Structure of a major immunogenic site on foot-and-mouth disease virus

Derek Logan*†, Robin Abu-Ghazaleh‡, Wendy Blakemore‡, Stephen Curry‡, Terry Jackson‡, Andrew King‡, Susan Lea*, Richard Lewis*, John Newman‡, Nigel Parry§, David Rowlands§, David Stuart*‖ & Elizabeth Fry*

* Laboratory of Molecular Biophysics, Rex Richards Building, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
‡ AFRC Institute for Animal Health, Ash Road, Pirbright, Woking GU24 0NF, UK
§ Department of Molecular Sciences, Wellcome Research Laboratories, Langley Court, Beckenham BR3 3BS, UK
† Present address: UPR de Biologie Structurale, IBMC du CNRS, 15 rue René Descartes, 67084 Strasbourg Cedex, France
‖ To whom correspondence should be addressed

ATTACHMENT of foot-and-mouth disease virus (FMDV) to its cellular receptor involves a long and highly antigenic loop containing the conserved sequence, Arg-Gly-Asp, a motif known to be a recognition element in many integrin-dependent cell adhesion processes[1–7]. In our original crystal structure of FMDV the Arg-Gly-Asp-containing loop ('the loop'), located between β-strands G and H of capsid protein VP1, was disordered and hence essentially invisible. We previously surmised that its disorder is enhanced by a disulphide bond linking the base of the loop (Cys 134) to Cys 130 of VP2 (ref. 8). We report here the crystal structure of the virus in which this disulphide is reduced. Reduced virus retains infectivity and serological experiments suggest that some of the loop's internal structure is conserved[8]. But here its structure has become sufficiently ordered to allow us to describe an unambiguous conformation, which we relate to some key biological properties of the virus.

To break the disulphide bond, crystals of the virus[9] were soaked in 10 mM dithiothreitol before data collection. Experimental details are presented in Table 1. In the refined structure of the reduced virus, the loop folds snugly against the surface of the particle (Fig. 1), filling a minor surface depression and maintaining the relatively smooth external surface of the virus[10]. On breaking the disulphide bond, there is a large rearrangement of the residues leading into the loop itself; the Cα-atom of residue Cys 134 of VP1 moves by some 12 Å and the released βE–αB loop of VP2 also moves significantly (the Cα of residue 131 moves by over 2 Å), settling into a more ordered conformation. Beyond Cys 134 of VP1, the polypeptide chain adopts an extended conformation, sweeping out over the G–H loop of VP3 (which as a result refolds, with movements of up to 10 Å). The chain then passes back across VP2 towards the 3-fold axis of the icosahedron, residues 144–146 forming an additional strand of β-sheet overlying residues 80–82 in strand C of VP2. An open turn then leads into a 3_{10} helix which heads back

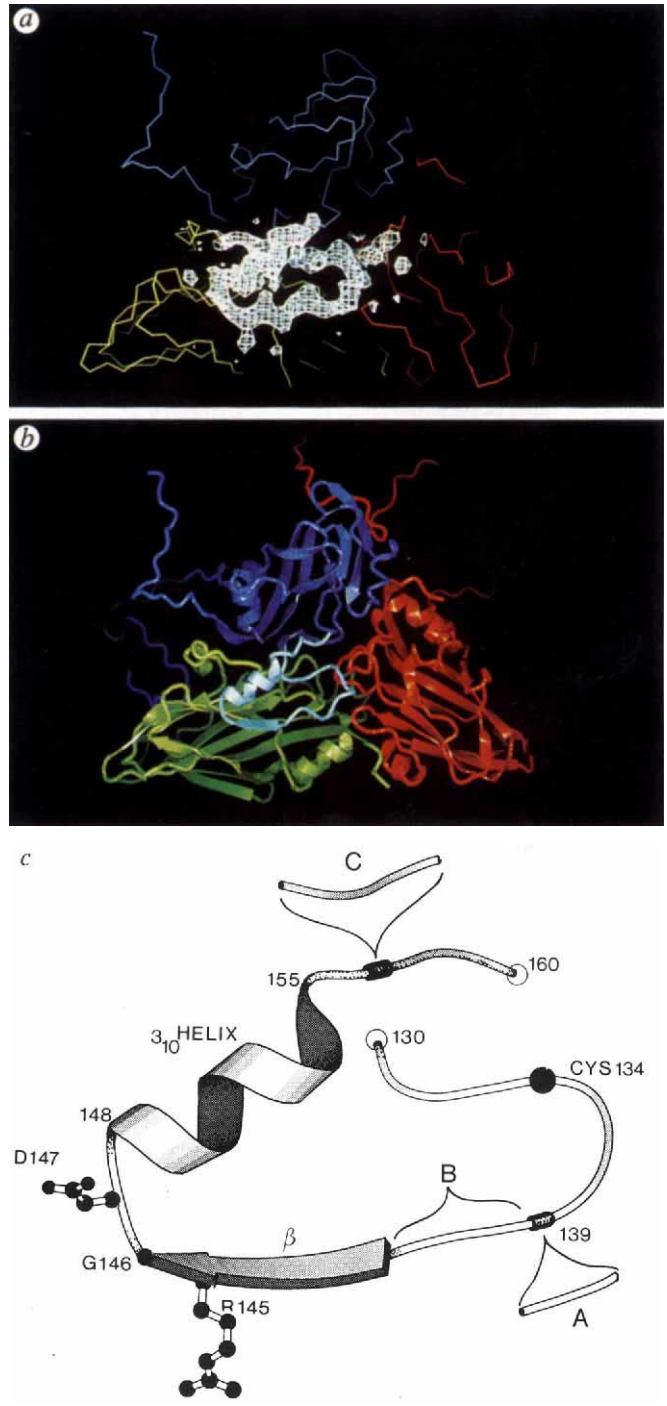


FIG. 1 The structure of the G–H loop of VP1 of FMDV. As in all picornaviruses, the capsid consists of three surface-oriented proteins, VP1–VP3, each with an eight-stranded β-barrel. The surface loops joining the β-strands are biologically important, none more so in FMDV than the G–H loop of VP1. *a*, Experimental evidence for the G–H loop structure. A difference electron density map calculated at 4 Å resolution with amplitudes $F_{h,obs}^{reduced} - F_{h,obs}^{native}$ and native phases (contoured at an arbitrary level and shown as a white mesh) showing clear continuous density defining the structure for the whole of the VP1 G–H loop. In the crystallographic analysis of the native virus, the electron density for the VP1 G–H loop was diffuse or absent between residue 134 (which participates in a disulphide bridge with residue 130 of VP2) and residue 156. The only other features in the difference map correspond to a refolding of the G–H loop of VP3, which interacts with the VP1 G–H loop in the reduced virus, a relaxation of the βE–αB loop of VP2 and the partial reduction of the less accessible disulphide bonds linking residue Cys 7 of VP3 around the 5-fold axis. Otherwise, the native and reduced viruses are almost indistinguishable. The polypeptide chain is coloured: VP1, blue; VP2, green, VP3, red. An electron density map with coefficients $2F_{h,obs}^{reduced} - F_{h,obs}^{native}$ and native phases showed clear side-chain density correlating exactly with the sequence for the loop residues and allowed a model to be built. The view here, and in the other figures, is from the outside of the capsid. *b*, A cartoon sketch[31] illustration locating the loop (pale blue) on the icosahedral asymmetric unit with the proteins colour coded as in *a*. *c*, A molscript[32] ribbon diagram of the VP1 G–H loop with the side chains of the RGD sequence (residues 145–147) depicted using a 'ball and stick' representation. The position of Cys 134 is shown. Insertions and deletions in other serotypes of the virus are roughly located on the basis of sequence alignments: (A), up to three residues inserted in SAT serotypes; (B), up to four residues deleted in C, A and Asia serotypes; (C), up to four residues inserted in SAT serotypes.

towards the 5-fold axis and ends the loop (Fig. 1*b*). Models constructed using predictive[3,11] and spectroscopic methods[12] for several serotypes of FMDV share few of these structural features. Although the G–H loop varies greatly both in sequence and length between the different serotypes of FMDV, the nature of these differences suggests that some general structural features of the loop are likely to be conserved (Fig. 1*c*). But only the O_1 viruses have both Cys residues required to form the disulphide bond between the VP1 G–H loop and VP2, and it will be interesting to see if structural analyses of other types of FMDV reveal an ordered structure for the loop.

Although the structure of the loop in the reduced virus is unambiguous, it remains one of the most mobile regions of the virus surface (as judged from the crystallographic *B*-factors). To assess the stability of the reduced form of the loop, we investigated the kinetics of reoxidation by doing a time-course analysis of the reformation of the disulphide. Two further three-dimensional structures were determined, 3 and 6 days after repeatedly washing the crystals in thiol-free medium to allow reoxidation (data not shown). The electron density maps indicate that the half-time for reformation *in vitro* at room temperature is of the order of 4 days. In purified virus the loop is predominantly disulphide-bonded (Fig. 2); thus this structure is not an artefact of crystallization. To determine at what stage in the life-cycle of the virus this disulphide bond is formed, we investigated the proportion of loops disulphide-bonded in virus at different times after release from infected cells (Fig. 2). These experiments show that, in freshly released virus, the reduced structure predominates and is therefore likely to be a biologically important species; however it seems that within 28 h the majority of the VP1–VP2 bonds have formed (a little more rapid than the timescale for reoxidation observed in the crystals). This agrees with the observation that the infectivity of the reduced virus is not significantly altered, demonstrating that it can still bind to the cellular receptor (data not shown and ref. 8), although the kinetics of cell binding in the reduced and oxidized states have not been compared.

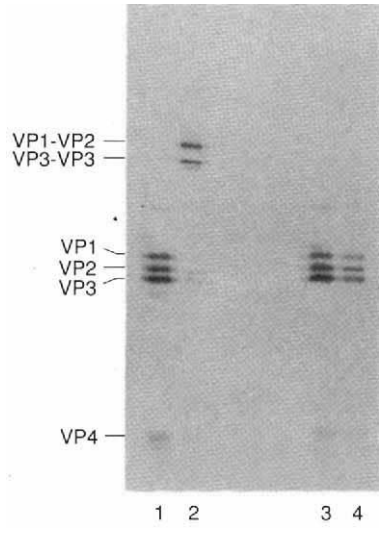


FIG. 2 The formation of disulphide bonds occurs after release of virus from the host cell. Virus-infected BHK-21 cells, labelled with ^{35}S-methionine and cysteine, were lysed in 0.5% Nonidet P-40, 0.04 M $NaPO_4$ pH 7.6, 0.1 M NaCl, and the viruses purified[9] from these lysates were analysed by SDS–PAGE. The virus analysed in lanes 1 and 3 was carboxymethylated by including 50 mM iodacetamide in the lysis solution, so preventing the formation of any disulphide bonds after release from the cell, whereas oxidation was allowed to proceed normally in the virus samples shown in lanes 2 and 4 (analysed 28 h post-release). To reveal disulphide-linked protein dimers, viruses were denatured in the absence of reducing agent (lanes 1, 2). Before loading on the gel, the nonreduced samples were also treated with 50 mM iodacetamide for 2 h to prevent free thiols in the protein exchanging with disulphide bonds during electrophoresis (a reaction that can otherwise lead to serious underestimation of the amount of dimer). Lanes 3, 4 show controls denatured in the presence of 0.4% (v/v) mercaptoethanol.

TABLE 1 Summary of crystallographic data collection and analysis

Data set	Number of filmpacks processed	Completeness to 2.6 Å (%)	$R(I)$ (%) (all data)	R_c (%)	Stereochemistry
Native	88	84.6	14.1%	18.6 (6.0–2.8 Å)	r.m.s. Δ_{bonds}: 0.025 Å r.m.s. Δ_{angles}: 4.2° r.m.s. ΔB_{bonds}: 3.3 Å^2 r.m.s. ΔB_{angles}: 5.3 Å^2
Reduced	14	53.0	16.2%	20.8 (52.0–2.6 Å)	r.m.s. Δ_{bonds}: 0.023 Å r.m.s. Δ_{angles}: 3.7° r.m.s. ΔB_{bonds}: 7.7 Å^2 r.m.s. ΔB_{angles}: 9.6 Å^2

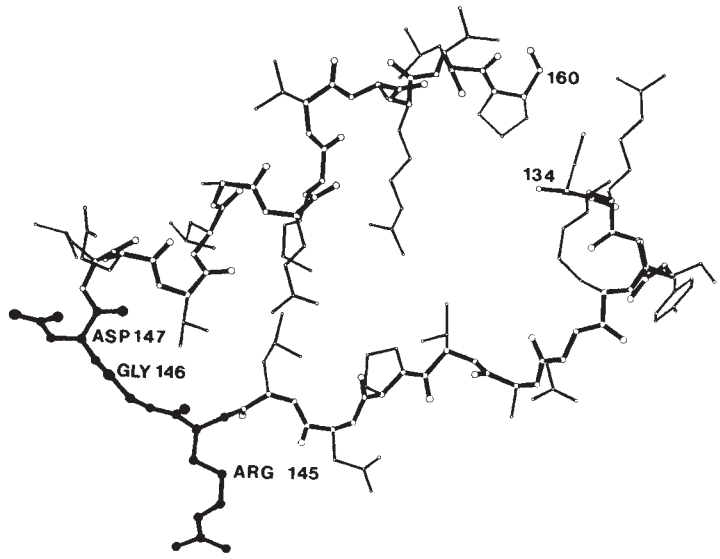
r.m.s.Δ_{bonds} refers to the root mean square deviation from ideal covalent bond lengths. r.m.s. Δ_{angles} refers to the root mean square deviation from ideal covalent bond angles. r.m.s. ΔB_{bonds} refers to the root mean square difference in isotropic *B*-value between covalently bonded atoms. ΔB_{angles} refers to the corresponding difference between covalently bonded next nearest neighbours.

$$R(I) = \frac{\sum_h \sum_i |(I_h - I_{hi})|}{\sum_h \sum_i I_{hi}} \times 100$$

$$R_C = \frac{\sum_h |(|F_{h,obs}| - |F_{h,calc}|)|}{\sum_h |F_{h,obs}|} \times 100$$

I_h is the weighted mean measured intensity of the observations I_{hi}. $F_{h,obs}$ is the observed structure amplitude of reflection *h*. $F_{h,calc}$ is the structure of reflection *h* calculated from the appropriate model. Crystals of the virus[9] were soaked for some 2 h in crystallization mother liquor to which 10 mM dithiothreitol had been added and quickly mounted in sealed quartz capillaries to minimize reoxidation. Data were collected at the SERC Daresbury synchrotron radiation source station 9.6, $\lambda = 0.89$ Å on 0.5° oscillation photographs. The size of the crystals soaked (up to 0.4 mm in the longest dimension) generally exceeded those used in the determination of the native structure[10], and so diffraction data were collected to a resolution of 2.6 Å. The mean fractional isomorphous difference between the native and reduced data was 20.7% on all data. Extra native data were also collected and processed, and compared with those used in the original structure determination as a control; this showed that the new native structure was indistinguishable from the old. The protocols for data collection and processing were as previously described[28] except that in this case partially recorded reflections on films from adjacent oscillation ranges were summed (two film packs were obtained from several crystals), and those reflections more than 80% recorded on a single film were scaled to 100%. The virus structure was refined against the data collected for the reduced virus with the package X-PLOR, version 3.0[29]. No account was taken of the scattering from bound waters; however, a bulk solvent correction was applied, hence the good agreement of the model over a wider resolution range than the native. A tight solvent mask was defined, with a sufficiently high average value for the solvent density to account for bulk solvent and RNA. All data including those recorded as negative intensities were retained throughout the processing and included in the refinement[30]. Exact 5-fold non-crystallographic symmetry constraints were included throughout the refinement of atomic positions and individual *B*-values. Using the full resolution range of the data and a suitably weighted bulk solvent mask it was possible to determine the relative occupancy of the G–H loops of VP1 and VP3 in the reduced structure (0.85 ± 0.15), giving us greater confidence in the *B*-factors of these regions (methods will be published elsewhere). Coordinates for the native and reduced virus structures have been deposited with the protein data bank[18].

FMDV was the first virus for which an integrin was implicated in cell attachment[4–6,13], although the exact identity of its receptor remains unknown. Recent results indicate that some other picornaviruses use receptors of this family[14,15]. Unfortunately there are as yet no direct structural results on integrin–ligand interactions. The residues Arg-Gly-Asp (RGD) are of particular importance in integrin recognition[16,17], and in FMDV this sequence (residues 145–147) is highly conserved and centrally placed in the G–H loop. We have compared the conformation of this triplet in FMDV with nine other occurrences of the motif in the Protein Data Bank[18] (Fig. 3). Three of these are reported to have been assayed for integrin binding activity *in vitro*[17]. The RGD motif in FMDV is very similar to that of γII-crystallin[19], the only one to bind an integrin actively, and differs from those of the inactive proteins[20,21] (for details see Fig. 3). Of the other RGD motifs in the Protein Data Bank, the one located in an external loop of the coat protein VP1 of human rhinovirus 1A (ref. 22) (whose receptor is unknown) is the most similar in structure to the motif in FMDV, ranking alongside γII-crystallin. In both γII-crystallin and FMDV, the acidic and basic side



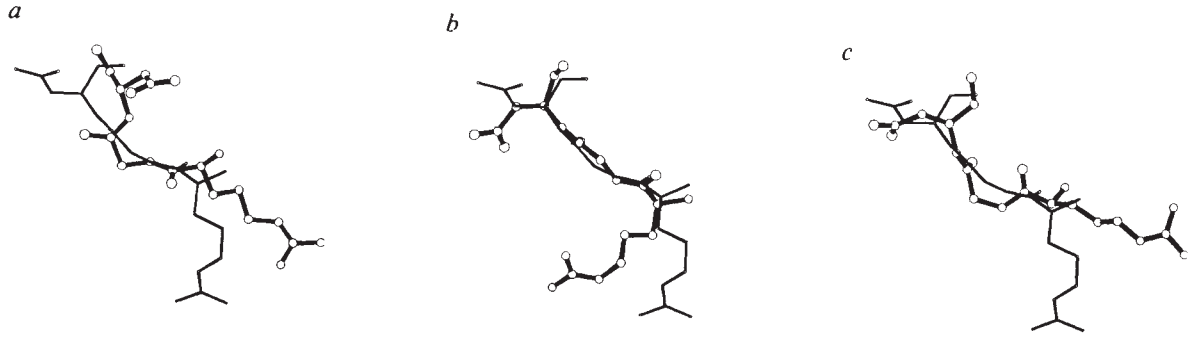


FIG. 3 An all-atom trace of the residues 134–160 of the G–H loop of VP1 is shown with the spheres marking the atoms of the RGD residues filled in. Comparisons of the RGD motif in FMDV are then made with α-lytic protease (*a*), γ-crystallin (*b*) and thermolysin (*c*), where the latter three structures have spheres defining the atom positions. A superimposition of the Cαs only for each of the nine RGD structures from the protein data base[18] with that in FMDV gave the following r.m.s. deviations: γII-crystallin[19] (0.09 Å), human rhinovirus 1A[22] VP1 (0.11 Å) and VP2 (0.16 Å), triacylglycerol acylhydrolase[33] (0.53 Å), thermolysin[21] (0.53 Å), xylose isomerase[34] (0.55 Å), tryptophan synthase[35] (0.77 Å), α-lytic protease[20] (0.89 Å) and *Eco* RI endonuclease[36] (1.16 Å).

chains of this tripeptide are in an 'open' conformation, well separated from each other. Nuclear magnetic resonance (NMR) studies on peptide inhibitors of integrins suggest that such a conformation enhances their potency[23,24]. Recently, structural information has become available for the disintegrins, a family of polypeptide integrin inhibitors[25,26]. A common feature of these is the presence of an RGD motif in a mobile loop, a situation similar to that in FMDV. This mobility has made it difficult to determine the structure of the RGD region in the disintegrins. We are fortunate that in FMDV the RGD triplet is stabilized by the involvement of residues 145–146 in β-sheet interactions (the mean *B*-factors for the backbone atoms of the RGD being lower than the average for the loop). A further active RGD has recently been visualized, that of the fibronectin type III domain from tenascin[27]. This is also found to be in an open conformation at the turn of an extended loop although it is not yet possible to make a detailed comparison.

The VP1 G–H loop seems to act as a self-contained unit, probably with some internal flexibility, hinged to the rest of the structure. Because the loop structure alters after release of virus from cells, time-dependent antigenic variation seems to be occurring in this virus. The structure for this loop suggests that part of the specificity of integrin–ligand interaction derives from the particular conformation of the polypeptide backbone of the RGD residues of the ligand existing on a mobile exposed loop. □

Received 30 November 1992; accepted 3 February 1993.

1. Strohmaier, K., Franze, R. & Adam, K. H. *J. gen. Virol.* **59,** 295–306 (1982).
2. Bittle, J. L. *et al. Nature* **298,** 30–33 (1982).
3. Pfaff, E., Mussgay, M., Bohm, H. O., Schulz, G. E. & Schaller, H. *EMBO J.* **1,** 869–874 (1982).
4. Surovoi, A. Y., Ivanov, V. T., Chepurkin, A. V., Ivanyuschenkov, V. N. & Dryagalin, N. N. *Soviet J. Bioorg. Chem.* **14,** 572–580 (1989).
5. Acharya, R. *et al.* in *Vaccines 89: Modern Approaches to New Vaccines Including Prevention of AIDS* (Cold Spring Harbor Laboratory Press, New York, 1989).
6. Fox, G. *et al. J. gen. Virol.* **70,** 625–637 (1989).
7. Baxt, B. & Becker, Y. *Virus Genes* **4,** 73–80 (1990).
8. Parry, N. R. *et al. Nature* **347,** 569–572 (1990).
9. Fox, G. *et al. J. molec. Biol.* **196,** 591–597 (1987).
10. Acharya, R. *et al. Nature* **337,** 709–716 (1989).
11. Rowlands, D. *et al. Nature* **306,** 694–697 (1983).
12. Siligardi, G. *et al. Eur. J. Biochem.* **199,** 545–551 (1991).
13. Geysen, H. M., Barteling, S. J. & Meloen, R. H. *Proc. natn. Acad. Sci. U.S.A.* **82,** 178–182 (1985).
14. Chang, K. H., Day, C., Walker, J., Hyypiä, T. & Stanway, G. *J. gen. Virol.* **73,** 621–626 (1992).
15. Bergelson, J. M., Shepley, M. P., Chan, B. M. C., Hemler, M. E. & Finberg, R. W. *Science* **255,** 1718–1720 (1992).
16. Piersbacher, M. D. & Ruoslahti, E. *Nature* **309,** 30–33 (1984).
17. Ruoslahti, E. *A. Rev. Biochem.* **57,** 375–413 (1988).
18. Bernstein, F. C. *et al. J. molec. Biol.* **112,** 535–542 (1977).
19. Wistow, G. *et al. J. molec. Biol.* **170,** 175–202 (1983).
20. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. *J. molec. Biol.* **184,** 479–502 (1985).
21. Holmes, M. A. & Matthews, B. W. *J. molec. Biol.* **160,** 623–639 (1982).
22. Kim, S. *et al. J. molec. Biol.* **210,** 91–111 (1989).
23. Aumailley, M. *et al. FEBS Lett.* **291,** 50–54 (1991).
24. Reed, J. *et al. Eur. J. Biochem.* **178,** 141–154 (1988).
25. Saudek, V., Atkinson, R. A. & Pelton, J. T. *Biochemistry* **30,** 7369–7372 (1991).
26. Adler, M., Lazarus, R. A., Dennis, M. S. & Wagner, G. *Science* **235,** 445–448 (1991).
27. Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P. *Science* **258,** 987–991 (1992).
28. Fry, E., Acharya, R. & Stuart, D. *Acta crystallogr.* A**49,** 45–55 (1993).
29. Brunger, A. T., Kuriyan, J. & Karplus, M. *Science* **235,** 458–460 (1987).
30. French, S. & Wilson, K. *Acta crystallogr.* A**34,** 517–525 (1978).
31. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* A**47,** 110–119 (1991).
32. Kraulis P. J. *J. appl. Crystallogr.* **24,** 946–950 (1991).
33. Brady, L. *et al. Nature* **343,** 767–770 (1990).
34. Henrick, K., Blow, D. M., Carrell, H. L. & Glusker, J. P. *Protein Eng.* **1,** 467–469 (1987).
35. Hyde, C. C. & Miles, E. W. *Biotechnology* **8,** 27–31 (1990).
36. Kim, Y., Grable, J. C., Love, R., Greene, P. J. & Rosenberg, J. M. *Science* **249,** 1307–1309 (1990).

ACKNOWLEDGEMENTS. We thank D. Goodrich for his time on data collection trips as disease security officer; the staff of the Synchrotron Radiation Source, SERC Daresbury Laboratory for practical assistance; R. Acharya, F. Brown and G. Fox for interest in the project; R. Bryan, R. Esnouf, L. Johnson, S. Lee and others from the Laboratory of Molecular Biophysics and Institute for Animal Health for their support. D. L. and E. F. were supported by the Medical Research Council, D.L. also by the Carnegie Trust; S.L., R.A., W.B., S.C. and T.J. were supported by the AFRC; D.S. is a member of the Oxford Centre for Molecular Sciences.