Agrobacterium tumefaciens VirB8 structure reveals potential protein–protein interaction sites

Susan Bailey*, Doyle Ward†‡, Rebecca Middleton†, J. Gunter Grossmann*, and Patricia C. Zambryski†§

*Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720
†Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720
‡Present address: Program in Molecular Medicine, University of Massachusetts Medical School, 55 Lake Avenue North, S6-120, Worcester, MA 01655.
§To whom correspondence should be addressed. E-mail: zambrysk@nature.berkeley.edu.

Contributed by Patricia C. Zambryski, December 27, 2005

Bacterial type IV secretion systems (T4SS) translocate DNA and/or proteins to recipient cells, thus providing a mechanism for conjugative transfer of genetic material and bacterial pathogenesis. Here we describe the first structure of a core component from the archetypal Agrobacterium tumefaciens T4SS: the 2.2-Å resolution crystal structure of the VirB8 periplasmic domain (pVirB8AT). VirB8 forms a dimer in the crystal, and we identify residues likely important for stabilization of the dimer interface. Structural comparison of pVirB8AT with Brucella suis VirB8 confirms that the monomers have a similar fold. In addition, the pVirB8AT dimer superimposes very closely on the B. suis VirB8 dimer, supporting the proposal that dimer formation in the crystal reflects self-interactions that are biologically significant. The evolutionary conservation level for each residue was obtained from a data set of 84 VirB8 homologs and projected onto the protein structure to indicate conserved surface patches that likely contact other T4SS proteins.

**Results**

**Structure of the pVirB8AT Monomer.** The structure of pVirB8AT has been determined at 2.2-Å resolution by using the structure of the periplasmic domain of B. suis VirB8 (pVirB8SB) (6) as a model for molecular replacement. The final model contains a homodimer plus two molecules of 2-methyl-2,4-pentanediol and 170 water molecules in the asymmetric unit. The structure of each monomer (denoted monomer A and B throughout) comprises VirB8 residues 92–231 and additional N-terminal tag residues 88–91 (Gly-Pro-His-Met). Residues 232–237 are disordered. Residues 184–187 form a flexible loop that is sufficiently ordered to model the main chain into continuous 2Fo−Fc electron density for monomer A but displays discontinuous electron density at the 1σ level in monomer B. A cis-peptide bond was identified in each monomer at Pro-143.

Each monomer is composed of a single domain of approximate dimensions 20 × 30 × 45 Å that contains five helices clustered to one side of a four-strand antiparallel β-sheet (Fig. 1A). The β-sheet extends the length of the molecule and curves around helix α1. Helices α1 to α4 are at the N terminus of the crystallized fragment, whereas helix α5 is inserted in a loop between strands β3 and β4 of the sheet (Fig. 2A). The protein fold is similar to that of the NTF2-like family, including NTF2 (17), scytalone dehydratase (18), ketosteroid isomerases (19), and one domain of CaMKII (20), although the sequence identity is low (4–13%). NTF2 has a cone-shaped structure with a substantial cavity at the base of the cone. pVirB8AT differs by the presence of helix α5 and the absence of two short β-strands. The absence of the β-strands results in a deep groove on one face of pVirB8AT that is flanked by β4 and α3 at the top, extends from midway along the molecule length, and widens at the base (Fig. 2B).

Conflict of interest statement: No conflicts declared.

Abbreviations: HMM, hidden Markov model; T4SS, Type IV secretion system.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2cc3).

© 2006 by The National Academy of Sciences of the USA
The opposite face of pVirB8 AT is relatively flat and corresponds with the dimer interface.

**VirB8 Dimer Interface.** The presence of a homodimer of pVirB8 AT in the asymmetric unit allows us to describe the dimer interactions in detail. The pVirB8 AT monomers forming the dimer can be superimposed with a root-mean-square deviation of 0.175 Å between all 144 Cα atoms. The dimer interface buries a total of 1,720 Å², corresponding to 10% of the accessible surface area of each monomer, and is the largest contact area between two monomers in the crystal. Dimer interactions are divided into three distinct contact regions (Fig. 1D, see also Fig. 4, which is published as supporting information on the PNAS web site). The largest contact area, shown in detail in Fig. 1E, buries 1,180 Å² and is dominated by hydrophobic interactions between 13 surface residues. This contact region involves residues from the N terminus of α1, α5, and the residues connecting α5 to β4. There is only one hydrogen bond, between Glu-94 of monomer A and Gln-93 of monomer B. Two residues at the C-terminal end of α5 (Arg-210 and Pro-214) are highly conserved across VirB8 sequences and form a conserved surface patch at the dimer interface (Fig. 2A and D).

The second and third contact regions are significantly smaller with interactions predominantly between hydrogen bonding partners. The smallest interaction site in the central region of the dimer buries only 200 Å² and involves the side chains of Arg-107 and Tyr-119, which protrude from the surface of each monomer (Fig. 1D and F).

At the base of the dimer interface, the third contact point buries 340 Å² and is mediated by Asp-113, Asp-115, Thr-116, and Arg-157 from each monomer (Fig. 1G). The four aspartates have well defined electron density and contribute five intersubunit hydrogen bonds. Thus, our model implies that the Asp-115 residues are at least partially protonated. The aspartates are surrounded by the positively charged side chains of Arg-157 and Lys-154, albeit the electron density for these side chains is poor and it is not possible to confidently identify specific interactions. Arg-157 has the potential to form a salt bridge with either Asp-113 of the same monomer, or Asp-115 of the 2-fold related monomer, and may play a role in stabilization of this aspartate cluster. The crystals were grown at pH 4.6 and, depending on the pKa of the aspartates, it is unlikely that the Asp–Asp interactions are present at higher pH. Potentially, dimerization is influenced by pH.

**Comparison of pVirB8 AT and pVirB8 BS.** pVirB8 AT and pVirB8 BS share 30% sequence identity, and as expected, the overall fold of
pVirB8AT is very similar to the published structure of pVirB8BS (6). Both structures correspond to the periplasmic fragment of the VirB8 protein, composed of residues 92–231 of the A. tumefaciens protein and 97–234 of the B. suis protein. Superposition of monomer A or B of pVirB8AT on one of the five monomers of pVirB8 BS present in the asymmetric unit gives a root-mean-square deviation between 1.33 and 1.42 Å for 125–127 C\textsubscript{α} atoms, depending on the monomers superimposed. The biggest differences in regions of superposition, highlighted in Fig. 1C, are seen at the N- and C-termini, and at three surface loop regions, comprising residues 139–152, 167–169, and 205–215.

Only 4 of 19 residues involved in dimer interactions are conserved between pVirB8AT and pVirB8BS (Fig. 2A). Nevertheless, dimer formation, in both VirB8 proteins, involves packing of the two monomers in a very similar fashion and the pVirB8AT dimer superimposes on the pVirB8 BS BD and CE dimers with a root-mean-square deviation of 1.52 Å (251 C\textsubscript{α} atoms) and 1.65 Å (252 C\textsubscript{α} atoms), respectively (Fig. 4A). Several residues that interact at the largest dimer interface (Fig. 1E) correspond to bulkier side chains in the B. suis protein (A. tumefaciens Gln-93, Val-97, Ala-100, and Gly-215 are replaced by B. suis Tyr-98, Met-102, Tyr-105, and Leu-222) and, as a consequence, the N-termini of α in the A and B monomers pack closer in pVirB8BS. In the central dimer region, residues Arg-107 and Tyr-119 of pVirB8AT are replaced, in the B. suis protein, by smaller side chains Ile-112 and Lys-124, respectively, with the consequence that the opening observed between the two pVirB8BS monomers is much less pronounced in pVirB8AT (Fig. 4B and C).

Mapping of Conserved Residues. The biological importance of a residue often correlates with its level of evolutionary conservation within the protein family (15). We analyzed a broad range of VirB8 sequences and determined highly conserved residues that are potentially necessary for either structural integrity or biological activity. These residues were then mapped onto the structure of pVirB8AT.

The multiple sequence alignment used for the CONSURF analysis (15) was composed of 84 sequences (Fig. 5, which is published as supporting information on the PNAS web site). Representative sequences illustrate the degree of conservation (Fig. 2A) and this conservation was mapped onto the protein structure (Fig. 2B–D). The most highly conserved residues are either buried or contribute to one of two clusters exposed at the surface of the protein. Thirteen residues in the two most conserved CONSURF levels were found to have little or no solvent accessibility at the monomer surface: Tyr-105, Val-106, Tyr-112, Val-124, Ser-128, Tyr-136, Ser-147, Ile-175, Ala-195, Asn-213, Gly-216, Leu-217, and Val-219. Further analysis of these buried conserved residues revealed that they also appear to form clusters that are located adjacent to the conserved surface patches and, thus, are presumably important for maintaining the structural integrity of these surface patches. It seems highly likely that these conserved surface regions are important for maintaining the functionality of VirB8 (16).
One conserved region, formed by residues 210–219, is defined by the sequence motif RxxNxGxxV (Fig. 2 A and D) and appears to be important to stabilize and position the loop containing helix α5 between strands β3 and β4. This loop appears to be relatively rigid in our crystal structure, reflected by lower-than-average temperature factors. The local fold is stabilized by extensive hydrogen bonds between the side chain of Arg-210 and main-chain atoms of residues Thr-203 and Pro-214, and the side chain of Asn-213 and main-chain atoms of residues Asn-213, Gly-216, and Leu-217 (Fig. 3). Other stabilizing influences are hydrophobic interactions between residues Leu-217 and Val-219 of this loop, which pack against the N terminus of helix α1. In addition, side-chain and main-chain atoms of residues Arg-210, Leu-211, Pro-214, and Gly-215 interact with the second pVirB8AT monomer at the dimer interface (Fig. 1E). The positioning of this loop is likely essential for maintenance of the dimer interface. The highly conserved residues (Arg-210, Asn-213, Pro-214, Gly-216, and Val-219) of this loop are largely buried in our crystal structure, only Arg-210 and Pro-214 are exposed at the monomer surface, and Pro-214 becomes buried on dimer formation. It is not clear whether Arg-210 would be exposed at the surface of the native protein. In our structure, Arg-210 is predominantly protected from solvent by nonnative residues N-terminal to Thr-92, the first VirB8AT residue of the exposed at the surface of the native protein. In our structure, bridge in the pVirB8BS structure (6), but in VirB8 AT, these VirB8 homologs as an aspartate and an arginine, form a salt

groove. Adjacent to the salt bridge, Ala-164 is a highly conserved serine in most VirB8 homologs.

Mapping of Functional Residues. Kumar and Das (14) introduced random mutations into pTiA6-derived virB8 of A. tumefaciens strain A348 and identified five mutations, G78S, S87L, A100V, R107A, and T192M, that produce stable proteins but fail to complement a virB8-deletion mutant in tumor-forming assays. The S87L mutant exhibited a semidominant phenotype providing early evidence that VirB8 forms a multimer (14). Two mutations, G78S and S87L, are in a region of the polypeptide (residues 60–90) linking the predicted transmembrane helix with the VirB8 periplasmic domain and are not present in the construct we have expressed. This linking peptide is disordered in the crystal structure of pVirB8AT and does not appear to be a structural component of the periplasmic domain, which is stable in its absence. The S87L mutant prevents interaction with VirB9 in the yeast two-hybrid assay and abolishes DNA transfer function (11). Although Ser-87 is not part of our structure, it will be close to the observed N terminus, suggesting that the VirB9 interaction site is close to the top of the VirB8 molecule as viewed in Fig. 1A.

The three remaining mutations, A100V, R107A, and T192M, can be mapped onto the presented structure (Fig. 2 A and D). Although these residues are not highly conserved among VirB8 homologs, it is useful to consider the role of these residues in more detail. All three residues are located on the surface of a VirB8 monomer and are close to the conserved surface patches and likely to be involved in interactions with other proteins. Both Ala-100 and Arg-107 are involved in dimer interactions in the crystal. The side chain of Ala-100 is in close contact with the side chain of Pro-214 (3.7 Å) and the carbonyl oxygen of Leu-211 (3.2 Å) of the opposing monomer. Thus, it is possible that the A100V mutation could disrupt formation of the VirB8 dimer. However, yeast two-hybrid assays show that the mutated protein continues to interact with itself and with VirB9 and VirB10. Yeast two-hybrid assays also show that the R107A mutant has retained the ability for self-interaction but is not able to interact with VirB9 or VirB10. In our structure, Arg-107 is hydrogen-bonded to Tyr-119 from the noncrystallographic symmetry-related monomer (Fig. 1F), and replacement with an alanine is unlikely to disrupt dimer formation. VirB8 may not always be in the dimeric state in vivo, allowing interaction with other T4SS components such as VirB9 or VirB10. The apparent conflict between yeast two-hybrid and tumorigenesis data may be explained by innate differences in each of the assays. Tumour formation requires VirB8 activity and may involve dynamic interactions with many T4SS components and T4SS substrates. In contrast, two-hybrid interactions that are compromised by single amino acid changes may be stabilized by other contacts along the length of bait and prey peptides.

A T192M mutant of VirB8 is unable to complement a VirB8 deletion mutant but does not affect VirB8 interactions with itself or with VirB9 and VirB10 (14). Thr-192 is in close contact with the highly conserved salt bridge between Arg-176 and Asp-227 and with Ile-163 and Lys-178. Possibly mutation to a methionine could disrupt the local tightly fitted structure or remove required flexibility in residues surrounding the groove.

Solution Characterization of VirB8. The apparent molecular mass of pVirB8AT was 23 kDa by SDS/PAGE and 70 kDa by native PAGE, suggesting dimer or trimer formation. To resolve the discrepancy, pVirB8AT was subjected to gel filtration at pH 8.0 and pH 4.5. At both pH values, pVirB8AT was eluted as a monomer with a molecular mass of 20 kDa. Similarly, x-ray scattering experiments provided results that are consistent with a monomer in solution with a radius of gyration of 18.8 ± 0.2 Å and maximum dimension 63 ± 3 Å. Furthermore, simulations
based on the monomer and dimer taken from the crystal structure of \textit{pVirB8}AT, suggest the monomer best fits the scattering curve, yielding a goodness-of-fit value ($\chi$) of 2.33 (in contrast to $\chi = 8.72$ for the dimer, data not shown). Thus it appears that the protein is monomeric in solution and runs anomalously on native PAGE.

**Discussion**

VirB8 is a core component of the \textit{A. tumefaciens} T4SS. Our structure suggests specific residues involved in dimer formation and allows us to predict surface features that could direct its interaction with other VirB proteins. We believe that the observed dimer is physiologically significant. Yeast two-hybrid studies have shown that VirB8 is able to interact with itself (11) and mutagenesis of residues at the dimer interface, A100V and R107A, reduces VirB8 function (14). Significantly, the crystal structure of \textit{pVirB8}AT also indicates a nearly identical dimeric structure (Fig. 4A); the similarity between quaternary structures would be very unlikely if dimer formation were solely a consequence of crystallization.

Data mining of protein structure databases can indicate correlative properties such as geometric features, chemical features, and residue conservation that distinguish a biological interface from a nonspecific crystal contact (21 and references therein). Despite wide variation in the values of these properties, the size of the surface interface and, to a lesser extent, the degree of hydrophobicity or the presence of conserved residues, can aid differentiation of biological interfaces from crystal contacts (21). The \textit{pVirB8}AT dimer interface buries 1,720 Å² of accessible surface area, corresponding to 10% of the total monomer accessible surface area. This interface is larger than the other crystal contacts in our structure and larger than the average crystal contact in the Protein Data Bank (PDB), although it is slightly smaller than the average dimer interface (21). In addition, we observe highly conserved residues participating in dimer formation. Assuming VirB8 dimerization is of biological significance, it would be helpful to investigate the conditions that promote dimerization. It is possible that dimer formation in solution is energetically unfavorable for the isolated periplasmic domain, requiring additional interactions formed by the missing N-terminal polypeptide. Alternatively, VirB8 may be required to interact with itself only under certain conditions, for example, in the presence or absence of other T4SS components.

We identify two highly conserved surface regions that are likely of functional significance. Although the precise biological role of VirB8 is unknown, there is good evidence that VirB8 interacts with a number of other VirB proteins and with DNA during translocation (11). Therefore, we suggest that the clustered, solvent-accessible, conserved residues correspond to sites of interaction of VirB8, either with itself, with other VirB proteins, or with DNA. The smaller conserved region corresponds to (210-RxxNPxSxxV-219) and likely functions to position the extended loop containing helix $\alpha_5$, Arg-210, and Pro-214 for proper dimer-forming interactions. The VirB9 interaction site may also be close to this conserved region.

The larger conserved region includes surface-exposed residues Arg-109, Glu-110, Val-132, Ala-164, Arg-176, Tyr-222, Gln-223, Ser-225, and Asp-227, which can be seen as dark red and pink conserved patches in two views of VirB8AT (Fig. 2B and C). The extended nature of this conserved surface patch, and that VirB8 interacts with multiple VirB components, suggests this region may enable specific protein interactions. The valley of the groove flanked by $\beta 4$ and $\alpha 3$ likely forms one interaction site. Structurally, the groove corresponds to the hydrophobic cavity of NTF2, which comprises the protein interface region of the NTF2–Ran complex (22). Examination of the NTF2-like family also reveals that the central region of the $\beta$-sheet serves as a common site for protein–protein interactions. This region contributes to the homodimer interfaces of NTF2, CaMKII, and ketosteroid isomerase, and to the heterodimer interface of the two NTF2-like mRNA export factors Mtr2 and Mex67 (23).

Thus, we suggest that the surface patch centered on highly conserved residues Ala-164, Arg-176, and Ser-225 on the face of the $\beta$-sheet of \textit{pVirB8}AT may form a second specific interaction site.

Mapping the level of conservation onto the structure of \textit{pVirB8}AT has enabled us to distinguish between residues likely essential for structural integrity and surface residues that may be important for VirB8 activity. The discovery of a highly conserved surface groove is especially intriguing, because it may provide a site for interactions with other T4SS components. The identification of conserved residues at the dimer interface and surface of VirB8 provokes the design of a mutagenesis program to further investigate the importance of these interfaces in the structure and function of the T4SS.

**Materials and Methods**

**Cloning and Protein Preparation.** A fragment encoding residues 92–237 of the VirB8 peptide was amplified from \textit{pTiCS8} and cloned into a modified pET28 vector. The resulting clone, pDW134, was used to transform \textit{Escherichia coli} strain BL21(DE3). His-tagged \textit{pVirB8}AT was expressed and purified on a Ni-NTA column, followed by anion exchange chromatography. Purity of \textit{pVirB8}AT was checked by SDS/PAGE and native PAGE.

**Gel Filtration and Electrophoresis.** SDS/PAGE and native PAGE were run on a PHAST system by using standard protocols. Gel filtration experiments used a Superose-75 column with buffer systems 50 mM Tris-HCl pH 8.0/50 mM NaCl or 50 mM sodium citrate, pH 4.5/50 mM NaCl.

**Solution X-Ray Scattering.** Scattering data were collected (station 2.1 of the Daresbury Synchrotron Radiation Source) of protein samples (at concentrations between 0.5 and 12 mg/ml) and buffer in the momentum transfer interval 0.03 Å⁻¹ ≤ $q ≤ 0.71$ Å⁻¹, where $q = 4\pi \sin(\theta)/\lambda$ (20 is the scattering angle and $\lambda$ is the x-ray wavelength 1.5 Å) according to published procedures (24).

The distance distribution function $p(r)$, and the radius of gyration were evaluated with the program GNOM (25), which also provides a reliable estimation of the overall particle size. Scattering pattern simulations based on crystal structure information were carried out with the program CRYSTOL (26). The goodness-of-fit value ($\chi$) is a measure of how well the structure fits the experimental data (a reliable agreement between experiment and simulation is generally obtained for $\chi < 3$).

**Crystallization and Data Collection.** Crystals of \textit{pVirB8}AT (10–13 mg/ml) were obtained at 4°C by vapor diffusion against a reservoir of 0.1 M sodium acetate, pH 4.6/0.2–0.4 M sodium bromide/20% vol/vol 2-methyl-2–4 pentanediol. Crystals belonged to the space group C2 with cell dimensions $a = 112.3$ Å, $b = 72.3$ Å, $c = 48.8$ Å, and $\beta = 110.7$ Å and diffracted anisotropically, to 1.8 Å in one direction and to 2.2 Å in the orthogonal direction. Data were collected at the Daresbury Synchrotron Radiation Source and processed by using programs in the CCP4 SUITE (27) to yield a 2 Å native data set. Data collection and processing statistics are presented in Table 1, which is published as supporting information on the PNAS web site.

**Structure Determination.** The crystal structure was determined by molecular replacement by using the structure of a periplasmic domain of VirB8 from \textit{B. suis} (pVirB8B), PDB entry 2BHM (6), as a search model. The positions of two monomers, related by a noncrystallographic dyad axis, were identified. Maps were im-

---

*Bailey et al.*
proved by using two-fold noncrystallographic symmetry averaging, solvent flattening, and extensive model building. Refinement using REFMACS (28) converged to a model with R and R_free factors of 23.5% and 29.8% with good stereochemistry, and refinement statistics are given in Table 1. Noncrystallographic symmetry restraints were applied throughout the entire refinement process.

**Sequence Comparisons.** Several methodologies were followed for comparison of homologous sequences and identification of conserved sites on the surface of VirB8 and all methodologies yielded similar results. For the analysis described here, a Hidden Markov Model (HMM) was constructed by using the program HMMPRE on a pairwise structural alignment of A. tumefaciens VirB8 residues 78–231 and B. suis VirB8 residues 77–238.
