Structure
Previews

Assembly of a Tightly Interwound DNA Recombination Complex Poised for Deletion

Reid C. Johnson[1,*] and John K. Heiss[1]
[1]Department of Biological Chemistry, David Geffen School of Medicine at UCLA, 615 Charles Young Drive South, Los Angeles, CA 90095-1737, USA
*Correspondence: rcjohnson@mednet.ucla.edu
DOI 10.1016/j.str.2008.04.005

In a recent issue of *Molecular Cell*, Mouw et al. (2008) report a crystal structure of a serine recombinase bound to a regulatory DNA site in an unexpected synaptic complex configuration, which forms the framework for a new model of the entire 12 subunit, 186 bp deletion complex.

Most site-specific DNA recombinases fall into two evolutionary and mechanistically distinct groups even though the basic reaction, recombination between two DNA sites that are specifically bound by a recombinase, is identical. Members of the tyrosine recombinase family, like Cre and FLP, are well known, in part because the fundamentals of their structure and enzymology are well understood and because they have been utilized extensively for genetic engineering. Serine recombinases, so named because a serine rather than a tyrosine is the active site residue for DNA cleavage, make up a second large but evolutionarily distinct group (Grindley et al., 2006; Smith and Thorpe, 2002). Serine recombinases, like tyrosine recombinases, function in many diverse reactions involving DNA (see Figure 1 for a partial list). Serine recombinases can also function in eukaryotic systems and are gaining increasing usage as genetic engineering tools (Keravala et al., 2006).

A hallmark of many site-specific recombination systems, particularly those of the serine recombinase family, is that the reactions are exquisitely regulated. This, of course, makes sense given that a chromosome is transiently broken during the reaction; serine recombinases induce double strand breaks in each of the recombination sites prior to DNA exchange. Regulation not only involves the timing of recombination, usually through control of recombinase levels, but more interestingly, by the assembly of elaborate synaptic complexes prior to initiation of any DNA chemistry. In the case of resolvases, such assemblies insure that: (1) only specific DNA segments (called *res*) on the same DNA molecule can participate, and (2) only *res* elements oriented on a DNA molecule in a directly repeated configuration can productively synapse into a catalytically active complex. These constraints are mediated by the arrangement and functions of multiple protein binding sites within *res* and by the energetics of DNA supercoiling, which is required to assemble a productive synaptic complex (Grindley et al., 2006).

In a recent paper, the laboratories of Rice and Stark have combined to give us new insights into how the synaptic complex of the Sin resolvase reaction is organized (Mouw et al., 2008). Sin is a relative newcomer to the limited number of well-studied model site-specific recombination reactions. Sin is found on a large plasmid from *Staphylococcus aureus* where it probably functions to resolve plasmid multimers into monomers to enable faithful partitioning during cell division. Rowland et al. (2002) showed that the products of Sin recombination between *res* elements on a supercoiled plasmid substrate are two circular DNA molecules linked by exactly one DNA passage (Figure 2G). This result implies that three negative supercoils are trapped within the synaptic complex. This is the same product structure as generated from the well-studied resolvase reactions from transposons Tn3 and γδ, yet the nature of the recombining sites are remarkably different between the Sin and Tn3/γδ systems (Figure 1). Sin recombines 86 bp *res* elements that contain two binding sites for the dimeric recombinase. Site I is where DNA exchange occurs, and site II performs an essential regulatory function; no deletions occur between *res* elements lacking site II. Interestingly, the DNA sequences of the half-sites at site I exhibit inverted symmetry, whereas the half-sites at site II are oriented in a head-to-tail fashion (Figure 1). Unlike the Tn3/γδ systems, an additional DNA

Diversity among serine recombinases

Serine recombinase	Biological function	Substrate organization	Regulatory proteins	Mode of synapsis
Sin resolvase	Monomerize plasmids	site I, HU, site II	HU + Sin	3 (-) interwound supercoils
Tn3, γδ resolvase	Resolve transposition intermediates	site I, site II, site III	Resolvase	3 (-) interwound supercoils
Hin DNA invertase Gin DNA invertase	Control transcription by inversion of promoter Phage tail-host receptor diversity	Fis I, Fis II	Fis (HU)	2 (-) nodes at supercoil branch
ϕC31 integrase Bxb1 integrase	Phage integration & excision		RDFs (excision)	simple collision
TnpX, IS607 transposases	DNA transposition		none known	unknown

Figure 1. Examples of DNA Recombination Reactions Catalyzed by Serine Recombinases
Serine recombinases are related by a conserved ~120 amino acid residue catalytic core domain. Additional residues, which function in DNA binding and sometimes other roles, typically extend from the C-terminal end, although occasionally can be found at the N terminus. Full length resolvases and DNA invertases are ~200 residues, but some serine recombinases extend up to ~800 residues.

bending protein, which could be supplied by the HU protein from *Bacillus subtilis* or *Escherichia coli*, is absolutely required for Sin recombination. Although HU generally binds DNA nonspecifically, footprinting experiments showed that HU specifically binds to the 31 bp DNA segment between sites I and II in the presence of Sin (Rowland et al., 2002). By contrast, each 114 bp *res* element from the Tn3/γδ systems contains three binding sites for the resolvase, with the half-sites from each unit displaying inverted symmetry but with different numbers of base pairs separating the half-sites. The differences within resolvase binding units in the various systems imply an unusual degree of flexibility of DNA binding.

Mouw et al. (2008) have obtained a 3.2 Å cocrystal structure of Sin bound to a 29 bp site II. Two Sin-DNA complexes are present in the asymmetric unit revealing the structural nature of the complex that probably nucleates assembly of the entire recombination complex. The structure of the 125 residue Sin catalytic-dimerization domain is quite similar to that of γδ resolvase in solution or bound to site I (Yang and Steitz, 1995). The three helix bundle comprising the C-terminal DNA binding domain (DBD) is related to those of γδ resolvase and Hin DNA invertases, but their head-to-tail orientation and connections to the catalytic domains in the dimer are different. The Sin catalytic domains are positioned in a remarkably asymmetric manner over the DNA with the polypeptide chains linking the catalytic and DBDs adopting distinct structures to accommodate the locations of the DBDs in the two subunits.

Perhaps the biggest surprise of the Sin-site II cocrystal structure is that resolvase interactions responsible for mediating synapsis of the regulatory sites involve residues from the DBDs (Mouw et al., 2008). The head-to-tail oriented DBDs from each dimer associate to form an extensive interface that buries 1865 $Å^2$ of surface area between the two site II complexes, positioning the site II DNA segments close to each other in an orientation favored by (−) DNA supercoiling (Figure 2C). This arrangement is in contrast to the situation at the crossover site I where many lines of evidence from other resolvases and DNA invertases show the opposite configuration; the DNA duplexes are located on the outside, and synapsis is stabilized by

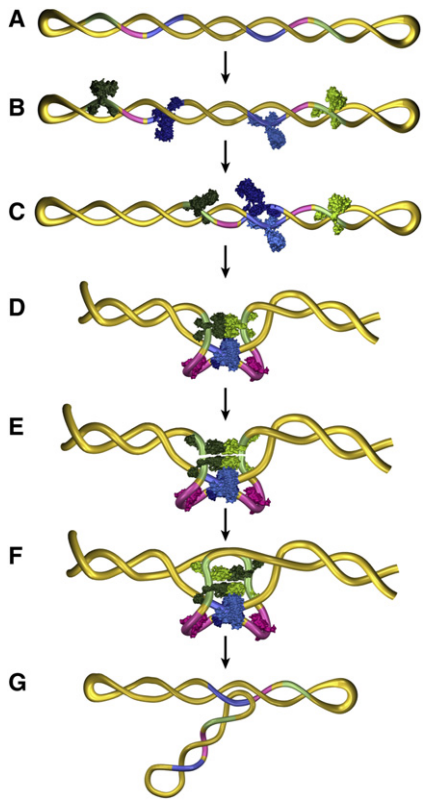

Figure 2. Assembly Pathway for Deletion Formation by Sin

(A) Supercoiled DNA with two *res* elements in direct repeat orientation. Each *res* element contains three protein binding sites: crossover site I (green), regulatory site II (blue), and an intervening segment (magenta) where the DNA bending protein HU binds.

(B) Sin resolvase dimers bind sites I and II (illustrated by PDB IDs 1GDT and 1R0Q, respectively) with similar affinities even though the orientation of half-sites are different.

(C) Residues within the DNA binding domains of Sin dimers bound at site II of each *res* interact to form a stable site II-site II complex (1R0Q), as revealed in the crystal structure of Mouw et al. (2008).

(D) Binding of HU (or IHF [1IHF] to modified *res* elements) bends the DNA to position the Sin dimers at sites I close to each other in a reaction enhanced by DNA supercoiling.

(E) Sin proteins at site I are remodeled into an "activated" tetramer (1ZR4), and all four DNA strands at the center of site I are then cleaved with each resolvase subunit covalently associated with a 5′ end through a serine-phosphodiester bond (Li et al., 2005). The activated synaptic complex model (Mouw et al., 2008) contains three trapped DNA supercoils, consistent with the earlier topological studies of Rowland et al. (2002).

(F) DNA exchange is mediated by a rotation of the top pair of subunits, together with their linked DNA strands, about a flat and largely hydrophobic interface (illustrated as a gap) that is present in the site I tetramer (Dhar et al., 2004; Li et al., 2005).

(G) Reversal of the phosphoserine linkages restores the DNA phosphodiester backbones resulting in singly linked deletion circles.

remodeling of dimeric catalytic domains into an interconnected tetramer (Grindley et al., 2006). Mouw et al. (2008) employed an elegant genetic screen to corroborate the functional importance of residues within the DBD for stabilizing site II synapsis. Mutations in many of the interacting residues were obtained that reduced or abolished site II synapsis and similarly impacted Sin-catalyzed recombination between full *res* elements. However, these mutations did not affect recombination efficiency by a hyperactive Sin mutant that can catalyze recombination between *res* elements containing only site I. These results confirm that synaptic interactions involving the DBDs are only required at the regulatory site II.

Two other potentially significant interfaces are revealed from analysis of the crystal packing. A small interface between symmetry-related Sin molecules is formed by stacking interactions involving the side chains of Phe52 and Arg54, and it was shown that a glutamic acid at residue 54 strongly inhibited site II-dependent recombination. Moreover, interactions over analogous residues in γδ resolvase occur in one crystal form and genetic studies have shown that such interactions between γδ resolvases bound at sites I and III (site III in γδ is probably functionally analogous to site II in Sin) are important (Grindley et al., 2006). Finally, an interface between adjacent catalytic domains in the lattice could reflect initial presynaptic interactions between dimers bound at site I (Figure 2D) prior to remodeling into the catalytically active tetramer.

Mouw et al. (2008) used their crystal structure as a starting point to construct a new model of the entire synaptic complex. This was facilitated by earlier work showing that the sequence specifically binding *E. coli* IHF protein could effectively substitute for the related HU protein if an IHF cognate sequence (the H′ site from phage λ) was optimally positioned between the Sin binding sites (Rowland et al., 2006). Docking of the IHF-H′ crystal structure (Rice et al., 1996) onto the appropriate ends of the Sin-site II complex sharply redirects the DNA paths such that the DNA segments at the crossover site I are proximal to each other (Figure 2D). A model of a Sin tetramer bound to cleaved site I DNAs, which was derived from a recent crystal structure of the site I synaptic complex of γδ resolvase (Li et al., 2005), was

then added, culminating in an atomic model (schematically represented in Figure 2E) that fits all available data regarding the topology and structure of the Sin recombination complex. Importantly, the complex traps three negative supercoils and positions the DNA segments at the crossover site (site I) in the correct configuration to generate singly linked deletion circles upon DNA strand exchange. DNA exchange (Figure 2F) is accomplished by a 180° rotation of a pair of subunits about a largely flat and hydrophobic interface created upon formation of the activated tetramer (Dhar et al., 2004; Li et al., 2005).

The model of the ~350 kDa Sin recombination complex beautifully demonstrates how regulatory and catalytic subunits collaborate with architectural (DNA bending) proteins and DNA supercoiling to generate a tightly interwound nucleoprotein complex. There are no direct interactions between the DNA bending protein and Sin, consistent with the observation that HU/IHF proteins from a variety of sources function effectively. In the Tn3/$\gamma\delta$ reactions, it seems likely that the resolvase dimers bound at site II (Figure 1) may also be performing a strictly architectural role in the assembly of a similar synaptic complex structure. Residues 52 and 54 from proximal subunits bound to sites I and II in the Sin model are not close enough to interact. However, Mouw et al. (2008) argue that conformational adjustments in the linker regions of the proteins and DNA segments could enable these residues to contact each other. How these specific interactions modulate the activity of the catalytic subunits bound at site I remains to be determined.

REFERENCES

Dhar, G., Sanders, E.R., and Johnson, R.C. (2004). Cell *119*, 33–45.

Grindley, N.D., Whiteson, K.L., and Rice, P.A. (2006). Annu. Rev. Biochem. *75*, 567–605.

Keravala, A., Groth, A.C., Jarrahian, S., Thyagarajan, B., Hoyt, J.J., Kirby, P.J., and Calos, M.P. (2006). Mol. Genet. Genomics *276*, 135–146.

Li, W., Kamtekar, S., Xiong, Y., Sarkis, G.J., Grindley, N.D., and Steitz, T.A. (2005). Science *309*, 1210–1215.

Mouw, K.W., Rowland, S.J., Gajjar, M.M., Boocock, M.R., Stark, W.M., and Rice, P.A. (2008). Mol. Cell *30*, 145–155.

Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996). Cell *87*, 1295–1306.

Rowland, S.J., Stark, W.M., and Boocock, M.R. (2002). Mol. Microbiol. *44*, 607–619.

Rowland, S.J., Boocock, M.R., and Stark, W.M. (2006). Mol. Microbiol. *59*, 1730–1743.

Smith, M., and Thorpe, H. (2002). Mol. Microbiol. *44*, 299–307.

Yang, W., and Steitz, T.A. (1995). Cell *82*, 193–207.

LRR Domain Folding: Just Put a Cap on It!

Stephanie M.E. Truhlar[1] and Elizabeth A. Komives[1,*]
[1]Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0378, USA
*Correspondence: ekomives@ucsd.edu
DOI 10.1016/j.str.2008.04.002

In this issue of *Structure*, Courtemanche and Barrick (2008) describe the role of helical capping motif in nucleating the folding of leucine-rich repeat (LRR) domains.

Nearly 20% of the human genome encodes repeat proteins which are made up of 20–40 amino acid repeats. The leucine-rich repeat (LRR) domain is one such domain formed from tandem arrays containing a leucine-rich consensus sequence (Figure 1; Marino et al., 1999). The repeat contains a β strand and loop usually followed by a 3_{10} helix. The β strands of each repeat stack to form a parallel β sheet in the domain. Variable sequences connecting the consensus provide functional diversity for binding. LRR domains mediate macromolecular interactions in processes as diverse as bacterial invasion of host cells, the plant immune response, and inhibition of RNA binding. The LRR domain of Internalin B (InlB) is critical for the pathogenesis of Lysteria monocytogenes by binding to the hepatocyte growth factor receptor and activating the Ras-MAP kinase pathway (Marino et al., 2000). Sea lampreys have even evolved a primitive immune system based on the LRR domain scaffold (Binz et al., 2005). Designed proteins based on the LRR consensus sequence have been successfully produced and some of these show excellent inhibitory properties. The design not only incorporates the consensus sequence but also requires careful attention to the ''capping'' domain (Stumpp et al., 2003). Despite the tremendous utility and versatility of LRR domains, little is known about how they fold and what controls their folded stability. Barrick and coworkers have brought the field several years ahead in one single study, published in this issue of *Structure* (Courtemanche and Barrick, 2008).

Two main experimental approaches have been used to determine the folding landscapes of repeat proteins: dissection and mutation. In dissection, repeats are eliminated one at a time and the folding of the domain in the absence of one or more repeats is measured. Barrick and coworkers used this approach to define the energy landscape of the ankyrin repeat domain of Notch (Mello and Barrick, 2004). A more subtle approach is to introduce mutations that remove an interaction either within or between repeats. If