Structure
Article

Variable Internal Flexibility Characterizes the Helical Capsid Formed by Agrobacterium VirE2 Protein on Single-Stranded DNA

Tanmay A.M. Bharat,[1,6,8] David Zbaida,[1] Miriam Eisenstein,[2] Ziv Frankenstein,[3] Tevie Mehlman,[4] Lev Weiner,[2] Carlos Oscar S. Sorzano,[7] Yoav Barak,[2] Shira Albeck,[5] John A.G. Briggs,[6] Sharon G. Wolf,[2] and Michael Elbaum[1,*]
[1]Department of Materials and Interfaces
[2]Chemical Research Support
[3]Department of Structural Biology
[4]Biological Research Support
[5]Israel Structural Proteomics Center
Weizmann Institute of Science, Rehovot 76100, Israel
[6]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany
[7]Biocomputing Unit, Centro Nacional Biotecnología (CSIC), c/Darwin, 3 Campus University Autónoma, Cantoblanco, 28049 Madrid, Spain
[8]Present address: Structural Studies Division, MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK
*Correspondence: michael.elbaum@weizmann.ac.il
http://dx.doi.org/10.1016/j.str.2013.04.027

SUMMARY

Agrobacterium is known for gene transfer to plants. In addition to a linear ssDNA oligonucleotide, *Agrobacterium tumefaciens* secretes an abundant ssDNA-binding effector, VirE2. In many ways VirE2 adapts the conjugation mechanism to transform the eukaryotic host. The crystal structure of VirE2 shows two compact domains joined by a flexible linker. Bound to ssDNA, VirE2 forms an ordered solenoidal shell, or capsid known as the T-complex. Here, we present a three-dimensional reconstruction of the VirE2-ssDNA complex using cryo-electron microscopy and iterative helical real-space reconstruction. High-resolution refinement was not possible due to inherent heterogeneity in the protein structure. By a combination of computational modeling, chemical modifications, mass spectroscopy, and electron paramagnetic resonance, we found that the N-terminal domain is tightly constrained by both tangential and longitudinal links, while the C terminus is weakly constrained. The quaternary structure is thus rigidly assembled while remaining locally flexible. This flexibility may be important in accommodating substrates without sequence specificity.

INTRODUCTION

Agrobacterium tumefaciens is a common soil bacterium with a unique capability to transform its plant hosts genetically. In nature it causes the "crown gall" disease, while in recent decades it has been adapted widely for plant genetic engineering and biotechnology (Gelvin, 2012). The basic mechanism of gene transfer appears to be similar to bacterial conjugation (Christie et al., 2005), also based on a type IV secretion system (T4SS), yet in this case the host is a eukaryote rather than another bacterial cell. A primary adaptor of conjugation to the plant host is a single-stranded DNA (ssDNA) binding protein that is expressed in and secreted from the bacterium: the virulence protein VirE2 (Christie et al., 1988; Citovsky et al., 1989; Sen et al., 1989). VirE2 molecules coat the single-stranded transferred DNA strand (T-strand), creating a helical structure known as the "T-complex." VirE2 is essential for plant cell transformation and is thought to play several roles, in particular protection of the T-strand from cytosolic nucleases, and recruitment of the necessary macromolecules for nuclear uptake. These roles are similar to those of an invading viral capsid. It has also been proposed that VirE2 acts as a translocon of the T4SS in providing a channel through the host plasma membrane (Dumas et al., 2001). Interestingly, however, VirE2 and the T-strand need not arrive from the same bacterium, leading to the paradigm that the two components meet in the cytoplasm to form a T-complex (Citovsky et al., 1992). Once inside the nucleus, VirE2 is removed from the T-strand (Anand et al., 2012; Tzfira et al., 2004; Zaltsman et al., 2010), freeing the oligonucleotide for further processing stages of second-strand synthesis, transient transcription, and ultimately genomic integration.

The interaction of VirE2 with ssDNA has been studied extensively in vitro (Christie et al., 1988; Citovsky et al., 1989; Sen et al., 1989). A saturating stoichiometry of 10:1 wt:wt (protein:DNA) was first established by gel shift assays (Citovsky et al., 1989). With a molecular weight of 63.5 kDa, this indicates a ratio of approximately one protein unit per 20 bases. Next, the mass per length in the T-complex filament was determined by scanning transmission electron microscopy (STEM) measurements (Citovsky et al., 1997), and later a three-dimensional (3D) structure of the VirE2-ssDNA complex in negative stain was obtained (Abu-Arish et al., 2004) by iterative helical real space reconstruction from transmission electron micrographs (Egelman, 2000). These studies revealed a structure wherein protein units are arranged in a spiral staircase surrounding a hollow core, i.e., a solenoid or "telephone cord" structure. Helical parameters determined by the STEM and the 3D reconstruction differed by approximately

one unit of VirE2 per turn of the helix, the former reporting an outer diameter of 12.8 nm with 3.4 protein units per turn, and the latter finding a diameter of 15.7 nm with 4.3 units per turn. In both cases the rise per turn was 5 nm. The studies found a nucleotide:protein stoichiometry of 16 and 19:1, respectively, in satisfactory agreement with the earlier biochemistry. In the negative staining 3D structure, three distinct subdomains could be discerned in each protein unit, although it was not possible to distinguish the boundaries between proteins. The position of the ssDNA could be deduced (by geometrical considerations) to run along the inner wall of the solenoid, where it would indeed be well protected from enzymatic attack.

Within the Agrobacterium, VirE2 is co-expressed with a small acidic protein, VirE1. Mutation of VirE1 inhibits secretion of VirE2, hence its identification as a secretion chaperone. The complex of VirE1 with VirE2 was expressed in *Escherichia coli* and the structure determined by X-ray diffraction (Dym et al., 2008). VirE2 was found to contain two folded domains between amino acids 112 and 517. The two domains are TIM barrel-like but with altered topology—$\beta\beta\alpha\alpha$ rather than $\alpha\beta\alpha\beta$. The fact that both domains have very similar structure suggests a gene duplication event, although they bear very little sequence homology to each other. A single unstructured linker of 10 amino acids connects the two domains. The C-terminal half of VirE1 forms a single α helix that binds both folded domains of VirE2, primarily via electrostatic interactions with basic amino acids. The flexibility of the VirE2 linker permits the protein to close like a clam shell around the rod-like VirE1. The missing 39 amino acids at the C terminus of VirE2 correspond closely to the arginine-rich secretion signal defined earlier by biochemical and genetic assays (Simone et al., 2001; Vergunst et al., 2000). The missing 111 residues at the N terminus were more enigmatic, and it could not be determined unambiguously if they had been partly truncated or were simply disordered. In examining the structure with DNA, however, three domains are distinctly seen, suggesting that the far N terminus might fold only in contact with the C terminus of the neighboring protein around the ring.

A structure-based understanding of interactions between the invading T-complex and proteins of the plant host cell requires a merge of the information from crystallography and EM to characterize the outer surface of the complex. Initial attempts to relate the high-resolution crystal structure of VirE2 from the VirE1-VirE2 complex to that of the negative staining EM map produced in complex with ssDNA were frustrated by the low resolution of the latter. It was nonetheless clear that while either of the two high-resolution domains could be rigid-body-fitted into the EM map, the second would protrude. This immediately suggested a rearrangement around the flexible region that links the two; in complex with ssDNA, the VirE2 molecule is more extended. A simple assembly model was proposed by which the N- and C-terminal contacts are critical to cooperative assembly of the ordered solenoid structure (Dym et al., 2008). In complex with VirE1, the VirE2 molecule is too compact to continue end-to-end interactions around the circumference of the helix, while in the absence of DNA it is too "open" to form the extended solenoid, possibly due to un-neutralized electrostatic repulsion between facing basic patches. A tentative docking arrangement put these basic VirE1-binding (and presumably DNA-binding) patches toward the center. This orientation left the nuclear localization signal (NLS) domains facing outward, where they would be available for interaction with host plant proteins.

The aim of the present work is to generate a docking model for the atomic structure within the EM envelope and to validate the model in a data-driven manner using lateral experiments. This required, first of all, improved resolution in the VirE2-ssDNA complex structure. For that reason we repeated the 3D structural study using cryo-electron microscopy (cryoEM) and real-space helical reconstruction. However, even with extensive variation of experimental conditions, data classification, and tuning of the reconstruction algorithms, we were unable to obtain a high-resolution map in which secondary structure elements could be unambiguously assigned. This suggested that structural heterogeneity might be intrinsic, i.e., that variability among the asymmetric units appears on a length scale comparable to details of the unit itself. Indeed it is not uncommon for helical complexes to resist high-resolution structure determination by single-particle-based image processing (Bharat et al., 2011; Desfosses et al., 2010). Therefore we turned to a hybrid approach combining extensive image classification, molecular modeling, cysteine modification, lysine cross-linking, mass spectroscopy, and electron paramagnetic resonance (EPR) analysis to confirm the docking results with independent information. While each such piece of evidence is clearly incomplete independently, the integrated structural biology approach provided a coherent insight to understand VirE2 assembly and stability even though high-resolution structural information could not be obtained directly. Taken together, the results describe a helical assembly stabilized by azimuthal N- to C-terminal contacts around the turn of the helix, with N- to N-domain contacts axially around the perimeter. This constrains the possible movement of the N-domain along two axes, while the C-domain is free to rotate around the axis between adjacent lateral N-domains. We can suggest that this allows for inherent flexibility of the filament, which could provide for accommodation of variable DNA sequences with different local rigidities, a key characteristic necessary for this bacterial-plant gene transformation process.

RESULTS

3D Reconstruction from CryoEM Data

VirE2-ssDNA complexes were prepared for cryoEM using 26-mer ssDNA substrates. Typical EM images appear in Figure 1A. Three-dimensional reconstructions were performed by iterative helical real-space reconstruction (IHRSR) (Egelman, 2000), similar to an earlier study that used negative staining and M13 ssDNA (Abu-Arish et al., 2004). Particle selection was easier using the short substrates because the filaments were longer and straighter than those formed on circular DNA, with fewer obvious breaks. Altogether, a total of 44,311 particles were collected for analysis in boxes of 98 × 98 pixels (scale 0.43 nm/pixel) at 90% overlap, representing a total length of T-complex of 1.87×10^5 nm. Particles were aligned rotationally and centered in-plane, recropped to 64 × 64 pixels, and then classified using the KerDenSOM method of Xmipp (Sorzano et al., 2004). A smaller data set of 8,019 particles was retained after this procedure, representing what appeared to be the highest quality data. A more permissive data set containing an additional 9,276 particles was also selected; these included

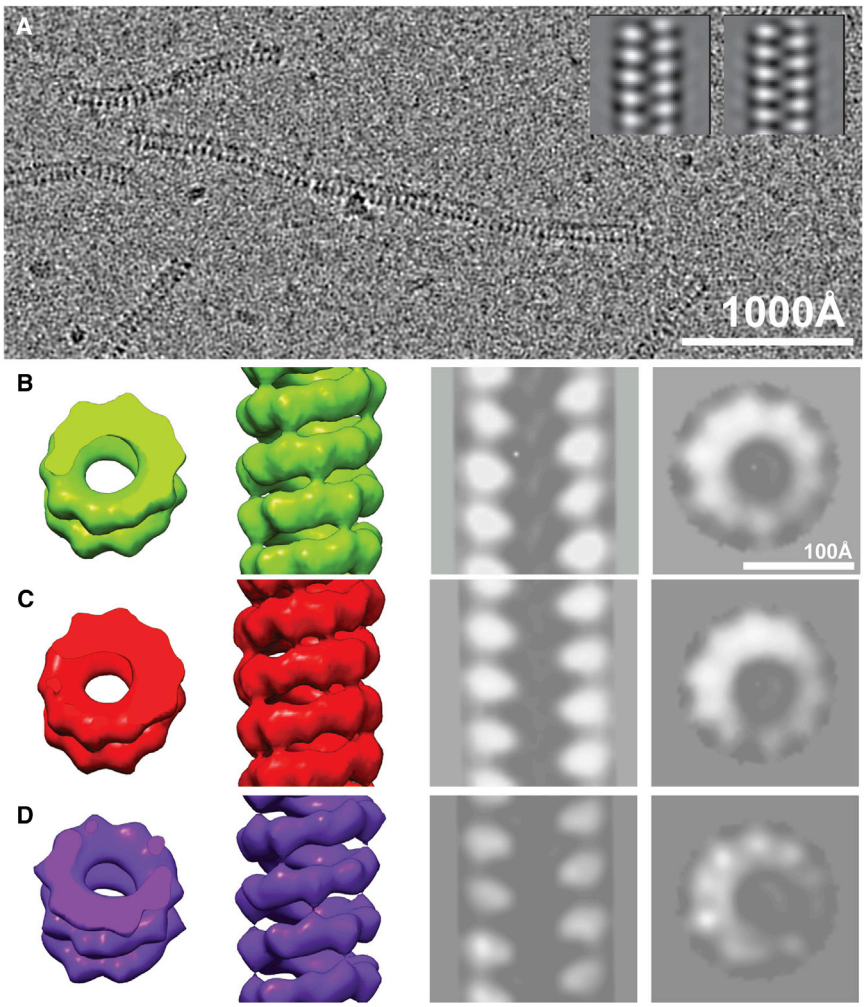

Figure 1. 3D Reconstruction from CryoEM Data Using Iterative Helical Real-Space Reconstruction

(A) Examples of raw data, showing extended linear complexes of VirE2 with short ssDNA. Insets show typical class averages.

(B) Isosurface representations of the reconstruction from the "strict" data set of 8,019 particles, as described in the text, in two different orientations (left). The same reconstruction is shown as a slice through the density in two orthogonal orientations along and perpendicular to the helical axis (right). See also Figures S1–S3.

(C) A comparative reconstruction from a more permissive data set (see Figure S1 for details).

(D) Reconstruction from exhaustively classified and selected data. The reconstruction was generated from only 280 raw particles.

particles displaying noticeable scalloping along the side. Code vectors representing the two selected sets are shown in Figure S1 (available online). We noted that the filaments were narrower than those obtained earlier in negative stain, with an outer diameter of approximately 13 nm instead of 15.7 nm. Preliminary reconstructions indicated a helical symmetry of 3.27 protein units per turn and rise of 4.8 nm per turn of the helix. This finding fit the earlier STEM data (Citovsky et al., 1997) better than the negative staining EM reconstruction and suggested alternative modes of complex assembly. After 42 cycles of refinement using the strict data set, the structure had converged to that shown in Figure 1B. Formal resolution was determined at 2.0 nm by Fourier shell correlation (FSC) between even-odd data set reconstructions (0.5 FSC criterion), shown in Figure S2. Inclusion of the scalloped particles in the permissive data set did not lead to any improvement in the resolution, as seen in Figure 1C. While features are similar qualitatively to the reconstruction from the strict set, the resulting reconstruction was smeared around the circumference (compare Figures 1B and 1C). Refinement of data using different helical parameters (4.25 or 5.25 subunits per turn) resulted in final volumes with a "stacked-disc" appearance that were discontinuous at realistic isosurface threshold levels. These symmetries were thus inconsistent with the raw data, which clearly indicates a single start helix with a Bessel order of 1 (Figure S3G).

As a further test for proper symmetry assignment, we allowed the final reconstruction to iterate five more cycles without intermediate imposition of symmetry (Penczek et al., 2011), i.e., without the averaging of symmetry-related voxels that is included in the standard IHRSR protocol. The search for helical symmetry parameters returned essentially unchanging values during this procedure. Moreover, symmetry-related averaging at the end of the five cycles yielded a structure very similar to the original reconstruction (seen in Figure S3). This gives confidence that the structure is not an artifactual result due to imposition of a spurious symmetry per cycle.

Many of the classes rejected from the strict data set showed a distinct scalloping along one or the other side of the structure that would suggest a tilting of filaments in the ice. Efforts to reconstruct the structure with assignment to multiple reference models that differed by out-of-plane tilts were not fruitful; we found that only the low tilt model was substantially populated. Therefore we took an alternative approach to ask whether the scalloping may be intrinsic. We prepared the VirE2-ssDNA complex on circular M13 ssDNA and deposited it on thin carbon (Figure 2). Images were recorded at high defocus to enhance visibility because high resolution was not the aim. A tomogram was collected, which confirmed that the filaments lay flat (Figure 2A). Particles were selected as described previously and a KerDenSOM classification was performed. A similar scalloping was observed. We then divided the code vectors into those that appeared symmetric side to side, i.e., flat, and those that were scalloped, suggesting tilt (Figure 2B), and remapped the assigned images back onto the original micrograph (Figure 2C). Had the filaments been tilted we should have seen consistent assignments along their length representing a slow variation of the tilt. Instead the assignments were mixed at a short scale,

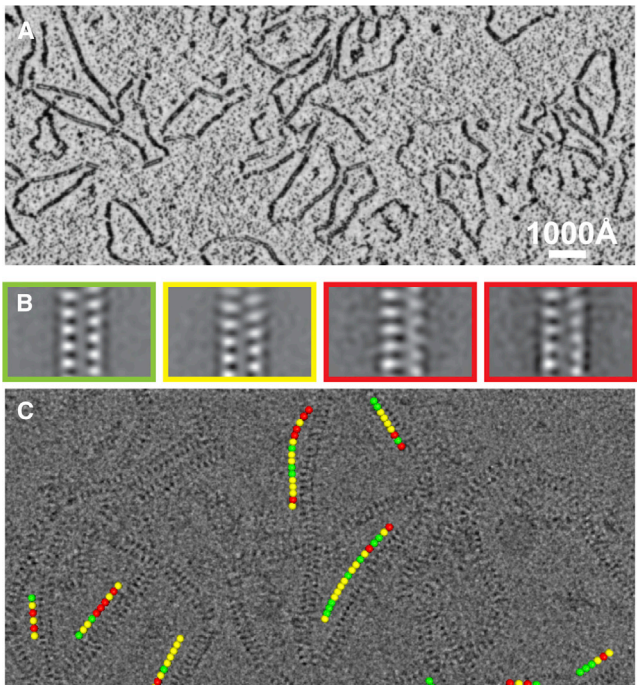

Figure 2. Evaluation of Scalloping for Tilt or Intrinsic Heterogeneity
Complexes of VirE2 with circular M13 ssDNA were prepared and deposited on thin carbon before blotting and plunging to liquid ethane for cryo-preservation. Both single-particle and tomographic analyses were performed.
(A) A tomographic reconstruction (central slice) shows that the protein-DNA rings lie in a plane and there is no significant tilting on this grid.
(B) Code vectors were selected that suggested cleanly "in-plane" (green), "marginally in plane" (yellow), and "tilted" (red) data according to the appearance of scalloped edges.
(C) Images corresponding to the above three categories were remapped onto the original micrograph. The mixing of categories locally within a single filament points to local heterogeneity rather than overall out-of-plane tilt as the origin of the scalloped edge.

confirming that the scalloping may originate from local heterogeneity rather than out-of-plane tilt.

We next attempted to improve the 3D resolution by further increasing the stringency in classification. To that end, we selected fewer code vectors for the final reconstruction, carried out subclassification of selected code vectors, and also carried out multimodel refinement to account for slightly differing helical symmetries, but ultimately we obtained very similar structures to that of Figure 1B. Classification efforts culminated in an extreme selection of only 280 particles out of the 44,311 in the original data set, by exhaustive search using the CL2D method (Sorzano et al., 2010), from which an iterative reconstruction was performed using the predetermined helical parameters (full details appear in the Supplemental Experimental Procedures). The result appears in Figure 1D and is again qualitatively very similar to that of Figure 1B. The fact that such a clearly recognizable structure could be obtained with so little data indicates that noise reduction is not the limiting factor in resolution improvement.

At this attainable resolution, the asymmetric unit in the cryoEM reconstruction is similar to the one obtained earlier from negatively stained data (compare Figures 3A and 3B). The outer diameters differ by 3.0 nm (Figures S4B and S4D), corresponding to 9.2 nm in circumference; this matches well to the extended length of the protein around the solenoid. This diameter difference is not a result of flattening of the negatively stained filaments. This was confirmed by subtomogram averaging of individual negatively stained T-complex filaments (Figure S4E), complementing a previous analysis of tetramers of VirE2 that presented themselves on the electron microscopy grid in an orientation corresponding to cross-sections of the reconstructed solenoid (Abu-Arish et al., 2004). Again there were three distinct sub-domains in each asymmetric unit, providing an additional confirmation of the structural data. Similar ring-to-ring connections appear between protein units at symmetry-related locations around the solenoid: units n and n+3, and n and n+4, respectively (Figures 3A and 3B). Such a result, a tighter spiral providing similar contacts, suggests internal flexibility of the protein subunits. Most likely the angle around the linker between the two major domains of VirE2 is more acute in the cryoEM structure with smaller diameter. Within either sample set, however, this flexibility could well result in a local structural heterogeneity that prevents the extraction of high-resolution information from averaged projections.

Molecular Modeling

The robust connectivity obtained by EM, in spite of limited resolution, suggested a repeated, regular organization of the VirE2 structure around the solenoid. Manual docking showed that either N or C-terminal domain of the high-resolution structure could potentially fit independently into either of the larger domains seen in the cryoEM envelope and so yielded little additional insight. Therefore a comprehensive approach of computational modeling was taken, requiring that both domains should fit the envelope in a manner consistent with their connectivity at the linker and without clash to each other or to domains of neighboring protein units around the solenoid. The modeling began with independent docking of each of the folded domains by exhaustive search, employing the FFT-based fitting procedure, *fit*PDB2EM, described in the Supplemental Experimental Procedures and Table S1. Then the intermediate results were filtered according to progressively stricter steric and biochemical criteria. In view of the limited resolution of the EM map we considered only the shape of the molecules in the computational fitting, disregarding electron density data except in the determination of the cryoEM envelope.

The slightly larger N-terminal domain (amino acids [aa] 112–341) was examined most extensively. Normal mode analysis was performed using the ElNemo server (Suhre and Sanejouand, 2004), and four confomers of each of the three lowest frequency modes were prepared, yielding a total of 13 models including the original unmodified structure. These appear in Figure 4A. Each of these was fit to the cryoEM envelope of approximately five helix turns, using *fit*PDB2EM with a rotational and translational search grid of 12.0° and 0.28 nm. Note that the effective search grid was finer because these parameters are incommensurate with the helical symmetry of the EM map, with rotation and axial translation of 109.1° and 1.47 nm, respectively. The C-terminal domain (aa 345–517) was independently fitted into the EM envelope using the same search parameters. However, only the unmodified conformer was used for this domain because some of the loops are not resolved in the X-ray structure. The combination

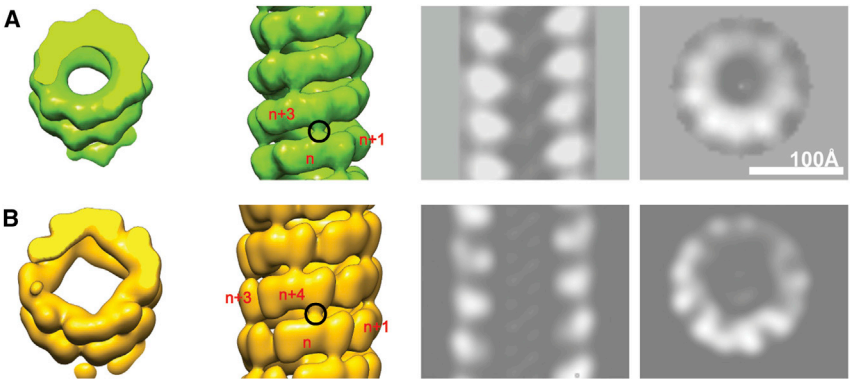

Figure 3. Comparison of the CryoEM and Previous Negative-Stain Reconstructions

(A) The same reconstruction as in Figure 1B, shown as an isosurface (left) and as a slice through the density in two orthogonal orientations along and perpendicular to the helical axis (right). See also Figure S3.

(B) Corresponding images of the earlier reconstruction from negatively stained data (Abu-Arish et al., 2004). Axial ring-to-ring contacts in both reconstructions are marked with a black circle (see isosurface representation). In the cryoEM reconstruction, every VirE2 monomer interacts with the third monomer along the helix, whereas in the negative-staining EM reconstruction, it interacts with the fourth. See also Figure S4.

of the 6,289 and 13,542 acceptable poses for the N and C termini, respectively, produced 85,165,638 models. These were selected based on specific structural criteria, first for absence of overlap between the two domains or between neighboring proteins (rather than subdomains) in the solenoidal arrangement and then for a maximum distance of 2.0 nm between the Cα atoms of residues 341 and 345. Finally, models in which single proteins spanned the solenoid axially rather than tangentially were rejected, as well as models in which the NLS motifs of the N-terminal domain faced into the central core. These were the only external criteria that entered the modeling. The remaining 7,221 models were scored and clustered as described in the Supplemental Experimental Procedures and divided into four sets according to the position of the N-terminal domain within the EM map. Representative models for the four sets appear in Figure S5. Set 1 was clearly dominant, representing 6,686 models or 93% of the total. Moreover, this set included all the high scoring models that consisted of high-scoring fits for both the folded N- and C-terminal domains evaluated independently. Representative models within set 1 appear in Figure 4B, with the N-terminal domain shown as a ribbon and the C-terminal domain as a wire diagram. Visual inspection within this set shows a strong convergence in placements of the N domain, with little constraint on the relative orientation of the C domain. Notably, the exposed and relatively immobile V-shaped loop of the N domain is located at the longitudinal junction between the solenoid turns. This loop contains in particular the residue Cys134 (Figure 5A), whose modification provided an illuminating corroboration of the docking model.

Cysteine Labeling, Lysine Cross-Linking, Mass Spectroscopy, and EPR

In an independent strategy to orient the crystal structure within the DNA-containing complex, we tested cysteine labeling as a measure of chemical availability. Nopaline-strain VirE2 contains three isolated cysteine residues: C134, C149, and C191, all in the folded N-terminal domain of aa 112–342 (see Figure 5A). We first attempted to detect cysteine modifications by mass spectrometry, using adaptations of the standard protocol for protein identification. This involves heating to 65°C with reduction by DTT, alkylation by iodoacetamide, and trypsin digestion, followed by electrospray ionization mass spectrometry. This procedure assists particularly in identification and sequencing of peptides containing cysteine residues. Data were analyzed by Mascot and MassMatrix software (Xu and Freitas, 2007); full details appear in the Supplemental Experimental Procedures and in Table S2.

Three protein forms of VirE2 were analyzed: (1) expressed and purified in isolation as soluble filamentous aggregates, (2) as a soluble heterodimer with VirE1, and (3) as a protein-ssDNA complex (the T-complex). In case 1 of isolated VirE2, all three cysteines were detected as carbamidomethyl derivatives. With VirE1-VirE2 heterodimers, only the C134 and C149 residues were modified and detected. In sharp contrast, none of the peptides containing cysteine residues were identified from the T-complex filaments (case 3 above). In this case the heating step had to be omitted because it disrupted the protein binding to DNA. (We verified that peptides containing IAA-modified C134 and C149 were also identified on VirE2, case 1, without heating.) These experiments implied that interaction with the DNA somehow protects or hides the three cysteines from reaction with the iodoacetamide reagent.

Cysteine-based modification was tested in addition using N-ethyl maleimide and, for the purpose of EPR spectroscopy, using the nitroxyl radical spin label MTSL S-(2, 2, 5, 5-tetramethyl-2, 5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate. For both reagents, robust modification of C134 was detected in cases 1 and 3, with modification of C149 detected depending on specific preparation conditions (described in Supplemental Experimental Procedures). Only in case 2 of soluble VirE1-VirE2 heterodimers were all three cysteines modified. Most strikingly, no cysteine modification was detected for any reagent when the VirE2 protein was in complex with ssDNA.

To test the docking model specifically via localization of the N-terminal domain, we performed a cross-linking assay using the lysine-lysine homo cross-linker BS^3 that can cross-link lysine side chains within 3 nm of each other (as measured from the polypeptide backbone). For completeness, we again investigated the three forms of VirE2 described above: isolated VirE2, VirE1-VirE2 heterodimer, and VirE2-ssDNA complex. The reacted protein samples were digested using chymotrypsin and fragments were analyzed by mass spectrometry. Results are shown in Table S3. Using the VirE1-VirE2 Protein Data Bank (PDB) structure as a guide, most of the cross-links were found to be intramolecular. Additional links were found in the VirE2 and VirE2-ssDNA samples between residues K219 or K222 and K298. The distance between these residues is too large to attribute to cross-links within the individual protein. Figure 5A

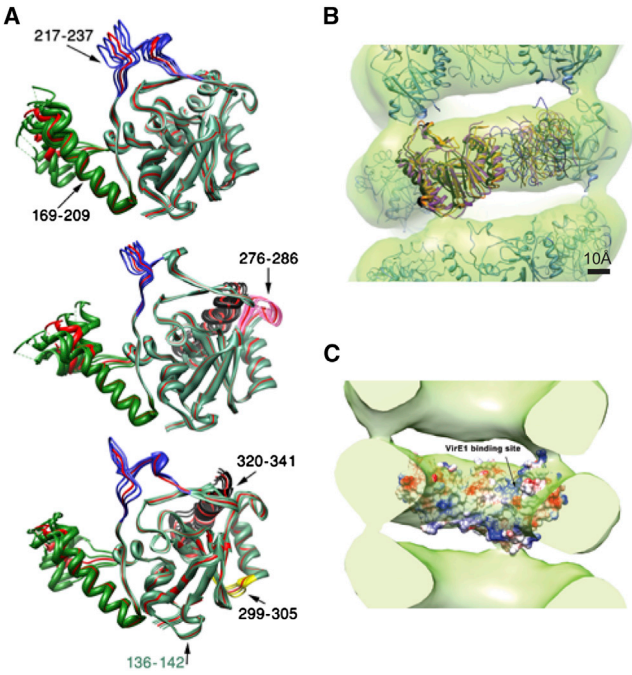

Figure 4. Docking of the High-Resolution PDB Structure of VirE2 into the Low-Resolution CryoEM Envelope

(A) Normal modes analysis shows high amplitude movements in the N-terminal domain of VirE2. Flexible loops and the large mobile arm consisting of residues 169–209 are highlighted (black text). In comparison, the loop consisting of residues 136–142 is relatively immobile (green text). See also Figure S5.

(B) The folded N-terminal domain (shown in ribbon form) fits tightly into the EM envelope with only minor variations due to the normal mode oscillations and the stepwise scanning employed in *fit*PDB2EM. In contrast, the C-terminal domain shown via the backbone line is poorly constrained in the modeling.

(C) Accessible surface representation of the inner surface cut-away shows predominantly positive (blue) charge, consistent with electrostatic binding to DNA.

shows that they lie within 3 nm between vertically adjacent protein units in the models of set 1 described in the previous section, where they straddle the location of Cys134 at the interface. The cross-linker does not further distinguish between the models of sets 1–4, however, as acceptable distances may be found in examples of every set. Interestingly, the same intermolecular links are found in isolated VirE2 of filamentous aggregate form. This suggests that the disordered filaments are quite similar locally to the complex with ssDNA; indeed, we reported earlier that a small fraction of such aggregates can be found in a solenoidal form even without addition of ssDNA (Frenkiel-Krispin et al., 2007).

We next examined the availability of the free SH groups in VirE2 and VirE2-ssDNA complexes for chemical modification with EPR using the disulfide-containing biradical bis (1-oxy-2,2,5,5-tetramethyl-3-imidazo-line-4-yl) disulfide (RSSR). The thiol-disulfide exchange reaction was monitored via the released monoradical R-SH as described in the Supplemental Experimental Procedures. Figure 5B shows the increase in EPR signal representing the reaction between the label and the VirE2 protein. The reaction kinetics follows a single exponential on the time scale examined (green curve). The signal did not reach a full plateau, suggesting the possibility of an additional slower reaction, although longer measurements were not possible due to decomposition of the reagent in aqueous medium. In contrast, the VirE2 complex with ssDNA was not reactive and the signal obtained is similar to that of the buffer alone.

We further addressed the stoichiometry of VirE2 reaction using the nitroxyl radical spin label MTSL. (Spectra justifying the following conclusions appear in the Supplemental Experimental Procedures.) In the case of VirE2 alone, double integration of the EPR signal yielded a molar ratio of 1.3 ± 0.15 MTSL per molecule. This is consistent with the previous suggestion of one reaction that takes place to completion and a second slower one that remains partially reacted, as well as with the mass spectrometry finding of MTSL-modified peptides containing C134 and C149. The stoichiometry obtained by similar analysis for VirE2-VirE1 heterodimers was 2.8 ± 0.35, again consistent with the mass spectroscopy showing all three cysteines modified. The most striking result was that for VirE2-ssDNA complexes. Similar to the biradical reagent and consistent with the mass spectroscopy, no reaction with MTSL was detected. Thus it appears that C134 is buried and therefore unreactive in the complex of VirE2 with ssDNA. The models of set 1 all place this residue at the ring-to-ring interface between (vertically) adjacent N-terminal domains, where they would indeed be buried at the interface in the VirE2-ssDNA complex.

Given that the VirE2-ssDNA complex was unreactive to RSSR and MTSL, we tested whether premodified MTSL-C134-VirE2 would interact with ssDNA. No T-complex was formed when prepared in this order (Figure S6). These converse observations, that MTSL-modified VirE2 did not bind ssDNA while VirE2 prebound to ssDNA could not be modified by MTSL, led to a repeat study using a much smaller reagent, N-ethylmaleimide (NEM). Figure 5C summarizes these interactions of modified or unmodified VirE2 with ssDNA by gel retardation. In isolated VirE2 (case 1), mass spectroscopy identified C134 as the single NEM-modified residue. As in the case of MTSL, ssDNA-bound protein (case 3) could not be modified by NEM confirming results obtained above with MTSL.

Altogether it is clear from these data that chemical modification of C134 interferes with T-complex formation, even though the residue lies far from the positively charged face that binds VirE1 and that could be expected to bind ssDNA (see Figure 4C). In comparison with the chosen set 1 from the modeling, we can interpret this result as a failure of the protein to form longitudinal links in the solenoid. Thus ssDNA binding is determined by the ability of the protein to assume the suitable quaternary structure.

DISCUSSION

The cryoEM structure, exhaustive modeling, mass spectrometry after chemical modification and cross-linking, and EPR analysis allow us to propose a detailed structural model for the T-complex, building on the high-resolution data from X-ray crystallography. Such integrated approaches in structural biology can effectively characterize the organization of complex assemblies (Lasker et al., 2012).

The VirE2 protein exists in at least three distinct forms: as a soluble heterodimer with its small chaperone VirE1; as a poorly soluble, disordered filamentous aggregate in the absence of

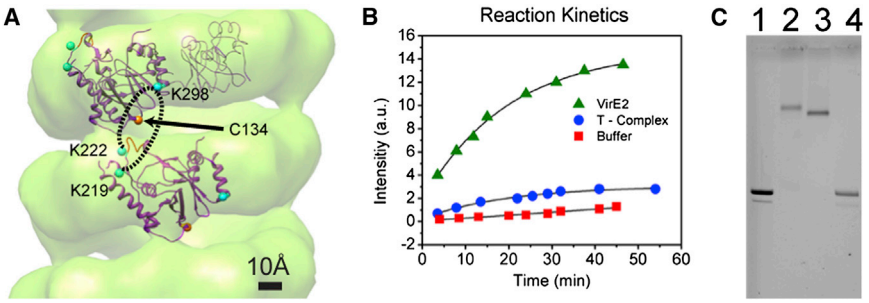

Figure 5. Residue Reactivity and Its Influence on VirE2-ssDNA Complex Assembly

(A) Blue spheres denote the lysine residues K219 and K222 that cross-link to K298 via the bifunctional reagent BS^3, confirming the computed fitting model of the X-ray structure into the cryoEM map. Orange spheres depict the cysteine residue C134. Note the position of C134 at the longitudinal interface between protein units.

(B) Cysteine reaction kinetics by EPR show clearly that VirE2 (as filamentous aggregates) is modified by the biradical reagent ṜS-SṜ. The fit shows first-order kinetics. VirE2-ssDNA (T-complex) is not modified. The buffer alone is shown as a control. See also Figure S7.

(C) VirE2-ssDNA is not reactive to NEM, while NEM-modified VirE2 does not bind ssDNA. Lane 1: M13 ssDNA in agarose (0.8%) electrophoresis gel; lane 2: electrophoretic mobility of M13 ssDNA bound to VirE2 is retarded; lane 3: mobility of NEM added to M13 ssDNA bound to VirE2—according to mass spectroscopy no reaction takes place; and lane 4: VirE2 premodified with NEM (confirmed with mass spectroscopy) does not bind M13 ssDNA. See also Figure S6 and Tables S2 and S3.

binding partners; and in the ordered solenoid form described as the T-complex. Each of these quaternary forms should be reflected in somewhat different tertiary structures. We have seen at the outset that the structure determined by X-ray crystallography requires at least a global bend in the flexible linker region to fit within the EM map of the T-complex. The disordered filament represents an intermediate state, most likely with a greater variability in the relative orientation around the flexible linker than would be present in either of the other cases. There is also a clear transition of structure between the forms. For example the outer diameter of the solenoid and disordered filaments were similar (Frenkiel-Krispin et al., 2007), and even a short incubation (~1 min) of VirE2 with ssDNA transformed these aggregates into solenoidal complexes. More surprisingly, incubation of VirE2 with short ssDNA oligomers also led to the transition from disordered filaments to ordered solenoids. This indicates that the transition in the quaternary protein assembly is catalyzed locally by the presence of the DNA, rather than by dissolution and repolymerization of the VirE2 along a continuous DNA substrate (Frenkiel-Krispin et al., 2007).

The solenoidal structure has been the primary focus of this study. The cryoEM map of the VirE2-ssDNA complex shows three distinct domains per asymmetric unit, but individual protein boundaries are not clearly delimited. Our aim has been to orient the high-resolution information available from X-ray crystallography within the constraint of the cryoEM map, thereby identifying critical points of contact for the complex assembly. To this end, we used structural information and one point of biochemical data: that the nuclear localization signal motifs must face the outer surface of the T-complex to interact with host cell proteins (Bhattacharjee et al., 2008; Guralnick et al., 1996; Salman et al., 2005). This orientation puts the positively charged, VirE1-binding face on the inside, where it should be available to bind ssDNA. Figure 6A summarizes the docking in detail as determined by modeling and confirmed by chemical analysis.

The solenoidal arrangement of a one-start helix around a hollow core presents azimuthal connections around the circumference, in which the C terminus of one protein unit contacts the N terminus of the neighboring unit around the perimeter. The propensity of VirE2 to form an alternative closed tetrameric structure (Abu-Arish et al., 2004) supports this assignment of the protein orientation rather than a hypothetical form running across adjacent turns, which would be formally acceptable on symmetry grounds. The VirE2-VirE2 junction is made, then, between the "pre-N" domain of 111 aa and the "post-C" 39 aa, both of which are unseen in the crystal structure but that can be attributed to the third lobe in the EM map. Folding of the pre-N domain might be induced by the constraint provided by the neighboring proteins, although we cannot comment on the degree of order present, nor on the extent to which this domain is folded in the filamentous aggregates of VirE2 alone. Consistent with the folding on contact, fusions at either the N or C terminus interfere with VirE2 interaction with ssDNA, as seen earlier with EM (Frenkiel-Krispin et al., 2007). In the soluble heterodimer formed with the secretion chaperone VirE1, the termini are unfolded and thus the C-terminal secretion signal remains available for recognition by the bacterial secretion system.

In the longitudinal direction, i.e., around the circumference and parallel to the helical axis, the only contacts between VirE2 molecules are between N-terminal domains. The immobile V-shaped loop at the lower side, formed by aa 132–142, contacts the highly mobile upper side of the N-terminal domain in the next turn down the solenoid, near residues aa 222–230 as seen in Figure 5A. Longitudinal contacts around the solenoid are also implied by a micro-mechanical study based on optical tweezers, in which it was found that the spring-form complex undergoes a stiffening upon "annealing" by repeated cycles of stretch and relaxation (Grange et al., 2008). The turns were adherent and not simply elastic, as they would be in a telephone cord for example. Interestingly, while the earlier structure in negative stain showed different helical parameters, the difference between the structures corresponds to precisely one additional protein unit per turn. The outer diameter increases accordingly, and the longitudinal contacts thus remain the same. This change in symmetry implies an important degree of flexibility within the molecular unit.

The most unexpected finding in this study was the super-allosteric regulation of DNA binding, via interruption of quaternary structure in the VirE2 solenoid. The critical residue C134 is not directly involved in DNA binding, yet it is blocked from modification by small molecule reagents exclusively in the

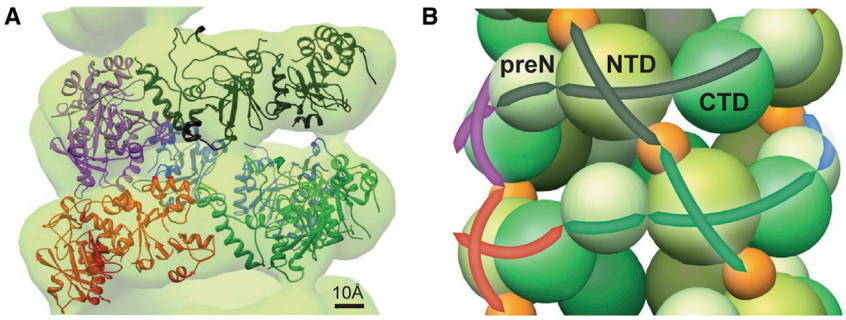

Figure 6. High-Resolution Model of the Solenoidal Shell of the T-Complex

(A) Extended view of the docked protein domains within the cryoEM envelope. Different monomers of VirE2 are represented by different colored ribbons and the cryoEM density is light green. See also Table S1.

(B) A cartoon view showing the N-terminal domain (NTD) held in place by two circumferential links to the pre-N (representing amino acids 1–111) and C-terminal domain (CTD) and by two longitudinal links, marked by the arrow bands and orange spheres. Thus the NTD is held rigidly in the assembled structure. Orientation of the CTD could not be determined effectively by the modeling approach and so remains suggestive. According to the general architecture of the T-complex, however, the CTD and adjoining pre-N domain of the adjacent protein are held only by circumferential links around the perimeter. In comparison with the NTD held by four links, these domains may be relatively free to rotate or bend within the assembled solenoid.

DNA-associated T-complex form. Because C134 remains reactive to IAA and other small molecule modifiers in the filamentous aggregate, it is most likely in a dynamic balance there between open and closed or buried forms at the longitudinal N-terminal interfaces. We can suggest that in open form C134 is reactive, but only in the buried form is the BS^3 lysine cross-linker able to react on both sides with distinct VirE2 molecules. ssDNA binding converts C134 to the buried form exclusively and inhibits reaction with reactive small molecules. We conclude from the sensitivity to C134 modification that ssDNA binding depends on cooperative assembly of the VirE2 shell, consisting of many protein units. In this respect the VirE2 assembly resembles a helical viral capsid in which both protein and oligonucleotide play an essential role in stabilizing the ordered assembly.

It is also interesting to note that octopine-strain VirE2 (accession number AAF77177.1), which shares a very high sequence similarity to the nopaline-strain protein studied here, has a glycine residue at the corresponding location (see alignment in Dym et al., 2008). Lacking C134, an early study of octopine VirE2 found its DNA binding insensitive to NEM (Sen et al., 1989). The interference at C134 thus appears to be steric rather than chemical in nature; generation of larger residues by chemical modification clash with the longitudinal neighbor and interfere with DNA complex formation.

In contrast to the N-terminal domain, our data provide little direct information on the orientation of the C-terminal domain. The cryoEM map is broad enough to accommodate it with rather little constraint. Also, the fact that the C-terminal domain interacts with the pre-N region whose structure is unavailable and was not included in the modeling limits our ability to localize it. Yet, given the architecture of the complex, we can expect this domain to be flexibly oriented. Figure 6B shows a schematic view of the solenoidal assembly emphasizing the presence of both longitudinal and circumferential constraints on the N-terminal domain of VirE2 but only two circumferential constraints on the C-terminal domain. One is the flexible interdomain linker of aa 337–346; on the opposite side is the pre-N domain (aa 1–111) of the adjacent protein. Thus one domain of VirE2 may be rigidly held while the remainder remains relatively flexible. Interaction between the post-C and pre-N regions, which are conserved in the two EM structures, as well as DNA binding at the basic face, corresponding to the VirE1-binding surface, may restrain the free rotation, but this should be less restrictive than the longitudinal protein-protein constraints on the N-domain. The possibility of variation among pre-N- and C-domain orientations can also explain, in hindsight, the failure of classification schemes to improve resolution in the 3D reconstruction. It is worth noting that the reconstruction of Figure 1D was obtained from only 280 particles chosen by the most severe selection criteria. If the C-terminal domains are indeed heterogeneously oriented, then little improvement will be obtained from further averaging.

We can propose that the flexibility at the C terminus allows the narrow, regular shell of VirE2 to accommodate DNA sequences with different local rigidities that arise due to base stacking. This is an essential factor in the sequence independence of Agrobacterium-mediated gene transfer. In this respect, the VirE2 assembly may be compared to flexible capsid shells of helical viruses. The ssRNA rabies virus, Marburg virus, Ebola virus, and measles virus were similar to VirE2 in the sense of resisting high-resolution structure determination by electron microscopy and 3D image processing (Albertini et al., 2008; Bharat et al., 2011, 2012; Schoehn et al., 2004). The nucleocapsids of these viruses formed a flexible macromolecular helix with many different symmetries. Additionally, in the case of the measles virus nucleoprotein, intrinsic flexibility was found to be important for proper protein function (Jensen et al., 2011). While VirE2 is not encoded by the enclosed oligonucleotide, the co-assembly they form otherwise bears a strong resemblance to a helical viral capsid both in function and in structure.

EXPERIMENTAL PROCEDURES

CryoEM and Image Processing

The VirE2:ssDNA sample (see Supplemental Experimental Procedures) was applied to freshly glow-discharged Quantifoil grids with circular holes and plunge-frozen in liquid ethane. Frozen-hydrated samples were transferred to a Gatan 626 cryo-holder, and imaged under low-dose conditions ($\leq$ 20 $e^-/Å^2$) at a magnification of 50,000× and at a defocus range from −0.8 to −2.8 μm, using an FEI Tecnai F20 FEG transmission electron microscope (TEM) operating at 200 KV and equipped with a Tietz F415 charge-coupled device (CCD) camera (pixel size 24 μm) with an effective sampling of 2.16 Å/pixel in the images. The procedure for collecting images on film (Kodak S0163) was exactly the same except that magnification used was 62,000× and the defocus range was −1.0 to −3.2 μm. The electron micrographs were scanned using a Nikon Coolscan 8000ED scanner at an effective sampling of 2.16 Å/pixel. The scanned micrographs were divided into smaller (128 × 128 pixels) boxes computationally, and an averaged power spectrum of these boxes

was determined for each micrograph using Bsoft (Heymann and Belnap, 2007). Fitting of the contrast transfer function (CTF) curve was carried out interactively in Bsoft. Only micrographs in which Thon rings could be observed at least up to ~13 Å resolution and where the CTF curve fit was reasonable by visual inspection were retained for further processing. Helical segments were picked from selected micrographs and CTF correction was carried out by phase-flipping implemented in Bsoft. The CTF-corrected helical segments were rotated to vertical orientation using Spider (Frank et al., 1996). Classification of the vertical images was then carried out using the neural networks-based Kernel density self-organizing maps (KerDenSOM) approach (Pascual-Montano et al., 2001) available in the Xmipp package (Sorzano et al., 2004). Helical reconstruction was carried out based on the IHRSR scheme (Egelman, 2000). Analysis of 3D data was performed using the Matlab software suite and visualization was carried out using UCSF Chimera (Pettersen et al., 2004).

Molecular Modeling

Docking of the individual domains of VirE2, taken from the crystal structure of VirE2 in complex with VirE1, into the cryoEM map was carried out using the program *fit*PDB2EM. *Fit*PDB2EM is a computational tool for geometrical rigid body fitting of atomic resolution structures to low-resolution EM maps (see Supplemental Experimental Procedures). The docking solutions for the N-terminal and C-terminal domains were combined and filtered, first based on physical structural criteria and then based on external criteria. Each model was scored and the models were clustered based on the root-mean-square deviation of Cα atoms. Further details of the docking and filtering are given in the Supplemental Experimental Procedures.

Mass Spectrometry Analysis

Samples were analyzed in an LTQ-Orbitrap (Thermo Fisher Scientific, Bremen, Germany) operated in the positive ion mode. Nano-liquid chromatography-electrospray chromatography-tandem mass spectrometry (LC-ESI-MS/MS)-peptide mixtures were separated by online reverse phase nanoscale capillary LC and analyzed by ESI-MS/MS. For the LC-MS/MS, the samples were injected onto an in-house made 15 cm reverse phase fused-silica capillary column (inner diameter 75 m), packed with 3 um ReproSil-Pur C18 A18 media (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany), using an UltiMate 3000 Capillary/Nano LC System. The LC setup was connected to the LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nano-electrospray ion source (Thermo Fisher Scientific, Bremen, Germany).

Electron Paramagnetic Resonance Analysis

All EPR measurements presented in this study were performed on an ELEXSYS 500 spectrometer (Bruker), equipped with a Euroterm ER 4113VT temperature unit (Bruker) with an accuracy of ± 1 K, in a 120 μl flat cell. Details and analysis of the resulting data appear in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The Electron Microscopy Data Bank (EMDB) accession number for the final negative-staining EM map of the T-complex is 2368. The EMDB and PDB accession numbers for the final cryoEM map and the fitted N-terminal domain of VirE2 are 2339 and 4BLF, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Information, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.str.2013.04.027.

ACKNOWLEDGMENTS

M.E. and S.G.W. acknowledge valuable discussions with Pawel Penczek. T.A.M.B. was supported initially by the FEBS summer fellowship program. We thank Dalit Merhav for her enthusiastic support in mass spectroscopy measurements. This work was supported in part by grants from the Israel Science Foundation-Bikura program and the European Research Council. Electron microscopy was performed at the Irving and Cherna Moscowitz Center for Nano and Bio- Nano Imaging of the Weizmann Institute of Science. We acknowledge support from the Gerhardt M.J. Schmidt Minerva Center for Supramolecular Architecture and the historical generosity of the Harold Perlman family.

Received: December 31, 2012
Revised: March 28, 2013
Accepted: April 18, 2013
Published: June 13, 2013

REFERENCES

Abu-Arish, A., Frenkiel-Krispin, D., Fricke, T., Tzfira, T., Citovsky, V., Wolf, S.G., and Elbaum, M. (2004). Three-dimensional reconstruction of Agrobacterium VirE2 protein with single-stranded DNA. J. Biol. Chem. *279*, 25359–25363.

Albertini, A.A., Schoehn, G., Weissenhorn, W., and Ruigrok, R.W. (2008). Structural aspects of rabies virus replication. Cell. Mol. Life Sci. *65*, 282–294.

Anand, A., Rojas, C.M., Tang, Y., and Mysore, K.S. (2012). Several components of SKP1/Cullin/F-box E3 ubiquitin ligase complex and associated factors play a role in Agrobacterium-mediated plant transformation. New Phytol. *195*, 203–216.

Bharat, T.A.M., Riches, J.D., Kolesnikova, L., Welsch, S., Krähling, V., Davey, N., Parsy, M.-L., Becker, S., and Briggs, J.A.G. (2011). Cryo-electron tomography of Marburg virus particles and their morphogenesis within infected cells. PLoS Biol. *9*, e1001196.

Bharat, T.A., Noda, T., Riches, J.D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoka, Y., and Briggs, J.A. (2012). Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. Proc. Natl. Acad. Sci. USA *109*, 4275–4280.

Bhattacharjee, S., Lee, L.Y., Oltmanns, H., Cao, H., Veena, Cuperus, J., and Gelvin, S.B. (2008). IMPa-4, an Arabidopsis importin alpha isoform, is preferentially involved in agrobacterium-mediated plant transformation. Plant Cell *20*, 2661–2680.

Christie, P.J., Ward, J.E., Winans, S.C., and Nester, E.W. (1988). The Agrobacterium tumefaciens virE2 gene product is a single-stranded-DNA-binding protein that associates with T-DNA. J. Bacteriol. *170*, 2659–2667.

Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. *59*, 451–485.

Citovsky, V., Wong, M.L., and 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.

Citovsky, V., Zupan, J., Warnick, D., and Zambryski, P. (1992). Nuclear localization of Agrobacterium VirE2 protein in plant cells. Science *256*, 1802–1805.

Citovsky, V., Guralnick, B., Simon, M.N., and Wall, J.S. (1997). The molecular structure of agrobacterium VirE2-single stranded DNA complexes involved in nuclear import. J. Mol. Biol. *271*, 718–727.

Desfosses, A., Goret, G., Estrozi, L.F., Ruigrok, R.W., and Gutsche, I. (2010). Nucleoprotein-RNA orientation in the measles virus nucleocapsid by three-dimensional electron microscopy. J. Virol. *85*, 1391–1395.

Dumas, F., Duckely, M., Pelczar, P., Van Gelder, P., and Hohn, B. (2001). An Agrobacterium VirE2 channel for transferred-DNA transport into plant cells. Proc. Natl. Acad. Sci. USA *98*, 485–490.

Dym, O., Albeck, S., Unger, T., Jacobovitch, J., Branzburg, A., Michael, Y., Frenkiel-Krispin, D., Wolf, S.G., and Elbaum, M. (2008). Crystal structure of the Agrobacterium virulence complex VirE1-VirE2 reveals a flexible protein that can accommodate different partners. Proc. Natl. Acad. Sci. USA *105*, 11170–11175.

Egelman, E.H. (2000). A robust algorithm for the reconstruction of helical filaments using single-particle methods. Ultramicroscopy *85*, 225–234.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J. Struct. Biol. *116*, 190–199.

Frenkiel-Krispin, D., Wolf, S.G., Albeck, S., Unger, T., Peleg, Y., Jacobovitch, J., Michael, Y., Daube, S., Sharon, M., Robinson, C.V., et al. (2007). Plant transformation by Agrobacterium tumefaciens: modulation of single-stranded DNA-VirE2 complex assembly by VirE1. J. Biol. Chem. *282*, 3458–3464.

Gelvin, S.B. (2012). Traversing the Cell: Agrobacterium T-DNA's Journey to the Host Genome. Front Plant Sci *3*, 52.

Grange, W., Duckely, M., Husale, S., Jacob, S., Engel, A., and Hegner, M. (2008). VirE2: a unique ssDNA-compacting molecular machine. PLoS Biol. *6*, e44.

Guralnick, B., Thomsen, G., and Citovsky, V. (1996). Transport of DNA into the nuclei of xenopus oocytes by a modified VirE2 protein of Agrobacterium. Plant Cell *8*, 363–373.

Heymann, J.B., and Belnap, D.M. (2007). Bsoft: image processing and molecular modeling for electron microscopy. J. Struct. Biol. *157*, 3–18.

Jensen, M.R., Communie, G., Ribeiro, E.A., Jr., Martinez, N., Desfosses, A., Salmon, L., Mollica, L., Gabel, F., Jamin, M., Longhi, S., et al. (2011). Intrinsic disorder in measles virus nucleocapsids. Proc. Natl. Acad. Sci. USA *108*, 9839–9844.

Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. USA *109*, 1380–1387.

Pascual-Montano, A., Donate, L.E., Valle, M., Bárcena, M., Pascual-Marqui, R.D., and Carazo, J.M. (2001). A novel neural network technique for analysis and classification of EM single-particle images. J. Struct. Biol. *133*, 233–245.

Penczek, P.A., Kimmel, M., and Spahn, C.M. (2011). Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. Structure *19*, 1582–1590.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Salman, H., Abu-Arish, A., Oliel, S., Loyter, A., Klafter, J., Granek, R., and Elbaum, M. (2005). Nuclear localization signal peptides induce molecular delivery along microtubules. Biophys. J. *89*, 2134–2145.

Schoehn, G., Mavrakis, M., Albertini, A., Wade, R., Hoenger, A., and Ruigrok, R.W. (2004). The 12 A structure of trypsin-treated measles virus N-RNA. J. Mol. Biol. *339*, 301–312.

Sen, P., Pazour, G.J., Anderson, D., and Das, A. (1989). Cooperative binding of Agrobacterium tumefaciens VirE2 protein to single-stranded DNA. J. Bacteriol. *171*, 2573–2580.

Simone, M., McCullen, C.A., Stahl, L.E., and Binns, A.N. (2001). The carboxy-terminus of VirE2 from Agrobacterium tumefaciens is required for its transport to host cells by the virB-encoded type IV transport system. Mol. Microbiol. *41*, 1283–1293.

Sorzano, C.O., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. J. Struct. Biol. *148*, 194–204.

Sorzano, C.O., Bilbao-Castro, J.R., Shkolnisky, Y., Alcorlo, M., Melero, R., Caffarena-Fernández, G., Li, M., Xu, G., Marabini, R., and Carazo, J.M. (2010). A clustering approach to multireference alignment of single-particle projections in electron microscopy. J. Struct. Biol. *171*, 197–206.

Suhre, K., and Sanejouand, Y.H. (2004). ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res. *32*(Web Server issue), W610-4.

Tzfira, T., Vaidya, M., and Citovsky, V. (2004). Involvement of targeted proteolysis in plant genetic transformation by Agrobacterium. Nature *431*, 87–92.

Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuïnk, T.J., and Hooykaas, P.J. (2000). VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science *290*, 979–982.

Xu, H., and Freitas, M.A. (2007). A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics *8*, 133.

Zaltsman, A., Krichevsky, A., Loyter, A., and Citovsky, V. (2010). Agrobacterium induces expression of a host F-box protein required for tumorigenicity. Cell Host Microbe *7*, 197–209.