In vitro assembly of physiological cohesin/DNA complexes

Itay Onn\textsuperscript{a,b} and Douglas Koshland\textsuperscript{a,c,1}

\textsuperscript{a}Howard Hughes Medical Institute, \textsuperscript{b}Department of Embryology, Carnegie Institution, 3520 San Martin Drive, Baltimore, MD 21218; and \textsuperscript{c}Department of Molecular and Cell Biology, University of California, Berkeley, 16 Barker Hall #3202, Berkeley, CA 94720-3202

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2010.

Contributed by Douglas Koshland, May 17, 2011 (sent for review April 6, 2011)

Cohesin is a member of the Smc family of protein complexes that mediates higher-order chromosome structure by tethering different regions of chromatin. We present a new in vitro system that assembles cohesin-DNA complexes with in vivo properties. The assembly of these physiological salt-resistant complexes requires the cohesin holo-complex, its ability to bind ATP, the cohesin loader Scc2p and a closed DNA topology. Both the number of cohesin molecules bound to the DNA substrate and their distribution on the DNA substrate are limited. Cohesin and Scc2p bind preferentially to cohesin associated regions (CARs), DNA sequences with enriched cohesin binding in vivo. A subsequence of CARC1 promotes cohesin binding to neighboring sequences within CARC1. The enhancer-like function of this sequence is validated by in vivo deletion analysis. By demonstrating the physiological relevance of these in vitro assembled cohesin-DNA complexes, we establish our in vitro system as a powerful tool to elucidate the mechanism of cohesin and other Smc complexes.

Cohesin mediates higher-order chromosome structure by binding and tethering different regions of chromatin. These cohesin activities are important for sister chromatid cohesion, condensation, transcription, and DNA repair (recently reviewed in refs. 1 and 2). Despite the importance of cohesin to genome integrity and chromosome dynamics, the molecular mechanism by which cohesin binds and tethers chromatin fibers remains to be elucidated. The complex architecture of cohesin and the complex features of its in vivo chromatin binding have impeded attempts to elucidate its molecular function.

At the core of cohesin’s architecture are the Smc1p and Smc3p subunits, which are orthologs of the structural maintenance of chromosomes (SMC) ATPases family. Each Smc protein contains a central hinge domain flanked by two long coiled-coil domains that culminate in a head domain. The Smc proteins form a large ring by dimerization of their hinge domains at one end and the their head domains at the other end. The dimerization of the heads also generates two ABC-like ATPase domains. The third subunit, Mcd1p/Scc1p/Rad21p, links the heads together, whereas a fourth subunit Scc3p/Irr1p/SAP binds to Mcd1p (3, 4). This complex architecture coupled with the fact that none of the cohesin subunits have known DNA binding motifs indicated early on that cohesin binds DNA by a mechanism distinct from other DNA binding proteins.

Additional evidence for a different mechanism of binding came from in vivo studies of cohesin function. First, its DNA binding is not attributable to a single subunit but rather requires the assembly and function of all the subunits. In vivo assembled cohesin/chromosome complexes can be very stable as evidenced by the inability to elute cohesin from isolated chromosomes with 1.6 M salt (5). Binding to chromosomes also requires ATP binding to both Smc1p and Smc3p (6–8). In all organisms, cohesins bind to specific regions on chromosome arms as well as to large domains in the pericentric region (9, 10). In budding and fission yeast, cohesins bind to highly AT-rich intergenic regions covering 0.5–2.0 Kb (9–12). These broad binding sites are called cohesin associated regions (CARs) (9). The broad size of CARs and their AT richness have led researchers to assume that cohesin binding to chromosomes is not dictated by DNA sequence. Finally, the association of cohesins with chromosomes requires the Scc2p/Scc4p complex that loosely interacts with cohesin but whose molecular function is unknown (5). For example, Scc2p/Scc4p complex may act directly on cohesin or recruit chromatin remodeling factors to allow cohesin access to DNA. Taken together, the unusual features of CARs, the requirement for ATP and multiple subunits, and the extreme salt resistance all make cohesin different from canonical DNA binding proteins.

Many of the unusual properties of cohesin binding to chromosomes can be explained by the ring or embrace model. In this model chromosomes become topologically entrapped within the large ring of cohesin through an ATP-dependent mechanism (13–15). This model is consistent with the requirement for all subunits of cohesin for DNA binding. Indeed, cleavage of the protein backbone of cohesin subunits or the DNA backbone of isolated circular minichromosomes can release cohesin from chromosomes. The topological binding rather than direct contact with DNA would also allow it to bind to anywhere on chromosomes and yet be enriched at CARs by indirect mechanisms of chromosome metabolism. For example, one in vivo study suggests that cohesin is loaded at Scc2p/Scc4p binding sites distinct from CARs and then is pushed to CARs by transcription (16). Although the simplicity of the embrace model is attractive, a number of studies suggest that the simple entrapment of DNA and its derivative hypotheses may need to be modified. For example, in vivo cleavage of the DNA backbone next to a CAR fails to release cohesin from the DNA (17) suggesting cohesin may bind chromosomes by additional or alternative mechanisms than topological entrapment. Indeed, in vivo FRAP studies suggest cohesin can bind to chromosomes in two distinct modes that differ in residence time on chromosomes (18–21). A recent in vivo study indicates that Scc2p/Scc4p binding sites lie within CARs, indicating that cohesin enrichment at CARs may not be indirect (22). Colocalization of cohesin and its loader was also detected in human and drosophila (23, 24). Finally in the embrace model, the thinness of the ring and the absence of direct contact with DNA infer that the number of cohesins bound to chromosomes should be limited only by the number of cohesins in the cell, a concept that has been assumed but not tested.

Author contributions: D.K. and I.O. designed research; D.K. and I.O. performed research; I.O. contributed new reagents/analytic tools; D.K. and I.O. analyzed data; and D.K. and I.O. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

\textsuperscript{1}To whom correspondence should be addressed. E-mail: koshland@berkeley.edu.

This article contains supporting information online at www.pnas.orglookup.suppl/doi:10.1073/pnas.1107504108/-/DSSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1107504108
Clearly many of these uncertainties about cohesin binding to DNA could be resolved by in vitro biochemical studies. Indeed, a number of studies have utilized cohesin assembled from purified subunits to analyze its interaction with generic DNA. These studies document cohesin’s ability to bind linear DNA independent of DNA sequence and its high affinity for DNA secondary structures (25–27). However, the physiological relevance of these observations is not clear as these cohesin/DNA complexes fail to mimic known properties of in vivo assembled complexes like ATP and Scc2p dependence, salt resistance, topology dependence, or preferential binding to different DNA regions. Furthermore, the biological relevance of these in vitro properties has not been validated by a diagnostic in vivo test. Biochemical studies of other SMC complexes had similar drawbacks (28–31). Therefore, the study of cohesin and all SMC proteins have suffered from the absence of a physiologically relevant in vitro system, coupled with the failure to assess the physiological relevance of in vitro observations. With this in mind, we developed an in vitro system that recapitulates the critical properties of in vivo assembled complexes. Furthermore, we successfully employ this system to address fundamental questions about cohesin-CAR interactions in vitro and demonstrate its relevance in vivo.

Results
Assaying Cohesin Binding to DNA in Vitro. The starting DNA substrate in our assay was prepared by amplifying a 5-kb fragment that has a 3.2-kb sequence containing CARC1 (Fig. 1A). One or both ends of the PCR product were labeled with 5’ biotin (Fig. 1B). A PCR product labeled at one end should bind streptavidin-coated beads with linear topology (C1-L), whereas a PCR product labeled at both ends should allow both ends to bind to the same bead generating a closed topology (C1-C) (Fig. 1B). Analyses of C1-L and C1-C with restriction endonucleases confirm that greater than 93% have the linear and closed topology, respectively (SI Text). By elution of the bound DNA from C1-L and C1-C, we estimate that approximately 1,000–2,000 DNA molecules are bound per bead.

We assembled cohesin onto C1-L and C1-C, as described in Fig. 1C. We incubated the substrate with crude extracts from yeast cells arrested in S phase using hydroxyurea. We choose S-phase extracts for two reasons. First, in vivo experiments showed that S-phase cells have fully assembled cohesin complexes, are competent for cohesin loading onto DNA, and also are competent to convert chromatin-bound cohesin to its cohesive state, the latter being a future goal of our approach. Second, cohesin from a crude extract seemed more likely than purified or recombinant cohesin to have all the activities needed for loading cohesin onto DNA in a physiologically relevant manner. After incubation, beads were immobilized using a magnet, and then unbound proteins were removed by multiple washes. The bound fraction was eluted from DNA, and the presence of bound cohesin subunits were analyzed by Western blot with antibodies against cohesin subunits. For most analyses we used antibodies directed against the Mcd1p subunit of cohesin as a marker for cohesin binding to the beads. Previous in vivo analyses have shown that Mcd1p binds chromatin/DNA only in the context of the other subunits of cohesin. In addition, the binding of all the other subunits to chromatin/DNA requires functional Mcd1p (4, 7).

Cohesin Binding in Our In Vitro Assay Recapitulates Its Properties in Vivo. We first asked whether cohesin would associate with C1-L, and if so, whether this DNA binding would mimic the salt-resistant binding of in vivo assembled complexes. C1-L was incubated in S-phase yeast extracts; then the beads were washed with buffer containing 0.1, 0.5, 0.8, or 1.2 M KCl to remove unbound cohesin and to assess cohesin binding under conditions of increasing stringency (Fig. 2A). Western blot analyses show that Mcd1p binds to beads only when beads contain DNA. However Mcd1p is readily removed at salt concentrations above 0.1 M. Therefore, linear DNA does not allow the assembly of physiologically stable cohesin-DNA complexes.

Substrate topology has been implicated as an important feature of cohesin binding to DNA based on in vitro assays of circular minichromosomes where cohesin had been preassembled in vivo. Therefore, we wondered whether our in vitro assay would recapitulate salt-resistant cohesin binding if we used the topologically closed C1-C substrate. Incubation of equal amounts of C1-C and C1-L with the S-phase extract led to roughly twice as much Mcd1p binding as with C1-L in 0.1 M salt, suggesting that topology enhanced overall binding. Moreover, 60% of the C1-C Mcd1p remained bound to C1-C even at 1.2 M KCl, in contrast to the quantitative depletion we observed from C1-L. To further characterize the relationship between the salt-sensitive and salt-resistant cohesin/C1-C complexes, we measured the kinetics of their formation (Fig. 2B). Salt-sensitive complexes form first, followed by the salt-resistant complexes. This difference in kinetics supports distinct modes of binding. Hereafter we define salt-resistant cohesin-DNA complexes as those that remain intact after washes with 0.5 M KCl or greater. The topologically modulated assembly of a salt-resistant Mcd1p/DNA complex in
Fig. 2. In vitro cohesin binding recapitulates that of cohesin bound in vivo. (A) Protein extract was prepared from 3131-V13 cells arrested in S and mixed with C1-L or C1-C under the standard assay conditions with the following modification: Unbound proteins were washed with 0.1, 0.5, 0.8, or 1.2 M KCl. Bound proteins were analyzed with antibody against Mcd1p. Error bars represent SD of two independent experiments. (B) Protein extract was prepared from 3131-V13 cells arrested in S and mixed with C1-C under the standard assay conditions with the following modification: Samples were incubated for 2, 15, or 30 min before the unbound fraction was washed. Bound proteins were analyzed with antibody against Mcd1p. Error bars represent SD of two independent experiments. (C) Protein extract was prepared from YIO71 cells arrested in G1 or S and mixed with C1-C under the standard assay conditions. Bound proteins were analyzed with antibodies against HA (Smc3-6HA, K38I-6HA). Error bars represent SD of two independent experiments. (D) Protein extract was prepared from YIO5031 and YIO5032 cells arrested in S and mixed with C1-C under the standard assay conditions. Bound proteins were analyzed with antibodies against HA (Smc3p-6HA and smc3p K38I-6HA). Error bars represent SD of two independent experiments.

Our system mimics two important parameters of cohesin/chromosome complexes assembled in vivo, indicating that our in vitro assembled complexes are physiologically relevant.

Two alternative explanations might explain the preferred binding of cohesin to C1-C over C1-L. First, C1-L may be degraded by the extract. However, we observe no loss of either C1-C or C1-L DNA after 30 min of incubation with our extracts (SI Text). Alternatively, the less constrained DNA of C1-L may allow the binding of an inhibitor that precludes cohesin binding like the assembly of nucleosomes. The assembly of nucleosomes seemed unlikely because it had been shown that chromatin assembly in yeast extract requires specific conditions that are not present in our assay (32). Nevertheless, we compared the presence of nucleosomes on C1-L and C1-C by detecting the presence of histones H2B and H3 in the bound fraction. Neither template binds H2B or H3, indicating that neither assembles nucleosomes (SI Text). Although we cannot rule out some other unforeseen inhibitor, the preferential binding of cohesin to the C1-C is consistent with the topological requirements observed in vivo studies.

To assess further the physiological relevance of our in vitro system, we explored whether the in vivo regulation of cohesin binding to DNA is also recapitulated. In G1 cells, Mcd1p is degraded and its gene expression repressed, so very little if any Mcd1p is present (33, 34). Consequently, only a partial cohesin complex of the Smc1p/Smc3p heterodimers forms, and this subcomplex is not bound to chromosomes. To test whether this was also true in our system, we incubated extracts from G1 or S-phase arrested cells with C1-C and followed the binding of an epitope-tagged Smc3p to the beads (Fig. 2C). Almost no Smc3p from a G1 extract is detected in the DNA-bound fraction, whereas both Smc3p and Mcd1p from the S-phase extract are recovered in the salt-resistant DNA-bound fraction. Furthermore, in the S-phase extract, the ratio of the salt-sensitive to salt-resistant fraction of Smc3p and Scc3p is similar to the ratio observed for Mcd1p (Fig. 2C and SI Text). These results have three major implications that validate our in vitro assay. First, we recapitulate with our extracts the normal cell cycle control of cohesin binding to chromosomes. Second, we demonstrate that binding of Smc3p to DNA requires Mcd1p as it does in vivo. Third, the cofractionation of Smc3p, Scc3p, and Mcd1p in the high salt-resistant fraction validates Mcd1p as a marker for the cohesin complex and strongly supports the assembly of cohesin/DNA complexes in our in vitro system.

Another important in vivo parameter of cohesin binding to chromosomes is its dependency on binding and hydrolysis of ATP by Smc1p and Smc3p (6–8). To test whether our in vitro system also mimicked this dependency, we reduced cohesin affinity to ATP with a K38I mutation in the ATP-binding cassette of Smc3p. This mutation inhibits both ATP binding to Smc3p and in vivo cohesin binding to chromosomes (6, 8). Yet, it does not affect the integrity of the complex (6, 8). We prepared S-phase extracts...
from cells expressing Smc3p and either wild-type Smc3p-6HA or smc3p-6HA K38I, and assayed C1-C substrate binding via Western blot using HA antibodies. Although smc3p-6HA K38I binding to C1-C is comparable to the Smc3p-6HA in low salt, its ability to form salt-resistant complexes is reduced fivefold compared to the Smc3p-6HA (Fig. 2D). Our data suggest that ATP binding to Smc3p is essential for the formation of the high-affinity nucleo-protein complex consistent with in vivo analyses. However, the ability to bind at low salt in an ATP-independent manner may represent an undescribed intermediate of cohesin binding. In summary, in our in vitro assembled cohesin–DNA complexes recapitulate key attributes of in vivo assembled complexes, including salt resistance, dependence on DNA topology, and ATP binding of Smc3p.

In Vitro Assembled Cohesin–DNA Complexes Recapitulate Preference for CARs. Previous studies used generic DNA sequences to study cohesin–DNA binding properties. However, cohesin is known to enrich at CARs in vivo. We tested whether the assembly of our physiological cohesin–DNA complexes also recapitulates this key in vivo property. First, we tested if we can assemble salt-resistant cohesin/DNA complexes by using a different CAR sequence. We prepared a circular substrate L1-C that contains 3.1 kb of CAR L1 from chromosome XII (9). Similar to C1-C, cohesin showed high affinity to L1-C suggesting that cohesin association with CAR DNA in our system is not a unique property of C1-C (Fig. 3A).

Next, we replaced the CAR sequences with 3-kb sequences with chromosomal sequences that exhibit low cohesin binding in vivo. These non-CAR sequences, N1-C and N2-C, lie immediately adjacent to CAR1C or the MAT locus, respectively (Fig. 3). Then we compared the ability of C1-C, N1-C, and N2-C to form salt-sensitive and salt-resistant complexes with cohesin (Fig. 3 B and C). Salt-sensitive complexes with N1-C and N2-C are reduced 2-fold compared to C1-C, and salt-resistant complexes are reduced even further, 7.5-fold for N1-C and 15-fold for N2-C. The results show that CARs have a cis element that promotes the formation of high-affinity cohesin/DNA complex in vitro.

The Number and Position of Cohesins Bound to C1-C is Limited. In the simple ring model where the cohesin ring entraps but does not directly bind DNA, the potential number of cohesins that could load on DNA should be very large because the ring is very thin and it does not interact with specific DNA sequences. Furthermore, once loaded the cohesin should diffuse along DNA. However, in vivo, both the location and number of cohesins on chromosomes appear limited as they are enriched at CARs with approximately 3–20 cohesins per CAR (7, 10, 35). This finite ratio may simply reflect that almost all of the active cohesin in the cell is chromosome bound. Furthermore, cohesin enrichment at CARs has been proposed to be a consequence of transcription. Alternatively features of the complex itself may limit the inherent capacity of cohesins to bind CARs and to spread.

Our in vitro assay allowed us to address whether cohesin binding to DNA is limited by active cohesin molecules. First the ratio of cohesin molecules per DNA molecule in our extract is about 500:1. Incubation of C1-C in the extract longer than 30 min shows no significant increase in the total amount of cohesin/C1-C complexes. At this steady state only about 2% of the Smc3p and Mcd1p molecules in the extract are bound to C1-C beads. With this and the knowledge of the number of DNA molecules on each bead, we calculate that the ratio of bound cohesin per DNA molecule is only between 1 to 10 despite the vast excess of cohesin (see Materials and Methods). Thus the number of cohesin molecules that can bind to DNA in vitro is limited. This limitation is not a result of limiting active cohesin molecules. When we double the amount of C1-C in the reaction, it doubles the amount of cohesin binds to DNA (Fig. 4A). This limitation is likely physiologically relevant given the striking similarity between ratio of cohesin per CAR in our in vitro system and that measured in vivo.

To test further whether the number and the position of cohesins bound to C1-C are limited, we mapped cohesin binding to C1-C. We used a chromatin immunoprecipitation protocol similar to that used to map in vivo cohesin binding to chromosomes (Fig. 4B). First, cohesin loading is highly enriched on CAR1C and absent from neighboring vector sequences. Poor binding to bacterial sequences has been observed in vivo (9, 36). Second, a peak of cohesin binding occurs in a subregion of CAR1C on the C1-C substrate as it does in vivo chromosomes. The maximal peak level of cohesin cross-linked to CAR1C on C1-C (3%) is very similar to that seen in vivo (2.5–5%) consistent with the assembly of a physiological relevant cohesin/chromosome complex. The limited distribution of cohesin binding to C1-C strongly supports our conclusion based upon average stoichiometry of cohesin to DNA molecule that only one or a few cohesin can bind to C1-C.

The limited distribution of cohesin on C1-C is unlikely due to transcription or exclusion by chromatin. As shown above, C1-C does not assemble into chromatin. Transcription is unlikely to be supported by our extracts as no rNTPs are added. Furthermore, when transcription in the extract is blocked by the addition of transcription inhibitor thiolutin, no significant difference in cohesin/DNA complex assembly is observed (SI Text). Interestingly, the distribution of cohesin binding of CAR1C in our extracts is not identical to that observed in vivo. This suggests that our in vitro system may be missing some cis boundary elements or activities that control the distribution of cohesin binding within CAR1C.
The limited distribution of cohesin on C1-C suggests that the diffusion of cohesin along C1-C, if it occurs at all, is constrained after forming the salt-resistant complex with C1-C. To test this possibility we performed a standard binding reaction to allow cohesin to form a salt-stable complex with C1-C, then added PstI restriction endonuclease to the reaction. When DNA was digested, beads were immobilized and washed with low salt buffer as before (Fig. 1B). No significant difference in cohesin binding is detected after cutting the DNA, indicating that the salt-resistant complexes remain stably bound to the now linear C1-L-like substrate (Fig. 4C and D). The ability of cohesin to maintain a stable complex with the linear DNA C1-C after it is cut but not with C1-L infers that the complex matures. This maturation may involve more intimate interactions with C1-C during or after the formation of the salt-resistant complex preventing cohesin from disengaging from or diffusing along C1-C once it becomes linear.

**Dissection of CARC1 Identifies Subregions That Modulate Cohesin Affinity to DNA.** To begin to identify elements in CARC1 that contribute to formation of salt-resistant cohesin we divided CARC1 DNA into four regions marked a–d (Fig. 5A). We constructed a panel of deletions of these four regions. Formation of salt-sensitive complexes is only moderately affected by these deletions. The formation of salt-resistant complexes is reduced by about 2-fold for the aΔ-C and abΔ-C substrates compared to full-length C1-C. However, it is reduced 13-fold for cdΔ-C substrate, suggesting that a critical element for cohesin binding lies within region cd (Fig. 5B). Interestingly, region cd lies to the right of the binding peak identified by the in vitro ChIP (Fig. 4B). This suggests that sequences within CARC1 lead to asymmetric assembly of cohesin. Individual deletions of c and d regions significantly reduce cohesin binding, identifying at least two regions important for assembly of salt-resistant complex (Fig. 5C).

Next we tested the in vivo relevance of the CARC1 sequences that modulate cohesin binding in vitro. We constructed deletions in chromosome III in two steps. First, we deleted c, d or cd and replaced them with the *URA3* gene flanked by bacterial sequences. Then we selected for loss of the *URA3* gene by recombination between flanking loxP sites leaving the desired deletion with just one loxP site (Fig. 6A). We compared in vivo enrichment of cohesin at the wild-type CARC1 and deletion derivatives by chromatin immunoprecipitation (Fig. 6C).

Region cd of CARC1 lies to the right of the peak of in cohesin binding to CARC1 both in vitro and in vivo. Deleting the corresponding sequence in chromosome III (cdΔ) reduces this peak by 3.5-fold compared to wild-type cells. To exclude the possibility that reduction in cohesin association with the cdΔ is due to shortening of the distance between flanking elements such as transcription terminators, we tested the strain in which we replaced the cd region with the *URA3* gene (Fig. 6B). We compared in vivo enrichment of cohesin at the wild-type CARC1 and deletion derivatives by chromatin immunoprecipitation (Fig. 6C).
However, deletion of c and d regions individually did not impact cohesin binding in vivo, indicating that the cd region has at least two elements, either of which is sufficient for cohesin binding (Fig. 6C). In contrast, our in vitro analyses also identified two elements in cd but both are necessary (Fig. 5C). It would appear that the two elements in c and d regions are more redundant in vivo for assembly of stable cohesin complexes presumably reflecting a current limitation in our in vitro system. However, we are strongly encouraged by the fact both the in vivo and in vitro data suggest that CARC1 contains elements within both c and d that regulate cohesin assembly on DNA.

**In Vitro Assembly of Cohesin/DNA Complex Is Scc2 Dependent.** Cohesin loading on chromosomes depends on the Scc2p-Scc4p complex, yet the exact role of Scc2p is unclear. First, we tested if the in vitro assembled cohesin/DNA complexes in our system depend on Scc2p. We immunodepleted Scc2p from the extract to a about 25% of its initial level (Fig. 7A). This immunodepletion does not reduce the amount of cohesin as monitored through Mcd1p. Mock-depleted and depleted extracts showed no difference in the assembly of cohesin/DNA complex in low salt (Fig. 7B). However, the formation of salt-resistant C1-C cohesin complexes is reduced about 50% in the Scc2p depleted extract. Complementation of this defect by the addition of recombinant Scc2p is not currently feasible because we and others have been unable to purify Scc2p either directly from yeast or in a recombinant system. Because of this limitation we conclude that Scc2p and/or an associated factor is required for the efficient formation of the salt-resistant cohesin/DNA complex but not for the formation of the salt-sensitive complex.

We then asked whether the ability of different substrates to bind Scc2 might explain their relative proficiency to assemble salt-resistant complexes with cohesin. Substrates were mixed with extract made from cells containing Scc2p-3V5 and arrested in S phase, as before. After 30 min unbound proteins were washed with 0.1 M KCl and the amount of Scc2p-3V5 in the bound fraction was determined using antibodies against V5. Scc2p-3V5 is bound to beads with C1-C but not to beads lacking DNA. The binding of Scc2p-3V5 by C1-C is consistent with the functional requirement of Scc2p for formation of the stable complex (Fig. 7B). Compared to C1-C, the binding of Scc2p-3V5 to non-CAR substrates NC1-C and NC3-C is reduced by 6- and 20-fold, respectively (Fig. 7C). Indeed, this decrease in Scc2p-3V5 binding is very similar to the reduced ability of NC sequences to form salt-resistant DNA complexes with cohesin. (Fig. 3B and C). In contrast, the binding of Scc2-3V5 to the linear C1-L and the cdΔ-C is reduced 2-fold (Fig. 7D and E). This reduction is not sufficient to cause the 5- to 10-fold reduction in formation of the salt-resistant cohesin complex on C1-L and even greater reduction on cdΔ-C (Figs. 2A and 5B). Thus these two substrates reveal the existence of additional steps beyond Scc2p recruitment to assemble stable cohesin-DNA complexes.

**Discussion**

Here we describe the development of a biochemical system to study the interaction of cohesin with DNA in vitro. The in vitro assembled cohesin/DNA complex shares many physiological properties with cohesin/ chromosome complexes assembled in vivo. Both in vivo and in vitro complexes exhibit resistance to high salt (5). As predicted from in vivo analyses, the formation of stable complexes in vitro is dependent upon the topology of the DNA substrate, the presence of Mcd1p, cell cycle stage of the extract, and ATP binding of Smc3p (reviewed in ref. 1). Also as predicted from in vivo studies, cohesin–DNA complexes form preferentially on CARs (9, 10, 36). By recapitulating physiological properties, our DNA–cohesin complexes are the only in vitro assembled DNA–Smc complex to be validated as biologically relevant. Clearly this biological relevance provides a critical foundation to perform advanced mechanistic studies with high-resolution structural and single molecule methodology.

Our studies also reveal that cohesin associates with DNA by more than a single mode of DNA binding. As indicated above one mode, the salt-resistant complex exhibits all the properties of the stable in vivo complex. The slow formation of the salt-resistant complex is consistent with significant rearrangement of the cohesin structure during DNA binding. Indeed, in vivo studies infer rearrangements of the head and distal hinge domains of Smc1p and Smc3p for proper DNA binding (37, 38). The second mode, the salt-sensitive complex, forms much more rapidly and independent of substrate, topology, substrate sequence (CAR), and Scc2p. Recent studies have identified distinct populations of cohesin that have dramatically different residence time on chromosomes in vivo (18, 21). Our results suggest that the low residence
time complexes may in fact be due to our low-salt nonspecific complex. Thus in the future, it will be interesting to determine which of the cohesin chromosome complexes, defined by different binding half lives, are ScC2p dependent.

The binding of cohesin to our substrate contradicts two predictions of the simple embrace model. First, the number of cohesins that can bind to our substrate is limited to an average of 2

\[
\frac{1}{2} \text{tions of the simple embrace model. First, the number of cohesins that can bind to our substrate is limited to an average of 2–10 even when active cohesin is in vast excess in the extract. Second, cohesin is unable to bind or subsequently to diffuse randomly along our substrate as evidenced by our ChIP analysis. In agreement with the ChIP study, once the slow formation of the salt-resistant complex between cohesin and the topologically closed DNA substrate is complete, cohesin remains bound even upon linearization of the DNA.

In vivo studies to assess cohesin’s ability to diffuse along DNA have yielded different results. Introduction of a double-strand break proximal to CARs on chromosome III does not release chromatin-bound cohesin, consistent with a constraint on diffusion (17). Furthermore, the ability to constrain cohesin diffusion on chromosomes and our substrate is attractive as more and more studies implicate cohesin function in transcription regulation where regional specific modulation of chromatin is essential for specificity (23, 39). However, cleavage of the DNA of an isolated minichromosome releases about 80% of the cohesin associated with the minichromosome’s pericentric region indicating that at least some cohesin can diffuse along DNA (15, 40).

Fig. 7. In vitro assembly of cohesin/DNA complexes depends on ScC2p. (A) ScC2p-3V5 was either mock or immunodepleted from protein extract prepared from JHS257 cells arrested in S. About 25% ScC2p-3V5 remained in the depleted extract (d) in comparison with the mock-depleted extract (m). Error bars represent SD of two independent experiments. (B) Protein extract was prepared from JHS257. ScC2p-3V5 was either mock or immunodepleted and the extract was mixed with C1-C under the standard assay conditions. Bound proteins were analyzed with antibodies against Mcd1p. Statistical difference between mock and depleted samples was confirmed with a t-test. Error bars represent SD of two independent experiments. (C) Protein extract was prepared from JHS257 cells arrested in S phase and mixed with C1-C or NC1-C. The reaction was done as described. Bound proteins were analyzed with antibodies against V5. Proteins were normalized to the DNA on the beads. Error bars represent SD of two independent experiments. (D) Protein extract was prepared from JHS257 cells arrested in S phase and mixed with C1-C or C1-L. The reaction was done as described. Bound proteins were analyzed with antibodies against V5. Proteins were normalized to the DNA on the beads. Error bars represent SD of two independent experiments. (E) Protein extract was prepared from JHS257 cells arrested in S phase and mixed with C1-C or cd4-C. The reaction was done as described. Bound proteins were analyzed with antibodies against V5. Proteins were normalized to the DNA on the beads. Error bars represent SD of two independent experiments.

There are a number of possible differences to explain this apparent discrepancy. Cohesin may slowly transit from simple topological binding to a more intimate interaction as part of its maturation to a tethering complex. The closed, but not the linear, DNA substrate prevents cohesin diffusion off the ends before the maturation of the complex. Such nontopological binding could involve a surface within the cohesin ring, similar to the mechanism of DNA binding for the Smc-like Msh2 and RecR proteins (41, 42). In vitro DNA binding assays with recombinant Smc hinge domain from bacteria and mouse condensin showed nontopological DNA binding by these domains (28, 43). The fraction of cohesin in the more mature state may vary between CARs and pericentric regions. Alternatively all cohesin may have the same mode of binding but cohesin’s diffusion along DNA may be constrained specifically at euchromatic CARs by its association with conventional DNA binding proteins to regulate its function in transcription (44–46). In this model the assembly of cohesin with DNA binding proteins must also proceed slowly to explain the difference between the linear and closed substrate.

Further analysis of complex assembly with this in vitro system should provide a means to distinguish between these possibilities.

Finally, we use our in vitro system to elucidate two factors that are responsible for the preference to form physiological cohesin–DNA complexes on CARs. Although cohesin itself forms low-salt complexes with both CAR and non CAR DNA, ScC2p binds preferentially to CARs. Furthermore, depletion of ScC2p impedes the formation of salt-resistant cohesin–CAR complexes. Together these results suggest that the specificity for the formation of high salt complexes on CARs may result from the preferential binding of ScC2p to CARs. Neither ScC2p nor its partner, ScC4p, have known DNA binding motifs. Mechanisms of targeting ScC2p are suggested from studies implicating histone modifications or auxiliary factors like Sir2, Ctfc, and Atrx in the targeting of cohesin to specific chromosomal sites or regions (17, 44–49). Because our substrates lack histones, we suggest that a transcription-like factor, to be discovered, may target ScC2p to CARs in yeast, which in turn acts on cohesin to promote the formation of the salt-resistant cohesin–CAR.

A second factor that promotes formation of physiological cohesin-CAR complexes is the 1 Kb cd sequence of CAR1. We show that deletion of cd dramatically reduces the ability of CAR1 to form salt-resistant complexes with cohesin in vitro and to bind cohesin in vivo. In both cases, cd deletion reduces cohesin binding to CAR1 sequences adjacent to cd as well as cd itself. The discovery of cd suggests that sequences important for cohesin loading to CARs lie within the CARs themselves. Interestingly the deletion of cd does partially reduce ScC2p binding. These observations are consistent with a recent study localizing ScC2p-ScC4p complexes to CARs as opposed to a previous study that concluded cohesins migrate to CARs from distal loading.
The cd sequence may contain specific protein binding motifs or structures (like cruciform) necessary for Scs2p binding. Unfortunately, identification of a small sequence with potential bioinformatic utility has been confounded by our subsequent deletion analysis of cd that has revealed functional redundancy. Clearly the reduction in Scs2p binding by the cd deletion is not sufficient to explain the dramatic effect on formation of the salt-resistant cohesin-CAR complex (this study). This result implies that additional factors beyond simple binding of Scs2p regulate formation of a stable cohesin-CAR complex. The existence of these factors has eluded previous in vivo analyses and therefore, identifying the cd-dependent, Scs2p-independent factor will be an exciting future direction.


Materials and Methods

Detailed information about reagents and molecular methods used in this work are presented in *SI Materials and Methods*. We describe growth conditions for yeast, the preparation of DNA substrates, and the preparation of the yeast extracts from cells harboring wild-type and smc3-K388p as well as extracts from cells staged in G1. We provide detailed protocols for assembly and detection of cohesin-DNA complexes on magnetic beads. We describe the immunodepletion of Scs2p, the in vitro and in vivo detection of *CAR1* sequences, and the detection of cohesin binding to these DNA variants both in vitro and in vivo by chromatin immunoprecipitation.

ACKNOWLEDGMENTS. We thank Vinny Guacci, Jill Heidinger-Pauli, Fred Tan, and all other members of the Koshland lab for constructive comments and technical support. This work was funded by Howard Hughes Medical Institute and National Institutes of Health Grant GM092813.
Supporting Information

Onn and Koshland 10.1073/pnas.1107504108

SI Materials and Methods

Yeast Strains, Media, and Cell Synchronization. Yeast strains used in this study are listed in Table S1. All strains are derivatives of the A364A genetic background. Plasmid information is also given within the strain table, written as a part of the relevant strain containing the plasmid. Exponentially dividing cultures were arrested in G1 using 15 mM alpha factor (Sigma) or in S using 130 mM hydroxyurea (HU) (Sigma) or G2/M using 1.5 mg/mL nocodazole (Sigma) for 3 h in YePD (1% yeast extract, 2% peptone, and 2% glucose, and 0.01 mg/mL adenine).

DNA Substrates. pIO1 was constructed by inserting a 2.09-kb Smal-KpnI fragment containing LacO repeats into pTNT (Promega). Yeast chromosomal DNA fragments described in Table S2 were amplified by PCR from yeast genomic DNA and cloned into pIO1 sequentially to the lacO array KpnI site. Substrates were prepared by PCR with one of the following sets of pTNT universal primers: 5′-biotin-tg-GGCTAGACTTAAATGCT AC and 5′-GTATAGGCCCCAGGGGTGTATGC (linear substrates, -L); Or 5′-biotin-tg-GGCTAGACTTAAATGCTACGTCAC and 5′-biotin-tg-GTTAGAGGCCCCAAGGGGTATGC (closed substrates, -C).

Substrate Preparation. Streptavidin-coupled Dynabeads (M-280, Invitrogen) were washed 3 times with DBB (50 mM Tris-HCl (pH 7.5, 1 M NaCl, 1 mM EDTA)) and resuspended in 1 volume of DBB. Five hundred nanograms biotin-labeled DNA was added for every 20 μL beads and beads were incubated at room temperature for 2 h with gentle shaking. For no DNA control, 5 μL of 0.1 M Spermidine (Fluka) was added to the beads. Beads were washed 5 times with PBB (50 mM Hepes pH 8.0, 10% glycerol, 100 mM KCl 2.5 mM MgCl2, 0.4% NP-40, 5 mM 2-mercaptoethanol, 50 mM Na Botyrate, protease inhibitor cocktail (Roche)) and resuspended in 55 μL PBB for every 20 μL beads. Five microliters of BSA 100 mg/mL (Sigma), poly-dIdC-dIdC (Sigma) and acetyl-coA (Roche) were added for every 20 μL beads and beads were split to 70-μL reactions.

Protein Extract Preparation and Cohesin Binding. Cells were grown in 5 mL YePD for 8–10 h at 30 °C, diluted into 0.5 L, and grown for 16 h at 30 °C to a density of approximately 0.6 × 10^7 cells/mL then arrested in S (unless mentioned otherwise). Cells were washed once with water and pellets from 1 × 10^7 cells were frozen in liquid nitrogen and stored at −80°C. To prepare extracts, cell pellets were thawed briefly on ice. Three hundred fifty microliters cold PBB and chilled 0.6 g of 425–600 μm glass beads (Sigma) were added, then cells lysed by bead beating at 3 × 30 s pulses using the Biospec™ mini bead beater. The soluble fraction was separated from the insoluble fraction by two consecutive centrifugations. Depleted extract was used to assemble cohesin/DNA complexes as described above.

In Vitro Deletion. Δ1a and Δ1b were created by PCR of pIO2 with one universal oligo 5′-biotin-teg-GTTTAGAGGCCCCACAAGGGGTATGC and an internal oligos 5′-biotin-tg-A GTGAGGGACTAAAACG C T AACTACTG or GAACTGCGCTGTTGGTGGCGCCAGAG, respectively. pIO2 was constructed by cutting pIO2 with KpnI and MluI, fill-in and re-ligation. Δ1c was amplified from pIO2 as described under DNA Substrates.

Half kilobase deletion was done by QuikChange II XL site-directed mutagenesis kit (Stratagene) according to the manufacturer’s procedure with the following oligos: Δ2a: 5′-TGCTCTCTTGAGTTGGTTGTATGCTGCAGCGTGGTTGGTGGTGAGCCAGAGGACCAAtAGCAACAAATCTTTATTTAAGTACGCG, Δ2b CGCCAAAATCTTTAAGTACGCG, Δ2c 5′-TGCTCTCTTGAGTTGGTTGTATGCTGCAGCGTGGTTGGTGGTGAGCCAGAGGACCAAtAGCAACAAATCTTTATTTAAGTACGCG, Δ2d CGCCAAAATCTTTAAGTACGCG, Δ2e 5′-TGCTCTCTTGAGTTGGTTGTATGCTGCAGCGTGGTTGGTGGTGAGCCAGAGGACCAAtAGCAACAAATCTTTATTTAAGTACGCG, Δ2f 5′-TGCTCTCTTGAGTTGGTTGTATGCTGCAGCGTGGTTGGTGGTGAGCCAGAGGACCAAtAGCAACAAATCTTTATTTAAGTACGCG.

In Vivo Deletion and ChIP. Deletion of chromosome III coordinates 101,220–102,333 was done by PCR-based integration of hisG:URA3::hisG cassette flanked with LoxP sites into the genome to create YIO604/Δcd::URA3. To pop out the URA3 insertion, cells were transformed with pFVL57 Cre recombinase and grown in YeP supplemented with 2% galactose (Sigma). The deletion was confirmed by PCR and sequencing. Δ2a (chromosome III coordinates 101,282–101,789) and Δ2b (chromosome III coordinates 101,739–102,333) were constructed as above. ChIP was performed as described before (2).

In Vitro ChIP. Cohesin binding assay was performed as described above. After the last wash with PBB, beads were resuspended in 100 μL PBB and proteins were cross-linked to the DNA by adding formaldehyde to a final concentration of 1% for 30 min at room temperature. Beads were washed 3 times with PBB and resuspended in 0.5 mL PBB. DNA was sheared to about 500-bp fragments by six sonication bursts for 10 s each, using the Branson 250 sonicator with a 3/16” tapered microtip, at constant duty cycle and L5 output setting. Beads were immobilized on a magnet, and the soluble fraction was transferred to a new tube and the total volume was increased to 1 mL with PBB and 5 μL of anti-HA antibodies (12CA5, Roche) was added. Samples were rotated over night at 4°C. Antibodies were collected on 30 μL protein A agarose beads (Roche), consecutively washed for 5 min. with 1 mL of PBB, PBB0.5, LiCl detergent (0.25 M LiCl 1%, NP-40, 1% DOC, 1 mM EDTA, 10 mM Tris-HCl pH 8.1), and twice with TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). Proteins were eluted in 300 μL (1% SDS, 0.1 M NaHCO3). Reversal of cross-linking...
and recovery of DNA were performed following the in vivo ChIP procedure above. One reaction was used as input control. Beads were resuspended in 0.5 mL PBB, sonicated, and treated as described above and kept at −20°C. Reversal of cross-linking and DNA recovery was performed by following the in vivo ChIP procedure above.


**Fig. S1.** DNA labeled with 2 biotins form a substrate with closed topology. (A) Schematic of linear and closed substrates of PstI (P) and XhoI (X) restriction sites and the expected fragment sizes in kb are indicated. (B) The DNA substrates C1-C and C1-L (Fig. 1B) were cleaved with Pst I (P) or Pst I and Xho I (P + X). Beads were collected on a magnet and the DNA content of the soluble fraction was analyzed by agarose gel electrophoresis.

**Fig. S2.** DNA substrates are not degraded. (A) C1-L and C1-C were recovered from beads after incubation with protein extract from 3131-13V under the standard reaction conditions. (B) Protein extract was prepared from JH5257 cells arrested in S and mixed with C1-L or C1-C under the standard assay conditions with the following modification: Unbound proteins were washed with 0.1 M KCl. Bound proteins were analyzed with antibody against histones H2B and H3 (Active Motif). Proteins were normalized to the DNA on the beads. (C) Protein extract was prepared from VG3176-2A cells arrested in G1 or S and mixed with C1-C under the standard assay conditions. Bound proteins were analyzed with antibodies against HA (Scc3p-3HA) and Mcd1p (Lower).

**Fig. S3.** Cohesin enrichment on MAT CAR is not affected by CARC1 deletion. ChIP analysis of the MAT CAR from 2185-6B, YIO6041, and YIO604.
**Fig. S4.** Transcription does not affect cohesin binding in vitro. Protein extract was prepared from YIO71 cells arrested in S and mixed CC1-C under the standard assay conditions with the following modification: Extract was supplemented with 3 μg/mL thiolutin (Axxora). Bound proteins were analyzed with antibodies against Mcd1p.

Table S1. Yeast strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3131-V13</td>
<td>MATa pep4Δ::g418::TRP1 trp1-1 leu2-3,112 ura3-52 his3-11,15 bar1 GAL+</td>
</tr>
<tr>
<td>YIO5031</td>
<td>MATapep4Δ::g418::TRP1 trp1-1 leu2-3,112 ura3-52 his3-11,15 bar1 GAL+ SMC3 SMC3-6HA::URA3</td>
</tr>
<tr>
<td>YIO5032</td>
<td>MATa pep4Δ::g418::TRP1 trp1-1 leu2-3,112 ura3-52 his3-11,15 bar1 GAL+ SMC3 SMC3-6HA K38I::URA3</td>
</tr>
<tr>
<td>YIO71</td>
<td>MATa pep4Δ::g418::TRP1 trp1-1 leu2-3,112 ura3-52 his3-11,15 bar1 GAL+ SMC3Δ::clonat SMC3-6HA::URA3</td>
</tr>
<tr>
<td>2185-6B</td>
<td>MATa ura3-12 lys2-801 his3Δ200 bar1 gal MCD1-6HA</td>
</tr>
<tr>
<td>YIO604/Δcd</td>
<td>MATa ura3-12 lys2-801 his3Δ200 bar1 gal MCD1-6HA URA3::Δchr.III101,220-102,333</td>
</tr>
<tr>
<td>YIO6041/Δcd</td>
<td>MATa ura3-12 lys2-801 his3Δ200 bar1 gal MCD1-6HA Δchr. III 101,220-102,333pFVL57::HIS3</td>
</tr>
<tr>
<td>YIO6051/Δc</td>
<td>MATa ura3-12 lys2-801 his3Δ200 bar1 gal MCD1-6HA Δchr.III 101,282-101,789 pFVL57::HIS3</td>
</tr>
<tr>
<td>YIO6052/Δd</td>
<td>MATa ura3-12 lys2-801 his3Δ200 bar1 gal MCD1-6HA Δchr. III 101,739-102,333 pFVL57::HIS3</td>
</tr>
<tr>
<td>YIO971</td>
<td>MATa pep4Δ::g418::TRP1 trp1-1 leu2-3,112 ura3-52 his3-11,15 bar1 GAL+ MCD1-TEV-3HA SMC3Δ::clonat SMC3-6HA::URA3</td>
</tr>
<tr>
<td>JHS257</td>
<td>MATa LacO(DK)-NAT:Lys4 pHIS3-GFP Lacl-HIS3::his-3-11,15 trp1-1 leu2-3,112 ura3-52 bar1 GAL+ SCC2-3V5-G418</td>
</tr>
<tr>
<td>VG3176-2A</td>
<td>MATa SCC3-3HA-H35MX trp1-1 leu2-3,112 ura3-52 lys2-801 his3-11,15 bar1 GAL+</td>
</tr>
</tbody>
</table>

Table S2. In vitro substrates

<table>
<thead>
<tr>
<th>Name</th>
<th>Coordinates</th>
<th>Plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC1</td>
<td>Chr. III 99,096–102,380</td>
<td>pIO2</td>
</tr>
<tr>
<td>CC2</td>
<td>Chr XII 441,109–444,191</td>
<td>pIO30</td>
</tr>
<tr>
<td>N1</td>
<td>Chr. III 95,500–98,488</td>
<td>pIO15</td>
</tr>
<tr>
<td>N2</td>
<td>Chr. III 204,025–206,994</td>
<td>pIO25</td>
</tr>
<tr>
<td>N3</td>
<td>pBR322 45-4215</td>
<td>pIO29</td>
</tr>
</tbody>
</table>