Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017

Biophysical Journal
Article

Biophysical Society

# Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms

Marvin Freitag,[1] Sigrun Jaklin,[1] Francesco Padovani,[2] Ecaterina Radzichevici,[2] Sarah Zernia,[1] Kurt M. Schmoller,[2,3] and Johannes Stigler[1,*]

[1]Gene Center Munich, Ludwig-Maximilians-Universität München, Munich, Germany; [2]Institute of Functional Epigenetics, Helmholtz Zentrum München, Neuherberg, Germany; and [3]German Center for Diabetes Research (DZD), Neuherberg, Germany

ABSTRACT Structural maintenance of chromosome (SMC) complexes form ring-like structures through exceptional elongated coiled-coils (CCs). Recent studies found that variable CC conformations, including open and collapsed forms, which might result from discontinuities in the CC, facilitate the diverse functions of SMCs in DNA organization. However, a detailed description of the SMC CC architecture is still missing. Here, we study the structural composition and mechanical properties of SMC proteins with optical tweezers unfolding experiments using the isolated Psm3 CC as a model system. We find a comparatively unstable protein with three unzipping intermediates, which we could directly assign to CC features by crosslinking experiments and state-of-the-art prediction software. Particularly, the CC elbow is shown to be a flexible, potentially non-structured feature, which divides the CC into sections, induces a pairing shift from one CC strand to the other and could facilitate large-scale conformational changes, most likely via thermal fluctuations of the flanking CC sections. A replacement of the elbow amino acids hinders folding of the consecutive CC region and frequently leads to non-native misalignments, revealing the elbow as a guide for proper folding. Additional in vivo manipulation of the elbow flexibility resulted in impaired cohesin complexes, which directly link the sensitive CC architecture to the biological function of cohesin.

SIGNIFICANCE The detailed understanding of the molecular mechanisms of SMC complexes is important to identify the origin of chromosomal defects. The coiled-coil (CC) domains were recently found to undergo large-scale conformational changes, which enable diverse functions of these proteins. However, detailed structural information of the CCs is still elusive due to their high flexibility. By measuring mechanical responses, we identified thermodynamically relevant features in the CC that structure the CC in segments. One of these features, the elbow, additionally ensures proper alignment of flanking CC sections and can be viewed as a folding guide. Furthermore, in vivo manipulation of the elbow resulted in impaired complex function, showing the relevance of the CC features for SMC mechanics.

## INTRODUCTION

Structural maintenance of chromosome (SMC) complexes are essential for the organization and regulation of chromosomes. They comprise a conserved architecture consisting of two SMC subunits forming a heterodimer bridged by a flexible kleisin unit (1) (Fig. 1 *A*). This tripartite ring is able to entrap DNA and form loops, leading to DNA compaction, chromosome segregation, and gene regulation (2). Each SMC protein consist of long anti-parallel coiled-coils (CCs), flanked by a hetero-dimerization domain (hinge) and an ATP-binding domain (head) at opposing ends (3). Recently, electron microscopy (EM) and imaging atomic force microscopy (AFM) studies have revealed that SMC complexes can adopt a wide range of shapes, ranging from open circles (O shaped) to collapsed (B shaped) or half-collapsed conformations (4–6). This started a discussion as to whether these conformational changes are relevant for the loop extrusion mechanism.

Single-molecule studies on the conformation of the SMC complex cohesin during DNA binding have shown that cohesin can bypass small obstacles bound to DNA, but not obstacles larger than 20 nm (7,8). This is surprising as the cohesin ring has a diameter of up to 50 nm (9). In addition, cohesin can be pushed by even smaller motors, such as T7 RNA polymerase or FtsK, while being topologically or pseudo-topologically bound to DNA (7,8). These findings suggest that collapsed conformations are preferred over

*Submitted May 6, 2022, and accepted for publication October 12, 2022.*

*Correspondence: stigler@genzentrum.lmu.de

Editor: Michael T. Woodside.

https://doi.org/10.1016/j.bpj.2022.10.017

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



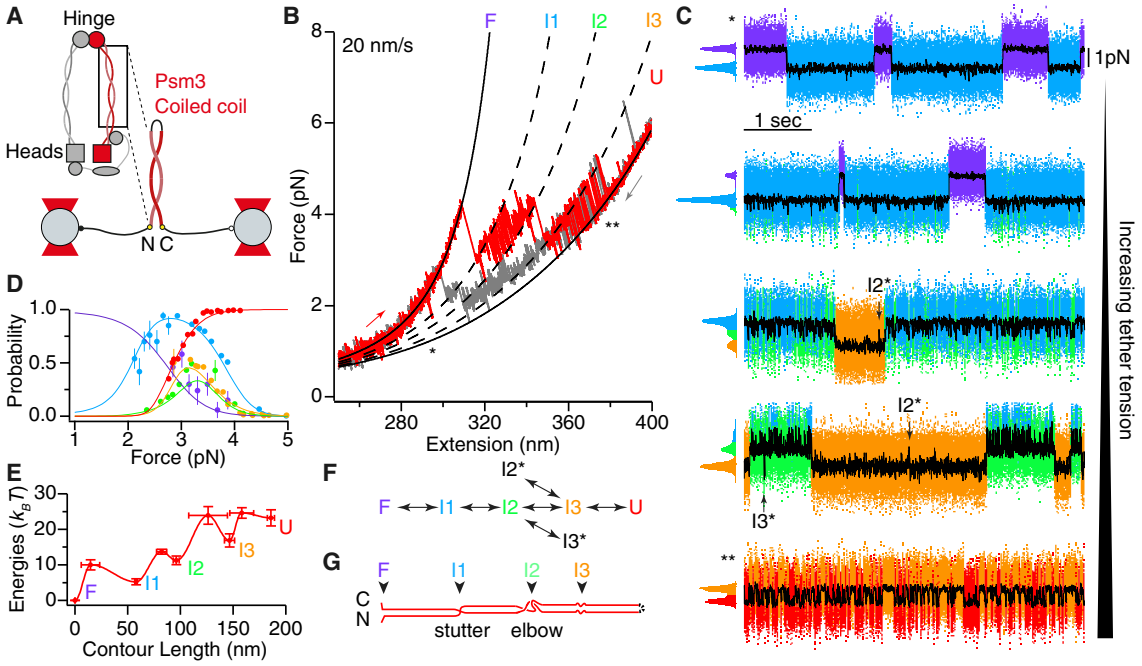


FIGURE 1 Measuring the thermodynamic stability profile of Psm3 CCs with optical tweezers. (*A*) Cartoon representation of the unfolding experiments: a large single-peptide stretch of the anti-parallel CC (D245-T505 and G658-K908) was attached to DNA handles and tethered between optically trapped beads for manipulation. (*B*) A force-extension curve of the WT construct shows a typical unfolding (red curve) and refolding trace (gray curve) containing transitions over the three obligatory intermediates (I1, I2, and I3) to the unfolded protein (U). Asterisks indicate the minimal and maximal distances used for passive mode measurements (see *C*). (C) Passive mode force versus time traces recorded at constant trap distances, demonstrating transition kinetics and the rarely populated, non-productive states I2* and I3*. Color code as in (*B*). With increased tether tension the protein populates higher energetic states. The minimal extension (single asterisk) corresponds to an average applied force of 2.9 pN, the maximal extension (double asterisk) corresponds to an average force of 4.1 pN. (*D*) Global energy fit over the force-dependent state probabilities. Error bars indicate standard deviations and result from bootstrapping of the corresponding passive mode trace (see section "materials and methods"). Color code as in (*B*). (*E*) Interpolated energy landscape consolidating contour lengths and energies of the folding states. The energy barriers are taken from force-dependent transition kinetics (Fig. S3), n = 5 molecules, error bars indicate standard deviations. (*F*) Folding network for the WT construct showing a zipper-like unfolding pattern with two non-productive misfolded states. (*G*) Identification of structural features from folding intermediates, containing the unfolded ends, the stutter, the elbow, and a possible feature located at the third intermediate (Fig. S1). To see this figure in color, go online.

fully circular ones, as also confirmed by recent cryogenic electron microscopy (cryo-EM) structures of holo complexes (10–12). However, a dynamic transition between O and B shapes is likely and was reported for condensin (4,6) and cohesin (13) in recent high-speed AFM studies. Hence, SMC CC domains have to undergo large-scale conformational changes to allow this transition.

Canonical CCs follow a heptad repeat pattern $(abcdefg)_n$, where "a" and "d" denote hydrophobic amino acids that promote pairing of the CC strands by constructing a hydrophobic core (14). CC prediction software (15,16) can be used to analyze this repeat pattern and create CC propensity profiles to identify deviations. It was found that all tested SMC proteins exhibit discontinuities in their CC domains (5,17), which provide flexibility and could enable large-scale conformational changes. Crystallography studies together with crosslink data identified two general discontinuities for bacterial and eukaryotic SMC proteins (5,18), which were named "joint" (17,19) and "elbow"(5). The joint is located near the SMC heads and increases local CC flexibility (11,20,21), while the elbow functions as a turning point, where the CCs are able to bend and fold back onto themselves. Therefore, the complex folding of the CCs is able to mediate novel interactions between the hinge and head domains, inspiring different loop extrusion models (13,22,23). Besides the identification of the features, little is known about the mechanical properties of SMC CCs, which would provide more detailed information about the capabilities of the SMC complexes in DNA organization.

Optical tweezers allow the study of protein folding features with high temporal and spatial resolution on a single-molecule basis. A precise CC unzipping pattern reveals the internal CC construction and the position of certain folding features, which appear as unfolding intermediates during protein stretching (24–26). Additionally, folding kinetics can be used to obtain the stability profile along the CC (27). For example, optical tweezers revealed that the CCs of vimentin contain several CC sections with varying stability (28). Therefore, an unfolding study of SMC CCs could pinpoint structurally and thermodynamically relevant

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



features based on the protein unfolding pattern and specify the stability of found CC sections.

Here, we study the structural composition and mechanical properties of SMC proteins with optical tweezers unfolding experiments using the non-canonical Psm3 CC as a model system. We find a relatively unstable protein compared with other CCs. We show that the selected CC region unfolds through three obligatory intermediates: a stutter, the SMC elbow, and a not-predicted feature at the hinge-proximal site, and is divided by these features into four sections. Misalignments of the wild-type (WT) CC strands are rarely observed. However, replacing the elbow induces frequently appearing metastable and non-productive misfoldings, which highlights the elbow as a proper guide for further CC alignment. Furthermore, in vivo experiments using changes in cell volume as an indicator for impaired cohesin complexes revealed the elbow flexibility as crucial for cohesin's biological function.

## MATERIALS AND METHODS

All reagents were purchased from New England Biolabs (NEB, Ipswich, MA), Merck (Darmstadt, Germany), or Thermo Fisher Scientific (Waltham, MA) unless otherwise stated.

### Protein expression and sample preparation

A pET28a-based plasmid containing the CC construct was transformed in Rosetta (DE3). Cells were grown at 37°C to optical density 600 ($OD_{600}$) = 0.4–0.6. The expression was induced by 0.4 mM IPTG and carried out at 18°C overnight. Harvested cells were resuspended in lysis buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole) supplemented with 1 mM PMSF and 0.1 mg/mL lysozyme, sonicated, and centrifuged. The supernatant was loaded onto a HisTrap FF-columns (GE Healthcare, Chicago, IL), washed with 10 CV lysis buffer, and eluted in lysis buffer containing 200 mM imidazole. Pooled fractions were gel filtrated using a Sephacryl S300 (GE Healthcare) or Superdex S200 (GE Healthcare) column. Eluted proteins were equilibrated in measuring buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl) for ybbR-tagged proteins or in measuring buffer at pH 7.2 for proteins with terminal cysteines.

From 5 to 15 $\mu$M proteins containing terminal ybbR-tags (WT, ΔEB, $\Delta EB_N$, $\Delta EB_C$) were coupled with 3′-coenzyme A (CoA)-oligos (Biomers, Ulm, Germany) in 2 molar excess with twofold excess of Sfp (made in-house, plasmids were kindly provided by Prof. H.E. Gaub, LMU Munich) and 10 mM $MgCl_2$ for 120 min at room temperature and reduced with 0.5 mM TCEP for 30 min before purifying with a size-exclusion column (Superdex S200 or Yarra 3000; Phenomex, Torrance, CA). The TR construct containing terminal cysteines (10–20 $\mu$M) was reduced with 0.5 mM TCEP 30 min before coupling to 3′-maleimide-oligos (Biomers, Ulm, Germany) in at least 2.5-fold molar excess. After 120 min at room temperature or an overnight reaction at 4°C, the mix was loaded onto the size-exclusion column. Fractions containing the protein coupled to two oligos were identified as described in (29) by SDS-PAGE stained first with SYBR Gold and afterward with SimplyBlue (Fig. S2).

DNA handles containing a dual-biotin or dual-digoxigenin modification and the single-stranded DNA (ssDNA) overhang were produced by PCR (30) and suitable ratios of handles to protein-DNA-chimeras were determined as reported earlier (31). Streptavidin-coated and anti-digoxigenin-coated silica beads were functionalized from carboxyl-functionalized 1-$\mu$m beads (Bangs Laboratories, Fischers, IN) as reported previously (32). Handles and protein-DNA chimeras were incubated with anti-digoxigenin beads and mixed with streptavidin beads in a passivated custom-made chamber, as described previously (33). Measurements were performed in measuring buffer containing the glucose oxidase oxygen scavenging system (0.65% w/v glucose, 13 U/mL glucose oxidase, and 8500 U/mL catalase).

### Cysteine-based crosslinking of CCs

BMOE-crosslinks were made analogous to (34). The proteins (5 $\mu$M) were reduced for 60 min on ice, then BMOE (20 mM stock solution in DMSO) was added 1:20 and the reaction was quenched with excess of $\beta$-mercaptoethanol. Additionally, proteins were labeled with a fluorescent CoA-LD555 dye (Lumidyne, New York, NY, custom synthesis) according to the above-stated coupling procedure using the CoA-dye instead of CoA-oligo.

### Prediction software and replacement of the elbow feature

Discontinuities in the Psm3 heptad pattern were identified using CC prediction software (15,16). The amino acids of the N-terminal Psm3 CC (F385-W394) were replaced with the amino acid sequence QMQRINSEISD for the corresponding ΔEB mutants. Additionally, the heptad repeat of the C-terminal CC was continued by deletion of P780 for Psm3 for the ΔEB mutants.

In addition, the CC propensity profile was generated with DeepCoil (35). N-terminal CC and C-terminal CC were analyzed separately due to length limitations of DeepCoil. The AlphaFold model was derived from the AlphaFold Protein Structure Database (36,37). Models for the WT elbow region and the ΔEB elbow region were generated using ColabFold with I337-I406 and N824–L760 of Psm3 linked by GGSGGSGGSGGS for the WT construct and the same amino acid sequence with the above-described mutations for the ΔEB construct (36,38–41). Illustrations were made with ChimeraX (42).

### Strain construction, microscopy, and cell volume measurements

For the Smc3 CC, the replacement mutations were designed analogously to the Psm3 mutations. The N-terminal section (F393-W402) was replaced with QMQRINSEISD and the replacement of Q791, E792, and F793 on the C-terminal CC to threonine continued the Smc3 heptad repeat pattern. The single mutations only contained the respective ΔEB mutants.

All yeast strains used in this work are based on W303 and were constructed using standard methods. Full genotypes are listed in the Supporting Materials (Table S3). The DNA used to transform each strain was obtained by PCR amplification of plasmids listed in the Supporting Materials (Table S4). The PCR product was then integrated into the *SMC3* endogenous locus. All plasmids carry the C terminus of (mutated or WT) *SMC3* with the *CglaTRP1* marker. The correctness of the mutated *SMC3* sequence was confirmed by Sanger sequencing.

Cell cultures (4 mL) were grown at 30°C in synthetic complete medium containing 2% glucose (SCD) for exactly 24 h. The $OD_{600}$ (measured with the Thermo Fisher Scientific, NanoDrop 2000 spectrophotometer) was maintained below 1.0 through appropriate dilutions during culture growth. Cell volume distributions were measured with a Coulter counter (Beckman Coulter, Brea, CA, Z2 Particle Counter). The measurements are based on at least four independent biological replicates. Each biological replicate consisted of two technical replicates. We note that, to reproduce our results, it is important to measure cell volume after 24 h of growth in SCD. This is because we noticed that the size distribution of the strains carrying the single N-terminal mutation shifted toward WT-like distribution when keeping the cells in exponential growth ($OD < 1.0$) for longer than 24 h. To exclude an artifact due to contamination, we re-plated cultures after 72 h of growth

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



in SCD, and picked single clones to repeat the experiment. After 24 h of growth, we found a size phenotype consistent with our original result shown in Fig. 5.

For microscopy, cells were grown with the same conditions as for the Coulter counter measurements. Next, wells of chambered coverslips (μ-Slice 8 Well, ibidi) were covered with 200 μL of 1 mg/mL (in water) Concavalin A and incubated for 10 min. The wells were subsequently washed twice with water and twice with SCD medium. Then, cells were sonicated for 10 s and 200 μL of the sonicated suspension was transferred to the coated well and left to stabilize for about 5 min, after which the medium was removed and the unbound cells were washed away twice with SCD medium. Finally, the well was covered with 200 μL of SCD medium and mounted onto the microscope. Images were taken on a Zeiss LSM 800 microscope (Carl Zeiss, Oberkochen, Germany, Zen 2.3, blue edition software) with an Axiocam 506 camera using a Plan-Apochromat 63×/1.4 Oil DIC objective. The cells were imaged in bright-field mode. The images were then segmented with YeaZ (43) embedded in the bioimage analysis software Cell-ACDC (44), which also automatically estimates the cell volume from 2D segmentation masks.

## Optical trap setup and measurement modes

Optical tweezers experiments were performed on a LUMCIKS C-trap (LUMICKS, Amsterdam, the Netherlands) with embedded smooth motion update by active pulling on the tethered molecule with a fixed pulling and relaxation speed. In passive mode, the molecule was held at a constant trap distance in distinct steps. Data were recorded at a sampling rate of 78.125 kHz and downsampled by a factor of 4 for additional analysis.

## Polymer models and length coordinates

Force-extension curves recorded by active pulling on the tethered molecule follow the stretching response of the DNA handles, which was modeled with the extensible worm-like chain eWLC model:

$$F(\xi_D) = \frac{k_B T}{p_D}\left(\frac{1}{4}\left(1-\frac{\xi_D}{L_D}+\frac{F}{K}\right)^{-2}-\frac{1}{4}+\frac{\xi_D}{L_D}-\frac{F}{K}\right) \quad (1)$$

with DNA extension $\xi_D$, force $F$, thermal energy $k_B T$, DNA persistence length $p_D$, DNA contour length $L_D$, and the stretch modulus $K$.

The additional force-extension of an unfolded protein was modeled by using a worm-like chain (WLC) model in series with the DNA handles' response (above):

$$F(\xi_p) = \frac{k_B T}{p_p}\left(\frac{1}{4}\left(1-\frac{\xi_p}{L_p}\right)^{-2}-\frac{1}{4}+\frac{\xi_p}{L_p}\right) \quad (2)$$

with unfolded protein extension $\xi_p$, protein persistence length $p_p$ and protein contour length $L_p$.

The curves were fitted using a temperature $T$ of 296 K, the DNA stretch modulus $K$ was fixed to 800 pN/nm, and the protein persistence length $p_p$ was fixed to 0.7 nm. Resulting DNA persistence lengths were typically $p_D \sim$ 10–20 nm and resulting DNA contour lengths $L_D$ were ~350–370 nm.

## Total free-energy estimation based on the Crooks fluctuation theorem

The Crooks fluctuation theorem is applicable for "stochastic, microscopically reversible dynamic" (45) and provides an estimation of the total free energy $\Delta G^0_{F\text{-}U}$ from non-equilibrium force versus distance curves (46):

$$\frac{P(W)}{P(-W)} = e^{\frac{W-\Delta G^0}{k_B T}} \quad (3)$$

with the work distribution during unfolding $P(W)$ and the work distribution during folding $P(-W)$. The work was determined as the area below a force versus distance curve and the contributions from the DNA handles were subtracted with respect to the linker models (27).

## Equilibrium energies, barrier heights, rate models, and contour length transformation

Passive mode measurements were analyzed as previously described (28,47). In brief, the different passive mode traces were modeled using hidden Markov modeling (48). The errors of the force-dependent state probabilities were gained from bootstrapping from each individual constant-distance trace with 200 iterations (28) , and resulting standard deviations are shown. The equilibrium energies $\Delta G^0_{ij}$ between folding states i,j were calculated by global fits to an equilibrium model:

$$P_i(F_i) = \left(1+\sum_{j\neq i} exp\left(\frac{-\Delta G^0_{ij} - \Delta G^{sys}_{ij}\left(F_i, F_j\right)}{k_B T}\right)\right), \quad (4)$$

where $P_i(F_i)$ is the force-dependent state probability and $\Delta G^{sys}_{ij}$ is the energy contribution resulting from the bead deflection from the trap center (assuming a Hookean spring model), the stretching of the DNA handles, and the stretching of the unfolded polypeptide (integrals over eWLC and WLC models, see above) and can be determined based on the recorded data (27), whereas the equilibrium energy $\Delta G^0_{ij}$ can be determined by fitting (for a detailed description, please see (49)).

The force-dependent rate models with missed events correction and the contour length transformations were done for the protein-dumbbell system as described in (47), and the energy barrier heights were determined to construct an interpolated energy landscape using

$$\frac{\Delta G^{\dagger}}{k_B T} = \ln\left(\frac{k_0}{A}\right),\ A = 1.2 \times 10^4\ s^{-1} \quad (5)$$

with energy barrier height $\Delta G^{\dagger}$, extrapolated zero-force folding rate $k_0$ and an Arrhenius-factor $A$, which estimates the frequency of folding attempts of the construct (27).

## Classification of force-extension curves for the ΔEB constructs

For classification of the ΔEB F-I1 folding transition, the rupture of the initial DNA stretch was monitored in 500 nm/s force-extension curves of all constructs. If the rupture value was below 2 pN, corresponding force-extension curves lacking a folded F-I1 section were classified as non-native-like pulling traces. Traces were then transformed into contour space (50), where the population of non-productive, metastable states was clearly distinguishable from WT-like folds due to their longer residence time. Consequently, pulling traces showing intermediates resisting higher forces than 8.5 pN, which were never reached for the WT construct at that pulling speed, were classified as misaligned configurations. The ΔEB analysis includes 11 molecules with at least 20 cycles summing up to a total of 436 cycles of 500 nm/s, $\Delta EB_N$ analysis includes five molecules with 769 cycles of 500 nm/s, and $\Delta EB_C$ analysis includes seven molecules with 486 cycles

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



of 500 nm/s. Errors were given as standard deviations of the single-molecule distribution.

### Simulation of force-extension curves from interpolated energy landscapes

The Monte Carlo simulations were performed as described previously (28,51). In brief, in the modeled system two trapped beads are linked with eWLC handles to the protein. The transition probabilities were taken from the force-dependent rates (Table S2; Fig. S3). The bead deflection and protein contour length were updated at each time step ($\Delta t = 10^{-5}$ s) and sampled at 20 kHz, mimicking experimental conditions. Force-extension curves were generated for the WT and TR construct. To generate force-extension traces for ΔEB, the WT energy landscape was iteratively adjusted such that the resulting force-extension curves matched the experimental WT-like force-extension curves of ΔEB (Fig. S6 *B*). To achieve this, the I2 state had to be eliminated and the resulting energy of the I3 state and the energy barrier of the I1-I3 transition were adjusted.

## RESULTS

### The Psm3 CC shows low-force unfolding over three intermediates

To study the mechanics of SMC CC domains, we engineered a single-peptide construct of the isolated CC region of the fission yeast cohesin subunit Psm3 (Fig. 1 *A*). Lacking high-resolution structural information about the entire CC, we identified endpoints of a suitable construct relying on crosslinking studies (52) and CC prediction software (15,16,35). In detail, the head-proximal endpoints (D245-K908) of Psm3 were found from sequence alignment of human Smc3 crosslinks. Using prediction software (Fig. S1 *A*), the CC fold was verified and spans at least to T505 and G658 at the hinge-proximal site. The utilized construct containing the selected stretches was equipped with terminal ybbR-tags, which served as attachment points for dsDNA handles (Figs. 1 *A*, S1 *A*, and S2). A 16-amino-acid (aa) flexible linker replaced the hinge region between T505 and G658. We then manipulated this construct in an optical tweezers instrument using a zipper geometry (Fig. 1 *A*).

Force-extension curves recorded at a pulling speed of 20 nm/s showed first unfolding events between ∼2 and 4 pN. This initial transition is followed by subsequent un- and refolding transitions with three obligatory intermediates close to equilibrium (red curve in Fig. 1 *B*). We used polymer models (53) to determine the length of the folded CC sections and found a length of 185.9 ± 3.4 nm for the unfolded construct U (red). All errors represent standard deviations, unless stated otherwise. This corresponds to 509 ± 10 aa, 19 aa shorter than the 528 aa contained in the construct. This suggests that the head-proximal amino acid of the CC construct does not contribute to the stably folded part of the CC, since no further unfolding event corresponding to an unfolding of these 19 aa was observed. The intermediates, which we termed I1 (blue), I2 (green), and I3 (orange in Fig. 1 *B*), consistently appeared at contour lengths of 58.0 ± 1.3 nm, 96.0 ± 2.9 nm and 146.5 ± 4.9 nm, respectively (see Table S1). The construct readily refolded into the native conformation during relaxation (gray curve in Fig. 1 *B*).

To characterize the exchange between the intermediates in greater detail, we next recorded equilibrium fluctuation traces in passive mode with a constant trap distance for each trace (27). Different applied tensions revealed the kinetics and energetics of the protein folding transitions (Fig. 1 *C*). At low distances (single asterisk in Fig. 1 *B* and *C*), the protein was most present in the folded configuration F (native state, purple), while partial unfolding into the I1 intermediate state was observed. At higher tensions, the construct was most present in the I1 intermediate, while first transitions into the I2 state could be resolved. Following traces demonstrate further unfolding from I2 into the I3 state. Finally, the population of states shifted toward U for higher trap distances (double asterisk in Fig. 1 *B* and *C*). From these traces, the stability profile of the Psm3 CC can be revealed. First, we used hidden Markov models (48) to identify the force-dependent populations (Fig. 1 *A*) and rate constants (Fig. S3 *A*). Next, we applied mechanical linker models to determine the thermodynamic energies of the states (27). The resulting energy of the fully folded CC is $\Delta G^0_{\text{F-U}} = (23.2 \pm 2.3)\ k_BT$. The intermediates are at 5.2 ± 0.9 $k_BT$, 11.2 ± 1.2 $k_BT$ and 16.9 ± 1.9 $k_BT$ measured from the native state (see Table S1), which agrees with the total free energy derived from the Crooks fluctuation theorem (45) of 19.1 ± 2.8 $k_BT$ based on non-equilibrium force-distance curves (Fig. S4 *A* and *B*). Furthermore, we determined the barrier heights from the kinetics of the states by using the extrapolated zero-force unfolding rates (Fig. S3 *A*) and assumed an Arrhenius factor of $1.2 \times 10^4\ s^{-1}$ (27) (see Table S2). This allowed us to generate a putative energy landscape for the unzipping of the Psm3 CC (Fig. 1 *E*), which revealed faster kinetics for the I1-I2 section compared with the F-I1 or I2-I3 section due to the lower energy barrier height between I1 and I2. The I3-U energy profile is comparatively flat after the energy barrier. However, the profile in this region could be masked by the introduction of the hinge-replacing flexible linker. By analyzing $\Delta G^0$ per contour length for each intermediate stretch, we found that the energy stored inside this CC construct is quite uniformly distributed and averages to 0.12 $k_BT$/nm (Table S1; Fig. S4 *C*), which means that the Psm3 CC is a thermodynamically less stable fold compared with other CC proteins (25–28).

Generally, folding and unfolding transitions proceeded sequentially in the scheme F-I1-I2-I3-U, suggesting that the Psm3 CC folds and unfolds in a zipper-like fashion. However, a detailed analysis of many molecules also revealed rarely populated and short-lived off-pathway states, which we called I2* (mean dwell time, $\tau_{I2}^* = 5.2 \pm 4.7$ ms) and I3* ($\tau_{I3}^* = 4.0 \pm 2.4$ ms), respectively (Fig. 1 *C* and *F*). Note that their contour lengths of

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



98.7 ± 3.4 nm and 146.6 ± 5.1 nm are indistinguishable from the on-pathway states I2 and I3. However, because these I2* and I3* are significantly less stable than I2 and I3, respectively ($\Delta G^0_{I2\text{-}I2}{}^* = 8.4 \pm 2.2\ k_BT$, $\Delta G^0_{I3\text{-}I3}{}^* = 5.5 \pm 2.3\ k_BT$) and always refolded into the on-pathway states, we suspect that they represent misaligned configurations of the I2-I3 section. Taken together, we found for the Psm3 CC a general zipper-like transition pattern between F and U, with occasional off-pathway deviations (Fig. 1 *G*).

## Crosslinking corroborates unfolded amino acids in the head-proximal region

To solve the conundrum of the missing 19 aa in the folded conformation (see above) and unambiguously determine the pairing of the CC in the head-proximal region, we created a series of cysteine point mutations to precisely identify adjacent amino acids in the folded conformation by BMOE crosslinking experiments (34) (Fig. 2 *A*). Based on CC prediction software (15,16) we chose the most head-proximal C-terminal "d" amino acid of the CC (D906) as an anchor point and tested its association to seven "a" amino acids on the N-terminal strand (R246, N253, F260, I267, I274, L281, K288). We expressed and purified seven proteins containing double cysteine mutations and tested whether these constructs were crosslinked by the homo-bifunctional zero-length crosslinker BMOE. The efficiencies were evaluated by gel electrophoresis and fluorescence imaging (Figs. 2 *B* and S5).

We found a significant crosslinking efficiency of D906C to R246C, N253C, F260C, and I267C (Fig. 2 *B*, top panel). Surprisingly, the range of identified pairs exceeds the length of BMOE (8.0 Å). We conclude that the region around D906 must be flexible and hence not in a well-folded CC (17). This agrees with the observation of the missing 19 aa in pulling experiments (see above). To identify the actual start of the well-folded CC section, we performed an additional series of crosslink experiments with an amino acid (L892) further inside the CC as an anchor point (Fig. 2 *B*, bottom panel). In contrast to D906, L892C only paired with two amino acids on the N-terminal strand (F260C and I267C), which is compatible with the BMOE length and consistent with a well-folded CC close to L892. We conclude that the head-proximal CC region of our construct is well folded until I267 and L892 but is likely unfolded around R246 and D906 (Fig. 2 *C*).

## Intermediates occur at non-canonical CC features

We next sought to use the length information obtained from our measurements to pinpoint the structural location of the intermediate states to features inside the CC. This assignment is straightforward if the CC adopts a regular fold without discontinuities. However, structural and biochemical studies have indicated that SMC CCs deviate from a regular CC fold (5,17) and may contain several non-canonical features such as loops or linker regions (54), which lead to shifts in pairing along the hydrophobic core of the protein (Fig. S1 *A*). In our tweezers experiments, we measure the number of amino acids that free up during an unfolding event independently of a symmetric or shifted pairing of the CC, which results in ambiguous assignments of folding features (Fig. S1 *B*).

We therefore used computational prediction software to identify non-canonical features of the CC. First, we studied the predicted CC propensities by DeepCoil (35,55,56) to identify possible separated CC sections (Fig. S1 *A*). The minima in the profile indicate three or four non-structured linkers between sections of continuous CC (Fig. S1 *A*). In addition, PCOILS (16) and MARCOIL (15), which estimate the CC heptad repeat pattern, showed two sites with shifts in CC pairing for the selected Psm3 CC region. Next, we used the atomic model of the whole Psm3 protein predicted by AlphaFold (36,37) and identified the elbow and a stutter as non-canonical CC features of the selected Psm3 region. In the model, the stutter is located close to P831, while the elbow is predicted as non-alpha helical linkers on both CC strands, located at F385-T387 and S773-D779 (Fig. S1 *C*, top), which is close to the elbow predicted by crosslinks K388-Q398 and D779-L789 (18). Interestingly,

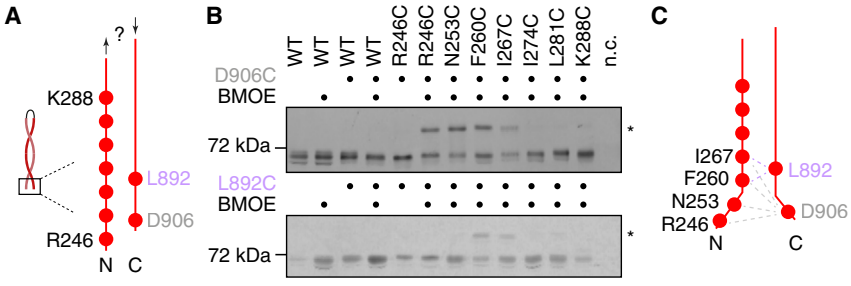


FIGURE 2 Crosslinking of terminal amino acids reveals fraying at the termini. (*A*) Cysteine-based BMOE crosslinking was performed for the anchor amino acids D906 (gray) and L892 (light purple), which are located inside the predicted hydrophobic core and at the C-terminal end of the WT construct. Seven positions inside the hydrophobic core of the N-terminal end, ranging from R246 to K288, were tested as possible crosslink partners. (*B*) BMOE-crosslinked proteins bearing denoted cysteine point mutations were additionally labeled with CoA-dyes via their ybbR-tags and further investigated by SDS-PAGE and fluorescent imaging. Top, D906 crosslinks; bottom, L892 crosslinks. WT, C-terminal single mutations, controls without BMOE, and negative control (n.c.) lacking a CC construct display no crosslinks (only lower band or no band, respectively), whereas certain double mutations show a high fraction of crosslinked proteins (higher band, asterisks). A quantitative analysis of the crosslink efficiency can be found in Fig. S5. (*C*) Summary of the crosslink series. Pairing partners of D906 (gray) and L892 (light purple) indicated with dashed lines. To see this figure in color, go online.

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



the prediction programs agree on a pairing shift between the CC strands of two amino acids at the stutter and five amino acids at the elbow.

To assign our measured intermediate contour lengths to paired amino acids in the protein, we used the pairing at the CC termini identified by our crosslinking experiments (Fig. 2 *B*) to find the folded state F and extrapolate further CC alignment. The measured contour length of the folded construct (F – U: 509 ± 10 aa) significantly exceeds the crosslink-based minimal paired amino acids I267 and L892, which includes only 488 aa (238 aa from the N-terminal coil I267-T505, 16 aa from the linker, and 234 aa from the C-terminal coil G658-L892). Therefore, the folded state starts around S257 and I903, which agrees with the pairing of I267-L892 and includes 509 aa (F: S257-I903 ± 10 aa). We hence used the measured number of unfolded amino acids to locate the intermediates in the protein with respect to the predicted pairing shifts (I1, I337-N824 ± 4 aa; I2, D392-M775 ± 8 aa; I3, R461-D706 ± 14 aa; also see Fig. S1 *C*, bottom). Interestingly, the stutter and the elbow lie close to the detected intermediates I1 and I2.

## A truncated CC shows regular elbow folding

In contrast to I1 and I2, which coincided with sites of predicted non-canonical CC features, the intermediate I3 was clearly identified in pulling experiments but not predicted. We therefore wondered if I3 folding is a necessary prerequisite for productive folding and designed a truncation mutant TR, where the hinge-proximal region was shortened to G445 and T733, eliminating I3 (Fig. 3 *A*). Pulling cycles of TR showed I1 and I2 features at the same contour lengths as for WT (Fig. 3 *B*), but lacked I3 features. In addition, passive-mode measurements revealed that the population and energetics of I1 and I2 are identical to WT (Figs. 3 *C–E* and S3 *B*). Interestingly, the I2-$U_{TR}$ section showed faster kinetics than the I2–I3 section of the WT construct. Furthermore, TR did not misfold into the intermediates I2* and I3*, which both require a functional I3 according to the stated on-pathway folding network (Fig. 1 *F*). We conclude that Psm3 folds consecutively and proper F-I1 and I1-I2 folding does not require a folded I3.

## Replacing the elbow results in rare non-productive, metastable folds

Since the elbow is a central part of SMC CCs and allows collapsed conformations of SMC complexes (13,22), we wondered how the CC will be aligned without the elbow. Recent findings showed that altered SMC CC length and even point mutations of aromatic amino acids inside the elbow region can be lethal (5,20). Therefore, we first identified conserved aromatic amino acids in the N-terminal elbow region of Psm3 and found a non-helical sequence that causes a heptad discontinuity. We replaced this elbow sequence with a suitable CC section on the N-terminal CC strand ($\Delta EB_N$). For the C-terminal strand, a deletion of P780 was sufficient to complete the heptad repeat pattern ($\Delta EB_C$). Furthermore, a combination of $\Delta EB_N$ and $\Delta EB_C$ excludes the elbow on both strands (ΔEB), and possesses the same number of amino acids as WT. In accordance with our design, prediction software confirmed the improved CC propensity and indicates a straightening of the CCs in the double mutant ΔEB (Fig. 4 *A*).

In contrast to WT, we found that all mutants (ΔEB, $\Delta EB_N$, $\Delta EB_C$) displayed more diverse force-extension curves. Stretching traces often started at intermediate contour lengths, with many traces exhibiting unexpectedly high unfolding forces, indicating that, during relaxation, the mutants frequently misfolded. The misfolded configurations occurred in 2% ± 2% of the traces for $\Delta EB_N$ and in 3% ± 2% of the traces for $\Delta EB_C$, whereas, in the double mutation ΔEB, we observed 17% ± 8% misfolded configurations (Fig. 4 *B*). This suggests an additive effect of both mutations on the CC folding. Hence, we characterized the misfolding of the ΔEB mutant in more detail. In 84% ± 7% of ΔEB pulling traces recorded at 500 nm/s we found a folded F-I1 section similar to WT and no intermediate comparable with the WT I2 state. The remaining pulling traces exposed configurations that had not folded into the native state F during previous relaxation (Fig. S6 *C*, blue). Therefore, these configurations can be considered as misfolded kinetic traps. However, here, the consecutive relaxation (Fig. S6 *C*, gray) showed refolding into the native state, suggesting that the protein is able to frequently exchange between folded and misfolded states. For a closer look, we display representative cycles at 20 nm/s showing unzipping of folded and misfolded configurations (Fig. S6 *B* and *C*). Next, we focused on the novel ΔEB-specific intermediates that ruptured at unexpectedly high forces (Fig. 4 *B*, blue traces). Here, the protein adopts half-folded conformations, which are non-productive as further pairing of unfolded, terminal amino acids is not observed. In contrast to WT and TR, these configurations of ΔEB still resisted high forces at 20 nm/s (Fig. S6 *C*), which concealed information about the consecutive I3-U section. A closer look by reducing the pulling speed to 2 nm/s revealed various misaligned configurations of the I1-I3 section, which are clearly different from the WT I2 state and are followed by the familiar WT-like I3-U transitions (Fig. S6 *D*). Interestingly, the misfolded states exchanged over the I3 intermediate (arrowhead in Fig. S6 *E*) and not over the fully unfolded protein U. We conclude that the heptad discontinuity induced by the Psm3 elbow, which is present in WT but has been removed in ΔEB, ensures a native-like alignment of the head-proximal CCs, and that the replacement of the elbow by a continuous CC section frequently leads to non-native configurations. The elbow, albeit not being a canonical CC feature, is therefore an

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



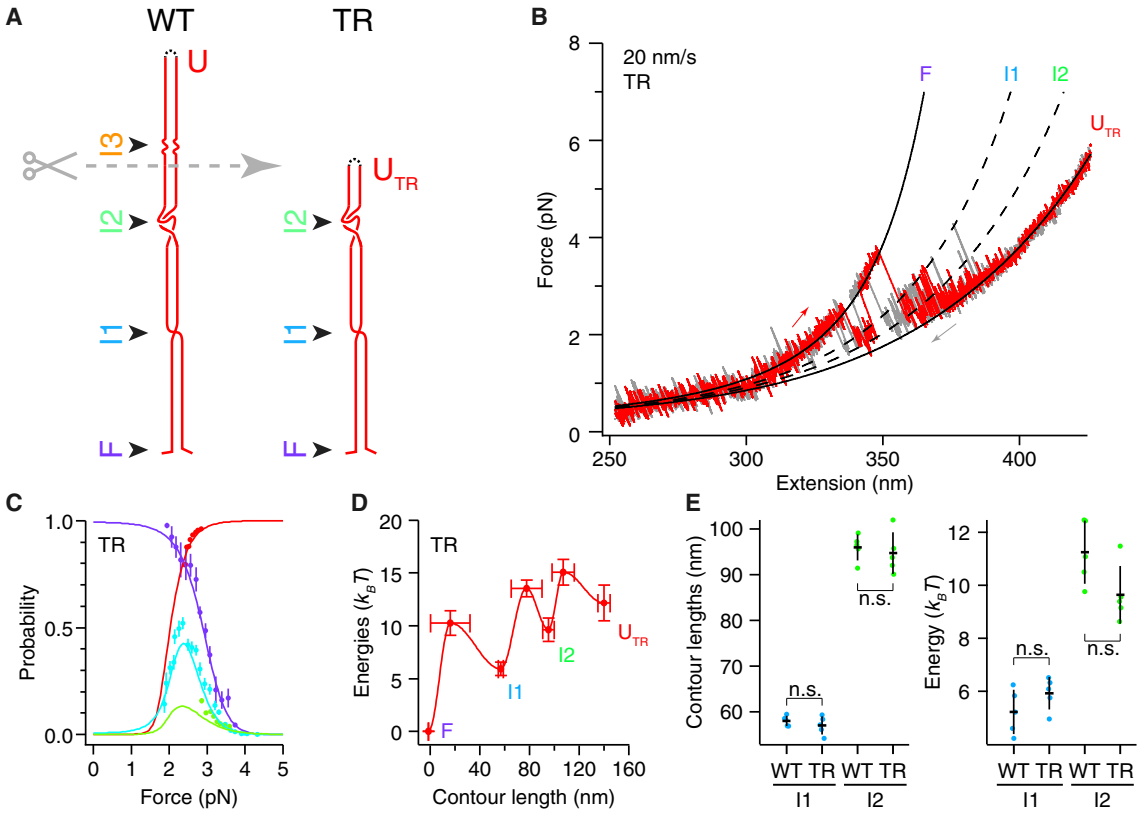


FIGURE 3 Unzipping of a truncated Psm3 CC. (*A*) A truncation of the WT construct removes I3. (*B*) Force-extension curve of the truncated construct. Pull, red curve; relaxation, gray curve. Two folding intermediates (I1 and I2) and the unfolded protein ($U_{TR}$) were identified. (*C*) Global energy fit of the force-dependent state probabilities; error bars indicate standard deviations and result from bootstrapping of the corresponding passive mode trace (see section "materials and methods"). Color code as in (*B*). (*D*) Energy landscape constructed from the contour lengths and energies of the folding states. Barrier heights were estimated from force-dependent transition kinetics (Fig. S3 *B*), n = 5 molecules, error bars indicate standard deviations. (*E*) Comparison of the measured contour length and energies of the intermediates I1 (blue) and I2 (green) found for the WT and the TR construct, n = 5, error bars indicate standard deviations; n.s., not significant based on a Student's *t*-test. To see this figure in color, go online.

integral component of the Psm3 CC region and ensures proper folding of the CC sections.

To validate if the elbow replacement had a stabilizing effect on the CC, we were interested in the energy landscape of the WT-like ΔEB configuration. Unfortunately, passive-mode measurements of the ΔEB construct were ambiguous to analyze since transitions between native-like and misfolded configurations, the variety of misfolded states, and their longevity masked the folding network, which was also reported for a misfolding hairpin (57). This prevented us from directly obtaining an energy landscape of ΔEB. Instead, we used an approach based on Monte Carlo simulations. Simulated force-extension curves matched experimental data for WT and TR (Fig. S7 *A–F*). We then generated a likely ΔEB landscape by altering the WT landscape around the region of the elbow replacement to match simulated and experimental force-extension curves (Fig. S7 *G–I*). This analysis showed that the ΔEB elbow replacement resulted in a removal of I2 and a destabilization of the I1–I3 section by ∼7 $k_BT$.

## Replacing the elbow in vivo causes cell growth defects

To test the influence of the elbow feature on the cohesin complex in a biological context, we introduced similar mutations of the elbow regions into the budding yeast cohesin subunit Smc3. As cohesin plays a major role in cell cycle progression and chromatin organization and because mitosis defects often manifest in increased cell volume (58), we chose cell volume measurements as a sensitive read-out for biological cohesin function. An elbow replacement corresponding to ΔEB showed significantly increased cell volume compared with the WT strain or a control, where the WT *SMC3* sequence was re-integrated in the same locus using an analogous construct to exclude unspecific effects (Fig. 5 *A*). We next sought to pinpoint the phenotype of the replacement to the individual CCs and repeated the experiment with only N- or C-terminal mutations ($\Delta EB_N$ and $\Delta EB_C$, respectively). To our surprise, we found that $\Delta EB_N$ had an even stronger phenotype than the double mutation, whereas $\Delta EB_C$ was not significantly different from

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



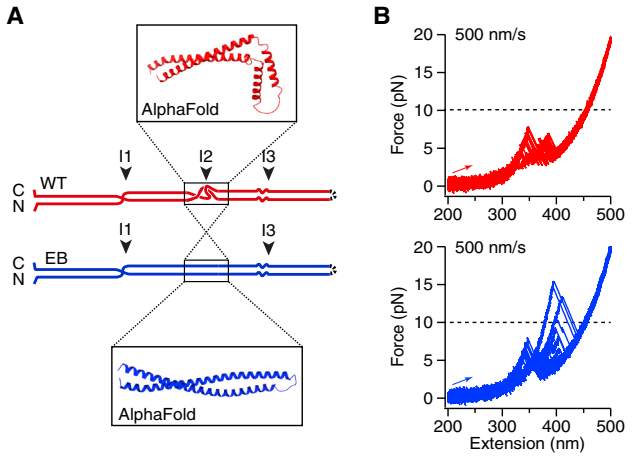


FIGURE 4 Unzipping of the elbow replacement mutant ΔEB. (*A*) The ΔEB mutant (blue) is constructed from WT (red) by replacing the elbow at I2 with a continuous heptad repeat pattern that encodes a canonical CC. The AlphaFold prediction created with ColabFold indicates a straightening of the ΔEB CC and the removal of the elbow. (*B*) Comparison of 20 overlaid pulling traces from the WT construct (red) and the ΔEB construct (blue) at a pulling speed of 500 nm/s. The WT construct follows the described folding pathway with moderate unfolding forces. The ΔEB mutant shows altered, halfway-folded conformations, which resist unexpected high forces >10 pN (dashed lines to guide the eye, see also Figs. S6 and S7 for a detailed comparison between WT and ΔEB). To see this figure in color, go online.

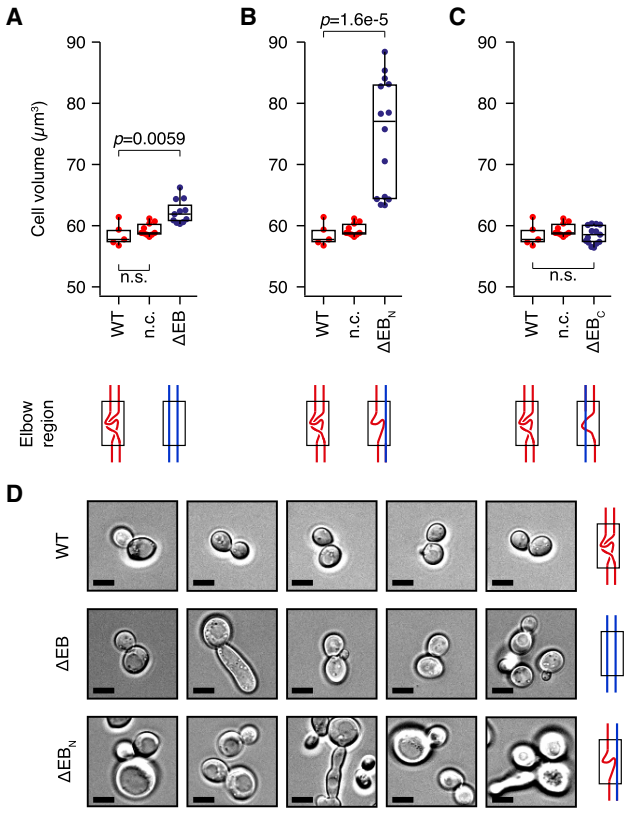


FIGURE 5 In vivo investigation of budding yeast elbow replacement mutations (ΔEB) for Smc3 via cell volume analysis. The *SMC1* gene was unmodified in all strains. (*A*) Average yeast cell volumes for WT, negative control, in which the targeted *SMC3* gene and the following untranslated region is replaced by the WT sequence (n.c., two independent clones), and the double mutation containing N- and C-terminal ΔEB mutations in Smc3 (two independent clones). The double mutation causes a significantly increased cell volume. Lower part: corresponding cartoons of the mutated (blue) or unmodified elbow region (red). (*B*) A more severe phenotype was found for the single N-terminal elbow replacement ($\Delta EB_N$, three independent clones). Lower part: cartoons as in (A). (*C*) The single C-terminal elbow replacement mutation ($\Delta EB_C$, four independent clones) does not influence the cell volume. Lower part: cartoons as in (A). *p*-values based on Student's two-sample *t*-test. n.s., not significant. (*D*) Exemplary microscopy images of yeast cells carrying WT, ΔEB, and $\Delta EB_N$ mutations in the *SMC3* gene. Compared with WT, the mutants ΔEB and $\Delta EB_N$ show cell size and shape defects. Scale bar in all images, 5 μm (see also Fig. S8). To see this figure in color, go online.

WT (Fig. 5 *A* and *C*). To conclude, while cohesin tolerates a small mutation that continues the heptad in the Smc3 C-terminal CC, the heptad continuation of the Smc3 N-terminal CC results in severe size and shape defects (Figs. 5 *D* and S8). Interestingly, a simultaneous continuation of both CCs partially rescues the phenotype, but still shows a cell volume increase consistent with a mitosis defect.

## DISCUSSION

Here, we have characterized a long region of Psm3's CC as a model protein for eukaryotic SMC proteins by optical tweezers with supporting in vitro and in vivo experiments. We identified structural features from their mechanical response, which uncovered the arrangement of the CC.

CCs are a ubiquitous and frequently occurring folding motif of proteins, where the amino acids follow a heptad repeat pattern (14). Besides facilitating oligomerization, CCs are known to serve as molecular spacers that separate functional domains of proteins. While the length of SMC CCs is conserved in eukaryotes (17), which is consistent with a role as spatial separators, the relatively high sequence conservation (14) compared with spacer CCs and the transitions between different large-scale conformations (O and B shape; Fig. S9) contradict a function as rigid spacers and suggest a more subtle role within SMC mechanics.

SMC CCs were found to contain kinks in rotary-shadow experiments (4,59), and recent crystal structures of SMC CCs exhibit interrupted helices (5,12,60), which suggest the presence of different segments in the CC. Numerous short breaks were already found in other elongated CCs, such as MRN (61), Myosin II (62), or Golgins (63). However, a segmented architecture does not necessarily result in high flexibility, as MRN (64) or Myosin II (65) show dramatically higher persistence length than SMC CCs (4). Our unfolding experiments supports the notion of a modular arrangement revealing that the Psm3 discontinuities are structurally and thermodynamically relevant features. In addition, the free energies per length of the individual segments (0.1–0.2 $k_BT$/nm) (Figs. 1 *E* and 3 *D*) are significantly lower than the reported values for canonical CC proteins, such as vimentin (0.54 $k_BT$/nm) (28) or GCN4 (1.20

$k_BT$/nm) (26,27). Therefore, the modular arrangement, but also the weak fold of the CC segments, promotes flexible polymer properties and facilitates the different conformations of SMC complexes.

SMC complexes contain an ATPase domain; hence, it stands to reason that CC bending at the elbow is an active process triggered by ATP hydrolysis and was already suggested for condensin (22). For cargo-carrying motor proteins such as dynein (52), conformational changes were found to be promoted by ATP-cycle-dependent helix sliding of the CC. These proteins contain, exactly like SMC proteins, kinked CCs attached to ATPase domains. In SMC CCs, different heptad pairings inside the hydrophobic core could lock the SMC complex in either the straight CC conformation or in the bent CC conformation. The transition between these conformations would then be promoted by helix sliding, whose energetic cost is estimated to be in the order of ~2–3 $k_BT$ (66), which could be afforded by ATP hydrolysis (67). However, our crosslinking experiments of the Psm3 CC showed no indications of a flexible heptad pairing (Fig. 2 *B*), which would be a prerequisite of helix sliding. In addition, passive mode traces did not show typical features of staggered helices, which were reported for GCN4 (26). In summary, our data support a model where elbow bending is driven by passive thermal fluctuations (13) of the CC sections identified in this study.

A passive CC bending mechanism critically depends on a structurally flexible elbow, but recent cryo-EM densities lack resolution at this crucial component (11,12). The simplest possible feature would be a kink, such as in GreA (68), or a loop, such as in Ndc80 (69) (Fig. S10 *A*), that is only present on one CC strand. Since AlphaFold and CC propensity profiles revealed a non-helical linker on both CC strands (Fig. S1 *A*), it is more likely that the elbow forms an entangled linker (e.g., as in Omp100 (70) or KKT4 (71); see Fig. S10 *B*) or a "knuckle motif," as in bacterial SMC or SMC-like proteins, such as MukB (5,17), RecN (72), or SbcCD (73) (Fig. S10 *C*), both of which are consistent with the found pairing shift of five amino acids (Fig. S1 *C*). Replacing the elbow by creating a continuous canonical heptad repeat pattern destabilized the CC (Fig. S7). Intuitively, one would expect that a continuation of the regular heptad repeat stabilizes the CC. However, a reason for destabilization might be that the flexible linkers in the elbow are required for CC strand movement and proper alignment, while a regular heptad continuation locks them in an unfavorable configuration. Without the separating effect of the elbow linkers, these misaligned configurations then propagate over the elbow region and result in non-native alignments of the amino acids close to I1, which might result in the metastable intermediates found in the ΔEB unzipping traces (Fig. S4 *C*). In contrast to continuous short CCs such as GCN4 that need to completely unfold to remove misfolds before starting a new folding attempt (26), unfolding to the previous intermediate is sufficient for Psm3 CC (Fig. S6 *E*). Taken together, the elbow guides proper folding of the entire Psm3 CC and prevents misfolding. As other long CCs, such as Golgins (63) or MRN (61), also contain interrupted helices, we propose that the modular composition is a general feature of elongated CCs to reliably enable correct folding and functional proteins.

In in vivo studies, alterations of the CC length (20) and even point mutations inside the elbow (5) were found to lead to lethality. We observed severe size and shape phenotypes of the N-terminal replacement mutant ($\Delta EB_N$), which were tempered by the addition of the C-terminal deletion (Fig. 5). This indicates that the altered cohesin complexes are still partially functional, but their ability to perform the stated large-scale conformational changes seemed to be significantly restricted by the alteration in elbow flexibility. Maybe the mutated CC arrangement is furthermore unfavorable for additional binding partners. Hence, the elbow is not only a folding guide but also likely ensures proper cohesin function.

## CONCLUSIONS

Using optical tweezers, we have found four relatively unstable sections inside the Psm3 CC. In contrast to many other CCs, and especially to homodimers, the non-canonical elbow feature comprises non-structured linkers, which induces a pairing shift from one CC strand to the other, enabling potentially unique CC mechanics. We found that the elbow is a key feature inside Psm3s CC, which ensures proper CC arrangement and further complex functions. As discontinuities are common CC features, these findings might translate to other elongated CCs and especially to other SMC or SMC-like proteins. Additionally, this study provides new details for SMC arm flexibility, which could be used to improve future modeling approaches of the SMC loop extrusion mechanism.

## SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.10.017.

## AUTHOR CONTRIBUTIONS

Conducting experiments, M.F., F.P., E.R., and S.J.; data analysis, M.F., F.P., E.R., K.M.S., and J.S.; concept of the study, J.S.; writing, M.F., S.Z., and J.S. with input from all authors.

## ACKNOWLEDGMENTS

J.S. and K.M.S. acknowledge support by the LMU Center for NanoScience CeNS. J.S. is supported by a DFG Emmy Noether grant STI673/2-1 and an ERC starting grant 758124.

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017



## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES


1. Haering, C. H., J. Löwe, …, K. Nasmyth. 2002. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell.* 9:773–788.
2. Davidson, I. F., and J.-M. Peters. 2021. Genome folding through loop extrusion by SMC complexes. *Nat. Rev. Mol. Cell Biol.* 22:445–464.
3. Yatskevich, S., J. Rhodes, and K. Nasmyth. 2019. Organization of chromosomal DNA by SMC complexes. *Annu. Rev. Genet.* 53:445–482.
4. Eeftens, J. M., A. J. Katan, …, C. Dekker. 2016. Condensin smc2-smc4 dimers are flexible and dynamic. *Cell Rep.* 14:1813–1818.
5. Bürmann, F., B.-G. Lee, …, J. Löwe. 2019. A folded conformation of MukBEF and cohesin. *Nat. Struct. Mol. Biol.* 26:227–236.
6. Ryu, J.-K., A. J. Katan, …, C. Dekker. 2020. The condensin holocomplex cycles dynamically between open and collapsed states. *Nat. Struct. Mol. Biol.* 27:1134–1141.
7. Davidson, I. F., D. Goetz, …, F. Uhlmann. 2016. Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J.* 35:2671–2685.
8. Stigler, J., G. Ö. Çamdere, …, E. C. Greene. 2016. Single-molecule imaging reveals a collapsed conformational state for DNA-bound cohesin. *Cell Rep.* 15:988–998.
9. Lee, H., H. Noh, and J.-K. Ryu. 2021. Structure-function relationships of SMC protein complexes for DNA loop extrusion. *BioDesign.* 9:1–13.
10. Bürmann, F., L. F. H. Funke, …, J. Löwe. 2021. Cryo-EM structure of MukBEF reveals DNA loop entrapment at chromosomal unloading sites. *Mol. Cell.* 81:4891–4906.e8.
11. Lee, B.-G., F. Merkel, …, C. H. Haering. 2020. Cryo-EM structures of holo condensin reveal a subunit flip-flop mechanism. *Nat. Struct. Mol. Biol.* 27:743–751.
12. Petela, N. J., A. Gonzalez Llamazares, …, K. A. Nasmyth. 2021. Folding of cohesin's coiled coil is important for Scc2/4-induced association with chromosomes. *Elife.* 10:e67268.
13. Bauer, B. W., I. F. Davidson, …, J.-M. Peters. 2021. Cohesin mediates DNA loop extrusion by a "swing and clamp" mechanism. *Cell.* 184:5448–5464.e22.
14. Truebestein, L., and T. A. Leonard. 2016. Coiled-coils: the long and short of it. *Bioessays.* 38:903–916.
15. Delorenzi, M., and T. Speed. 2002. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics.* 18:617–625.
16. Gruber, M., J. Söding, and A. N. Lupas. 2006. Comparative analysis of coiled-coil prediction methods. *J. Struct. Biol.* 155:140–145.
17. Waldman, V. M., T. H. Stanage, …, M. G. Oakley. 2015. Structural mapping of the coiled-coil domain of a bacterial condensin and comparative analyses across all domains of life suggest conserved features of SMC proteins. *Proteins.* 83:1027–1045.
18. Higashi, T. L., P. Eickhoff, …, F. Uhlmann. 2020. A structure-based mechanism for DNA entry into the cohesin ring. *Mol. Cell.* 79:917–933.e9.
19. Diebold-Durand, M.-L., H. Lee, …, S. Gruber. 2017. Structure of full-length SMC and rearrangements required for chromosome organization. *Mol. Cell.* 67:334–347.e5.
20. Bürmann, F., A. Basfeld, …, S. Gruber. 2017. Tuned SMC arms drive chromosomal loading of prokaryotic condensin. *Mol. Cell.* 65:861–872.e9.
21. Vazquez Nunez, R., Y. Polyhach, …, S. Gruber. 2021. Gradual opening of Smc arms in prokaryotic condensin. *Cell Rep.* 35:109051.
22. Takaki, R., A. Dey, …, D. Thirumalai. 2021. Theory and simulations of condensin mediated loop extrusion in DNA. *Nat. Commun.* 12:5865.
23. Nomidis, S. K., E. Carlon, …, J. F. Marko. 2022. DNA tension-modulated translocation and loop extrusion by SMC complexes revealed by molecular dynamics simulations. *Nucleic Acids Res.* 50:4974–4987, gkac268.
24. Bornschlögl, T., G. Woehlke, and M. Rief. 2009. Single molecule mechanics of the kinesin neck. *Proc. Natl. Acad. Sci. USA.* 106:6992–6997.
25. Gao, Y., S. Zorman, …, Y. Zhang. 2012. Single reconstituted neuronal SNARE complexes zipper in three distinct stages. *Science.* 337:1340–1343.
26. Xi, Z., Y. Gao, …, Y. Zhang. 2012. Single-molecule observation of helix staggering, sliding, and coiled coil misfolding. *Proc. Natl. Acad. Sci. USA.* 109:5711–5716.
27. Gebhardt, J. C. M., T. Bornschlögl, and M. Rief. 2010. Full distance-resolved folding energy landscape of one single protein molecule. *Proc. Natl. Acad. Sci. USA.* 107:2013–2018.
28. Ramm, B., J. Stigler, …, M. Rief. 2014. Sequence-resolved free energy profiles of stress-bearing vimentin intermediate filaments. *Proc. Natl. Acad. Sci. USA.* 111:11359–11364.
29. Mukhortava, A., and M. Schlierf. 2016. Efficient formation of site-specific protein–DNA hybrids using copper-free click chemistry. *Bioconjugate Chem.* 27:1559–1563.
30. Freitag, M., D. Kamp, …, J. Stigler. 2021. Identification and correction of miscalibration artifacts based on force noise for optical tweezers experiments. *J. Chem. Phys.* 155:175101.
31. Synakewicz, M., D. Bauer, …, L. S. Itzhaki. 2019. Bioorthogonal protein-DNA conjugation methods for force spectroscopy. *Sci. Rep.* 9:13820.
32. Tych, K., and G. Žoldák. 2019. Stable substructures in proteins and how to find them using single-molecule force spectroscopy. *In* Protein Supersecondary Structures: Methods and Protocols. A. E. Kister, ed Springer, pp. 263–282.
33. Synakewicz, M., R. S. Eapen, …, J. Stigler. 2022. Unraveling the mechanics of a repeat-protein nanospring: from folding of individual repeats to fluctuations of the superhelix. *ACS Nano.* 16:3895–3905.
34. Soh, Y.-M., J. Basquin, and S. Gruber. 2021. A rod conformation of the Pyrococcus furiosus Rad50 coiled coil. *Proteins.* 89:251–255.
35. Ludwiczak, J., A. Winski, …, S. Dunin-Horkawicz. 2019. DeepCoil—a fast and accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics.* 35:2790–2795.
36. Jumper, J., R. Evans, …, D. Hassabis. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature.* 596:583–589.
37. Varadi, M., S. Anyango, …, S. Velankar. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50:D439–D444.
38. Mirdita, M., K. Schütze, …, M. Steinegger. 2022. ColabFold: making protein folding accessible to all. *Nat. Methods.* 19:679–682.
39. Mirdita, M., M. Steinegger, and J. Söding. 2019. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics.* 35:2856–2858.
40. Mirdita, M., L. von den Driesch, …, M. Steinegger. 2017. Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* 45:D170–D176.
41. Mitchell, A. L., A. Almeida, …, R. D. Finn. 2020. MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res.* 48:D570–D578.
42. Pettersen, E. F., T. D. Goddard, …, T. E. Ferrin. 2021. UCSF ChimeraX : structure visualization for researchers, educators, and developers. *Protein Sci.* 30:70–82.
43. Dietler, N., M. Minder, …, S. J. Rahi. 2020. A convolutional neural network segments yeast microscopy images with high accuracy. *Nat. Commun.* 11:5723.
44. Padovani, F., B. Mairhörmann, …, K. M. Schmoller. 2022. Segmentation, tracking and cell cycle analysis of live-cell imaging data with Cell-ACDC. *BMC Biol.* 20:174.

Please cite this article in press as: Freitag et al., Single-molecule experiments reveal the elbow as an essential folding guide in SMC coiled-coil arms, Biophysical Journal (2022), https://doi.org/10.1016/j.bpj.2022.10.017




45. Crooks, G. E. 1999. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E.* 60:2721–2726.

46. Collin, D., F. Ritort, …, C. Bustamante. 2005. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature.* 437:231–234.

47. Bauer, D., S. Meinhold, …, G. Žoldák. 2018. A folding nucleus and minimal ATP binding domain of Hsp70 identified by single-molecule force spectroscopy. *Proc. Natl. Acad. Sci. USA.* 115:4666–4671.

48. Stigler, J., and M. Rief. 2012. Hidden Markov analysis of trajectories in single-molecule experiments and the effects of missed events. *ChemPhysChem.* 13:1079–1086.

49. Žoldák, G., J. Stigler, …, M. Rief. 2013. Ultrafast folding kinetics and cooperativity of villin headpiece in single-molecule force spectroscopy. *Proc. Natl. Acad. Sci. USA.* 110:18156–18161.

50. Puchner, E. M., G. Franzen, …, H. E. Gaub. 2008. Comparing proteins by their unfolding pattern. *Biophys. J.* 95:426–434.

51. Stigler, J., F. Ziegler, …, M. Rief. 2011. The complex folding network of single calmodulin molecules. *Science.* 334:512–516.

52. Huis in 't Veld, P. J., F. Herzog, …, J.-M. Peters. 2014. Characterization of a DNA exit gate in the human cohesin ring. *Science.* 346:968–972.

53. Wang, M. D., H. Yin, …, S. M. Block. 1997. Stretching DNA with optical tweezers. *Biophys. J.* 72:1335–1346.

54. Hons, M. T., P. J. Huis In 't Veld, …, F. Herzog. 2016. Topology and structure of an engineered human cohesin complex bound to Pds5B. *Nat. Commun.* 7:12523.

55. Gabler, F., S.-Z. Nam, …, V. Alva. 2020. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr. Protoc. Bioinformatics.* 72:e108.

56. Zimmermann, L., A. Stephens, …, V. Alva. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430:2237–2243.

57. Alemany, A., A. Mossa, …, F. Ritort. 2012. Experimental free-energy measurements of kinetic molecular states using fluctuation theorems. *Nat. Phys.* 8:688–694.

58. Schmoller, K. M. 2017. The phenomenology of cell size control. *Curr. Opin. Cell Biol.* 49:53–58.

59. Anderson, D. E., A. Losada, …, T. Hirano. 2002. Condensin and cohesin display different arm conformations with characteristic hinge angles. *J. Cell Biol.* 156:419–424.

60. Soh, Y.-M., F. Bürmann, …, S. Gruber. 2015. Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol. Cell.* 57:290–303.

61. Park, Y. B., M. Hohl, …, Y. Cho. 2017. Eukaryotic Rad50 functions as a rod-shaped dimer. *Nat. Struct. Mol. Biol.* 24:248–257.

62. Heissler, S. M., A. S. Arora, …, K. Chinthalapudi. 2021. Cryo-EM structure of the autoinhibited state of myosin-2. *Sci. Adv.* 7:eabk3273.

63. Ishida, R., A. Yamamoto, …, N. Nakamura. 2015. GM130 is a parallel tetramer with a flexible rod-like structure and N–terminally open (Y-shaped) and closed (I-shaped) conformations. *FEBS J.* 282:2232–2244.

64. de Jager, M., J. van Noort, …, C. Wyman. 2001. Human rad50/mre11 is a flexible complex that can tether DNA ends. *Mol. Cell.* 8:1129–1135.

65. Schwaiger, I., C. Sattler, …, M. Rief. 2002. The myosin coiled-coil is a truly elastic protein structure. *Nat. Mater.* 1:232–235.

66. Gomez, D., Y. Gavrilov, and Y. Levy. 2019. Sliding mechanism at a coiled-coil interface. *Biophys. J.* 116:1228–1238.

67. Rosing, J., and E. C. Slater. 1972. The value of $\Delta G^{\circ}$ for the hydrolysis of ATP. *Biochim. Biophys. Acta.* 267:275–290.

68. Stebbins, C. E., S. Borukhov, …, S. A. Darst. 1995. Crystal structure of the GreA transcript cleavage factor from Escherichia coli. *Nature.* 373:636–640.

69. Ciferri, C., S. Pasqualato, …, A. Musacchio. 2008. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. *Cell.* 133:427–439.

70. Hartmann, M. D., C. T. Mendler, …, B. Hernandez Alvarez. 2016. $\alpha/\beta$ coiled coils. *Elife.* 5:e11861.

71. Ludzia, P., E. D. Lowe, …, B. Akiyoshi. 2021. Structural characterization of KKT4, an unconventional microtubule-binding kinetochore protein. *Structure.* 29:1014–1028.e8.

72. Pellegrino, S., J. Radzimanowski, …, J. Timmins. 2012. Structural and functional characterization of an SMC-like protein RecN: new insights into double-strand break repair. *Structure.* 20:2076–2089.

73. Käshammer, L., J.-H. Saathoff, …, K.-P. Hopfner. 2019. Mechanism of DNA end sensing and processing by the mre11-rad50 complex. *Mol. Cell.* 76:382–394.e6.