Title

Quasi-equivalence in site-specific recombinase structure and function: crystal structure and activity of trimeric cre recombinase bound to a three-way lox DNA junction11Edited by K. Morikawa

Journal

Journal of Molecular Biology, 313(1)

ISSN

0022-2836

Authors

Woods, Kevin C
Martin, Shelley S
Chu, Victor C
et al.

Publication Date

2001-10-01

DOI

10.1006/jmbi.2001.5012

Peer reviewed

doi:10.1006/jmbi.2001.5012 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 313, 49–69

JMB

Quasi-equivalence in Site-specific Recombinase Structure and Function: Crystal Structure and Activity of Trimeric Cre Recombinase Bound to a Three-way Lox DNA Junction

Kevin C. Woods[1], Shelley S. Martin[1], Victor C. Chu[1] and Enoch P. Baldwin[1,2]*

[1]*Section of Molecular and Cellular Biology, University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA*

[2]*Department of Chemistry, University of California, Davis 1 Shields Ave, Davis CA 95616, USA*

**Corresponding author*

The crystal structure of a novel Cre-Lox synapse was solved using phases from multiple isomorphous replacement and anomalous scattering, and refined to 2.05 Å resolution. In this complex, a symmetric protein trimer is bound to a Y-shaped three-way DNA junction, a marked departure from the pseudo-4-fold symmetrical tetramer associated with Cre-mediated LoxP recombination. The three-way DNA junction was accommodated by a simple kink without significant distortion of the adjoining DNA duplexes. Although the mean angle between DNA arms in the Y and X structures was similar, adjacent Cre trimer subunits rotated 29° relative to those in the tetramers. This rotation was accommodated at the protein-protein and DNA-DNA interfaces by interactions that are ''quasi-equivalent'' to those in the tetramer, analogous to packing differences of chemically identical viral subunits at non-equivalent positions in icosahedral capsids. This structural quasi-equivalence extends to function as Cre can bind to, cleave and perform strand transfer with a three-way Lox substrate. The structure explains the dual recognition of three and four-way junctions by site-specific recombinases as being due to shared structural features between the differently branched substrates and plasticity of the protein-protein interfaces. To our knowledge, this is the first direct demonstration of quasi-equivalence in both the assembly and function of an oligomeric enzyme.

Keywords: Y junction; protein-DNA complex co-crystal; DNA-binding protein; Holliday junction; integrase

Introduction

Site-specific recombinases catalyze breakage and rejoining of DNA at defined motifs.[1] One enzyme, phage P1 Cre recombinase, promotes recombination between two 34 bp LoxP recognition sequences.[2] Cre belongs to the Int superfamily of recombinases and type IB topoisomerases, a divergent group of enzymes that share a conserved active site architecture and chemical mechanism.[3–5] The Cre-LoxP system functions efficiently *in vitro*[6] and in a variety of cultured cells[7] and living organisms[8] without any additional factors. As a result, it is an invaluable tool for the manipulation of genomic DNA[9] particularly for targeted insertion of transgenes[10,11] and construction of conditional gene ''knockouts''.[12] The relative abundance of biochemical and structural data

Abbreviations used: 13 bp repeats, two 13 bp inverted repeats that harbor the determinants for LoxP protein binding; 8 bp spacer, 8 bp sequence that separates the two 13 bp repeats; HJ, Holliday junction; YJ, Y-shaped three-way junction; $3HS_{1or2}$, YJs that contain one or two unpaired nucleotides inserted at the junction; 3H, YJs with completely paired bases at the junction; AFM, atomic force microscopy; Cre*, His-tagged Cre; CRY, novel Cre-LoxA complex discussed here; 1CRX, covalent cleavage intermediate; 4CRX, precleavage intermediate; 3CRX, HJ-bound intermediate; EMSA, electrophoretic mobility shift assay; Lox3Y, oligonucleotide YJ substrate with three arms; Lox2Y, YJ substrate with four mutations in one of the 13 bp repeats; Cre* (Y324F), Cre* with Phe substituted at position 324.

E-mail address of the corresponding author: epbaldwin@ucdavis.edu

0022-2836/01/010049–21 $35.00/0

make the self-contained Cre-Lox system ideal for mechanistic and protein engineering studies of site-specific recombination.

Two 34 bp LoxP sites are the minimal and sufficient recombination substrate.[13,14] LoxP consists of two 13 bp inverted repeats that harbor the determinants for protein binding (''13 bp repeats''), separated by an 8 bp intervening sequence (''8 bp spacer''). Cutting and rejoining take place between the first and second nucleotides within the 8 bp spacer (arrows, Figure 1(a)).[6,15] Biochemical experiments by Hoess and co-workers[6,14–16] and seminal crystallographic studies by Guo, Gopaul, and van Duyne[17–19] revealed an elegant mechanism for Cre-mediated recombination (Figure 1(c)). These crystal structures defined the three key reaction intermediates, the precleavage intermediate (4CRX[19]), the covalent cleavage intermediate (1CRX[17]), and the Holliday junction-bound (HJ-bound) intermediate (3CRX[18]). In the complexes, four Cre molecules are bound to a nearly planar four-way DNA junction (Figure 2(b)). Adjacent Cre subunits are in different conformations, A and B, and the unique AB dimers generate a pseudo-4-fold tetramer through crystallographic symmetry. The Lox DNA arms form an unequally splayed X-shaped structure with 102-106° and 65-72° angles between them (Figures 1(c), 2(b), and 6(b)). The ''cleaving'' B subunit[17] performs the DNA cleavage and rejoining reactions and has slightly different conformations in the intermediate structures, while the ''non-cleaving'' A subunit changes little. Positional shifts in the active site, the 8 bp spacer, the three quaternary protein-protein interfaces, and the relative orientations of the subunits and DNA arms, account for the majority of differences between the intermediates.

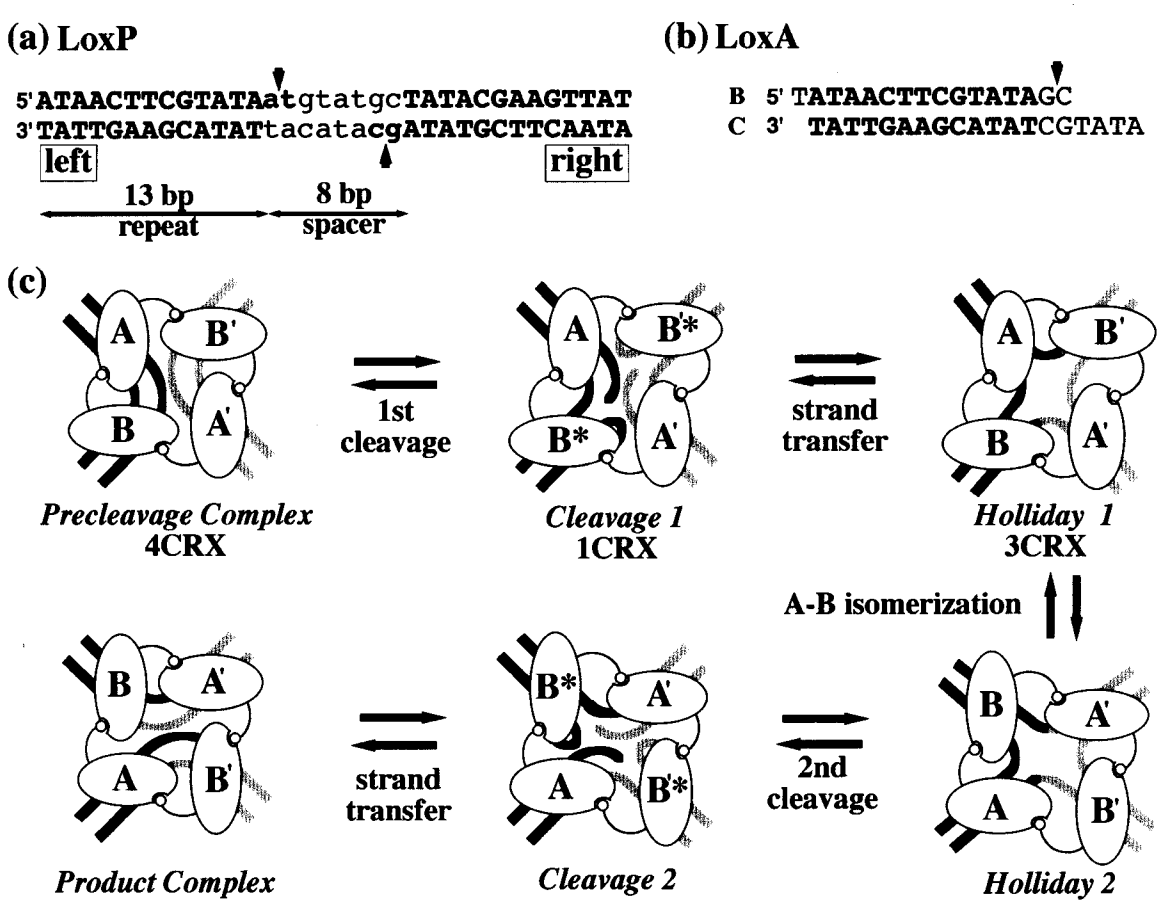

Figure 1. Lox DNA sequences and the mechanism of site-specific recombination. (a) Both strands of the LoxP site showing the left and right 13 bp repeats (upper case, bold), the 8 bp spacer (lower case), and the scissile sites (lower case, bold, arrowheads). (b) LoxA DNA duplex used for crystallization of CRY and 1CRX complexes.[17] The ''B'' and ''C'' strand designations correspond to the chain references in the structure. The 13 bp repeats are indicated in bold lettering. Two LoxA half-sites anneal at the 4 nt overhang to form a complete symmetric Lox sequence. (c) Proposed mechanism of site-specific recombination showing the correspondence between the 1CRX,[17] 3CRX,[18] and 4CRX[19] structures and putative reaction intermediates (adapted from Guo *et al*,[17]). The AB and A′B′ protein dimers are related by 2-fold symmetry. Analogous intermediates in the upper and lower pathways are aligned vertically, and differ in the relative orientation of the Lox DNA with respect to the A and B subunits. The covalently attached Cre monomers are marked with an asterisk (*).

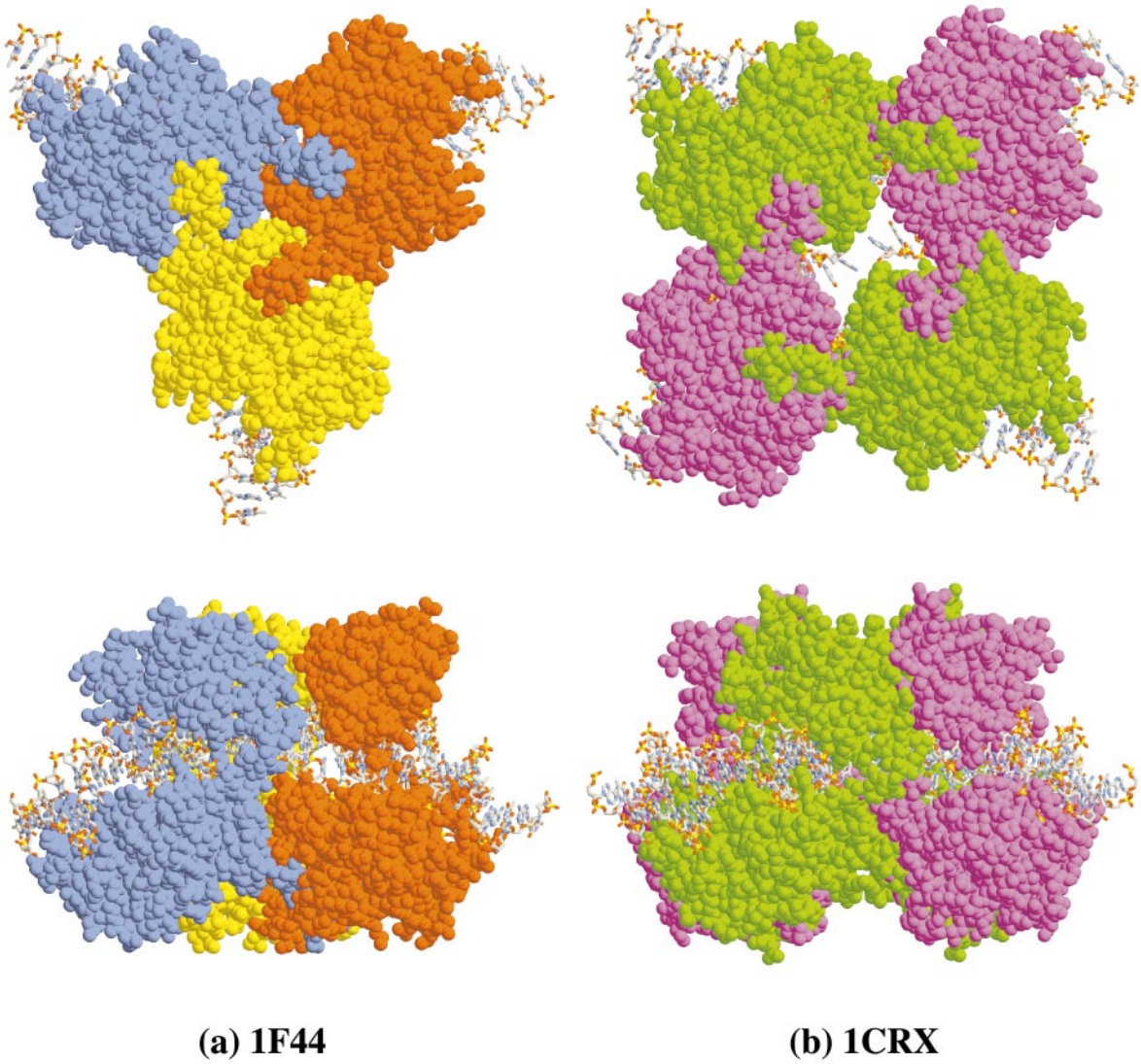

Figure 2. Architecture of trimeric and tetrameric Cre-Lox synaptic complexes. (a) The 1F44 CRY trimer structure with three identical Cre subunits in yellow, red, and blue. (b) 1CRX tetramer structure, with the ''non-cleaving'' A subunit in purple and the ''cleaving'' B subunit in green.[17] Complexes are shown in C-terminal (upper panels) and side (lower panels) views. The protein is represented as space-filling atoms and the DNA in balls-and-sticks.

Recombination is initiated when four inactive Cre monomers[6] and two LoxP sites assemble cooperatively[20] into the active tetrameric synaptic complexes (Figure 1(c)). In the precleavage complex, the bound DNA is bent, bringing the 8 bp spacers into proximity in the large solvent-accessible ''strand transfer channel''. In the cleavage complexes, each Lox site is nicked and covalently attached to Tyr324[15,21] of the B subunit *via* a 3′ phosphodiester linkage. The cleaved 5′ ends diffuse across the channel (Figure 2(b)) and transesterify to the 3′ phosphate groups of the opposite strands creating a Cre-HJ complex. Cleavage and rejoining of the second set of strands is initiated by a global conformational switch[22] during which the HJ isomerizes and the A and B subunits exchange conformations.[17,18] The reaction is completed by an analogous set of steps (Figure 1(c), lower path). A recent crystal structure of the FLP-FRT Holliday junction complex[23] supports the generality of the Cre-Lox mechanism for Int family recombination, but with significant variations in the details.[21,24–26]

Watson-Crick base-pairing not only allows for the formation of double-stranded DNA, but also the creation of branched structures in which multiple strands are intertwined to form junctions[27–29] that join duplexes. The well-studied four-way DNA junction (HJ, Figure 3(a)) is a central DNA intermediate in recA/rad51-mediated[30] homologous recombination as well as in site-specific recombination. Less commonly investigated are Y-shaped three-way junctions (YJs, Figure 3(a)). YJs arise as recombination intermediates from double-strand break repair, retroviral integration, transposition, and during DNA replication at the meeting of leading and lagging strands.[27,28] They accumu-

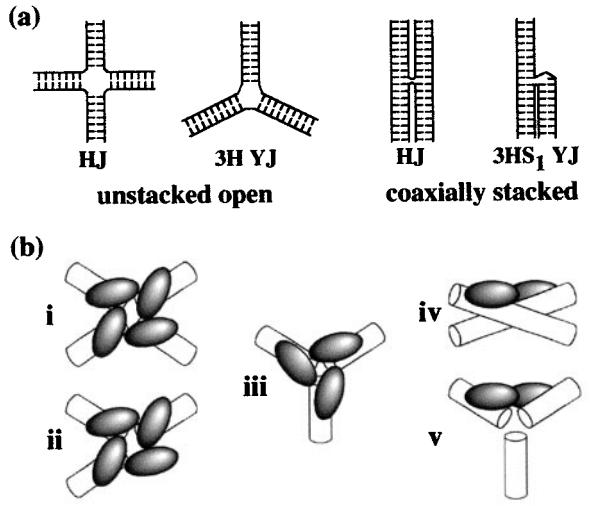

Figure 3. Holliday and Y-DNA junctions and their recognition by oligomeric proteins. (a) Comparison of ''unstacked open'' and ''coaxially stacked'' forms of Holliday junctions (HJs) and three-way junctions (YJs). The open form is favored by YJs with completely paired bases (3H YJ, left), and contains unstacked bases and arms that are equally separated. In the coaxially stacked form, one pair of arms forms a continuous helix. This configuration is particularly favored by one or two unpaired bases at the junction which facilitates the insertion of the third helix (the $3HS_1$ YJ contains one inserted base, right). (b) Models for recognition of YJs and HJs by oligomeric proteins; (i) tetramer bound to an HJ with one subunit on each arm; (ii) tetramer bound to a YJ with one empty binding site; (iii) YJ binding by a trimer; (iv) HJ endonuclease dimer binding two arms of coaxially stacked HJ. Open forms may also be bound analogously; and (v) endonuclease dimer binding to adjacent arms of a YJ.[52]

late in some prokaryotic[31] and eukaryotic[32] viral DNAs *in vivo*. DNA YJs are implicated in hereditary diseases that arise from the genetic instability of repeating short DNA sequences. These repeating sequences induce DNA slippage during replication, possibly through the formation of YJs,[33] leading to their expansion and the associated changes in gene expression. However, whether YJ DNAs are transitory intermediates or are key to a particular aspect of DNA metabolism remains to be demonstrated. In contrast, YJs are a common motif in structural and catalytic RNAs and RNA-protein complexes.[29]

Most structural studies have focussed on YJs that contain unpaired nucleotides inserted at the junction (''3HS'',[34] Figure 3(a)), the form most commonly found in RNA. The insertion allows these molecules to favor the ''coaxially stacked''[28] configuration, in which two arms form one continuously stacked duplex interrupted by the third arm[40–42] (Figure 3(a)). The base stacking at the junction likely confers the relatively high stability of these molecules. Free HJs also favor the analogous coaxially stacked form[35–38] but are found in the open form in protein-HJ complexes.[18,23,39,40] In contrast, ''3H'' YJs, those that have completely paired bases at the junction, are less stable. Their structures have been investigated using NMR,[41] fluorescence resonance energy transfer,[42] ligation kinetics,[43] and atomic force microscopy (AFM).[44] The emerging view is that they adopt a dynamic pyramidal ''open'' geometry consisting of three duplex arms joined at the ends (Figure 3(a)). The angles between arms apparently fluctuate between 60 and 80°,[44] and the bases at the duplex termini are unstacked. Until now, no high-resolution structure of any YJ DNA has been reported.

The potential recognition of DNA YJs by cellular components is largely unexplored. YJ-specific DNA binding proteins have not been identified, but several proteins that recognize HJs, including site-specific recombinases,[45,46] also interact with YJs. FLP recombinase binds and promotes efficient resolution of YJs harboring FRT recognition sequences.[45] Similarly, phage λ integrase also resolves *att*-containing YJs but less efficiently.[46] Another tetrameric protein, human p53, binds YJs within threefold affinity as it does HJs.[47] Without structural information, two possible models could account for tetramer recognition of YJs (Figure 3(b)). A tetrameric enzyme might bind a YJ, leaving one active site empty (ii, Figure 3(b)). Alternatively, a trimer could be assembled to conform to the symmetry of the substrate (iii, Figure 3(b)). This second view might seem unlikely, since the protein would have to assemble and retain activity in two different oligomerization states.

Several HJ-specific endonucleases also recognize YJs and cleave them position-specifically.[48–52] These enzymes were proposed to span the HJ as a dimer to interact with two adjacent arms across the junction (iv, Figure 3(b)), and to bind analogously to adjacent arms in a YJ (v, Figure 3(b)).[52] Unless these enzymes are capable of forming trimers (iii, Figure 3(b)), they must be flexible enough to recognize structural differences between the two substrates. Alternatively, HJs and YJs may have enough structural similarity to allow for cross-reactivity.

In our crystallographic investigations of Cre-Lox complexes, we identified a novel three-way synapse (Figure 2(a)). In this structure, a Cre trimer is bound to an open three-way Y-shaped DNA structure representing a symmetry variant of previously characterized tetrameric complexes[17–19] (Figure 2(b)). As suggested by the structure, Cre binds and cleaves the corresponding symmetry variant Lox YJ substrate. We conclude that Cre and other site-specific recombinases can form active ''quasi-equivalent'' complexes to conform to the different branch numbers of DNA substrates. The different quaternary interactions in trimeric and tetrameric complexes are analogous to the adaptation of subunit structures to non-equivalent

packing environments at different symmetry positions in icosahedral capsids.[53–56]

Results

Structure solution

In our hands, Cre, His-tagged Cre (Cre*) and several mutants readily crystallized with a number of Lox half-site variants using PEG8000 as a precipitant, but these crystals typically diffracted poorly. Singularly, the inactive Tyr324-Phe mutant (Cre*Y324F) and the ''LoxA'' half-site from 1CRX,[17] (Figure 1(b)) yielded crystals that diffracted to better than 3 Å using Cu-Kα radiation. Less well diffracting isomorphous crystals (~4 Å) were obtained using the Cre* or a Tyr324-His Cre* variant and LoxA. Solving the structure of a variant Cre-LoxA complex offered the advantage of direct comparison with the 1CRX structure. Crystals of Cre*Y324F/LoxA, 0.4 mm × 0.4 mm × 0.4 mm in size, grew from PEG8000, $LiSO_4$, and Pipes buffer at pH 6.1. The space group was cubic, *I*23, and crystals diffracted to 2.8-3.0 Å with CuKα radiation at 21 °C and to 2.55 Å at 100 K using methane-2,4-pentanediol (MPD) as a cryo-protectant.

Native and three derivative data sets were collected at 100 K to 2.55 Å and to 2.8-4.0 Å resolution, respectively. A readily interpretable electron density map was obtained using phases calculated from isomorphous and anomalous differences to 3 Å (Table 1), followed by solvent-flattening and phase-extension to 2.8 Å (overall Figure of merit = 0.51). The initial model was independently traced and refined against 2.55 Å data without explicit comparison with other Cre structures, which were unavailable at that time. The final 2.55 Å model had an *R*-factor of 18.4 % and an R_{free} of 26.7 % for 5-2.55 Å data. Subsequently, a 2.05 Å data set at 100 K was obtained at SSRL beamline 9-1, from a crystal grown in the above conditions supplemented with 9 % MPD, and used

Table 1. Data collection, phasing, and refinement statistics

Data set	Resolution (last shell)[a] (Å)	Coverage (%)	R_{merge} (last shell)[a] (%)	Source	Wavelength (Å)	Device
Native 1	2.55 (2.65-2.55)	99.7	6.2 (29.8)	CuKa	1.54	Raxis4 IP
Native 2	2.05 (2.12-2.05)	98.6	4.4 (33.4)	SSRL 9-1	0.98	MAR IP
A. *Heavy Atoms*						
Derivative	Resolution (iso/ano) (Å)	Coverage (%)	R^{a}_{merge} (%)	R^{b}_{iso} (%)	R^{c}_{anom} (%)	Phasing power[d] (cen/acen)
Iodine-1	3.0/3.3	99.6	5.3	8.9	3.3	1.33/1.08
Iodine-2	3.0/3.3	95.4	6.4	14.9	4.0	0.79/0.61
Thimersol	4.0/4.0	99.1	3.0	11.1	2.2	0.85/0.62
Figure of merit:			28-3.00 Å MLPHARE			0.531
			28-2.80 Å DM/phase extension			0.505

B. *Refinement*	1DRG	1F44
Resolution (Å)	5-2.55	5-2.05
Cell (Å)	*I*23 *a*,*b*,*c*=160.99	*I*23 *a*,*b*,*c*=160.54
No. reflections (completeness)		
Working set	18,634 (95%)	37,741 (94%)
Free set	981 (5%)	2063 (5%)
Total no. atoms	3331	3537
No. protein atoms	2485	2516
No. DNA atoms	711	711
No. solvent atoms	114	310
Average *B*-factor	49.7	44.8
Protein	49.6	43.9
DNA	50.1	45.8
Solvent	49.7	49.6
rms bond lengths (Å)[e]	0.009	0.010
rms bond angles (deg.)[e]	1.28	1.25
rms *B* (overall) (Å^2)[e]	3.4	2.94
R_{factor}[f]	0.186	0.176 (28-2.05 Å=0.195)
R_{free}[f]	0.268	0.224 (28-2.05 Å=0.240)
Ramachandran plot[g]		
Allowed (%)	98.2	99.6
Generously allowed (%)	99.6	99.6

[a] $\Sigma|I - \langle I\rangle|/\Sigma\langle I\rangle$ for all of the data.[89]
[b] $\Sigma|\langle I\rangle_{native} - \langle I\rangle_{deriv}|/\Sigma\langle I\rangle_{native}$ for all of the data.[89]
[c] $\Sigma|\langle I\rangle^{+} - \langle I\rangle^{-}|/\Sigma\langle I\rangle$ for all of the data.[89]
[d] $\langle F_H\rangle/\langle F_{PH} - F_P - F_H\rangle$, MLPHARE.[92]
[e] Calculated by TNT[94] using Engh & Huber parameters.[95]
[f] $\Sigma|F_o - F_c|/\Sigma|F_o|$.
[g] From PROCHECK.[101]

for further model building and refinement. The improved model differed somewhat from the lower resolution model in the α-helical region residues 21-34 (see below). This model had *R*-factor and R_{free} values of 17.5 % and 22.4 % for 5-2.05 Å data, respectively.

These structures represent an independent solution for Cre recombinase, but agreed well with the CRX structures except where noted herein. The 2.05 Å model contains a single Cre/LoxA complex in the asymmetric unit, residues 20-198 and 208-343 and all of the DNA. Electron density for residues 1-19 and 199-207 was not visible in our MIRAS or refined maps. Residues and 199-208 are absent in several CRX subunits, while residues 1-18 and 342-343 are absent in all of the structures.

Overall trimer structure

The novel complex, hitherto designated ''CRY'', contains three identical Cre*Y324F/LoxA complexes arranged in a 3-fold symmetric trimeric synapse. In Figure 2, the CRY trimer and covalent cleavage intermediate (1CRX) tetramer are compared in similar views. The complex resembles a flat-topped trigonal pyramid, 110 Å on a side and 70 Å thick. The trimer 3-fold axis is coincident with the crystallographic symmetry axis along the *I*23 body diagonal. As in the other structures, the Cre monomer consists of the N-terminal domain (residues 20-126), connected by an inter-domain linker (residues 127-133) to the C-terminal domain 134-343 (Figure 4(a)). The monomer forms a clamp surrounding the Lox DNA duplex. The N-terminal domains form the top of the pyramid, the C-terminal domains form the base, and the Lox half-site DNA molecules define the vertices. Protein and DNA at the junction are more densely packed than in the tetramers. The protein-DNA interface is closed to solvent on the C-terminal side of the complex (Figure 2(a)), while on the N-terminal side, a channel is formed into the center of the complex, lined by the 103-111 loop from each N-terminal domain and by the junction DNA base-pairs. The relative inaccessibility of the junction contrasts that in the tetrameric structures, in which the strand transfer channel extends from the top to the bottom of the complexes (Figure 2(b)).

Both the DNA and protein participate in complex assembly through a combination of tetramer-derived and novel contacts (Figure 5). At the three-way DNA junction, the DNA half-sites are linked through Watson-Crick base-pairs (Figure 6(a)). The angle between the mean helical axis of each half-site duplexes (CURVES[57]) in the three-way junction is 107°, similar to the 102-106° bending angles of the Lox duplexes in the CRX structures[57] (Figure 6(b)). On the other hand, the 60° angle between the protein monomers in the trimer is more acute than that between those in the tetramers[58] (Figure 5(a)). These angular differences result in a 29° rotation and a 6 Å shift of the adjacent CRY monomer compared to the analogous A monomer in 1CRX[58] (Figure 4(b)). This repositioning prevents the formation of two of the three tetrameric contacts (Figures 4(b) and 5(a)) but is compensated by the formation of a new 3-fold interface (Figure 5(b)).

The CRY Cre-LoxA complex is similar to the B subunit-DNA complexes in all CRX structures with regards to the backbone conformations and DNA-protein contacts, but particularly resembles the B subunit of 1CRX (1CRX-B). The main chain rmsd for residues 40-198 and 208-328 is 0.73 Å compared to 1.0-1.2 Å for the B subunits and 1.2-1.5 Å for the A subunits. The 14 bp of DNA duplex outside the junction are essentially identical (rmsd = 0.50 Å). In Figure 4(a), the CRY N-terminal and C-terminal domains are shown superimposed on the corresponding 1CRX-A (magenta) and 1CRX-B (green) domains, respectively. The largest position and mobility differences map to the altered subunit interfaces, as a direct result of the trimeric oligomerization state. The details of these differences are depicted in Figures 4 and 5 and are described below.

Three-way DNA junction

This structure represents the first crystallographic view of a three-way DNA junction. Three LoxA half-sites duplexes meet in a symmetric ''open'' Y,[34] that is essentially three *B*-DNA segments joined at their ends. The palindromic ATAT 5′ extensions of the 19-mer C strand of LoxA (Figure 1(b)) are intertwined in a cyclic arrangement of reciprocal Watson-Crick base-pairs in which the first 5′ AT dinucleotide on one arm pairs with the second AT dinucleotide in the adjacent arm (Figure 6(a)). The base-pairs at the duplex termini are unstacked and line the triangular solvent-accessible junction center. The sharp 60° turn of the junction is accomplished primarily by a rotation in the phosphodiester backbone linkage that bridges the duplexes ($\varepsilon = -98°$, compared to a mean of $-161°(\pm 25)°$ for all of LoxA), resulting in a −25° tilt and 63° roll angle between the bases (FREEHELIX98[59]). The base positions and helical parameters are very similar to the non-crossover strand in the 3CRX Cre-HJ complex, with an rmsd of 0.95 Å for the sugar-phosphate backbone of the four bridging residues. The two DNA structures deviate greatly outside the junction.

Overall, the CRY YJ is not planar, but forms a shallow dome from the downward curvature of the DNA arms (Figure 2(a), bottom panel, Figure 6(c)). The outer ends of the Lox half-sites are displaced 22 Å and 22° out-of-plane relative to the junction center. The apparent angle between the arms differs depending on the DNA residues used for the calculation.[57] Using the entire half-site, the inter-arm angle is approximately 107°, less than the expected 120°, and quite similar to the obtuse angle in 1CRX (106°) (Figure 6(b)). However, this angle approaches 115°-120° using the two to four base-steps adjacent to the junction to

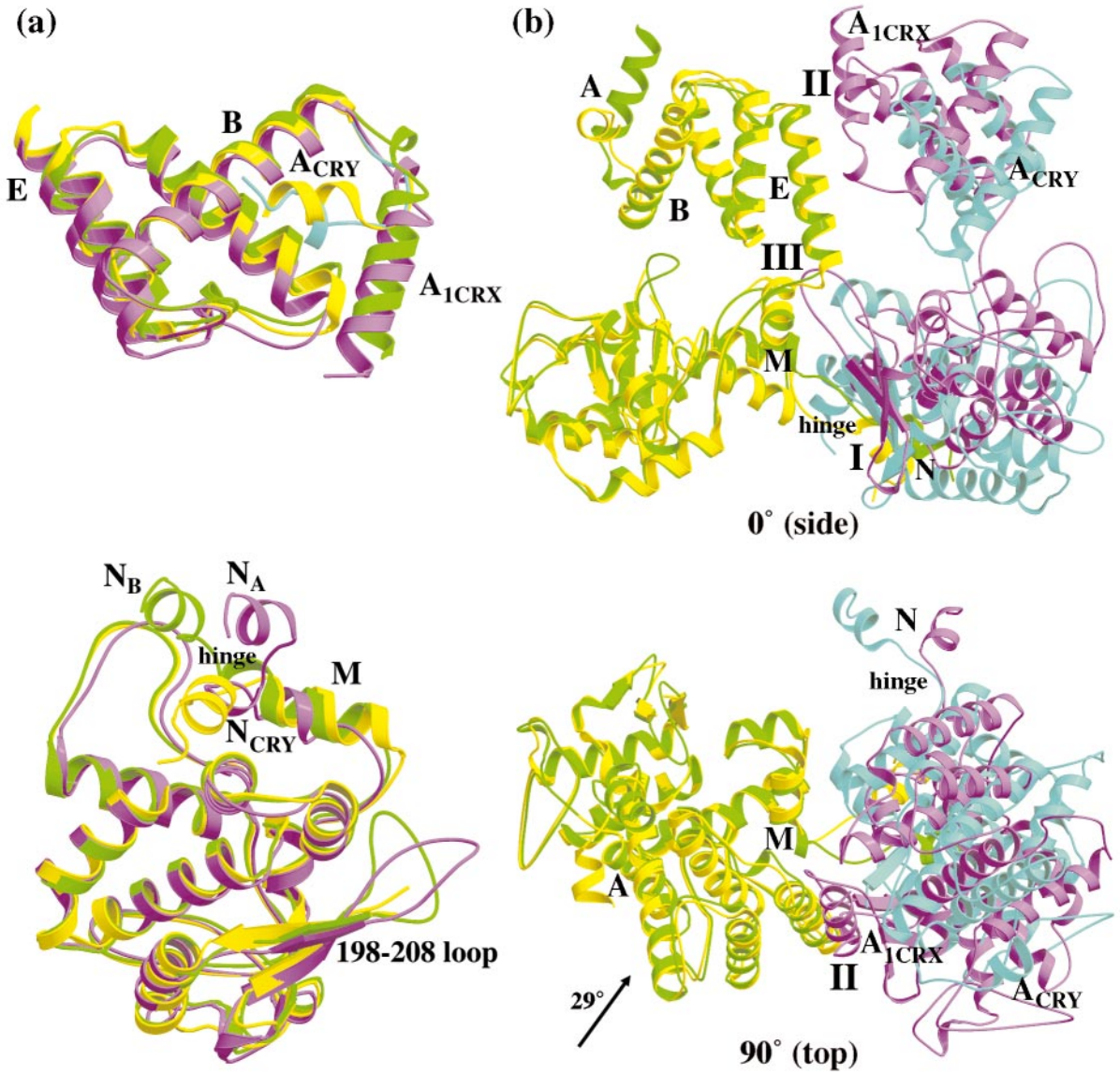

Figure 4. Structural comparison of Cre protein in trimeric and tetrameric complexes. (a) Ribbon diagram comparisons of N-terminal domains (upper panel) and C-terminal domains (lower panel) of 1F44 (yellow), 1CRX-A (magenta) and 1CRX-B (green) subunits. Upper case letters A, B, E, M, and N indicate the positions of particular helices[17] with subscripts designating the associated Cre monomer, while ''hinge'' indicates residues 328-333 which connect the ''domain-swapped'' helix N to the remainder of the C-terminal domain. For the N-terminal domain comparison (upper panel), the main chain atoms of residues 40-126 were superimposed. Note the large positional shift of helix A, residues 20-38. In the 2.55 Å resolution structure, this region had a somewhat different position than in the 2.05 Å resolution structure (cyan ribbon). For the C-terminal domain comparison (lower panel), the main chain atoms of residues 134-326 were superimposed. Note the similarity of helix M from CRY and 1CRX-B, and the heterogeneity of helix N positions. (b) Ribbon diagrams showing the relative positioning of adjacent subunits in CRY and 1CRX in two orthogonal views. One CRY subunit (yellow) was superimposed on the 1CRX-B *(green)* using the main chain atoms of residues 40-198 and 208-326 and the resulting matrix was used to transform a CRY dimer. Three quaternary interfaces are indicated; ''I'', the helix N-C-terminal domain contact; ''II'', inter-subunit pairing of helices A and E; and ''III'', the contact of the A subunit 198-208 loop to the adjacent B-subunit (see the text). The second CRY subunit (cyan) rotated 29° and shifted 6 Å with respect to the 1CRX-A subunit, substantially increasing the distance between A and E helices on adjacent molecules. The 29° rotational axis is indicated (arrow, lower panel). It is oriented 54° from the superimposed monomer and 6° above the trimer plane.

define the helix axes. Thus, the pyramidal nature of the complex is less due to the presence of a pyramidal junction,[41–44] but results primarily from the approximate 25° bend in the protein-bound portion of the LoxA half-site. In the CRX structures the curve is oriented more in the complex plane, while in CRY it points towards the C-terminal side of the pyramid. This relative rotation of the YJ arms compared to the arms of the Cre-Lox Holliday junction leads to a ~26° angular offset

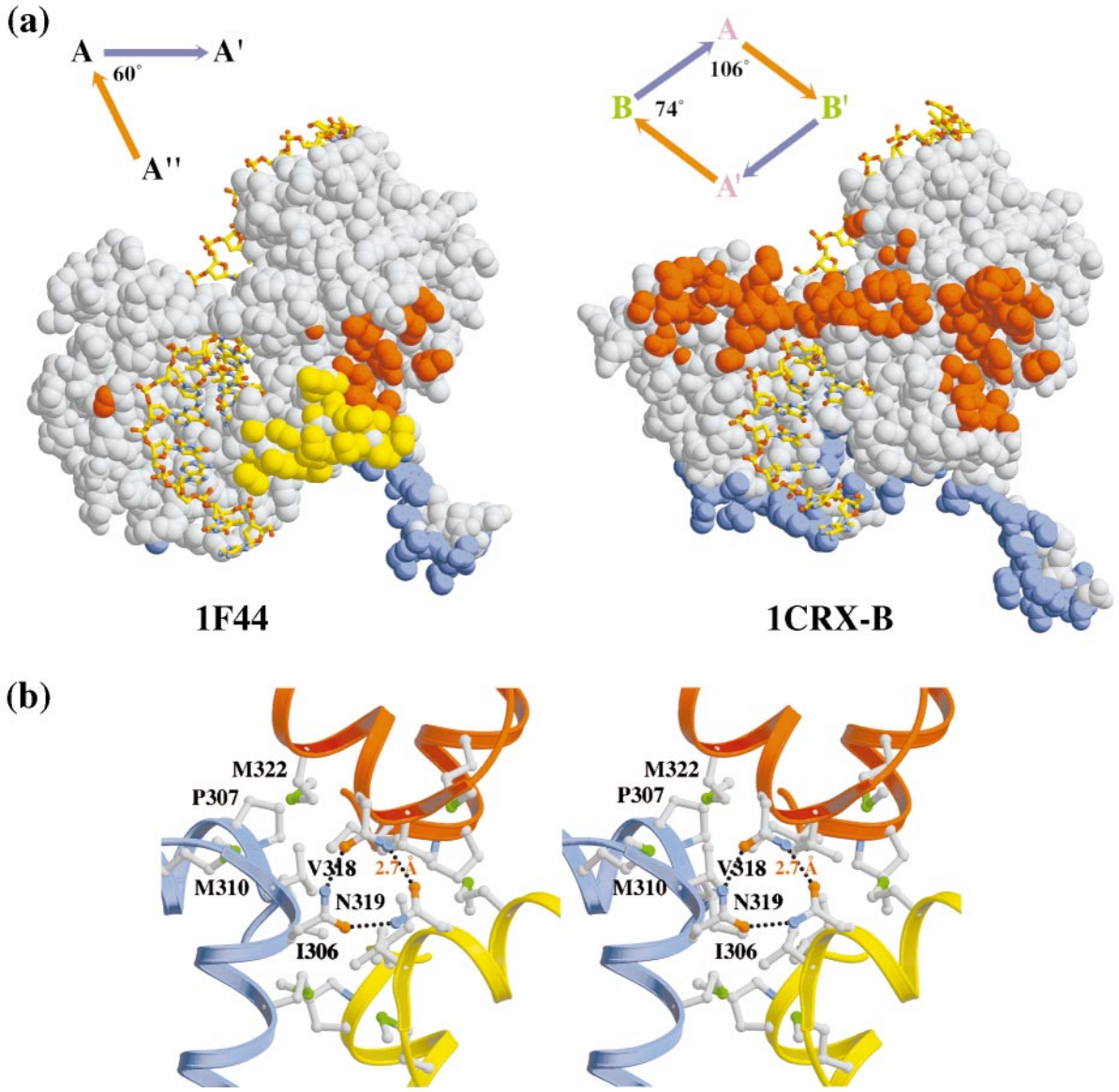

Figure 5. Comparison of trimeric and tetrameric protein-protein interfaces. (a) Quaternary contacts in CRY (right) and 1CRX-B (left). Colored surfaces are those that are inaccessible to a 1.4 Å probe when in contact with the adjacent subunit in the complex. Color-coding denotes the contacting partner as depicted in the upper left diagrams. The approximate rotational angles between adjacent subunits are also indicated. The new trimeric interface in 1F44 is shown in yellow. (b) Stereoview of the three-way interface, showing the layers of hydrophobic residues formed by 3-fold interactions of Ile306 and Val318, the packing of Met322 into a pocket formed by Pro307 and Met310, and the Asn319 hydrogen bond network. The main chain paths of residues 306-322 for each chain are shown in red, yellow, and blue ribbons. The side-chains at the interface are shown as ball-and-sticks.

between adjoining DNA arms after superposition of 1CRX-B and CRY half-sites (Figure 6(c)). This relative rotation also corresponds to a similar difference in dispositions of the adjacent protein subunits (Figure 4(b), lower panel).

Subunit interactions

The trimeric nature of the CRY complex leads to ''quasi-equivalent'' packing of the Cre monomers in the complex. In the discussions below, monomer comparisons are between CRY and 1CRX-B, while comparisons of quaternary contacts are between adjacent CRY subunits and interactions of 1CRX-B with its neighboring 1CRX-A subunits. Because of the cyclic nature of these contacts, a monomer has distinct ''donor'' and ''acceptor'' surfaces for each interface, corresponding to the red and blue surfaces in Figure 5(a). Buried surface areas are given for one monomer and, for simplicity, the donor and acceptor surface areas for a each contact are combined.

In the CRX tetramers there are three major subunit interfaces, I, II and III (Figure 4(b)). For 1CRX-

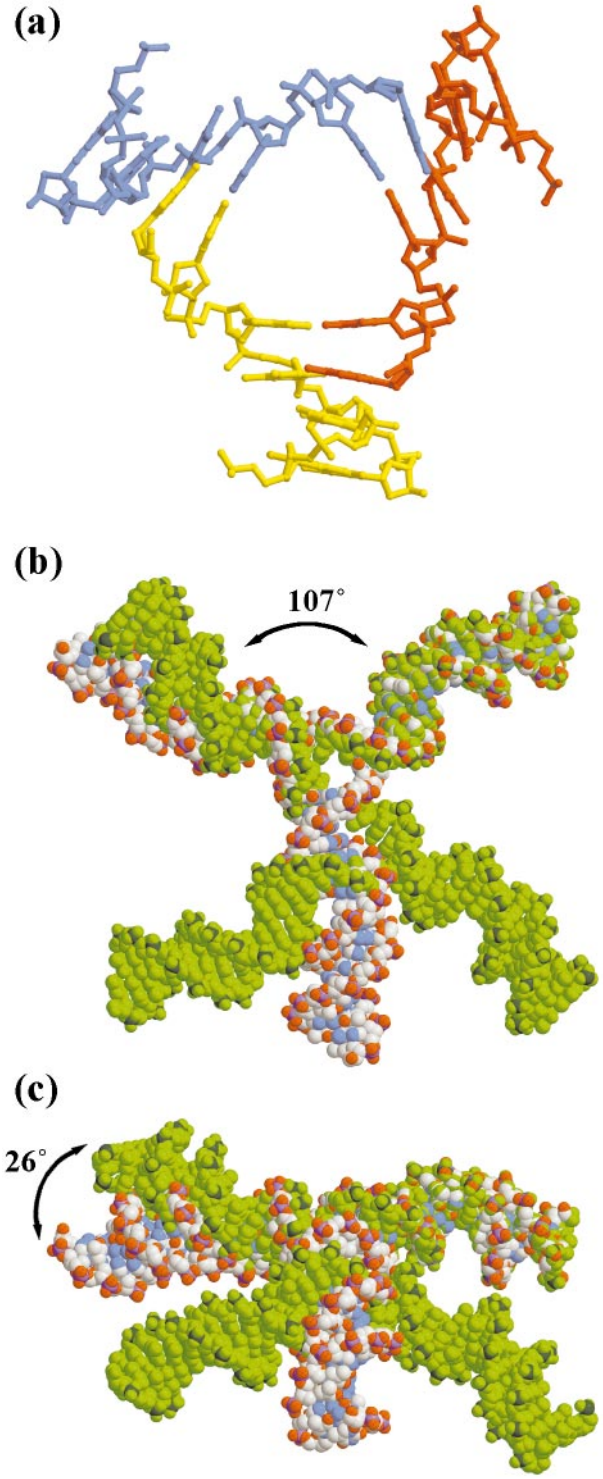

Figure 6. Structural details of the Y DNA junction. (a) Closeup view of the interlocking base-pairs of the junction showing the first seven residues of each of the C strands in yellow, red, and blue. (b) Comparison of the DNA arm geometries of CRY and 1CRX junctions from top view and (c) 60° from the 3-fold axis. All atoms of the 13 bp repeat and scissile nucleotides from 2.05 Å resolution LoxA DNA were superimposed on the equivalent atoms of LoxA bound by 1CRX-B (rmsd 0.5 Å), and used to transform the entire Y structure (atom colors) for comparison to 1CRX (green). Although the angles between arms are similar, 106° *versus* 107°, the relative positions of adjacent arms are offset by a ~26° rotation. The pyramidality of the junctions is apparent in the 60° view.

B, a total of 3187 Å^2 solvent-accessible surface is buried,[58] 1787 Å^2 of which is hydrophobic. The largest surface, interface I, 1428 Å^2, is created by ''domain-swapping'' of the C-terminal ''hinge'' and helix N (residues 327-341), which pack in a groove in the C-terminal domain of the adjacent monomer. Interface I is conserved in Int class recombinases but represents a point of divergence for the *trans* and *cis*-cleaving enzymes.[23,60] Its configuration in Cre and HP1[61] is presumably representative of the *cis*-cleaving sub-group. For the *trans*-cleaving proteins, illustrated by the FLP-FRT structure,[23] the helix containing the catalytic tyrosine (helix M in Cre) is exchanged between domains rather than the C-terminal helix. Interface II is created by inter-subunit pairing of helices A and E. At interface III, the A subunit 198-208 loop packs into a pocket formed by residues 85, 86, 130-132, 326 and 327 of the B subunit. The B subunit 198-208 loop contacts the scissile DNA base in its own active site through the Lys201 side-chain, but also packs against residues 125-131 in the other A subunit (A′, see Figure 5(a)).

The contact surfaces of CRY and 1CRX-B are compared in Figure 5(a). Overall, CRY buries a total of 1954 Å^2 surface area per monomer, 1199 Å^2 of which is hydrophobic. This corresponds to about 61% and 67% of the 1CRX-B contact surface. Only interface I is preserved and buries the majority of surface area, about 1314 Å^2. The small reduction of buried surface is due to the different positioning of the hinge residues 327-332. The CRY hinge is well-defined but traces a distinct path from those in the tetramers to accommodate the large relative rotation of the adjacent monomer (Figure 4(b)). The hinge position is highly variable between reaction intermediates, as well as between A and B subunits within a tetramer (Figure 4(a)).

The relative orientation of the CRY subunits increased the gaps at both interfaces II and III, and prevented their formation (Figure 4(b)). In the absence of tetramer contacts, these regions are disordered in CRY. There is no visible electron density for the 198-208 loop in MIRAS and $F_o - F_c$ maps. Although marginal, sufficient electron density was observed to trace helix A. In spite of the blurred density, the placement of the main chain was unambiguous in both the 2.55 Å and 2.05 Å resolution structures. Residues 20-34 adopted a very different conformation from those in the tetramer. Residues 30-34 unfolded to form a loop connecting to the remaining helical residues 20-29. The helical segment translated 15 Å and rotated nearly 90° to pack a hydrophobic patch on helix B of its own subunit. This shift converted a parallel inter-subunit helix-helix contact into an anti-parallel intra-subunit helix-helix pair. The packing partner of helix A, helix E, was relatively well-defined in

CRY but also had markedly increased mobility with an average of 30 Å^2 higher main-chain *B*-factors. The marginal density may be due to conformational heterogeneity. The positions of residues 29-34 differed somewhat between 2.55 Å resolution structure and the 2.05 resolution structure (compare yellow and cyan ribbons in Figure 4(a), rmsd main chain = 2.8 Å), but are clearly distinct in each structure, suggesting two possible stable conformations. The Trp42 side-chain appears as two rotamers in the 2.05 Å structure, perhaps also reflecting such disorder. The differences between the two data sets may be a function of crystallization conditions, which could have favored alternate conformers.

The CRY trimer contains a novel three-way interface, indicated by the yellow surface in Figure 5(a). This interface accounts for 40 % of the total and hydrophobic buried surface, 787 Å^2 and 462 Å^2, respectively. It is created by the packing of side-chains from the N terminus of helix M, the C terminus of the preceding helix L and the connecting loop (302-323, Figure 5(b)). Two layers of hydrophobic and van der Waals interactions are formed by 3-fold interactions between Ile306 and Val318 residues from each chain. The Met322 side-chain packs into a pocket formed by side-chains of Pro307, Met310 and Val318 from the adjacent subunit. A 3-fold polar network is created by a cyclic arrangement of reciprocal hydrogen bonds between the Asn319 side-chains. Overall, the quality of the interactions suggest that this interface stabilizes the trimeric complex.

Active site

Cre has six active site residues that are conserved in the Int/TopoIB family of enzymes.[4] Arg173, Arg292 and Trp315 coordinate the scissile phosphate group, and His289 appears poised to act as a general base for deprotonating the hydroxyl groups of the Tyr324 nucleophile during strand cleavage. Lys201 contacts the scissile base,-[17,19] but has been proposed to protonate the 5′ hydroxyl group of the leaving strand during cleavage.[62] While the disorder of the 198-208 loop precluded visualization of Lys201, the other five conserved side-chains were clearly defined in the CRY active site. Their positions were most similar to 1CRX-B subunit (Figure 7) with the main chain atoms differing 0.45 Å rms after superposition of the five residues, compared to 1.3-1.5 Å for other CRX subunits. Helix M, which contains the catalytic tyrosine (Figure 4(a)), is distinct in 1CRX-B from the other complexes in both length and position. This configuration is imitated in CRY. The rmsd for residues 318-326 after superposition of residues 138-198 and 208-317 from 1CRX, 3CRX, and 4CRX, was 1.3 Å, 3.2 Å and 3.5 Å, respectively. The side-chain of Phe324 was rotated 85° about χ1, away from the phosphate backbone. The rotation mimicked the phosphotyrosine linkage by allowing helix M to pack close to the DNA as it does in 1CRX-B. In other intermediates, the Tyr324 hydroxyl group sterically hinders this close approach (magenta side-chain, Figure 7). The side-chain conformations of His289, Arg292 and Trp315 altered slightly, preserving their interaction with the scissile phosphate group, which was displaced 1.5 Å in CRY compared to 1CRX. In contrast, Arg173 rotated 95° about χ3 shifting the guanidi-

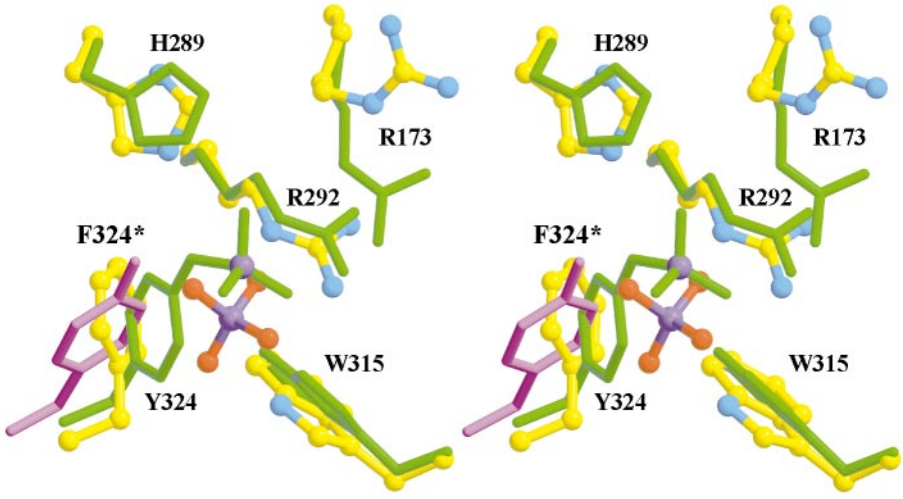

Figure 7. Stereo comparison of CRY and 1CRX B subunit active sites. The conserved catalytic residues of 1CRX-B (green) were superimposed on those in the 2.05 Å structure (atom colors), using the transformation from superposition of the main-chain atoms of residues 173, 289, 292, 315 and 324 (rmsd 0.45 Å). The scissile phosphate atoms are indicated (purple). The rotation of the Phe324 side-chain (''F324*'') allowed the close approach of the 324 main chain to mimic the covalent intermediate. This approach is sterically blocked in other structures by the Tyr324 hydroxyl. The Tyr324 side-chain of 3CRX-B after active site superposition (rmsd 1.3 Å) is also shown (magenta).

nium head group 4.1 Å, reflecting its inability to maintain the phosphate contact in the uncleaved state.

Cre binds, cleaves and catalyzes the resolution of YJ substrates

The CRY structure suggested that Cre could react with YJ substrates, in analogy to its HJ resolution activity.[14] Moreover, the Int enzymes FLP recombinase[45] and phage λ integrase[46] promote YJ resolution. The proposed mechanism,[45] shown in Figure 8(a), is analogous to resolution of a Holliday junction (Figure 1(c), lower path). Transesterification of the scissile phosphate to Tyr324 in the first step would generate a covalent complex. A subsequent reaction on an adjacent strand generates a free 5′ end to be transferred and religated to the tyrosine-bound phosphate group, yielding a Lox duplex and a covalent half-site complex[45] as products. To determine if Cre binds and reacts with YJs, an oligonucleotide substrate was designed. A

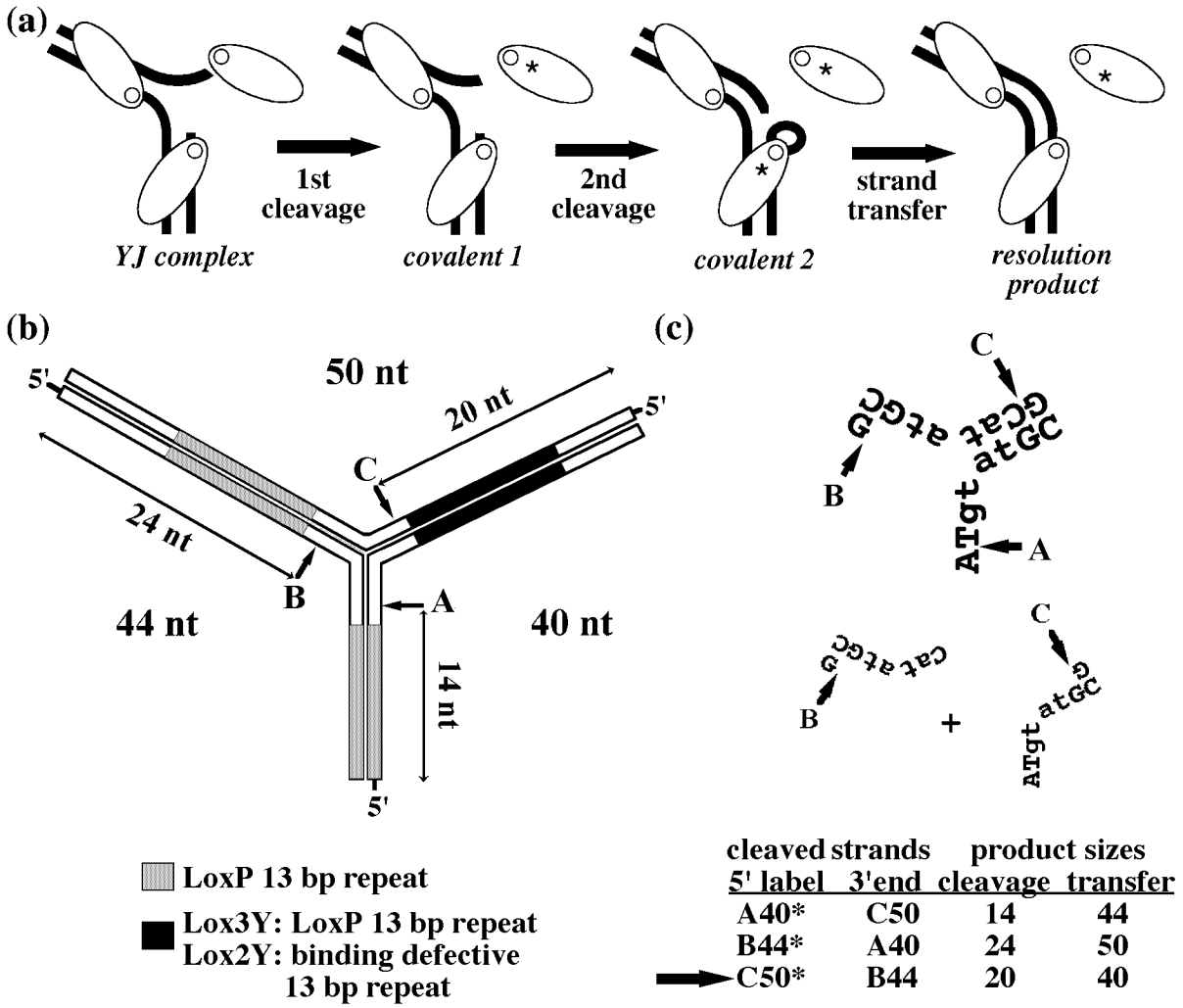

Figure 8. Mechanism of YJ resolution by Cre recombinase. (a) Proposed mechanism for Cre resolution of a Lox Y junction into a Lox duplex and covalently attached half-site, adapted from Lee *et al.*[45] After trimer assembly on the YJ (first panel), two cleavages on adjacent subunits occur (second and third panels). The covalently attached subunits are marked with an astersik (*). A single-strand transfer leads to a duplex product, but because of the odd number of cleavage sites, the remaining covalent intermediate cannot be resolved. (b) YJ substrates consist of three oligonucleotides, each of which spans the three-way junction; A40, B44, and C50 for Lox3Y; or 40 × , B44, and 50× for Lox2Y (See Materials and Methods). In Lox2Y, the C arm contains a binding defective 13 bp repeat. The lengths of the complete oligonucleotides are given along with those of the 5′-labelled cleavage products corresponding to segments 5′ of the cut site (arrows). (c) Substrate junction region sequence showing the cleavage sites on the A, B, and C strands (arrows) and scissile bases (upper case). Only the products of cleavage at B and C have no mismatches in the spacer region and are shown. The transferred strand from B cleavage is highlighted in grey. The other two products would contain at least two mismatches. The table below lists the sizes of the expected products arising from each combination of cleavage strands (''5′ label'') and the 3′ strand donor (''3′ end'') that is resolved by them. The arrow indicates the products of the favored BC cleavage reaction.

substrate with three arms (Lox3Y) contained 13 bp repeats and unique extensions in each arm (Figure 8(b)). These extensions stabilized the Y configuration over alternative duplexes and stem-loops that could be formed by the palindromic strands.[45] Distinct 8 bp spacer sequences were incorporated for perfect pairing in the junction (Figure 8(c)). This pairing scheme also allowed for only one set of stable resolution products. Of the three possible pairs of cleavages, only those arising from B44 and C50 cleavage do not create mismatches. In all experiments, one oligonucleotide was labeled and mixed with a fivefold excess of unlabeled complementary arms to maximize formation of the YJ.[45] The results were refractory to the amount of unlabeled DNA above the fivefold excess.

Cre binding to Lox3Y was probed in an electophoretic mobility shift assay (EMSA) using non-denaturing TBE-polyacrylamide gels (Figure 9(a)). Titration of Lox3Y with 0-2 μM Cre*Y324F at room temperature led to a single shifted band, which accumulated with increasing Cre concentration and a C_{50} of 0.7 μM (Figure 9(c)). A similar band was observed in titrations with wild-type Cre* (Figure 9(b)) but lower concentrations were required to achieve the same degree of shift, $C_{50} \sim 0.3$-0.4 μM, (Figure 9(c)) suggesting that Cre* binds more tightly than Cre*-Y324F. The additional smeared band above the common shifted one may indicate a transient covalent intermediate. Formation of insoluble Cre aggregates precluded using higher protein concentrations (>3 μM) to achieve complete saturation (note the counts in the gel wells, Figure 9(a) and (b)). Identical results were obtained when the other oligonucleotides were labeled (data not shown). All three oligonucleotides are required to generate bands corresponding to both the putative three-way junction and the shifted complex. These bands are not observed in mixtures containing any combination of two oligos (Figure 9(a)), indicating that the shifted species contains a YJ. The C_{50} values provide only a qualitative comparison with LoxP data given the differences in binding conditions and the difficulty of extracting equilibrium binding constants with a ninefold excess of competing unlabeled oligonucleotides present. These unlabeled palindromic DNAs can form stem-loops that are competitive Cre binding substrates.

The lack of intermediate bands containing one or two Cre molecules suggested that formation of a stable Cre-Lox3Y complex required binding to all three arms. Previous EMSA studies with LoxP show two shifted species believed to indicate one or two bound Cre molecules,[20,63] but these experiments were done under conditions using lower salt or at 4 °C. To test the correspondence of the shifted species to a trimeric complex, we generated a YJ substrate in which four mutations were introduced into one of the 13 bp repeats to prevent Cre binding,[20] Lox2Y. Although YJs formed under the same conditions, no Cre binding to Lox2Y was detected for these complexes at Cre concentrations up to 3 μM (Figure 9(b)). This result strongly implies that stable Cre-YJ complex formation requires occupancy of all three arms, and that the mobility shifts are attributable to specific binding rather than non-specific protein-DNA interactions.

To test the reactivity of the Cre-Lox3Y complex, reactions were probed for covalent intermediates. Three separate reactions were performed with each strand ^{32}P-labeled at the 5′ end to determine the strand origin of products (Figure 10). The 3′ covalent intermediates were detected as protein-shifted bands when electrophoresed through urea/SDS-containing gels (Figure 10(a)). In all three reactions, covalent attachment was observed. About 7 % of the label was incorporated into the covalent intermediate from A cleavage reactions, and ~1-2 % for B and C reactions. The putative ~40 nt strand transfer product was observed in the C50* reaction, accounting for 3 % of the incorporated radioactivity (labeled ''T''). The mobility of this band is consistent with the thermodynamically favored strand-transfer product arising from cleavage of C50 and B44 (Figure 8(c) and table therein).

To determine the cleavage site positions, the reactions were deproteinized with proteinase K and electrophoresed through a denaturing urea-containing gel (Figure 10(b)). All three reactions generated Cre-dependent labeled bands, labeled ''C'', but with the expected relative mobilities of hydrolytic cleavage products (see table, Figure 8(c)). In addition, each reaction also contained slightly higher molecular mass products, apparently 4-5 nt longer, that we have tentatively assigned as Cre-derived peptide-DNA complexes, labeled ''CP''. The presence of hydrolysis products in the proteinase K-digested reactions may reflect instability of the tyrosine ester under those conditions and was observed by Hoess & Abremski.[15] The A40 cleavage gave the highest yield, approximately 7 % of the input DNA, while B44 and C50 complexes accumulated to 2 %. Again a ~40 nt product, 5 %, was observed in the C50* reaction (''T'' in Figure 10(b)). In both analyses, the amount of covalent intermediate was in good agreement with the ''steady-state'' amount observed in Cre-LoxP reactions.[15,64,65] The low yield of strand transfer product suggests that the reaction does not facilely proceed past the first step, in part because Lox3Y is not an optimal substrate (see Discussion). The concentration dependence of cleavage and strand-transfer activity corresponded reasonably to the amount of Cre-induced Lox3Y shifting in the EMSA titration experiment. Contaminating activities are unlikely sources of cleavage given the lack of endonuclease activity in the Cre preparation, the absence of magnesium and presence of EDTA in the reaction, and that no such bands are observed when Cre*Y324F is utilized (data not shown).

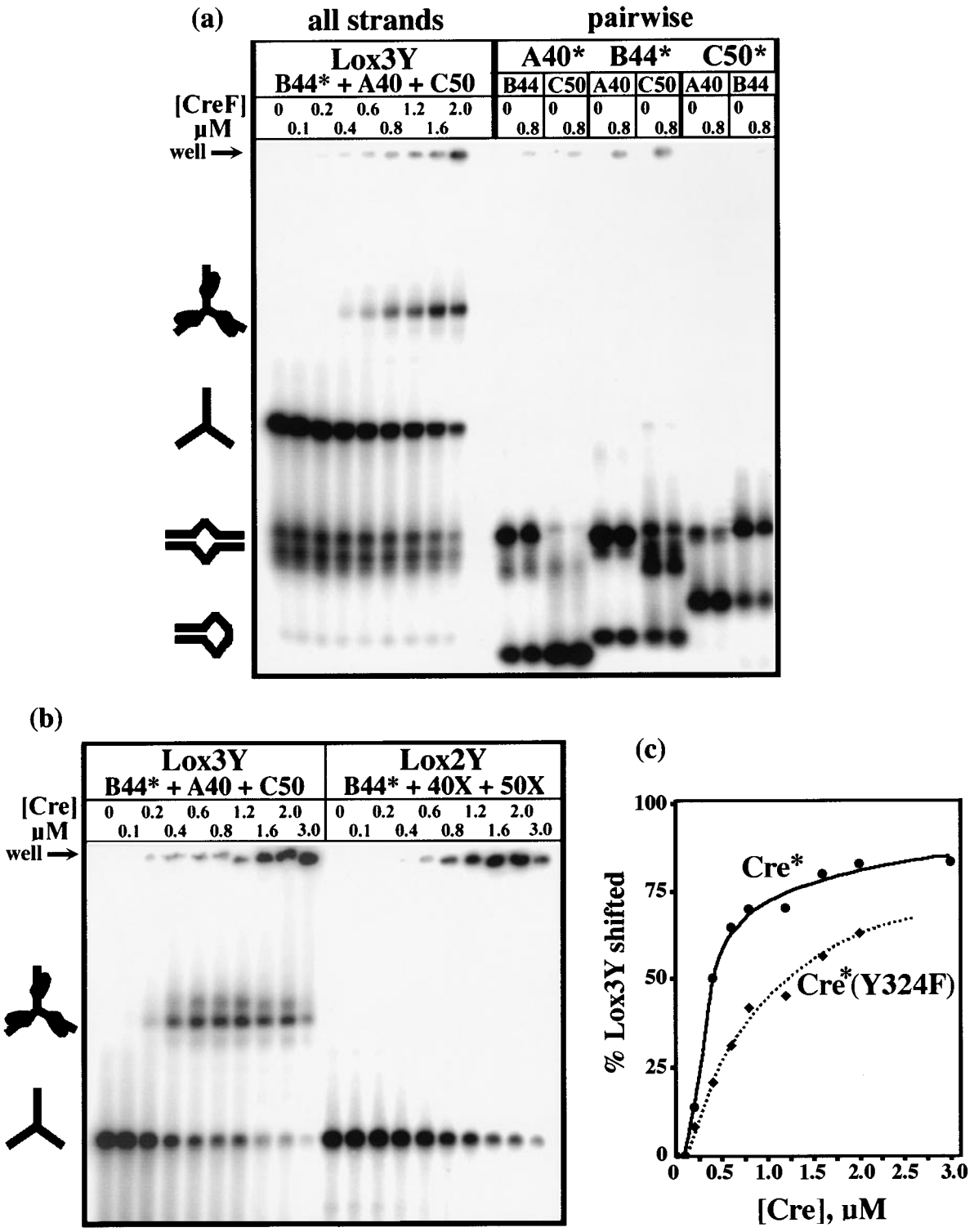

Figure 9. EMSA of Cre binding to Lox Y junctions. Cre (0-3 μM) was incubated with 50 nM ^{32}P-labeled YJ substrates at 21 °C, electrophoresed through non-denaturing 8% (w/v) polyacrylamide gels, and analyzed as described in Materials and Methods. (a) Titration of complete Lox3Y using labeled B44 (B44*) and A40 and C50 with 0-2 μM inactive Cre*Y324F (left nine lanes), led to a Cre-dependent shifted DNA species band. The top-most bands were in the wells of the gel, suggesting aggregation. In control experiments (right 12 lanes), each pair-wise combination of A40, B44 and C50 was incubated with 0 or of 0.8 μM Cre*Y324F. There were no bands with the same mobilities as the putative YJ or Cre*Y234F-YJ complex. (b) Titration of Lox3Y, A40, B44* and C50, (left ten lanes) and Lox2Y, 40X, B44*, and 50X, with 0-3 μM active Cre* (right ten lanes). In Lox2Y, one 13 bp repeat was substituted to prevent Cre binding. While Lox2Y YJ formed, no shifted species were detected. However, DNA was still trapped in the wells at high protein concentrations suggesting non-specific aggregation. (c) Quantification of Cre-YJ binding using phosphorimaging. Plot of percentage YJ shifted *versus* concentration of Cre* (continuous line) or Cre*Y324F (broken line). The percentage YJ shifted was taken as the number of counts in the shifted bands divided by the sum of the counts in free YJ and shifted bands. The curves were hand drawn.

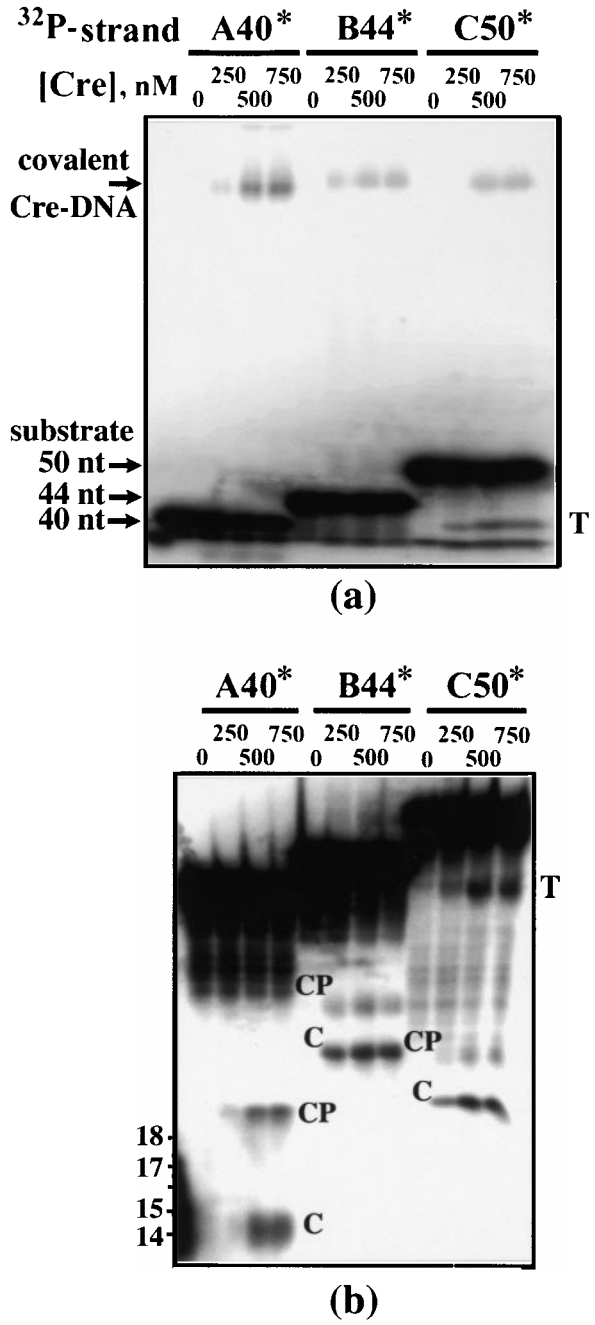

Figure 10. Reactivity of Cre* towards Lox3Y. End-labeled oligonucleotides A40*, B44*, or C50* were assembled into Lox3Y with excess cold partners and reacted with 250, 500, and 750 nM Cre* for four hours at 21 °C, quenched with SDS and analyzed as described in Materials and Methods. (a) Visualization of Cre* covalent attachment to each Lox3Y strand by SDS-PAGE. The upper shifted bands correspond to the covalent intermediates. The size of putative strand-transfer product from ligation of strands B44 and C50 (marked ''T'') was close to the predicted 40 nt long. (b) Size analysis of DNA products. Samples were deproteinized and analyzed using PAGE (15% (w/v) acrylamide) in 8 M urea. Expected product sizes are given in Figure 8(c). All reactions contained a putative cleavage product (''C'') presumably arising from hydrolysis of the covalent intermediate and a longer product tentatively assigned as the cleavage product attached to a peptide (''CP''). The expected strand transfer product ''T'' was again observed in the C50* reaction. The results of two completely independent experiments were identical (data not shown). While a saturated gel is shown for clarity, a second sub-saturating exposure was used for quantification.

Discussion

The principle of quasi-equivalence was first proposed by Caspar & Klug[53] to reconcile the packing of chemically identical capsid subunits into non-equivalent positions in the coats of icosahedral viruses. In addition to the application to viruses, quasi-equivalence was more recently invoked to describe the packing of dodecahedral and cubic assemblages in different crystal forms of pyruvate dehydrogenase,[66] and different packing modes in various phage T4 lysozyme space groups.[67] In these situations, subunits form oligomers with different symmetries[54] because interfaces adapt to the necessarily different packing environments.[55] This adaptation was initially envisioned to require deformation of the protein's internal structure to maintain the same interactions in different contexts.[53] More often, alternative packing arrangements are achieved through a combination of common and unique interactions at the quasi-equivalent interfaces.[54–56] In viruses, the ''switch'' between the different symmetry forms is typically accompanied by order changes in polypeptide segments that generate the unique contacts.[54] Such switches may be a general feature of protein-protein interactions in which different partners share a common contact surface.[68]

Here, these signatures are illustrated in a 3-fold trimeric Cre-Lox synapse that is quasi-equivalent to the pseudo-4-fold tetramer central to the site-specific recombination mechanism. The modular unit is a Cre monomer-Lox half-site complex. Analogous to quasi-equivalent packing of viral capsid proteins, the Cre-Lox trimer utilized a combination of novel interactions and those existent in the CRX structures to maintain contacts. The change in oligomerization number was accompanied by order changes of two polypeptide segments that unfold or refold in the trimer, reminiscent of the viral ''switches''. The two DNA junctions could be considered quasi-equivalent because of the similarities of the entwined DNAs. This structural quasi-equivalence was reflected in function, since Cre is active against both HJ and YJ substrates. That this is an apparently synthetic situation suggests that the signatures of quasi-equivalence may be general responses to variations in oligomeric packing rather than being necessarily evolutionarily selected.

In the CRY complex, the overall contact surface is less extensive than in the tetramers. The largest interface, interface I, was retained, suggesting that it is the primary protein-protein interaction driving

complex assembly. Its plasticity is likely functionally important and is observed to a more limited extent in the different CRX structures and subunits (compare A and B subunits, Figure 4(a)). A seeming by-product is the ability to form oligomeric species of varied multiplicities. In CRY, hinge flexibility appeared key to the adaptation to different subunit orientations. This idea is further illustrated by the structure of the catalytic domain of homologous HP1 integrase.[61] Although the hinge and C-terminal helix are similar in length and contacts with the C-terminal domain, a 2-fold HP1 Int dimer is formed and the hinge makes a correspondingly different angle. From entropic considerations, the smallest feasible oligomer would be expected to predominate. However, analogous dimer formation by Cre is likely blocked by anticipated steric clashes between residues that form the trimeric interface.

Interfaces II and III could not form in the trimer and their component interaction surfaces became reordered or disordered. The dispositions of these un-utilized contacts lend a view of their states in unbound Cre. In contrast to the complete disorder of the 198-208 loop, helix A rearranged to interact loosely with helix B of its own subunit. This shift suggests that helix A switches positions in response to the proximity of helix E from the adjacent subunit. This switch in helix pairing presumably reduces the possibility of non-specific protein-protein interactions in the absence of the cognate binding surface, and so specifically stabilizes only the particular geometries which enforce that proximity.

The absent tetramer contacts are partially compensated by the new three-way interface which contributes nearly 40 % of the trimeric contact area (Figure 5(b)). The good fit and the fractional amount of buried hydrophobic surface are consistent with oligomeric protein-protein contacts in general and less comparable with crystal contacts.[69–71] Whether this interface is gratuitous or contributes in some way to Cre function is yet unknown. Mutations in this region, residues 317, 318 and particularly 319, relax the sequence-specificity of both binding and recombination.[72] The most effective mutation, Asn319Leu, might be expected to stabilize the trimer interface by exchanging a polar contact for a hydrophobic one. Speculatively, this interface may play a transient role in tetramer assembly by stabilizing an intermediate inter-subunit geometry that closely matches that of the trimer. This may also be the case for the HP1 integrase dimeric interface.

CRY offers the first crystallographic views of a three-way DNA junction, albeit in the context of bound protein. As predicted by computational[43] and experimental work,[73,74] the open three-way junction was accommodated with little disruption of the *B*-DNA arms. The turn of the junction was created by a simple kink in the phosphodiester backbone. This backbone geometry was nearly identical with that of the ''non-crossover'' strand of the Cre-Lox Holliday complex (3CRX) indicating that junctions may have limited conformational options to effect a bend. Although the duplexes joined at an angle near 120°, the orientation of the DNA curvature yielded a pyramidal structure with a smaller ~107° average angle. The rotation of the Lox half-sites and the pyramidal nature of the YJ permitted the formation of interface I, which would have been too greatly separated if the YJ were flat. A much more pronounced pyramidal structure was favored by free 3H YJs.[42–44,75]

As suggested from the structure, Cre bound Lox-derived YJs, formed covalent intermediates, and performed strand transfer in the thermodynamically expected direction. However, Lox3Y, and perhaps YJs in general, are sub-optimal substrates for both binding and catalysis. This behavior is in contrast to FLP recombinase, which is much more active towards FRT-derived YJs, but more resembles the low activity of λ integrase.[46]

Stable binding of Cre to Lox3Y required occupancy by all three molecules, since no intermediate species were observed in the EMSA experiments and disrupting binding to one arm completely blocks trimer assembly (Figure 9(b)). Although the trimeric complex formed at sub-micromolar concentrations, this is several hundred-fold higher than the affinity of Cre for LoxP,[20] indicating that the YJ is a poorer ligand. The reason for reduced binding was not apparent in the CRY structure, as there is no indication of geometric or steric strain. However, model building indicated that substantial steric clashes with DNA occur if the inter-arm angle is less than 90°, somewhat greater than the 60-80° angle observed in free YJs.[44] Perhaps the cost of flattening the YJ to allow Cre binding contributes negatively. However, a second, less populated distribution of flatter YJs was also observed by AFM[44] indicating that this geometry is, at least, accessible.

That CRY most resembled the B subunit indicated that ''active'' Cre conformations, particularly the covalently bound one, are attainable in the trimer. This similarity to 1CRX-B in particular may be a function of rotation of the substituted Phe324 side-chain which allowed helix M to closely mimic the position in the covalent intermediate. Helix M and the active site were disordered in one precleavage intermediate structure[19] that also contained a Phe324 substitution (5CRX). One possible explanation is that the mutation allowed this segment to fluctuate between the precleavage and covalently bound positions. Consistent with this ability to achieve the ''cleaving'' conformation in the YJ-bound context, Cre readily formed covalent intermediates at levels comparable to those observed with LoxP.[64,65] Cre does not form covalent complexes with single half-sites in the context of a larger duplex[15,20] implying that oligomerization is required for activity. The cleavage activity of CRY supports this idea but also suggests that the precise position of the C-terminal hinge is not crucial. Apparently, formation of interface I is sufficient to

allow cleavage, either by ordering the C-terminal domain or restricting the degrees of freedom of helix M and Tyr324. Clearly, additional factors such as the conformation of the 198-208 loop modulate the activity differences between A and B subunits.

The thermodynamically favored strand transfer product is formed, but at low levels. The low efficiency is not surprising because of several suboptimal aspects of the structure and substrate. First, the conformational switches that coordinate cleavage order and position in the tetramer may not occur in the trimer. The situation is made even less favorable because of the clear preference for forming covalent intermediates at the left arm ApT base-step. This is in contrast to earlier published suggestions that the right arm GpC base-step was cleaved preferentially.[15] Unfortunately, the required cleavages in Lox3Y are at these less reactive sequences (Figure 8(b)). This bias was recently documented,[65] but subsequent to our substrate design. If the active sites are independent and, given the observed levels of 1-2 %, 1 in 10^3-10^4 molecules would have the appropriate configuration of covalent intermediates at any one time. We are presently investigating whether the substrate can be improved through re-design. Second, diffusion of the cleaved strands may be hindered in CRY because the strand-transfer channel is substantially contracted and the new trimeric interface sequesters the scissile nucleotides in the active site. Third, the loop containing essential Lys201[4,19] is not visible in CRY. The active site is geometrically accessible, but lacking the loop interaction with the adjacent subunit, Lys201 may not be not be held rigidly enough to maintain high occupancy for this contact. Finally, given the perfect symmetry of CRY, the trimer may disfavor the required asymmetric intermediates for YJ resolution[45] (Figure 8(a)). If the YJ-bound form is the most energetically favorable intermediate in this pathway, pre-assembled Cre trimers could potentially promote the reverse reaction of synthesizing YJs from a duplex and a half-site. This reaction is performed by λ integrase.[76]

CRY structure provides the structural bases for recognition of three-way junctions by Int family site-specific recombinases. That three recombinase monomers can minimally form an active complex has been suggested,[45,77] although more recently in the context of the tetramer.[78,79] Cox and co-workers also characterized the assembly of covalent FLP-DNA complexes containing three FRT half-sites.[77] Jayaram first suggested that recombinase trimers were the active species in FLP-mediated YJ resolution.[45] However, the oligomeric state of the protein species was not demonstrated in these cases. Our structural data furnish strong rationales for the idea that Int class recombinases trimerize quasi-equivalently to recognize such substrates (Figure 3(b), iii), rather than binding them as a tetramer with one empty DNA site (Figure 3(b), ii). The obtuse DNA branch angle of the tetramer is similar enough to that in the trimer to preserve the cyclic arrangement of the major quaternary interactions (Figure 2, top view). The FLP-FRT structure[23] suggests a commonality in YJ recognition, as two arms of the synapsed FRT DNA overlay the CRY arms similarly to 1CRX (Figure 6(b)). The two major quaternary contacts in the FLP complex are mediated by helices connected to flexible loops suggesting that they could be tolerant of the greater distance between subunits required for trimer formation. This greater flexibility and the larger strand-transfer channel expected in a FLP trimer, based on our modeling studies, gives a rationale for the higher activity of FLP towards YJs substrates. In contrast, the p53 tetramer seems unlikely to form trimers, but the flexibility of the connection between the DNA-binding and tetramerization domains[80] make it likely that YJs are bound with one DNA binding domain empty (v, Figure 3(b)). Similarly, based on steric considerations, recombinase-promoted resolution of HJ substrates containing three or less binding-competent arms were likely promoted by the tetrameric form.[79,81]

While analogous trimer formation by the Holliday junction cleaving enzymes RuvC, phage T4 endonuclease VII and T7 endonuclease I in YJ recognition cannot be ruled out, the crystal structures are 2-fold dimers that seem unlikely to trimerize.[82–84] These enzymes appear either to rearrange from their crystallographic dimers or deform the HJs significantly to bind DNA in both active sites. Hydroxyl radical cleavage experiments suggest that T4 EndoVII recognizes the coaxially stacked Holliday junction,[85] but DNA-binding surfaces appear to recognize DNA bent approximately 120°,[82] smaller than the ~140° angle between the arms of the coaxially stacked HJ.[36,38] The observed 120° angle of the CRY YJ supports the model proposed by Bhattacharyya *et al*[52] in which the nuclease dimers bind to adjacent arms (v, Figure 3(b)). In contrast, RuvC and yeast EndoX3 induce a more planar unstacked form[86,87] analogous the CRX synapse DNA, with its aforementioned similarities to the CRY YJ. However, our unsuccessful attempts to dock the CRY YJ to any the nuclease structures suggest that these enzymes have marked flexibility in their dimer contacts or are able to substantially deform the YJ substrate.

In summary, we have demonstrated that Cre recombinase site-specifically binds, cleaves, and resolves the three-way DNA substrates, by forming a quasi-equivalent oligomer to the HJ-reactive form. However, the origin of this cross-reactivity, shared with other HJ-recognizing proteins, remains speculative. The simplest hypothesis is that flexibility of subunit interfaces allows these proteins to recognize and adapt to common structural determinants present in both substrates as we have suggested here. The similar dispositions of the DNA branches in the substrates seem the most likely point of commonality, with finer differences accommodated by the flexibility of quaternary con-

tacts. The unanswered question is whether this flexibility in recognition is mechanistically required to account for changes in oligomer geometry that occur during recombination, or whether YJ-recognition is a ''built-in'' feature indicative of a yet undetermined or vestigial biological function.

Materials and Methods

Protein production and purification

The Cre recombinase gene amplified using PCR from the P1 lysogen BM25.8 (Novagen) with oligonucleotides (5′-desMet, TCCAATTTACTGACCGTACACCAA; 3′TAA, cgaattcatTAATCGCCATCTTCCAGCAGGC; non-P1 nucleotides are in lower case). The phosphorylated blunt-ended fragment was ligated into *Nco*I-digested and filled pET28b(+) (Novagen) to give pET28b-Cre. Authentic Cre recombinase was purified from induced lysates of BL21(::DE3, pET28b-Cre) by ammonium sulfate fractionation, selective precipitation in low salt, followed by cation exchange and gel filtration chromatographies (E.P.B., unpublished results). The mutation Tyr324-Phe was introduced using the method of Kunkel[88] with the mutagenic oligonucleotide GGTTACGGATAaAGTTCATaACAATATTTAC (mutations in lower case). Protein expressed from BL21(DE3) cells containing the mutant plasmid, pET28bCre-Y324F, was purified in an identical manner to the wild-type. To facilitate purification, six histidine residues were added to the N terminus of the wild-type and Tyr324Phe mutant using inverse PCR with the 5′desMet primer above, a primer containing the His-tag and vector sequences ATGATGATGATGATGATGGTGCATATGTATATCTCCTTCTTAAAGTTAAAC), and pET28bCre as template. Purification steps were carried out at 4 °C and all solutions were buffered by 20 mM Tris-HCl at pH 7.8-7.9 and contained 0.05 % (w/v) sodium azide. All Cre proteins are poorly soluble at below 300 mM NaCl at pH 7.8 (50-100 μg/ml). However, they are soluble and monomeric to 100 mg/ml in 700 mM NaCl, or to at least 250 μM in the presence of stoichiometric Lox half-site DNA. Six liters of BL21(DE3) cells harboring the expression plasmids, pET28b-His6Cre and pET28b-His6Cre-Y324F, were grown at 30 °C in 2 × LB broth containing 50 μg/ml kanamycin to a density of 1.0-1.4 A_{600}, induced with 1 mM IPTG, and grown for three hours at 30 °C. Chilled cells were lysed using a French press in buffer 1 (0.7 M NaCl) containing 5 mM imidazole and 1 mM PMSF, and clarified by centrifugation (39,000 ***g***, 30 minutes). Nucleic acids were removed by addition of 4 % (w/v) streptomycin sulfate and centrifugation. The lysate was dialyzed against buffer 1, applied to a 15 ml His-Bind column (1-2 ml/minute, Novagen), washed with three volumes of buffer 1 with 0.1 % (v/v) Triton X-100 and then ten to 15 volumes of buffer 1. Cre recombinase was eluted with a gradient of imidazole (5 mM-100 mM) in buffer 1. The purest eluates were dialyzed against buffer 1 containing 10 mM EDTA. The dialysate was diluted with 4.5 volumes of 20 mM Tris-HCl (pH 7.8), 200 mM NaCl, 2 mM DTT, and immediately applied to a 5 ml S-Sepharose column (Pharmacia, 5 ml, 4-6 ml/minute). The column was washed with 20 mM Tris-HCl (pH 7.8), 250 mM NaCl, 2 mM DTT, and Cre was eluted with a minimal volume of buffer 2 (buffer 1 with 2 mM DTT and 0.5 mM EDTA). The pooled fractions were adjusted to 1 M NaCl, concentrated to 20-100 mg/ml using a Centricon-10 microconcentrator (Millipore), dialyzed against buffer 2, filtered, and stored at 4 °C. Protein concentrations were determined using an extinction coefficient of 1.21 mg^{-1} or 49,000 M^{-1}. The yield was 5-20 mg of purified protein per liter of cells. Cre* had similar activities to wild-type using an integration assay *in vitro* and both His-tagged and untagged proteins behaved identically in crystallization experiments. All preparations were free of measurable endo- and exonuclease activities using supercoiled and linear DNA substrates, respectively at 1 mg/ml Cre.

Crystallization and data xollection

Half-site duplexes based on the LoxA site (Figure 1(b)) from Guo *et al.*[17] were designed to yield Lox sequences 32-40 bp in length containing either blunt outside ends, or 1-4 nt external 5′ extensions composed of alternating A and T residues. These duplexes were screened for crystallization with Cre*Y324F using vapor diffusion under 12 different conditions that routinely give Cre/Lox complex crystals (E.P.B., unpublished results). Both the 32-mer and 35-mer (identical with LoxA) versions yielded the crystal form described here. The 35-mer complex was chosen for comparison with the CRX structures and for its greater diffraction.

The oligonucleotide pair (B strand, TATAACTTCGTATAGC; C strand, ATATGCTATACGAAGTTAT, 1 μmol scale, RPC purified, DNA Express) were annealed by mixing at room temperature, cooled to 4 °C, and then mixed with Cre (35-45 mg/ml), diluted to 150 μM duplex and 75 μM protein, and allowed to stand for ten minutes. The 3 μl of protein-DNA solution was mixed in 1:1 (v/v) ratio with well solution and inverted over a 500 μl reservoir containing 4 % (w/v) PEG8 K, 125 mM $LiSO_4$, 20 mM Na-Pipes (pH 6.1), 60 mM NaCl at 20-22 °C. Crystals appeared by seven days, reaching maximum size (400 μm × 400 μm × 400 μm) within two weeks. The crystals were cubic, space group *I*23, $a = 161$ Å with one Cre*Y324/LoxA per asymmetric unit. For iodinated heavy-atom derivatives, iodine-1 and iodine-2,[17] 5-iododeoxyuridine bases (dU^I) replaced selected thymidine residues in the oligonucleotides: iodine-1 (dU^I A dU^I AACTTCGTA dU^I AGC), iodine-2 (ATATGC dU^I A dU^I ACGAAGTTAT). The mercury derivative was obtained by soaking a crystal in 1 mM thiomersal plus mother liquor for ten minutes. For the 2.05 Å resolution crystals, growth conditions contained 9 % (v/v) MPD.

All data were collected using crystals that were flash-cooled to 100 K. Prior to placing in the cold stream, crystals were rinsed briefly with reservoir buffer to which 20-25 % (v/v) MPD had been added. The Native 1 and heavy-atom derivative data sets were collected on an RAXIS IV using Cu-Kα radiation. The Native 2 data set was collected at SSRL beamline 9-1 (special thanks to Andrew Fisher, U.C. Davis). Diffraction data were indexed and processed with XDSIPLAY and DENZO and scaled using SCALEPACK.[89] Data collection statistics are given in Table 1.

X-ray structure solution and refinement

Candidate heavy-atom sites were determined using SOLVE[90] and verified by inspection of Patterson maps and cross-phased Fourier difference maps.[91] Phases and statistics were determined using MLPHARE.[92] A map at 2.8 Å was generated using 3.0 Å MLPHARE multiple

isomorphous replacement and anomalous scattering phases which were extended to 2.8 Å using density modification and solvent flattening with DM.[92] Data collection and phasing statistics are summarized in Table 1. The initial model was traced from the 2.8 Å map assisted by O,[93] using a 3.5 Å unmodified map for comparison. The model was refined with TNT[94] using all of the data to 2.55 Å. and observed structure factors were scaled to the model using the scaling parameters B 0.0 KSOL 0.9 BSOL 150.0. The geometry restraints of Engh & Huber were used.[95] The atomic *B*-factors were restrained using a BCORRELS library[96] modified by us to contain information from high-resolution DNA structures. For verification, $F_o - F_c$ difference maps were generated from models in which segments of 30 residues were omitted prior to refinement.[97] Five percent of the data were set aside prior to refinement for the free *R* calculation. In the final round of model building and refinement, only the 5-2.55 Å data and no solvent model were used because of poor scaling of the low-resolution data. The final model contained a single Cre/half-Lox complex in the asymmetric unit, residues 21-198 and 209-343, the entire LoxA half-site DNA, and 114 water molecules. A second 2.05 Å resolution data was obtained from SSRL beamline 9-1 and used for further refinement and rebuilding. The second model contained residues 20-199 and 208-343, all of the DNA, and 310 solvent molecules and differed from 1DRG significantly in region of residues 21-34, which is somewhat poorly defined. The final refinement statistics are given in Table 1. Nine percent of the protein or DNA atoms have high *B*-factors of 80-100 Å^2 but are included in the model because there is corresponding electron density visible in omit maps. Superpositions, solvent-accessible surfaces with a 1.4 Å probe, and other parameters for model analysis were performed using EDPDB.[58] Figures were created with MOLSCRIPT[98] followed by rendering with Raster3D.[99]

DNA binding and cleavage assays with YJ substrates

YJ substrates were created by annealing three oligonucleotides, each of which contained Lox 13 bp repeats, unique external extensions, and spacer sequences to direct the formation of a unique three-way junction. The following oligonucleotides were synthesized (DNA Express, 1 μmol scale, RPC-purified); the 13 bp repeats are underlined, deviations from LoxP are in lower case, "/" represents the center of the junction. For the substrate with three Lox arms, Lox3Y: A40 (40-mer, 5′-ATAACTTCGTATAatgT/AtgcTATACGAAGTTATGTCAAT); B44 (44-mer, 5′-TTCTACATGGATAACTTCGTATAgcaT/AcatTATACGAAGTTAT); and C50, (50-mer, 5′-ATTGACATAACTTCGTATAgcaT/AtgcTATACGAAGTTATCCATGTAGAA); and for the substrate containing two Lox arms, Lox2Y: 40X (40-mer, 5′-ATAACTTCGTATAatgT/AtgcT**gc**AC**cg**AGTTATGTCAAT); B44, and 50X (5′-ATTGACATAACT**cg**GT**gc**AgcaT/AtgcTATACGAAGTTATCCATGTAGAA), changes from Lox3Y in bold.

In all experiments, one oligonucleotide was 5′-labeled with [γ^{32}P]ATP (3000 Ci/mmol, Amersham) using T4 polynucleotide kinase (NEB) in the supplied buffer followed by purification using a Quiagen Nucleotide Purification Kit. The labeled oligo, 50 nM, was annealed with 250 nM each of unlabeled partners and used without further treatment as these were not stable to gel purification. The majority (>90 %) of the labeled oligonucleotide was incorporated into the Y-junction. Polyacrylamide gels (20:1 (w/w) acrylamide/bis-acrylamide) contained and were run in Tris-borate-EDTA buffer (TBE) with or without 8 M urea, or using SDS-PAGE buffers[100] supplemented with 8 M urea. Gels without urea were dried directly, while those containing 8 M urea were rinsed with gel running buffer and then dried. An image plate (Fuji BAS-MS 2040) was exposed, scanned with an image plate reader (Molecular Dynamics Storm 860) and visualized and analyzed with ImageQuant (Molecular Dynamics).

EMSA titrations were performed as follows: Cre* or Cre*Y324 was mixed with pre-annealed YJ containing ~5 nM ^{32}P-labeled B44 mixed with 45 nM unlabeled B44 (B44*), and 250 nM each A40 or 40× and C50 or 50×) in 200 mM NaCl, 1 mM Na-EDTA, 10 mM Tris-HCl (pH 7.8), at 21 °C. After standing for one hour, one fourth volume of loading buffer, 30 % (w/v) Ficoll, 0.005 % (w/v) bromophenol blue, 1 mM Na-EDTA, and 10 mM Tris-HCl (pH 7.8), was added and the mixture was applied to a 6 % (w/v) polyacrylamide gel usingTBE buffer, electrophoresed, dried and quantified.

Covalent complex formation and strand transfer activity was assayed by incubating Cre* with pre-annealed labeled YJ substrate containing ~5 nM ^{32}P-labeled ^{32}P 5′ end and 45 nM unlabeled A40, B44, or C50 (A40*, B44*, or C50*) and 250 nM each unlabeled strand, for four hours at 21 °C in optimized Cre reaction buffer (300 mM LiOAc, Tris-acetate, 1 mM Na-EDTA, pH 8.3, 50 μg/ml BSA). The reaction was quenched with 2 % (w/v) SDS, and split into two portions. One portion was electrophoresed through a Laemmli SDS containing gel[100] (8 % (w/v) acrylamide, TBE and 8 M urea). The other portion was digested with proteinase K (0.5 mg/ml, one hour, 65 °C) and electrophoresed through an 8 M urea/TBE containing gel (15 % (w/v) acrylamide). The gels were processed and quantified as described above. The results of two completely independent experiments were identical. Product sizes were determined from band mobilities of the proteinase K-treated samples. The relationship of DNA size and gel mobility was obtained from linear regression of the log of the length *versus* the mobility of markers (14, 15, 16 (not indicated), 17 and 18 nt) and substrate strands (40, 44 and 50 nt).

Protein Data Bank accession numbers

The coordinates of the 2.55 Å model were deposited in the Protein Data Bank, accession number 1DRG. The 2.05 Å model was also deposited in the PDB, accession number 1F44. The abbreviations used for the reaction intermediates are their PDB accession numbers, e.g. the precleavage intermediate has the PDB accession number 4CRX.

Acknowledgments

The authors thank Professor Andrew Fisher for collecting the 1F44 data set at SSRL, Dr James Endrizzi (U.C. Berkeley) for many helpful discussions and synchrotron access at ALS, and Professors Matt Redinbo (U.N.C.-Chapel Hill) and Andy Fisher (U.C. Davis) for insights into quasi-equivalence. This work was funded, in part, by an American Cancer Society Institutional Research Grant (No. 36769) through U.C. Davis Cancer Center and U.C. Davis startup funds.

References

1. Sadowski, P. (1986). Site-specific recombinases: changing partners and doing the twist. *J. Bacteriol.* **165**, 341-347.
2. Sternberg, N. (1981). Bacteriophage P1 site-specific recombination. III. Strand exchange during recombination at lox sites. *J. Mol. Biol.* **150**, 603-608.
3. Esposito, D. & Scocca, J. J. (1997). The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucl. Acids Res.* **25**, 3605-3614.
4. Cheng, C., Kussie, P., Pavletich, N. & Shuman, S. (1998). Conservation of structure and mechanism between eukaryotic topoisomerase I and site-specific recombinases. *Cell,* **92**, 841-850.
5. Nunes-Duby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T. & Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucl. Acids Res.* **26**, 391-406.
6. Abremski, K. & Hoess, R. (1984). Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J. Biol. Chem.* **259**, 1509-1514.
7. Kilby, N. J., Snaith, M. R. & Murray, J. A. (1993). Site-specific recombinases: tools for genome engineering. *Trends Genet.* **9**, 413-421.
8. Metzger, D. & Feil, R. (1999). Engineering the mouse genome by site-specific recombination. *Curr. Opin. Biotechnol.* **10**, 470-476.
9. Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. *Genesis,* **26**, 99-109.
10. Sauer, B. & Henderson, N. (1990). Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. *New Biol.* **2**, 441-449.
11. Albert, H., Dale, E. C., Lee, E. & Ow, D. W. (1995). Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J.* **7**, 649-659.
12. Chambers, C. A. (1994). TKO'ed: lox, stock and barrel. *Bioessays,* **16**, 865-868.
13. Hoess, R. H., Ziese, M. & Sternberg, N. (1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc. Natl Acad. Sci. USA,* **79**, 3398-3402.
14. Hoess, R., Wierzbicki, A. & Abremski, K. (1987). Isolation and characterization of intermediates in site-specific recombination. *Proc. Natl Acad. Sci. USA,* **84**, 6840-6844.
15. Hoess, R. H. & Abremski, K. (1985). Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J. Mol. Biol.* **181**, 351-362.
16. Hoess, R. H., Wierzbicki, A. & Abremski, K. (1986). The role of the loxP spacer region in P1 site-specific recombination. *Nucl. Acids Res.* **14**, 2287-2300.
17. Guo, F., Gopaul, D. N. & van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature,* **389**, 40-46.
18. Gopaul, D. N., Guo, F. & Van Duyne, G. D. (1998). Structure of the Holliday junction intermediate in Cre-loxP site- specific recombination. *EMBO J.* **17**, 4175-4187.
19. Guo, F., Gopaul, D. N. & Van Duyne, G. D. (1999). Asymmetric DNA bending in the Cre-loxP site-specific recombination synapse. *Proc. Natl Acad. Sci. USA,* **96**, 7143-7148.
20. Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R. & Stewart, A. F. (1998). Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination. *J. Mol. Biol.* **284**, 363-384.
21. Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1992). Catalysis by site-specific recombinases. *Trends Genet.* **8**, 432-439.
22. Nunes-Duby, S. E., Azaro, M. A. & Landy, A. (1995). Swapping DNA strands and sensing homology without branch migration in lambda site-specific recombination. *Curr. Biol.* **5**, 139-148.
23. Chen, Y., Narendra, U., Iype, E. L., Cox, M. M. & Rice, A. P. (2000). Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol. Cell,* **6**, 885-897.
24. Yang, W. & Mizuuchi, K. (1997). Site-specific recombination in plane view. *Structure,* **5**, 1401-1406.
25. Jayaram, M. (1997). The cis-trans paradox of integrase. *Science,* **276**, 49-51.
26. Lee, J., Jayaram, M. & Grainge, I. (1999). Wild-type Flp recombinase cleaves DNA in trans. *EMBO J.* **18**, 784-791.
27. Seeman, N. C. & Kallenbach, N. R. (1994). DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53-86.
28. Altona, C., Pikkemaat, J. A. & Overmars, F. J. (1996). Three-way and four-way junctions in DNA: a conformational viewpoint. *Curr. Opin. Struct. Biol.* **6**, 305-316.
29. Lilley, D. M. (2000). Structures of helical junctions in nucleic acids. *Quart. Rev. Biophys.* **33**, 109-159.
30. DasGupta, C., Wu, A. M., Kahn, R., Cunningham, R. P. & Radding, C. M. (1981). Concerted strand exchange and formation of Holliday structures by E. coli RecA protein. *Cell,* **25**, 507-516.
31. Mosig, G. (1987). The essential role of recombination in phage T4 growth. *Annu. Rev. Genet.* **21**, 347-371.
32. Severini, A., Scraba, D. G. & Tyrrell, D. L. (1996). Branched structures in the intracellular DNA of herpes simplex virus type 1. *J. Virol.* **70**, 3169-3175.
33. Pearson, C. E. & Sinden, R. R. (1998). Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Curr. Opin. Struct. Biol.* **8**, 321-330.
34. Altona, C. (1996). Classification of nucleic acid junctions. *J. Mol. Biol.* **263**, 568-581.
35. Pikkemaat, J. A., van den Elst, H., van Boom, J. H. & Altona, C. (1994). NMR studies and conformational analysis of a DNA four-way junction formed in a linear synthetic oligonucleotide. *Biochemistry,* **33**, 14896-14907.
36. Ortiz-Lombardia, M., Gonzalez, A., Eritja, R., Aymami, J., Azorin, F. & Coll, M. (1999). Crystal structure of a DNA Holliday junction. *Nature Struct. Biol.* **6**, 913-917.
37. Nowakowski, J., Shim, P. J., Stout, C. D. & Joyce, G. F. (2000). Alternative conformations of a nucleic acid four-way junction. *J. Mol. Biol.* **300**, 93-102.
38. Eichman, B. F., Vargason, J. M., Mooers, B. H. & Ho, P. S. (2000). The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl Acad. Sci. USA,* **97**, 3971-3976.
39. Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H. & Morikawa, K. (2000). Crystal structure of the holliday junction DNA in complex with a single RuvA tetramer. *Proc. Natl Acad. Sci. USA,* **97**, 8257-8262.

40. Roe, S. M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I. R. & Pearl, L. H. (1998). Crystal structure of an octameric RuvA-Holliday junction complex. *Mol. Cell,* **2**, 361-372.
41. Duckett, D. R. & Lilley, D. M. (1990). The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking. *EMBO J.* **9**, 1659-1664.
42. Stuhmeier, F., Welch, J. B., Murchie, A. I., Lilley, D. M. & Clegg, R. M. (1997). Global structure of three-way DNA junctions with and without additional unpaired bases: a fluorescence resonance energy transfer analysis. *Biochemistry,* **36**, 13530-13538.
43. Shlyakhtenko, L. S., Rekesh, D., Lindsay, S. M., Kutyavin, I., Appella, E., Harrington, R. E. & Lyubchenko, Y. L. (1994). Structure of three-way DNA junctions. 1. Non-planar DNA geometry. *J. Biomol. Struct. Dynam.* **11**, 1175-1189.
44. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R., Gall, A. A. & Lyubchenko, Y. L. (2000). Structure and dynamics of three-way DNA junctions: atomic force microscopy studies. *Nucl. Acids Res.* **28**, 3472-3477.
45. Lee, J., Whang, I. & Jayaram, M. (1996). Assembly and orientation of Flp recombinase active sites on two, three and four-armed DNA substrates: implications for a recombination mechanism. *J. Mol. Biol.* **257**, 532-549.
46. Nunes-Duby, S. E., Yu, D. & Landy, A. (1997). Sensing homology at the strand-swapping step in lambda excisive recombination. *J. Mol. Biol.* **272**, 493-508.
47. Lee, S., Cavallo, L. & Griffith, J. (1997). Human p53 binds Holliday junctions strongly and facilitates their cleavage. *J. Biol. Chem.* **272**, 7532-7539.
48. Takahagi, M., Iwasaki, H. & Shinagawa, H. (1994). Structural requirements of substrate DNA for binding to and cleavage by RuvC, a Holliday junction resolvase. *J. Biol. Chem.* **269**, 15132-15139.
49. Jensch, F., Kosak, H., Seeman, N. C. & Kemper, B. (1989). Cruciform cutting endonucleases from Saccharomyces cerevisiae and phage T4 show conserved reactions with branched DNAs. *EMBO J.* **8**, 4325-4334.
50. Kupfer, C. & Kemper, B. (1996). Reactions of mitochondrial cruciform cutting endonuclease 1 (CCE1) of yeast *Saccharomyces cerevisiae* with branched DNAs *in vitro*. *Eur. J. Biochem.* **238**, 77-87.
51. Dickie, P., McFadden, G. & Morgan, A. R. (1987). The site-specific cleavage of synthetic Holliday junction analogs and related branched DNA structures by bacteriophage T7 endonuclease I. *J. Biol. Chem.* **262**, 14826-14836.
52. Bhattacharyya, A., Murchie, A. I., von Kitzing, E., Diekmann, S., Kemper, B. & Lilley, D. M. (1991). Model for the interaction of DNA junctions and resolving enzymes. *J. Mol. Biol.* **221**, 1191-1207.
53. Caspar, D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-4.
54. Johnson, J. E. (1996). Functional implications of protein-protein interactions in icosahedral viruses. *Proc. Natl Acad. Sci. USA,* **93**, 27-33.
55. Johnson, J. E. & Speir, J. A. (1997). Quasi-equivalent viruses: a paradigm for protein assemblies. *J. Mol. Biol.* **269**, 665-675.
56. Rossmann, M. G. (1984). Constraints on the assembly of spherical virus particles. *Virology,* **134**, 1-11.
57. Lavery, R. & Sklenar, H. (1988). The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dynam.* **6**, 63-91.
58. Zhang, X.-J. & Matthews, B. W. (1995). EDPDB: a multi-functional tool for protein structure analysis. *J. Appl. Crystallog.* **28**, 624-630.
59. Dickerson, R. E. (1998). DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucl. Acids Res.* **26**, 1906-1926.
60. Gopaul, D. N. & Duyne, G. D. (1999). Structure and mechanism in site-specific recombination. *Curr. Opin. Struct. Biol.* **9**, 14-20.
61. Hickman, A. B., Waninger, S., Scocca, J. J. & Dyda, F. (1997). Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell,* **89**, 227-237.
62. Krogh, B. O. & Shuman, S. (2000). Catalytic mechanism of DNA topoisomerase IB. *Mol. Cell,* **5**, 1035-1041.
63. Mack, A., Sauer, B., Abremski, K. & Hoess, R. (1992). Stoichiometry of the Cre recombinase bound to the lox recombining site. *Nucl. Acids Res.* **20**, 4451-4455.
64. Shaikh, A. C. & Sadowski, P. D. (2000). Chimeras of the Flp and Cre recombinases: tests of the mode of cleavage by Flp and Cre. *J. Mol. Biol.* **302**, 27-48.
65. Tribble, G., Ahn, Y. T., Lee, J., Dandekar, T. & Jayaram, M. (2000). DNA recognition, strand selectivity, and cleavage mode during integrase family site-specific recombination. *J. Biol. Chem.* **275**, 22255-22267.
66. Izard, T., Aevarsson, A., Allen, M. D., Westphal, A. H., Perham, R. N., de Kok, A. & Hol, W. G. (1999). Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl Acad. Sci. USA,* **96**, 1240-1245.
67. Zhang, X. J., Wozniak, J. A. & Matthews, B. W. (1995). Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. *J. Mol. Biol.* **250**, 527-552.
68. Harrison, S. C. (1996). Peptide-surface association: the case of PDZ and PTB domains. *Cell,* **86**, 341-343.
69. Carugo, O. & Argos, P. (1997). Protein-protein crystal-packing contacts. *Protein Sci.* **6**, 2261-2263.
70. Jones, S. & Thornton, J. M. (1995). Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **63**, 31-65.
71. Lijnzaad, P., Berendsen, H. J. & Argos, P. (1996). Hydrophobic patches on the surfaces of protein structures. *Proteins: Struct. Funct. Genet.* **25**, 389-397.
72. Hartung, M. & Kisters-Woike, B. (1998). Cre mutants with altered DNA binding properties. *J. Biol. Chem.* **273**, 22884-22891.
73. Zhong, M. & Kallenbach, N. R. (1993). Conformation and thermodynamics of DNA ''necks''. Models for three-arm branch formation in a duplex. *J. Mol. Biol.* **230**, 766-778.
74. Lu, M., Guo, Q. & Kallenbach, N. R. (1991). Effect of sequence on the structure of three-arm DNA junctions. *Biochemistry,* **30**, 5815-5820.
75. Fenley, M. O., Manning, G. S., Marky, N. L. & Olson, W. K. (1998). Excess counterion binding and ionic stability of kinked and branched DNA. *Biophys. Chem.* **74**, 135-152.
76. Nunes-Duby, S. E., Matsumoto, L. & Landy, A. (1989). Half-att site substrates reveal the homology

independence and minimal protein requirements for productive synapsis in lambda excisive recombination. *Cell,* **59**, 197-206.
77. Qian, X. H., Inman, R. B. & Cox, M. M. (1990). Protein-based asymmetry and protein-protein interactions in FLP recombinase-mediated site-specific recombination. *J. Biol. Chem.* **265**, 21779-21788.
78. Lee, J. & Jayaram, M. (1997). A tetramer of the Flp recombinase silences the trimers within it during resolution of a Holliday junction substrate. *Genes Dev.* **11**, 2438-2447.
79. Qian, X. H. & Cox, M. M. (1995). Asymmetry in active complexes of FLP recombinase. *Genes Dev.* **9**, 2053-2064.
80. Waterman, J. L., Shenk, J. L. & Halazonetis, T. D. (1995). The dihedral symmetry of the p53 tetramerization domain mandates a conformational switch upon DNA binding. *EMBO J.* **14**, 512-519.
81. Kho, S. H. & Landy, A. (1994). Dissecting the resolution reaction of lambda integrase using suicide Holliday junction substrates. *EMBO J.* **13**, 2714-2724.
82. Raaijmakers, H., Vix, O., Toro, I., Golz, S., Kemper, B. & Suck, D. (1999). X-ray structure of T4 endonuclease VII: a DNA junction resolvase with a novel fold and unusual domain-swapped dimer architecture. *EMBO J.* **18**, 1447-1458.
83. Ariyoshi, M., Vassylyev, D. G., Iwasaki, H., Nakamura, H., Shinagawa, H. & Morikawa, K. (1994). Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from *E. coli*. *Cell,* **78**, 1063-1072.
84. Hadden, J. M., Convery, M. A., Declais, A. C., Lilley, D. M. & Phillips, S. E. (2001). Crystal structure of the Holliday junction resolving enzyme T7 endonuclease I. *Nature Struct. Biol.* **8**, 62-67.
85. Parsons, C. A., Kemper, B. & West, S. C. (1990). Interaction of a four-way junction in DNA with T4 endonuclease VII. *J. Biol. Chem.* **265**, 9285-9289.
86. White, M. F. & Lilley, D. M. (1997). The resolving enzyme CCE1 of yeast opens the structure of the four-way DNA junction. *J. Mol. Biol.* **266**, 122-134.
87. Bennett, R. J. & West, S. C. (1995). RuvC protein resolves Holliday junctions via cleavage of the continuous (noncrossover) strands. *Proc. Natl Acad. Sci. USA,* **92**, 5635-5639.
88. Kunkel, T. A., Bebenek, K. & McClary, J. (1991). Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol.* **204**, 125-139.
89. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
90. Terwilliger, T. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D,* **55**, 849-861.
91. Furey, W. & Swaminathan, S. (1997). PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. *Methods Enzymol.* **277**, 000-.
92. Collaborative Computational Project Number 4 (1994). The CCP4 suite programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.
93. Jones, T. A., Zou, J., Cowan, S. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.
94. Tronrud, D. (1997). TNT refinement package. *Methods Enzymol.* **277**, 306-319.
95. Engh, R. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A,* **47**, 392-400.
96. Tronrud, D. (1996). Knowledge-based *B*-factor restraints for the refinement of proteins. *J. Appl. Crystallog.* **29**, 100-104.
97. Brunger, A. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and Applications. *Acta Crystallog. sect. D,* **49**, 24-36.
98. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
99. Merritt, E. A. & Murphy, M. E. P. (1994). Raster3d version 2.0 - a program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869-873.
100. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680-685.
101. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

Edited by K. Morikawa

(Received 23 May 2001; received in revised form 10 August 2001; accepted 14 August 2001)