Structures of the Arm-type Binding Domains of HPI and HAI7 Integrases*

Received for publication, August 25, 2009 Published, JBC Papers in Press, September 8, 2009, DOI 10.1074/jbc.M109.059261

Aleksandra Szwagierczak‡, Uladzimir Antonenka§, Grzegorz M. Popowicz‡, Tomasz Sitar‡, Tad A. Holak‡, and Alexander Rakin†

From the ‡Max Planck Institute for Biochemistry, D-82152 Martinsried, Germany and the §Max von Pettenkofer Institute, Ludwig Maximilians University, D-80336 Munich, Germany

The structures of the N-terminal domains of two integrases of closely related but not identical DNA-associated genomic islands, Yersinia HPI (high pathogenicity island; encoding siderophore yersiniabactin biosynthesis and transport) and an Erwinia carotovora genomic island with yet unknown function, HA17, have been resolved. Both integrases utilize a novel four-stranded β-sheet DNA-binding motif, in contrast to the known proteins that bind their DNA targets by means of three-stranded β-sheets. Moreover, the β-sheets in InttHPI and InttHAI7 are longer than those in other integrases, and the structured helical N terminus is positioned perpendicularly to the large C-terminal helix. These differences strongly support the proposal that the integrases of the genomic islands make up a distinct evolutionary branch of the site-specific recombinases that utilize a unique DNA-binding mechanism.

Genomic islands, together with temperate phages, integrative plasmids, transposons, and integrative conjugative elements, make up the group of mobile genetic elements that play an important role in bacterial quantum leap evolution and adaptation (1–3). Genomic islands that carry clustered genes encoding vital functions supply bacteria with additional capabilities to withstand and overcome host defenses and to improve fitness. Genomic islands are integrative elements that are not able to self-transfer and replicate. Typically, they are composed of functional and recombination modules. The recombination module consists of a tyrosine family integrase and two attachment sites involved in recombination. The integrase promotes attP×attB site-specific DNA recombination of the genomic islands into highly conserved tRNA-encoding genes (attB recombination targets) of the host genome and subsequent excision (4, 5).

It has been demonstrated that the N-terminal domain of phage integrases is responsible for specific recognition of the arm-type site sequence of the attachment sites, a step that is essential for activity of the catalytic C-terminal domain responsible for the strand exchange. Three-dimensional structures of the N-terminal domains of two prokaryotic integrases, namely bacteriophage λ integrase and Tn916 transposon integrase, have been determined. Although they do not share significant sequence homology, both adopt similar structures and recognize the arm-type DNA site by inserting their N-terminal domain into a major groove of DNA (6–9). The N-terminal domain, consisting of a three-stranded antiparallel β-sheet, is proposed to be a new DNA-binding motif whose residue composition and position within the major DNA groove varied to alter specificity (6). Nevertheless, the genomic islands are evolutionarily divergent from phages and other mobile elements and represent a distinct mobile genetic element class. Moreover, island-encoded integrases are not closely related to phage integrases, as was expected previously (3, 10).

Four closely related genomic islands, Yersinia HPI (high pathogenicity island; encoding siderophore yersiniabactin biosynthesis and transport) (11), Ecoc54N or the pks island (encoding the cytotoxic polyketide colibactin in uropathogenic Escherichia coli CFT073) (12), and two genomic islands with yet unknown functions in Erwinia carotovora, HA17 and HA113, contain highly similar but not identical integrases that recognize asn tDNA genes as their attB integration sites (13). In contrast to the highly conserved bacterial attB attachment site, the element-encoded attP sites and integrases are subjected to sequence fluctuations. Parallel evolution of the corresponding integrase with its cognate attP sites results in inability of the evolved integrase to support recombination of the heterologous attP and to substitute heterologous integrase that has the same bacterial target. In this work, we present new insights into the evolution of two site-specific genomic island integrases through the three-dimensional structures of their N-terminal domains that utilize a novel four-stranded β-sheet DNA-binding motif.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The construct of the arm-type binding domain of HPI integrase comprised residues 1–80, whereas a corresponding domain of the HAI7 integrase covered residues 1–100. Both constructs were cloned into the pET21b vector (Novagen). Native proteins were expressed in the BL21(DE3) strain grown on standard LB medium, whereas for the selenomethionine labeling, the B834(DE3) strain was used in a defined modified M9 medium with methionine being replaced by selenomethionine. For all preparations, the medium was inoculated with freshly transformed overnight cultures, grown at 37 °C until A 600 = 0.6–0.7 was reached, and
induced with 1 mM isopropyl β-D-thiogalactopyranoside. Afterward, the temperature was lowered to 27 °C, and the cells were grown for ~16 h. After harvesting by centrifugation, cells were resuspended in lysis buffer (300 mM NaCl, 50 mM Na₂HPO₄ and 10 mM imidazole), broken by sonication, and centrifuged to remove cell debris. Clear supernatant was applied to a preparative gel filtration column in crystallization buffer (50 mM NaCl and 10 mM Tris, pH 7.6). Purity of the proteins was confirmed by SDS-PAGE and mass spectroscopy.

Electrophoretic Mobility Shift Assay Experiments—The ability of recombinant Int₄₂₀ and Int₄₁₀ to bind putative arm-type DNA was evaluated by electrophoretic mobility shift assays using FAM™-labeled probes generated by annealing two DNA oligonucleotides for each probe: attHPI_DR_for (5'-FAM-CGCAACTATTTGCGTGATTGTC-TTGGACATAATCATAGGGTGGTGGTC-TTGGACAGATATCCTACCAATTCTAAAATAACCTAA-3') and attHPI_DR_rev (5'-TAATTAAGCTAAACCTGAATT TTGGTATAAAAACCTGGAAAAAACAATCTGATGAT-3') for the HPI integrase and attHAI7_DR_for (5'-GTTATCATTGAGTTTGTTTTTTTGTTTTTTGAACACATACCTACACCATATT-3') and attHAI7_DR_rev (5'-GTATATTGGTGGTCATTATGGTGGTC-ATTACACTATTGATTGAT-3') for the HAI7 integrase.

Crystallization—Crystallization was performed with a sitting drop vapor diffusion method by mixing equal volumes of a drop containing 0.1 M MES, pH 6.3, 250 mM NaCl and 10 mM ammonium sulfate, and crystals optimized for data collection were measured on the PXII beamline of the Swiss Light Source (Villigen, Switzerland). Data were indexed, integrated, and scaled with the XDS package (14). The native data set from the Int₄₂₀ crystals was collected at 0.9873 Å, and native crystals diffracted up to 1.3 Å. The selenomethionine derivative was measured at 0.9796 Å for the peak, 0.9796 Å for the inflection point, 0.972 Å for the high remote, and 0.9875 Å for the low remote data sets; all data sets were collected up to 2.0 Å. Two selenomethionine crystals were measured, and corresponding data sets were merged to obtain higher redundancy. Data sets were of high quality and showed strong anomalous signals. Anomalous scatterers were found using SHELDX software (15). Interestingly, two sites, instead of the expected one site, were found. Map analysis showed that the selenium atom is very close to a special position and that the selenomethionine side chain assumes different conformations in each asymmetric unit, thus locally breaking the crystallographic symmetry. (This is also the case in the native data for the sulfur atom.) Two initial atom positions were refined using the autoSHARP software package (16). The resulting phases were improved by the DM program (17) and used for automated model building with ARP/wARP software (18). The resulting model of ~80% completeness was inspected and finished manually with the Xfit program (19). Restrained refinement enforced by Refmac5 software was then performed using the native data, followed by the addition of water molecules by ARP/wARP (20). Data collection, phasing, and refinement statistics are presented in Table 1. Most of the model has a clear and well interpretable electron density with the exception of a few solvent-exposed side chains. These parts were omitted in the final model. The R-factor of the presented structure of the HPI integrase domain is 20.8%, and R_free is 23.5%.

Data for the HAI7 integrase domain crystal, collected to 1.6 Å, were integrated, scaled, and merged by XDS and XSCAL programs. The structure was determined by molecular replacement using the Molrep program from the CCP4 suite with the structure of the previously solved Int₄₂₀ used as a probe. The model was then refined by Refmac5 and rebuilt by XtalView/ Xfit and by a subsequent Refmac5 refinement. Water molecules were added by the ARP/wARP program. Side chains of Lys7, Lys16, Lys46, Lys47, Arg49, Lys80, Gln93, Lys95, and Arg96 had no interpretable electron densities and therefore were omitted in the model. The final R-factor of the structure of the HA17 integrase domain is 14.7%, and R_free is 21.2%.

RESULTS

N-terminal Domains of Int₄₂₀ and Int₄₁₀ Bind Short Direct Repeats in Heterologous attP Sites—Previously, we reconstituted the attP site of HPI and defined several copies of imperfect direct repeats located in the regions that can be designated “arms” by analogy with integrases of temperate phages (5). To determine whether an N-terminal part of Int₄₂₀ indeed specifically recognizes attP_HPI DNA, Int₄₂₀ was tested in the electrophoretic mobility shift assay for its ability to bind the attP_HPI fragment containing direct repeats. This Int₄₂₀ protein (spanning Met1–Asn80) was able to bind only its cognate attP_HPI but not attP_HAI7 fragment (Fig. 1). Similarly, the
Structures of HPI and HAI7

IntHAI7100 protein (Met1–Arg100), representing the N-terminal domain of IntHAI7, was able to form retarded bands with only attP HAI7 but not attP HPI DNA (Fig. 1).

Overall Structures—The structure of the arm-type binding domain of the HPI integrase (residues 1–80) comprises a four-stranded antiparallel β-sheet preceded and followed by α-helices (Fig. 2A). The N-terminal α-helix H1 covers Asp5–Thr10 and is positioned perpendicularly to the C-terminal larger helix H2. Helix H1 is followed by a 7-residue-long loop connecting it to the first β-strand B1 (Phe18–Ser23). β-Strands are parallel in the following pairs: B1 with B2 and B3 with B4. The first two strands (B1 and B2) form angle of 45° with respect to strands B3 and B4. The β-sheet is connected to helix H2 (Leu61–Ala76) by the 5-amino acid loop. This helix is positioned almost parallel to strands B3 and B4. The β-sheet, together with helix H2, forms the L-shaped hydrophobic core of the protein.

The structure of the HPI integrase domain shows a high resemblance to that of the arm-type binding domain of the HAI7 integrase (Fig. 2B). The positioning of helices and β-strands is identical, and the root mean square deviation of superimposed structures (aligned taking into account C-α atoms) is 0.81 Å. The largest difference in the main chain tracing is localized in the loop between strands B3 and B4, with the IntHAI7100 main chain bending toward helix H1. This segment

### TABLE 1

|                | HPI          | HAI7         |
|----------------|--------------|--------------|
| Data collection|              |              |
| Space group    | P4,2,2       | P4,2,2       |
| a (Å)          | 48.75        | 52.56        |
| b (Å)          | 78.75        | 52.53        |
| c (Å)          | 74.20        | 79.47        |
| Resolution range (Å) | 20–1.3       | 6–1.6        |
| Wavelength (Å) | 0.9873       | 1.0015       |
| Observed reflections | 391,656      | 183,210      |
| Unique reflections | 19,463       | 15,089       |
| Whole resolution range (%) | 85           | 91.6         |
| Completeness (%) | 93.1         | 96.6         |
| Rmerge (%)     | 30.7         | 2.7          |
| I/σI (%)       | 29.07        | 45.32        |
| Resolution range (Å) | 1.3–1.4       | 9.31         |
| Completeness (%) | 65.5         | 76           |
| Rmerge (%)     | 1.52         | 10.7         |
| I/σI (%)       | 29.85        | 30.67        |
| Phasing        |              |              |
| No. of sites found/present | 2/2 (each Se site as 50% occupancy) | 2/2          |
| Phasing power anomalous/isomorphous | 4.2/1.622 | 0.65         |
| FOM*           | 0.65         |              |
| Refinement     |              |              |
| No. of reflections | 18,404       | 14,113       |
| Resolution (Å) | 10.13        | 6.16         |
| R-factor (%)   | 20.8         | 14.76        |
| Rmerge (%)     | 21.21        | 6.02         |
| Average B (Å²) | 16.7         | 23.3         |
| r.m.s. bond length (Å) | 0.08         | 0.024        |
| r.m.s. angles  | 1.28°        | 2.051°       |
| Content of asymmetric unit |              |              |
| No. of protein molecules | 1            | 1            |
| No. of protein residues/atoms | 77/612       | 95/866       |
| No. of solvent atoms | 125          | 130          |
| No. of heavy atom sites identified | 2            | 2            |
| Ramachandran statistics |              |              |
| Most favored regions (no./%) | 59/91        | 81/95        |
| Additionally allowed regions (no./%) | 6/9          | 4/5          |
| Generously allowed regions (no./%) | 0/0          | 0/0          |
| Disallowed regions (no./%) | 0/0          | 0/0          |

* FOM, figure of merit; r.m.s., root mean square.

**FIGURE 1.** Electrophoretic mobility shift assay of different purified integrase N-terminal domains: IntHAI7(1–100) and IntHPI(1–80) with attP HAI7 (A) and attP HPI (B) arm DNAs.
shows also a difference in its primary structure with the large residues (Glu\textsuperscript{45} and Lys\textsuperscript{46}) in Int\textsubscript{HAI7}\textsuperscript{100} exchanged for small ones in Int\textsubscript{HPI}\textsuperscript{80} (serine and glycine, respectively) (Fig. 3). The crystallized Int\textsubscript{HAI7} construct is longer than Int\textsubscript{HPI} (spanning residues 1–100). The residues, located C-terminally to helix H2, are visible in the crystal structure, but this region is largely unstructured, with only residues 81–83, 86–88, and 92–95 forming very short helices. The electron density is, however, well defined, and the polypeptide chain is arranged antiparallel to helix H2.

A comparison of the amino acid sequences of the arm-type binding domains of Int\textsubscript{HPI} and Int\textsubscript{HAI7} is shown in Fig. 6. Conserved residues can be divided into several groups. The amino acids of the N-terminal helix H1 and the following loop form a large cluster of conserved residues (8 being identical and 4 being similar). Several of these residues stabilize the “vertical position” of helix H1 through formation of hydrogen bonds. For example, Arg\textsuperscript{65} binds the carbonyl group of Leu\textsuperscript{3}, and Leu\textsuperscript{3} also seems to create a hydrogen connection with the conserved Asp\textsuperscript{22}. Additionally, in Int\textsubscript{HPI}, the hydroxyl group of Tyr\textsuperscript{56} binds the carboxyl group of the conserved Asp\textsuperscript{29} from the top of helix H1.

The interactions that connect helix H2 to strand B1 seem to be of special importance because all residues responsible for this interaction are conserved. Asp\textsuperscript{22}, Arg\textsuperscript{65}, and Arg\textsuperscript{68} (exclusively in Int\textsubscript{HPI}) create a network of hydrogen bonds between one another. Arg\textsuperscript{68} in Int\textsubscript{HAI7}, although conserved, does not contribute to the interactions. Another group of well conserved residues are those present in or in the proximity of the turns: Gly\textsuperscript{25} in the turn between strands B1 and B2; Gly\textsuperscript{74} and Ser\textsuperscript{25} in the proximity of the turn between strands B2 and B3; and Gly\textsuperscript{54}, Pro\textsuperscript{57}, and Ala\textsuperscript{58} in the turn between strand B4 and helix H2.

The Int\textsubscript{HAI7}\textsuperscript{100} structure gives additional insight into the positioning of the C-terminal part of the protein. A dense network of hydrogen bonds formed by Lys\textsuperscript{73}–Ser\textsuperscript{82} caps the C terminus of helix H1 and thus fixes the region with respect to this helix. The following part of the C terminus does not form hydrogen interactions with the rest of the protein, with the exceptions of the carbonyl oxygen of Asn\textsuperscript{89} bonding Arg\textsuperscript{67} and the side chain of Asp\textsuperscript{64} interacting with NH of Ile\textsuperscript{62} (Fig. 4). It should be mentioned here that, as for proteins of such relatively small sizes (10–12 kDa), the arm-type binding domains of integrases and especially Int\textsubscript{HPI} possess a rich H-bonding network.

**Comparison with Structures of Three-stranded β-Sheet DNA-binding Proteins**—Several structures of arm-type binding domains of integrases have been solved (6, 21, 22). The align-
ment of the arm-type binding domain of HPI with the \( \lambda \) integrase and the GCC box-binding protein is shown in Fig. 5; the root mean square factors are 1.51 and 1.48 Å for the backbone atoms, respectively. These proteins bind their DNA recognition sites by means of the three-stranded \( \beta \)-sheet. Our structures differ from these published structures by a prominent extra \( \beta \)-strand, which is an integral part of the \( \beta \)-sheet. Furthermore, all \( \beta \)-strands in the structure of the arm-type binding domains of HPI and HAI7 integrases are significantly longer than in any other arm-type DNA-binding domains, with as many as 11 residues projecting from the \( \beta \)-sheet in the direction of the potential DNA-binding cavity (with the Tn916 transposon integrase having 9 of them and the \( \lambda \) integrase only 7) (Fig. 6).

The other distinguishing feature of our integrase domain is its structured helical N terminus positioned perpendicularly to the large C-terminal helix. Among all known integrases, a small N-terminal helix is present only in a newly described structure of the \( \lambda \) integrase in complex with DNA (9). The positioning of this helix is, however, different from that in the HPI and HAI integrases, with the N terminus of the \( \lambda \) integrase pointing away from the main body of protein (Fig. 5). Because the structure of the N terminus of the \( \lambda \) integrase proved to be significantly different in the free versus DNA-bound state, it cannot be excluded that the N-terminal helix of HPI and HAI7 integrases could change its conformation upon binding to the attachment site.

It is still unconfirmed whether the mode of DNA binding is conserved between the integrases possessing three-stranded \( \beta \)-sheets and the HPI and HAI7 integrases. The space created by a concave structure beneath the \( \beta \)-sheet is, however, large enough to fit the major groove of DNA. Furthermore, the charge distribution in IntHPI80 and IntHAI7100 is similar to that in other arm-type binding proteins, with positive charges gathered on the concave surface of the \( \beta \)-sheet (Fig. 7).

Comparison of DNA-binding Cavities of N-terminal Domains of HAI7 and HPI Integrases—The largest observed differences between our solved structures appear, as expected, on the putative DNA-binding interface. The majority of the residues conserved between IntHPI80 and IntHAI7100 present in the \( \beta \)-sheet have their side chains directed toward the core of the interface between the \( \beta \)-sheet and helix H2 (Leu\(^{28}\), Val\(^{30}\), Trp\(^{38}\), Leu\(^{40}\), and Tyr\(^{42}\)). Three conserved residues are facing the putative DNA-binding cavity, namely Lys\(^{19}\), Leu\(^{29}\), and Arg\(^{43}\). Other residues pointing in the same direction differ substan-
tially between the HPI and HAI7 integrases, as well as (although to a lesser extent) between the HA113 and HA17 integrases.

Comparison of the putative DNA-binding cavities of IntHPI80 and IntHA17100 shows several differences. First, the proximal border of the cavity (composed of loop L1, parts of β-strands B1 and B2, and the turn between them) in IntHPI80 bears a strong positive charge, introduced by Lys16, Lys19, and Lys31 (Fig. 7). The situation differs in the case of IntHA17100, which not only lacks Lys31 (substituted with histidine, which remains buried) but also presents the negative charge from Glu17 protruding from the surface of the cavity. An exact position of the side chain of Lys16, which is present in both proteins, could not be detected in the IntHA17100 structure, probably due to its high mobility.

The central part of the cavity (formed by the middle parts of strands B3 and B4) in IntHPI80 is dominated by the presence of the positive charge of Lys16, next to the negative charge supplied by Glu48. In IntHA17100, none of these residues is present (Lys31 being substituted with Ser, and Glu48 with Gln). However, Leu50 protrudes into the cavity approximately in the same place as the charged residues in IntHPI80. The distal rim of the cavity in IntHA17100 (formed by the C-terminal end of strand B3, the N-terminal part of strand B4, and the loop between them) seems larger and bulkier than that in IntHPI80.

Putative Residues Interacting with DNA—The IntHPI and IntHA17 DNA-binding interfaces bear too many differences to unambiguously link specific residues with the DNA base recognition. Therefore, we used a related integrase, IntHA13, which shows a strong sequential resemblance to IntHA17, to obtain more insight into the residues’ functions. A model of the IntHA13 DNA-binding cavity was constructed based on the structure of HA17 with surface-exposed residues mutated in the Swiss-PdbViewer (Fig. 8). Comparison of the crystal and modeled structures shows that differences group in a cluster in the central part of the cavity (residues mutated: H27T, R39Q, Lys31 being substituted with Ser, and Glu48 with Gln). However, Leu50 protrudes into the cavity approximately in the same place as the charged residues in IntHPI80. The distal rim of the cavity in IntHA17100 (formed by the C-terminal end of strand B3, the N-terminal part of strand B4, and the loop between them) seems larger and bulkier than that in IntHPI80.

Putative Residues Interacting with DNA—The IntHPI and IntHA17 DNA-binding interfaces bear too many differences to unambiguously link specific residues with the DNA base recognition. Therefore, we used a related integrase, IntHA13, which shows a strong sequential resemblance to IntHA17, to obtain more insight into the residues’ functions. A model of the IntHA13 DNA-binding cavity was constructed based on the structure of HA17 with surface-exposed residues mutated in the Swiss-PdbViewer (Fig. 8). Comparison of the crystal and modeled structures shows that differences group in a cluster in the central part of the cavity (residues mutated: H27T, R39Q,
S41G, and L50V). This results in a slight change of the surface of the cavity, mostly induced by the substitution of Leu at with a smaller Val side chain. Furthermore, the hydrogen bond pattern between the protein and the target DNA fragment can be changed because a histidine, which can serve both as a donor and an acceptor, is substituted with a donor residue, tyrosine. The reverse is the case for the R39Q substitution, where a double donor is exchanged for the donor + acceptor residue. Another significant change in the DNA-binding interface is the smaller Val side chain. Furthermore, the hydrogen bond pattern between the protein and the target DNA fragment can be attributed to the recognition of base 4.

FIGURE 9. Attachment sites of HPI, HAI7, and HAI13 integrases. Bases that are different for each integrase are indicated in blue. Bases that are identical in HAI7 and HAI13, although different from those in HPI, are indicated in green.

DISCUSSION

Two closely related but still non-identical asp tDNA-associated integrases of the four genomic islands give us a unique opportunity to superimpose differences in their amino acid structure and the cognate co-evolved attP recombination sites to address specific protein-DNA recognition. The structures of the arm-type binding domains of both Int_HPI and Int_HAI7 integrases resemble each other and comprise a four-stranded antiparallel β-sheet preceded and followed by α-helices. The main difference can be localized in the loop between strands B3 and B4, with the HAI7 main chain bending toward helix H1.

Comparison with the structures of the arm-type binding domains of other integrases demonstrates that the arm-type binding domains of Int_HPI and Int_HAI7 differ by the presence of an extra β-strand. In contrast, all known proteins bind their DNA recognition sites by means of three-stranded β-sheets (6, 21, 22). Additionally, all β-strands in the structures of the arm-type binding domains of HPI and HAI7 integrases are significantly longer than those in three-stranded arm-type binding domains, which enables mimicking the shape of the major groove of a longer DNA fragment. HPI and HAI7 attachment sites comprise eight nucleotides, the same number found as DNA-interacting for the λ integrase (9). In the λ integrase, however, only six nucleotides are recognized through the interaction of the β-sheets with the major groove, whereas the remaining two are bound by the N-terminal helix inserting into the minor groove (9). Therefore, it can be speculated that the long β-sheets in Int_HPI and Int_HAI7 are sufficient for recognition without a contribution of the N-terminal helix. The other distinguishing feature of the Int_HPI and Int_HAI7 N-terminal domains is their structured helical N terminus positioned perpendicularly to the large C-terminal helix. These differences support the proposal that integrases of the genomic islands form a distinct evolutionary branch of the site-specific recombinases different from that of phage integrases (3). On the other hand, integrases of the genomic islands are bidirectional recombinases that efficiently support both integrative and excisive recombination and do not require an additional recombination directionality factor (RDF), in contrast to the phage integrases. For example, the excise activity of HPI by its RDF is supported at-fold (23), in contrast to the highly efficient RDF of bacteriophage λ. Moreover, not all asp tDNA-associated islands possess an RDF.

Evolutionary integrases of the temperate phages might represent a further step in function speciation to increase the efficiency of prophage rescue under unfavorable conditions. In contrast, genomic islands, being unable to replicate and to transfer, are long term passengers of the bacterial chromosome with the dominating integrative function. In addition, the inefficient RDF is not clustered with the cognate integrase, as in the case of temperate phages. This supports the proposal of their independent and consecutive acquisition by the integrative modules of genomic islands. Likewise, differences in the structures of Int_HPI and Int_HAI7 might be more suited for the recognition of both attP and recombinant attL/attR sites to support efficiently both types of site-specific recombination. Until now, it also had not been demonstrated whether the mode of DNA binding is conserved between the integrases possessing three-stranded β-sheets and Int_HPI and Int_HAI7.

Our resolved structures of the arm-type binding domains of the two asp tDNA-associated integrases show the largest observed differences between two structures on putative DNA-binding interfaces. Although 3 conserved residues are facing the putative DNA-binding cavity (Lys, Leu, and Arg), all other residues pointing in the same direction differ substantially between the HPI and HAI7 integrases, as well as (although to a lesser extent) between Int_HAI13 and Int_HAI17. Such β-sheet sequence-specific DNA binding is very rare and has been demonstrated only for three proteins that form complexes with DNA, namely the bacteriophage λ integrase, the Tn916 transposon integrase, and the ethylene-responsive factor AtERF1 from Arabidopsis thaliana.

At the moment, it is impossible to demonstrate what DNA nucleotides are actually interacting with the amino acid residues in the DNA-binding interface. However, co-crystallization of the arm-type binding domains with the cognate DNA arms might be the next step to solve this problem.
REFERENCES

1. Dobrindt, U., Hochcut, B., Hentschel, U., and Hacker, J. (2004) Nat. Rev. Microbiol. 2, 414–424
2. Hacker, J., Blum-Oehler, G., Mühldorfer, I., and Tschäpe, H. (1997) Mol. Microbiol. 23, 1089–1097
3. Boyd, E. F., Almagro-Moreno, S., and Parent, M. A. (2009) Trends Microbiol. 17, 47–53
4. Williams, K. P. (2002) Nucleic Acids Res. 30, 866–875
5. Rakin, A., Noelting, C., Schropp, P., and Heesemann, J. (2001) Mol. Microbiol. 39, 407–415
6. Wojciak, J. M., Sarkar, D., Landy, A., and Clubb, R. T. (2002) Proc. Natl. Acad. Sci. U.S.A. 99, 3434–3439
7. Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005) Nature 435, 1059–1066
8. Connolly, K. M., Wojciak, J. M., and Clubb, R. T. (1998) Nat. Struct. Biol. 5, 546–550
9. Fadeev, E. A., Sam, M. D., and Clubb, R. T. (2009) J. Mol. Biol. 388, 682–690
10. Balding, C., Bromley, S. A., Pickup, R. W., and Saunders, J. R. (2005) Environ. Microbiol. 7, 1558–1567
11. Rakin, A., Schubert, S., Pelludat, C., Brem, D., and Heesemann, J. (1999) in The High-Pathogenicity Island of Yersinia (Kaper, J. B., and H. J., eds) ASM Press, Washington, DC
12. Nougayrède, J. P., Homburg, S., Starke, T., Brem, D., and Heesemann, J. (2005) in The High-Pathogenicity Island of Yersinia (Kaper, J. B., and H. J., eds) ASM Press, Washington, DC
13. McRee, D. E. (1999) J. Struct. Biol. 125, 156–165
14. Lamzin, V. S., and Wilson, K. S. (1993) Acta Crystallogr. D Biol. Crystallogr. 49, 129–147
15. Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M. (1998) EMBO J. 17, 5484–5496
16. Wojciak, J. M., Connolly, K. M., and Clubb, R. T. (1999) Nat. Struct. Biol. 6, 366–373
17. Antonenka, U., Nöltling, C., Heesemann, J., and Rakin, A. (2006) Int. J. Med. Microbiol. 296, 341–352