the cryo-EM structure, and future mutational studies of these residues may determine if they affect IstA/IstB assembly, DNA binding, and transposition activities.

Collectively, the results presented here allow a new mechanism for transposition to be proposed (Figure 1). Previous results showed that transposases generally oligomerize to bring transposon ends together. These new findings show that IstA specifically interacts with IstB in its ATP-bound large oligomeric form, not the ADP-bound dimer, to stimulate its ATPase activity. ATPase stimulation accelerates ATP-turnover by IstB, triggering the disassembly of IstB decamers to dimers. The IstB dimers likely dissociate from DNA following ADP release, although this needs to be tested. Following IstB dimer dissociation from target DNA, IstA-DNA can facilitate the strand transfer process with target DNA and complete the transposition reaction (Figure 1B, bottom).

That IstB forms a decameric clamshell architecture extends the known assembly states for both proteins involved in transposition and the AAA+ ATPase superfamily members. The unique and stable complex of IstB with target DNA, which is recognized by IstA, generates an enabling system to capture the IstA transposase and IstB-DNA complexes together. Thus, this system promises new structures of IstA/IstB macromolecular assemblies to reveal how the transposition process is completed.

ACKNOWLEDGMENTS

The authors' efforts are supported in part by NIH National Cancer Institute grant P01 CA092584 (to J.A.T.), and National Institute Of Environmental Health Sciences grant R21ES025895 (to G.J.W.).

REFERENCES

Arias-Palomo, E., and Berger, J.M. (2015). Cell *162*, this issue, 860–871.

Berger, B., and Haas, D. (2001). Cell. Mol. Life Sci. *58*, 403–419.

Curcio, M.J., and Derbyshire, K.M. (2003). Nat. Rev. Mol. Cell Biol. *4*, 865–877.

Erzberger, J.P., and Berger, J.M. (2006). Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

Jang, S., Sandler, S.J., and Harshey, R.M. (2012). PLoS Genet. *8*, e1002642.

Mills, R.E., Bennett, E.A., Iskow, R.C., and Devine, S.E. (2007). Trends Genet. *23*, 183–191.

Mizuno, N., Dramićanin, M., Mizuuchi, M., Adam, J., Wang, Y., Han, Y.-W., Yang, W., Steven, A.C., Mizuuchi, K., and Ramón-Maiques, S. (2013). Proc. Natl. Acad. Sci. USA *110*, E2441–E2450.

Montaño, S.P., Pigli, Y.Z., and Rice, P.A. (2012). Nature *491*, 413–417.

Mott, M.L., Erzberger, J.P., Coons, M.M., and Berger, J.M. (2008). Cell *135*, 623–634.

Putnam, C.D., Clancy, S.B., Tsuruta, H., Gonzalez, S., Wetmur, J.G., and Tainer, J.A. (2001). J. Mol. Biol. *311*, 297–310.

A CTCF Code for 3D Genome Architecture

Michael H. Nichols[1] **and Victor G. Corces**[1,*]
[1]Department of Biology, Emory University, 1510 Clifton Road NE, Atlanta, GA 30322, USA
*Correspondence: vcorces@emory.edu
http://dx.doi.org/10.1016/j.cell.2015.07.053

The architectural protein CTCF plays a complex role in decoding the functional output of the genome. Guo et al. now show that the orientation of a CTCF site restricts its choice of interacting partner, thus creating a code that predicts the three-dimensional organization of the genome. We propose a DNA extrusion model to account for orientation-specific loop formation.

CTCF is a DNA-binding protein known to play a variety of roles in the regulation of transcription by forming loops in which distant elements of the genome are brought into spatial proximity within the nucleus (Ong and Corces, 2014). The formation of these loops is believed to involve homodimerization of the CTCF proteins bound to their bases. By mediating contacts between distant sequences, CTCF regulates enhancer-promoter interactions throughout the genome and appears to play a key role in the formation of topologically associating domains (TADs) (Nora et al., 2012). Analysis of genome-wide interaction data obtained by Hi-C suggests that CTCF-mediated contacts occur much more frequently when the binding sites for this protein are present in the convergent forward and reverse orientations (Rao et al., 2014). Interactions between binding sites arranged in the same forward-forward or reverse-reverse orientation still occur, although less frequently, and interactions between CTCF sites in a divergent reverse-forward orientation rarely take place. In this issue of *Cell*, Guo et al. (2015) carry out a detailed functional analysis of the role of CTCF binding site orientation in the regulation of enhancer-promoter choice underlying stochastic expression of specific protocadherin isoforms.

The protocadherin genes are subject to alternative splicing, and each variable exon contains an upstream promoter, transcription from which depends on interaction with a downstream enhancer via DNA looping. Each variable exon and enhancer has a CTCF binding site. Guo et al. (2015) noticed that the CTCF binding

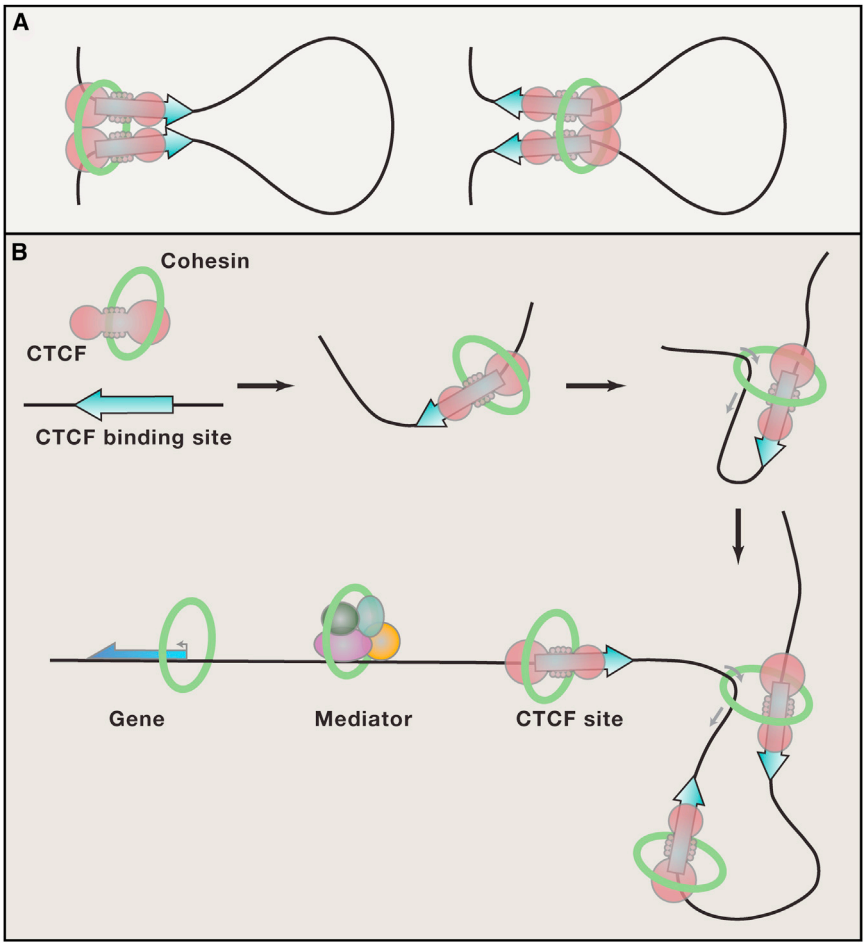

Figure 1. Model of Orientation-Biased CTCF Looping

(A) CTCF-mediated loops in convergent and divergent orientations only differ in how they are connected by the DNA. The loop on the left occurs much more frequently than the loop on the right, suggesting that the mechanism of loop formation must be able to distinguish the two cases.

(B) A loop-extrusion model would explain the orientation bias seen in CTCF-mediated looping. CTCF bends DNA and could be capable of forming a loop on one side of its binding site only, due to the manner in which the DNA is bent. This loop could then be expanded in one direction via the action of cohesin and possibly also transcription, causing the CTCF site to contact other DNA elements such as other CTCF sites, cohesin-associated Mediator complexes, and cohesin-associated gene promoters more frequently in one orientation. Homodimerization of CTCF complexes in anti-parallel orientations may not be favored, leading to continued rather than completed loop formation when two CTCF binding sites encounter each other during loop extrusions, accounting for the paucity of these interactions observed in genome interaction data.

sites that form loops between promoters and enhancers are arranged in a convergent orientation. Using the CRISPR-Cas9 genome editing system, they create inversions of key CTCF binding sites, switching their orientation. The authors then use 4C to show that the inverted CTCF binding sites now have an inverted interaction bias. This confirms the causal relationship between DNA binding site orientation and the direction of looping. Furthermore, the change in looping directionality is accompanied by changes in transcription, indicating a functional role for the CTCF-mediated interactions in regulating gene expression.

The authors then expand their investigation to the entire genome using published CTCF ChIA-PET data. They find the same orientation bias in interactions between CTCF sites as previously shown with Hi-C data. These observations solidify what now appears to be one of the underlying principles by which the orientation of the DNA sequence in CTCF binding sites shapes 3D genome organization. However, this new finding raises a series of questions as to the mechanisms underlying the specificity of interactions between CTCF sites in the genome. CTCF binding sites in divergent and convergent orientations are molecularly identical and impossible to distinguish outside of the larger context of the DNA molecule. Figure 1A shows two theoretical CTCF-mediated loops. The only difference between the two loops is which side of the CTCF sites the looped-out DNA is on. Despite this, the loop depicted on the left occurs much more frequently than the loop depicted on the right. This means that the mechanism by which CTCF forms loops must be aware of this context and be capable of discriminating between CTCF sites in convergent and divergent orientations. A simplistic model of loop formation that relies on random collisions in the nuclear space between CTCF bound to DNA in different orientations to form interactions is incompatible with the observations, as it could not be aware of the relative positions or orientations of the CTCF binding sites.

One potential explanation for the directionality in loop formation is that the bias is created by the binding of CTCF to its recognition site, which causes a 90 degree bending in the DNA, resulting in the formation of an unusual, oriented structure that could be interpreted as a loop (MacPherson and Sadowski, 2010). As this DNA structure is formed in the same orientation as the bias in looping, it seems likely that the two phenomena are causally linked. Several potential processes could then contribute to the expansion of the initial loop (Figure 1B). Since one end of the loop would be defined by CTCF binding, cohesin, which frequently co-binds with CTCF, might function to translocate DNA on the other side of the CTCF-induced "kink" to expand the loop. This is supported by results showing that cohesin is able to extrude a loop, perhaps using energy from its ATPase activity (Alipour and Marko, 2012; Strick et al., 2004). Transcriptional activity could also contribute to the cohesin-based translocation of the DNA into the loop (Lengronne et al., 2004). The observed frequency of interactions between CTCF sites with the same orientation is relatively low (Guo et al., 2015). Perhaps as two sites with the same orientation encounter

each other during loop extrusion, the antiparallel orientations of the CTCF proteins disfavor dimerization, and loop extrusion would continue until a convergent site was met (Figure 1B). In addition, the directionality imposed by this DNA bending-initiated loop extrusion model results in a CTCF site interacting more frequently with the DNA on one side of it, explaining why divergent CTCF sites interact very infrequently (Guo et al., 2015; Rao et al., 2014). This would also explain the finding that TAD boundaries, i.e., the generally "non-looped" stretch of DNA between two TADs, are enriched in CTCF sites arranged in divergent orientations (Vietri Rudan et al., 2015; Guo et al., 2015), since these border-associated divergent sites will tend to loop toward the interior of each adjacent TAD. This finding helps explain why only a subset of CTCF sites in the genome is able to form these boundaries and reinforces the functional relevance of CTCF to the formation of TADs. Finally, the loop extrusion model also imposes directionality on the interactions between CTCF and transcriptional complexes and/or gene promoters (Phillips-Cremins et al., 2013). The plethora of genome-probing tools that are constantly emerging should allow rigorous experimental testing of this model, stimulated by the results of Guo et al. (2015) and others in the near future.

REFERENCES

Alipour, E., and Marko, J.F. (2012). Nucleic Acids Res. *40*, 11202–11212.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). Cell *162*, this issue, 900–910.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Nature *430*, 573–578.

MacPherson, M.J., and Sadowski, P.D. (2010). BMC Mol. Biol. *11*, 101.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Nature *485*, 381–385.

Ong, C.T., and Corces, V.G. (2014). Nat. Rev. Genet. *15*, 234–246.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Cell *153*, 1281–1295.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). Cell *159*, 1665–1680.

Strick, T.R., Kawaguchi, T., and Hirano, T. (2004). Curr. Biol. *14*, 874–880.

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Cell Rep. *10*, 1297–1309.

Lighting Up Genes in Single Cells at Scale

Zhe Liu[1,*]

[1]Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA

*Correspondence: liuz11@janelia.hhmi.org

http://dx.doi.org/10.1016/j.cell.2015.07.052

In this issue, Shachar et al. report a high-throughput imaging position mapping platform (HIPmap) enabling large-scale, high-resolution localization of 3D gene positions in single cells. Coupling loss-of-function screens with HIPmap, the authors identify DNA replication rather than mitosis as a major determinant of genome positioning.

Deciphering how the genome is structurally organized and dynamically functions in the nucleus (4D nucleome) represents a remarkable challenge in the post-genome era (Bickmore, 2013; Dekker et al., 2013; Levine et al., 2014; Misteli, 2007). Two orthogonal approaches are commonly used to study genome folding—chromatin conformation capture (3C) and microscopy techniques such as fluorescence in situ hybridization (FISH) (Bickmore, 2013; Dekker et al., 2013). Combined with the recent developments of high-throughput sequencing, 3C-based methods provide insights into long-distance chromatin looping, genome folding, and topological domains in the context of whole genomes (Dekker et al., 2013). However, due to the nature of cell-population-based measurements, it is still challenging to interrelate 3C-derived genomic interactions with spatial distances inside the nucleus (Belmont, 2014; Williamson et al., 2014). Microscopy-based techniques directly measure physical distances but are usually applied to a few loci at a time and thus suffer from scalability and throughput limitations. Here, Shachar et al. (2015) report a high-throughput DNA FISH platform—high-throughput imaging position mapping platform (HIPmap)—with a fully automated liquid-handling FISH protocol, automated 3D confocal imaging, and a custom-designed analysis pipeline. Streamlined DNA FISH experiments can be performed in the 384-well format, allowing quantitative determination of the position of multiple endogenous loci in single cells with high accuracy and speed. In conjunction with large-scale perturbation screens, this platform should be suitable for single-cell analysis and systematic investigation of 3D spatial genome organization (Figure 1A).