Acta Cryst. (1997). D53, 186–194

Structure of the Fab Fragment from a Neutralizing Monoclonal Antibody Directed Against an Epitope of gp41 from HIV-1

CHRISTOPHER DAVIES,† JEREMY C. BEAUCHAMP,‡ DAVID EMERY,§ AHMAD RAWAS AND HILARY MUIRHEAD*

Department of Biochemistry and Molecular Recognition Centre, University of Bristol, Bristol BS8 1TD, England. E-mail: muirhead@bsa.bristol.ac.uk

(*Received 9 July 1996; accepted 27 September 1996*)

Abstract

The structure of a Fab fragment of a monoclonal antibody (1583) that neutralizes a broad range of HIV-1 isolates has been solved by X-ray crystallography. This antibody is directed against a poliovirus/HIV-1 chimaera which presents a conserved epitope of the envelope protein gp41. Crystals of 1583 were obtained in the space group $P2_12_12_1$ and the structure solved by molecular replacement. The model has been refined against all data in the range 10–2.9 Å to a final crystallographic R factor of 0.198. The antigen-binding site features a well defined groove, typical of antibodies that bind to small antigens, created in part by a relatively short CDR H3. The variable regions of 1583 were sequenced and, given the hydrophilic nature of the epitope, revealed a surprising lack of charged residues in the CDR's. However, the antigen-binding cleft is indeed very polar, due in part to the presence of two charged residues that emanate from outside the recognised CDR's.

1. Abbreviations

AIDS, acquired immune deficiency syndrome; HIV-1, human immunodeficiency virus type 1; HIV-2, human immunodeficiency virus type 2; ELISA, enzyme-linked immunosorbant assay; Mab, monoclonal antibody; PND, principal neutralizing domain; Fab, fragment of antibody; r.m.s., root mean square; CDR, complementarity determining region; SDS–PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; FPLC, fast protein liquid chromatography.

2. Introduction

The production of neutralizing antibodies directed against specific regions of the HIV-1 envelope has raised hopes that it may be possible to design vaccines based on the structure of the epitopes involved or perhaps to generate neutralizing antibodies which have therapeutic value. Most of the studies thus far have involved sites of the external glycoproteins of HIV-1, namely gp120 and gp41, which are formed following cleavage of the precursor protein gp160. In particular, the PND of gp120, the so-called V3 loop, has been extensively studied (Rini, Schulze-Gahmen & Wilson, 1992; Ghiara, Stura, Stanfield, Profy & Wilson, 1994). However, the variability of the epitopes of HIV-1 amongst the various isolates of the virus has hampered progress since the neutralization has tended to be strain specific. Clearly, an epitope which elicits more broadly reactive antibodies may be of more importance. Two such regions have been identified in the transmembrane glycoprotein of the viral envelope, gp41. Both of these are promising targets since they are relatively well conserved. One of these is the so-called 'Kennedy' epitope comprising residues 735–752 (Kennedy *et al.*, 1986) (numbering according to isolate HTLV-III) and the other is within residues 647 to 671 (Broliden *et al.*, 1992). Any structural information regarding these epitopes would be of great importance in order to identify those amino acids primarily responsible for the elicitation of neutralizing antibodies. The work presented here concerns the Kennedy epitope.

The Kennedy epitope forms the cytoplasmic anchor of the transmembrane segment of gp41 (Modrow *et al.*, 1987) and probably remains hidden within the virion particle for much of the viral life cycle. However, since it is recognized by antibodies from HIV-1 positive sera (Kennedy *et al.*, 1986; Broliden *et al.* 1992), it presumably becomes exposed after CD4 binding. Antibodies raised against a synthetic peptide corresponding to this sequence both recognize gp160 in an ELISA (Kennedy *et al.*, 1986) and neutralize HIV-1 infectivity *in vitro* (Chanh *et al.*, 1986; Kennedy *et al.*, 1986). The mechanism of neutralization, therefore, may be to block virion-cell fusion. Evidence that these antibodies manifested broad reactivity was revealed in a further study where the monoclonal antibodies were found to neutralize six isolates of HIV-1 (Dalgleish *et al.*, 1988). The Kennedy epitope has since been engineered into antigenic site 1 of the viral capsid (VP1) of Sabin 1 poliovirus (Evans *et al.*, 1989). Antibodies raised against this chimaera (S1/env/3) also neutralize a broad range of HIV-1 isolates (Evans *et al.*, 1989; Vella *et al.*, 1993).

† Present address: Department of Microbiology, Duke University Medical Center, Durham, NC 27710, USA.

‡ Present address: Department of Chemistry, University of Glasgow, Scotland.

§ Present address: National Institute for Medical Research, Mill Hill, London NW7 1AA, England.

© 1997 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 © 1997

```
tccgctagcatgacccaaactccactcactttgtcggttaccattggacaaccagcctcc
SerAlaSerMetThrGlnThrProLeuThrLeuSerValThrIleGlyGlnProAlaSer
AspValValMetThrGlnThrProLeuThrLeuSerValThrIleGlyGlnProAlaSer
1                              10                              20

atctcttgcaagtcaagtcagagcctcttagatagtgatggaaagacgtatttaaattgg
IleSerCysLysSerSerGlnSerLeuLeuAspSerAspGlyLysThrTyrLeuAsnTrp
IleSerCys
          24 25 26 27 a  b  c  d  e  28 29 30 31 32 33 34
          <------------------- CDR L1 ------------------>
ttgttacagaggccaggccagtctccaaagcgcctaatctatctggtgtctaaactggac
LeuLeuGlnArgProGlyGlnSerProLysArgLeuIleTyrLeuValSerLysLeuAsp
             40             45             50 51 52 53 54 55
                                           <----- CDR L2 ---

tctggagtccctgacaggttcactggcagtggatcagggacagatttcacactgaaaatc
SerGlyValProAspArgPheThrGlySerGlySerGlyThrAspPheThrLeuLysIle
 56          60             65             70             75
-->

agcagagtggaggctgaggatttgggagtttattattgctggcaaggtacacattttcct
SerArgValGluAlaGluAspLeuGlyValTyrTyrCysTrpGlnGlyThrHisPhePro
             80             85          89 90 91 92 93 94 95
                                       <-------- CDR L3 ----

cggacgttcggtggaggcaccaagctggaaatcaaacgggctgatgctgcaccaactgta
ArgThrPheGlyGlyGlyThrLysLeuGluIleLysArgAlaAspAlaAlaProThrVal
 96 97      100            105            110            115
----->

tccatcttcccaccatccacctgggatcc
SerIlePheProProSerThrTrpAspPro
            120            125
```

(*a*)

```
caggtcaagctgcagcagtctggacctggcctagtgcagccctcacagagcctgtccatc
GlnValLysLeuGlnGlnSerGlyProGlyLeuValGlnProSerGlnSerLeuSerIle
  1                              10                              20

acctgcacagtctctggtttctcattaacttgctatggtgtacactgggttcgccagtct
ThrCysThrValSerGlyPheSerLeuThrCysTyrGlyValHisTrpValArgGlnSer
                                 31 32 33 34 35              40
                                 <---CDR H1 --->

ccaggaaagggtctggagtggctgggagtgatatggagtggtggagacacagactataat
ProGlyLysGlyLeuGluTrpLeuGlyValIleTrpSerGlyGlyAspThrAspTyrAsn
                               50 51 52 53 54 55 56 57 58 59 60
                              <------------------- CDR H2 ------

gccgctttcatatccagactgagcatcaccaaggacaattccaagagccaagttttcttt
AlaAlaPheIleSerArgLeuSerIleThrLysAspAsnSerLysSerGlnValPhePhe
 61 62 63 64 65              70                              80
-------------->

aaaatgaacagtctgcaacctaatgacagagccatatattactgtgccagacgggggggg
LysMetAsnSerLeuGlnProAsnAspArgAlaIleTyrTyrCysAlaArgArgGlyGly
       A  B  C                        90             95 96 97
                                                    <-- CDRH3

gacttctggggccaagggaccacggtcaccgtctcctca
AspPheTrpGlyGlnGlyThrThrValThrValSerSer
101102                         110
----->
```

(*b*)

Fig. 1. The cDNA sequence of the (*a*) light and (*b*) heavy chains of the variable domain with the translated amino-acid sequence. Numbering refers to the amino-acid sequence and is based on Kabat, Wu, Perry, Gottesman & Foeller (1991). The sequences marked in italics correspond to the PCR primers including those amino acids partially or completely encoded by primer. The first 23 residues of the kappa light chain, which were determined by protein sequencing of purified 1583, are also shown underneath. Residues in the complementarity determining regions as defined by Kabat *et al.* (1991) are underlined and numbered.

That the peptide, irrespective of the means of presentation, induces neutralizing antibodies is evidence that it readily adopts a structure which is conformationally close to that occurring in the native gp41 epitope. Given this, it was surprising that NMR experiments in our laboratory showed no detectable structure for the peptide in solution (unpublished results).

As a prelude to an investigation of this important epitope, we present here the three-dimensional structure of a Fab fragment derived from one of the anti-S1/env/3 antibodies (termed 1583), which has been crystallized and solved by molecular replacement at a resolution of 2.9 Å. It reveals an antigen-binding site which is highly polar in character and suggests residues that might interact with the peptide antigen.

3. Materials and methods

3.1. *Sequencing of the heavy- and light-chain variable regions*

An adaptation of the method of Orlandi, Gussow, Jones & Winter (1989) was followed to obtain the sequence of the heavy and light chains of the variable domain of 1583. Total mRNA was extracted from approximately 10^7 confluent 1583 hybridoma cells by the method of Sambrook, Fritsch & Maniatis (1989) and used to produce full-length cDNA. The cDNA was used as a template DNA for 30 cycles of PCR amplification. Variable domain DNA was purified by agarose-gel purification and subcloned into pUC19 vector for sequencing by a DuPont Genesis-2000 automated DNA sequencer. The entire procedure was repeated in order to verify the resulting sequences.

3.2. *Antibody and Fab purification*

The antibody 1583 (IgG2a-K) was purified from ascitic fluid kindly provided by Dr Morag Ferguson of the National Institute for Biological Standards and Control (NIBSC), South Mimms, Hertfordshire, UK. Ascitic fluid was diluted by a factor of four in 50 m*M* sodium phosphate–HCl buffer (pH 7) and applied to a Protein G HR 10/2 FPLC column (Pharmacia) *via* a 0.45 µm filter. Fractions containing IgG were eluted by 0.1 *M* glycine, pH 2.7 and pooled. Glycerol to 20%(*v*/*v*) was added immediately followed by the addition of 0.1 *M* sodium hydroxide dropwise in order to bring the sample to neutral pH.

The Fab fragments were generated using insolubilized papain (Sigma) according to the method of Porter (1959). The papain beads were washed several times in deionized water and then added to the antibody in a ratio of 1:1 by weight. Typically 40–50 mg of antibody were digested in one experiment. The progress of the reaction was monitored by FPLC ion-exchange chromatography (Mono S, Pharmacia) and the reaction terminated by centrifugation of the reaction mixture to remove the papain beads. Experience showed that the generation of Fab isoforms could be virtually eliminated by termination just prior to complete digestion of the mAb, generally after 2–3 h of digestion at 310 K.

The Fab fragments were purified from the reaction mixture in two steps. Firstly, the reaction mixture was applied to a Sephacryl HR-200 (Pharmacia) gel-filtration column equilibrated in 50 m*M* sodium acetate, pH 5.0. The fractions which contained Fab fragments were determined by SDS–PAGE and further purified, to remove residual Fc fragments, using a Mono S HR

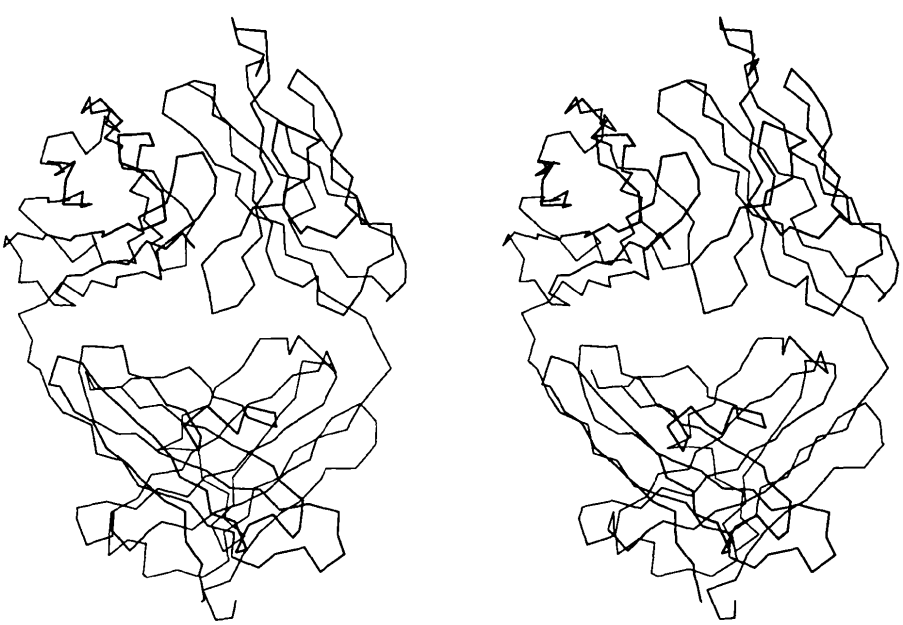

Fig. 2. A stereo Cα trace of the refined model of 1583 showing the immunoglobulin fold.

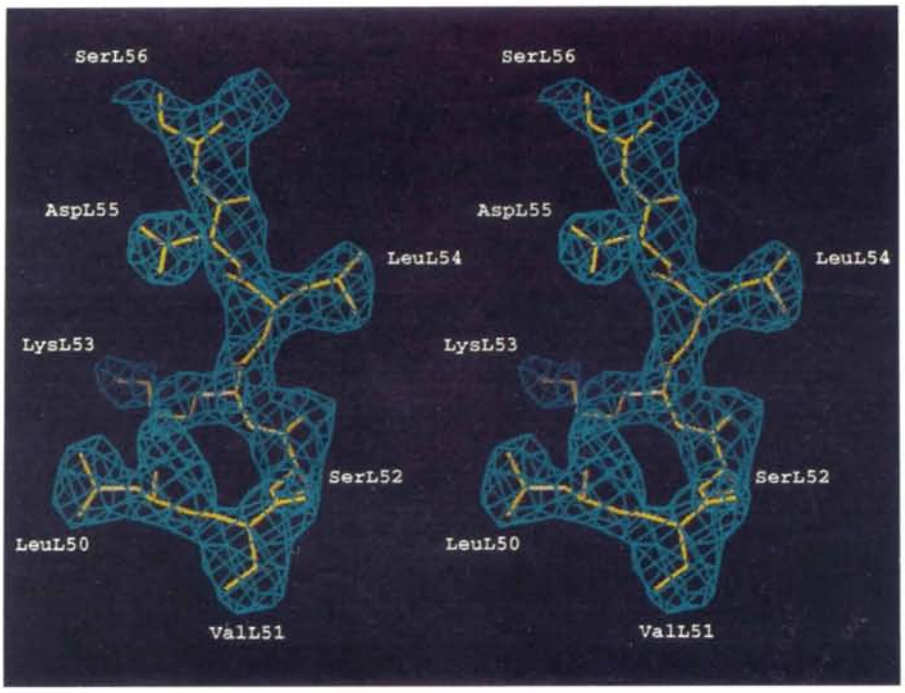

Fig. 3. A portion of the final $2F_o - F_c$ electron-density map contoured at 1.0σ. The region shown is the CDR L2 comprising residues L50–L56.

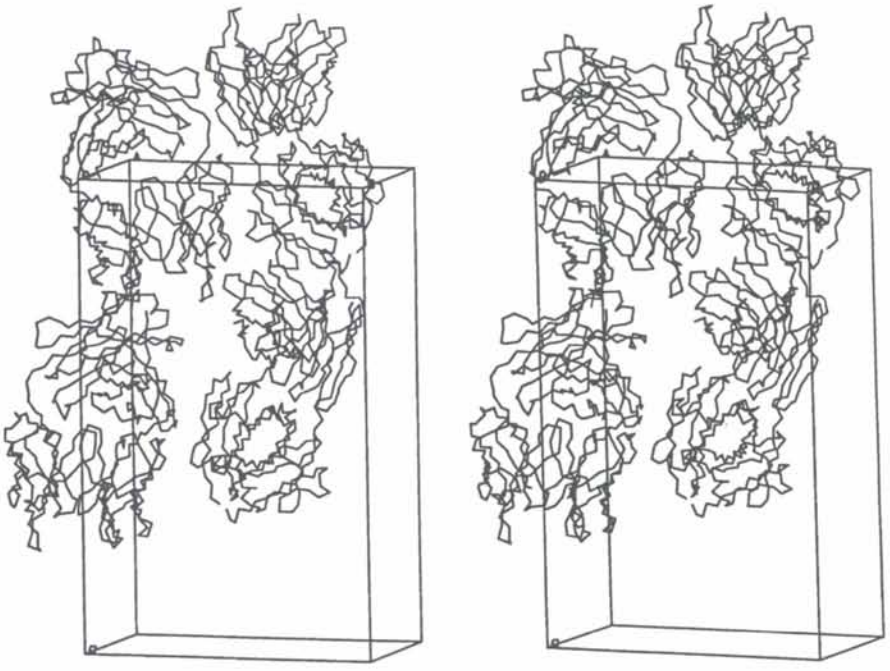

Fig. 4. Stereo drawing of the Cα trace of four 1583 Fab molecules showing their packing in the crystallographic unit cell. Along the *c* axis the molecules adopt a head-to-tail arrangement where the variable domain of one Fab packs against the constant domain of an adjacent molecule.

735 752

D R P E G I E E E G G E R D R D R S

Fig. 5. The amino-acid sequence of the Kennedy epitope (Kennedy *et al.*, 1986). The numbering along the top corresponds to that in gp160 of the HIV-1 isolate HTLV-III. The underlined residues correspond to the minimal epitope recognized by 1583 (Vella *et al.*, 1993).

5/5 ion-exchange chromatography column (Pharmacia), equlibrated in the same buffer.

3.3. *Crystallization*

The purified Fab fragments were dialysed against distilled water and concentrated to 5 mg ml^{-1} using

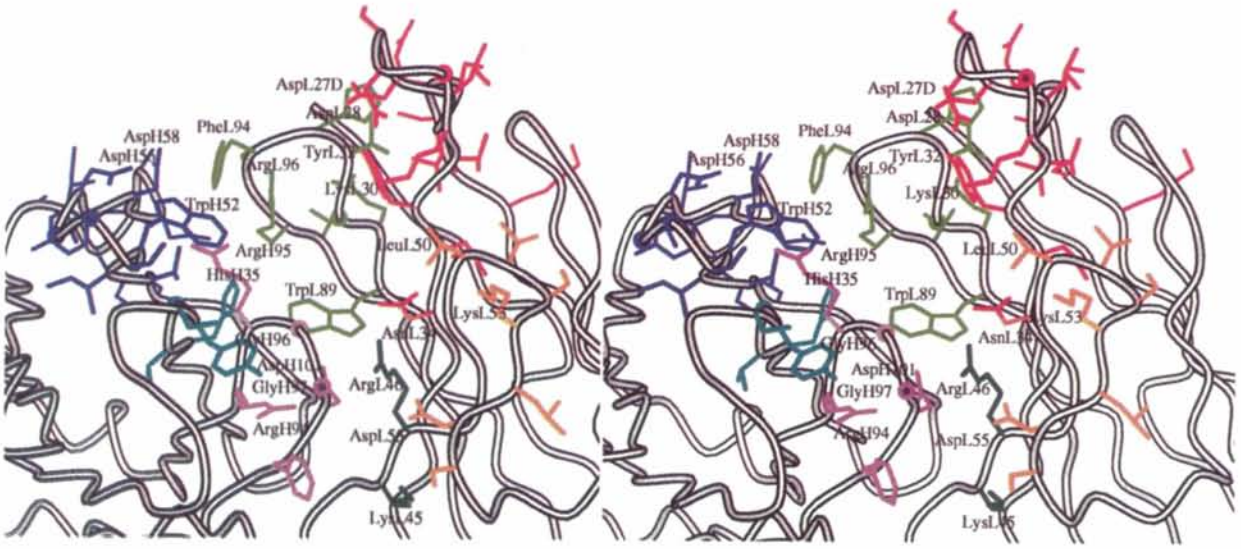

Fig. 6. A stereo representation of the antigen-binding site. All of the residues present in the CDR's are shown in stick form with the addition of ArgH94, LysL45 and ArgL46 which originate from outside the recognized CDR's. Residues mentioned in the text are labelled. Residues are coloured differently according to the CDR: red = L1, orange = L2, green = L3, blue = H1, purple = H2 and pink = H3.

a Centricon-10 microconcentrator (Amicon). Crystallization trials were performed using the hanging-drop vapour-diffusion method. A wide range of conditions in terms of pH, concentration, type of precipitant and temperature were tested. In each case 3 µl of the protein solution was mixed with 3 µl of the well solution on plastic cover slips which were inverted and sealed over the wells. Crystals were obtained using PEG 4000 at well concentrations in the range 10–25%(*w*/*v*) and a pH range of 5.0–5.5 (200 m*M* MES) and 6.0–6.5 (200 m*M* BisTris). The crystals were plate-like in appearance and grew to a size of 0.3 × 0.2 × 0.05 mm following a period of 2–3 d incubation at 291 K.

3.4. *X-ray analysis*

The crystals used for the diffraction experiments were harvested from the wells containing 20% PEG 4000, 200 m*M* BisTris, pH 6.5. These were characterized using a Siemens multiwire area detector mounted on a Siemens rotating anode operating at 45 kV and 90 mA. The crystal-to-detector distance was 120 mm and a graphite monochromator produced Cu $K\alpha$ X-rays ($\lambda = 1.54$ Å). Using the *Xengen* processing software package (Howard *et al.*, 1987), the cell dimensions were determined to be $a = 38.15$, $b = 82.55$ and $c = 131.0$ Å in an orthorhombic space group, $P2_12_12_1$. The cell dimensions together with the assumption of a molecular weight of 50 kDa for the Fab indicated that there was one Fab molecule per crystallographic asymmetric unit. The volume per unit mass, V_m, was 2.1 Å^3 Da^{-1} which falls within the range (1.68–3.53 Å^3 Da^{-1}) observed for protein crystals (Matthews, 1968). This corresponds to a solvent content of 46%.

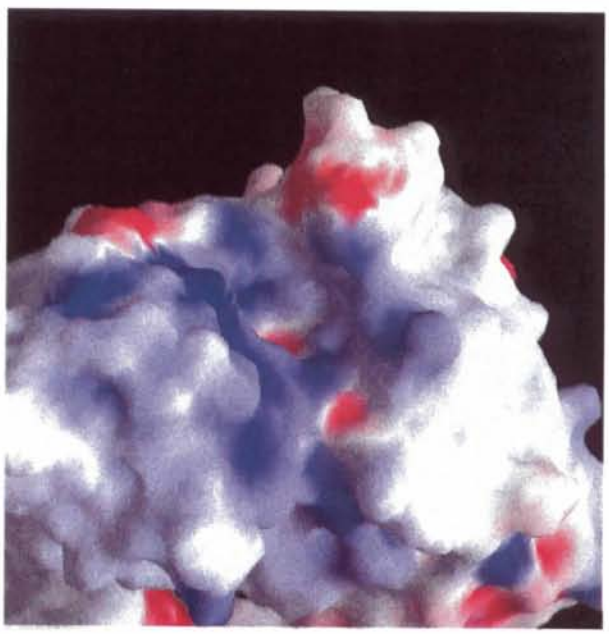

Fig. 7. The electrostatic surface potential of the 1583 antigen-binding region calculated using the program *GRASP* (Nicholls, Sharp & Konig, 1991). The view is the same as that in Fig. 6. Regions that are basic in character are coloured blue and those of acidic character are red.

Native data were collected at room temperature from a single crystal by the oscillation method at two settings

Table 1. *The significance statistics of the native data*

Resolution (Å) (upper limit)	R_{sym}*	Total No. of observed reflections	No. of unique reflections	Completeness of the data (%)	Completeness of the data with $I > 2\sigma$ (%)
5.24	6.60	8342	1695	94.9	90.8
4.16	7.07	5447	1651	99.5	94.3
3.63	11.57	4667	1631	100.0	92.4
3.30	14.08	4263	1626	100.0	87.0
3.06	20.95	3753	1609	99.2	78.1
2.88	25.10	1172	746	51.3	32.7
Totals	10.16	27644	8958	91.6	83.1

*$R_{sym} = \sum |I(h)i - \langle I(h)i \rangle| / \sum \langle I(h)i \rangle$, where $I(h)i$ is the observed intensity and $\langle I(h) \rangle$ is the mean intensity of reflection h over all measurements.

of 2θ, 0 and 8°. In each case 180° of data were measured consisting of 720 frames (exposure time = 4 min) where $\delta\varphi$ was 0.25°. The data were reduced using *Xengen*. The scaling agreement factor, $R_{sym}(I) = 10.16\%$ on intensities for 27 644 observations giving rise to 8958 unique reflexions for all data to 2.88 Å. These data are 91.6% complete to 2.88 Å and 83.1% of the reflections have $I > 2\sigma$ (Table 1). The weaker statistics in the final shell indicate that the effective resolution of the structure determination is approximately 3.0 Å.

Table 2. *Refinement parameters of the final model*

Resolution range (Å)	8.0–2.9
σ cut-off applied	0.0
No. of reflections	8042
R factor (%)	19.8
R free (%)	32.6
R.m.s. deviations in bond lengths (Å)	0.008
R.m.s. deviations in bond angles (°)	1.7
Mean B factor (main chain) (Å^2)	16.32
R.m.s. deviation in B factor (main chain) (Å^2)	3.26
Mean B factor (side chain) (Å^2)	17.54
R.m.s. deviation in B factor (side chain) (Å^2)	3.87
No. of non-H atoms	3312

3.5. *Structure determination*

The structure was solved by molecular replacement using the refined structure of FabD1.3 (Fischmann *et al.*, 1991) as a search model in *X-PLOR* (Brünger, 1990). This model was selected on the basis of the closest homology of the heavy chains (at that stage the light chain had not yet been sequenced). The model was modified by removing the parts of the structure corresponding to the CDR's. At first a solution was attempted without any modification to the existing elbow angle (the angle made by the pseudo-twofold axes relating V_H to V_L and C_H1 to C_L) which was 176° (Amit, Mariuzza, Philips & Poljak, 1986). The rotation function was calculated using data in the range 15–4 Å where the model was transformed in a $P1$ box of dimensions 120 × 120 × 150 Å. The resulting 308 peaks were tested by Patterson correlation (PC) refinement where the groups selected were (*a*) the whole molecule; (*b*) the variable (V) and constant (C) domains; and (*c*) the individual V_L, V_H, C_L and C_H1 domains. One peak (number 141/338) rose above the background and this orientation was tested in a translation function. A clear peak arose with a Tf value of 0.25 and an R factor of 0.527 (15–4 Å) which reduced to 0.504 (10–3 Å) upon rigid-body minimization using the same groups as the PC refinement. However, an examination of the packing of this solution revealed areas of overlap between symmetry-related molecules. It became evident that a modification to the elbow angle of the Fab would relieve the bad contacts and, therefore, a new search was performed using two models with modified elbow angles, 155 and 130°, respectively. Following the same procedure as above these two structures converged to an identical solution, albeit with different origins in the unit cell. In the case of the 130° search model rotation function three peaks (out of 332) rose significantly above background in the PC refinement with a function value around 0.17. These were in fact symmetry-related solutions and the highest peak (number 37) was selected for the translation function. This generated a solution of Tf value of 0.535 with an R factor of 0.425 after rigid-body refinement. The packing was free of any bad contacts and, therefore, this solution was judged to be correct.

3.6. *Refinement*

The 130° model was refined by *X-PLOR* with a standard protocol using simulated annealing (Brünger, 1992) which generated a model with an R factor of 0.21 against all data in the range 10–2.9 Å with good geometry. Using the resulting $2F_o - F_c$ map the model was inspected and corrected using *FRODO* (Jones, 1985) based on the known sequence for the variable domain light and heavy chains. The six CDR's, which had been omitted from the search model, were also fitted to the electron density. After several rounds of model building and refinement, using both *TNT* (Tronrud, Ten Eyck & Matthews, 1987) and *X-PLOR*, the final R factor was 0.193 and $R_{free} = 0.32$ with excellent geometry. The parameters of the final model are shown in Table 2. The viability of the model was verified by *PROCHECK*

(Laskowski, MacArthur, Moss & Thornton, 1993). The numbering used throughout this paper corresponds to that of Kabat, Wu, Perry, Gottesman & Foeller (1991).

4. Results and discussion

4.1. *Sequencing*

The sequences of the heavy and light chain of the Fv domain were determined by cloning and sequencing of the cDNA and are shown in Fig. 1. The hypervariable loops as defined by Kabat *et al.* (1991) are underlined. The sequences for the constant domain were not determined but instead a consensus sequence based on homology with other known sequences was used. A check of the accuracy of this sequence was made by comparison with the electron-density map. Protein sequencing of the N-terminal part of the light chain was performed in order to confirm its κ identity and to ensure that the correct κ-derived cDNA corresponding to 1583 had indeed been isolated from the hybridoma cell line.

4.2. *Structure determination*

A Cα trace of the Fab structure is shown in Fig. 2. For the majority of the Fab structure the electron density is interpretable with most of the poorly defined regions occurring on solvent-exposed loops in the constant domain, particularly in the heavy chain. The hypervariable loops could all be visualized and modelled without ambiguity. A portion of the $2F_o - F_c$ map of the L2 region of 1583 is shown in Fig. 3. In the crystal packing arrangement the molecules lie in a head-to-tail fashion along the c axis. The hypervariable region packs into a space between the base of the constant region of one molecule and the hypervariable region of another (Fig. 4). The majority of the CDR residues do not form crystal contacts and are exposed to solvent. Only residues of L1 form any close contacts. In L1 residues the Oγ of SerL27A and the carbonyl O atom of LeuL27C pack against the side chains of AspH72 and LysH75, respectively. None of these crystal contacts are likely to affect the conformation of these hypervariable loops. The majority of atoms with high B factors are located in the constant domain, clustered at the base of the CH1 and CL domains. Most notably residues H124–H131 and L217–L219 have poor density and are presumably disordered. These areas are frequently disordered in reported Fab structures.

The elbow angle, which defines the angle between the variable and constant domains, is 125° and is at the lower range of that observed in Fab structures. This explains the failure of the molecular replacement using the search model with an elbow angle of 180°. Clearly the switch from 180 to 125° is beyond the radius of convergence of PC refinement. However, the fact that each of the modified models with elbow angles of 130 and 155° converged to the same solution illustrates that the radius of convergence in the Patterson correlation method is nevertheless very substantial.

4.3. *Canonical forms of the CDR's*

Four of the hypervariable loops, L2, L3 (type 1), H1 and H2 (type 1), can be classed according to the canonical structures of Chothia *et al.* (1989). The exceptions are L1 and H3. L1 most closely resembles a type 4 canonical structure, as represented by Fab 4-4-20 which was originally solved at 2.7 Å resolution (Herron, He, Mason, Voss & Edmundson, 1989). However, L1 of 1583 differs from that of 4-4-20 reported by Herron *et al.*, most notably in the position of residue 30 (Kabat numbering) which in 4-4-20 is an asparagine and points to the inside of the L1 loop whereas in 1583 this residue is a lysine and points outside. Our structure for L1 agrees more closely with that of the recent 1.85 Å determination of 4-4-20 (Whitlow, Howard, Wood, Voss & Hardman, 1995). The large size of L1 in 1583 means that it contributes significantly to the antigen-binding region. CDRH3 forms a (type IV) β-turn structure where the apex consists of an arginine (ArgH95) followed by two glycine residues. This lack of side chains together with its small length means that the space occupied by H3 in the antigen-binding region is comparatively small. The principal contribution of H3 to the antigen-binding site is through ArgH95.

4.4. *The antigen-binding site*

The antigen-binding surface of 1583 forms a pronounced groove of approximate dimensions 20 Å long, 10 Å wide and 15 Å deep parallel to the heavy- and light-chain interface. This is similar to other antibodies where the antigen is a single-chain molecule and contrasts with conformational epitopes which tend to bind to a flat surface on the Fab (Davies, Padlan & Sheriff, 1990). The cleft is formed largely as result of the combination of a relatively small contribution from CDRH3 and a particularly large CDRL1. The centre of the cleft is especially deep where it forms a hollow. It is likely that the peptide antigen will bind along the length of the groove in an analogous fashion to that observed in the 50.1 antibody–peptide complex (Rini *et al.*, 1993) which also has a short CDRH3 and a long CDRL1.

The antigen used to generate 1583 antibodies consisted of an 18-residue peptide genetically inserted into the VP3 loop of the Sabine type I poliovirus. This sequence is shown in Fig. 5. It is highly hydrophilic along its whole length and would be expected to bind in a charged environment. An analysis of escape mutants of S1/env/3 from neutralization by mAbs together with an analysis of the binding of overlapping peptides has

shown that the epitope can be subdivided into two minimal epitopes that correspond to residues 740–743 and residues 746–750 of gp160 (Vella *et al.*, 1993). 1583 binds to the second of these (ERDRD), a sequence which is highly charged. However, examination of the CDR regions of 1583 (Fig. 1) reveals that these are unexpectedly hydrophobic in character, particularly those of the heavy chain which contains only four CDR residues that are likely to be charged. The residues present in the antigen-binding site of 1583 are shown in Fig. 6. The deep, central region of the antigen-binding groove is indeed predominantly hydrophobic with two tryptophan residues (TrpL89 and TrpH52) and a histidine (HisH35) comprising the base of the cleft and a tyrosine (TyrL32), a leucine (LeuL50) and the aliphatic portion of an arginine side chain (ArgH95) forming the walls of the cleft. The large number of aromatic residues at the antigen-binding region is in common with many other antibodies (Davies & Cohen, 1996). The only charged residues present in this central region are ArgL96 and AspH101. This latter residue is especially buried and, in the model, is observed to form a hydrogen bond with the Nε of TrpH103. However, surrounding this are numerous charged residues, showing that although the CDR's appear to contain relatively few charged residues, those that are present are invariably clustered around the cleft. Furthermore, two of the charged residues, LysL45 and ArgL46, emanate from well outside the CDR's. These two residues are located within a patch of charged residues at the entrance of the groove which also comprises residues from and adjacent to CDRH3 and CDRL2, namely, AspH101, ArgH94, LysL53 and AspL55. Most of the remaining charged residues present are clustered in two patches high on either side of the groove. On the heavy-chain side are AspH56, AspH58 and ArgH95 and on the light-chain side are AspL27D, AspL28 and LysL30. Overall there are some 13 potentially charged residues which are suitably positioned to bind the antigen. Further clues as to how the peptide antigen may bind to 1583 are provided by the electrostatic surface plot (Fig. 7) which shows that the antigen-binding region is indeed polar and predominantly basic in character. The basicity is contributed in part by three arginines, ArgL96, ArgH95 and ArgL46, which are evenly spaced along the length of the groove. It is possible that the side chains of the three negatively charged residues of the peptide, Glu746, Asp748 and Asp750, point into the groove and interact with these arginines. It is likely that the two glycines present on H3 facilitate the close approach of the antigen. Either of the two arginines of the peptide might then interact with the aspartates present at the top of the groove. An asparagine residue, AsnL34, is located adjacent to ArgL46 and could form hydrogen bonds to the antigen. It is hoped that precise details of these interactions will revealed by a crystal complex of 1583 with its peptide antigen.

We wish to thank Drs Morag Ferguson and Philip Minor at the National Institute for Biological Standards and Control, South Mimms, Hertfordshire, England, for the samples of ascitic fluid containing 1583 and the 1583 hybridoma cell line and Dr Cherelyn Vella, also at NIBSC, for providing valuable data in advance of publication. We also wish to thank Dr Philip Marsh, Department of Genetics, Kings College, London, for the generous gift of κ-chain primers. This work was supported by the MRC as part of the AIDS directed programme. The X-ray data collection facilities used in this work were provided by the Bristol Molecular Recognition Centre which is supported by the BBSRC. The coordinates and structure factors have been deposited with the Protein Data Bank.*

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1NLD, R1NLDSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: AD0020).

References

Amit, A. G., Mariuzza, R., Phillips, S. E. V. & Poljak, R. J. (1986). *Science*, **233**, 747–758.

Broliden, P. A., von Gegerfelt, A., Clapham, P., Rosen, J., Fenyo, E. M., Wahren, B. & Broliden, K. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 461–465.

Brünger, A. T. (1990). *Acta Cryst.* A**46**, 46–57.

Brünger, A. T. (1992). *X-PLOR. Version* 3.1 *Manual. A System for X-ray Crystallography and NMR*. Yale University, New Haven, CT, USA.

Chanh, T. C., Dreesman, G. R., Kanda, P., Linette, G. P., Sparrow, J. T., Ho, D. D. & Kennedy, R. C. (1986). *EMBO J.* **5**, 3065–3071.

Chothia, C., Lesk, A. M., Tramantano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D. R., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M. & Poljak, R. J. (1989). *Nature (London)*, **342**, 877–833.

Dalgleish, A. G., Chanh, T. C., Kennedy, R. C., Kanda, P., Chapman, P. R. & Weiss, R. A. (1988). *Virology*, **165**, 209–215

Davies. D. R. & Cohen, G. H. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 7–12.

Davies, D. R, Padlan, E. A. & Sheriff, S. (1990). *Annu. Rev. Biochem.* **59**, 439–473.

Evans, D. J., McKeating, J., Meredith, J. M., Burke, K. L., Katrak, K., John, A., Ferguson, M., Minor, P. D., Weiss, R. A. & Almond, J. W. (1989). *Nature (London)*, **339**, 385–388.

Fischmann, T. O., Bentley, G. A., Bhat, T. N., Boulot, G., Mariuzza, R. A., Phillips, S. E. V., Tello, D. & Poljak, R. J. (1991). *J. Biol. Chem.* **266**, 12915–12920.

Ghiara, J. B., Stura, E. A., Stanfield, R. L., Profy, A. T. & Wilson, I. A. (1994). *Science*, **264**, 82–85.

Herron, J. N., He, X.-M., Mason, M. L., Voss, E. W. & Edmundson, A. B. (1989). *Proteins*, **5**, 271–280.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.

Jones, A. (1985). *Methods Enzymol.* **115**, 157–171.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*. Bethesda: National Institutes of Health.

Kennedy, R. C., Henkel, R. D., Pauletti, D., Allan, J. S., Lee, T. H., Essex, M. & Dreesman, G. R. (1986). *Science*, **231**, 1556–1559.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wang-Staal, F. & Wolf, H. (1987). *J. Virol.* **61**, 570–578.

Nicholls, A., Sharp, K. A. & Konig, B. (1991). *Proteins*, **11**, 281–296.

Orlandi, R. Gussow, D. H., Jones, P. T. & Winter, G. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 3833–3837.

Porter, R. R. (1959). *Biochem. J.* **73**, 119–126.

Rini, J. M., Schulze-Gahmen, U. & Wilson, I. A. (1992). *Science*, **255**, 959–965.

Rini, J. M., Stanfield, R. L., Stura, E. A., Salinas, P. A., Profy, A. T. & Wilson, I. A. (1993). *Proc. Natl. Acad. Sci.* USA. **90**, 6325–6329.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). In *Molecular Cloning – A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Tronrud, T. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* **A43**, 489–500.

Vella, C., Ferguson, M., Dunn, G., Meloen, R., Langedijk, H., Evans, D. & Minor, P. D. (1993). *J. Gen. Virol.* **74**, 2603–2607.

Whitlow, M., Howard, A. J., Wood, J. F., Voss, E. W. & Hardman, K. D. (1995). *Protein Eng.* **8**, 749–761.