COMMENTARY

Insights into the energy landscapes of chromosome organization proteins from coevolutionary sequence variation and structural modeling

Ronald M. Levy[a,b,1]

Uncovering mechanisms of protein function is challenging when structural characterization of the functionally relevant states is elusive, for instance for large and flexible proteins which resist crystallization. This is the case for structural maintenance of chromosomes (SMC) proteins and kleisin subunits which are crucial for the segregation of chromosomes during cell division but whose mechanism to organize chromosomes for replication is not known. The two most prominent multisubunit SMC complexes are cohesin and condensin, which form large complexes (more than 650 kDa for condensin) containing long antiparallel coiled coils, making them difficult to study by conventional crystallography or cryoelectron microscopy. Integrative structural biology techniques have arisen in response to these challenges which attempt to solve the puzzle by combining incomplete information from diverse experimental data with computational and theoretical results (1). In PNAS, Krepel et al. (2) use an integrative approach to understanding the mechanism of SMC action which incorporates evolutionary sequence analysis with molecular simulations based on an energy landscape optimized effective potential (AWSEM, associative memory, water-mediated, structure and energy model) along with crystallographic structural information and cross-linking experimental data to map features of the conformational and energy landscapes of SMC proteins. A complete atom-level structural model of several kleisin complexes is constructed consistent with the structural and evolutionary data, with particular focus on the structure and energetics of the long antiparallel coiled coils. The work by Krepel et al. (2) suggests a dynamical model for SMC proteins. The integrative analysis enables the authors to suggest that SMC complexes may act as "topological motors" where the SMC topology couples to the topology of DNA when entwined. The results suggest a mechanism by which movements of the ATPase motors of the head domains can lead to propagation of the chromosomal DNA toward the hinge region when the DNA is entwined with the braided coil regions of the SMC–kleisin complexes. These results are a demonstration of what can be achieved by models which integrate evolutionary, structural, and physiochemical analysis.

Structural Modeling Which Combines DCA with AWSEM

Direct coupling analysis (DCA) is a method to quantify the strength of the direct interaction between two positions on a biological sequence. It has been used to infer structural information about intra- and intermolecular protein–protein contacts, as well as information related to protein fitness, and uses maximum entropy inference to produce a "Potts" Hamiltonian model of the observed sequence variation inspired by statistical physics (3–5). Using DCA together with crystallographic data (6, 7), Onuchic and coworkers (8) were previously able to construct an atomic-scale model of the whole condensin complex, which provides a starting point for the energy landscape analysis of chromosome organization proteins ref. 2 reports. Molecular dynamics (MD) simulations of the bacterial and eukaryotic cohesion and condensin were performed using the coarse-grained AWSEM (9). It should be noted that the currently available structural information is insufficient to unambiguously determine the distribution of braiding topologies of chromosomal organization proteins, or other features of the heterogeneity of the coiled-coil regions which are likely to be important for function. The MD simulations with AWSEM were used in two key ways to supplement the experimentally determined structural and coevolutionary information. First, it was possible to show that the proportion of minimally frustrated contacts and highly frustrated contacts in the starting models were consistent with those generally observed in protein crystal structures. Second, the coarse-grained MD simulations

[a]Center for Biophysics and Computational Biology, Temple University, Philadelphia, PA 19122; and [b]Department of Chemistry, Temple University, Philadelphia, PA 19122

Author contributions: R.M.L. wrote the paper.

The author declares no competing interest.

Published under the PNAS license.

See companion article 10.1073/pnas.1917750117.

[1]Email: ronlevy@temple.edu.

Downloaded at Temple University on January 14, 2020

with AWSEM were able to recapitulate contacts observed in cross-linking experiments (10) and those predicted from DCA, but importantly the structures become better defined, leading to tight braiding of the coils which cannot be inferred from the crystallographic data alone.

SMC Complexes Can Form Different Braiding Topologies

The SMC complexes consist of globular domains (head and hinge domains) which are connected by coiled-coil regions. The SMC dimer forms a circular unit. While the SMC dimer by itself is not covalently closed, once the head and the hinge domains bind together the ring does become closed. DCA indicated contacts between the coiled-coil regions of the two SMC monomers. The most strongly coevolving contacts appear near the hinge and at the bottom portion of the coiled coil, which is kinked. Seven isomers of the eukaryotic cohesion protein differing in braiding topologies characterized by twist angle T_w and writhe angle W_r were used to initiate AWSEM simulations. All of the topologically distinct isomers were found to form coiled-coil structures. A major finding of this work is that a distribution of metastable ring conformations involving different braiding topologies is necessary to rationalize experimental studies involving the observed cross-linking of pairs of residues (10–13). Barysz et al. (10), using a combination of amino acid selective cross-linking and mass spectroscopy, identified a total of 15 SMC1–SMC3 cross-linking sites as being of high confidence. The ensemble of predicted structures observed in the AWSEM simulations captured all 15 of the high-confidence cross-linking sites and suggests that many more Lys–Lys pairs can be identified from more extensive cross-linking experiments. Confirmation of these predicted cross-linked pairs would further validate the heterogeneous nature of the supercoiled regions of SMC–kleisin complexes.

While it is not clear precisely how the topological diversity of the supercoiled regions of SMC proteins is used in a functional sense, it may be that topological isomerism allows the motor domains to couple the supercoiling of the SMC proteins to that of the DNA if the braided coils become entwined with DNA. The authors propose a possible mechanism by which rotation of the head domain associated with its ATPase function can be transmitted through the braided region to the hinge region in a way that allows for propagation of DNA strands along the SMC ring complex. A combined approach probably involving single-molecule experiments and molecular simulations of the kind described in PNAS (2) appears to be a promising path forward to further develop the mechanistic ideas concerning the role of SMC proteins in chromosome organization the authors propose.

1 A. B. Ward, A. Sali, I. A. Wilson, Biochemistry. Integrative structural biology. *Science* **339**, 913–915 (2013).
2 D. Krepel, A. Davtyan, N. P. Schafer, P. G. Wolynes, J. N. Onuchic, Braiding topology and the energy landscape of chromosome organization proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 10.1073/pnas.1917750117 (2019).
3 F. Morcos *et al.*, Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1293–E1301 (2011).
4 S. Ovchinnikov, H. Kamisetty, D. Baker, Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* **3**, e02030 (2014).
5 R. M. Levy, A. Haldane, W. F. Flynn, Potts Hamiltonian models of protein co-variation, free energy landscapes, and evolutionary fitness. *Curr. Opin. Struct. Biol.* **43**, 55–62 (2017).
6 C. H. Haering, J. Löwe, A. Hochwagen, K. Nasmyth, Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell* **9**, 773–788 (2002).
7 F. Bürmann *et al.*, An asymmetric SMC-kleisin bridge in prokaryotic condensin. *Nat. Struct. Mol. Biol.* **20**, 371–379 (2013).
8 D. Krepel, R. R. Cheng, M. Di Pierro, J. N. Onuchic, Deciphering the structure of the condensin protein complex. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11911–11916 (2018).
9 M. Y. Tsai *et al.*, Electrostatics, structure prediction, and the energy landscapes for protein folding and binding. *Protein Sci.* **25**, 255–269 (2016).
10 H. Barysz *et al.*, Three-dimensional topology of the SMC2/SMC4 subcomplex from chicken condensin I revealed by cross-linking and molecular modelling. *Open Biol.* **5**, 150005 (2015).
11 K. Nasmyth, How are DNAs woven into chromosomes? *Science* **358**, 589–590 (2017).
12 M. L. Diebold-Durand *et al.*, Structure of full-length SMC and rearrangements required for chromosome organization. *Mol. Cell* **67**, 334–347.e5 (2017).
13 M. T. Hons *et al.*, Topology and structure of an engineered human cohesin complex bound to Pds5B. *Nat. Commun.* **7**, 12523 (2016).

Downloaded at Temple University on January 14, 2020