Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the *Yersinia pestis* tip protein LcrV refined to 1.65 Å resolution

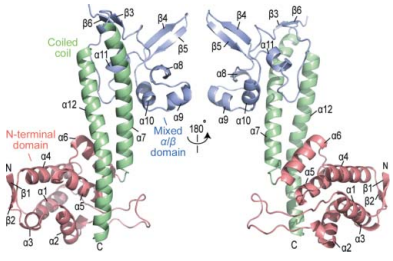
Sukanya Chaudhury,[a] Kevin P. Battaile,[b] Scott Lovell,[c] Gregory V. Plano[d] and Roberto N. De Guzman[a]*

[a]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA, [b]IMCA-CAT, Hauptman–Woodward Medical Research Institute, 9700 South Cass Avenue, Argonne, IL 60439, USA, [c]Protein Structure Laboratory, Del Shankel Structural Biology Center, University of Kansas, 2034 Becker Drive, Lawrence, KS 66047, USA, and [d]Department of Microbiology and Immunology, University of Miami, Miami, FL 33136, USA

Correspondence e-mail: rdguzman@ku.edu

Received 22 February 2013
Accepted 28 March 2013

PDB Reference: *Yersinia pestis* tip protein LcrV, 4jbu



The human pathogen *Yersinia pestis* requires the assembly of the type III secretion system (T3SS) for virulence. The structural component of the T3SS contains an external needle and a tip complex, which is formed by LcrV in *Y. pestis*. The structure of an LcrV triple mutant (K40A/D41A/K42A) in a C273S background has previously been reported to 2.2 Å resolution. Here, the crystal structure of LcrV without the triple mutation in a C273S background is reported at a higher resolution of 1.65 Å. Overall the two structures are similar, but there are also notable differences, particularly near the site of the triple mutation. The refined structure revealed a slight shift in the backbone positions of residues Gly28–Asn43 and displayed electron density in the loop region consisting of residues Ile46–Val63, which was disordered in the original structure. In addition, the helical turn region spanning residues Tyr77–Gln95 adopts a different orientation.

1. Introduction

The Gram-negative bacterium *Yersinia pestis*, which is the causative agent of bubonic plague, has caused pandemics that have resulted in large-scale social devastation. The type III secretion system (T3SS) plays a critical role in the pathogenicity of *Y. pestis* as well as that of other human pathogens such as *Shigella flexneri*, *Salmonella typhimurium* and *Burkholderia pseudomallei* (Cornelis, 2006). The T3SS is a complex nanomachine that is utilized by pathogenic bacteria to deliver bacterial effector proteins into the cytosol of targeted host cells (Cornelis, 2002, 2006). The structural component of the T3SS is a needle apparatus which is assembled from over a dozen different proteins and is composed of a basal-membrane-embedded structure, an external needle, a tip complex and a translocon. The external needle of the *Y. pestis* T3SS (also known as the Ysc–Yop system) is formed by polymerization of YscF and ends in a tip complex formed by multiple copies of LcrV which serves as a platform for the assembly of the translocon proteins YopB and YopD (Cornelis, 2002). Prior to assembly at the needle tip, LcrV is chaperoned by a smaller protein, LcrG (DeBord *et al.*, 2001; Matson & Nilles, 2001).

The 2.2 Å resolution crystal structure of LcrV (residues Leu24–Lys326) was reported in 2004 using a construct containing triple mutations (K40A/D41A/K42A) in a C273S background (Derewenda *et al.*, 2004). Since then, crystal structures of other T3SS tip proteins such as the *Shigella* IpaD (Johnson *et al.*, 2007), the *Salmonella* SipD (Chatterjee *et al.*, 2011; Lunelli *et al.*, 2011) and the *Burkholderia* BipD (Erskine *et al.*, 2006; Johnson *et al.*, 2007) tip proteins have been determined. The T3SS tip proteins share common structural motifs as well as variations among bacterial species. Here, the refined structure of LcrV is reported at 1.65 Å resolution, which highlights structural similarities as well as differences that exist between two constructs of the same protein.

2. Methods

2.1. Protein expression and purification

The DNA region coding for *Y. pestis* LcrV residues Gly28–Asp322 with a C273S point mutation was subcloned as a fusion protein with a His_6-tagged streptococcal GB1 domain and a tobacco etch virus

(TEV) protease site. The subcloning introduced a three-residue (GHM) cloning artifact at the N-terminus of LcrV following cleavage of the fusion tag. The expression plasmid was transformed into *Escherichia coli* BL21(DE3) DNAY and a 10 ml overnight starter culture was used to inoculate a 1 l LB culture. The cells were grown at 310 K to an OD_{600} of 0.6–0.8 and were induced with 1 m*M* IPTG overnight at 288 K. The cells were harvested, resuspended in 30 ml binding buffer (500 m*M* NaCl, 20 m*M* Tris–HCl, 5 m*M* imidazole pH 8.0) and stored at 253 K. To purify LcrV, the frozen cells were warmed to 277 K, lysed by sonication and centrifuged to remove the cellular debris. The supernatant was loaded onto a 5 ml nickel-affinity column (Sigma) and washed with the binding buffer. The protein was eluted using 500 m*M* NaCl, 20 m*M* Tris–HCl, 250 m*M* imidazole pH 8.0. Fractions containing LcrV were combined and dialyzed overnight against 1 l TEV digestion buffer (50 m*M* Tris, 0.5 m*M* EDTA, 1 m*M* DTT, 20 m*M* NaCl pH 8.0) in the presence of 300 µl 0.06 m*M* recombinant TEV protease (Geisbrecht *et al.*, 2006). The mixture was passed over a nickel-affinity column to separate LcrV from the GB1 fusion partner and TEV protease. Fractions containing LcrV were pooled, dialyzed against 50 m*M* NaCl, 20 m*M* Tris pH 8.0 and concentrated to 22 mg ml^{-1} (0.6 m*M*) using an Amicon Ultra 3K concentrator (Millipore).

2.2. Crystallization and data collection

Crystallization screening was conducted at 293 K in Compact Clover (Emerald BioSystems) sitting-drop vapor-diffusion plates using 0.5 µl protein solution and 0.5 µl crystallization solution equilibrated against 75 µl crystallization solution. Initial crystals, which formed plate clusters, were obtained in approximately 24 h from Index HT screen (Hampton Research) conditions D8 (25% PEG 3350, 100 m*M* HEPES pH 7.5) and D9 (25% PEG 3350, 100 m*M* Tris pH 8.5). The crystallization conditions were refined using the pH Buffer Screen (Emerald BioSystems) and the prismatic crystals used for data collection were obtained using 25% PEG 3350, 100 m*M* Bicine pH 7.0 (Fig. 1). Single crystals were transferred to a fresh drop of a cryoprotectant solution consisting of 80% crystallization solution and 20% PEG 400. X-ray diffraction data were obtained using a PILATUS 6M pixel-array detector (Dectris) on beamline 17ID (IMCA-CAT) at the Advanced Photon Source.

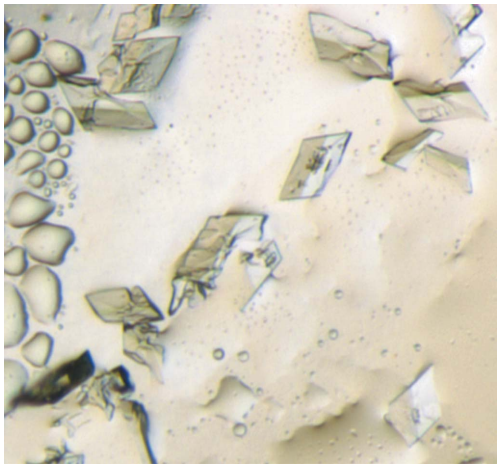


Figure 1
Prismatic crystals of LcrV (Gly28–Asp322, C273S). Initial crystals were observed from Index HT screen conditions D8 and D9. Refined crystals were obtained using the pH Buffer Screen.

Table 1
Crystallographic data for LcrV (Gly28–Asp322, C273S) refined to 1.65 Å resolution.

Values in parentheses are for the highest resolution shell.

Data collection	
Unit-cell parameters (Å, °)	a = 39.51, b = 51.44, c = 65.18, β = 92.00
Space group	$P2_1$
Resolution (Å)	51.44–1.65 (1.74–1.65)
Wavelength (Å)	1.0000
Temperature (K)	100
Observed reflections	103407
Unique reflections	31392
Mean $I/\sigma(I)$	18.0 (2.2)
Completeness (%)	99.3 (99.8)
Multiplicity	3.3 (3.3)
R_{merge}† (%)	3.4 (50.5)
R_{meas}‡ (%)	4.0 (60.3)
$R_{p.i.m.}$§ (%)	2.2 (32.6)
Refinement	
Resolution (Å)	34.33–1.65
Reflections (working/test)	29790/1582
R factor/R_{free}¶ (%)	17.7/22.0
No. of atoms	
Protein	2216
PEG 400	10
Water	149
Model quality	
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.196
Average B factor (Å^2)	
All atoms	28.6
Protein	28.0
PEG 400	48.9
Water	35.4
Coordinate error based on maximum likelihood (Å)	0.24
Ramachandran plot	
Most favored (%)	98.9
Additionally allowed (%)	1.1

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the ith reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices hkl. ‡ R_{meas} is the redundancy-independent (multiplicity-weighted) R_{merge} (Evans, 2006). § $R_{p.i.m.}$ is the precision-indicating (multiplicity-weighted) R_{merge} (Diederichs & Karplus, 1997; Weiss, 2001). ¶ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$; R_{free} is calculated in an identical manner using a randomly selected 5% of the reflections, which were not included in the refinement

2.3. Structure solution and refinement

The diffraction data were integrated and scaled using *XDS* (Kabsch, 1988) and *SCALA* (Evans, 2006), respectively. The Laue class was checked with *POINTLESS* (Evans, 2006), which indicated that 2/*m* was correct and that the likely space group was $P2_1$. Structure solution was conducted with *Phaser* (McCoy *et al.*, 2007) *via* the *PHENIX* (Adams *et al.*, 2010) interface using a previously determined structure of LcrV (PDB entry 1r6f; Derewenda *et al.*, 2004) as the search model. Searches were conducted in space groups $P2$ and $P2_1$, with the top solution being obtained in the latter. Following manual model building with *Coot* (Emsley & Cowtan, 2004), the structure was improved by automated model building with *ARP/wARP* (Langer *et al.*, 2008). Structure refinement and additional model building were performed with *PHENIX* and *Coot*, respectively. Disordered side-chain residues were truncated at the point to which electron density could be observed. Structure validation was conducted with *MolProbity* (Chen *et al.*, 2010). Crystallographic data are provided in Table 1.

3. Results and discussion

The final model contained residues Gly28–Leu319, a PEG 400 molecule and 149 water molecules (Fig. 2). Residues Lys261–Pro279

could not be fitted to the electron density owing to disorder. The overall structure of LcrV is similar to that reported previously (PDB entry 1r6f; Derewenda *et al.*, 2004), with an r.m.s.d. between C^{α} atoms (244 residues) of 1.62 Å as determined using secondary-structure matching (Krissinel & Henrick, 2004) in *Coot* (Emsley *et al.*, 2010). Similar to 1r6f (Derewenda *et al.*, 2004), LcrV is mainly an α-helical protein, with 55% of the residues present in 12 α-helices (α1–α12), 7% of the residues present in six β-strands and 32% of the residues present in random coils. The overall shape of the protein structure can be described as 'dumbbell-shaped' and is composed of a central coiled-coil region connecting two globular domains (Derewenda *et al.*, 2004). The N-terminal globular domain (residues Gly25–Gly147) is formed by two antiparallel β-strands (β1 and β2) and six helices (α1–α6) and packs at one end of a long central coiled coil formed by helix α7 and helix α12. The central coiled coil connects to a second globular domain formed by a mixture of α-helices and β-strands. This mixed α/β domain is less structured compared with the N-terminal globular domain and contains four short helices (α8–α11) and four antiparallel β-strands (β3–β6). The bulk of the α/β domain is unstructured and is composed of coils and turns connecting the short helices and the β-strands.

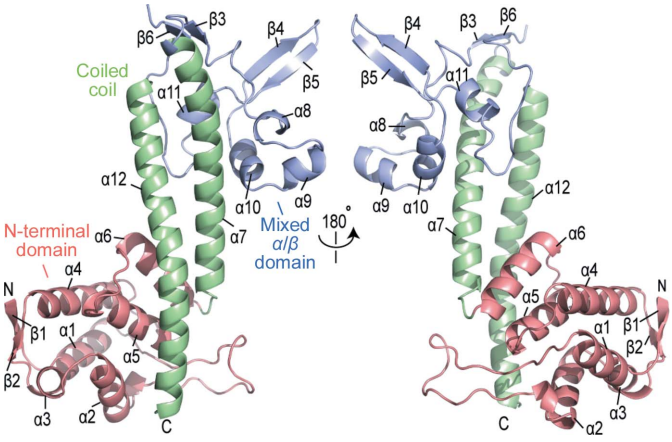


Figure 2
Crystal structure of LcrV refined to 1.65 Å resolution. LcrV can be divided into three domains: the N-terminal globular domain (pale pink), the central coiled-coil domain (green) and the less structured mixed α/β domain (light blue).

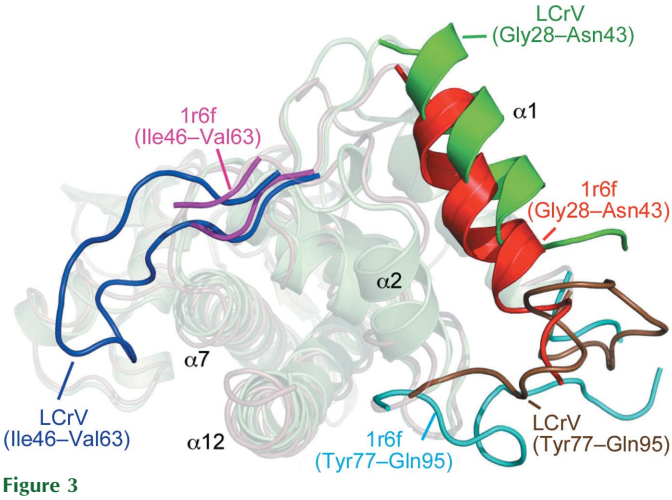
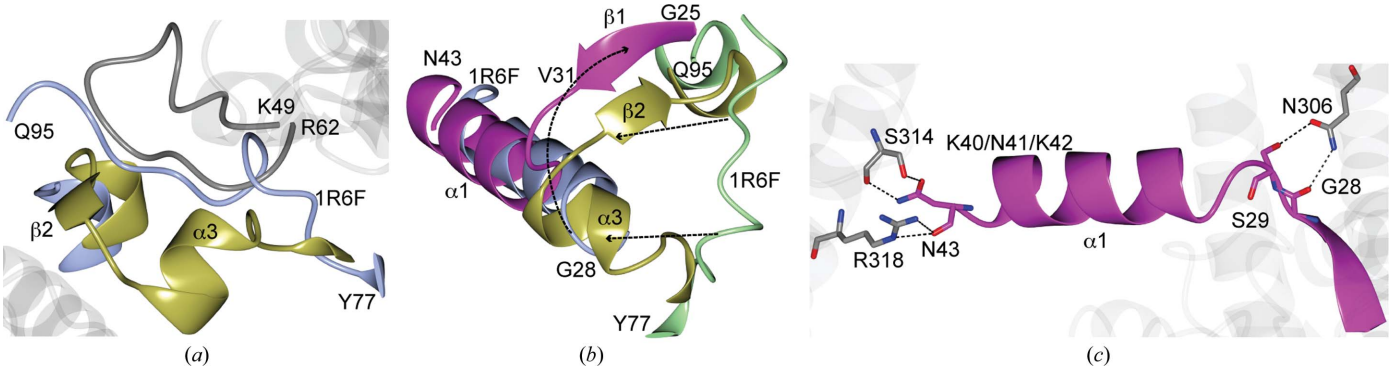


Figure 3
The differences in the N-terminal domain of the refined stucture of LcrV from that of PDB entry 1r6f.

However, there are significant differences between 1r6f and the current structure. In the N-terminal domain, differences exist at residues Gly28–Asn43, Ile46–Val63 and Tyr77–Gln95 (Fig. 3). Residues Gly23–His27 could not be modeled in PDB entry 1r6f; however, in the current structure residues Gly28–Ser30 formed a random coil, with helix α1 starting at Val31 (helix α1 starts at Gly28 in 1r6f). The relative orientation of helix α1 is also slightly shifted compared with the equivalent helix in 1r6f. In 1r6f, residues Lys49–Val63 were not visible in the electron-density map owing to disorder (Derewenda *et al.*, 2004), whereas in the current structure Ile46–Val63 form an ordered loop that packs at the C-terminal end of one of the helices (helix α12) of the central coiled coil. Additionally, the helical turn spanning residues Tyr77–Gln95 adopts different orientations in the two structures and Tyr90, which was not visible in 1r6f, could be modeled in the current structure. In 1r6f the residues Asn260–Asp275 could not be modeled. Similarly, in our model the residues Lys261–Pro279 could not be modeled owing to disorder.

As noted above, the 1r6f structure was solved using the triple mutant K40A/D41A/K42A in a C273S background, whereas the LcrV structure reported here only has the background C273S mutation. The C273S point mutation facilitated protein purification and the triple mutation (K40A/D41A/K42A) at the terminal end of helix α1 facilitated crystallization (Derewenda *et al.*, 2004). Most of the significant differences between the 1r6f structure and the current structure are in the helix containing the triple mutation and in the region spanning Tyr77–Gln95. The differences in these regions of the two structures may be a consequence of crystal packing. It should be



(*a*) (*b*) (*c*)

Figure 4
Superposition of the refined structure of LcrV with that of PDB entry 1r6f. (*a*) A loop spanning Lys49–Arg62 from a symmetry-related molecule (gray) in the current structure would clash with the Tyr77–Gln95 region (blue) of PDB entry 1r6f. (*b*) Conformational changes in the Gly25–Asn43 region are shown by the curved arrow and those in the Tyr77–Gln95 region are indicated by straight arrows. (*c*) Hydrogen-bond crystal contacts (dashed lines) observed in the current structure. Symmetry-related molecules are colored gray. The Lys40–Lys42 region from 1r6f is indicated.

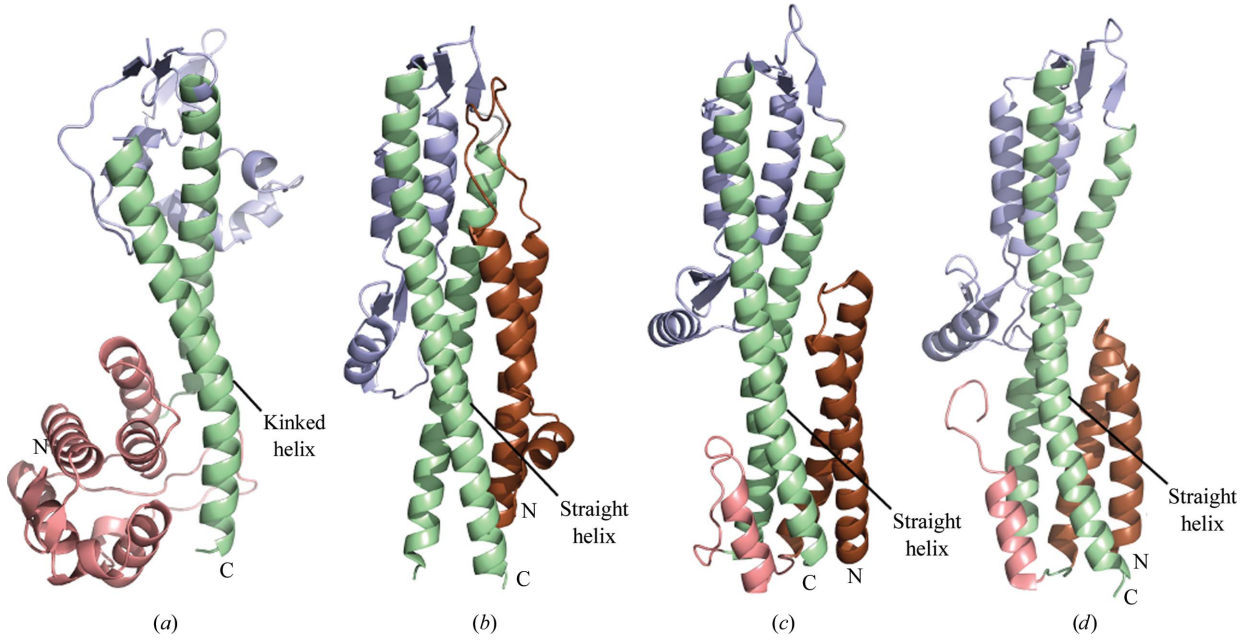


Figure 5
The most prominent structural feature of the T3SS tip proteins is the central coiled coil (green), which is kinked in *Y. pestis* LcrV (*a*) but is straight in *B. pseudomallei* BipD (*b*), *S. flexneri* IpaD (*c*) and *S. typhimurium* SipD (*d*).

noted that the 'tightness' of the crystal packing, as gauged from the solvent content (39.1% for 1r6f *versus* 37.0% for the current structure) and the Matthews coefficient (2.02 Å^3 Da^{-1} for 1r6f *versus* 1.95 Å^3 Da^{-1} for the current structure), is similar in the two structures. However, the Tyr77–Gln95 region appears to be precluded from adopting a conformation similar to that observed for 1r6f owing to a nearby loop from a symmetry-related molecule (Fig. 4*a*). Thus, Tyr77–Gln95, which is random coil in 1r6f, is composed of a helix and a sheet (α3 and β2) in the current structure. This in turn appears to result in movement of the N-terminus from Gly25 to Val31 to form the β1 strand, which adopts an antiparallel orientation to β2, and a slight shift in helix α1 relative to 1r6f (Fig. 4*b*). It was noted that the site of the triple mutant in 1r6f forms a crystal contact with Lys176 (Derewenda *et al.*, 2004), although it does not appear to form a direct hydrogen bond but is rather in proximity to the mutation sites. In the current structure, new hydrogen bonds between symmetry-related molecules are formed at the ends of helix α1, as shown in Fig. 4(*c*). The triple mutations in 1r6f are located in the N-terminal domain of LcrV, which forms the base of the tip complex and is suggested to be involved in protein–protein interaction with the translocon (Broz *et al.*, 2007). Thus, the refined crystal structure of LcrV reported here with native residues in the N-terminal domain will be useful in future structural studies of the LcrV–translocon interaction.

There is low sequence identity between LcrV and the tip proteins IpaD (Johnson *et al.*, 2007), BipD (Erskine *et al.*, 2006; Johnson *et al.*, 2007) and SipD (Chatterjee *et al.*, 2011; Lunelli *et al.*, 2011). LcrV is more closely related to PcrV, the tip protein of *Pseudomonas aeruginosa* (Derewenda *et al.*, 2004), for which a crystal structure is not available. Nevertheless, despite poor sequence similarity to IpaD, BipD and SipD, the LcrV central coiled coil (forming the handle of the dumbbell) is conserved and forms a common structural feature among T3SS tip proteins (Fig. 5; Erskine *et al.*, 2006; Johnson *et al.*, 2007; Lunelli *et al.*, 2011; Chatterjee *et al.*, 2011). The structural conservation of the central coiled coil among T3SS tip proteins suggests an important role in T3SS function. Others have hypothesized that the central coiled coil plays an important role in the assembly of the T3SS tip complex (Deane *et al.*, 2006; Blocker *et al.*, 2008). Recently, it has been shown that the central coiled coil of SipD is the site of protein–protein interactions with the needle protein and is important in the assembly of the T3SS needle apparatus (Rathinavelan *et al.*, 2011; Lunelli *et al.*, 2011).

With respect to structural differences among T3SS tip proteins, the LcrV N-terminal globular domain is unique to LcrV and is absent in IpaD (Johnson *et al.*, 2007), BipD (Erskine *et al.*, 2006; Johnson *et al.*, 2007) and SipD (Chatterjee *et al.*, 2011; Lunelli *et al.*, 2011) (Fig. 5). In IpaD, BipD and SipD the N-terminal region forms an α-helical hairpin that has been proposed to be a self-chaperone for these tip proteins (Johnson *et al.*, 2007; Erskine *et al.*, 2006; Blocker *et al.*, 2008). However, in LcrV a small 95-residue protein, LcrG, functions as a chaperone *via* the formation of a protein complex (Matson & Nilles, 2001, 2002; Lawton *et al.*, 2002; Blocker *et al.*, 2008). Another difference between LcrV and the self-chaperoned tip proteins (IpaD, BipD and SipD) is that the C-terminal helix of the central coiled coil (helix α12) is bent and kinked, whereas in the other tip proteins the equivalent helix is straight (Fig. 5; Blocker *et al.*, 2008). These structural differences suggest that the needle-tip assembly process may vary between T3SS family members.

We are grateful to Srirupa Chatterjee for helpful discussions. This research was supported by NIH grants AI074856 (RND), AI039575 (GP), 5P20RR017708-10 and 8P20GM103420-10 (University of Kansas Protein Structure Laboratory). Use of the IMCA-CAT beamline 17-ID at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with the Hauptman–Woodward Medical Research Institute. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract No. DE-AC02-06CH11357.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Blocker, A. J., Deane, J. E., Veenendaal, A. K., Roversi, P., Hodgkinson, J. L., Johnson, S. & Lea, S. M. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 6507–6513.
Broz, P., Mueller, C. A., Müller, S. A., Philippsen, A., Sorg, I., Engel, A. & Cornelis, G. R. (2007). *Mol. Microbiol.* **65**, 1311–1320.
Chatterjee, S., Zhong, D., Nordhues, B. A., Battaile, K. P., Lovell, S. & De Guzman, R. N. (2011). *Protein Sci.* **20**, 75–86.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cornelis, G. R. (2002). *Nature Rev. Mol. Cell Biol.* **3**, 742–752.
Cornelis, G. R. (2006). *Nature Rev. Microbiol.* **4**, 811–825.
Deane, J. E., Roversi, P., Cordes, F. S., Johnson, S., Kenjale, R., Daniell, S., Booy, F., Picking, W. D., Picking, W. L., Blocker, A. J. & Lea, S. M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 12529–12533.
DeBord, K. L., Lee, V. T. & Schneewind, O. (2001). *J. Bacteriol.* **183**, 4588–4598.
Derewenda, U., Mateja, A., Devedjiev, Y., Routzahn, K. M., Evdokimov, A. G., Derewenda, Z. S. & Waugh, D. S. (2004). *Structure*, **12**, 301–306.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Erskine, P. T., Knight, M. J., Ruaux, A., Mikolajek, H., Wong Fat Sang, N., Withers, J., Gill, R., Wood, S. P., Wood, M., Fox, G. C. & Cooper, J. B. (2006). *J. Mol. Biol.* **363**, 125–136.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Geisbrecht, B. V., Bouyain, S. & Pop, M. (2006). *Protein Expr. Purif.* **46**, 23–32.
Johnson, S., Roversi, P., Espina, M., Olive, A., Deane, J. E., Birket, S., Field, T., Picking, W. D., Blocker, A. J., Galyov, E. E., Picking, W. L. & Lea, S. M. (2007). *J. Biol. Chem.* **282**, 4035–4044.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 67–72.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Lawton, D. G., Longstaff, C., Wallace, B. A., Hill, J., Leary, S. E., Titball, R. W. & Brown, K. A. (2002). *J. Biol. Chem.* **277**, 38714–38722.
Lunelli, M., Hurwitz, R., Lambers, J. & Kolbe, M. (2011). *PLoS Pathog.* **7**, e1002163.
Matson, J. S. & Nilles, M. L. (2001). *J. Bacteriol.* **183**, 5082–5091.
Matson, J. S. & Nilles, M. L. (2002). *BMC Microbiol.* **2**, 16.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Rathinavelan, T., Tang, C. & De Guzman, R. N. (2011). *J. Biol. Chem.* **286**, 4922–4930.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.