One-sentence summary: Solving the structure of the hinge domain of an epigenetic regulator highlights residues involved in dimerization and function.

Editor's summary:
Illuminating dimerization
Proteins of the SMC family are epigenetic regulators involved in sister chromatid cohesion, chromosome condensation, and DNA repair. Unlike other family members, SMCHD1 forms homodimers rather than heterodimers and has a distinct domain architecture. Dysregulation of SMCHD1 function results in a form of muscular dystrophy and a developmental disorder. Chen *et al*. solved the x-ray crystal structure of the SMCHD1 hinge domain, which is important for homodimerization and nucleic acid binding. Site-directed mutagenesis studies identified critical residues involved in SMCHD1 function in cells. Together, these data suggest how mutations in the SMCHD1 hinge domain contribute to human disease.

Crystal structure of the hinge domain of Smchd1 reveals its dimerization mode and nucleic acid–binding residues

Kelan Chen[1,2*], Richard W. Birkinshaw[1,2*], Alexandra D. Gurzau[1,2*], Iromi Wanigasuriya[1,2], Ruoyun Wang[1,2], Megan Iminitoff[1,2], Jarrod J. Sandow[1,2], Samuel N. Young[1], Patrick J. Hennessy[1], Tracy A. Willson[1,2], Denise A. Heckmann[1,2], Andrew I. Webb[1,2], Marnie E. Blewitt[1,2,3†], Peter E. Czabotar[1,2†], James M. Murphy[1,2†]

[1]The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Melbourne, VIC 3052, Australia. [2]Department of Medical Biology, University of Melbourne, Melbourne, VIC 3010, Australia. [3]School of Biosciences, University of Melbourne, Melbourne, VIC 3010, Australia.

[*]These authors contributed equally to the work.

[†]Corresponding author. Email: jamesm@wehi.edu.au (J.M.M.); blewitt@wehi.edu.au (M.E.B.); (czabotar@wehi.edu.au) (P.E.C.)

ABSTRACT

Structural maintenance of chromosomes flexible hinge domain containing 1 (SMCHD1) is an epigenetic regulator in which polymorphisms cause the human developmental disorder, Bosma arhinia and micropthalmia syndrome, and the degenerative disease, facioscapulohumeral muscular dystrophy. SMCHD1 is considered a noncanonical SMC family member because its hinge domain is C-terminal, because it homodimerizes rather than heterodimerizes, and because SMCHD1 contains a GHKL- rather than an ABC-type

ATPase domain at its N-terminus. The hinge domain has been previously implicated in chromatin association; however, the underlying mechanism involved and the basis for SMCHD1 homodimerization are unclear. Here, we used X-ray crystallography to solve the three-dimensional structure of the Smchd1 hinge domain. Together with structure-guided mutagenesis, we defined structural features of the hinge domain that participated in homodimerization and nucleic acid binding, and we identified a functional hotspot required for chromatin localization in cells. This structure provides a template for interpreting the mechanism by which patient polymorphisms within the SMCHD1 hinge domain could compromise function and lead to facioscapulohumeral muscular dystrophy.

INTRODUCTION

Structural maintenance of chromosomes (SMC) proteins are an evolutionarily conserved family responsible for regulating chromatin and chromosomal organization (1). Conventional SMC proteins contain several structural motifs, including a bipartite ABC-type ATPase domain comprising subdomains contributed by the N- and C-termini, a central-located hinge domain, and a coiled-coil arm formed by the two intervening segments (1-3). Whereas prokaryotic SMC proteins exist as homodimers, there are three exclusive pairs of SMC heterodimers in eukaryotes for which the hinge domain constitutes the dimerization interface. Together with auxiliary subunits, SMC proteins form large protein complexes that play distinct roles in sister chromatid cohesion, chromosome condensation, and DNA repair.

SMCHD1 is considered a noncanonical member of the SMC protein family because it has a domain architecture that is distinct from those of other SMC proteins, including an N-terminal GHKL-type ATPase domain (4-6) and a hinge domain located at the C-terminus (7-9). Disruption of SMCHD1 function is as an underlying factor in two human disorders: the muscular dystrophy, facioscapulohumeral dystrophy (FSHD) (10-12); and the developmental disorder, Bosma arhinia micropthalmia syndrome (BAMS) (10, 13, 14). Therefore, there is increasing interest in uncovering how SMCHD1 functions in development and pathogenesis. At the molecular level, our previous studies demonstrated that SMCHD1 occupies defined genomic regions to exert epigenetic silencing effects (15, 16). We and others reported that Smchd1 mediates long-range chromatin interactions and appears to insulate the chromatin from the action of other DNA-binding proteins (16-18). The hinge domain plays a crucial role in exerting these functions, because it is this domain that mediates SMCHD1 homodimerization and nucleic acid binding in vitro (9, 15). This functionality is clearly critical, because when compromised by the R1866G substitution present in an FSHD kindred, the silencing capacity of SMCHD1 is reduced (9, 15). In contrast to the antiparallel intramolecular coiled-coil present in canonical SMC proteins, our previous SAXS studies indicate that SMCHD1 adopts an intermolecular parallel coiled-coil arrangement flanking the hinge domain, consistent with a predicted head-to-head arrangement in full-length SMCHD1 dimers (9). However, the atomic-resolution structure of the SMCHD1 hinge domain, and the mechanism underlying its association with nucleic acids, remain to be elucidated.

Several crystal structures of SMC hinge dimers have been previously reported, including those of the prokaryotic SMC homodimer (19), the SMC1-SMC3 heterodimer of the cohesin complex (20), SMC2-SMC4 of the condensin complex (21), and the SMC5-SMC6 heterodimer (22). In this study, we present the crystal structure of the homodimer of the hinge domain of mouse Smchd1, which adopts a fold grossly resembling the canonical SMC hinge and forms a homodimer through a similar dimerization interface. Smchd1 possesses distinctive features compared to those of canonical hinge domains, such as fewer, longer helices, the absence of a pair of β-strands at the interface between the two subdomains, and different secondary structural elements at the termini. By performing structure-guided mutagenesis, we defined the dimerization interface, implicated the extended linker region and residues surrounding the "central channel" in nucleic acid binding, and demonstrated the importance of the hinge domain and a functional hotspot within this domain for chromatin localization in cells. Together, our findings provide insights into SMCHD1 function at the molecular level and provide a basis for interpreting how mutations in the SMCHD1 hinge domain might affect function and contribute to human disease.

RESULTS

The Smchd1 hinge domain contains an unconventional subdomain I structure

We solved the structure of the mouse Smchd1 hinge domain (residues 1683 to 1899) by X-ray crystallography to 3.3 Å resolution following anomalous phasing using mercury derivatized crystals (fig. S1). The overall architecture of the Smchd1 hinge domain has similarities to previously published SMC hinge domain dimers, forming a doughnut shape

with the N- and C- termini on the same face (Fig. 1, A-D). A DALI structure search identified the *Thermotoga maritima* SMC (*Tm*SMC) hinge domain structures as the closest structural homologs of the Smchd1 hinge domain (PDB 1GXK and 1GXJ; RMSD: 3.1 Å aligning 131 Cα and 3.7 Å aligning 131 Cα, respectively) (19, 23). The *Tm*SMC hinge was the first hinge domain structure reported, and, as a homodimer, 1GXJ provides a reference point for discussing the features of the Smchd1 structure. Similar to the *Tm*SMC and other canonical SMC proteins, the SMCHD1 hinge monomer comprises subdomains I and II connected by a flexible linker. The *Tm*SMC subdomains are pseudosymmetric, with subdomain I containing a three-stranded β-sheet (S1 to S3) and subdomain II a five-stranded β-sheet (S4 to S8) with a pair of β-strands (SI and SI′) contributing to an auxiliary intersubdomain interface (Fig. 1E to G). Within the *Tm*SMC homodimer, β-sheets contributed by subdomains I and II of dimer-opposed protomers contribute to a composite extended β-sheet. In contrast, Smchd1 lacks the N-terminal strand, with only S1 and S2 from subdomain I contributing to the inter-subdomain β-sheet. In addition, it lacks the two antiparallel β-strands (SI and SI′) that are central to the intersubdomain interface in *Tm*SMC, which are replaced by extended subdomain interface helices in Smchd1 (Fig. 1B and C). Furthermore, the Smchd1 hinge domain is distinct because of a short loop insertion in S7 that generates an additional β-strand, S8, at the C terminus. As a result, S7 loops out of the plane of the inter-subdomain β-sheet and instead S8 forms a β-sheet with strands S3 to S6 from subdomain II as part of the inter-subdomain extended β-sheet. There are also differences between the α-helices present in the *Tm*SMC and Smchd1 hinge domains. In the *Tm*SMC hinge, subdomain I is flanked by H1 to H2 and H6, and subdomain II by H11 (Fig. 1F). However, in Smchd1,

the equivalent of the helices H1 to H2 is not observed in the density even though the corresponding region was present in the construct crystallized. As a result, the subdomain I and II β-sheets in the hinge domain of Smchd1 are flanked by only one α-helix denoted H3 in subdomain I and H6 in subdomain II (Fig. 1, B and C), which relative to their counterparts in *Tm*SMC, H6 and H11, are shortened and lengthened by one helical turn, respectively. Grossly, the inter-subdomain orientations of the Smchd1 and *Tm*SMC hinge domains differ markedly, with a rotation of the Smchd1 subdomain II by 18 or 27° relative to the closed (1GXK) or open (1GXJ) *Tm*SMC reference structures. Such a rotation places the two subdomains I further apart and the two subdomains II closer within the Smchd1 hinge dimer (Fig. 1, B, D, F, H and I).

Our previous small-angle X-ray scattering (SAXS) analyses of extended hinge domain constructs (9), and an electron microscopy study of full-length SMCHD1 (4), support a role for the coiled-coil regions immediately flanking the hinge domain in homodimerization in a head-to-head orientation. However, these coiled-coil regions were not observed in our crystal structure. The construct subjected to crystallization encompassed residues 1683 to 1899; however, density was scarce for the N-terminal (1683 to 1709) and C-terminal (1887 to 1899) regions despite mass spectrometry analysis confirming that the crystallized protein was intact (fig. S2A). As a result, densities for between 27 and 32 amino acids at the N-terminus and 13 to 15 at the C terminus were unresolved, varying among the six hinge domain copies in the asymmetric unit. It was possible to assign two additional helices in the electron density of 12 and 21 amino acids within the asymmetric unit (fig. S2, B and C). However, because these helical segments contained no

selenomethionines and had poor sidechain density, sequence assignment was not possible, and they were modelled as polyalanine helices. The absence of sequence information prevented us from deducing their connectivity to the hinge domain, and it remains unclear why only two additional helices were present in an asymmetric unit containing six copies of the hinge domain. One possibility is that these helices could contribute to the coiled-coil regions (comprising residues 1683 to 1709 and residues 1885 to 1899) predicted to flank the core hinge domain. Whether these flanking coiled-coils adopt a closed conformation, as observed in structures of the yeast SMC2-SMC4 complex and *Pyrococcus furiosis* SMC hinge domain (21), or an open conformation, as seen in the *Geobacillus stearothermophilus* (24) and *Escherichia coli* MukB SMC hinge domain (25) structures (fig. S3), in the structure presented herein remains unclear in the absence of ordered coiled-coils.

The Smchd1 hinge domain differs from other SMC hinge domain structures

Several structures of eukaryotic heterodimeric SMC complexes have been reported, including mouse cohesin SMC1-SMC3 (20), human condensin SMC2-SMC4 (26), and yeast SMC5-SMC6 (22) (fig. S4, A to H). The sequence of the Smchd1 hinge domain shows little homology to these proteins, with amino acid identity ranging from 10 to 20%. Structural alignments identified mouse SMC3 as the closest heterodimeric SMC hinge domain structural homolog, with an RMSD of 3.4 Å (aligning 124 Cα), despite a modest 12% sequence identity for the aligned regions. Human SMC4 was the next most similar structurally, with an RMSD of 3.8 Å (aligning 135 Cα) and a 22% sequence identity.

Although the overall fold of Smchd1 is similar to that of the prokaryotic *Tm*SMC and its eukaryotic orthologs SMC1 to SMC6, the hinge domain of Smchd1 possesses several divergent structural features. The loss of H1/H2 and the N-terminal β-strand from the Smchd1 hinge is a distinguishing feature, because whereas H1 is absent from human SMC2 (26), concurrent loss of the N-terminal β-strand has not been observed in any structures reported to date (fig. S4, A to I). Furthermore, there are differences between the inter-subdomain interfaces between Smchd1 and canonical SMC hinge domains. In Smchd1, the subdomain I helices at the interface, H1 and H2, are extended by 1.5 and 1 turns, respectively, relative to their counterparts in *Tm*SMC (H3 and H4; Fig. 1, E and F, Fig. 2), with the equivalent of H5 absent from Smchd1. The subdomain II helices at the interface are conserved in Smchd1, but H5 is one turn shorter than H10, its counterpart in *Tm*SMC (Fig. 1, C and G). Another distinguishing feature of the Smchd1 hinge domain is the extended linker region that bridges subdomains I and II (residues 1778 to 1805) between S2 and S3 (Fig. 1, C and G, and Fig. 2). Among copies of Smchd1 in the asymmetric unit, this was the structural element that exhibited the greatest heterogeneity, consistent with loop flexibility. Whereas the length of this linker varies between SMC hinge domains from one residue shorter in human SMC2 to ten residues shorter in *Tm*SMC (Fig. 2), it is notable that the linker of Smchd1 possesses the greatest abundance of positively charged amino acids, suggesting a possible role in interactions, such as with nucleic acids.

The Smchd1 hinge dimer interfaces comprise canonical and noncanonical interactions

The Smchd1 hinge domain exists as an obligate homodimer with two-fold rotational symmetry (Fig. 1, B and F). As for canonical SMC proteins, dimerization is mediated by two key interfaces between SMC pairs (fig. S4). In Smchd1, two helices contributed by subdomain I (H3) and subdomain II (H6) of opposing Smchd1 protomers form the first interface, which includes a salt bridge between Arg^{1762} (R1762) of H3 and Asp^{1842} (D1842) of H6 (Fig. 3A and fig. S4A). These helices cap a second interface comprising an extended β-sheet formed by β-strands S1 and S2 from the subdomain I of one protomer (which also contributes H3 to the dimer interface) and β-strands S3 to S8 of subdomain II from the symmetrical protomer. This interaction involves backbone-backbone hydrogen bonds between Val^{1774} (V1774) of S2 in subdomain I and Phe^{1874} (F1874) of S7 in subdomain II of the dimeric partner (Fig. 3A). Broadly, a similar interface is formed in *Tm*SMC in which H6 of subdomain I of one protomer and H11 from the other form a cap over a β-sheet containing an additional β-strand in addition to those in Smchd1 (Fig. 1F).

Furthermore, in addition to the canonical dimer interfaces noted earlier, several distinct interactions may also contribute to the Smchd1 hinge dimer interface. A noncanonical β-strand, S8, present at the C-terminus of Smchd1 forms part of the extended β-sheet with strands S1 to S6, which displaces the S7 strand out of the plane and in so doing imparts a defining feature of the SMCHD1 hinge domain (Fig. 1, C and D, and fig. S4, A and B). This feature is likely stabilized by distinct contacts between H3 of subdomain I and S6 and S7 of subdomain II from the other protomer. These contacts include a hydrogen bond between the hydroxyl of Tyr^{1765} (Y1765) in H3 and the S6 Asp^{1866} (D1866) carboxylate, and aliphatic interactions between Tyr^{1765} of H3 and Phe^{1874} of S7 (Fig. 3A). Moreover,

S7 not only mediates intersubunit interactions with S2 and H3, but also interacts with H2 from the other protomer through a salt bridge between the side chains of Lys^{1873} (L1873) of S7 and Asp^{1749} (D1749) of H2 (Fig. 3A). As is typical of other SMC hinge domains, Smchd1 contains a triglycine motif at the C terminus of the domain that is involved in intersubunit interactions and houses the aforementioned Lys^{1873} and Phe^{1874}. Note, however, that the Smchd1 motif is GX_6GX_2GG rather than the more typical SMC motif GX_6GX_3GG. This difference contributes to the kink between S7 and S8, which in turn is an important determinant of the unconventional interaction between protomers. Furthermore, Arg^{1848} (R1848) within H6 of subdomain II interacts with the sidechain of Ser^{1779} at the start of the extended linker of the opposing monomer, potentially strengthening its canonical interaction with H3 of subdomain I from the other protomer.

Mutagenesis of dimer interface residues does not alter the overall structure of the Smchd1 hinge dimer

To examine whether residues located at the crystallographic dimer interfaces were important to dimer formation, we generated ten mutants in which interface residues were substituted, including a triple mutant replacing all three glycine residues at the consensus site. All of the mutants were successfully expressed and purified analogously to the wild-type (WT) protein, with all mutants eluting from size-exclusion chromatography (SEC) as a single peak that resembles the profile of the WT hinge dimer (fig. S5). This indicates that all interface mutants remained dimeric in solution under these purification conditions.

To confirm the dimerization state and examine gross structural changes caused by these mutations, we used SAXS to study five representative mutants: Y1765A, V1774G, R1848A, K1873A, and G1872A/G1875A/G1876A from within H3, S2, H6, S7, and loop regions flanking S7, respectively. Background-subtracted scattering profiles of WT and mutant Smchd1 hinge domains (Fig. 3, B to G, top) were subjected to Guinier (Fig. 3, B to G, top, inset) and interatomic distance distribution (Fig. 3, B to G, bottom) analyses. The linearity of the Guinier plots indicates that the samples were monodisperse, enabling us to perform more detailed analyses of the size and shape of the mutants. Our previous studies established that the radius of gyration values (R_g) of WT Smchd1 is ~27.0 Å, with a maximum particle dimension (D_{max}) of 100 Å (9). These parameters correspond to a homodimeric configuration, as validated with orthogonal analytical ultracentrifugation studies (15). Here, we observed similar parameters for the Smchd1 hinge domain mutants (Fig. 3, C to G, and table S1. Together with our SEC data, our SAXS analyses support the idea that Smchd1 interface mutants form stable homodimers in solution, with overall sizes and shapes grossly resembling those of the WT protein.

Dimer interface residues contribute to protein thermal stability

We reported previously that a mutation related to FSHD2, R1867G (R1866G in humans), did not change the overall conformation of the Smchd1 hinge domain homodimer, but reduced the thermal stability of the protein as assessed by differential scanning fluorimetry (DSF) (9). Because the residue neighboring Arg^{1867}, Asp^{1866}, mediates dimerization contacts, we postulated that mutation of the dimer interface residues would affect dimer stability. The melting temperature (T_m) for WT and R1867G mutant dimers

were measured as 53 and 42°C, respectively (Table 2; fig. S6), comparable to our previous results (9), whereas most of the dimer mutants showed compromised stability in parallel assays. The exceptions were D1842A, which exhibited a T_m comparable to that of the WT dimer, and K1873A, which showed an increased thermal stability (T_m: 68°C). Six of the eight mutants that showed decreased thermal stability exhibited T_m values between 41 and 47°C, consistent with a role for the mutated residues in stabilizing the dimer contacts proposed based on our hinge domain crystal structure. Introduction of V1774G or the triple glycine mutant even further decreased thermal stability, with these mutants having T_m values of 35 and 38°C, respectively.

The introduction of glycines into ordered regions destabilizes secondary structures (31). Additionally, V1774G mutation within S2, will result in loss of van der Waals interactions to the adjacent Phe^{1874} in S7 from the other protomer at the dimerization interface. The combination of these two factors likely contributes to the observed decrease in the thermostability of the V1774G mutant. Note that the D1842A mutation did not compromise thermal stability, but mutation of a residue with which it forms a salt bridge, Arg^{1762} (fig. S1D), did. Whereas Asp^{1842} is involved exclusively in an interaction with Arg^{1762}, the R1762 mutation would be expected to abrogate multiple interactions, which would have a deleterious effect on thermal stability. For example, Arg^{1762} is sandwiched between Asp^{1766} and Asp^{1842}; thus, loss of Arg^{1762} would be expected to destabilize the protein due to the repulsion between the two negatively charged aspartates. Additionally, mutation of Arg^{1762} would preclude the formation of a charged pair with Asp^{1766}, one helical turn away, and likely contributes to destabilization of this helix. Because Gly^{1872}, Gly^{1875}, and Gly^{1876}

are in loop regions flanking S7, their mutation to alanine likely destabilizes the S7 to S8 loop by interfering with backbone hydrogen bonding to Gln^{1878} and Asn^{1879} (fig. S4B), thereby reducing the thermostability. Therefore, these results support an integral role for multiple intersubunit contacts, including between residues in S2 and S7, in stabilizing hinge dimers. Although it appears that, overall, these mutations do not block dimer formation, likely due to a role for the coiled-coil regions flanking the core hinge domain, destabilization of the homodimer was evident from the reductions in thermal stability.

Note that the substitution of another S7 residue, Lys^{1873}, with alanine unexpectedly increased the T_m to ~68 °C. The side chain of Lys^{1873} is orientated in parallel to the positively charged side chain of the neighboring Arg^{1869} in the adjacent S6 strand, with both side chains directed towards the central cavity of the Smchd1 dimer. We hypothesized that the K1873A substitution might eliminate an electrostatic repulsive interaction between Lys^{1873} and Arg^{1869}, thus increasing thermal stability. We tested this possibility by generating an R1869A mutant, which we found also exhibited an increase in T_m to that of K1873A (Table 2). Therefore, the observed stabilizing effect of these alanine substitutions is likely explained by attenuation of repulsive steric interactions between the positively charged side chains.

The Smchd1 hinge domain binds to nucleic acids through two positively charged patches

Analysis of the electrostatic surface potential of the Smchd1 hinge domain structure revealed the domain to be highly basic in charge, as befitting a domain previously shown

to bind to nucleic acids in vitro (15). We identified three clusters of positive charge: cluster 1, cluster 2 on the opposite face, which is contributed by the extended linker, and cluster 3 surrounding the central channel (Fig. 4A and B). We substituted the positive residues in these clusters with alanine to deduce nucleic acid binding residues using a fluorescence polarization assay. None of the tested cluster 1 mutants had a marked bearing on nucleic acid binding (Fig. 4C). In contrast, the mutations R1790A, R1796A, and K1799A in cluster 2 led to respectively 5-, 2.8- and 2.2-fold reductions in nucleic acid–binding affinity, suggesting that residues within cluster 2 contribute to nucleic acid binding (Fig. 4D). Our previous work using EMSA and analytical ultracentrifugation assays established the importance of the FSHD2-associated substitution, R1867G, in DNA binding (15). Here, using a fluorescence polarization assay, we validated the importance of Arg^{1867} (residing in cluster 3) for nucleic acid binding (Fig. 4, A and E, and Table 2). Our assay examined the binding of the recombinant Smchd1 hinge domain to a single-stranded DNA sequence containing the HS1-b binding site sequence harboring a methylated cytosine, although we would expect comparable binding to double-stranded DNA and unmethylated DNA based on our previous studies (15). In addition to R1867G, the cluster 3 mutants, R1869A and K1873A, exhibited respectively 15- and 5-fold reductions in nucleic acid binding affinity relative to that of the WT Smchd1 hinge domain. Note that among the interface residues, besides K1873A, which also contributes to cluster 3, the only tested mutations that showed deficits in DNA binding were R1762A and R1848A, with 2.1- and 7.6-fold reductions in K_d, respectively (Fig. 4F). This suggests that these residues perform dual roles in mediating interprotomer contacts and ligand binding. We then examined whether binding might occur through the torus of the hinge domain

dimer by introducing mutations to occlude DNA from entering the central hole of the dimer (Fig. 4G). We selected Ser^{1870} for substitution because it is not engaged in crucial structural interactions in our dimer structure and we chose His^{1856} because it lies on the rim of the torus. Introduction of methionine, asparagine, or glutamine in place of Ser^{1870}, or substitution of His^{1856} with the bulkier residue tryptophan, did not measurably affect nucleic acid binding (Fig. 4G), consistent with nucleic acid binding not occurring by threading through the central channel. Together, these analyses implicate two sites surrounding the Smchd1 hinge domain torus in nucleic acid binding (Fig. 4, B and H to J): binding site 1 centered around the linker (cluster 2) and binding site 2 on the opposite face (cluster 3).

The hinge domain targets Smchd1 to chromatin

Mutational studies identified clusters 2 and 3 and the interface residue, Arg^{1848}, as important contributors to nucleic acid binding in vitro. It remained of interest whether these residues contributed to the cellular function of Smchd1. Accordingly, we introduced the cluster 2 mutations, R1790A and R1796A, the interface mutations, D1842A and R1848A, and the cluster 3 mutations, R1867G, R1869A and K1873A, into full-length mouse Smchd1 and transiently expressed these constructs in HEK 293 cells in which endogenous SMCHD1 was silenced by an shRNA (Fig. 5, and figs. S7 and S8A). shRNA-mediated silencing was restricted to the human ortholog, thereby enabling us to examine the effects of hinge domain mutations on the function of mouse Smchd1. Note that all of the transfections were performed in parallel under identical conditions, including transfecting the cells with identical amounts of each expression construct and using the

same imaging parameters for each sample. Consequently, differences in the localization of mutant Smchd1 proteins relative to that of WT Smchd1 should reflect intrinsic differences in protein behavior, which were observed across multiple independent experiments (fig. S8). D1842 was included as an additional control because this mutation did not alter either the T_m or the nucleic acid binding activity (Fig. 4F, Table 2, fig. S6). In nonsilenced control 293 cells, endogenous SMCHD1 formed nuclear foci corresponding to localization to the inactive X chromosome (Fig. 5A and fig. S7A), as reported previously (32). Foci were depleted upon knockdown of endogenous SMCHD1 (Fig. 5B and fig. S7B), and successful knockdown was evidenced by the loss of SMCHD1 staining in individual cells across the microscope field. Expression of WT Smchd1 restored focal Smchd1 enrichment in the nuclei, although these foci were distinct from those formed by endogenous SMCHD1 (Fig. 5C). A similar phenotype to overexpression of WT Smchd1 was observed for cells expressing R1790A, R1796A, D1842A (control), R1869A, and K1873A (fig. S7), but cells expressing R1848A and the FSHD-associated mutant R1867G showed diffuse localization and abrogation of focus formation (Fig. 5, D and E). These data suggest that, on the whole, mutations that affected nucleic acid binding in vitro did not compromise chromatin binding, likely owing to some compensation from residues in the adjacent nucleic acid binding cluster. It was clear, however, that Arg^{1848} and Arg^{1867} contributed to a functional hotspot in the hinge domain that played an essential role in localizing Smchd1 to chromatin, and that their loss could not be compensated for by residues in the adjacent nucleic acid–binding clusters.

Because the R1848A and R1867A mutant hinge domains exhibited reduced thermal stability in vitro, which could be attributed to the loss of several interactions (fig. S8, B and C), we further examined whether hinge domain integrity contributed to Smchd1 chromatin localization. Introduction of mutations that compromised thermal stability within hinge domain-coiled coil dimers, V1774G and G1872A/G1875A/G1876A (fig. S7, K and L), or deletion of the hinge domain (ΔL1710-M1884; Fig. 5F, and fig. S7M), in full-length Smchd1 led to abrogation of focus formation. These data support an essential role for the Smchd1 hinge domain in chromatin localization and demonstrate that mislocalization can arise when hinge domain integrity is compromised.

DISCUSSION

Here, we present the crystal structure the Smchd1 hinge domain. The overall topology of this domain grossly resembles the previously described *T. maritima* SMC hinge domain homodimers and the eukaryotic SMC hinge domains of the cohesin SMC1-SMC3 (20), human condensin SMC2-SMC4 (26), and yeast SMC5-SMC6 (22) heterodimers, despite having sequence identities of only 10 to 20%. Common to these structures, the Smchd1 hinge domain contains a conserved, extended β-sheet formed by subdomains I and II contributed by the two protomers in the dimer, which is capped by an α-helix from each subdomain. However, there are several distinguishing features of the Smchd1 hinge domain compared to the other solved SMC hinge domain structures. Further to the absence of the H1-H2 and S1 and the addition of the C-terminal S8, the relative orientations of the two component subdomains in hinge of Smchd1 differ from conventional SMC hinge domains. The first indications of differences in subdomain

orientations arose during unsuccessful attempts to solve the structure by molecular replacement, which necessitated experimental phasing. The likely basis for this is the rotation of subdomain II by 18 or 27° relative to the respective closed or open *Tm*SMC reference structures. These rotations arise from the presence of three key nonconserved residues in the intersubunit interface: Val^{1774} (S2) and Tyr^{1765} (H3) in subdomain I, and Phe^{1874} (S7) in the dimer-opposed subdomain II. In addition to their roles in mediating dimer formation, as evidenced by the decreased thermal stability of the V1774G, Y1765A, and F1874A mutants, these residues likely determine the preference of Smchd1 for self-association rather than forming heterodimers with other SMC proteins. The atypical subdomain orientations in the Smchd1 hinge domain structure is enabled by the unconventional GX_6GX_2GG motif, which leads to S7 dropping out of the plane of the extended β-sheet and the β-strand peculiar to Smchd1, S8, forming part of the β-sheet.

Our structure enabled us to visualize the core hinge domain of Smchd1, but remarkably we did not observe electron density for the flanking coiled-coil regions except for two unassigned helices in the asymmetric unit. This was surprising considering the coiled-coil forming sequences were present within the crystallized protein, and because previous SAXS analyses of the hinge domain and flanking coiled-coils (9) and electron microscopy of full-length SMCHD1 (4) indicated the propensity for the coiled-coil regions flanking the core hinge domain to interact in a head-to-head manner. In contrast, structures of previously reported SMC domains form coiled-coils through head-to-tail (N- to C-terminal coil) coiled-coils that may arrange into either an open or closed conformation (fig. S3) (21, 24, 25). The lack of electron density for the flanking coiled-coil regions in the Smchd1

structure indicates that these regions are likely to be flexible, although whether they adopt an open or closed conformation cannot be deduced from the structure reported herein. Furthermore, the N terminus of the Smchd1 hinge domain lacks the H1-H2 typical of canonical SMC domains. One explanation for this absence could be that this is a site of crystal contacts in both the tetragonal and rhombohedral crystal forms, but more broadly is suggestive of lability at this site that would enable exposure of the N-terminal region of subdomain I to facilitate crystal contact formation. Although the atomic detail of the coiled-coils was not revealed in our crystal structure, the importance of these regions in augmenting homodimer stability is supported by our studies of dimer interface mutants, which led to reduced stability but no evidence of abrogated dimer formation. Mutations at the intersubunit interface identified from our hinge domain structure, including at the variant hinge domain GX_6GX_2GG motif, did not compromise dimer assembly, but did affect stability. This was surprising because mutation of this motif, identified as putative dimerization interface residues based on homology to other SMC hinge domains, was concluded to play a key role in Smchd1 dimerization in a previous study (4). In that study, triple mutation to alanine caused a modest shift in elution time in analytical size exclusion chromatography (similar to our own observations, fig. S5a), although the proposed disruption remained to be established biophysically. In our study, mutation of the three interface glycines or other interface residues led to reductions in thermal stability of the dimer, while not breaking the Smchd1 hinge homodimer, with SAXS analyses consistent with retention of the homodimer configuration. These observations are consistent with the flanking coiled-coils augmenting dimer formation, even when the core hinge domain intersubunit interface is compromised by mutation.

Like other SMC dimer structures and modelled counterparts (20, 21), the Smchd1 hinge domain possesses highly basic surfaces, of which only clusters 2 and 3, but not cluster 1 (residues adjacent to H1 and H3 in subdomain I), markedly contributed to nucleic acid interaction in vitro. The Smchd1 DNA-binding clusters were not inside the donut of the dimer structure, because mutation of torus residues to residues with bulkier side chains that would occlude DNA binding had no effect on nucleic acid binding. Instead, nucleic acid binding was mediated by two surface patches (clusters 2 and 3). This contrasts with the cohesin SMC1-SMC3 heterodimer central pore, where mutation of three positively charged SMC1 residues in the lumen of the hinge donut abolished DNA loading, but not recruitment to chromatin or cohesin catalytic activity (33). Surprisingly, most of the residues that we have implicated in DNA binding here are not well conserved between SMC domains, whereas all are highly conserved in vertebrate Smchd1 sequences, with only charge-conservative substitutions observed. Consistent with its high conservation among Smchd1 orthologs, the DNA-binding residue Arg^{1867} in S6 (cluster 3) is mutated in patients with FSHD, and is associated with loss of SMCHD1 function (15). Additionally, the cluster 2 residues, Arg^{1790}, Arg^{1796} and Lys^{1799}, all of which reside in the extended linker and contribute to DNA binding, are only individually present in human SMC4, mouse SMC1A, and mouse SMC1A/*Tm*SMC, respectively. Our studies implicate the interface residue Arg^{1848} and residues in clusters 2 and 3 in nucleic acid binding in vitro, although most of these residues were dispensable for chromatin localization in cells when introduced into full-length Smchd1. Only R1848A and the cluster 3 mutant R1867G [originally identified in an FSHD patient (15)], compromised the formation of nuclear foci.

Both Arg^{1848} and Arg^{1867} form multiple hydrogen bonds to surrounding residues proximal to, and including, the Smchd1 dimer interface (fig. S8, B and C), which may contribute to their key role in chromatin localization. Together, these data indicate that other residues in nucleic acid–binding clusters can compensate for mutations that diminish nucleic acid affinity, except for a functional hotspot encompassing Arg^{1848} and Arg^{1867}. We cannot exclude the possibility that compromise of hinge domain stability upon introduction of these mutations contributes to the loss of chromatin localization, as was observed when we deleted the hinge domain or introduced destabilizing mutations within the hinge domain. The precise contributions of the highly basic nucleic acid–binding clusters and intradomain dynamics await further detailed examination.

The mode by which Smchd1 interacts with DNA remains unclear, although several lines of evidence enable us to propose a model. First, our previous work showed that Smchd1 is specifically associated with regions of the chromatin where it is necessary for transcriptional silencing, and that Smchd1 exhibits a slight preference for binding to methylated DNA in vitro (15, 16). By contrast, we found that any association with RNA occurs without sequence specificity (34), a finding reminiscent of the nucleic acid–binding capacity of another epigenetic regulator complex, polycomb repressive complex 2 (PRC2), which similarly lacks sequence specificity in nucleic acid interactions (35). Analogous to our findings with Smchd1, PRC2 exhibits reduced function when site-specific mutations decrease nucleic acid binding. Furthermore, in this study, the nucleic acid binding capacity results in prolonged residence time on the DNA and it was proposed that PRC2 is not targeted by the DNA binding activity, but rather that the non-sequence specific

affinity acts simply to retain the complex on chromatin long enough for it to enzymatically modify the histones (35). We previously found that the targeting of Smchd1 to chromatin is dependent on PRC1-mediated ubiquitylation of histone 2A Lys^{119} (H2AK119ub) (34), which was subsequently validated by others (36). Here, we propose that this pathway is responsible for targeting, but as for PRC2, the nucleic acid binding capacity of the hinge domain is required to stabilize the binding of Smchd1 to chromatin to enable productive silencing. The Smchd1 hinge domain structure reported here provides a foundation for understanding the molecular basis of how chromatin binding can be compromised in disease, such as by the FSHD2-associated SMCHD1 mutation R1866G, although the precise mechanism and dynamics underlying SMCHD1 chromatin binding and transcription silencing remain to be further explored in the future.

MATERIALS AND METHODS

Recombinant protein production

The DNA sequence encoding mouse Smchd1 residues 1683 to 1899 was subcloned into pPROEX Htb (Invitrogen) for expression. Mutations were introduced by oligonucleotide-directed mutagenesis PCR, and the inserted sequences were verified by Sanger sequencing (Micromon). Proteins were expressed and purified from BL21 CodonPlus *E. coli* as described previously (9, 15). Briefly, cells were cultured shaking in Super broth to an OD_{600} of ~0.6 to 0.8 at 37°C before the temperature was reduced to 18°C and protein expression was induced by treatment with 0.5 mM IPTG overnight. Cells were lysed by sonication in lysis buffer [0.5 M NaCl, 20 mM Tris-Cl (pH 8), 20% (v/v) glycerol, 5 mM imidazole (pH 8), 0.5 mM TCEP] supplemented with 1 mM PMSF. His-tagged hinge

domains were purified by Ni-chromatography (FastFlow resin, Qiagen) and, following extensive washing, were eluted in lysis buffer containing 250 mM imidazole (pH 8). After cleavage of the His tag with TEV protease (leaving an additional GAMGS overhang sequence encoded by the vector), the protein was buffer-exchanged into lysis buffer, subjected to further Ni chromatography, and then the unbound fraction containing the hinge domain was concentrated and further purified by Superdex-200 size exclusion chromatography (GE Healthcare) in 100 mM NaCl, 20 mM HEPES (pH 7.5). Protein purity was evaluated by reducing SDS-PAGE with SafeStain visualization and fractions were pooled, concentrated to 5 to 10 mg/ml as estimated from the A_{280} measurement, aliquoted, snap-frozen in liquid N_2, and stored at −80 °C until required. Selenomethionine (SeMet)-labeled recombinant proteins were produced similarly to the native recombinant protein, except that cells were cultured in a chemically defined culture medium as previously described with modification (37). One liter of the culture medium contains 100 ml of 10X stock solution (0.4 M MOPS, 40 mM Tricine, 100 μM $FeSO_4$, 2.76 mM K_2SO_4, 5 μM $CaCl_2$, 5.28 mM $MgCl_2$, 0.4 M NaCl, 3 μM Na_2MoO_4, 400 μM H_3BO_3, 30 μM $CoCl_2$, 10 μM $CuSO_4$, 80 μM $MnCl_2$, 10 μM $ZnSO_4$), 0.508 g of NH_4Cl, 0.23 g K_2HPO_4, 0.84 g of $NaHCO_3$, 0.05 g of thiamine, 20% (w/v) glucose, 100 mg of phenylalanine, 100 mg of threonine, 100 mg of lysine, 50 mg of valine, 50 mg of leucine, 50 mg of isoleucine, and 60 mg of selenomethionine.

Protein crystallization

Smchd1 protein (residues 1683 to 1899) at a concentration of 10 mg/ml was subjected to robotic crystal trials (C3 Facility, CSIRO) using the sitting-drop vapor diffusion method.

Initial crystals were obtained in a condition comprising 0.1 M sodium bicine (pH 9.35), 0.18 M $MgCl_2$, 20% (v/v) 2-methyl-2,4-pentanediol or 20% (v/v) 2-propanol at 8°C. Conditions were further optimized using the hanging-drop vapor diffusion method. Diffraction quality crystals with rhombohedral symmetry were grown by mixing 1 µl of protein (10 mg/ml) with 1 µl of reservoir solution [0.1 M Tris-Cl (pH 9.0), 0.21 M $MgCl_2$, 20% (v/v) 2-methyl-2,4-pentanediol] at 8°C. Crystals typically appeared after 20 days. Diffraction-quality crystals with tetragonal symmetry were grown by mixing 1 µl of protein (5 mg/ml) with 1 µl of reservoir solution [0.1 M Tris-Cl (pH 10.0), 0.24 M $MgCl_2$, 24% (v/v) 2-propanol] at 8°C. Crystals typically appeared after 7 days. Crystals were transferred to a cryoprotectant solution containing the reservoir solution and 30% (v/v) glycerol and were flash-cooled for X-ray diffraction data collection. SeMet-derivatized crystals with tetragonal symmetry were obtained from SeMet-labeled recombinant protein crystallized under the same conditions as for the native protein.

Structure determination

The protein showed modest sequence homology to SMC hinge domains for which previous structures have been solved, with the *Thermotoga maritima* SMC hinge domain (PDB 1GXJ) providing the greatest identity at 21% (19). Attempts to solve the structure by molecular replacement with these models failed, and so anomalous phasing techniques were used. The SeMet derivative crystallized in the tetragonal ($P4_12_12$) spacegroup and diffracted to 3.3 Å; however, the anomalous signal was too weak to identify the seven potential selenium sites. To solve the phase problem, we derivatized the rhombohedral (R32) crystal form by soaking native crystals with a cryoprotectant

solution supplemented with 100 μM methylmercury chloride and obtained a dataset that diffracted to 4.2 Å. This dataset had sufficient anomalous signal at 7 Å to identify three mercury atoms covalently attached to the Cys^{1752} thiol in the asymmetric unit, revealing three copies of the Smchd1 hinge domain in the asymmetric unit (fig. S1, A to C). Through a combination of density modification and noncrystallographic symmetry averaging in PHENIX (38), it was possible to identify secondary structure features in the maps that were aligned to the *T. maritima* SMC hinge domain [PDB 1GXJ; (19)], generating a partial dimer model. The partial dimer was used as a search model for molecular replacement searching for a single copy in the higher resolution SeMet dataset (TFZ 21.1) (39). This solution was used as a partial model for MR-SAD searching for selenium sites (38). The phases from MR-SAD were used in combination with the partial dimer model for a phased MR placing the remaining two dimers in the asymmetric unit (40). The resulting model with three Smchd1 hinge domain dimers was used in another MR-SAD protocol to identify the selenium locations and assign sequence to the model (41). The final $P4_12_12$ model contained six copies of the hinge domain arranged into three dimers. The final model refined in buster (version 2.10.3) (42) with R_{fact} and R_{free} values of 21.4 and 24.1%, respectively, indicating a fit to the data comparable to other structures of similar resolution in the Protein Databank.

Small-angle X-ray scattering (SAXS)

SAXS data were collected at the Australian Synchrotron using the in-line size exclusion chromatography (SEC) setup (43, 44), as previously described (5, 9, 10, 45). Data collection and analysis statistics are shown in table S1. Recombinant protein (50 μl of

recombinant protein at a concentration of ~5 mg/ml) was resolved by injection onto an in-line Superdex-200 5/150 column (GE Healthcare) in 200 mM NaCl, 20 mM Tris-Cl (pH 8), 10% (v/v) glycerol, 0.5 mM TCEP (Pierce), and eluted in the path in the beam through a quartz capillary, as described previously (9, 10, 45). Two-second exposures of diffraction data were collected with a 1M Pilatus detector (Pilatus 2M detector for G1872A/G1875A/G1876A) and radially-averaged. Buffer-only scattering collected earlier in the data collection were averaged and used to subtract background scatter from data from the apex of the SEC peak were using Scatterbrain software (Stephen Mudie, Australian Synchrotron). Data analyses were performed with the ATSAS suite (46) as described previously (5, 9, 10, 45). Guinier analyses were performed with PRIMUS (29) to examine scatter at very low Q ($qRg \leq 1.3$) to estimate the radius of gyration, *Rg*, and zero angle intensity (*I(0)*), with linearity indicating the absence of both high molecular weight aggregates and interparticle interference. The real space interatomic distance distribution function *P(r)* and the maximum dimension of the particle D_{max} were computed by indirect Fourier transform with GNOM (30), which also enabled estimation of *Rg* and *I(0)*.

Differential Scanning Fluorimetry (DSF)

Analysis of the thermal stabilities of WT and mutant Smchd1 proteins (residues 1683 to 1899) was performed by DSF, essentially as described previously (47). Briefly, protein at a concentration of 0.05 or 0.1 mg/ml in 100 mM NaCl, 20 mM Tris-Cl (pH 8) was subjected to thermal denaturation over 20 to 100°C with a Biorad CFX96 RT-PCR machine in a 20-μl volume containing 0.3 μl of SYPRO Orange (Sigma) with fluorescence measured at

530 nm. The protein melting temperature (T_m) was calculated as the midpoint for the protein unfolding transition by fitting the sigmoidal melt curves of quadruplicate samples to the Boltzmann equation using GraphPad Prism, as described previously (15, 48), and is reported as the mean of the T_m values calculated at each of the protein concentrations (0.05 and 0.1 mg/ml). Data points after the fluorescence intensity maximum were excluded from fitting.

Nucleic acid binding assays

6-Fam fluorescently labeled single-stranded DNA (12.5 nM; HPLC-purified, IDT) corresponding to the HS5-1b sense methylated oligonucleotides (5'-6-Fam-ATCTGCCACCTGGTGTTC*GA, where the asterisk denotes methylation) as previously reported (15) was incubated with recombinant Smchd1 hinge at final concentrations of 0 to 100 μM diluted in protein sample buffer [100 mM NaCl, 20 mM HEPES (pH 7.5)] in 10-μl reactions. Reactions were set up in 384-well low-flange black flat bottom plates (Corning) in duplicate and incubated at room temperature for 15 min in the dark. The emission polarization values were measured with the Envision Multilabel plate reader (PerkinElmer) with a 480 nm excitation filter, a 535 nm static and polarized filter, and FITC FP dual mirror.

shRNA-mediated knockdown of SMCHD1 in 293 cells

For knockdown of human SMCHD1 in 293 cells, we designed shRNA nucleotides targeting the 3'UTR of *SMCHD1* (3'UTR_forward: TCGAGAAGGTATATTGCTGTTGACAGTGAGCGCTTGGTTGTCACTACCTT

GCAATAGTGAAGCCACAGATGTATTGCAAGGTAGTGACAACCAAATGCCTACTGCC TCGG; 3'UTR_reverse: AATTCCGAGGCAGTAGGCATTTGGTTGTCACTACCTTGCAATACATCTGTGGCTTC ACTATTGCAAGGTAGTGACAACCAAGCGCTCACTGTCAACAGCAATATACCTTC). These were subcloned into the LMPEBFP2 vector (LTR miR30 Puromycin IRES EBFP2) vector (49). Retroviruses were prepared as previously described (50). 293T cells were cultured in DMEM supplemented with 10% FBS at 37°C in a humidified atmosphere with 5% (v/v) CO_2. For retrovirus production, cells at 80% confluence were transfected with calcium phosphate with the MD1-gag-pol structural vector, CAG-Eco, and the shRNA retroviral construct in the ratio 8:24:1. Plasmid DNA was made up in 250 mM $CaCl_2$ and precipitated in 2X HBS solution, and then was added to the 293T cells in medium containing 25 μM chloroquine (Sigma-Aldrich). The medium was changed 8 and 24 hours after transfection and collected 48 and 72 hours after transfection by centrifugation to remove residual 293T cells. For transduction of 293 cells with shRNA-expressing retroviral constructs, either a nonsilencing control shRNA or shRNAs designed to knockdown SMCHD1 expression, retroviral supernatant was prepared 1:10 in medium containing polybrene (4 μg/ml Sigma-Aldrich) and added to 293 cells at 50% confluence. After 24 hours, the medium was changed and puromycin (5 μg/ml, Sigma-Aldrich) was added for selection of transduced cells.

Transfection of SMCHD1-knockdown 293 cells with plasmids expressing Smchd1 hinge domain mutants

The introduction of point mutations in mouse full-length Smchd1 was accomplished by PCR-mediated site-directed mutagenesis, and deletion of the hinge domain (residues L1710 to M1884) or generation of the mutations G1872A/G1875A/G1876A or L1774G were introduced by Gibson assembly, before cloning into pcDNA3 vector at the BamHI and XbaI restriction enzyme sites. All inserted sequences were verified by Sanger sequencing (AGRF, Melbourne). For transfection of SMCHD1-knockdown 293 cells with constructs expressing full-length Smchd1 point mutant variants, ~2 x 10^4 293 cells were seeded in a 12-well plate on a 13-mm coverslip (Marienfield Superior). Twenty-four hours later, the cells were at ~80% confluency and were transfected with 1.2 μg of the appropriate construct by calcium phosphate-mediated transfection. Immunofluorescence microscopy was performed 24 hours after transfection.

Immunofluorescence microscopy

Immunofluorescence microscopy studies were performed as detailed by Chaumeil *et al*. (51). Briefly, the cells were washed in PBS and fixed in PBS, 3% (w/v) paraformaldehyde for 10 min at room temperature. Cells were washed three times with PBS for 5 min each, then permeabilized on ice with 0.5% (v/v) Triton X-100 in PBS, followed by three washes in PBS for 5 min each. Cells were blocked in 1% (w/v) bovine serum albumin (BSA; Life Technologies) for 15 min, followed by a 45-min incubation in a dark and humid chamber at room temperature with a primary anti-Smchd1 antibody [in-house clone 1D6, previously reported (34)] diluted 1:100 in 1% (w/v) BSA. Cells were washed three times in PBS for 5 min each and incubated for 40 min at room temperature in a dark and humid chamber with a secondary anti-rat-568 antibody (Life Technologies, A-11077) diluted 1:500 in 1%

(w/v) BSA. Cells were washed three times in PBS for 5 min each and stained with DAPI for 10 min at room temperature, followed by another two washes with PBS. Coverslips were mounted in Vectashield H1000 mounting medium (Vector Laboratories). Cells were visualized with a Zeiss LSM 880 NLO microscope at 63x magnification, and z-stacks were acquired. Images were analyzed with the open source ImageJ distribution package, FIJI.

Intact protein analysis of crystals using mass spectrometry

Crystals were dissolved in 3% acetonitrile (ACN) and 2% formic acid (FA). Samples were analyzed by nanoflow LC-MS on a nanoAcquity UPLC system (Waters) coupled to a Maxis II ETD mass spectrometer (Bruker) through a Captivespray ion source with nitrogen gas doped with ACN using a NanoBooster (Bruker). Proteins were directly injected onto a 75 μm X 500 mm column (3.6 μm C4, packed emitter tip, Ion Opticks) at 3% buffer B, and separated by reverse-phase chromatography on a 45 min linear gradient from 5% to 60% buffer B (A: Milli-Q water, 0.1% FA; B: 99.9% ACN, 0.1% FA) at a 500 nl/min constant flow rate. The Maxis II ETD recorded at 1 Hz (300 to 3000 m/z, profile mode) with the instrument controlled using otofControl version 4.0 (Bruker). Data analysis was performed using Compass DataAnalysis version 4.4 (Bruker). Spectra corresponding to Smchd1 were averaged across each LC-MS run before mass list generation (SNAP).

SUPPLEMENTARY MATERIALS

Fig. S1. The Smchd1 rhombohedral crystal structure used to solve the phase problem by single anomalous dispersion and electron density around Arg^{1762} in the tetragonal crystal structure.

Fig. S2. Coiled coils flanking the core Smchd1 hinge domain are observed as two unassigned helices in the asymmetric unit, despite being present in intact mass spectrometry analysis of SMCHD1 hinge dimer crystals.

Fig. S3. Crystal structures of SMC proteins with coiled coils.

Fig. S4. Dimer interfaces of Smchd1 and canonical SMC hinge domains.
Fig. S5. Size-exclusion chromatograms of Smchd1 hinge mutants.
Fig. S6. T_m values of wild-type and mutant Smchd1 hinge domains measured by DSF.
Fig. S7. The mutations R1848A and R1867G and deletion or compromise of hinge domain integrity alter the nuclear localization pattern of Smchd1.
Fig. S8. The mutations R1848A and R1867G perturb local interactions and abrogate formation of Smchd1 nuclear foci.
Table S1. Data collection and scattering parameters for SAXS analysis.

REFERENCES AND NOTES

1. T. Hirano, At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* **7**, 311-322 (2006)
2. T. Gligoris, J. Lowe, Structural Insights into Ring Formation of Cohesin and Related Smc Complexes. *Trends Cell Biol* **26**, 680-693 (2016)
3. K. Nasmyth, C. H. Haering, The structure and function of SMC and kleisin complexes. *Annu Rev Biochem* **74**, 595-648 (2005)
4. N. J. Brideau, H. Coker, A. V. Gendrel, C. A. Siebert, K. Bezstarosti, J. Demmers, R. A. Poot, T. B. Nesterova, N. Brockdorff, Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. *Mol Cell Biol* **35**, 4053-4068 (2015)
5. K. Chen, R. C. Dobson, I. S. Lucet, S. N. Young, F. G. Pearce, M. E. Blewitt, J. M. Murphy, The epigenetic regulator Smchd1 contains a functional GHKL-type ATPase domain. *Biochem J* **473**, 1733-1744 (2016)
6. L. C. Pedersen, K. Inoue, S. Kim, L. Perera, N. D. Shaw, A ubiquitin-like domain is required for stabilizing the N-terminal ATPase module of human SMCHD1. *Commun Biol* **2**, 255 (2019)
7. M. E. Blewitt, A. V. Gendrel, Z. Pang, D. B. Sparrow, N. Whitelaw, J. M. Craig, A. Apedaile, D. J. Hilton, S. L. Dunwoodie, N. Brockdorff, G. F. Kay, E. Whitelaw, SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* **40**, 663-669 (2008)
8. N. Jansz, K. Chen, J. M. Murphy, M. E. Blewitt, The Epigenetic Regulator SMCHD1 in Development and Disease. *Trends Genet* **33**, 233-243 (2017)
9. K. Chen, P. E. Czabotar, M. E. Blewitt, J. M. Murphy, The hinge domain of the epigenetic repressor Smchd1 adopts an unconventional homodimeric configuration. *Biochem J* **473**, 733-742 (2016)
10. A. D. Gurzau, K. Chen, S. Xue, W. Dai, I. S. Lucet, T. T. N. Ly, B. Reversade, M. E. Blewitt, J. M. Murphy, FSHD2- and BAMS-associated mutations confer opposing effects on SMCHD1 function. *J Biol Chem* **293**, 9841-9853 (2018)
11. R. J. Lemmers, R. Tawil, L. M. Petek, J. Balog, G. J. Block, G. W. Santen, A. M. Amell, P. J. van der Vliet, R. Almomani, K. R. Straasheijm, Y. D. Krom, R. Klooster, Y. Sun, J. T. den Dunnen, Q. Helmer, C. M. Donlin-Smith, G. W. Padberg, B. G. van Engelen, J. C. de Greef, A. M. Aartsma-Rus, R. R. Frants, M. de Visser, C. Desnuelle, S. Sacconi, G. N. Filippova, B. Bakker, M. J. Bamshad, S. J. Tapscott, D. G. Miller, S. M. van der Maarel, Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* **44**, 1370-1374 (2012)

12. S. Sacconi, R. J. Lemmers, J. Balog, P. J. van der Vliet, P. Lahaut, M. P. van Nieuwenhuizen, K. R. Straasheijm, R. D. Debipersad, M. Vos-Versteeg, L. Salviati, A. Casarin, E. Pegoraro, R. Tawil, E. Bakker, S. J. Tapscott, C. Desnuelle, S. M. van der Maarel, The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet* **93**, 744-751 (2013)
13. C. T. Gordon, S. Xue, G. Yigit, H. Filali, K. Chen, N. Rosin, K. I. Yoshiura, M. Oufadem, T. J. Beck, R. McGowan, A. C. Magee, J. Altmuller, C. Dion, H. Thiele, A. D. Gurzau, P. Nurnberg, D. Meschede, W. Muhlbauer, N. Okamoto, V. Varghese, R. Irving, S. Sigaudy, D. Williams, S. F. Ahmed, C. Bonnard, M. K. Kong, I. Ratbi, N. Fejjal, M. Fikri, S. C. Elalaoui, H. Reigstad, C. Bole-Feysot, P. Nitschke, N. Ragge, N. Levy, G. Tuncbilek, A. S. Teo, M. L. Cunningham, A. Sefiani, H. Kayserili, J. M. Murphy, C. Chatdokmaiprai, A. M. Hillmer, D. Wattanasirichaigoon, S. Lyonnet, F. Magdinier, A. Javed, M. E. Blewitt, J. Amiel, B. Wollnik, B. Reversade, De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat Genet* **49**, 249-255 (2017)
14. N. D. Shaw, H. Brand, Z. A. Kupchinsky, H. Bengani, L. Plummer, T. I. Jones, S. Erdin, K. A. Williamson, J. Rainger, A. Stortchevoi, K. Samocha, B. B. Currall, D. S. Dunican, R. L. Collins, J. R. Willer, A. Lek, M. Lek, M. Nassan, S. Pereira, T. Kammin, D. Lucente, A. Silva, C. M. Seabra, C. Chiang, Y. An, M. Ansari, J. K. Rainger, S. Joss, J. C. Smith, M. F. Lippincott, S. S. Singh, N. Patel, J. W. Jing, J. R. Law, N. Ferraro, A. Verloes, A. Rauch, K. Steindl, M. Zweier, I. Scheer, D. Sato, N. Okamoto, C. Jacobsen, J. Tryggestad, S. Chernausek, L. A. Schimmenti, B. Brasseur, C. Cesaretti, J. E. Garcia-Ortiz, T. P. Buitrago, O. P. Silva, J. D. Hoffman, W. Muhlbauer, K. W. Ruprecht, B. L. Loeys, M. Shino, A. M. Kaindl, C. H. Cho, C. C. Morton, R. R. Meehan, V. van Heyningen, E. C. Liao, R. Balasubramanian, J. E. Hall, S. B. Seminara, D. Macarthur, S. A. Moore, K. I. Yoshiura, J. F. Gusella, J. A. Marsh, J. M. Graham, Jr., A. E. Lin, N. Katsanis, P. L. Jones, W. F. Crowley, Jr., E. E. Davis, D. R. FitzPatrick, M. E. Talkowski, SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet* **49**, 238-248 (2017)
15. K. Chen, J. Hu, D. L. Moore, R. Liu, S. A. Kessans, K. Breslin, I. S. Lucet, A. Keniry, H. S. Leong, C. L. Parish, D. J. Hilton, R. J. Lemmers, S. M. van der Maarel, P. E. Czabotar, R. C. Dobson, M. E. Ritchie, G. F. Kay, J. M. Murphy, M. E. Blewitt, Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *Proc Natl Acad Sci U S A* **112**, E3535-3544 (2015)
16. N. Jansz, A. Keniry, M. Trussart, H. Bildsoe, T. Beck, I. D. Tonks, A. W. Mould, P. Hickey, K. Breslin, M. Iminitoff, M. E. Ritchie, E. McGlinn, G. F. Kay, J. M. Murphy, M. E. Blewitt, Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. *Nat Struct Mol Biol* **25**, 766-777 (2018)
17. C. Y. Wang, T. Jegu, H. P. Chu, H. J. Oh, J. T. Lee, SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. *Cell* **174**, 406-421 e425 (2018)
18. M. R. Gdula, T. B. Nesterova, G. Pintacuda, J. Godwin, Y. Zhan, H. Ozadam, M. McClellan, D. Moralli, F. Krueger, C. M. Green, W. Reik, S. Kriaucionis, E. Heard, J. Dekker, N. Brockdorff, The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. *Nat Commun* **10**, 30 (2019)

19. C. H. Haering, J. Lowe, A. Hochwagen, K. Nasmyth, Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell* **9**, 773-788 (2002)
20. A. Kurze, K. A. Michie, S. E. Dixon, A. Mishra, T. Itoh, S. Khalid, L. Strmecki, K. Shirahige, C. H. Haering, J. Lowe, K. Nasmyth, A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. *EMBO J* **30**, 364-378 (2011)
21. Y. M. Soh, F. Burmann, H. C. Shin, T. Oda, K. S. Jin, C. P. Toseland, C. Kim, H. Lee, S. J. Kim, M. S. Kong, M. L. Durand-Diebold, Y. G. Kim, H. M. Kim, N. K. Lee, M. Sato, B. H. Oh, S. Gruber, Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol Cell* **57**, 290-303 (2015)
22. A. Alt, H. Q. Dang, O. S. Wells, L. M. Polo, M. A. Smith, G. A. McGregor, T. Welte, A. R. Lehmann, L. H. Pearl, J. M. Murray, A. W. Oliver, Specialized interfaces of Smc5/6 control hinge stability and DNA association. *Nat Commun* **8**, 14011 (2017)
23. L. Holm, L. M. Laakso, Dali server update. *Nucleic Acids Res* **44**, W351-355 (2016)
24. K. Kamada, M. Su'etsugu, H. Takada, M. Miyata, T. Hirano, Overall Shapes of the SMC-ScpAB Complex Are Determined by Balance between Constraint and Relaxation of Its Structural Parts. *Structure* **25**, 603-616 e604 (2017)
25. Y. Li, A. J. Schoeffler, J. M. Berger, M. G. Oakley, The crystal structure of the hinge domain of the Escherichia coli structural maintenance of chromosomes protein MukB. *J Mol Biol* **395**, 11-19 (2010)
26. S. Uchiyama, K. Kawahara, Y. Hosokawa, S. Fukakusa, H. Oki, S. Nakamura, Y. Kojima, M. Noda, R. Takino, Y. Miyahara, T. Maruno, Y. Kobayashi, T. Ohkubo, K. Fukui, Structural Basis for Dimer Formation of Human Condensin Structural Maintenance of Chromosome Proteins and Its Implications for Single-stranded DNA Recognition. *J Biol Chem* **290**, 29461-29477 (2015)
27. F. Corpet, Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881-10890 (1988)
28. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324 (2014)
29. P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277-1282 (2003)
30. D. I. Svergun, Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J Appl Crystallogr* **25**, 495-503 (1992)
31. B. W. Matthews, H. Nicholson, W. J. Becktel, Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci U S A* **84**, 6663-6667 (1987)
32. R. S. Nozawa, K. Nagao, K. T. Igami, S. Shibata, N. Shirai, N. Nozaki, T. Sado, H. Kimura, C. Obuse, Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat Struct Mol Biol* **20**, 566-573 (2013)
33. M. Srinivasan, J. C. Scheinost, N. J. Petela, T. G. Gligoris, M. Wissler, S. Ogushi, J. E. Collier, M. Voulgaris, A. Kurze, K. L. Chan, B. Hu, V. Costanzo, K. A. Nasmyth, The Cohesin Ring Uses Its Hinge to Organize DNA Using Non-topological as well as Topological Mechanisms. *Cell* **173**, 1508-1519 e1518 (2018)
34. N. Jansz, T. Nesterova, A. Keniry, M. Iminitoff, P. F. Hickey, G. Pintacuda, O. Masui, S. Kobelke, N. Geoghegan, K. A. Breslin, T. A. Willson, K. Rogers, G. F. Kay, A. H. Fox, H. Koseki, N. Brockdorff, J. M. Murphy, M. E. Blewitt, Smchd1 Targeting to the

Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. *Cell Rep* **25**, 1912-1923 e1919 (2018)
35. J. Choi, A. L. Bachmann, K. Tauscher, C. Benda, B. Fierz, J. Muller, DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. *Nat Struct Mol Biol* **24**, 1039-1047 (2017)
36. C. Y. Wang, D. Colognori, H. Sunwoo, D. Wang, J. T. Lee, PRC1 collaborates with SMCHD1 to fold the X-chromosome and spread Xist RNA between chromosome compartments. *Nat Commun* **10**, 2950 (2019)
37. F. C. Neidhardt, P. L. Bloch, D. F. Smith, Culture medium for enterobacteria. *J Bacteriol* **119**, 736-747 (1974)
38. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010)
39. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007)
40. A. Vagin, A. Teplyakov, MOLREP: an automated program for molecular replacement. *J Appl Cryst* **30**, 1022-1025 (1997)
41. T. C. Terwilliger, P. D. Adams, R. J. Read, A. J. McCoy, N. W. Moriarty, R. W. Grosse-Kunstleve, P. V. Afonine, P. H. Zwart, L. W. Hung, Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601 (2009)
42. O. S. Smart, T. O. Womack, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C. Vonrhein, G. Bricogne, Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**, 368-380 (2012)
43. N. M. Kirby, S. T. Mudie, A. M. Hawley, D. J. Cookson, H. D. T. Mertens, N. Cowieson, V. Samardzic-Boban, A low-background-intensity focusing small-angle X-ray scattering undulator beamline. *J Appl Crystallogr* **46**, 1670-1680 (2013)
44. T. M. Ryan, J. Trewhella, J. M. Murphy, J. R. Keown, L. Casey, F. G. Pearce, D. C. Goldstone, K. Chen, Z. Luo, B. Kobe, C. A. McDevitt, S. A. Watkin, A. M. Hawley, S. T. Mudie, V. S. Boban, N. Kirby, An optimized SEC-SAXS system enabling high X-ray dose for rapid SAXS assessment with correlated UV measurements for biomolecular structure analysis. *J Appl Crystallogr* **51**, 97-111 (2018)
45. E. J. Petrie, J. J. Sandow, A. V. Jacobsen, B. J. Smith, M. D. W. Griffin, I. S. Lucet, W. Dai, S. N. Young, M. C. Tanzer, A. Wardak, L. Y. Liang, A. D. Cowan, J. M. Hildebrand, W. J. A. Kersten, G. Lessene, J. Silke, P. E. Czabotar, A. I. Webb, J. M. Murphy, Conformational switching of the pseudokinase domain promotes human MLKL tetramerization and cell death by necroptosis. *Nat Commun* **9**, 2422 (2018)
46. P. Konarev, Petoukhov, MV, Volkov, VV, Svergun, DI ATSAS 2.1, a program package for small-angle scattering data analysis. *J Appl Cryst* **39**, 277-286 (2006)
47. S. A. Seabrook, J. Newman, High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb Sci* **15**, 387-392 (2013)
48. J. M. Murphy, Q. Zhang, S. N. Young, M. L. Reese, F. P. Bailey, P. A. Eyers, D. Ungureanu, H. Hammaren, O. Silvennoinen, L. N. Varghese, K. Chen, A. Tripaydonis, N.

Jura, K. Fukuda, J. Qin, Z. Nimchuk, M. B. Mudgett, S. Elowe, C. L. Gee, L. Liu, R. J. Daly, G. Manning, J. J. Babon, I. S. Lucet, A robust methodology to subclassify pseudokinases based on their nucleotide binding properties. *Biochemical Journal* **457**, 323-334 (2014)
49. S. A. Kinkel, R. Galeev, C. Flensburg, A. Keniry, K. Breslin, O. Gilan, S. Lee, J. Liu, K. Chen, L. J. Gearing, D. L. Moore, W. S. Alexander, M. Dawson, I. J. Majewski, A. Oshlack, J. Larsson, M. E. Blewitt, Jarid2 regulates hematopoietic stem cell function by acting with polycomb repressive complex 2. *Blood* **125**, 1890-1900 (2015)
50. I. J. Majewski, M. E. Blewitt, C. A. de Graaf, E. J. McManus, M. Bahlo, A. A. Hilton, C. D. Hyland, G. K. Smyth, J. E. Corbin, D. Metcalf, W. S. Alexander, D. J. Hilton, Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. *PLoS Biol* **6**, e93 (2008)
51. J. Chaumeil, S. Augui, J. C. Chow, E. Heard, Combined immunofluorescence, RNA fluorescent in situ hybridization, and DNA fluorescent in situ hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. *Methods Mol Biol* **463**, 297-308 (2008)

Acknowledgements: We thank the MX and SAXS beamline staff at the Australian Synchrotron for assistance with data collection, the C3 Collaborative Crystallization Centre for assistance with robotic crystallization screening and DSF experiments, and P. Colman (Walter and Eliza Hall Institute) for helpful discussions. We thank A. Keniry (Walter and Eliza Hall Institute) for assistance with designing shRNAs. **Funding:** This study was supported by an Australian Research Training Program scholarship (A.D.G), a Melbourne Research Scholarship - International (I.W.), a Cancer Council Victoria fellowship (K.C.), the Bellberry-Viertel Senior Medical Research Fellowship (M.E.B), Australian National Health and Medical Research Council fellowships (1079700 to P.E.C.; 1105754 and 1172929 to J.M.M.) and grant (1098290 to P.E.C., M.E.B. and J.M.M.). Additional support was provided by the Victorian State Government Operational Infrastructure Support and the Australian National Health and Medical Research Council IRIISS grant (9000433). We thank the InSPIRE program for supporting R.W. **Author contributions:** K.C., R.W.B., and A.D.G. designed and performed experiments and analyzed data; I.W., R.W., M.I., J.J.S., S.N.Y., P.J.H., T.A.W., D.A.H., and A.I.W. performed experiments, analyzed data, or both; M.E.B., P.E.C., and J.M.M. led the study and wrote the paper with K.C., R.W.B., and A.D.G., as well as input from the other authors. **Competing interests:** The authors declare they have no competing interests. **Data and materials availability:** All sequences are available on request. Plasmids require a materials transfer agreement (MTA) from the Walter and Eliza Hall Institute. The refined coordinates for the $P4_12_12$ structure were deposited in the PDB (PDB ID: 6N64). All other data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

Fig. 1. The overall structure of the hinge domain of SMCHD1 reveals differences compared to canonical SMC proteins. (**A** and **B**) Schematic diagram of SMCHD1 (A)

and the canonical SMC (B) protein domain architecture. SMCHD1 and canonical SMC proteins dimerize through their hinge domain, which is composed of subdomain I (slate and cyan) and subdomain II (green). The canonical SMC dimer has long intramolecular antiparallel coiled-coil arms that are composed of N- (blue) and C-terminal (pink) residues flanking the hinge domain, and the ABC-type ATPase domain constituted of N- and C-terminal subunits. The Smchd1 dimer contains shorter intermolecular coiled-coils which are connected to its N-terminal region (gray) encapsulating a GHKL-type ATPase domain (red). (**C** and **D**) The crystal structure of the of Smchd1 hinge domain (C) and a canonical SMC hinge (PDB: 1GXK; D) as shown from above the torus in cartoon representation. The component protomers are demarked by a dashed border. The color scheme is the same as in Fig. 1A but with protomer 1 in purple (subdomain I) and orange (subdomain II) for contrast. (**E** and **F**) Diagrams depicting the topology of the Smchd1 monomer (residues 1710 to 1884) (E) and the canonical SMC hinge domain from *T. maritima* (F). (**G** and **H**) Cartoons comparing the hinge domain monomers of Smchd1 (G) and *T. maritima* (H). (**I**) Superimposition of the Smchd1 and *T. maritima* SMC hinge domains (from G and H, respectively). Secondary structure is annotated sequentially for Smchd1 and according to the original *Tm*SMC description (19); however, short helical segments (consisting of less than two hydrogen bonds) have been omitted from the renderings.

Fig. 2. Multiple sequence alignment of the hinge domains of the Smchd1 hinge domain and canonical SMC proteins. Protein sequences of *Mus musculus* (*Mm*) Smchd1 and canonical SMC proteins, including *Mm*SMC1 and SMC3, *Homo sapiens* (*Hs*) SMC2 and SMC4, and *Thermotoga maritima* (*Tm*) SMC were aligned using the

program MultAlin (27). α-helices (H) and β-strands (S) are colored in red and yellow, respectively, the 3_{10} helices are shaded in purple, and the loop connecting subdomain I and II is colored in blue. The secondary structures of mouse Smchd1 and *Tm*SMC as defined by DALI are annotated above and below the aligned sequences, respectively. The image was generated using ESPript3.0 (28). Gray text indicates regions of the SMCHD1 structure that could not be built due to unresolved electron density.

Fig. 3. Dimerization of the Smchd1 hinge domain occurs via canonical and non-canonical interfaces. (**A**) Magnified view of the dimerization interface of the Smchd1 hinge domain as highlighted with subdomain I in slate and subdomain II in green and the side view at 90° as indicated. Key residues coordinating the dimer interface are shown as sticks with numbers indicated and highlighted according to their subdomain. Electrostatic interactions between residues are indicated by dashed lines. (**B** to **G**) Small angle X-ray analysis of WT (B) [replotted from our earlier report (9)] and the representative dimer interface mutants Y1765A (C), V1774G (D), R1848A (E), K1873A (F), and G1872A/G1875A/G1876A (G). Top: Scattering intensity profiles are shown where the background-subtracted SAXS data are indicated as black circles representing mean intensity $I(q)$ as a function of momentum transfer q in $Å^{-1}$. Guinier plots for $qR_g \leq 1.3$ are shown as insets, where linearity indicates that high molecular weight and polydisperse particles do not measurably contribute to the scattering. The radius of gyration and initial scattering intensity $I(0)$ were approximated using the Guinier equation with PRIMUS (29), giving values as indicated in table S1. Bottom: Pair-distribution functions, $P(r)$ plots were calculated using GNOM (30). The R_g and maximum particle dimension D_{max} calculated

from the *P*(*r*) analysis are as reported in table S1. Data in (C) to (E) were collected with a Pilatus 1M detector; a Pilatus 2M detector was used to collect data for the G1872A/G1875A/G1876A mutant.

Fig. 4. The nucleic acid–binding activity of the SMCHD1 hinge homodimer is potentiated by two positively charged interaction sites. (**A**) Electrostatic surface potential representations of the Smchd1 hinge homodimer viewed down the side of the torus. Positively charged residues are shown as sticks in dark blue. Lys^{1718}, Arg^{1719}, and Arg^{1771} at the top side of the dimer are grouped as Cluster 1 (black). Lys^{1789}, Arg^{1790}, Arg^{1796}, and Lys^{1799} located at the bottom side of the dimer are grouped as Cluster 2 (cyan). Residues in close proximity to the pathogenic mutation Arg^{1867}, including Arg^{1869}, Lys^{1873}, and Lys^{1880} are grouped as Cluster 3 (blue). (**B**) Cartoon representation of the Smchd1 homodimer in the same orientation as that shown in (A) highlighting residues that interact with nucleic acid on one protomer. The thin red line indicates the possible position of nucleic acid binding as deduced from these experiments. (**C** to **G**) DNA binding fluorescence polarization (mP) of 12.5 nM 6-FAM–labeled, 20-bp single-stranded DNA by wild-type (WT) and Cluster 1 (C), Cluster 2 (D), Cluster 3 (E), dimer interface (F), and the torus (G) mutants of the Smchd1 hinge dimer at indicated protein concentrations. Data points are plotted for technical duplicates for three (wild-type and torus mutants) or two (all other mutants) independent experiments fitted to a hyperbolic binding curve. K_d values were calculated from the fitted curves are reported in Table 2. (**H** to **J**) The Smchd1 hinge dimer has two clusters of positively charged residues implicated in nucleic acid binding. The first interaction site, comprising Arg^{1790}, Arg^{1796}, and Lys^{1799} from cluster 2

(cyan), is positioned on the surface of the dimer where the inter-subdomain linker resides. The second interaction site is located at the central channel of the hinge dimer on the opposite side of the dimer, comprising Arg^{1867} and Lys^{1873} from cluster 3 (blue). Residue Arg^{1848} (colored in magenta in the structural figures), which is also positioned at the central channel, is located between interaction sites 1 and 2 and is also implicated in nucleic acid binding. (H) and (J) show the same view of the hinge domain dimer to show the torus residues selected for mutation (J) and cluster 2 residues (H). Distances between the residues located at the two nucleic acid interaction sites are marked beside corresponding dashed lines in (H) and (I).

Fig. 5. The mutations R1848A or R1867G or compromising the integrity of the hinge domain alter the nuclear localization pattern of Smchd1. (**A** to **F**) Immunofluorescence microscopy analysis of (A) control, nonsilencing shRNA–transduced or shRNA-mediated SMCHD1-knockdown 293 cells in the absence (B) or presence of wild-type (C) or mutant (D and E) full-length Smchd1 or a variant missing the hinge domain (F). Maximum intensity projection images are shown as representative of n>150 nuclei positive for Smchd1 overexpression per sample. Data are representative of three independent experiments. All images were obtained with identical settings between controls and all transfected cells to enable comparison between the images provided in the figure. Scale bars, 20 μm. DAPI and SMCHD1 staining and merged channels are displayed in fig. S7, and repeat experiments are shown in fig. S8.

Table 1. X-ray crystallography data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

Structural parameters		
	SMCHD1 SeMet P4	**SMCHD1 Hg R32**
Wavelength		
Resolution range	48.45 - 3.3 (3.418 - 3.3)	46.61 - 4.2 (4.35 - 4.2)
Space group	$P4_12_12$	$R32{:}H$
Unit cell dimensions	123.549 123.549 232.869 90 90 90	154.058 154.058 244.088 90 90 120
Total reflections	493800 (48634)	92383 (9255)
Unique reflections	27895 (2722)	8369 (821)
Multiplicity	17.7 (17.9)	11.0 (11.3)
Completeness (%)	99.87 (99.71)	99.38 (99.15)
Mean I/sigma(I)	14.81 (2.03)	8.38 (1.67)
Wilson B-factor	107.71	171.93
R-meas	0.2053 (1.956)	0.2137 (1.75)
R-pim	0.04826 (0.461)	0.06437 (0.5202)
Reflections used in refinement	27880 (2720)	8326 (814)
Reflections used for R-free	1396 (143)	406 (40)
CC1/2	0.999 (0.674)	0.998 (0.812)
R-work	0.2142 (0.2871)	0.2973 (0.3700)

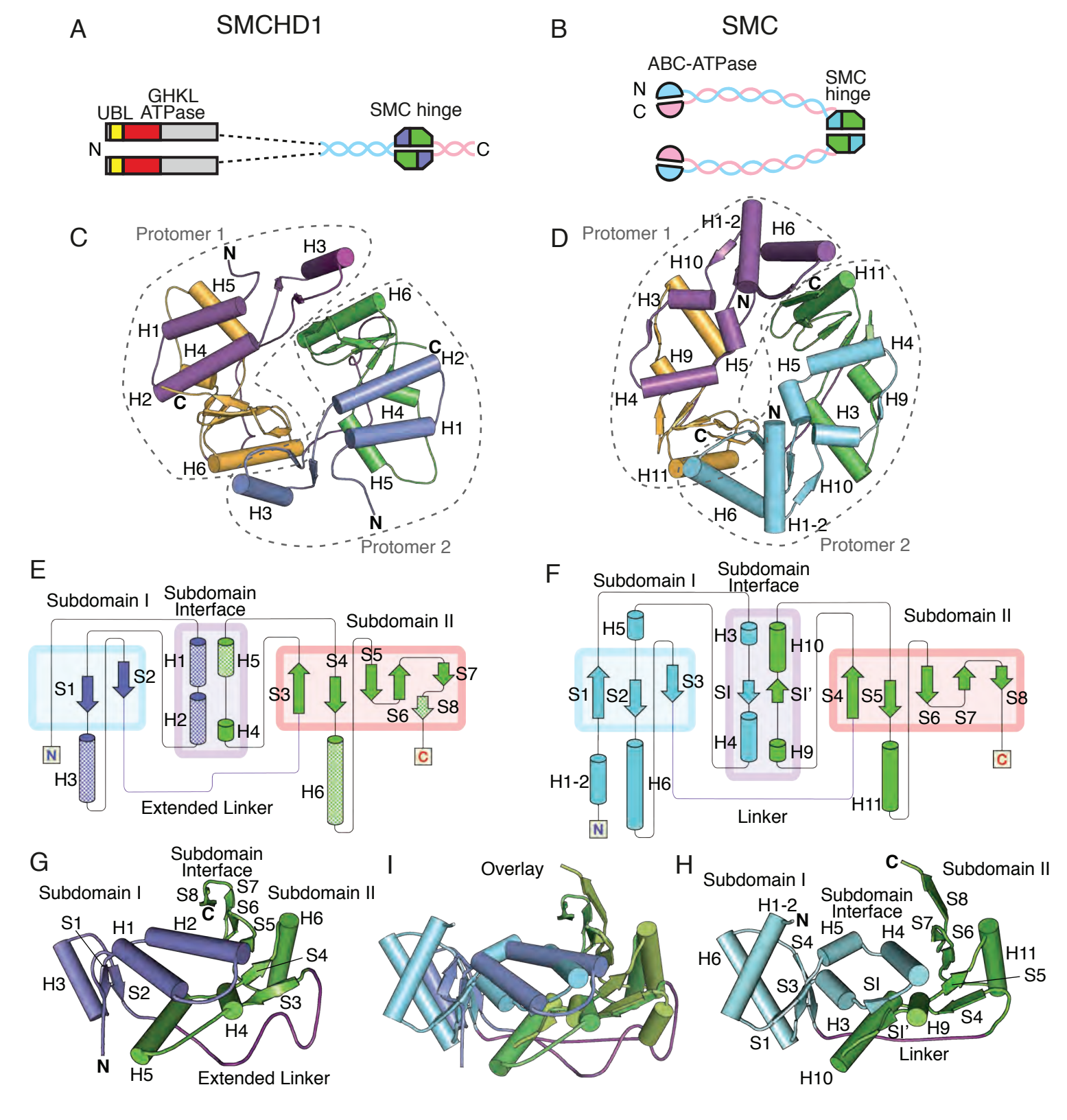
R-free	0.2414 (0.3245)	0.3098 (0.3368)
Macromolecules	8419	4104
RMS(bonds)	0.013	0.016
RMS(angles)	1.74	1.75
Ramachandran favored (%)	94.32	91.83
Ramachandran allowed (%)	4.92	6.42
Clashscore	6.66	17.35
Average B-factor	117.5	155.39

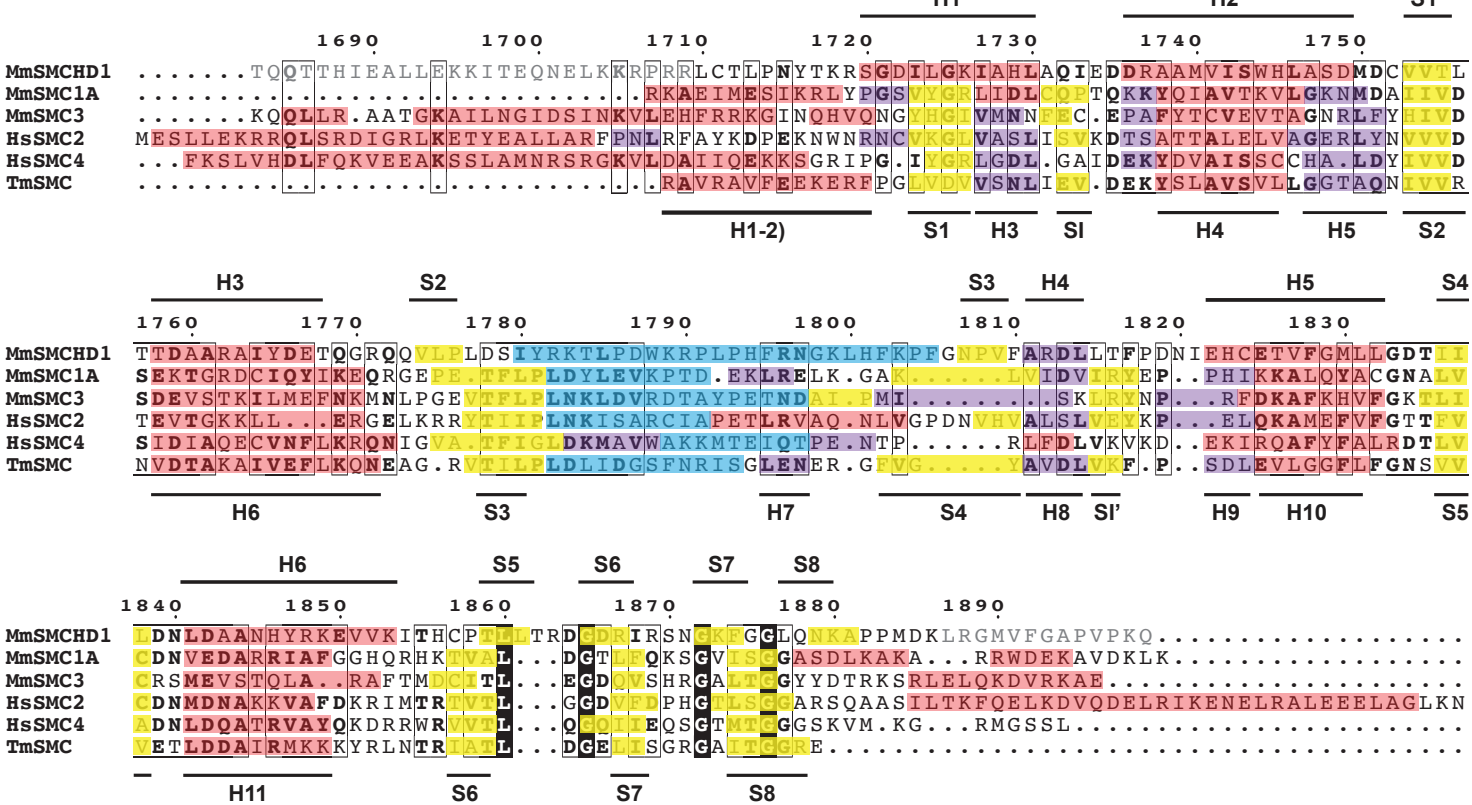
Table 2. Summary of melting temperature (T_m) and dissociation constant (K_d) values for the wild-type SMCHD1 hinge and for those of the indicated mutants.

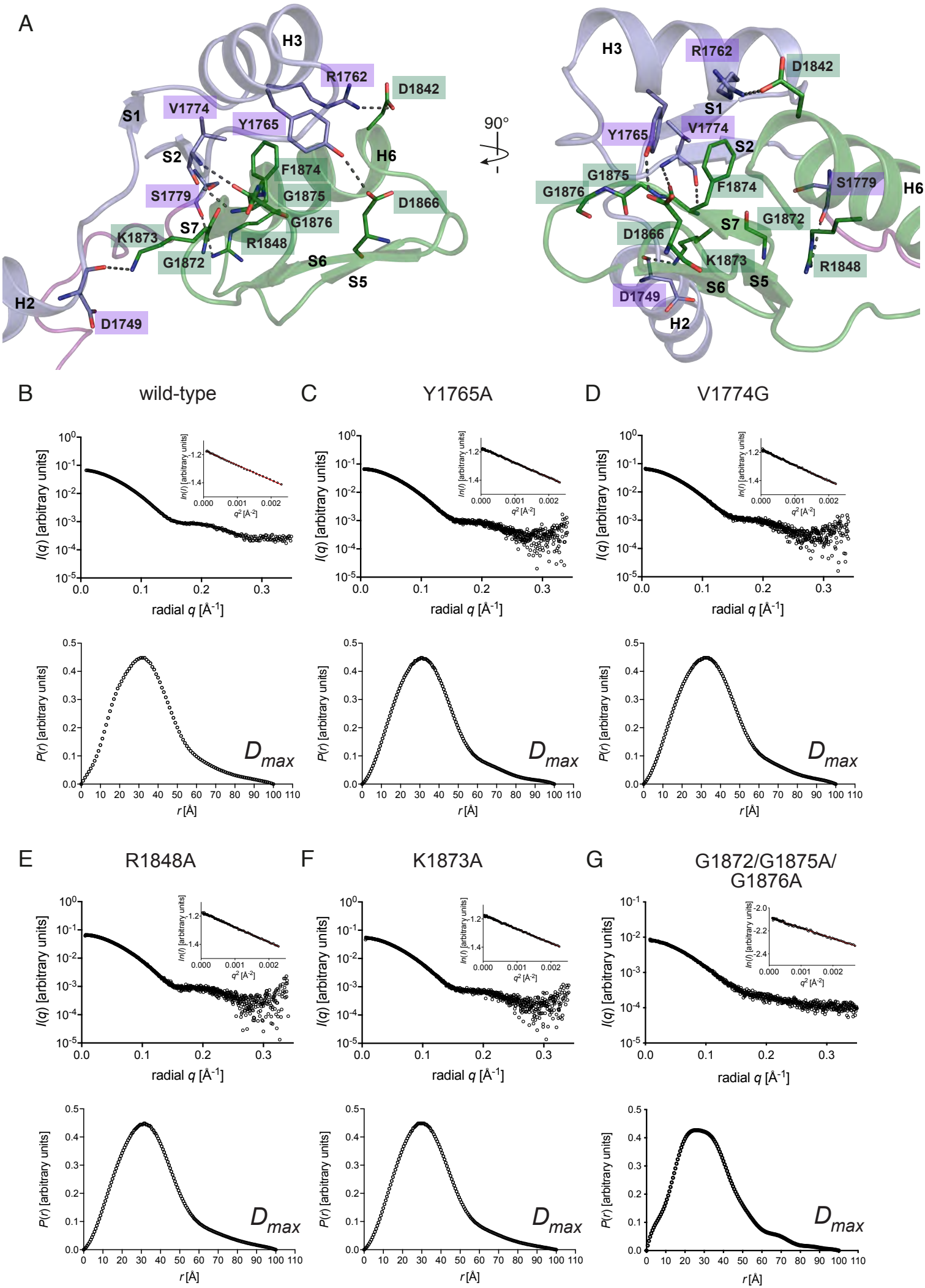
Property of SMCHD1 hinge	T_m (°C) at 0.05; 0.1 mg/ml[a]	K_d (μM)[b]
WT	53.6±0.1; 52.6±0.4	1.5±0.1
R1867G	42.4±0.1; 42.4±0.1	6.2±0.7
Dimer Interface mutants		
D1749A	46.1±0.1; 45.7±0.1	N.D.[c]
R1762A	47.1±0.1; 46.5±0.1	3.2±0.2
Y1765A	45.5±0.1; 45.0±0.1	1.1±0.1
V1774G	35.2±0.1; 34.7±0.1	N.D.
D1842A	52.0±0.2; 51.3±0.1	1.2±0.1
R1848A	42.5±0.1; 42.5±0.1	11.4±1.4
G1872A	46.1±0.1; 45.7±0.1	1.0±0.1
G1872A, G1875A, G1876A	39.1±0.1; 37.4±0.1	N.D.
K1873A[d]	68.1±0.1; 67.7±0.1	8.1±1.2
F1874A	41.6±0.3; 41.2±0.1	1.5±0.1
Cluster 1 Mutants		
K1718A	51.8±0.2; 50.9±0.1	1.4±0.1
R1719A	52.6±0.1; 51.5±0.1	1.8±0.1
R1771A	53.4±0.2; 52.0±0.2	2.4±0.2
Cluster 2 Mutants		
K1789A	52.0±0.1; 51.4±0.1	2.3±0.1
R1790A	55.6±0.1; 54.1±0.5	7.6±0.8

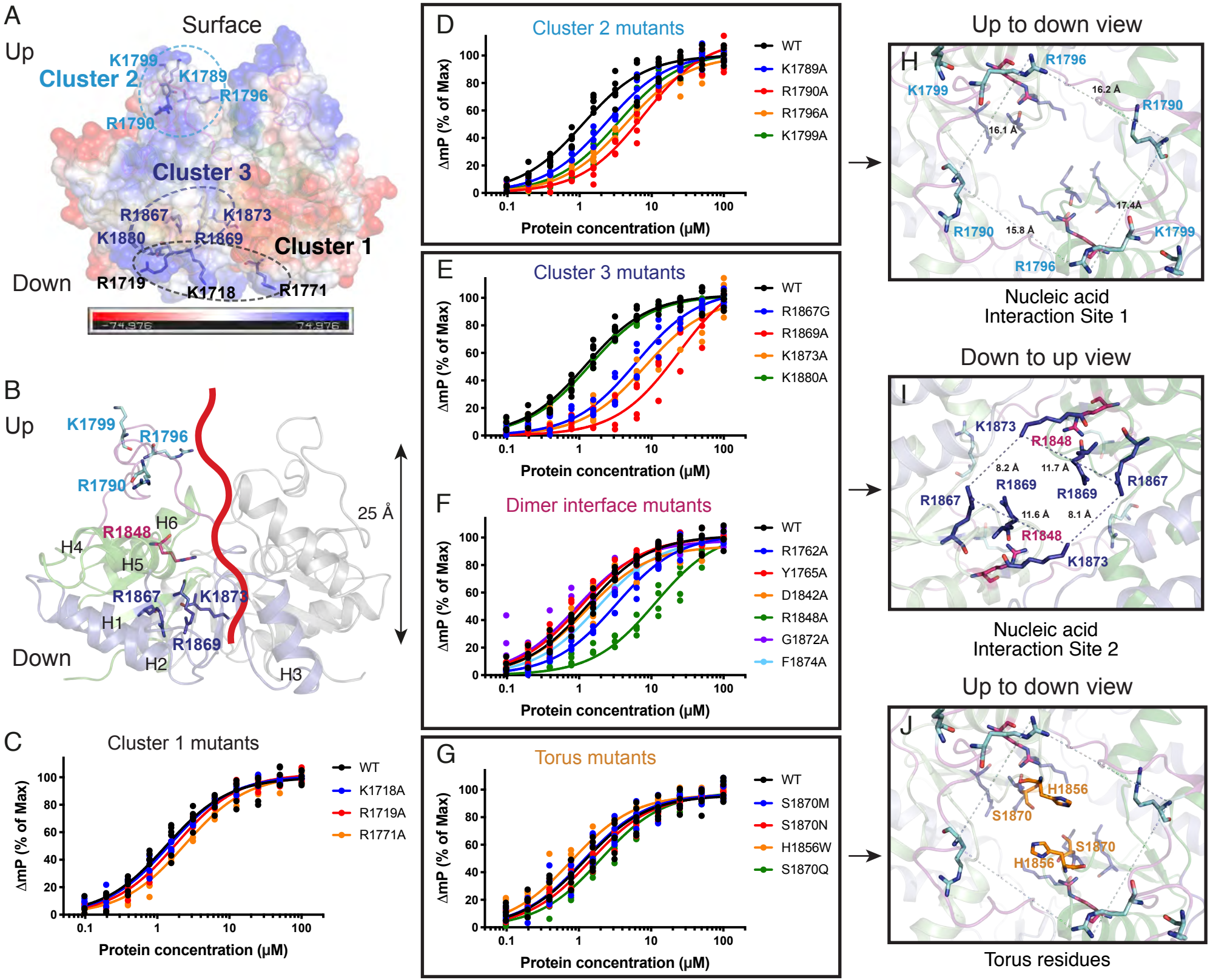
R1796A	54.2±0.2; 52.5±0.1	4.3±0.3
K1799A	52.1±0.2; 51.5±0.1	3.4±0.3
Cluster 3 Mutants		
R1869A	61.9±0.1; 61.6±0.1	23.6±4.6
K1873A[d]	68.1±0.1; 67.7±0.1	8.1±1.2
K1880A	44.9±0.1; 44.6±0.1	1.4±0.1
Torus Mutants		
S1870M	N.D.	1.2±0.1
S1870N	N.D.	1.4±0.1
S1870Q	N.D.	2.0±0.1
H1856W	N.D.	0.8±0.1

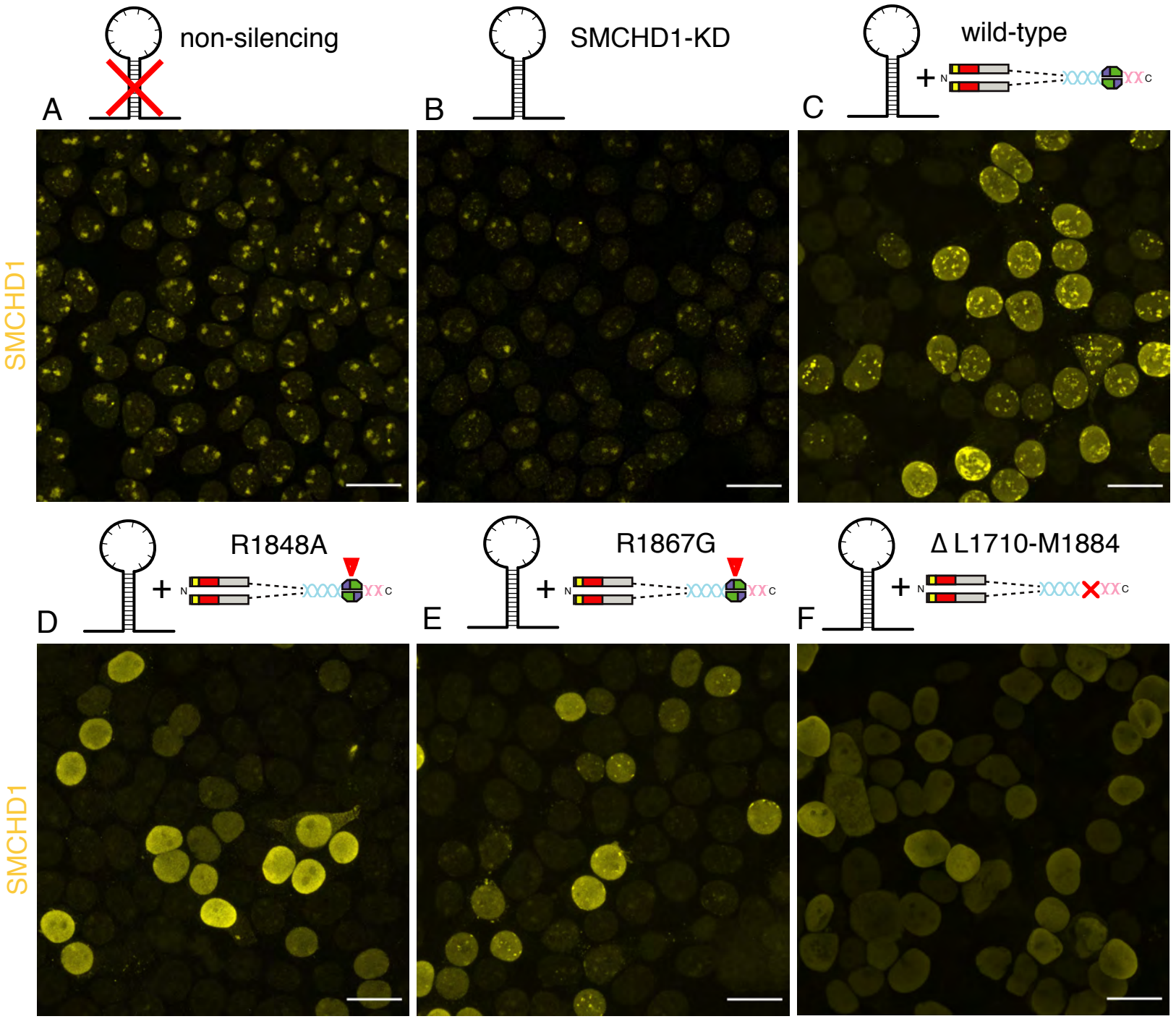
[a]T_m values are reported ± SEM for the fit calculated at protein concentrations of 0.05 and 1 mg/ml. The curves for quadruplicate measurements at each concentration are displayed in fig. S6. The T_m for the wild-type hinge domain is reported as mean ± SD of four independent experiments. [b]K_d values were estimated from the fits shown in Fig. 4 and are reported ± SEM for the fit. [c]N.D., not determined. [d]K1873A is categorized as both a dimer interface mutant and a Cluster 3 mutant.


A
SMCHD1
GHKL
UBL ATPase
N
SMC hinge
C
B
SMC
ABC-ATPase
N
C
SMC
hinge
C
Protomer 1
N
H3
H5
H6
H1
H4
C
H2
H2
C
H4
H1
H6
H5
H3
N
Protomer 2
D
Protomer 1
H1-2
H6
H10
C
H11
H3
N
H4
H9
H5
H5
H4
H4
N
H3
H9
C
H11
H10
H6
H1-2
Protomer 2
E
Subdomain I
Subdomain
Interface
Subdomain II
S1
S2
H1
H5
S4
S5
S7
S3
H2
H4
S6
S8
N
C
H3
H6
Extended Linker
F
Subdomain I
Subdomain
Interface
Subdomain II
H5
H3
H10
S1
S2
S3
SI
SI'
S4
S5
S6
S7
S8
H4
H9
H1-2
H6
N
C
Linker
H11
G
Subdomain
Interface
Subdomain I
S8
S7
Subdomain II
S6
C
S1
H1
H2
S5
H6
S4
H3
S2
S3
H4
N
H5
Extended Linker
I
Overlay
H
C
Subdomain II
Subdomain I
H1-2
Subdomain
Interface
S8
N
S4
H5
H4
S7
S6
H6
H11
S3
SI
S5
H3
S4
SI'
H9
S1
Linker
H10


H1
H2
S1
1690
1700
1710
1720
1730
1740
1750
MmSMCHD1 .....TQQTTHIEALLEKKITEQNELKKRPRRLCTLPNYTKRSGDILGKIAHLAQIEDDRAAMVISWHLASDMDCVVTL
MmSMC1A .....RKAEIMESIKRLYPGSVVGRLIDLCQPTQKKYQIAVTKVLGKNMDAIIVD
MmSMC3 .....KQQLLR.AATGKAILNGIDSINKVLEHFRRKGINQHVQNGYHGIVMNNFEC.EPAFYTCVEVTAGNRLFYHIVD
HsSMC2 MESLLEKRRQLSRDIGRLKETYEALLARFPNLRFAYKDPEKNWNRNCVKGLVASLISVKDTSATTALELVAGERLYNVVVD
HsSMC4 ...FKSLVHDLFQKVEEAKSSLAMNRSRGKVLDAIIQEKKSGRIPG.IVGRLGDL.GAIDEKYDVAISSCCHA.LDYIVVD
TmSMC .....RAVRAVFEEKERFPGLVDVVSNLIEV.DEKYSLAVSVLLGGTAQNIVVR
H1-2)
S1
H3
SI
H4
H5
S2
H3
S2
S3
H4
H5
S4
1760
1770
1780
1790
1800
1810
1820
1830
MmSMCHD1 TTDAARAIYDETQGRQQVLPLDSIYRKTLPDWKRPLPHFRNGKLHFKPFGNPVFARDLLTFPDNIEHCETVFGMLLGDTII
MmSMC1A SEKTGRDCIQYIKEQRGEPE.TFLPLDYLEVKPTD.EKLREELK.GAK.....LVIDVIRYEP..PHIKKALQYACGNALV
MmSMC3 SDEVSTKILMEFNKMNLPGEVTFLPLNKLDVRDTAYPETNDAI.PMI.....SKLRVNP...RFDKAFKHVFGKTLI
HsSMC2 TEVTGKKLL...ERGELKRRYTIIPLNKISARCIAPETLRVAQ.NLVGPDNVHVALSLVEYKP...ELQKAMEFVFGTTFV
HsSMC4 SIDIAQECVNFLKRQNIGVA.TFIGLDKMAVWAKKMTEIQTPE.NTP.....RLFDLVKVKD..EKIRQAFYFALRDTLV
TmSMC NVDTAKAIVEFLKQNEAG.RVTILPLDLIDGSFNRISGLENER.GFVG.....YAVDLVKF.P..SDLEVLGGFLFGNSVV
H6
S3
H7
S4
H8
SI'
H9
H10
S5
H6
S5
S6
S7
S8
1840
1850
1860
1870
1880
1890
MmSMCHD1 LDNLDAANHYRKEVVKITHCPTLLTRDGDRIRSNGKFGGLQNKAPPMDKLRGMVFGAPVPKQ.....
MmSMC1A CDNVEDARRIAFGGHQRHKTVAL...DGTLFQKSGVISGGASDLKAKA...RRWDEKAVDKLK.....
MmSMC3 CRSMEVSTQLA..RAFTMDCITL...EGDQVSHRGALTGGYYDTRKSRLELQKDVRKAE.....
HsSMC2 CDNMDNAKKVAFDKRIMTRTVTL...GGDVFDPHGTLSGGARSQAASILTKFQELKDVQDELRIKENELRALEEELAGLKN
HsSMC4 ADNLDQATRVAYQKDRRWRVVTL...QGQIIEQSGTMTGGGSKVM.KG...RMGSSL.....
TmSMC VETLDDAIRMKKKYRLNTRIATL...DGELISGRGAITGGRE.....
H11
S6
S7
S8


A
H3
R1762
D1842
S1
V1774
Y1765
S2
F1874
H6
S1779
G1875
D1866
G1876
S7
K1873
R1848
G1872
S6
S5
H2
D1749
90°
H3
R1762
D1842
S1
Y1765
V1774
S2
G1875
G1876
F1874
S1779
H6
S7
G1872
D1866
K1873
R1848
S6
S5
D1749
H2
B
wild-type
C
Y1765A
D
V1774G
I(q) [arbitrary units]
radial q [Å-1]
ln(I) [arbitrary units]
q2 [Å-2]
P(r) [arbitrary units]
Dmax
r [Å]
E
R1848A
F
K1873A
G
G1872/G1875A/
G1876A


A
Surface
Up
Cluster 2
K1799
K1789
R1796
R1790
Cluster 3
R1867
K1873
K1880
R1869
Cluster 1
Down
R1719
K1718
R1771
-74.976
74.976
B
Up
K1799
R1796
R1790
25 Å
R1848
H6
H4
H5
R1867
K1873
H1
R1869
Down
H2
H3
C
Cluster 1 mutants
ΔmP (% of Max)
Protein concentration (μM)
WT
K1718A
R1719A
R1771A
D
Cluster 2 mutants
WT
K1789A
R1790A
R1796A
K1799A
E
Cluster 3 mutants
WT
R1867G
R1869A
K1873A
K1880A
F
Dimer interface mutants
WT
R1762A
Y1765A
D1842A
R1848A
G1872A
F1874A
G
Torus mutants
WT
S1870M
S1870N
H1856W
S1870Q
Up to down view
H
R1796
K1799
16.2 Å
R1790
16.1 Å
17.4Å
R1790
15.8 Å
K1799
R1796
Nucleic acid
Interaction Site 1
Down to up view
I
K1873
R1848
8.2 Å
11.7 Å
R1869
R1867
R1869
R1867
11.6 Å
8.1 Å
R1848
K1873
Nucleic acid
Interaction Site 2
Up to down view
J
H1856
S1870
S1870
H1856
Torus residues


non-silencing
SMCHD1-KD
wild-type
A
B
C
SMCHD1
R1848A
R1867G
Δ L1710-M1884
D
E
F
SMCHD1
N
C