BIOCHEMISTRY

 Volume 46, Number 49 December 11, 2007

Articles

DNA Recognition via Mutual-Induced Fit by the Core-Binding Domain of Bacteriophage λ Integrase†

Hari B. Kamadurai and Mark P. Foster*

Graduate Program in Biophysics and Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210

Received May 21, 2007; Revised Manuscript Received September 13, 2007

ABSTRACT: Bacteriophage λ integrase (λ-Int), a phage-encoded DNA recombinase, cleaves its substrate DNA to facilitate the formation and later resolution of a Holliday junction intermediate during recombination. The core-binding and catalytic domains of λ-Int constitute a bipartite enzyme that mediates site-specific DNA cleavage through their interactions with opposite sides of the recognition sequence. Despite minimal direct contact between the domains, the core-binding domain has been shown to facilitate site-specific DNA cleavage when provided in trans, indicating that it plays a role beyond enhancing binding affinity. Biophysical characterization of the core-binding domain and its interactions with DNA reveal that the domain is poorly structured in its free form and folds upon binding to DNA. Folding of the protein is accompanied by induced-fit structural changes in the DNA ligand. These data support a model by which the core-binding domain plays a catalytic role by reshaping the substrate DNA for effective cleavage by the catalytic domain.

Bacteriophage λ integrase (λ-Int), the prototypical tyrosine-dependent DNA recombinase, enables the lysogenic and lytic lifestyles of bacteriophage λ by catalyzing site-specific incorporation and excision of the viral genome into and out of its bacterial host genome (*1*, *2*). These recombination reactions are mediated at specific DNA sequences termed "core sites", each consisting of an inverted pair of semiconserved λ-Int-binding regions (core half-sites). During recombination, two full core sites are brought together through a tetrameric assembly of four λ-Int protomers, each of which binds to one core "half-site". This tetrameric assembly then coordinates two series of DNA cleavage, strand exchange, and religation reactions to form and later resolve a Holliday junction intermediate.

Specific core-site DNA recognition and cleavage are mediated by two of three domains of λ-Int: the core-binding (Int^{CB}) and catalytic (Int^{Cat}) domains (*3*, *4*). (The small N-terminal domain is important for binding to DNA "arm" sites far from the site of cleavage and plays a role in controlling directionality of recombination (*5*−*8*).) When bound to DNA, the Int^{CB} and Int^{Cat} domains, connected by an extended linker, adopt a bipartite DNA-binding structure that nearly encircles the DNA (Figure 1) (*5*, *9*). Although both domains play a role in DNA site recognition (*4*, *9*), Int^{Cat} is necessary and sufficient for specific core-site DNA recognition and cleavage (*3*) and contains the critical catalytic residues, including a conserved tyrosine (Y342) that cleaves the DNA through nucleophilic attack to form a stable phosphotyrosine linkage. Solution and crystallographic structural studies indicate that Int^{Cat} undergoes significant induced-fit conformational changes upon binding to its cognate DNA (*9*−*11*).

† This material is based upon work supported by the National Science Foundation under Grant No. MCB-0092962.

* To whom correspondence should be addressed. E-mail: foster.281@osu.ed.; Phone: (614) 292-1377. Fax: (614) 292-6773.

10.1021/bi700974t CCC: $37.00
Published on Web 11/15/2007

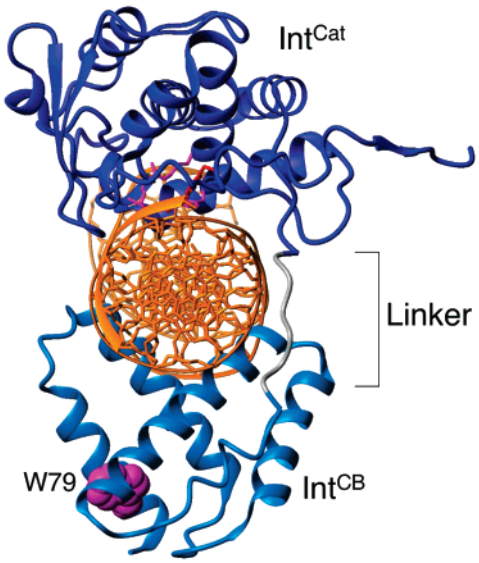

FIGURE 1: Crystal structure of Int^{CB+Cat} in a covalent complex with a DNA substrate (PDB accession code 1P7D). Int^{CB} and Int^{Cat} domains, and the linker between them, are indicated; DNA is shown as sticks. The two domains bind to opposite faces of the DNA with no significant direct contacts between them. The sole tryptophan in Int^{CB}, W79, is shown as spheres.

While an isolated construct of Int^{Cat} can specifically cleave a core half-site DNA substrate, cleavage activity is more pronounced with a protein construct containing both Int^{CB} and Int^{Cat}, here termed Int^{CB+Cat} (*3*, *4*). The higher activity of Int^{CB+Cat} compared to Int^{Cat} might reasonably be attributed to the bipartite DNA-binding architecture of Int^{CB+Cat} (Figure 1) that enables cooperative DNA binding via the "chelate effect" (*12*, *13*); the only base-specific contact between Int^{CB} and the DNA appears to be between the side chain of N99 and two conserved adenines in the core half-sites (*9*). If the enhancement of Int^{Cat} by Int^{CB} were adequately described by the chelate effect, one would expect that the cooperative effects would be lost if the two domains were separated. However, DNA cleavage assays with isolated Int^{Cat} and Int^{CB} constructs show that Int^{CB} is able to increase the cleavage activity of Int^{Cat} even when they are not expressed together as a single protein construct—that is, when provided in trans (*14*). This finding suggests that the role of Int^{CB} is not limited to anchoring Int^{Cat} near the site of cleavage, but also promotes DNA cleavage through some other mechanism. The observed boost in DNA cleavage in the presence of Int^{CB} does not appear to be due to direct Int^{CB}—Int^{Cat} interactions since crystallographic studies with the Int^{CB+Cat} construct do not reveal significant interdomain contacts (*9*), nor do NMR studies reveal any large perturbations in the conformation of Int^{Cat} due to the presence of Int^{CB} (*14*). Importantly, the stimulation of DNA cleavage by Int^{CB} can be selectively abolished through changes in the substrate DNA sequence without affecting the basal DNA cleavage activity by Int^{Cat}, suggesting a role for the substrate DNA in the Int^{CB}-promoted DNA cleavage activity (*14*).

To better understand the stimulatory effect of Int^{CB}, we studied the solution conformation of isolated Int^{CB} and its DNA-binding properties using NMR, CD, and fluorescence spectroscopies and isothermal titration calorimetry (ITC). We found that Int^{CB} is poorly structured on its own, but folds upon binding to DNA. Concomitant structural changes were observed in the interacting DNA ligand, indicating that Int^{CB} recognizes DNA through a mutual induced-fit conformational change. Indeed, in the crystal structure of the complex between Int^{CB+Cat} and DNA (*9*), the structure of the DNA at the Int^{CB}-binding site is altered from the canonical B-form: the DNA undergoes widening of the major groove at the site of helix insertion and corresponding narrowing of the minor groove. We propose that mutual induced-fit interaction provides the basis for the trans-stimulatory effect of Int^{CB} on DNA cleavage activity by promoting a DNA structure closely resembling the transition state of the cleavage reaction.

RESULTS

DNA-Induced Folding of Int^{CB}. The ^{15}N-edited HSQC spectrum of free Int^{CB} (Figure 2A) exhibits few well-resolved peaks, while also being marked by broad or less dispersed, intense peaks. The scarcity of peaks in the spectrum is indicative of severe resonance broadening due to exchange between multiple conformational states, suggesting that Int^{CB} is not a well-folded protein domain. Protein multimerization was not found to be responsible for the observed spectral features since equilibrium sedimentation studies using analytical ultracentrifugation were indicative of monomeric Int^{CB} in solution (*15*). These features of free Int^{CB} are not due to its isolation from Int^{Cat}, since solution (*14*) and crystallographic (*5*, *9*) studies reveal negligible contacts between the domains.

In contrast to the free protein, the HSQC spectrum of Int^{CB} recorded in the presence of a cognate DNA half-site exhibits a nearly complete set of well-resolved NMR signals (Figure 2B). The appearance of well-resolved and well-dispersed signals indicates that in the presence of DNA the backbone amide protons populate unique environments as would be expected in a well-folded protein, suggesting that Int^{CB} becomes better folded upon binding to DNA. Binding-coupled folding was not limited to interactions with the cognate DNA (a core half-site), as similar improvement in the spectrum was observed upon addition of a noncognate DNA ligand (*15*). These observations indicate that binding to DNA results in Int^{CB} folding.

The effect of DNA binding on the structure of Int^{CB} was investigated further using circular dichroism (CD) spectroscopy (Figure 3). The CD spectrum of free Int^{CB} exhibited signal minima at 208 and 220 nm, as expected for a helical protein (*12*), indicating the presence of significant secondary structure. The amplitude of the CD signal at 222 nm increased ~15% in the presence of equimolar amounts of DNA (i.e., compare the "Sum" and "Complex" spectra), indicating that DNA binding increases the secondary structure content of Int^{CB} (Figure 3A). Because DNA does absorb in this region, and changes in protein interhelical packing can further affect the signal in this region of the CD spectrum, we are hesitant to quantitate the change in protein secondary structure; however, we conclude that most of the protein secondary structure is already present in the free protein, but is stabilized by DNA binding.

The thermodynamic stability of the secondary structural elements in Int^{CB} was further probed by monitoring their thermal denaturation using the CD signal at 222 nm (Figure 3B). The melting curve of free Int^{CB} was very broad and

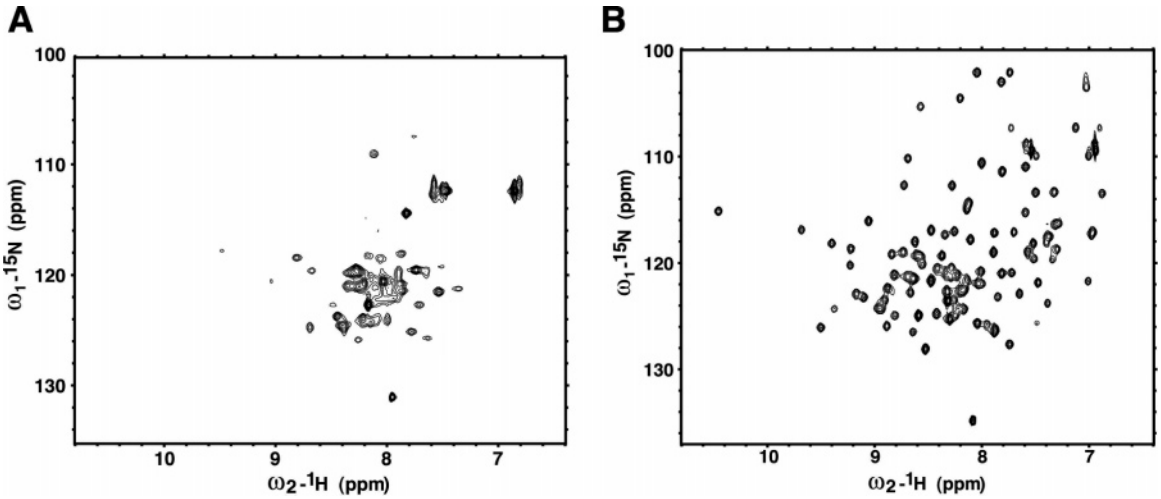

FIGURE 2: NMR spectra reveal DNA-induced folding of Int^{CB}. (A) ^{15}N-edited HSQC spectrum of the free Int^{CB}. Poor signal and limited dispersion in the spectrum are indicative of a poorly folded protein. (B) ^{15}N-edited HSQC spectrum of Int^{CB} in a 1:1 complex with the cognate DNA. An increase in the number of signals and their improved dispersion are consistent with binding-coupled folding of Int^{CB}.

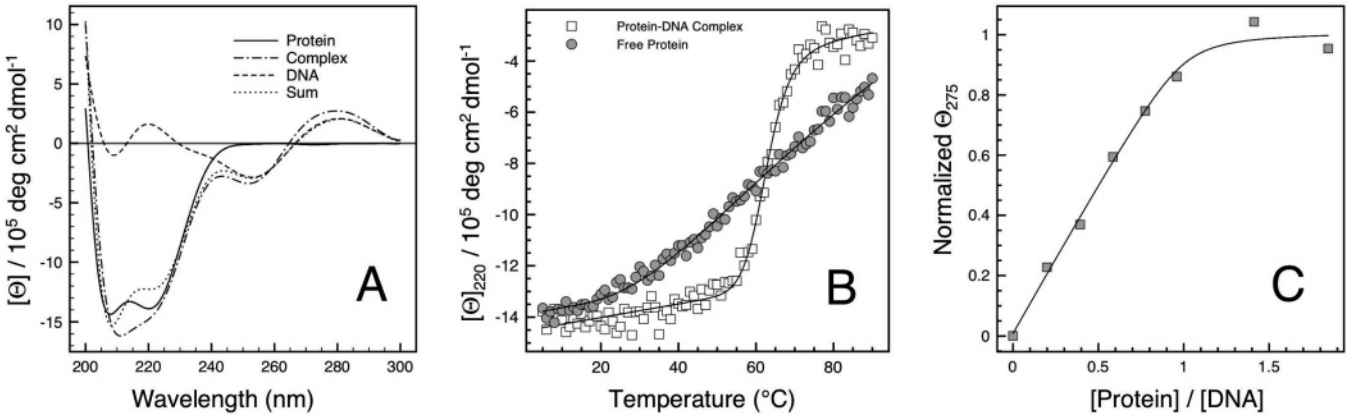

FIGURE 3: CD spectra reveal induced-fit structural changes in both protein and DNA. (A) Far-UV CD spectra of free Int^{CB}, DNA-bound Int^{CB}, and the free DNA. The sum of the CD signals of the free species is also shown. Free Int^{CB} has an α-helical signature that increases further with the addition of the DNA. Data were smoothed using a ninth-order linear spline function (http://plot.micw.eu). (B) Thermal denaturation of free and DNA-bound Int^{CB} monitored by molar ellipticity at 220 nm shows thermal unfolding of free Int^{CB} to be highly noncooperative, while unfolding of the Int^{CB}−DNA complex is highly cooperative. Solid lines represent the best fit to the integrated van't Hoff equation (see the Materials and Methods). (C) Normalized ellipticity of the DNA at a wavelength of 275 nm (which is sensitive to the helical structure), in the absence and presence of increasing concentrations of Int^{CB}. The line represents the best fit of the data to a binding quadratic (see the Materials and Methods).

incomplete over the sampled temperature range, indicative of noncooperative unfolding. In contrast, the Int^{CB}−DNA complex exhibits a classical sigmoidal feature during thermal denaturation (Figure 3B), indicative of a well-folded structure that unfolds cooperatively (*12*, *16*).

DNA Recognition via Mutual Induced Fit. Induced-fit interactions can involve conformational changes in both interacting partners. To investigate whether the DNA ligand undergoes similar structural alterations upon binding and folding of Int^{CB}, CD spectroscopy was also used to monitor the changes in the cognate DNA structure upon binding the protein. The near-UV region of the CD spectrum of DNA (∼240−300 nm) is sensitive to the helical geometry (*17*, *18*) and thus serves as a convenient means of monitoring changes in the DNA structure. Titration of DNA with Int^{CB} resulted in an incremental increase in intensity in the near-UV region of the spectrum, indicating structural changes in the DNA upon binding to Int^{CB} (Figure 3C). The spectral changes monitored using the intensity at 275 nm as a function of protein concentration saturate near equimolar amounts of Int^{CB}, indicating little additional structural changes in the DNA beyond 1:1 stoichiometry.

One-dimensional NMR spectra provided further evidence for protein-induced structural changes in the DNA structure. The imino protons of each DNA base pair report on the geometry and environment along the helical axis of the DNA, and their resonances are well resolved from regions of the spectra containing most protein signals (*19*). The imino proton spectrum of the free cognate DNA (15 base pairs) consisted of 12 resolved peaks, whose positions were sensitive to the presence of Int^{CB} (Figure 4), indicating mutual induced-fit structural changes in the DNA.

Equilibrium Binding Studies of the Int^{CB}−DNA Interaction. To quantitate the thermodynamics of the Int^{CB}−DNA interaction, a binding curve was constructed by plotting the changes in λ_{max} for the fluorescence emission of W79 with respect to the total DNA concentration (Figure 5). The wavelength for the fluorescence emission maximum (λ_{max}) of the unique tryptophan residue in Int^{CB} (W79) (see Figure 1) is shifted toward shorter wavelength (from 330 to 320

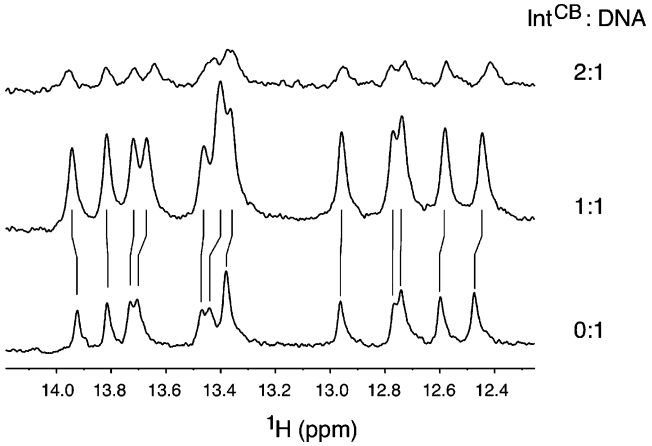

FIGURE 4: The effect of Int^{CB} on DNA imino proton NMR spectra reveals protein-induced structural changes. Imino proton spectra were recorded for the cognate DNA alone (bottom) and when Int^{CB} was present at 1:1 (middle) and 2:1 (top) molar equiv.

nm) in the presence of DNA (Figure 5A), indicative of burial in the interior of the protein and sequestration from solvent (*12*). Unexpectedly, this fluorescence titration, which monitors structural changes in the protein, reached saturation well before equimolar concentrations were reached, and the data could not be fit to a binding model that assumes 1:1 stoichiometry (Figure 5B). The data could be fit, however, to a model allowing for multiple Int^{CB}-binding sites on the DNA, yielding an apparent dissociation constant K_D of 34 $\pm$ 10 nM and 2 $\pm$ 0.03 binding sites (n) for the 15 bp cognate DNA. Titration experiments with a noncognate DNA of similar size yielded a similar value for n, but weaker binding ($K_D = 80 \pm 20$ nM); since the signal being monitored reports on protein folding, these data indicate Int^{CB} folding does not require a cognate DNA sequence. Experiments with a longer 29 bp noncognate DNA revealed similar weak binding ($K_D = 89 \pm 13$ nM), but n increased to ~6.5, reflecting the capacity of Int^{CB} for nonspecific DNA binding (Figure 5B).

To further characterize the thermodynamics of Int^{CB}−DNA interactions, we used ITC to monitor Int^{CB} binding to the 15 bp cognate DNA ligand. A binding isotherm was obtained by titrating DNA into a thermostatic cell containing protein (Figure 6A). This isotherm exhibited one clear binding event, but with an apparent stoichiometry of 0.5, and exhibited baseline disturbances as the titration progressed beyond 0.5 molar equiv of ligand, ruling out detailed thermodynamic analysis on those traces. On the other hand, when the titration was reversed (i.e., titrating protein into DNA), the binding isotherm clearly resolved into two phases, revealing two thermodynamically separable binding events, one approximately 30 times stronger than the other ($K_{D,1} = 1 \pm 0.3$ μM, $K_{D,2} = 34 \pm 2$ μM) (Figure 6B). The integrated values of heat exchanged for each injection were fit to a model with two independent binding sites (Origin 7.0). The stronger interaction was dominated by favorable entropy ($-T\Delta S$), while the weaker interaction was dominated by favorable enthalpy (ΔH) (see Figure 6B); favorable binding entropy is generally associated with release of solvent and concomitant burial of hydrophobic surface area that accompanies formation of specific complexes, while binding enthalpy is dominated by electrostatic interactions that can be nonspecific (*20*, *21*). Thus, these experiments demonstrate that Int^{CB} is capable of binding the 15 bp half-site DNA with a 2:1 stoichiometry when the DNA concentration is limiting, but with a 1:1 stoichiometry when the DNA is in excess. Importantly, the data establish thermodynamic parameters for the two classes of binding sites, which we classify as arising from "specific" (tighter) and "nonspecific" (weaker) protein−DNA interactions.

NMR Studies of 2:1 and 1:1 Int^{CB}−DNA Complexes. The trans-activating effect of Int^{CB} on Int^{Cat} could arise, in principle, from favorable protein−protein interactions that lead to assembly of tetrameric complexes, or through alterations of the DNA structure (*14*). Although the DNA substrate used in these experiments is too small to allow assembly of the type of higher order protein−DNA complexes observed in tetrameric recombigenic complexes (*5*), we nevertheless sought to investigate whether the 2:1 Int^{CB}−DNA complexes that form when DNA is limiting reflect favorable protein−protein interactions or nonspecific protein−DNA binding. To this end, backbone resonance assignments were obtained for Int^{CB} in a 1:1 complex with DNA (excess DNA), and a series of 2D ^{1}H−^{15}N correlation spectra of Int^{CB} were recorded through a DNA titration spanning Int^{CB}−DNA ratios from 2:1 through <1:1 (Figure 7A). Perturbations in the chemical shifts of backbone amide resonances of Int^{CB} were then mapped to the Int^{CB}−DNA structure to ascertain whether the interaction interfaces could be identified (*22*, *23*). Backbone resonance assignments could be obtained for 85% of the protein in a 1:1 complex with DNA using standard triple-resonance NMR methods (*24*).

A single set of resonances was observed for Int^{CB} when present at 2:1 and 1:1 protein−DNA stoichiometries. At the protein concentrations used for these experiments, the 1:1 complex is nearly quantitatively formed at equimolar DNA concentrations, while the 2:1 complex is ~70% formed with 0.5 molar equiv of DNA. Observation of one set of protein signals indicates that in the 2:1 complex the protein species are in fast exchange on the NMR time scale; for there to be a single set of resonances for a complex in slow exchange, the complex would have to be effectively C_2 symmetric, which is not allowed by the nonpalindromic DNA sequence. In interpreting the titration data, we reasoned that if the 2:1 protein−DNA complex were to result from protein−protein interactions in fast exchange between protein-bound and DNA-bound states, we would expect that, as the DNA concentration is increased and each protein binds a single DNA oligomer, significant chemical shift perturbations should map to the transiently formed protein−protein interface. On the other hand, if the 2:1 complex were to reflect specific and nonspecific binding to DNA, as DNA is titrated in shift perturbations would be expected to map to the protein−DNA interface, reflecting a shift from half of the protein being nonspecifically bound to all of the protein being bound to a cognate site.

Residues whose environments shifted during the titration from 0.5 to 1 equiv of DNA were detected through changes in peak positions in ^{15}N-edited HSQC spectra (Figure 7A). Backbone amide shift perturbations were quite modest, indicating that the average environment of the backbone is quite similar for the two species of binding sites. Resonance assignments enabled the identification of residues that experience measurable chemical shift perturbations during DNA titration (Figure 7B). Importantly, the largest shift perturbations map to residues in the DNA-binding face of

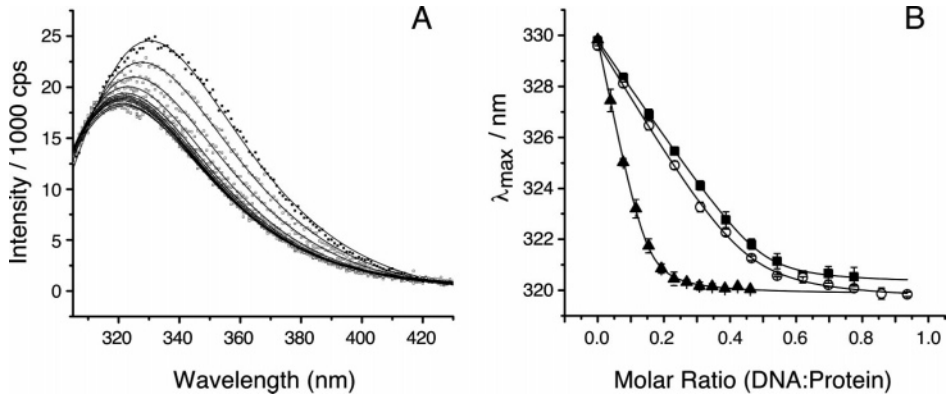

FIGURE 5: DNA binding monitored by tryptophan fluorescence. (A) Fluorescence emission spectra of W79 in free Int^{CB} (filled symbols) and in the presence of increasing concentrations of the cognate DNA ligand (empty symbols). The solid lines represent the best fit to the "Extreme" function (Origin 7.0). Addition of DNA results in a shift in λ_{max} toward shorter wavelengths. (B) λ_{max} for tryptophan fluorescence emission versus molar ratio (DNA:protein) for cognate and noncognate DNA ligands: filled squares, 15 bp cognate DNA; open circles, 15 bp noncognate DNA; filled triangles, 29 bp noncognate DNA. Error bars were obtained from triplicate measurements. Dissociation constants and stoichiometries from the data fitting were 34 ± 10 nM and 2 ± 0.03 for the 15 bp cognate DNA, 80 ± 20 nM and 2.2 ± 0.05 for the 15 bp noncognate DNA, and 89 ± 13 nM and 6.6 ± 0.12 for the 29 bp noncognate DNA, respectively.

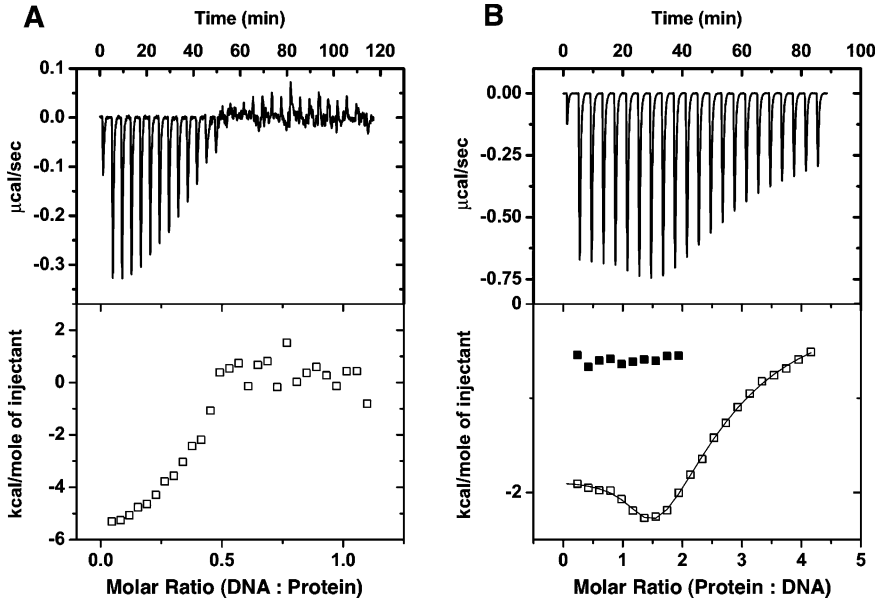

FIGURE 6: Isothermal titration calorimetry of DNA binding to Int^{CB}. (A) Top, heat exchanged during each injection of cognate DNA into a solution of Int^{CB}; bottom, the corresponding integrated heat of binding for each injection. This "reverse" titration of DNA into protein reaches the baseline with approximately 0.5 molar equiv of DNA, after which subsequent titrations result in very noisy data with little net heat exchanged. (B) Top, heat exchanged during each injection of Int^{CB} into a solution containing cognate DNA; bottom, corresponding thermogram with empty squares representing heat of binding and filled squares representing heat of dilution. The solid lines represent the fit to a two-site binding model (Origin 7.0); the resulting thermodynamic values were $n_1 = 1.26 \pm 0.04$, $K_{D,1} = 1 \pm 0.3$ μM, $\Delta H_1 = -1.9 \pm 0.1$ kcal mol^{-1}, and $-T\Delta S_1 = -6.3 \pm 0.2$ kcal mol^{-1}, and $n_2 = 1.1 \pm 0.2$, $K_{D,2} = 34 \pm 2$ μM, $\Delta H_2 = -5.7 \pm 1$ kcal mol^{-1}, and $-T\Delta S = -0.4 \pm 1$ kcal mol^{-1}, where n_i is the stoichiometry for each of the binding sites.

Int^{CB}, while residues that would be involved in protein−protein interactions in higher order recombigenic assemblies (*5*) were largely unperturbed. We conclude that the 2:1 Int^{CB}−DNA complexes reflect nonspecific protein−DNA interactions, not higher order specific protein−protein interactions.

DISCUSSION

Bipartite substrate binding is a common feature among enzymes that recognize and hydrolyze DNA. In the family of site-specific tyrosine recombinases, a domain containing the catalytically important residues is connected to an auxiliary domain via a flexible linker (*2*, *25*). Together, these linked domains wrap around the DNA substrate and cooperate in site recognition and cleavage. Although tyrosine recombinases exhibit a high degree of sequence specificity, crystallographic studies of their protein−DNA complexes (*5*, *26*−*28*) indicate their specificity is not well explained by direct contacts between amino acid side chains and DNA bases.

Core-site DNA cleavage studies using Int^{Cat} have shown the domain is capable of sequence-specific cleavage of cognate DNA sequences (*3*, *4*, *14*). However, cleavage experiments in which Int^{CB} is provided in trans show that Int^{CB} plays a role in controlling site recognition and cleavage by selectively promoting cleavage at cognate sites, without significantly affecting affinity for those sites (*14*). How is this specific enhancement achieved? Crystal structures of Int−DNA complexes (*5*, *9*) show that Int^{CB} makes a single base-specific contact to DNA, between the side chain of N99

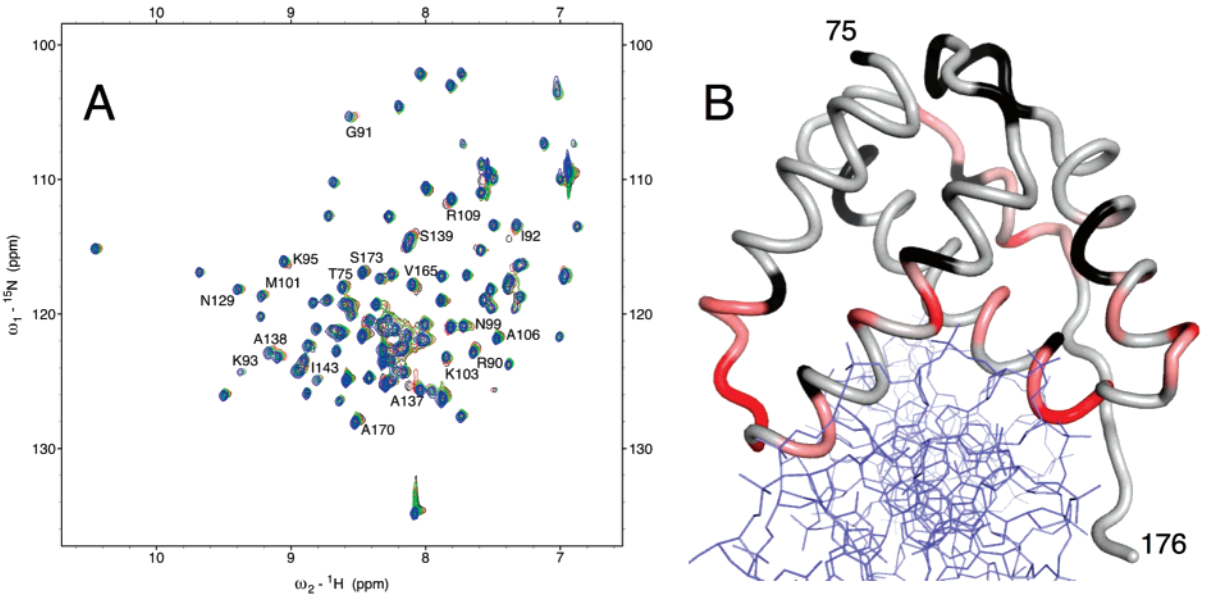

FIGURE 7: Perturbations in the ^{1}N−^{15}H resonance shifts of Int^{CB} during DNA titration from 2:1 to 1:1 molar equiv (protein:DNA) map to the protein−DNA interface. (A) ^{15}N-edited HSQC spectra were recorded with 0.5 (black), 0.67 (magenta), 0.8 (red), 1.0 (green) and 1.2 molar equiv of DNA (blue) with respect to Int^{CB}. The peaks that are significantly perturbed due to DNA addition are labeled. (B) Weighted averaged shift perturbations (*36*) mapped to a backbone cartoon of Int^{CB} (coordinates from 1Z19 (PDB accession code) (*5*)), with a linear color ramp from gray ($\Delta\delta \leq 0.02$ ppm) to red ($\Delta\delta \geq 0.06$ ppm). Unassigned residues are black, and DNA is blue.

and conserved adenines in the core half-site sequences. The present studies indicate relatively little sequence discrimination by the isolated domain, as binding-coupled folding was observed in the presence of cognate or noncognate DNA, with only modest differences in affinity (Figure 5). The observation that Int^{CB}−DNA site discrimination is small in comparison to its effect on catalysis (*14*) allows the proposal that induced shape recognition is an important determinant for its function.

It is clear that induced-fit conformational changes play important roles in directing and regulating the function of λ-Int (*1*, *5*, *7*, *9*−*11*, *29*−*31*). Previous studies illuminated induced-fit structural changes in the λ-Int catalytic domain and in DNA substrates upon forming protein−DNA complexes (*5*, *9*, *10*). Now, NMR, CD, and fluorescence spectra each indicate that the accessory domain Int^{CB} is not well folded on its own, but becomes structured upon binding to DNA. These Int^{CB}-induced changes in the CD and NMR spectra of the DNA are also consistent with deformation in the DNA structure at the Int^{CB}-binding site, as observed in the crystal structure of Int^{CB+Cat} bound to a half-site DNA substrate (*9*). In this crystal structure, Int^{CB} and Int^{Cat} are found to recognize the cognate DNA through many phosphate backbone contacts and few base contacts, suggesting that the shape of the DNA could play a role in DNA recognition. Thus, binding-induced structural changes in the protein are mirrored by induced-fit structural changes in the DNA substrate, as evidenced by NMR and CD spectra, and deviations from standard B-form helicity observed in the crystal structures of λ-Int−DNA complexes.

Binding-coupled protein folding has been shown to be particularly important in sequence recognition by DNA-binding proteins (*20*, *21*). Importantly, there is growing evidence that enzymes use induced fit as a mechanism for discriminating between cognate and noncognate DNA sites (*32*). It is evident from the ITC measurements (Figure 6) that the thermodynamics of specific binding of Int^{CB} to a cognate DNA differs from its nonspecific DNA-binding behavior. One possible implication of this difference is that the structure induced by binding to noncognate sequences is insufficiently similar to the "precleavage" structure to lead to DNA cleavage. Essentially, we suggest that Int^{CB} ensures selectivity not merely by promoting high-affinity binding of Int^{Cat} to a cognate site, but also by testing the fitness of the candidate substrate at the catalytic step for its ability to adopt a structure closely resembling the transition state.

Given the potential of site-specific DNA recombinases for applications in genetic engineering and gene therapy (*33*), it is essential that we understand the mechanisms that control site recognition and cleavage. The finding of mutual induced-fit binding between Int^{CB} and its DNA target has important implications for the design of modified recombinases. Int^{CB}−DNA interactions involve a mutual induced fit where conformational changes of protein and DNA accompany binding. Int^{CB}-induced structural changes in the DNA could act as stimulants for Int^{Cat}, leading to core-site DNA cleavage.

MATERIALS AND METHODS

The Int^{CB} construct used for these studies consisted of residues 75−176 of full-length λ-Int (NCBI accession P03700). The expression construct was generated by PCR-based deletion of residues 62−74 from an expression construct described previously (*14*) and is derived from the pRT3 vector of Tirumalai et al. (*4*). QuikChange site-directed mutagenesis (Strategene, Inc.) was performed in two consecutive steps, first using as primers a DNA duplex deleting residues 62−69 and encoding an *Aat*II restriction site (italicized) for screening positive clones, 5′-d(GGA GAT ATA CAT **ATG** *ACG TCT* AAT TCC GTT ACG TTA CAT TCA TGG), and then a second set of oligos that delete residues 70−74 and remove the *Aat*II site, 5′-d(GGA GAT ATA CAT **ATG** ACG TTA CAT TCA TGG CTT GAT CGC). Protein expression and purification were carried out as described previously (*14*). The pGRO-ArgUW helper

plasmid (*37*) was cotransformed with the λ-Int domain-expressing plasmids to overcome the codon bias problem in *Escherichia coli*. To prepare ^{13}C,^{15}N-labeled Int^{CB} samples, transformed *E. coli* BL21(DE3) cells were grown in rich media prior to induction, and exchanged into M9 minimal medium with ^{13}C-labeled glucose as the sole carbon source and ^{15}N-labeled ammonium chloride as the sole nitrogen source. The concentration of the protein was determined using its predicted extinction coefficient at 280 nm (9411 M^{-1} cm^{-1}).

Three duplex DNA ligands were used for these experiments: a 15 bp cognate ligand corresponding to the C′ half-site: (5′-dGCT CAA G**TT** AGT ACG-3′) (*1*), a 15 bp noncognate ligand based on the C′ site, with the conserved thymines altered to G (5′-dGCT CAA G**GG** AGT ACG-3′), and a 29 bp noncognate ligand (*4*) (5′-dTGA ACA GGT CAC TAT CAG TCA AAA TAA AA-3′). Underlined residues are invariant core sequences, and those in bold represent the differences between the ligands. Oligonucleotide duplexes were formed by heating a mixture of complementary single strands (IDT DNA, Inc.) to 95 °C and then cooling on ice for 15 min in 25 mM sodium phosphate (pH 7.2) and 0.02% sodium azide (buffer A) or in 25 mM Tris (pH 8.5) and 10 mM NaCl (buffer B). The duplexes were purified to remove the excess single strands using a HiTrap Q-column with a 0−2 M salt gradient. The purified samples were desalted by gel filtration (PD-10 columns, GE Healthcare) and concentrated by lyophilization. The lyophilized powder was resuspended and reannealed in either buffer A or buffer B. The concentrations of the DNA ligands were determined on the basis of their absorbance at 260 nm.

NMR data were recorded at 25 °C in buffer A containing 10% D_2O. NMR experiments were performed either on a Bruker DRX 600 spectrometer fitted with a cryoprobe or on a DMX 600 instrument with a room-temperature probe, each equipped with pulsed field gradient capabilities. Data were recorded at 25 °C and referenced to internal TSP. One-dimensional WATERGATE proton spectra were recorded on the free DNA (175 μM) and Int^{CB}−DNA complexes with molar ratios of 1:1 and 2:1 (350 μM Int^{CB}). DNA titration was monitored by recording ^{15}N-edited correlation spectra with a ^{13}C,^{15}N-labeled Int^{CB} sample using the 15 bp cognate DNA ligand at the following protein:DNA ratios: 2:1, 1.4:1, 1.3:1, 1:1, and 0.6:1. Precipitation was observed when the DNA was present at less than 0.5 molar equiv, but the solution clarified as the DNA concentration exceeded that concentration; no further spectral changes were observed beyond 1 molar equiv of DNA. Backbone resonance assignments were obtained from analysis of HNCO, HNCA, CBCA(CO)NH, and HNCACB spectra (*24*) recorded on the protein−DNA complex in a solution with excess DNA to ensure the observed protein signal derives from the 1:1 complex. The data were processed using NMRPipe (*34*) and analyzed using CARA (*35*) (http://www.nmr.ch/) to obtain resonance assignments through backbone sequential connectivity. Assignments were deposited at the BMRB (http://www.bmrb.wisc.edu/) under ID 15213.

Circular dichroism spectra were recorded with an Aviv 62A DS CD spectrophotometer at 25 °C in buffer A. Data were collected in a 1 mm quartz cuvette from 300 to 200 nm with a 1 nm step and 5 s averaging at each step. Protein and DNA concentrations were 25 μM; at these concentrations, given a K_D of 1 μM, the 1:1 complex is 80% formed. To compare the magnitude of the CD spectral differences between the free components and their complex, the Sum spectrum was obtained by adding the (concentration-adjusted) spectra of the free components; the Sum spectrum was compared with the Complex spectrum obtained from the sample containing 25 μM protein and DNA. If there is no change in structure upon binding, the Sum and Complex spectra will be identical; the magnitude of the difference is informative about the amount of structural change. In thermal denaturation experiments, signals from the free protein and the protein−DNA complex were monitored at 220 nm from 5 to 90 °C with an equilibration time of 2 min every 1 °C; denaturation data for the Int^{CB}−DNA complex were subjected to van't Hoff analysis by fitting to (*16*)

$$\theta_{220} = A - \frac{A}{1 + e^{-\Delta H/R(1/T - 1/T_m)}} + \text{baseline} + mT$$

where A is the amplitude of the signal change and m is a temperature-dependent linear offset, revealing an enthalpy of unfolding $\Delta H = -70$ kcal mol^{-1} and a melting temperature $T_m = 62$ °C. CD spectra of the cognate DNA (25 μM) were collected with the same instrumental settings with 5 μM incremental additions of protein up to 50 μM (2:1). The Extreme function available in Origin 7.0 software was used to fit the DNA spectra from 255 to 300 nm, and the intensity at 275 nm was obtained from the fit.

Tryptophan fluorescence spectra were recorded on a FluoroMax-3 spectropolarimeter at 25 °C with 2 μM protein in buffer B. The samples were excited at 295 nm, and the emission spectra were recorded by scanning from 315 to 450 nm. Each spectrum was fit to the Extreme function (Origin 7.0) to obtain a λ_{max} for fluorescence emission, and the experiments were carried out in triplicate to obtain statistically reliable values for λ_{max}. The effect of DNA concentration on the fluorescence λ_{max} was fit to a standard binding quadratic (*16*) to determine the dissociation constant K_D and number of Int^{CB}-binding sites on the DNA, n:

$$F_{obsd} = B + (A - B)\left[\frac{(P + nX + K_D) - [(P + nX + K_D)^2 - 4nXP]^{1/2}}{2P}\right]$$

where F_{obsd} is the observed λ_{max}, A and B are the λ_{max} values for the free and the fully bound protein, respectively, and P and X are the total protein and DNA concentrations.

ITC experiments were performed in buffer A. The protein and the DNA samples were dialyzed for 4 days in the same buffer container using 3500 and 100 Da cutoff dialysis membranes (Pierce, Inc.), respectively, to ensure minimal buffer mismatch between the samples. ITC experiments were carried out at 25 °C on a MicroCal VP-ITC instrument fitted with a 250 μL injection syringe. All the protein and DNA samples were thoroughly vacuum degassed. In the "forward titration", the protein concentration in the syringe was either 1 or 1.1 mM, and the DNA concentration in the cell was 40 μM. In the "reverse titration", the protein concentration in the cell was 50 μM, and the DNA concentration in the syringe was 250 μM. Each experiment consisted of an initial 3−5 μL injection followed by 28−29 10 μL injections. To estimate the heat of dilution during forward titrations,

separate experiments were performed by making 10 injections of the protein sample into the cell containing buffer alone. These control experiments revealed no events other than the dilution of the protein sample. After correction for the heat of dilution of the protein, the binding curve was fit to a model with two independent binding sites, in which the heat content after each injection is given by the sum of contributions from each site (Origin 7.0):

$$Q = M_t V_0 (n_1 \Theta_1 \Delta H_1 + n_2 \Theta_2 \Delta H_2)$$

where M_t is the total concentration of macromolecule in the cell (DNA, in this case), V_0 is the cell volume, n_i, Θ_I, and ΔH_i are the stoichiometry, fractional saturation, and molar heat of binding to site *i*. Fractional saturation Θ is related to the dissociation constant, K_D, and total ligand concentration, L_t, as

$$\frac{1}{K_D} = \frac{\Theta}{(1 - \Theta)(L_t - n\Theta M_t)}$$

Fitting of the isotherm to a "sequential" model (Origin 7.0) yielded similar results. Each reverse titration experiment was repeated in triplicate, and the thermodynamic parameters reported are the average and the standard deviation of the best fit values.

ACKNOWLEDGMENT

We thank C. Yuan, C. Cottrell, and M. Ibba for instrument access and support, A. Landy (Brown University), T. Ellenberger (Washington University), T. Wakagi (Aomori University), S. Subramaniam, C. Amero, and P. Mehta for materials and helpful discussions, and members of the Foster laboratory for critical reading of the manuscript.

SUPPORTING INFORMATION AVAILABLE

Two figures, ^{15}N-edited HSQC spectrum of the 1:1 IntCB−DNA complex with resonance assignments indicated and weighted-averaged chemical shift perturbations. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Landy, A. (1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination, *Annu. Rev. Biochem. 58*, 913−949.
2. Azaro, M. A., and Landy, A. (2002) Lambda Integrase and the Lambda Int Family, in *Mobile DNA II* (Lambowitz, A. M., Ed.) ASM Press, Washington, DC.
3. Tirumalai, R. S., Healey, E., and Landy, A. (1997) The catalytic domain of lambda site-specific recombinase, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6104−6109.
4. Tirumalai, R. S., Kwon, H. J., Cardente, E. H., Ellenberger, T., and Landy, A. (1998) Recognition of core-type DNA sites by lambda integrase, *J. Mol. Biol. 279*, 513−527.
5. Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005) A structural basis for allosteric control of DNA recombination by lambda integrase, *Nature 435*, 1059−1066.
6. Moitoso de Vargas, L., Pargellis, C. A., Hasan, N. M., Bushman, E. W., and Landy, A. (1988) Autonomous DNA binding domains of lambda integrase recognize two different sequence families, *Cell 54*, 923−929.
7. Radman-Livaja, M., Shaw, C., Azaro, M., Biswas, T., Ellenberger, T., and Landy, A. (2003) Arm sequences contribute to the architecture and catalytic function of a lambda integrase-Holliday junction complex, *Mol. Cell 11*, 783−794.
8. Wojciak, J. M., Sarkar, D., Landy, A., and Clubb, R. T. (2002) Arm-site binding by lambda -integrase: solution structure and functional characterization of its amino-terminal domain, *Proc. Natl. Acad. Sci. U.S.A. 99*, 3434−3439.
9. Aihara, H., Kwon, H. J., Nunes-Duby, S. E., Landy, A., and Ellenberger, T. (2003) A conformational switch controls the DNA cleavage activity of lambda integrase, *Mol. Cell 12*, 187−198.
10. Kamadurai, H. B., Subramaniam, S., Jones, R. B., Green-Church, K. B., and Foster, M. P. (2003) Protein folding coupled to DNA binding in the catalytic domain of bacteriophage lambda integrase detected by mass spectrometry, *Protein Sci. 12*, 620−626.
11. Subramaniam, S., Tewari, A. K., Nunes-Duby, S. E., and Foster, M. P. (2003) Dynamics and DNA substrate recognition by the catalytic domain of lambda integrase, *J. Mol. Biol. 329*, 423−439.
12. Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties*, 2nd ed., W. H. Freeman, New York.
13. Zhou, H. X. (2001) The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins, *Biochemistry 40*, 15069−15073.
14. Subramaniam, S., Kamadurai, H. B., and Foster, M. P. (2007) Trans cooperativity by a split DNA recombinase: the central and catalytic domains of bacteriophage lambda integrase cooperate in cleaving DNA substrates when the two domains are not covalently linked, *J. Mol. Biol. 370*, 303−314.
15. Kamadurai, H. B. (2007) Mechanistic Insights into Catalysis and Allosteric Enzyme Activation in Bacteriophage Lambda Integrase, Ph.D. Thesis, Biophysics Graduate Program, The Ohio State University, Columbus, OH.
16. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman, New York.
17. Johnson, B. B., Dahl, K. S., Tinoco, I., Jr., Ivanov, V. I., and Zhurkin, V. B. (1981) Correlations between deoxyribonucleic acid structural parameters and calculated circular dichroism spectra, *Biochemistry 20*, 73−78.
18. Bloomfield, V. A., Crothers, D. M., and Tinoco, I. (2000) *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito, CA.
19. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
20. Spolar, R. S., and Record, M. T., Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA, *Science 263*, 777−784.
21. Jen-Jacobson, L., Engler, L. E., and Jacobson, L. A. (2000) Structural and thermodynamic strategies for site-specific DNA binding proteins, *Struct. Folding Des. 8*, 1015−1023.
22. Foster, M. P., McElroy, C. A., and Amero, C. D. (2007) Solution NMR of large molecules and assemblies, *Biochemistry 46*, 331−340.
23. Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., and Wright, P. E. (1998) Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA, *J. Biomol. NMR 12*, 51−71.
24. Cavanagh, J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Academic Press, Amsterdam, Boston.
25. Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006) Mechanisms of site-specific recombination, *Annu. Rev. Biochem. 75*, 567−605.
26. Guo, F., Gopaul, D. N., and van Duyne, G. D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse, *Nature 389*, 40−46.
27. Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., and Rice, P. A. (2000) Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping, *Mol. Cell 6*, 885−897.
28. Ennifar, E., Meyer, J. E., Buchholz, F., Stewart, A. F., and Suck, D. (2003) Crystal structure of a wild-type Cre recombinase-loxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation, *Nucleic Acids Res. 31*, 5449−5460.
29. Kwon, H. J., Tirumalai, R., Landy, A., and Ellenberger, T. (1997) Flexibility in DNA recombination: structure of the lambda integrase catalytic core, *Science 276*, 126−131.

30. Tekle, M., Warren, D. J., Biswas, T., Ellenberger, T., Landy, A., and Nunes-Duby, S. E. (2002) Attenuating functions of the C terminus of lambda integrase, *J. Mol. Biol. 324*, 649−665.
31. Radman-Livaja, M., Biswas, T., Ellenberger, T., Landy, A., and Aihara, H. (2006) DNA arms do the legwork to ensure the directionality of lambda site-specific recombination, *Curr. Opin. Struct. Biol. 16*, 42−50.
32. Jen-Jacobson, L. (1997) Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state, *Biopolymers 44*, 153−180.
33. Groth, A. C., and Calos, M. P. (2004) Phage integrases: biology and applications, *J. Mol. Biol. 335*, 667−678.
34. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
35. Keller, R. (2004) *The Computer Aided Resonance Assignment Tutorial*, CANTINA Verlag, Goldau, Switzerland.
36. Garrett, D. S., Seok, Y. J., Peterkofsky, A., Clore, G. M., and Gronenborn, A. M. (1997) Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the Escherichia coli phosphotransferase system, *Biochemistry 36*, 4393−4398.
37. Imamura, H., Jeon, B.-S., Wakagi, T., and Matsuzawa, H. (1999) High level expression of *Thermococcus litoralis* 4-α-glucanotransferase in a soluble form in *Escherichia coli* with a novel expression system involving minor arginine tRNAs and GroELS, *FEBS Lett. 457*, 393−396.

BI700974T