Crystal Structure of Lyme Disease Variable Surface Antigen

VlsE of *Borrelia burgdorferi*

Christoph Eicken¹, Vivek Sharma¹, Thomas Klabunde¹*, Matthew B. Lawrenz², John M. Hardham²†, Steven J. Norris² and James C. Sacchettini¹

¹ The Center for Structural Biology, Texas A&M University, College Station, Texas 77843-2128 and Albert B. Alkek Institute of Biosciences and Technology, 2121 W. Holcombe Blvd., Houston, TX 77030-3303, USA. Phone: +1-979-8627636, FAX: +1-979-8627638, e-mail: sacchett@tamu.edu

² Departments of Pathology and Laboratory Medicine and Microbiology and Molecular Genetics, University of Texas Medical School at Houston, P.O. Box 2070, Houston, TX 77225, USA.

* current address: Aventis Pharma Deutschland GmbH, Chemical Research - Molecular Modeling, Building G838, D-65926 Frankfurt am Main, Germany.

† current address: Pfizer, Inc., Animal Health Biological Discovery, Eastern Point Road, Groton, CT 06340 USA

Correspondence should be addressed to James C. Sacchettini (e-mail: sacchett@tamu.edu, phone: 979-862-7636, FAX: 979-862-7638).

Running title: Crystal structure of Lyme Disease Antigen VlsE
The atomic coordinates and structure factors (code 1L8W) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

SUMMARY

VlsE is an outer surface lipoprotein of Borrelia burgdorferi that undergoes antigenic variation through an elaborate gene conversion mechanism, and is thought to play a major role in the immune response to the Lyme disease borellia. The crystal structure of recombinant VlsE1 at 2.3 Å resolution reveals that the six variable regions form loop structures that constitute most of the membrane distal surface of VlsE, covering the predominantly α-helical, invariant regions of the protein. The surface localization of the variable amino acid segments appears to protect the conserved regions from interaction with antibodies and hence contribute to immune evasion.
INTRODUCTION

Lyme disease is a multistage, tick-borne infection that is endemic to regions of the United States, Europe, and Asia (1). The causative bacteria are a family of closely related spirochetes, including *Borrelia burgdorferi*, *Borrelia garinii*, and *Borrelia afzelii*, and are transmitted from one mammal to another by *Ixodes* ticks. Lyme disease borrelia cause persistent infections and chronic neurologic, cardiovascular, and arthralgic manifestations that can last for months to years in humans and other mammals if not treated successfully, indicating that the spirochetes can effectively evade the host’s immune defenses.

The mechanisms of immune evasion are not well understood at this time, but are thought to include at least one form of antigenic variation. The VMP-like sequence (vls) locus of *B. burgdorferi* (2) is a complex antigenic variation system that in many ways resembles the more thoroughly characterized variable major protein system of relapsing fever borrelia (3). The vls locus of *B. burgdorferi* B31 consists of the expression site vlsE and a contiguous set of 15 vls silent cassettes. The exact function of the 35 kDa surface-exposed lipoprotein VlsE is unknown; however, it is thought that the vls system may play an important role in mammalian infection, because loss of the encoding linear plasmid lp28-1 results in reduced infectivity (4,5).

The overall primary structure of VlsE consists of N-terminal and C-terminal constant domains flanking a central cassette region. The cassette region exhibits ~90% sequence identity with the silent cassettes, and most of the sequence differences are concentrated in six variable regions (VR1-VR6) interspersed among six invariant regions (IR1-IR6). Both the vlsE cassette region and the silent cassettes are demarcated by 17-bp direct repeats (DR). During experimental infection of mice, segments of the silent cassettes recombine into the vlsE cassette region through a gene conversion mechanism (2,6). The sequence changes are detectable within
four days after experimental infection of mice; by 28 days every isolate from skin or other tissues is unique and contains roughly 9 to 13 recombination events (7). It has been hypothesized that the humoral immune response is unable to respond effectively to the seemingly continuous generation of new VlsE variants, thus permitting immune evasion and persistent infection (2,7). Paradoxically, VlsE induces a strong antibody response in Lyme disease patients and infected animals, and recombinant VlsE protein and a peptide corresponding to IR6 have been found to be useful in the serodiagnosis of Lyme disease (8-10). Thus a dilemma has existed between the high immunogenicity of VlsE and its possible role in immune evasion.

To better understand the relationship between VlsE structure and antigenicity, and to gain insight into its possible function(s), we have determined the three-dimensional structure of VlsE from *B. burgdorferi* B31. For these studies, a recombinant form of VlsE1, the first variant of VlsE identified (2), was utilized. The signal peptide sequence (corresponding to aa 1-19 of the full-length sequence) was omitted from the recombinant construct to permit isolation and crystallization of a protein lacking the lipid moiety. The resulting protein thus contained the N-terminal region (aa 20-115), the cassette region (aa 116-310), and the C terminal region (aa 311-346) of the mature protein. The recombinant protein also contained an N-terminal polyhistidine sequence to aid in the purification process (8).
EXPERIMENTAL PROCEDURE

Protein Expression and Purification. *Borrelia burgdorferi* B31, initially isolated from an *Ixodes scapularis* tick, was cultured in BSK II medium as described previously (11). B31-5A3 is a low-passage-number clone that has retained its ability to cause disease in C3H/HeN mice and harbors the VlsE-encoding linear plasmid, lp28-1 (2,5,11). The 1,227-bp vlsE gene of B31-5A3 was amplified by PCR, expressed as a polyhistidine fusion protein (VlsE1-His) from the pQE30 (Qiagen) expression system in *Escherichia coli* SURE2 (Stratagene), and purified as described before (8). The resulting fusion protein (VlsE-His) includes an N-terminal polyhistidine tag followed by the full-length, mature VlsE1 sequence, but lacks the native N-terminal lipoprotein signal sequence (8). Selenomethionylated VlsE1-His was produced by standard methods (12) and purified using the same procedure as described for the native protein.

Crystallization and data collection. Initial crystals were obtained using the micro-batch method in condition #43 of the Crystal Screen I from Hampton Research. High diffraction quality crystals were produced to a size of approximately $0.5 \times 0.3 \times 0.05$ mm in 10 mM Tris pH 8.0, 15 % (w/v) PEG 1,500 and with a protein concentration of 10-15 mg ml$^{-1}$. After soaking in cryoprotectant solution (10 mM Tris pH 8.0, 20 % PEG 400, 20 % PEG 1,500) or paratone oil for a few seconds, the crystals were flash cooled to 100 K. Crystals of the selenomethionine protein were observed in the space group $P2_1$ with $a = 85.2$ Å, $b = 59.2$ Å, $c = 116.2$ Å and $\beta = 104.6^\circ$, and contain four VlsE1 monomers in the asymmetric unit. MAD data were collected at beamline 14-BM-D at the Advanced Photon Source (APS), Argonne National Laboratory to 2.3 Å resolution on a charge-coupled-device area detector (ADSC Q4) at three wavelengths near the
selenium absorption edge (see Table I). All reflection data were processed and scaled using
DENZO and SCALEPACK (13).

Structure Determination and refinement. Initial phases were determined by multi-wavelength
anomalous dispersion (MAD) (14) using a selenomethionine (SeMet) derivative of VlsE1-6His.
The program SOLVE (15) was used to locate the eight selenium sites in the asymmetric unit.
Phases obtained from SOLVE had a mean figure of merit 0.42 to 2.8 Å (Table 1) and were
improved by solvent flattening and non-crystallographic symmetry (NCS) averaging as
implemented in DM (16) and CNS (17). The molecular coordinates were initially constructed
using NCS-averaged and phase-combined electron density maps with the computer program O
(18) and refined with CNS (17).
RESULTS AND DISCUSSION

Overall structure of VlsE1

The VlsE1 crystal structure (Fig. 1a) was refined to a resolution of 2.3 Å and a final R-factor of 20.4 % (R-free of 28.2 %). Both the N-terminal and C-terminal regions of the protein appear to be quite flexible in VlsE1, as we were unable to locate the first 16 N-terminal and the last 12 C-terminal residues in the electron density maps. In two out of the four molecules in an asymmetric unit, as many as 32 N-terminal residues are disordered. The final model exhibited good stereochemistry with over 90 % of the residues in the most favored regions. VlsE1 crystallizes such that four molecules (each 80 Å × 37 Å × 31 Å) are found in the asymmetric unit (Fig. 1b). The four molecules were refined without non-crystallographic symmetry restraints due to slight differences in their conformation. A single molecule of VlsE1 contains eleven α-helices and four short β-strands (Fig. 1a). Four helices α1 (aa 37-48), α2 (aa 68-87), and α3 (aa 114-139), and α11 (aa 306-341) are in close proximity to form the membrane proximal part of VlsE. In contrast, helices α4 through α10 define the core region of the membrane distal part of the protein. They are widely covered by connecting loop regions of which some show different conformations for each molecules. The four β-strand segments are short (3 aa each) and are located in the membrane distal region. The amino acid residues 93-112 connecting α2 and α3 are disordered in all four polypeptide chains and are missing in the final model.

Analytical ultracentrifugation and gel filtration chromatography experiments were performed in our laboratories in analogy to analyses of other surface proteins (19) and indicate that recombinant VlsE1 is primarily monomeric in solution (data not shown). In the crystal structure, the interface between neighboring VlsE1 molecules in the asymmetric unit buries
approximately 13% of the accessible surface area of each monomer (Fig. 2b). While this interface does not exhibit any significant hydrophobic patches, it does suggest a possibility that VlsE could exist as a dimer on the spirochete surface in the arrangement shown in Fig. 2b. Very recent thermal denaturation data support the formation of oligomeric vlsE under low ionic strength conditions (20).

The cassette region

The protein segment encoded by the first 17 base pair direct repeat (DR₁), representing the beginning of the cassette region, is part of helix α3 at the membrane proximal end of VlsE₁ and is exposed on the surface of the protein (Fig. 2). DR₂ at the end of the cassette region is also surface exposed, however it is located at the membrane distal portion. The cassette region forms the membrane distal surface and a portion of the lateral surface of VlsE₁ (Fig. 2,3).

All six VRs are predominantly loop structures entirely covering the surface of the membrane distal end of the protein (Fig. 2,3). Approximately 50% of the theoretical maximum surface area of the VRs is exposed on the surface of a subunit of VlsE₁. Amino acid positions that undergo changes during the antigenic variation by cassette exchange are concentrated in three major areas on the membrane distal portion of the protein surface (Fig. 3c). The variable amino acids constitute only a portion of the VR sequences; however, it is likely that changes in these residues alter the overall conformation of the random coil structures and hence reconfigure the epitopes throughout each VR. Although less than 26% of the primary sequence is located in VRs, these regions represent about 37% of the total surface area of the VlsE₁ monomer. The location of these variable regions on the most membrane distal end of the molecule suggests that
immune evasion may occur via the shielding of the conserved regions of VlsE from antibody binding.

In contrast to the VRs, the invariant regions (IRs) of the cassette consist of distinct elements of secondary structure ($\alpha$ helices and $\beta$ sheet structures) and have very limited surface exposure. As shown in Fig. 2, the IRs form helices $\alpha_3$ through $\alpha_{10}$, connected by the VR-loops. Although 59% of the cassette primary structure is located in IRs, these regions account for less than 40% of the cassette surface area as calculated with SPOCK (21). In particular, IR$_5$, and IR$_6$ and IR$_3$ exhibit only 6.5%, 13.7% and 17.2% of their theoretical maximum surface area, respectively. IR$_1$ (38.7%) and IR$_4$ (35.8%) are the invariable regions within the cassette that are the most solvent exposed.

**Structure and antibody reactivity**

A large body of literature is available regarding the antigenic properties of different portions of the VlsE primary structure. Because the N- and C-terminal unique conserved regions do not undergo sequence variation (see Fig. 2), they are the most obvious targets for use in diagnostic tests and vaccination experiments. Although no experiments examining the antigenicity of the invariant N-terminal region have been reported, data about the unique C-terminal region became available recently (22). Mice immunized with a 50 aa peptide called Ct corresponding to the C-terminal region of VlsE of *B. burgdorferi* B31 induced a strong antibody response, indicating that this region is highly immunogenic. Similarly, sera from mice and monkeys infected with *B. burgdorferi* B31 exhibited a high reactivity with Ct by enzyme immunoassay (22). Mouse antibodies against this peptide could immunoprecipitate extracted VlsE, but did not bind at detectable levels to the surface of intact *B. burgdorferi* (23). The three-
dimensional structure indicates that this region is membrane proximal and surface-exposed in monomeric VlsE (Fig. 2). However, the close packing of VlsE or binding of VlsE to other protein(s) on the membrane surface may block antibody binding to the lateral surfaces of the protein. In addition, Liang et al. (22) reported that antisera against the C-terminal peptide exhibited poor reactivity with VlsE of other Lyme disease *Borrelia* isolates, although antibodies against IR$_6$ reacted consistently with VlsE in these strains. The interpretation of these results was that the primary structure of C-terminal region of VlsE was poorly conserved among Lyme disease *Borrelia*.

The antigenic properties of the invariant regions (IRs) within the cassette region have been investigated in a series of studies by Liang et al (9,24-26). In peptide-based enzyme-linked immunosorbent assays using sera from infected humans, monkeys, and mice, IR$_6$ was found to be consistently reactive; IR$_2$ was also reactive, but only with sera from infected mice (24). In addition, anti-IR$_6$ mouse antiserum could immunoprecipitate detergent-solubilized VlsE, but did not bind at detectable levels to intact *B. burgdorferi* (9). The three-dimensional structure of VlsE1 reveals that IR$_6$ forms a helix ($\alpha_{10}$) that is almost entirely buried within the center of the membrane distal region. IR$_6$ is second only to IR$_5$ in terms of its low surface exposure to number of residues ratio (14 Å$^3$/residue). In comparison, the VRs have surface area exposure ratios of 46 to 66 Å$^3$/number residues. It is possible that this limited exposure is sufficient to cause immunoprecipitation, but that the IR$_6$ epitope(s) are not accessible in intact bacteria. In our modelling studies, even evoking a high degree of conformational flexibility to the covering loop regions does little to improve the surface accessibility of IR$_6$. Thus, based on the structure, interaction of anti-IR$_6$ antibodies with intact VlsE would be restricted just a few amino acid residues, primarily Lys 276, Gln 279, Lys 291, and Lys 294.
Crystal structure of Lyme Disease Antigen VlsE

Structural comparison

The overall fold of VlsE is very different from any other protein structures solved so far. However, comparison to the outer surface protein C from *Borrelia burgdorferi* (27,28) reveals that both proteins share some general features like neighbored N-/C-termini and long helices forming the membrane proximal portion of the molecule. It is important to note that OspC exhibits considerable sequence heterogeneity between different *borrelia* strains and that these sequence differences tend to be localized on the surface of the membrane distal region, as is the case with the VlsE variable regions. Thus antibody accessibility to regions on the surface of *B. burgdorferi* may be a driving force not only for the evolution of antigenic variation (as in the vls system), but also in the selection of heterogeneity among different strains.

Biological role of VlsE

The function of VlsE is currently unknown, but there is evidence that it is an important virulence factor in the mammalian host. Loss of the encoding plasmid, lp28-1, correlates with greatly reduced infectivity in the mouse model (4,5); studies are underway to determine whether this decrease in infectivity is attributable to the vls system or to other genes present on lp28-1. *vlsE* gene expression is increased following exposure of *B. burgdorferi* to endothelial cell membranes or intact endothelial cells *in vitro* (29), indicating that VlsE may fall into a growing class of genes that are up-regulated in the mammalian environment. Furthermore, *vlsE* recombination occurs at a rapid rate during infection of either immunocompetent or immunodeficient mice, but has not been detected during *in vitro* culture (7,30). It was reported
by Ohnishi et al. (31) that *B. burgdorferi* in ticks inoculated by feeding on infected mice contained *vlsE* sequence variants that either arose within the ticks or were selected during the second feeding. However, *vlsE* recombination was not detected in recent studies in which ticks were inoculated with *B. burgdorferi* clone 5A3 by capillary feeding (32). Thus the recombination mechanism appears to be up-regulated in mammalian tissues as compared to either the arthropod or *in vitro* culture environments.

It is possible that VlsE may fulfil some additional function during mammalian infection that requires surface exposure, and that its antigenic variation permits surface expression without resulting in destruction of the bacterium by the host’s antibody response. The three-dimensional structure of VlsE demonstrates that the membrane-distal surface of the protein is comprised primarily of the variable regions, which undergo rapid sequence changes during infection. The variable regions on the outermost surface may thus mask the invariant regions of the protein. In this manner, the Lyme disease spirochete may stay ‘one step ahead’ of the mammalian host by producing a myriad of variants that do not bind effectively to anti-VlsE antibodies elicited by previous versions of the protein. At present, we can only speculate on how the primarily invariant, lateral surfaces of the protein are protected from antibody binding. These surfaces may be sequestered by interactions with neighboring VlsE molecules or other proteins. A precedent for this type of interaction has been set by studies with the *borrelia* protein P66, in which the protein is protected from proteolytic cleavage in *B. burgdorferi* expressing high levels of the outer surface protein OspA (33). The structure of VlsE will be useful in determining its role in *B. burgdorferi* surface topology and host-parasite interactions.
REFERENCES

1. Steere, A. C. (2001) *N. Engl. J. Med.* **345**, 115-125.

2. Zhang, J. R., Hardham, J. M., Barbour, A. G., and Norris, S. J. (1997) *Cell* **89**, 275-285.

3. Barbour, A., and Restrepo, B. (2000) *Emerg. Infect. Dis.* **6**, 449-457.

4. Labandeira-Rey, M., Baker, E., and Skare, (2001) *J. Infect. Immun.* **69**, 446-455.

5. Purser, J. E., and Norris, S. J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 13865-13870.

6. Zhang, J. R., and Norris, S. J. (1998) *Infect Immun* **66**, 3698-3704.

7. Zhang, J. R., and Norris, S. J. (1998) *Infect. Immun.* **66**, 3689-3697.

8. Lawrenz, M. B., Hardham, J. M., Owens, R. T., Nowakowski, J., Steere, A. C., Wormser, G. P., and Norris, S. J. (1999) *J. Clin. Microbiol.* **37**, 3997-4004.

9. Liang, F. T., Alvarez, A. L, Gu, Y., Nowling, J. M., Ramamoorthy, R., and Philipp, M. T. (1999) *J. Immunol.* **163**, 5566-5573.

10. Liang, F. T., Steere, A. C., Marques, A. R., Johnson, B. J., Miller, J. N., and Philipp, M. T. (1999) *J. Clin. Microbiol.* **37**, 3990-3996.

11. Norris, S. J., Howell, J. K., Garza, S. A., Ferdows, M. S., and Barbour, A. G. (1995) *Infect. Immun.* **63**, 2206-2212.

12. Davies, C., Heath, R., White, S., and Rock, C. (2000) *Structure Fold. Des.* **8**, 185-195.

13. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307-326.

14. Hendrickson, W., J. R, H., and DM, L. (1990) *EMBO J.* **9**, 1665-1672.

15. Terwilliger, T., and Berendzen, J. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849-861.

16. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34-38.
17. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D Biol. Crystallogr. 54, 905-921.

18. Jones, T., Zou, J., Cowan, S., and Kjeldgaard, M. (1991) Acta Crystallogr. A 47, 110-119.

19. Zückert, W. R., Kerentseva, T. A., Lawson, C. L., and Barbour A. G. (2001) J. Biol. Chem. 276, 457-463.

20. Jones, K., Guidry, J., and Wittung-Stafshede, P. (2001) Biochem. Biophys. Res. Commun. 289, 389-394.

21. Christopher, J. A. SPOCK: The Structural Properties Observation and Calculation Kit (Program Manual). The Center for Macromolecular Design, Texas A&M University, College Station, TX, USA. (1998)

22. Liang, F., Bowers, L., and Philipp, M. (2001) Infect. Immun. 69, 3224-3231.

23. Liang, F., Jacobs, M., and Philipp, M. (2001) Infect. Immun. 69, 1337-1343.

24. Liang, F. T., and Philipp, M. T. (1999) Infect. Immun. 67, 6702-6706.

25. Liang, F. T., and Philipp, M. T. (2000) Infect. Immun. 68, 2349-2352.

26. Liang, F. T., Nowling, J. M., and Philipp, M. T. (2000) J. Bacteriol. 182, 3597-3601.

27. Eicken, C., Sharma, V., Klabunde, T., Owens, R. T., Pikas, D. S., Hook, M., and Sacchettini, J. C. (2001) J. Biol. Chem. 276, 10010-10015.

28. Kumaran, D., Eswaramoorthy, S., Luft, B. J., Koide, S., Dunn, J. J., Lawson, C. L., and Swaminathan, S. (2001) EMBO J. 20, 971-978.

29. Hudson, C., Frye, J., Quinn, F., and Gherardini, F. (2001) Mol. Microbiol. 41, 229-239.
30. Sung, S., McDowell, J., and Marconi, R. (2001) *J. Bacteriol.* **183**, 5855-5861.

31. Ohnishi J, P. J., and de Silva A. M. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 670-675.

32. Indest, K. J., Howell, J. K., Jacobs, M. B., Scholl-Meeker, D., Norris, S. J., and Philipp, M. T. (2001) *Infect. Immun.* **69**, 7083-7090.

33. Bunikis, J., and Barbour, A. G. (1999) *Infect. Immun.* **67**, 2874-2883.

34. Guex, N., and Peitsch, M. (1997) *Electrophoresis* **18**, 2714-2723.

35. Barton, G. (1993) *Protein Eng.* **6**, 37-40.
ACKNOWLEDGEMENTS

We are grateful to the APS beamline operators for assistance at beamline 14-BM-D and Julie C. Holding for excellent assistance in the lab. Supported by NIH, the Texas Advanced Technology Program, and the Welch Foundation.
FIGURE LEGENDS

FIG. 1 The structure of VlsE1. a, Ribbon diagram (α-helices in blue, β-strands are orange). The bottom of the figure represents the membrane proximal part. b, The asymmetric unit showing the crystal packing of the four VlsE1 molecules. c, Cα-trace in stereo view. Numbering corresponds to the residues in the full-length sequence of VlsE (see Fig. 3a) All figures have been prepared using the programs SwissPdbViewer (34) and POV-Ray (http://www.povray.org).

FIG. 2 VlsE primary structure domains and tertiary structure. a, The orientation of the molecule is similar to Fig. 1a. The unique conserved N- and C-terminal regions are colored grey, direct repeats are red, invariant regions of the cassette are blue while variable cassette regions are orange. b, Schematic representation of the primary structure. Color code as used for Fig. 2a. c, Dimeric model of VlsE based on the crystal structure (Fig. 1b), illustrating how the formation of potential dimers could effectively shield invariant regions at the monomer-monomer interface.

FIG. 3 Primary and secondary structure of VlsE1. a, Residue positions are numbered according to the VlsE1 B31 strain. The mature protein starts with residue 20. Sequence of recombinant VlsE1 (rvlsE1) used in this study is listed at the top. The secondary structure is superimposed on the bottom of the alignment. Orange triangles mark sequence positions that actually change through cassette exchanges upon antigenic variation. This figure was generated by ALSCRIPT (35). b, Location of the different VRs on the molecular surface of VlsE1. The orientation on the left is identical to Fig. 1a while the model in the middle is rotated by 90° around the vertical axis. The representation on the right shows the membrane distal part of the protein. c, The specific amino acid positions in the cassette region that change during the antigenic variation of VlsE in
*B. burgdorferi* are colored orange on the molecular surface. All sequence positions within VlsE that are strictly conserved through all antigenic variations based on cassette exchanges are grey. Solvent was excluded for all representations during the building of the molecular surface in the program SwissPdbViewer (34).
TABLES

TABLE 1 Data Collection and refinement statistics

Data collection$^a$

| Wavelength (Å) | Resolution (Å) | % Complete | $R_{\text{sym}}^b$ |
|----------------|----------------|------------|-------------------|
| MAD 0.9611 (remote) | 2.80 | 90.5 (79.0) | 0.060 (0.121) |
| 0.9797 (peak) | 2.80 | 97.9 (95.4) | 0.051 (0.083) |
| 0.9800 (inflection) | 2.80 | 90.5 (68.8) | 0.053 (0.121) |
| Refinement 0.9797 (peak) | 2.3 | 92.4 (75.0) | 0.051 (0.138) |

Refinement statistics

| No. of reflections | $R_{\text{cryst}}^c$ |
|--------------------|-----------------------|
| Working set 41,131 | 0.204 |
| Test set 4,597 | $R_{\text{free}}^c$ |
| Average B (Å$^2$) | 40.9 |

Model geometry

| r.m.s.d. from ideal | Protein | 7,984 |
|---------------------|---------|-------|
| Bond length | 0.005 Å | Water |
| Bond angle | 1.08° |

$^a$ Values in parentheses are for the highest resolution shell: 2.90-2.80 Å for the MAD data sets, 2.38-2.30 Å for the high remote data set that was used for refinement.

$^b$ $R_{\text{sym}} = \Sigma h_i \Sigma_i |I(h) - <I(h)>|/\Sigma h_i \Sigma_i <I(h)>$, where $I$ is the observed intensity, and $<I>$ is the average intensity of multiple observations of symmetry-related reflections.

$^c$ $R = \Sigma |F_o| - |F_c|/\Sigma |F_o|$. $R_{\text{cryst}}$ and $R_{\text{free}}$ were calculated using the working and the test reflection sets, respectively.
Figures

Fig. 1:

(a) [Diagram of a protein structure labeled with α and β chains]

(b) [Diagram showing different views of the structure with 90° rotations indicated]

(c) [Diagram of the 300 sequence]

Crystal structure of Lyme Disease Antigen VlsE
Fig. 2:

a, b

![Diagram showing crystal structure of Lyme Disease Antigen VlsE]

- Lipoprotein leader sequence
- Direct repeat
- vls cassette: invariable region (IR)
- Unique conserved regions
- Variable region (VR)

![Diagram showing 2x and 90° transformation]
Fig. 3

a
Fig. 3

b,c