Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a single-chain Fv fragment of an antibody that inhibits the HIV-1 and HIV-2 proteases

Julien Lescar,[a]† Jiri Brynda,[b] Milan Fabry,[b] Magda Horejsi,[b] Pavlina Rezacova,[b] Juraj Sedlacek[b] and Graham A. Bentley[c]*

[a]European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble, France, [b]Department of Gene Manipulation, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, 166 37 Prague 6, Czech Republic, and [c]Unité d'Immunologie Structurale (URA 2185 CNRS), Département de Biologie Structurale et Chimie, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris CEDEX 15, France

† Present address: School of Biological Sciences, Nanyang Technological University, 1 Nanyang Walk Blk 5 Level 3, Singapore 637616.

Correspondence e-mail: bentley@pasteur.fr

Received 10 January 2003
Accepted 12 February 2003

PDB Reference: Fv fragment of HIV protease inhibitor, 1n4x, r1n4xsf.

The monoclonal antibody 1696, which was raised against the HIV-1 protease, inhibits the catalytic activity of the enzyme from both the HIV-1 and HIV-2 strains. The antibody cross-reacts with peptides containing the N-terminus of the enzyme, which is highly conserved between these strains. The crystal structure of a single-chain Fv fragment of 1696 (scFv-1696) in the non-complexed form, solved at 1.7 Å resolution, is compared with the previously reported non-complexed Fab-1696 and antigen-bound scFv-1696 structures. Large conformational changes in the third hypervariable region of the heavy chain and differences in relative orientation of the variable domains are observed between the different structures.

1. Abbreviations and symbols

CDR, complementarity-determining region; L1, L2, L3, first, second and third CDR, respectively, of the light chain; H1, H2, H3, first, second and third CDR, respectively, of the heavy chain; HIV, human immunodeficiency virus; mAb, monoclonal antibody; r.m.s.d., root-mean-square deviation; scFv, single-chain Fv fragment; V_H, heavy-chain variable domain; V_L, light-chain variable domain.

Antibody residues are numbered using the Kabat convention, preceded by L or H for light and heavy chain, respectively.

2. Introduction

Proteins from HIV are expressed in the form of polyprotein precursors that are cleaved by a protease encoded in the viral genome. Because it plays an essential role in the infectious cycle of the virus, HIV protease is one of the primary targets for antiviral drug design (Wlodawer & Gutschina, 2000). We have produced monoclonal antibodies raised against HIV-1 protease in order to study the enzyme's activity and function and to aid in the design of novel inhibitors targeted to regions other than the active site. Two anti-HIV-1 protease mAbs, F11.2.32 and 1696, which inhibit the enzymatic activity of the protease have been characterized (Lescar *et al.*, 1996, 1997, 1999). We have recently reported the crystal structure at 2.7 Å resolution of a cross-reaction complex formed between scFv-1696 and a peptide, PQITLWQRR, whose sequence is derived from the N-terminus of the HIV-1 protease (Rezacova *et al.*, 2001). This structure supports the hypothesis that mAb-1696 inhibits the HIV protease by favouring the dissociation of the active homodimer. The crystal structure of the ligand-free scFv-1696 molecule that we describe here reveals a plasticity in the topology of the antigen-binding site of this mAb. Crystals of non-complexed scFv-1696 were obtained under acidic conditions (pH 3.5). Here, we compare this structure, determined at 1.7 Å resolution, with that of non-complexed Fab-1696 crystallized at pH 7.5 (Lescar *et al.*, 1999) and the antigen-bound scFv-1696 crystallized at pH 5.5 (Rezacova *et al.*, 2001).

3. Experimental

3.1. Crystallization and data collection

Needle-shaped crystals of approximate dimensions 0.15 × 0.01 × 0.01 mm were obtained by the hanging-drop technique using the following conditions: 5 µl of precipitant (0.1 *M* sodium citrate pH 3.5, 17.5% PEG 4000, 0.175 *M* ammonium sulfate) was mixed with an equal volume of a solution of scFv-1696 at a concentration of 5 mg ml^{-1} in 10 m*M* Tris–HCl pH 7.3. The drop was suspended over a reservoir containing 1 ml of the precipitant solution at 291 K and allowed to equilibrate. For data collection, crystals were briefly transferred to the reservoir buffer containing 25%(*v*/*v*) glycerol before being cryocooled to 100 K in a nitrogen-gas stream. Measurements were made with a MAR Research 345 imaging-plate system using an X-ray wavelength of 0.990 Å on the ID2 beamline at the ESRF (Grenoble). Owing to the small size of the crystals, it was not possible to obtain a complete data set from one crystal alone without incurring severe radiation damage. Three crystals were therefore used to obtain an acceptably complete

Table 1
Data set and refinement statistics.

Values in parentheses are for the highest resolution shell (1.79–1.70 Å).

Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	126.93
b (Å)	61.21
c (Å)	37.30
Resolution range (Å)	20.0–1.70
Unique reflections	45422 (5930)
Data redundancy	4.89 (3.34)
Completeness (%)	91.0 (84.2)
$I/\sigma(I)$	34.3 (6.6)
$R_{\rm merge}$ (%)	9.0 (36.0)
$V_{\rm M}$ (Å^3 Da^{-1})	2.07
$R_{\rm work}$ (%)	22.7
$R_{\rm free}$ (%)	25.6
No. of protein atoms	3656
No. of heteroatoms	2
No. of solvent molecules	530
R.m.s.d. bond lengths (Å)	0.005
R.m.s.d. bond angles (°)	1.34
Ramachandran plot	
Core region (%)	87.8
Additionally allowed regions (%)	11.2
Generously allowed regions (%)	0.5
Disallowed region (%)	0.5
Average *B* factors, all protein atoms (Å^2)	28.1
R.m.s.d. *B* factors, all protein atoms (Å^2)	1.7

data set. Integration, scaling and merging of the intensities were carried out using the program *XDS* (Kabsch, 1988), with data frames from the three crystals being scaled and merged simultaneously. Details of crystal parameters and data-collection statistics are given in Table 1.

3.2. Solution and refinement of the structure

A preliminary model of the structure was found by molecular replacement using the program *AMoRe* (Navaza, 1994). The rotation-function calculation was performed between resolution limits of 20.0 and 4.0 Å using the variable dimer (V_H/V_L) of Fab-1696 (PDB code 1lc7) as the search model and a Patterson integration radius of 20 Å. Two strong peaks in the rotation function (correlation coefficients of 0.173 and 0.171, compared with 0.138 for the third highest peak) clearly indicated the presence of the two anticipated molecules in the asymmetric unit. The two scFv molecules were then placed in the unit cell using one- and two-body translation searches. Rigid-body refinement of the two molecules yielded a correlation coefficient of 0.443 and an *R* factor of 0.450 for 3592 reflections in the resolution range 9.0–4.0 Å resolution.

The structure was refined using the program *CNS* (Brünger *et al.*, 1998). After omitting the hypervariable loops from the molecular-replacement model, the atomic parameters were refined using all measured data between resolution limits of 20 and 1.7 Å, with the exception of 2280 randomly chosen reflections (5% of the data) used to monitor the refinement progress by the free *R* factor ($R_{\rm free}$). As refinement proceeded, the hypervariable loops were progressively built into the electron-density maps as poly-alanine chains and identifiable side chains were modelled accordingly. Tight non-crystallographic symmetry restraints were applied during the initial phases of the refinement, but were relaxed during the final stages. Once the model for the scFv-1696 molecules had been completed, water and solute molecules were fitted to difference electron-density maps. Structure-refinement statistics are summarized in Table 1.

4. Results

4.1. Description of the structure

The final model includes 113 residues for the light chain and 120 residues for the heavy chain for both scFv-1696 molecules. All residues in the model fall into favourable or allowed regions of the Ramachandran plot except ValL51, which is contained in the γ-turn defining the unique canonical structure of CDR-L2. The flexible $(Gly_4Ser)_3$ linker connecting the C-terminus of the heavy chain to the N-terminus of the light chain was not visible for either molecule. Residues belonging to the hypervariable loops are generally well defined in the electron density. Several residues of CDR-H3 of molecule 1, however, have high temperature factors and their positions are less well defined than those of molecule 2, although the path of the main chain is unambiguous. A total of 530 well defined water molecules and two chloride ions were placed.

A strong positive peak in the difference electron density was located at the apex of the CDR-H3 loop of the antibody (Fig. 1). First interpreted as a water molecule, this density was subsequently modelled as a chloride ion (from the Tris–HCl buffer of the protein solution), coordinated by four nitrogen amide hydrogen-bond donors from CDR-H3 ($N-Cl^-$ distances < 3.5 Å) and one solvent molecule. The presence of a chloride ion could thus strongly influence the conformation of CDR-H3 observed in this crystal form. In addition, one intermolecular interaction between HisH97 $N^{\varepsilon 2}$ in the CDR-H3 of scFv-1696 molecule 2 and the carboxylic group of GluL79 of scFv-1696 molecule 1 could also stabilize the conformation of CDR-H3 of scFv-1696 molecule 2.

Table 2
Pairwise comparisons of the structures of the variable domains of non-complexed scFv-1696 (this work, two molecules), ligand-free Fab-1696 (PDB code 1cl7) and scFv-1696 in complex with the HIV-1 peptide (PDB code 1jp5).

For each structural comparison, the r.m.s. differences between the main-chain atom positions of equivalent residues of the variable domains together ($V_L + V_H$), the V_L domain only and the V_H domain only, respectively, are given in Å. Superpositions were carried out using program *LSQKAB* from the *CCP*4 program suite (Kabsch, 1988; Collaborative Computational Project, Number 4, 1994).

	scFv-1696-2	Fab-1696	scFv-1696 complex
scFv-1696-1	0.61, 0.40, 0.42	1.00, 0.74, 0.93	1.24, 0.55, 1.47
scFv-1696-2		1.22, 0.74, 0.86	1.39, 0.57, 1.44
Fab-1696			1.05, 0.56, 1.21

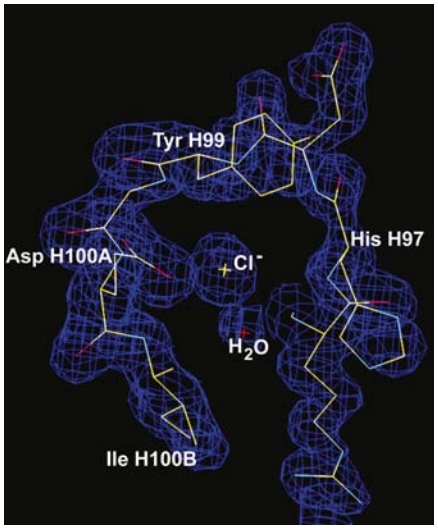

Figure 1
View of CDR-H3, with electron-density map, calculated with $(3F_o - 2F_c)$ coefficients and phases from the refined model, contoured at the 1.2 r.m.s. level. A prominent solvent peak, at hydrogen-bonding distances from four nitrogen amides of CDR-H3 and one water molecule, has been modelled as a chloride ion. Those residues of CDR-H3 for which the amide nitrogen is less than 3.5 Å from the putative chloride ion, together with the coordinating water molecule, are indicated.

4.2. Comparison of the ligand-free scFv-1696 at acid pH with ligand-free Fab-1696 and complexed scFv-1696 at neutral pH

After superimposing the V_L domains of the different forms, additional rotations and translations are required in order to optimize the superposition of V_H domains (Table 2). These differences in relative orientation of the variable domains give rise to antigen-binding sites with quite different topologies: the ligand-free scFv-1696 binding site is more open than that of the

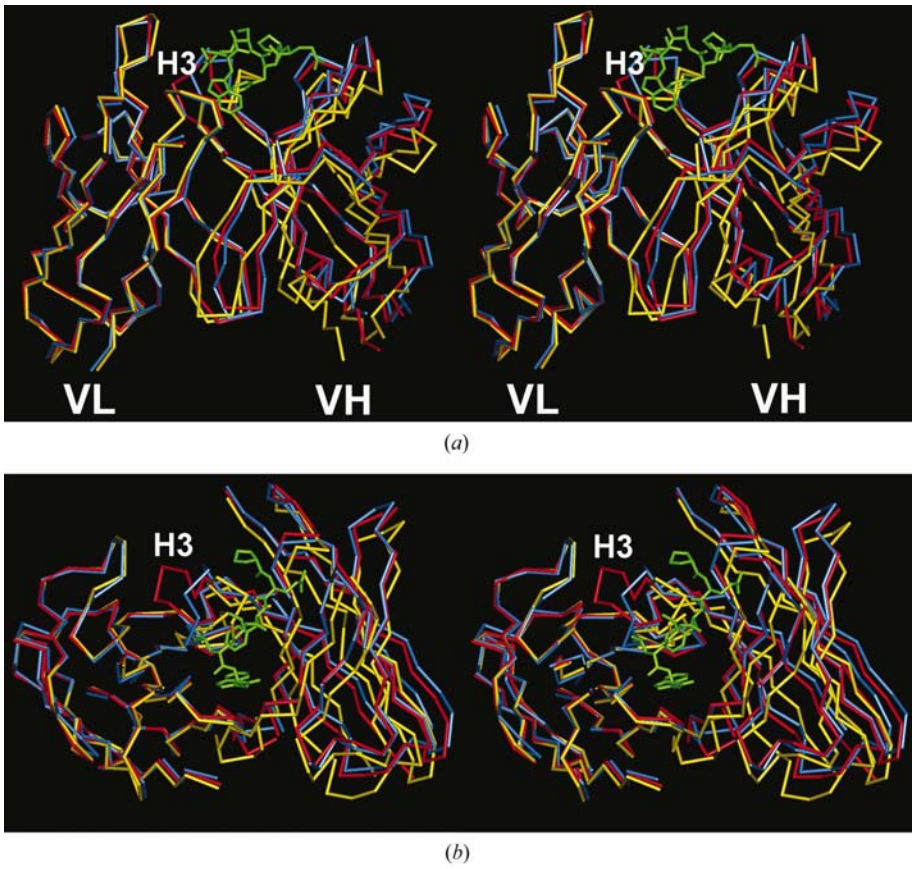

Figure 2
Comparison of the variable domains of non-complexed Fab-1696 (PDB code 1cl7, blue), molecule 1 of the non-complexed scFv-1696 (this work, yellow) and complexed scFv-1696 (PDB code 1jp5, red). The peptide is shown in green. Superposition has been made by optimizing the correspondence between the main-chain atoms of the V_L domains of the three structures. Fv1696 is shown from the side of (*a*) and from above (*b*) the antigen-bindng site. The V_L domain is shown on the left of the Fv fragment, while the V_H domain is on the right. The third CDR of V_H (H3) is indicated.

ligand-free Fab-1696, while the complexed scFv-1696 structure adopts a position intermediate between the two (Fig. 2). It is notable that the closed conformation adopted by the ligand-free Fab-1696 binding site (but not that of the ligand-free scFv1696 molecule) would introduce some steric hindrance between its CDR-H1 and the N-terminal end of the antigenic peptide, as seen in the complexed structure with scFv-1696.

4.3. Structure and flexibility of CDR-H3

The most conspicuous changes in the tertiary structure of mAb-1696 occur in CDR-H3. By contrast, the backbone conformations of the five other CDRs are quite conserved. The residues situated between ArgH96 and GluH100A, at the solvent-exposed end of CDR-H3, adopt strikingly different positions. In the peptide complex of scFv-1696, CDR-H3 moves towards the V_L domain, giving rise to a shallow groove at the centre of the antigen-binding site where the peptide is buried (Rezacova *et al.*, 2001). In the ligand-free scFv-1696 crystal structure obtained at pH 3.5, however, CDR-H3 folds back towards the V_H domain (Fig. 2). This position occludes the antigen-binding pocket by bringing residues AspH98 and TyrH99 into the space taken up by GlnP2 and IleP3 of the antigenic peptide. In the ligand-free Fab-1696 structure crystallized at pH 7.5, the observed conformation of CDR-H3 is maintained in part by an ion pair between AspH95 and ArgH96 (Lescar *et al.*, 1999). This salt bridge is lost in the ligand-free scFv-1696 structure of CDRH3, possibly because of protonation of the carboxylic group of AspH95 at pH 3.5. The acidic pH could also account for the disruption of the hydrogen-bond interaction between AspH95 and HisH35, which associates CDR-H3 with CDR-H1 in the Fab-1696 structure at neutral pH.

5. Conclusion

The structural changes observed in the binding site of mAb-1696 result in the formation of antigen-binding pockets with quite different shapes. The conformational changes observed between the three different crystal forms result from rigid-body movements of variable domains accompanied by large conformational changes in CDR-H3. Three different conformations for CDR-H3 of mAb-1696 have been identified: the antigen-bound conformation of CDR-H3 appears to be induced by the presence of the antigenic peptide, whereas the two alternative structures of CDR-H3 that we have observed in the non-complexed forms of 1696 point to the possible roles of pH and solute molecules in influencing the conformation.

This work has been financed by the ESRF, the Pasteur Institute and the Centre National pour la Recherche Scientifique. Assistance and useful discussions with members of the ESRF/EMBL Joint Structural Biology Group are gratefully acknowledged.

References

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. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Lescar, J., Brynda, J., Rezacova, P., Stouracova, R., Riottot, M. M., Chitarra, V., Fabry, M., Horejsi, M., Sedlacek, J. & Bentley, G. A. (1999). *Protein Sci.* **8**, 2686–2696.
Lescar, J., Stouracova, R., Riottot, M. M., Chitarra, V., Brynda, J., Fabry, M., Horejsi, M., Sedlacek, J. & Bentley, G. A. (1996). *Protein Sci.* **5**, 966–968.
Lescar, J., Stouracova, R., Riottot, M. M., Chitarra, V., Brynda, J., Fabry, M., Horejsi, M., Sedlacek, J. & Bentley, G. A. (1997). *J. Mol. Biol.* **267**, 1207–1222.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Rezacova, P., Lescar, J., Brynda, J., Fabry, M., Horejsi, M., Sedlacek, J. & Bentley, G. A. (2001). *Structure*, **9**, 887–895.
Wlodawer, A. & Gutschina, A. (2000). *Biochim. Biophys. Acta*, **1477**, 16–34.