antibodies (Earnshaw et al., 1989).

This method again reveals the presence of ScII along the entire length of the chromosome arms (Fig. 10). As in the formaldehyde-fixed cells, ScII appears to occupy a somewhat more restricted domain than the bulk chromatin, even

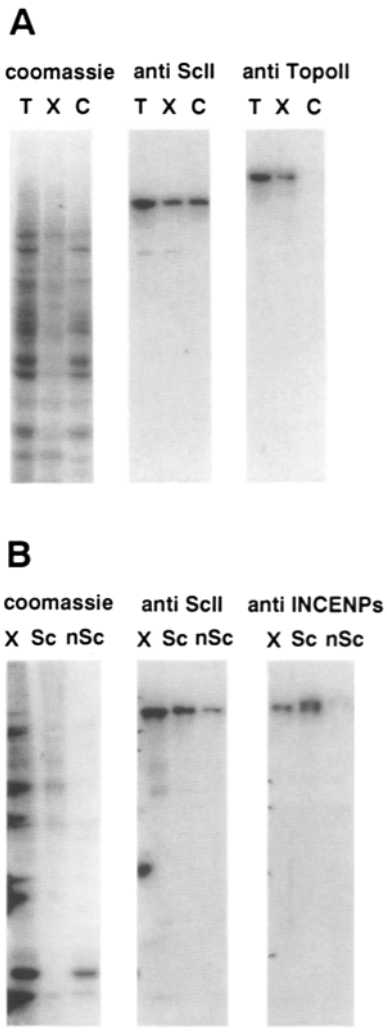


Figure 8. Determination of the sub-cellular localization of ScII in mitotic cells. (*A*) 50% of the total ScII is found in the chromosome fraction. Total (*T*), chromosomal (*X*), and cytoplasmic (*C*) proteins prepared from cells arrested at mitosis with colcemid were separated on SDS-PAGE, stained with Coomassie blue (*left panel*), or probed with anti-ScII or anti-topoisomerase II antibodies. (*B*) 70% of the chromosomal ScII is found in the chromosome scaffold fraction. Chromosomal proteins (*X*) were extracted with 2 M NaCl and fractionated to yield insoluble proteins (*Sc*, the chromosome scaffold) and soluble proteins (*nSc*, nonscaffold proteins). The proteins were separated on SDS-PAGE, stained with Coomassie blue, or blotted and probed with anti-ScII or anti-INCENP (a scaffold protein marker) antibodies.

though the latter is less swollen than in the case of spreading by centrifugation. Interestingly, HeLa chromosomes, which are larger than their chicken counterparts, often show a locally coiled morphology of ScII staining (Fig. 10 *D*) similar to that observed for topoisomerase II by Laemmli and coworkers (Boy de la Tour and Laemmli, 1988; Saitoh and Laemmli, 1994). Similar staining patterns were obtained using antibodies against antigens N, M, and C, or affinity-purified antibody against antigen A, suggesting strongly that the staining results from specific recognition of ScII. Thus, two independent methods of fixation and spreading confirm

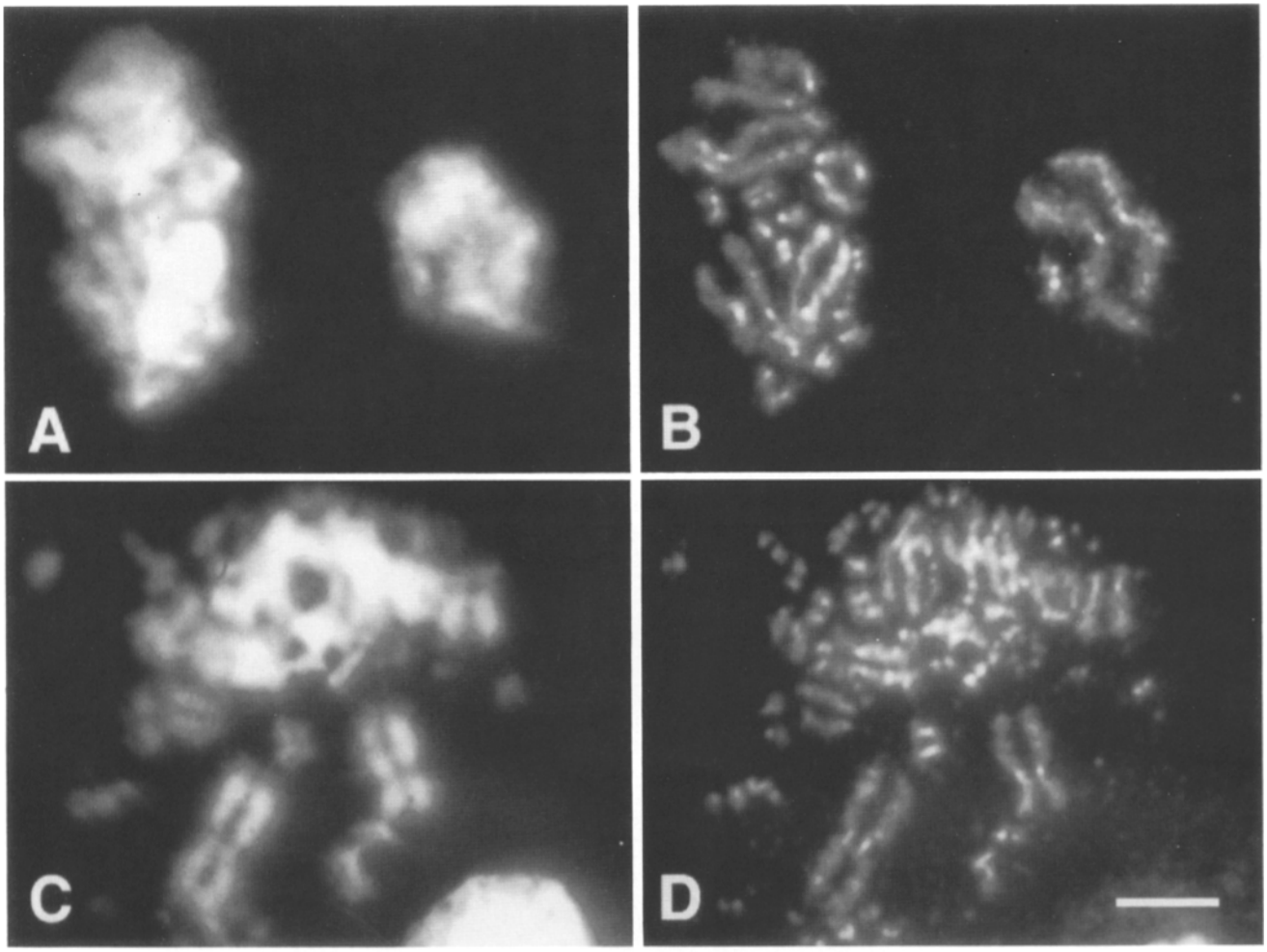


Figure 9. Chicken ScII localizes throughout the axial region of mitotic chromosome arms. Chicken MSB-1 cells were fixed with paraformaldehyde and stained with DAPI (*A* and *C*) or anti-ScII antibodies (*B* and *D*). Anti-ScII antibodies stained the interior of each chromatid arm. Coiled staining patterns are seen locally. Selective staining of centromere regions is often seen with this specimen preparation technique.

that ScII can be visualized along the entire length of the chromosome arms, with a distribution reminiscent of that seen previously with antibodies to topoisomerase II. To confirm colocalization of ScII and topoII, we tried double staining of chicken cells and indeed, we obtained an indistinguishable staining pattern of ScII and topoII on the mitotic chromosome (Fig. 11).

The localization of ScII to mitotic chromosomes was confirmed by immunofluorescence of cells processed in situ without hypotonic swelling or spreading. Unfortunately, with our present antibody reagents, we have been unable to demonstrate a convincing cellular localization of ScII during interphase under a variety of fixation conditions.

Discussion

ScII, the Second Most Abundant Component of the Mitotic Chromosome Scaffold, Is Not a Component of the Interphase Nuclear Matrix

ScII was originally identified as a component of the mitotic chromosome scaffold fraction (Lewis and Laemmli, 1982). [^{35}S]methionine autoradiography of HeLa cells suggests that ScII is second in abundance in this fraction only to topoisomerase II (Lewis and Laemmli, 1982). ScII is thus an excellent candidate for a protein that plays a major role in mitotic chromosome structure or function. As we will discuss below, ScII exists in at least one characterized complex with topoisomerase II, and it is therefore tempting to speculate that the protein might function in either chromosome condensation or disjunction, since both of these processes require topoisomerase II.

Unlike topoisomerase II, ScII is not a component of the interphase nuclear matrix fraction. Rather, the protein appears to leak out of nuclei during subcellular fractionation. We were only able to convincingly demonstrate the presence of ScII in nuclei when cells were fractionated into karyoplasts and cytoplasts by centrifugation in the presence of cytochalasin B.

A different situation is observed during mitosis, when standard fractionation procedures reveal that ~50% of the ScII is associated with isolated chromosomes. When these

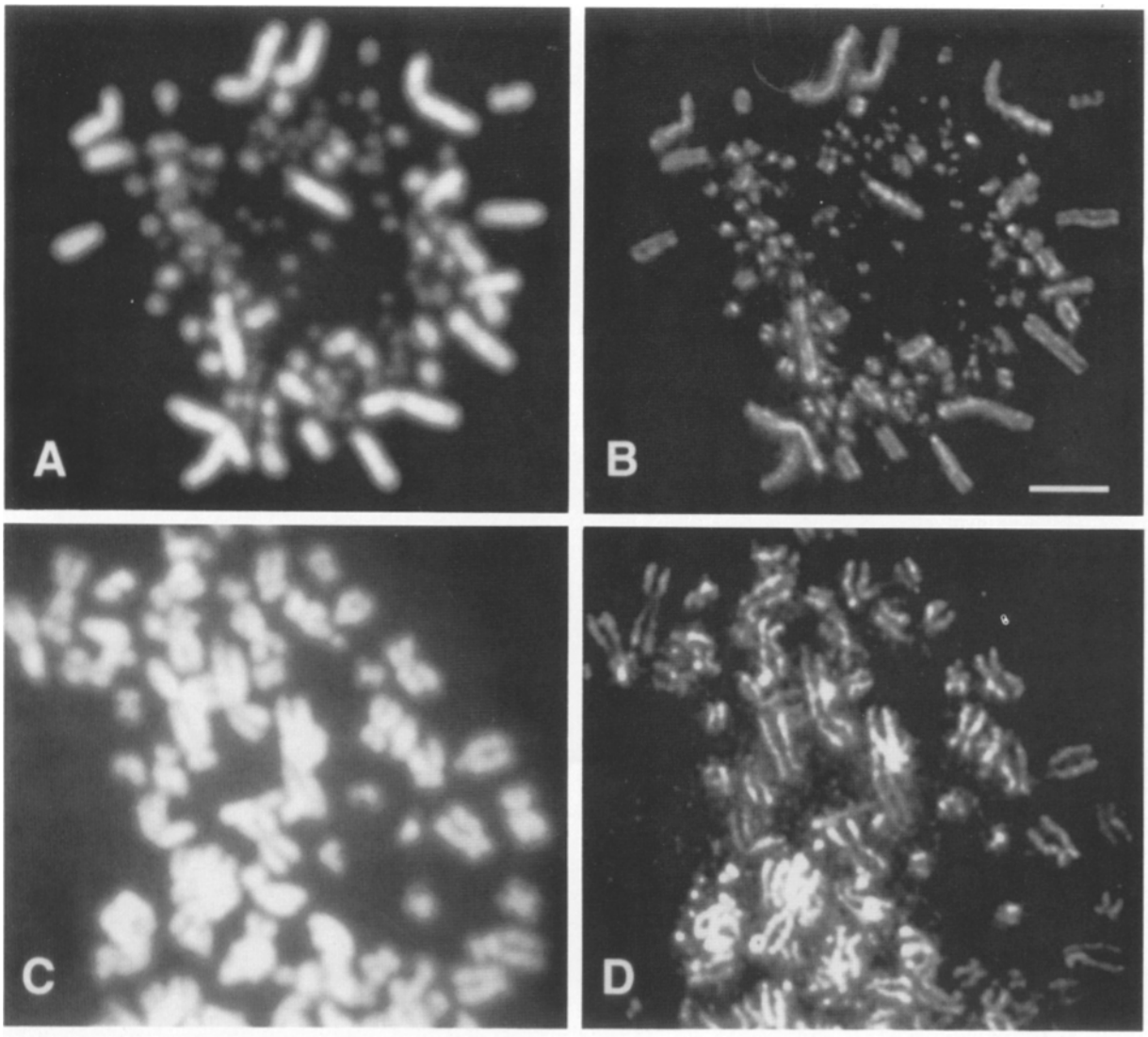


Figure 10. Localization of ScII in cytological spreads of chicken (MSB-1, *A* and *B*) and human (HeLa, *C* and *D*) cells. Cells were fixed with methanol–acetic acid, and cytological spreads were prepared by gentle blowing with an aquarium pump. Cells were stained with DAPI (*A* and *C*) or anti-ScII antibodies (*B* and *D*).

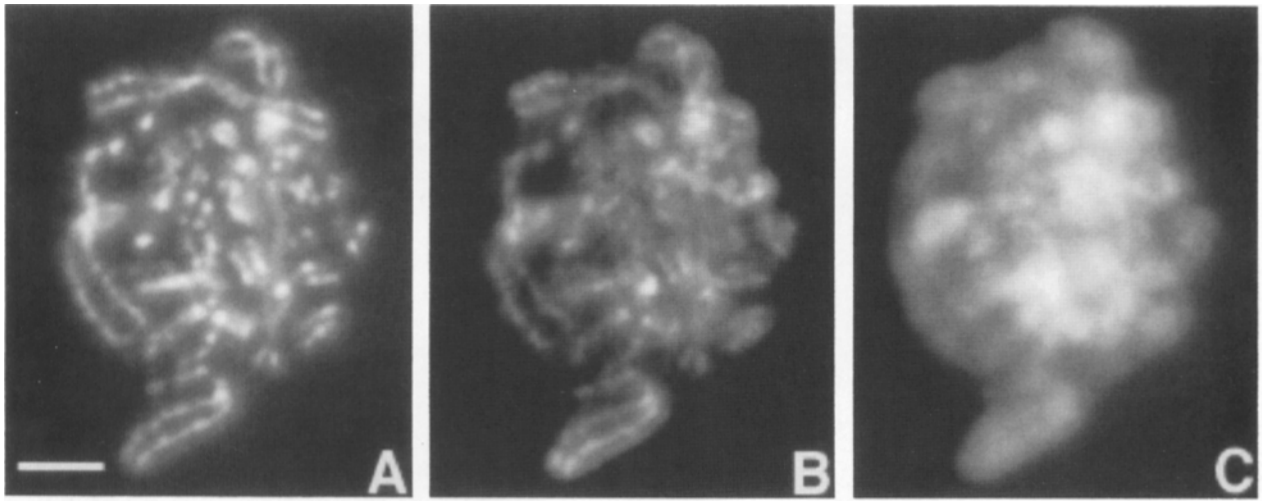


Figure 11. Colocalization of ScII and topoisomerase II throughout the axial region of mitotic chromosome arms. MSB-1 cells were fixed with methanol–acetic acid, the chromosomes were spread with an aquarium air pump, and stained with anti-ScII antisera (*A*), anti-topoisomerase II (*B*), or DAPI (*C*). We note that ScII appears to occupy a slightly more condensed distribution along the chromatid arms than does topoisomerase II.

chromosomes are subjected to DNase digestion and extraction to reveal the insoluble chromosome scaffold, ~70% of the chromosomal ScII remains in this fraction. In similar experiments, we previously observed that 72% of chromosomal topoisomerase II is retained in the chromosome scaffold fraction (Earnshaw et al., 1985).

Thus, ScII interacts very differently with chromosomes during interphase and in mitosis. The basis for this change is not understood. It is possible that phosphorylation of the ScII polypeptide during mitosis alters its interactions with one or more chromosomal components.

ScII Belongs to an Emerging Family of Polypeptides that have Two Internal Coiled-Coil Domains Flanked by NTP-binding A and B Sites at Either End

Our analysis of the ScII cDNA clones predicts a polypeptide with a molecular weight of 134,900 and a pI of 8.6. The deduced amino acid sequence shows limited similarity to a protein, Smclp, which is involved in chromosome segregation in the budding yeast. The *Smcl* mutation was first identified in a screen for mutations that resulted in increased frequencies of selective loss of a plasmid minichromosome (Strunnikov et al., 1993). Subsequent disruption of the gene resulted in a perturbation of the segregation of bona fide yeast chromosomes (Strunnikov et al., 1993). Thus, the genetic analysis of the *SMCl* gene and our biochemical analysis of ScII are consistent with the notion that these proteins play some role in chromosome structure or function during mitosis.

The *SMCl* protein is the founder member of an emerging family of related polypeptides in organisms from myco-

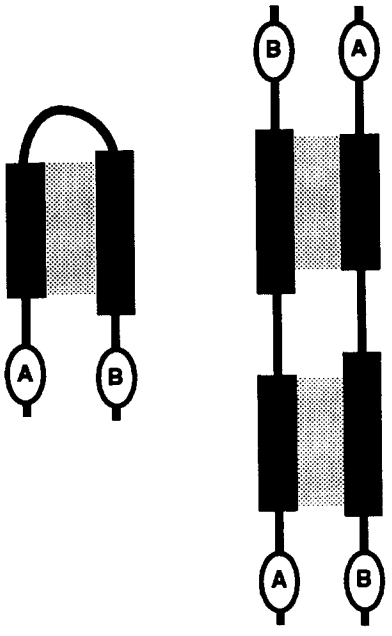


Figure 12. Proposed model for the structure of ScII. The circled A and B represent NTP-binding A and B sites. Those sites are brought close together by bending of the spacer between the coiled coils (*left*) or forming an antiparallel homodimer (*right*). Thick boxed regions represent the predicted coiled-coils, and the shaded regions represent the interactions between the coils that hold the structure together.

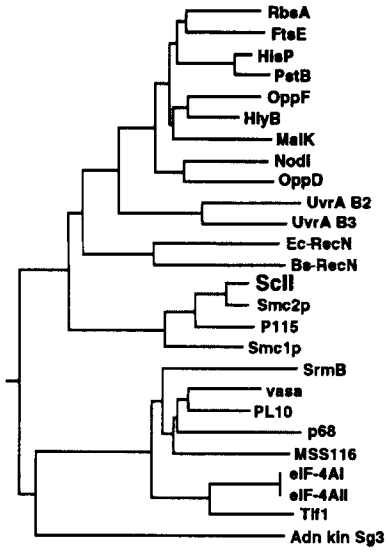


Figure 13. Unrooted phylogenetic tree showing distances between Walker B sites of selected ABC transporters, DNA repair and recombination enzymes, *SMCl* family members, RNA helicases, and adenylate kinase.

plasma to man, which share three common features (Strunnikov et al., 1993). These polypeptides, which range from 115 to 165 kD in size, all have a nucleotide-binding pocket (A site) very close to their amino terminus, and a second conserved region very close to their carboxy terminus. As described below, this second site resembles the NTP-binding B site found in several classes of ATPases. These two sites, potentially involved in nucleotide utilization, are widely separated by two regions predicted to form α-helical coiled coil (~250 and 350 amino acids in length, respectively, in ScII), and a central spacer. Overall comparison of the primary amino acid sequences for the entire molecules does not show a particularly high level of similarity among *SMCl* family members. This is partly because the regions of coiled-coil are divergent in sequence, even between the most closely related family members, ScII and Smc2p. If the coiled-coils serve primarily as spacers to position the Walker A and B sites with respect to one another, then the relative lengths of the coils may be more important than their detailed amino acid sequences.

It is unusual to have the NTP-binding A and B sites so far apart in an ATPase, since both must interact with the nucleoside triphosphate, a small molecule. This leads us to predict one of two conformations for members of the *SMCl* family. First, the proteins could function as monomers if the central region is a hinge (Fig. 12, *right*). In this case, the two regions of coiled-coil could fold back on one another in an antiparallel conformation, thus giving the molecule a hairpin shape and bringing the two parts of the NTP binding fold together. Alternatively, the proteins could function as extended antiparallel dimers, with the NTP-binding A site of one monomer associating with the B site from its partner (Fig. 12, *left*). Of course all of this speculation rests on the validity of the assumption that these proteins are, in fact, ATPases. It will be essential to test this hypothesis as purified proteins become available.

When the sequences of the *SMCl* family members are compared in detail, some similarity is also noted in the re-

gion of the central spacer that separates the two regions of coiled-coil. As suggested above, this region might act as a hinge, creating a characteristic fold so that two A and B sites are brought together. Alternatively, it may form a site for interaction with other proteins such as topoisomerase II (see below) or with DNA.

SMC1 Family Members Resemble a Bacterial DNA Repair and Recombination Enzyme

NTP-binding A and B sites were defined by Walker and colleagues by using amino acid sequences from several different ATPases, including adenylate kinase, Rec A protein, and myosin (Walker et al., 1982). A variety of studies, including site-directed mutagenesis and structural analysis by nuclear magnetic resonance, have led to the thinking that the positively charged lysine residue in the A site binds to the negatively charged phosphate, while the negatively charged aspartate in the B site coordinates the positively charged Mg^{2+}. This aspartate in the B site may be involved in hydrolysis of the phosphodiester bond (Serpersu et al., 1986; Pause and Sonenberg, 1992).

Although the core sequence of the A site is fairly well conserved, the B site is considerably more divergent. However, functionally related proteins tend to have B site sequences that are more closely related than the corresponding regions of functionally distinct proteins. To gain some insight about the possible function of ScII, we assembled a consensus sequence for the conserved COOH-terminal region common to the *SMC1* family members. In the original description of *Smc1*, the similarity of this conserved sequence to the Walker B site was not discussed, and the motif was called the DA box (Strunnikov et al., 1993). Using our consensus, we searched the entire coding region of representatives from the transporter and RNA helicase families, as well as other proteins whose ATPase activities were well known. The fact that the computer recognized the well-characterized B sites of those proteins as related to the DA box supports our hypothesis that the DA box may be an NTP-binding B site. Significantly, all of the B sites examined, including those of the *SMC1* family members, contained the above mentioned aspartate as the only absolutely conserved residue. Next, taking advantage of the alignments between the B sites of respective ABC transporters, RNA helicases, and adenylate kinase, we established a phylogenetic tree (Fig. 13) based on the distance between the sequences in the alignments. The *SMC1* family was positioned in the middle of this tree. Similarities between ABC transporters, UvrA, and recN (Doolittle et al., 1986; Gorbalenya and Koonin, 1990) that have previously been found by comparing the entire coding sequences were also found by our strategy.

We were very interested to find that the protein whose NTP-binding B site most closely resembles the *SMC1* family consensus also shares a second significant structural feature with the *SMC1* family members. In this protein, bacterial recN (a DNA repair and recombination enzyme), the two NTP-binding sites are also separated by a relatively long stretch of ~500 amino acids that has a significant probability of forming a coiled-coil. We emphasize that the phylogenetic tree of Fig. 13 only concerns the sequence of the NTP-binding B site, and therefore, probably only indicates proteins that hydrolyse ATP in a similar manner. However, this additional similarity between recN and the *SMC1* family members may indicate further similarities of function, and it serves as a starting point for future functional analysis of the role of ScII in the chromosome.

The Function of ScII is Unknown, but the Protein Has Been Shown To Be in at Least One Complex with Topoisomerase II

Although the function of ScII is unknown, evidence obtained in an earlier study indicated that the protein can exist in a complex with topoisomerase II. In that study, a protein complex called UB2 was purified from undifferentiated murine erythroleukemia cells based on its binding to a specific sequence motif found in the 3′ flanking region of the β-globin gene (Ma et al., 1993). This binding site also occurs near the carbonic anhydrase I and c-*myb* genes, as well as in the immunoglobulin heavy chain enhancer region (Ma et al., 1991). These regions shared two discernible sites that are important for UB2 binding. The UB2 DNA-binding activity was not present in MEL cells that had been induced to differentiate and had initiated transcription of the β-globin gene. Thus, it was postulated that the UB2 complex might be involved in regulation of β-globin expression.

The UB2 activity passed over DEAE sepharose and DNA affinity columns was found to copurify with three major polypeptides of 170, 116, and 48 kD, and a minor polypeptide of 135 kD (Ma et al., 1993). The 170-kD component was shown by immunoblotting to be topoisomerase II, and addition of anti–topoisomerase II to the nucleoprotein gels produced a super shift of the UB2 complex. The 135-kD component of this complex was found to be ScII. Addition of affinity-purified anti-ScII antibodies caused a specific increase in the mobility of the UB2 complex, consistent with dissociation of one or more components.

In addition to the presence of topoisomerase II and ScII in the UB2 complex, two other observations also support the notion that the two molecules might interact. First, we have noted that the distribution of ScII in spread mitotic chromosomes is essentially identical to that seen for topoisomerase II. Second, in an independent method of purifying topoisomerase II, nuclear extracts from a chicken lymphoblastoid cell line were passed sequentially through hydroxyapatite and phosphocellulose columns (Wood, E. R., unpublished data). The final fraction contained roughly equimolar amounts of topo II, ScII, and a third (unidentified) species of 116 kD. Although we have not yet shown that these proteins are actually associated in a complex, the similarity of this fraction to the UB2 complex is striking.

It will be very important in future studies to determine whether ScII is, as its sequence suggests, an ATPase, and whether the protein has an essential role (possibly enzymatic) in chromosome condensation or sister chromatid disjunction. It will also be interesting to confirm whether ScII and topoisomerase II do exist in specific complexes in chromosomes, and if so to assess the relevance of such complexes in vivo for chromosome structure and function.

The authors gratefully acknowledge Dr. Bjorn Vennstrom's gift of his chicken λgt11 expression library and Dr. Don Cleveland's gift of his anti-tubulin antibody. We thank Dr. Yuri Lazebnik, Dr. Alastair Mackay, Dr. Ann Pluta, Dr. John Tomkiel, and Dr. Charles Yang for many insightful

suggestions during the course of this work. We also thank Victor Chua for help in sequencing.

These experiments were supported by National Institutes of Health grant GM30985 to W. C. Earnshaw.

Received for publication 29 June 1994 and in revised form 29 July 1994.

Note Added in Proof. We have recently learned of two further *SMC1* family members that resemble ScII. Fission yeast cut 14p is 34% identical to ScII, and it is required for chromosome condensation and disjunction in mitosis (Saka, Y., T. Sutani, Y. Yamashita, S. Saitoh, M. Takeuchi, Y. Nakaseko, and M. Yanagida. *EMBO (Eur. Mol. Biol. Organ.) J.* In press). *Xenopus* XCAP-E is 68% identical to ScII, and it is required for chromosome condensation in cell-free mitotic egg extracts (Hirano, T., and T. J. Mitchison. A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell.* In press).

References

Adachi, Y., E. Käs, and U. K. Laemmli. 1989. Preferential cooperative binding of DNA topoisomerase II to scaffold-associated regions. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:3997–4006.

Adachi, Y., M. Luke, and U. K. Laemmli. 1991. Chromosome assembly in vitro: topoisomerase II is required for condensation. *Cell.* 64:137–148.

Adolph, K. W., S. M. Cheng, and U. K. Laemmli. 1977*a*. Role of nonhistone proteins in metaphase chromosome structure. *Cell.* 12:805–816.

Adolph, K. W., S. M. Cheng, J. R. Paulson, and U. K. Laemmli. 1977*b*. Isolation of a protein scaffold from mitotic HeLa cell chromosomes. *Proc. Natl. Acad. Sci. USA.* 11:4937–4941.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1991. Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Bell, A. W., S. D. Buckel, J. M. Groarke, J. N. Hope, D. H. Kingsley, and M. A. Hermodson. 1986. The nucleotide sequences of the rbsD, rbsA, and rbsC genes of *Escherichia coli* K12. *J. Biol. Chem.* 261:7652–7658.

Berrios, M., N. Osheroff, and P. A. Fisher. 1985. In situ localization of DNA topoisomerase II, a major polypeptide component of the *Drosophila* nuclear matrix fraction. *Proc. Natl. Acad. Sci. USA.* 82:4142–4146.

Boy de la Tour, E., and U. K. Laemmli. 1988. The metaphase scaffold is helically folded: Sister chromatids have predominantly opposite helical handedness. *Cell.* 55:937–944.

Clarke, D. J., R. T. Johnson, and C. S. Downes. 1993. Topoisomerase II inhibition prevents anaphase chromatid segregation in mammalian cells independently of the generation of DNA strand breaks. *J. Cell Sci.* 105:563–569.

Comings, D. E., and T. A. Okada. 1971. Fine structure of kinetochore in Indian muntjac. *Exp. Cell Res.* 67:97–110.

Compton, D. A., T. J. Yen, and D. W. Cleveland. 1991. Identification of novel centromere/kinetochore associated proteins using monoclonal antibodies generated against human mitotic chromosome scaffolds. *J. Cell Biol.* 112: 1083–1097.

Cooke, C. A., M. M. S. Heck, and W. C. Earnshaw. 1987. The INCENP antigens: movement from the inner centromere to the midbody during mitosis. *J. Cell Biol.* 105:2053–2067.

Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids res.* 12:387–395.

DiNardo, S., K. Voelkl, and R. Sternglanz. 1984. DNA topoisomerase II mutant of Saccharomyces cerevisiae: topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proc. Natl. Acad. Sci. USA.* 81:2616–2620.

Doolittle, R. F., M. S. Johnson, I. Husain, B. Van Houten, D. C. Thomas, and A. Sankar. 1986. Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. *Nature (Lond.).* 323: 451–453.

Downes, C. S., A. M. Mullinger, and R. T. Johnson. 1991. Inhibitors of DNA topoisomerase II prevent chromatid separation in mammalian cells but do not prevent exit from mitosis. *Proc. Nat. Acad. Sci. USA.* 88:8895–8899.

Earnshaw, W. C. 1991. Large scale chromosome structure and organization. *Curr. Topics Struct. Biol.* 1:237–244.

Earnshaw, W. C., and U. K. Laemmli. 1983. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol.* 96:84–93.

Earnshaw, W. C., N. Halligan, C. Cooke, and N. Rothfield. 1984. The kinetochore is part of the chromosome scaffold. *J. Cell Biol.* 98:352–357.

Earnshaw, W. C., B. Halligan, C. A. Cooke, M. M. S. Heck, and L. F. Liu. 1985. Topoisomerase II is a structural component of mitotic chromosome scaffolds. *J. Cell Biol.* 100:1706–1715.

Earnshaw, W. C., and M. M. S. Heck. 1985. Localization of topoisomerase II in mitotic chromosomes. *J. Cell Biol.* 100:1716–1725.

Earnshaw, W. C., and B. Migeon. 1985. A family of centromere proteins is absent from the latent centromere of a stable isodicentric chromosome. *Chromosoma (Berl.).* 92:290–296.

Earnshaw, W. C., and N. Rothfield. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma (Berl.).* 91:313–321.

Earnshaw, W. C., H. Ratrie, and G. Stetten. 1989. Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. *Chromosoma (Berl).* 98:1–12.

Evans, I. J., and J. A. Downie. 1986. The *nodI* gene product of *Rhizobium leguminosarum* is closely related to ATP-binding bacterial transport proteins: nucleotide sequence analysis of the *nodI* and *nodJ* genes. *Gene (Amst.).* 43:95–101.

Fath, M. J., and R. Kolter. 1993. ABC transporters: bacterial exporters. *Microbiol. Rev.* 57:995–1017.

Feinberg, A., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13.

Feinberg, A., and B. Vogelstein. 1984. Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 137:266–267.

Felmlee, T., S. Pellett, and R. A. Welch. 1985. Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. *J. Bacteriol.* 163:94–105.

Felsenstein, J. 1989. PHYLIP: phylogeny inference package (version 3.2). *Cladistics.* 5:164–166.

Finch, J. T., and A. Klug. 1976. Solenoidal model for superstructure in chromatin. *Proc. Natl. Acad. Sci. USA.* 73:1897–1901.

Ford, M. J., I. A. Anton, and D. P. Lane. 1988. Nuclear protein with sequence homology to translation initiation factor eIF-4A. *Nature (Lond.).* 332:736–738.

Frohman, M. A., M. K. Dush, and G. R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA.* 85:8998–9002.

Fry, D. C., S. A. Kuby, and A. S. Mildvan. 1986. ATP-binding site of adenylate kinase: mechanistic implication of its homology with *ras*-encoded p21, F_1-ATPase, and other nucleotide-binding proteins. *Proc. Natl. Acad. Sci. USA.* 83:907–911.

Gasser, S. M., B. B. Amati, M. E. Cardenas, and J. F.-X. Hofmann. 1989. Studies on scaffold attachment sites and their relation to genome function. *Intl. Rev. Cytol.* 119:57–96.

Gasser, S. M., T. Laroche, J. Falquet, E. Boy de la Tour, and U. K. Laemmli. 1986. Metaphase chromosome structure. Involvement of topoisomerase II. *J. Mol. Biol.* 188:613–629.

Gill, D. R., G. F. Hatfull, and G. P. Salmond. 1986. A new cell division operon in *Escherichia coli. Mol. Gen. Genet.* 205:134–145.

Gilson, E., H. Nikaido, and M. Hofnung. 1982. Sequence of the malK gene in *E. coli* K12. *Nucleic Acids Res.* 10:7449–7458.

Gorbalenya, A. E., and E. V. Koonin. 1990. Superfamily of UvrA-related NTP-binding proteins. Implications for rational classification of recombination/repair systems. *J. Mol. Biol.* 213:583–591.

Heck, M. M. S., and W. C. Earnshaw. 1986. Topoisomerase II: a specific marker for proliferating cells. *J. Cell Biol.* 103:2569–2581.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene (Amst.).* 28:351–359.

Herrick, G., B. B. Spear, and G. Veomett. 1976. Intracellular localization of mouse DNA polymerase-α. *Proc. Natl. Acad. Sci. USA.* 73:1136–1139.

Higgins, C. F., P. D. Haag, K. Nikaido, F. Ardeshir, G. Garcia, and G. F. Ames. 1982. Complete nucleotide sequence and identification of membrane components of the histidine transport operon of S. typhimurium. *Nature (Lond.).* 298:723–727.

Higgins, C. F., I. D. Hiles, K. Whalley, and D. J. Jamieson. 1985. Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems. *EMBO (Eur. Mol. Biol. Organ.). J.* 4:1033–1039.

Higgins, C. F., I. D. Hiles, G. P. C. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell, and M. A. Hermodson. 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. *Nature (Lond.).* 323:448–450.

Hiles, I. D., M. P. Gallagher, D. J. Jamieson, and C. F. Higgins. 1987. Molecular characterization of the oligopeptide permease of *Salmonella typhimurium. J. Mol. Biol.* 195:125–142.

Hirano, T., and T. J. Mitchison. 1993. Topoisomerase II does not play a scaffolding role in the organization of mitotic chromosomes assembled in *Xenopus* egg extracts. *J. Cell Biol.* 120:601–612.

Hoffmann, A., M. M. S. Heck, B. J. Bordwell, N. F. Rothfield, and W. C. Earnshaw. 1989. Human autoantibody to topoisomerase II. *Exp. Cell Res.* 180:409–418.

Holm, C. 1994. Coming undone: how to untangle a chromosome. *Cell.* 77: 955–957.

Holm, C., T. Goto, J. C. Wang, and D. Botstein. 1985. DNA topoisomerase II is required at the time of mitosis in yeast. *Cell.* 41:553–563.

Horowitz, R. A., D. A. Agard, J. W. Sdat, and C. L. F. Woodcock. 1994. The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal

ribbon. *J. Cell Biol.* 124:1–10.
Husain, I., B. Van Houten, D. C. Thomas, and A. Sancar. 1986. Sequences of *Escherichia coli* uvrA gene and protein reveal two potential ATP binding sites. *J. Biol. Chem.* 261:4895–4901.
Izaurralde, E., E. Käs, and U. K. Laemmli. 1989. Highly preferential nucleation of histone H1 assembly on scaffold-associated regions. *J. Mol. Biol.* 210:573–585.
Kornberg, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science (Wash. DC).* 184:868–871.
Krek, W., G. Maridor, and E. A. Nigg. 1992. Casein kinase II is a predominantly nuclear enzyme. *J. Cell Biol.* 116:43–55.
Kuby, S. A., R. H. Palmieri, A. Frischat, A. H. Fischer, L. H. Wu, L. Maland, and M. Manship. 1984. Studies on adenosine triphosphate transphosphorylases. Amino acid sequence of rabbit muscle ATP-AMP transphosphorylase. *Biochemistry.* 23:2393–2399.
Laemmli, U. K., S. M. Cheng, K. W. Adolph, J. R. Paulson, J. A. Brown, and W. R. Baumbach. 1978. Metaphase chromosome structure: the role of nonhistone proteins. *Cold Spring Harbor Symp. Quant. Biol.* 423:351–360.
Lasko, P. F., and M. Ashburner. 1988. The product of the Drosophila gene vasa is very similar to eukaryotic initiation factor-4A. *Nature (Lond.).* 335:611–617.
Leroy, P., P. Alzari, D. Sassoon, D. Wolgemuth, and M. Fellous. 1989. The protein encoded by a murine male germ cell-specific transcript is a putative ATP-dependent RNA helicase. *Cell.* 57:549–559.
Lewis, C. D., and U. K. Laemmli. 1982. Higher order metaphase chromosome structure: evidence for metalloprotein interactions. *Cell.* 29:171–181.
Linder, P., P. F. Lasko, M. Ashburner, P. Leroy,, P. Nielsen, K. Nishi, J. Schnier, and P. P. Slonimski. 1989. Birth of the D-E-A-D box. *Nature (Lond.).* 337:121–122.
Linder, P., and P. P. Slonimski. 1989. An essential yeast protein, encoded by duplicated genes TIF1 and TIF2 and homologous to the mammalian translation initiation factor eIF-4A, can suppress a mitochondrial missense mutation. *Proc. Natl. Acad. Sci. USA.* 86:2286–2290.
Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science (Wash. DC).* 252:1162–1164.
Ma, X., P. Fraser, and P. J. Curtis. 1991. A differentiation stage-specific factor interacts with mouse carbonic anyhydrase form I gene and a conserved sequence in mammalian β-globin genes. *Differentiation.* 47:135–141.
Ma, X., N. Saitoh, and P. J. Curtis. 1993. Purification and characterization of a nuclear DNA-binding factor complex containing topoisomerase II and chromosome scaffold protein 2. *J. Biol. Chem.* 268:6182–6188.
Mackay, A. M., D. M. Eckley, C. Chue, and W. C. Earnshaw. 1993. Molecular analysis of the INCENPs (inner centromere proteins): separate domains are required for association with microtubules during interphase and with the central spindle during anaphase. *J. Cell Biol.* 123:373–385.
Mirkovitch, J., M.Mirault, and U. Laemmli. 1984. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell.* 39:223–232.
Nielsen, P. J., G. K. McMaster, and H. Trachsel. 1985. Cloning of eukaryotic protein synthesis initiation factor genes: isolation and characterization of cDNA clones encoding factor eIF-4A. *Nucleic Acids Res.* 13:6867–6880.
Nielsen, P. J., and H. Trachsel. 1988. The mouse protein synthesis initiation factor 4A gene family includes two related functional genes which are differentially expressed. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:2097–2105.
Nishi, K., F. Morel-Deville, J. W. Hershey, T. Leighton, and J. Schnier. 1988. An eIF-4A-like protein is a suppressor of an *Escherichia coli* mutant defective in 50S ribosomal subunit assembly. *Nature (Lond.).* 336:496–498.
Notarnicola, S. M., M. A. Mcintosh, and K. S. Wise. 1991. A *mycoplasma hyorhinis* protein with sequence similarities to nucleotide-binding enzymes. *Gene (Amst.).* 97:77–85.
Pause, A., and N. Sonenberg. 1992. Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO (Eur. Mol. Biol. Organ.) J.* 11:2643–2654.
Pearson, W. R. 1990. Rapid, and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63–98.
Pienta, K. J., and D. S. Coffey. 1984. A structural analysis of the role of the nuclear matrix and DNA loops in the organization of the nucleus and chromosome. *J. Cell Sci.* (Suppl. 1):123–135.
Prescott, D. M., and J. B. Kirkpatrick. 1973. Mass enucleation of cultured animal cells. *Methods Cell Biol.* 7:189–202.
Rattner, J. B., and C. C. Lin. 1985. Radial loops and helical coils coexist in metaphase chromosomes. *Cell.* 42:291–296.
Rostas, K., S. J. Morton, S. M. Picksley, and R. G. Lloyd. 1987. Nucleotide sequence and LexA regulation of the *Escherichia coli* recN gene. *Nucleic Acids Res.* 15:5041–5049.
Saitoh, Y., and U. K. Laemmli. 1994. Mataphase chromosome structure: bands arise from a differential folding path of the high AT-rich scaffold. *Cell.* 76:609–622.
Seraphin, B., M. Simon, A. Boulet, and G. Faye. 1989. Mitochondrial splicing requires a protein from a novel helicase family. *Nature (Lond.).* 337:84–87.
Serpersu, E. H., D. Shortle, and A. S. Mildvan. 1986. Kinetic and magnetic resonance studies of effects of genetic substitution of a Ca^{++}-liganding amino acid in *Staphlococcus nuclease. Biochemistry.* 25:68–77.
Shamu, C. E., and A. W. Murray. 1992. Sister chromatid separation in frog egg extracts requires DNA topoisomerase II activity during anaphase. *J. Cell Biol.* 117:921–934.
Strunnikov, A. V., V. L. Larionoc, and D. Koshland. 1993. *SMC1*: an essential yeast gene encoding a putative head-rod-tail protein is required for nuclear division and defines a new ubiquitous protein family. *J. Cell Biol.* 123:1635–1648.
Surin, B. P., H. Rosenberg, and G. B. Cox. 1985. Phosphate-specific transport system of *Escherichia coli*: nucleotide sequence and gene-polypeptide relationships. *J. Bacteriol.* 161:189–198.
Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA.* 82:1074–1078.
Thomas, F., T. Koller, and A. Klug. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* 83:403–427.
Uemura, T., H. Ohkura, Y. Adachi, K. Morino, K. Shiozaki, and M. Yanagida. 1987. DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in *S. pombe. Cell.* 50:917–925.
Uemura, T., and M. Yanagida. 1986. Mitotic spindle pulls but fails to separate chromosomes in type II DNA topoisomerase mutants: uncoordinated mitosis. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:1003–1010.
Van Hoy, B. E., and J. A. Hoch. 1990. Characterization of the spoIVB and recN loci of *Bacillus subtilis. J. Bacteriol.* 172:1306–1311.
Walker, J. E., M. Sarste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO (Eur. Mol. Biol. Organ.) J.* 1:945–951.
Williamson, D. H., and D. J. Fennell. 1975. The use of fluorescent DNA-binding agent for detecting and separating yeast mitochondrial DNA. *Methods Cell Biol.* 12:335–351.
Yen, T. J., D. A. Compton, W. C. Earnshaw, and D. W. Cleveland. 1991. CENP-E, a human centromere-associated protein released from chromosomes at the onset of anaphase. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:1245–1991.
Young, R. A., and R. B. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194–1198.
Zhao, K., E. Käs, E. Gonzalez, and U. K. Laemmli. 1993. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:3237–3247.