# *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon–helix–helix DNA-binding fold

Jun Lu[a], Amke den Dulk-Ras[b], Paul J. J. Hooykaas[b,1], and J. N. Mark Glover[a,1]

[a]Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7; and [b]Institute of Biology, Leiden University, Clusius Laboratory, Wassenaarseweg 64, 2333 AL, Leiden, The Netherlands

Edited by Patricia C. Zambryski, University of California, Berkeley, CA, and approved April 29, 2009 (received for review December 2, 2008)

***Agrobacterium tumefaciens*** **VirC2 stimulates processing of single-stranded T-DNA that is translocated into plants to induce tumor formation, but how VirC2 functions is unclear. Here, we report the 1.7-Å X-ray crystal structure of its trypsin-resistant C-terminal domain, VirC2$^{82–202}$, which reveals a form of the ribbon-helix-helix (RHH) DNA-binding fold contained within a single polypeptide chain. DNA-binding assays and mutagenesis indicate that VirC2 uses this RHH fold to bind double-stranded DNA but not single-stranded DNA. Mutations that severely affect VirC2 DNA binding are highly deleterious for both T-DNA transfer into yeast and the virulence of *A. tumefaciens* in different plants including *Nicotiana glauca* and *Kalanchoe daigremontiana*. These data suggest that VirC2 enhances T-DNA transfer and virulence through DNA binding with its RHH fold. The RHH fold of VirC2 is the first crystal structure representing a group of predicted RHH proteins that facilitate endonucleolytic processing of DNA for horizontal gene transfer.**

crystal structure | Ti plasmid | DNA processing

Tumor-inducing (Ti) plasmid-carrying *Agrobacterium tumefaciens* is the causative agent of crown gall disease that leads to tumor formation in a wide range of plants. During the *A. tumefaciens* infection, a part of the Ti plasmid called T-DNA is translocated in a single-stranded mode (called transferred strand or T strand) into plant cells in a process similar to bacterial conjugation (1, 2). After translocation, T-DNA integrates into the plant genome. Expression of genes located in the T-DNA transforms the host cell into a tumor cell, resulting in tumor (crown gall) formation. Because of its ability to transfer DNA from bacteria to plants, the T-DNA transfer system of *A. tumefaciens* is the most common DNA delivery tool for genetic engineering of plants.

There are two major classes of Ti plasmids characterized by different opine synthesis genes in their T-DNA, which are defined as nopaline- and octopine-type Ti plasmids (1, 2). All Ti plasmids contain two separate regions, the T-DNA and the virulence (*vir*) region, which contribute to tumor formation. The T-DNA region contains phytohormone and opine synthesis genes that are expressed in plant cells. The T-DNA region is flanked by two 25-bp sequences designated left and right borders, which are highly homologous and are oriented in a direct repeat (3). Outside the T-DNA region, the *vir* region encodes proteins involved in T-DNA processing and translocation (4–6). The *vir* regions of different types of Ti plasmids are conserved in sequence and are functionally cross-reactive (7, 8).

The right border of the T-DNA, which is recognized by the VirD1–VirD2 endonuclease complex, is essential for T-DNA transfer (9). VirD2, the border-specific endonuclease, nicks the lower strand within the right border with the help of VirD1 (10). Similar to the relaxase in bacterial conjugation systems, VirD2 remains covalently bound to the 5′ end of the T strand through its tyrosine residue (Tyr-29) (11, 12). T-DNA processing is completed by a second VirD2 nick on the same strand within the left border. The T strand–VirD2 complex is translocated to the plant cell by a *vir* region-encoded type IV secretion apparatus consisting of VirD4 and 11 different VirB proteins (13). The single-stranded DNA-binding protein VirE2 is transported from *A. tumefaciens* by the same type IV secretion apparatus but independently of T-DNA, and in the host cell virE2 coats the T strand to protect it against nucleolytic degradation (14–16). Both VirD2 and VirE2 have nuclear localization signals that can facilitate nuclear import of the T strand (14, 17–21). In the nucleus, the T-DNA integrates into the host genome at random positions (22), and the genes carried by the T-DNA express enzymes mediating the production of plant hormones in the plant cell to cause tumor formation (23, 24).

Outside the T-DNA region, a *cis*-element designated the "overdrive" sequence near the right border is required for efficient T-DNA transfer (25, 26). VirC1 specifically binds to overdrive, and disruption of *virC1* results in a loss of virulence similar to deletion of overdrive, suggesting that VirC1 functions through binding to overdrive (27, 28). VirC1 belongs to the ParA/MinD ATPase superfamily containing a deviant Walker A box that is required for border nicking and for localizing the T-DNA to the cell pole (29, 30). VirC1 also interacts with all other known T-DNA-processing proteins, including VirD1, VirD2, and VirC2, and thus it was proposed that VirC1 nucleates relaxosome assembly at T-DNA borders and spatially positions the T-DNA at the cell pole (29).

VirC2, a 202-aa cytoplasmic protein, is the only other protein encoded by the *virC* operon. VirC2 from octopine- and nopaline-types of Ti plasmids are cross-reactive and share high amino acid sequence identity (31) [Fig. 1*A* and supporting information (SI) Table S1]. Disruption of either *virC1* or *virC2* in *A. tumefaciens* exhibits the same phenotype with reduced T-DNA processing and attenuated virulence, suggesting that VirC2 works in a pathway similar to that of VirC1 (29, 31). Another study suggested that VirC2 plays a role in correct T-DNA processing because it is required for single-copy T-DNA integration in fungi (32). However, little is known about how VirC2 performs these functions.

The VirC2 function in T-DNA processing resembles the function of a family of DNA-binding proteins accessory to endonucleases for ssDNA processing during bacterial conjugation (33); however, so far no evidence has been obtained to suggest that VirC2 actually interacts with DNA. To gain insight

Author contributions: J.L., P.J.J.H., and J.N.M.G. designed research; J.L. and A.d.D.-R. performed research; J.L., P.J.J.H., and J.N.M.G. analyzed data; and J.L., A.d.D.-R., P.J.J.H., and J.N.M.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2RH3).

[1]To whom correspondence may be addressed. E-mail: p.j.j.hooykaas@biology.leidenuniv.nl or mark.glover@ualberta.ca.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812199106/DCSupplemental.

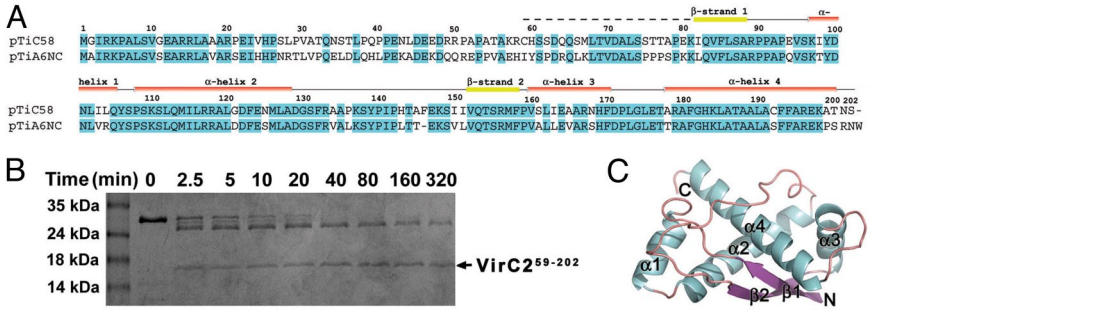



**Fig. 1.** Amino acid sequence, limited proteolysis, and overall structure of VirC2. (*A*) Alignment of the amino acid sequences of VirC2 from nopaline-type Ti plasmid pTiC58 (GenBank accession no. CAA68596) and octopine-type Ti plasmid pTiA6NC (GenBank accession no. NP_059811). Secondary structural elements were obtained from the crystal structure of VirC2$^{82–202}$. Connecting loops are represented by lines. The dotted line corresponds to the fragment with no electron density observed in VirC2$^{59–202}$ electron density map. (*B*) Limited proteolysis of $His_6$-tagged VirC2. Samples were assayed after the indicated time of trypsin digestion. (*C*) Overall structure of VirC2$^{82–202}$.

into how VirC2 functions, we used limited proteolysis to map out a stably folded C-terminal domain of VirC2 from residue 82 to 202 (VirC2$^{82–202}$). We crystallized VirC2$^{82–202}$ and solved its 3D structure at 1.7-Å resolution. The structure resembles the ribbon–helix–helix (RHH) DNA-binding fold commonly formed by homodimerization of transcription factors in the Arc/Met superfamily. Mutational, biochemical, and in vivo experiments indicate that VirC2 binds DNA through its RHH fold to enhance T-DNA transfer and virulence during *A. tumefaciens* infection.

## Results

**Limited Proteolysis and Crystallization of VirC2.** The *virC2* gene was cloned from pTiC58ΔT-DNA in a $His_6$-tagged version and was overexpressed in *Escherichia. coli* (*SI Materials and Methods*). Because attempts to crystallize full-lengthVirC2 were unsuccessful, we used limited proteolysis to identify stably folded domains of VirC2, which revealed a trypsin-resistant fragment of ≈16 kDa (Fig. 1*B*). MALDI-TOF mass spectrometry revealed its mass to be 15,817 (±5) Da, corresponding to the molecular mass of a VirC2 tryptic fragment from residue 59 to the C terminus of the protein (residue 202) (Fig. 1*A*). The amino acid sequence of this fragment was further confirmed by liquid chromatography-tandem MS. We found that after ≈3 months of storage at 4 °C, most of native VirC2$^{59–202}$ degraded to a smaller tryptic fragment corresponding to residues 82–202, probably because of proteolysis by the trace amount of trypsin left in the VirC2$^{59–202}$ preparation (*SI Materials and Methods*).

We produced crystals of a VirC2$^{59–202}$ selenomethionine (SeMet) derivative, which diffracted to 2.9-Å resolution under synchrotron X-ray radiation (Table S2). We also purified VirC2$^{82–202}$ from the VirC2$^{59–202}$ stock and produced its crystals, which belong to the same $P3_1$ space group but have a smaller unit cell than VirC2$^{59–202}$ crystals. VirC2$^{82–202}$ crystals diffract to much higher resolution than VirC2$^{59–202}$ crystals (Table S2). Initial phases and a preliminary structure of SeMet-VirC2$^{59–202}$ at 2.9-Å resolution were obtained by single anomalous diffraction, from which the structure of VirC2$^{82–202}$ was solved by molecular replacement and refined at a maximum resolution of 1.7 Å. The N-terminal polypeptide chain from residues 59 to 81 was not traceable in the VirC2$^{59–202}$ electron density map and is likely disordered.

**Structure of VirC2$^{82–202}$ Resembles a RHH DNA-Binding Fold.** VirC2$^{82–202}$ exists as a monomeric domain with dimensions of ≈19 Å × 28 Å × 38 Å. The structure is ≈2-fold symmetric because of the tandem repeat of a β-strand–helix–helix structural motif (Fig. 1*C*). A Protein Data Bank (PDB) search using DALI (34) showed that VirC2$^{82–202}$ shares structural homology with the structural superfamily of Arc/MetJ repressors (Fig. 2*A*). Each β-strand–helix–helix motif in VirC2 structurally resembles a RHH motif, often found in bacterial DNA binding proteins such as the Arc repressor and CopG (35, 36). Interestingly, all members of the RHH family bind DNA through a 2-fold symmetric dimer of RHH motifs. In contrast, VirC2 adopts a similar fold by using two RHH motifs within a single polypeptide chain; a large linker (residues 129–151) connects the N- and C-terminal RHH structural repeats. The C-terminal RHH repeat of VirC2$^{82–202}$ exhibits higher structural similarity to a typical RHH fold compared with the N-terminal RHH repeat, with Arc repressor and CopG having the highest DALI search Z-score at 4.1 (Z <2.0 regarded structurally dissimilar). The C-terminal RHH superimposes on the Arc repressor RHH with an rmsd of 2.0 Å over 43 matching $C^\alpha$ atoms from 49 residues, whereas the N-terminal RHH superimposes less well with the Arc repressor (rmsd of 2.8 Å over 45 matching $C^\alpha$ atoms from 49 residues).

Although the sequences of RHH proteins are in general poorly conserved (average amino acid sequence identity of 15.4%) (37, 38), positions of residues constituting a hydrophobic core and a glycine that defines a turn connecting the two helices are conserved. Both RHH motifs of VirC2$^{82–202}$ show the conserved pattern of hydrophobic residues within the secondary structure elements that constitute the protein hydrophobic core (Fig. 2*B*); however, some variations exist in VirC2$^{82–202}$. The β-strand of the C-terminal half of VirC2$^{82–202}$ has only one of the three conserved hydrophobic residues, probably indicating that less hydrophobic force is needed to support the packing of this strand in a single-polypeptide RHH fold than in a dimeric RHH fold. The N-terminal half of VirC2$^{82–202}$ does not have the conserved glycine that defines a turn in many RHH proteins, explaining its lower structural homology to other RHH family members.

**VirC2 Binds DNA Through Its RHH Fold.** Proteins that enhance plasmid-nicking reactions typically bind to DNA near the endonuclease-binding site (33). We therefore asked whether the region adjacent to the T-DNA right border (Fig. 3*A*), which contains a conserved overdrive core sequence recognized by VirC1, could contain the binding site for VirC2. Binding of VirC2 to this region together with VirC1 could potentially facilitate VirD1–VirD2 endonuclease-catalyzed nicking within the right border. Electrophoretic mobility shift assays (EMSAs) using three 20-bp DNA fragments covering the sequence adjacent to the right border of the Ti plasmid pTiBo542 were used to test the DNA binding ability of purified full-length VirC2 (Fig. 3*A*). VirC2 shifted all three 20-bp DNA fragments but bound the OD20 fragment with ≈3-fold higher affinity (apparent $K_d$ ≈125 nM) than either BO20 or OR20 fragments (Fig. 3*B*). VirC2$^{82–202}$

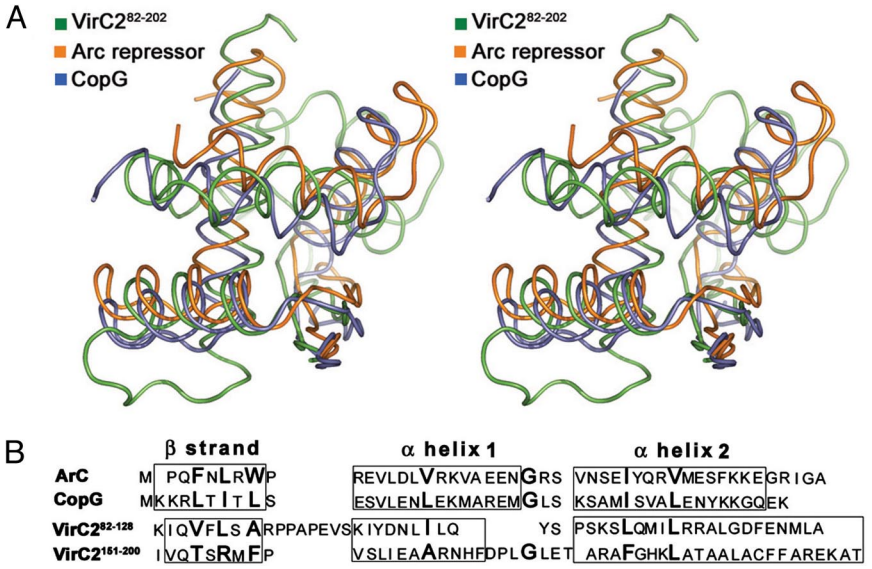
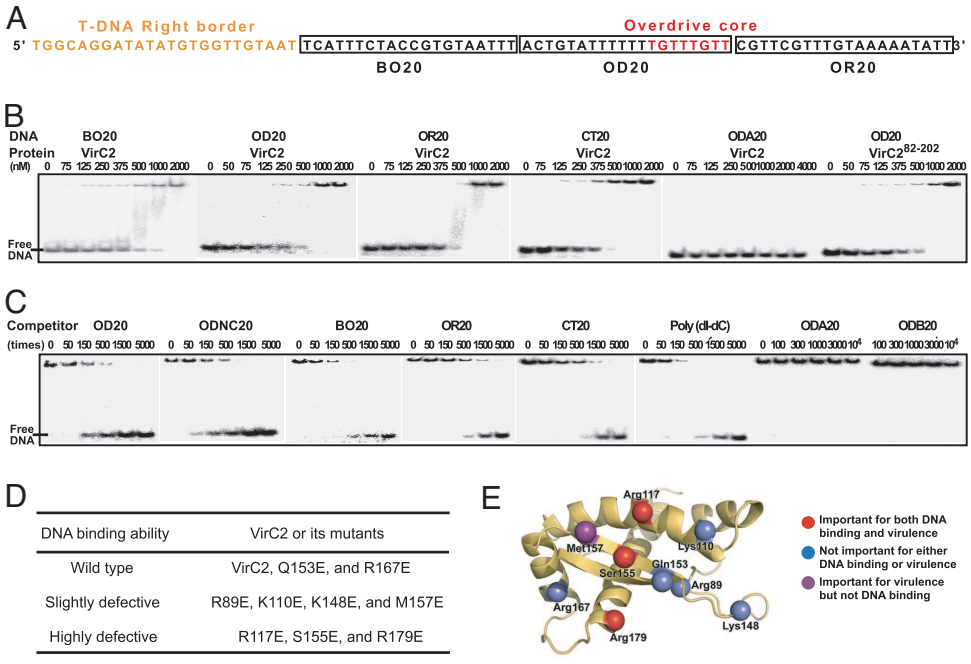



**Fig. 2.** Resemblance of VirC2$^{82–202}$ to RHH proteins. (*A*) Stereoview of the structural alignment of VirC2$^{82–202}$ with homodimers of RHH superfamily proteins Arc repressor (PDB ID code) and CopG (PDB ID code 2CPG). Secondary structure elements are labeled as in *B*. (*B*) Structure-based sequence alignment of Arc repressor, CopG, VirC2$^{82–128}$, and VirC2$^{151–200}$. Boxed sequences are equivalent secondary structures in an RHH motif as indicated above the sequence.

binds to OD20 with an affinity similar to that of full-length VirC2, suggesting that the RHH domain is responsible for the DNA binding of VirC2. VirC2 also binds a 20-bp control DNA of unrelated sequence (CT20) with an apparent $K_d$ of ≈250 nM, indicating a significant level of nonspecific DNA binding. A small amount of intermediate shifted species can be observed in every gel (most obvious in the gels of VirC2-BO20 and VirC2-OR20 binding), which could suggest nonspecific DNA binding or dissociation of protein–DNA complexes during electrophoresis (Fig. 3*B*). Both unlabeled OD20 and ODNC20 (OD20 without




| DNA binding ability | VirC2 or its mutants |
|---|---|
| Wild type | VirC2, Q153E, and R167E |
| Slightly defective | R89E, K110E, K148E, and M157E |
| Highly defective | R117E, S155E, and R179E |

**Fig. 3.** EMSA analysis of VirC2, VirC2$^{82–202}$, and full-length VirC2 mutants. (*A*) Nucleotide sequence near the right border of Ti plasmid pTiBo542 (GenBank accession no. NC_010929). The T-DNA right border is highlighted orange, and the 8-nucleotide overdrive core sequence highly conserved among Ti plasmids is highlighted red. Boxed sequences correspond to some of the 20-bp DNA fragments used for EMSA and competition assays. (*B*) EMSA analysis of VirC2 and VirC2$^{82–202}$ binding to sequences near the right border. BO20 (border to overdrive), OD20 (overdrive), and OR20 (overdrive right) correspond to the 20-bp DNA indicated in *A*. ODA20 is the 20-nucleotide top strand of OD20. CT20 (control) is used as a control with a 20-bp sequence within the T-DNA region that is presumably not a potential VirC2-binding site (Table S1). The numbers above each gel indicate increasing concentrations of a full-length VirC2 or VirC2$^{82–202}$ as indicated. (*C*) Competition assays. Increasing mass ratios of competitor to $^{32}$P-labeled OD20 are indicated above each gel. ODNC20 (overdrive without core) is equivalent to OD20 except that its conserved 8-nucleotide core was mutated (Table S1). ODA20 and ODB20 are the top and bottom strands of OD20, respectively. Poly(dI·dC) is used as a nonspecific double-stranded (ds) DNA competitor. (*D*) Summary of the DNA-binding ability of different full-length VirC2 mutants as determined by EMSA (Fig. S1). (*E*) Locations and effects of mutated residues in the structure of VirC2$^{82–202}$.

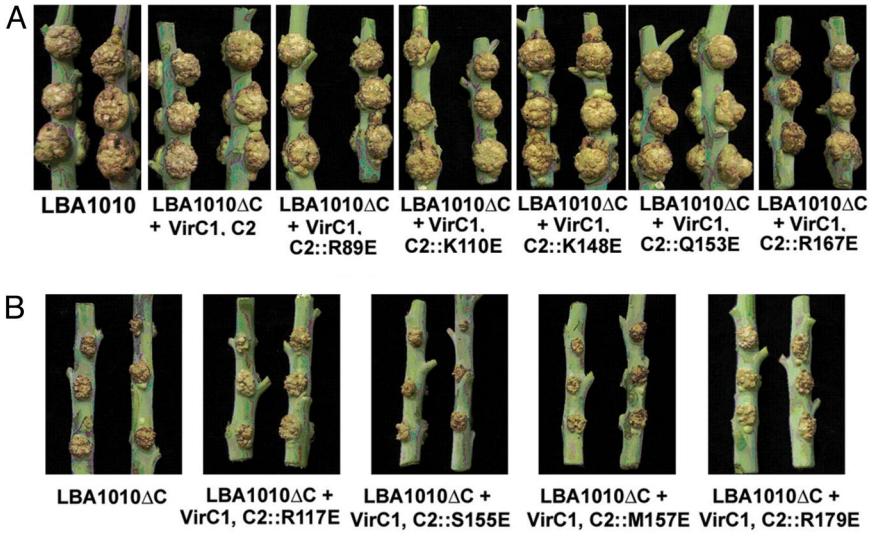



**Fig. 4.** Effects of *virC2* mutations on *A. tumefaciens* virulence in *N. glauca*. (*A*) Mutations that have no significant effect on tumor formation. (*B*) Mutations that affect tumor formation.

the conserved overdrive core sequence) can efficiently compete the VirC2–OD20 complex (Fig. 3*C*), indicating that VirC2–OD20 binding is not determined by the conserved overdrive core sequence. In comparison, ≈3-fold higher levels of BO20, OR20, CT20, or poly(dI·dC) dsDNAs are required to compete the VirC2–OD20 complex, indicating a modest degree of specificity of VirC2 for the OD20 DNA. Although VirC2 only shows a low level of specificity for sequences near the overdrive, these interactions absolutely require a double-stranded DNA target. No binding was observed to either strand of overdrive DNA (ODA20 or ODB20) as shown in both direct binding and competition assays (Fig. 3 *B* and *C*), indicating that VirC2 is a double-strand-specific DNA-binding protein.

RHH folds contact DNA backbone phosphates through the positively charged surfaces flanking the β-ribbon (35–37). Similar surface areas, which are formed by side chains of basic residues including Lys-110, Arg-117, and Arg-179, also exist in VirC2$^{82–202}$. To probe for sites that can affect VirC2 DNA binding, we tested the ability of a group of VirC2 missense mutants to bind DNA (Fig. 3 *D* and *E*). Lys-110, Arg-117, and Arg-179, which can potentially form electrostatic interactions with the DNA backbone phosphates, were mutated individually to glutamic acid. Similarly mutated residues include Arg-89 and Lys-148, which are in the loops extending from the β-strands, as well as Gln-153, Ser-155, and Met-157, which alternate with the conserved hydrophobic positions in the β-ribbon and can potentially contact DNA bases in the major groove. As a control, Arg-167, which is exposed on the protein surface but is distant from the potential DNA-contacting region, was also mutated. Each purified mutant protein was tested for DNA binding by EMSA using a 59-bp DNA (59BP; Table S1), corresponding to the region adjacent to the T-DNA right border of the nopaline-type plasmid pTiC58. VirC2 bound 59BP DNA with an affinity similar to the 20-bp DNAs (apparent $K_d$ ≈250 nM). However, the shifted species did not enter the gel, presumably because of the size of the complex (Fig. S1 *Left*). The Q153E and R167E mutants bound 59BP DNA and the wild type (Fig. 3*D* and Fig. S1 *Left*). The R89E, K110E, K148E, and M157E mutants are slightly defective for DNA binding (Fig. 3*D* and Fig. S1 *Center*), whereas the R117E, S155E, and R179E mutants are most defective (Fig. 3*D* and Fig. S1 *Right*). Thus, mutation of residues exposed on or near the antiparallel β-ribbon significantly impacts DNA binding, whereas mutation of residues distant from this surface does not (Fig. 3*E*). These results demonstrate that VirC2 binds DNA using the same surface used by other, dimeric RHH proteins.

**Function of VirC2 Mutants in Plant Tumor Formation and T-DNA Transfer to Yeast.** To determine whether the RHH fold of VirC2 is important for *A. tumefaciens* virulence, we constructed a *virC*-knockout strain LBA1010ΔC (*SI Materials and Methods*), which is inefficient in inducing tumors in plants (Fig. 4*B*). The above *virC2* mutations were introduced into a native pTiC58 *virC* operon cloned in an *E. coli-A. tumefaciens* shuttle vector to assess the ability of the mutants to complement LBA1010ΔC in inducing tumors. Strains with the R89E, K110E, K148E, Q153E, or R167E mutant caused similar levels of tumor formation in *Nicotiana glauca* as the wild-type strain LBA1010 or LBA1010ΔC complemented by wild-type VirC proteins (Fig. 4*A*), whereas strains with the R117E, S155E, M157E, or R179E mutant induced much smaller tumors, similar to LBA1010ΔC (Fig. 4*B*). To see whether the effects of different *virC* mutations on *A. tumefaciens* virulence are consistent in different plants, *Kalanchoe daigremontiana* was further tested, which yielded identical results (Fig. S1).

To be able to determine quantitatively the DNA transfer function of *virC2* mutants, we compared the frequency of T-DNA transfer into yeast from a helper strain containing wild-type *virC2* with that of similar helper strains containing instead the mutant *virC2* genes (Table 1). The *virC* knockout strain complemented by the wild-type VirC proteins can transfer T-DNA to yeast at a $2 \times 10^{-5}$ frequency, which is comparable with levels of T-DNA transfer from the VirC2 R89E, K110E, K148E, Q153E, and R167E mutant strains. Consistent with the results from the virulence assays, levels of T-DNA transfer from the M157E, R117E, S155E, and R179E mutant strains are 80- to >300-fold lower than that of the wild-type strain.

R117E, S155E, and R179E are most defective for DNA binding (Fig. 3 *D* and *E*), and strains with any of these mutations are highly defective for both virulence in plants and T-DNA transfer into yeast. Most of other VirC2 mutants are normal or only slightly defective for DNA binding and function normally for *A. tumefaciens* virulence in plants and T-DNA transfer into yeast. These correlations suggest that DNA binding through the RHH fold is required for VirC2 to function in T-DNA transfer and virulence.

**Table 1. Function of VirC2 mutants during T-DNA transfer into yeast**

| VirC2 | wt | R89E | K110E | R117E | K148E | Q153E | S155E | M157E | R167E | R179E |
|---|---|---|---|---|---|---|---|---|---|---|
| Frequency* | $2 \times 10^{-5}$ | $8 \times 10^{-6}$ | $5 \times 10^{-6}$ | $5 \times 10^{-7}$ | $2 \times 10^{-5}$ | $2 \times 10^{-5}$ | $8 \times 10^{-8}$ | $9 \times 10^{-7}$ | $7 \times 10^{-6}$ | $8 \times 10^{-8}$ |
| Normalized† | 1 | 0.38 | 0.24 | 0.02 | 1 | 1 | 0.004 | 0.04 | 0.36 | 0.004 |

*Frequency of T-DNA transfer into yeast (average of two independent experiments) was assayed by using an *A. tumefaciens virC* knockout strain containing one of the *virC2* mutant derivatives of pSDM3664.

†Transfer frequency was normalized to that of the cells containing pSDM3664 which expresses wild-type VirC proteins.

## Discussion

The X-ray crystal structure of the C-terminal domain (VirC2$^{82–202}$) of pTiC58 VirC2 reveals a pseudo-2-fold symmetric RHH fold in a single polypeptide chain (Fig. 1). Although bearing significant resemblance to dimeric RHH folds, VirC2$^{82–202}$ displays unique features in both its amino acid sequence and 3D structure (Fig. 2). Nevertheless, mutagenesis and DNA-binding assays indicate that the VirC2 RHH domain uses a positively charged surface centering on the $\beta$-ribbon to bind DNA, much like the way in which dimeric RHH domains bind double-stranded DNA (Fig. 3 and Fig. S1). The integrity of the DNA-binding surface is critical for VirC2 function in T-DNA transfer and *A. tumefaciens* virulence, suggesting an essential role for DNA binding in VirC2 function (Fig. 4, Table 1, and Fig. S2). Thus, VirC2 appears to belong to a family of DNA-binding proteins in bacterial conjugation systems, which bind near the origin of transfer to enhance the endonuclease-catalyzed nicking reaction that initializes single-stranded DNA transfer. In this family, extensive genetic and biochemical evidence has suggested that F plasmid TraY and R388 plasmid TrwA function through binding to cognate DNA with an RHH fold (39–41).

In the case of T-DNA transfer, VirC2 works together with VirC1 to stimulate the T-DNA right border nicking, which is catalyzed by the VirD1–VirD2 endonuclease complex (10, 27). VirC1 binds specifically to the overdrive sequence adjacent to the VirD1–VirD2-binding site and interacts with VirC2, coordinating assembly of the endonuclease complex on DNA (29). Our finding that VirC2 has a small preference for sequences adjoining overdrive is therefore consistent with its interactions with VirC1. From overdrive, VirC2 could potentially nucleate along DNA toward the VirD1–VirD2 complex at the right border to facilitate the nicking reaction. Similar to VirC1, VirC2 positions at the *A. tumefaciens* cell poles, where the single-stranded T-DNA is recruited by VirC1 for translocation (29). Although VirC2 does not bind single-stranded DNA efficiently, potential localized dsDNA regions generated by folding of the otherwise single-stranded T-DNA could be targeted by nonspecific VirC2 binding. Such interactions might silence gene expression from T-DNA, protect T-DNA from nucleolytic degradation, and/or perform some other functions for T-DNA transfer.

Unlike dimeric RHH proteins, the VirC2 RHH domain binds DNA with little sequence specificity. Unique features of the VirC2 DNA contact surface suggest that this protein likely contacts DNA in a significantly different manner than other RHH proteins. Residues alternating with the conserved hydrophobic positions in the $\beta$-ribbon form a protruding ridge that inserts into the DNA major groove to recognize specifically bases in the dimeric RHH proteins (35–37, 42) (Fig. 2*B*). In contrast, the $\beta$-ribbon of VirC2 is partially buried between the long helices 2 and 4 in VirC2$^{82–202}$ (Fig. S3*A*), which probably explains why residue Gln-153 on the buried side of the ribbon did not appear to be important for either DNA binding or the in vivo function of VirC2 (Fig. 3*E*). It is also unusual that VirC2$^{82–202}$ does not have basic residues alternating with the conserved hydrophobic positions in the $\beta$-ribbon, which make the central region of the potential base-contacting head unusually neutral (Fig. 2*B* and Fig. S3*B*). Other RHH folds have at least two residues (one on each strand) facing the DNA major groove (Fig. 2*B*) (38). The only positively charged side chain in the VirC2 ribbon is Arg-156. This residue is at a conserved hydrophobic position and is sequestered away from its potential DNA-binding face through salt-bridging interactions with Glu-124.

Interestingly, the mutation of a hydrophobic residue at one the major groove-facing positions, Met-157 to glutamic acid, only slightly reduces DNA-binding affinity but causes profound defects in virulence and T-DNA transfer. Other mutations that exhibit similar DNA-binding defects were tolerated in virulence and T-DNA transfer assays (Figs. 3*E* and 4*A*, Table 1, and Fig. S1 *Center*). Met-157 is located on the half of the $\beta$-ribbon that protrudes out of the DNA backbone-contacting surface (Fig. 3*E*). Although these large hydrophobic residues are seldom used for sequence-specific interactions with DNA via the major groove (43), they are commonly used by architectural, minor groove-binding proteins. For example, the minor groove-binding proteins TBP, SRY, LEF-1, and IHF all have hydrophobic residues that insert between adjacent DNA base pairs to distort the DNA double helix (44). Because VirD2 requires a single-stranded region for its endonuclease activity (12), a distorted or destabilized DNA region adjacent to the T-DNA right border could facilitate T-DNA processing by the VirD1–VirD2 complex. A similar molecular-wedge function has been suggested for MobC to facilitate the MobA relaxase in nicking during conjugative DNA transfer of plasmid R1162 (45).

Altogether, the crystal structure of the VirC2 RHH fold and its importance in *A. tumefaciens* virulence reveal a potential target for drug design to control crown gall disease. The involvement of the VirC2 RHH fold in DNA binding has shed light on the mechanism of T-DNA processing during *A. tumefaciens* infection. The structure of the VirC2 RHH fold not only serves as a starting point for the investigation of VirC2 cooperation with VirC1 in DNA binding and processing, but also will help to reveal common principles in DNA-processing mechanisms in other horizontal DNA transfer systems.

## Materials and Methods

Detailed descriptions of materials and methods, including strains, plasmids, oligonucleotides, and various assays used in this work, are presented in *SI Materials and Methods* and Tables S1 and S2. Standard techniques were used to perform mutagenesis and construct plasmids and strains. Proteins were expressed in *E. coli* and purified by affinity chromatography and size exclusion chromatography. Crystals were obtained by the vapor diffusion method, and the structure was solved by single-wavelength anomalous diffraction by using selenomethionyl VirC2$^{59–202}$. VirC2 DNA binding was characterized by EMSAs. The in vivo functions of various VirC2 mutants were characterized by virulence assays and yeast transformation assays.

**ACKNOWLEDGMENTS.** We thank Drs. Laura S. Frost and Lois M. Banta for helpful discussion, Dr. Michael Deyholos (University of Alberta) for providing *A. tumefaciens* strain GV3101::pMP90, Dr. A. Vergunst (French National Institute for Health and Medical Research) for providing vector pIN61, and Drs. Ross Edwards and Jonathan Parrish for assistance during X-ray diffraction data collection and processing. This work was supported by Canadian Institutes for Health Research (CIHR). The Alberta Synchrotron Institute (ASI) synchrotron access program is supported by grants from the Alberta Science and Research

Authority (ASRA) and the Alberta Heritage Foundation for Medical Research (AHFMR). J.L. is supported by fellowships from AHFMR, Canadian Foundation for Infectious Diseases (CFID), and CIHR. J.N.M.G. acknowledges support from the Howard Hughes International Scholar program. P.J.J.H. is supported by a TOP grant from Chemical Sciences of the Netherlands Organization for Scientific Research.

1. Gelvin SB (2003) Agrobacterium-mediated plant transformation: The biology behind the ''gene-jockeying'' tool. *Microbiol Mol Biol Rev* 67:16–37.
2. Bevan MW, Chilton MD (1982) T-DNA of the *Agrobacterium* Ti and Ri plasmids. *Annu Rev Genet* 16:357–384.
3. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci USA* 79:6322–6326.
4. Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. *Annu Rev Plant Physiol Plant Mol Biol* 51:223–256.
5. Klee HJ, Gordon MP, Nester EW (1982) Complementation analysis of *Agrobacterium tumefaciens* Ti plasmid mutations affecting oncogenicity. *J Bacteriol* 150:327–331.
6. Ooms G, Klapwijk PM, Poulis JA, Schilperoort RA (1980) Characterization of *Tn904* insertions in octopine Ti plasmid mutants of *Agrobacterium tumefaciens. J Bacteriol* 144:82–91.
7. Hooykaas PJ, Hofker M, den Dulk-Ras H, Schilperoort RA (1984) A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid, and an Ri plasmid by complementation analysis of *Agrobacterium tumefaciens* mutants. *Plasmid* 11:195–205.
8. Engler G, et al. (1981) Physical mapping of DNA base sequence homologies between an octopine and a nopaline Ti plasmid of *Agrobacterium tumefaciens. J Mol Biol* 152:183–208.
9. Wang K, Herrera-Estrella L, Van Montagu M, Zambryski P (1984) Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* 38:455–462.
10. Yanofsky MF, et al. (1986) The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell* 47:471–477.
11. Vogel AM, Das A (1992) The *Agrobacterium tumefaciens virD3* gene is not essential for tumorigenicity on plants. *J Bacteriol* 174:5161–5164.
12. Pansegrau W, Schoumacher F, Hohn B, Lanka E (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: Analogy to bacterial conjugation. *Proc Natl Acad Sci USA* 90:11538–11542.
13. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* 59:451–485.
14. Citovsky V, Wong ML, Zambryski P (1989) Cooperative interaction of *Agrobacterium* VirE2 protein with single-stranded DNA: Implications for the T-DNA transfer process. *Proc Natl Acad Sci USA* 86:1193–1197.
15. Das A (1988) *Agrobacterium tumefaciens* virE operon encodes a single-stranded DNA-binding protein. *Proc Natl Acad Sci USA* 85:2909–2913.
16. Vergunst AC, et al. (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290:979–982.
17. Citovsky V, Zupan J, Warnick D, Zambryski P (1992) Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science* 256:1802–1805.
18. Zupan JR, Citovsky V, Zambryski P (1996) *Agrobacterium* VirE2 protein mediates nuclear uptake of single-stranded DNA in plant cells. *Proc Natl Acad Sci USA* 93:2392–2397.
19. Herrera-Estrella A, Van Montagu M, Wang K (1990) A bacterial peptide acting as a plant nuclear targeting signal: The amino-terminal portion of Agrobacterium VirD2 protein directs a $\beta$-galactosidase fusion protein into tobacco nuclei. *Proc Natl Acad Sci USA* 87:9534–9537.
20. Howard EA, Zupan JR, Citovsky V, Zambryski PC (1992) The VirD2 protein of *A. tumefaciens* contains a C-terminal bipartite nuclear localization signal: Implications for nuclear uptake of DNA in plant cells. *Cell* 68:109–118.
21. Ziemienowicz A, Merkle T, Schoumacher F, Hohn B, Rossi L (2001) Import of *Agrobacterium* T-DNA into plant nuclei: Two distinct functions of VirD2 and VirE2 proteins. *Plant Cell* 13:369–383.
22. Alonso JM, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana. Science* 301:653–657.
23. Joos H, et al. (1983) Genetic analysis of T-DNA transcripts in nopaline crown galls. *Cell* 32:1057–1067.
24. Ooms G, Hooykaas PJ, Moolenaar G, Schilperoort RA (1981) Grown gall plant tumors of abnormal morphology, induced by *Agrobacterium tumefaciens* carrying mutated octopine Ti plasmids: Analysis of T-DNA functions. *Gene* 14:33–50.
25. Peralta EG, Hellmiss R, Ream W (1986) Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. *EMBO J* 5:1137–1142.
26. van Haaren MJ, Sedee NJ, de Boer HA, Schilperoort RA, Hooykaas PJ (1989) Mutational analysis of the conserved domains of a T-region border repeat of *Agrobacterium tumefaciens. Plant Mol Biol* 13:523–531.
27. Toro N, et al. (1989) The *Agrobacterium tumefaciens virC1* gene product binds to overdrive, a T-DNA transfer enhancer. *J Bacteriol* 171:6845–6849.
28. Toro N, Datta A, Yanofsky M, Nester E (1988) Role of the overdrive sequence in T-DNA border cleavage in *Agrobacterium*. *Proc Natl Acad Sci USA* 85:8558–8562.
29. Atmakuri K, Cascales E, Burton OT, Banta LM, Christie PJ (2007) *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *EMBO J* 26:2540–2551.
30. Koonin EV (1993) A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol* 229:1165–1174.
31. Close TJ, et al. (1987) Molecular characterization of the *virC* genes of the Ti plasmid. *J Bacteriol* 169:2336–2344.
32. Michielse CB, Ram AF, Hooykaas PJ, van den Hondel CA (2004) *Agrobacterium*-mediated transformation of *Aspergillus awamori* in the absence of full-length VirD2, VirC2, or VirE2 leads to insertion of aberrant T-DNA structures. *J Bacteriol* 186:2038–2045.
33. Lanka E, Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* 64:141–169.
34. Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* 233:123–138.
35. Raumann BE, Rould MA, Pabo CO, Sauer RT (1994) DNA recognition by $\beta$-sheets in the Arc repressor-operator crystal structure. *Nature* 367:754–757.
36. Somers WS, Phillips SE (1992) Crystal structure of the *met* repressor–operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* 359:387–393.
37. Gomis-Ruth FX, et al. (1998) The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J* 17:7404–7415.
38. Schreiter ER, Drennan CL (2007) Ribbon–helix–helix transcription factors: Variations on a theme. *Nat Rev Microbiol* 5:710–720.
39. Lum PL, Schildbach JF (1999) Specific DNA recognition by F factor TraY involves $\beta$-sheet residues. *J Biol Chem* 274:19644–19648.
40. Bowie JU, Sauer RT (1990) TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J Mol Biol* 211:5–6.
41. Moncalian G, de la Cruz F (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* 1701:15–23.
42. Knight KL, Sauer RT (1989) DNA binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. *Proc Natl Acad Sci USA* 86:797–801.
43. Luscombe NM, Laskowski RA, Thornton JM (2001) Amino acid–base interactions: A three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res* 29:2860–2874.
44. Bewley CA, Gronenborn AM, Clore GM (1998) Minor groove-binding architectural proteins: Structure, function, and DNA recognition. *Annu Rev Biophys Biomol Struct* 27:105–131.
45. Zhang S, Meyer R (1997) The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. *Mol Microbiol* 25:509–516.