Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the recombinant antibody Fab fragment f3p4

Daniel Frey, Thomas Huber, Andreas Plückthun and Markus G. Grütter*

Department of Biochemistry, University of Zürich, Switzerland

Correspondence e-mail: gruetter@bioc.unizh.ch

Received 14 January 2008
Accepted 17 March 2008

PDB Reference: Fab fragment f3p4, 2v7n, r2v7nsf.

The structure of the antibody Fab fragment f3p4, which was selected from a subset of the synthetic HuCAL antibody library to bind the sodium citrate symporter CitS, is described at 1.92 Å resolution. Comparison with computational models revealed deviations in a few framework positions and in the binding loops. The crystals belong to space group $P2_12_12$ and contain four molecules in the asymmetric unit, with unit-cell parameters a = 102.77, b = 185.92, c = 102.97 Å. These particular unit-cell parameters allowed pseudo-merohedral twinning; interestingly, the twinning law relates a twofold screw axis to a twofold axis.

1. Introduction

The determination of membrane-protein structures is still very challenging, as indicated by the limited number of structures that have been deposited in the PDB. One bottleneck is the generation of crystals that are suitable for high-resolution structure determination. This is a consequence of the inherent conformational inhomogeneity and flexibility of the proteins, including the loops connecting the transmembrane segments of the detergent-solubilized membrane proteins, which is not beneficial for crystallization. One approach to overcome these limitations is cocrystallization with antibody fragments (Hunte & Michel, 2002; Ostermeier & Michel, 1997) or other proteins such as designed ankyrin-repeat proteins (Binz *et al.*, 2004; Kohl *et al.*, 2005; Schweizer *et al.*, 2007; Sennhauser *et al.*, 2006). The binding protein provides additional polar surface area available for protein–protein interactions, thus increasing the chance of obtaining well ordered crystals. Furthermore, the binding protein may stabilize a certain conformation and thereby reduce the conformational inhomogeneity of the sample.

In all structures published to date, antibody fragments (scFv or Fab) have been obtained by immunization of animals. The main disadvantage in the generation of monoclonal antibodies by classical hybridoma technology is the inability to control the structural integrity of the membrane protein during the immunization process. The use of phage or ribosome display as *in vitro* selection methods is more advantageous for finding antibodies that recognize solubilized membrane proteins (Huber *et al.*, 2007; Röthlisberger *et al.*, 2004). Both techniques are faster compared with eliciting an immune response *in vivo*, but more importantly, defined buffer and detergent conditions which are favourable to the protein of interest are used during the selection. In this way, the structural integrity of the mem-

brane protein can be maintained and the likelihood of selecting binders to conformational epitopes of the native structure of the membrane protein is increased.

Fab fragments were selected from a subset of the human combinatorial antibody library (HuCAL GOLD; Morphosys, Germany), consisting of only the most stable and well expressed frameworks V_H3 and $V_\kappa3$ (Ewert *et al.*, 2003), by phage display to the detergent-solubilized sodium citrate symporter CitS of *Klebsiella pneumoniae* (Röthlisberger *et al.*, 2004). CitS is a multitopic membrane protein of the helix-bundle type with relatively short loops connecting the transmembrane segments. This well studied member of the 2-hydroxycarboxylate transporter (2HCT) family facilitates the symport of citrate and sodium ions, allowing the bacteria to use citrate as the sole carbon source. Detergent-solubilized CitS can be functionally reconstituted (Pos & Dimroth, 1996) and, therefore, it is likely that CitS preserves its natural conformation in the detergent. This allowed us to obtain Fab fragments by phage display that recognize a conformational epitope of the native state of CitS (Röthlisberger *et al.*, 2004).

The Fab fragment f3p4 was selected by phage display. It binds specifically to the CitS dimer with an affinity in the nanomolar range (Huber *et al.*, 2007; Röthlisberger *et al.*, 2004). The cocrystallization of CitS with the Fab fragment f3p4 may not only help to generate crystals that are suitable for high-resolution structure determination, but may also help in phasing the complex by molecular replacement. Therefore, the Fab fragment f3p4 was crystallized alone in order to obtain a high-quality molecular-replacement search model.

In general, crystals form a continuous assembly of unit cells that are well ordered throughout the crystal, although this perfect arrangement is often disturbed by minor crystal-growth defects and variations of the protein conformation, leading to higher crystal mosaicity. Under special circumstances, however, the crystal is made up of two or more crystalline domains which are related to each other by additional symmetry elements, a phenomenon that is called twinning.

In the case of merohedral twinning, two or more crystal domains are related by an additional symmetry element that belongs to the lattice but not to the crystal space group, creating a perfect superposition of the observed diffraction patterns. Since the individual lattices of each domain are related by additional symmetry, this symmetry element is observed in the combined diffraction pattern.

This is particularly true when the crystal has a nearly even composition of the different twin domains. Pseudomerohedral twinning is found in space groups where the unit-cell parameters suggest higher symmetry and a different Laue class to that actually present in the crystal. For example, an angle $\beta = 90^\circ$ in the monoclinic space group $P2$ suggests the orthorhombic space group $P222$; the Laue class of such a crystal is only $2/m$, whereas the lattice has mmm symmetry.

The general problem of twinning is that the intensities recorded do not arise from a single crystal but are the summation of intensities from individual domains. In the most common case, in which two domains are related by a 180° twofold-axis rotation, the measured intensities of such twofold-related reflections are therefore $J_{1,\rm obs} = I_1(1 - \alpha) + \alpha I_2$ and $J_{2,\rm obs} = \alpha I_1 + (1 - \alpha)I_2$, where I_1 and I_2 are the untwinned intensities of individual domains, $J_{1,\rm obs}$ and $J_{2,\rm obs}$ are the observed intensities and the twinning fraction α is the fractional volume of one domain in the crystal, with the second fractional volume corresponding to $1 - \alpha$. For $\alpha < 0.5$ it is possible to obtain the correct intensities of each reflection and detwin the data set, but since the error [which is proportional to $1/(1 - 2\alpha)$] increases rapidly when α approaches 0.5, detwinning is not possible under these conditions (Yeates & Fam, 1999; Yeates, 1997). Nevertheless, data from twinned crystals with α close to 0.5 and even data from perfectly twinned crystals with $\alpha = 0.5$ can be used to solve a structure by molecular replacement (Yeates & Fam, 1999).

2. Material and methods

2.1. Protein expression and purification

The Fab fragment f3p4 was cloned into the expression vector pMx9 (Rauchenberger *et al.*, 2003). The final construct carried a His_6 tag fused to the C-terminus of the heavy chain. Expression was performed in *Escherichia coli* SB356 (Bass *et al.*, 1996) in 2YT (10 g yeast extract, 16 g tryptone and 5 g NaCl per litre). 30 ml 2YT medium containing 30 µg ml^{-1} chloramphenicol and 1.0% glucose was inoculated with a single colony and incubated overnight at 298 K. 1 l 2YT medium containing 15 µg ml^{-1} chloramphenicol and 50 m*M* K_2HPO_4 was inoculated with the overnight culture to a starting OD_{600} of 0.1 and incubated in a 5 l shaking flask with baffles at 298 K. Expression was induced at an OD_{600} of 0.6 with 1 m*M* IPTG (final concentration) and continued for 18 h.

The harvested cells were resuspended in lysis buffer (50 m*M* Na_2HPO_4 pH 8.0, 150 m*M* NaCl) containing DNaseI and 1 mg ml^{-1} lysozyme and incubated on ice for 30 min. The cells were lysed by passage through a French press and the lysate was cleared by centrifugation (20 min at 40 000*g*) and filtration through a 0.22 µm membrane.

The cleared lysate was applied onto a BioCAD system with an IMAC column (Poros 20MC, 7 ml, Applied Biosystems) and an anion-exchange column (Poros 20HQ, 7 ml, Applied Biosystems) coupled inline. The protein was loaded onto the IMAC column and washed with 10 m*M* Tris pH 8.0, 150 m*M* NaCl, high salt (10 m*M* Tris pH 8.0, 1 *M* NaCl), low salt (10 m*M* Tris pH 8.0, 15 m*M* NaCl) and finally with a low concentration of imidazole (10 m*M* Tris pH 8.0, 15 m*M* NaCl, 5 m*M* imidazole). The protein was eluted with 10 m*M* Tris pH 8.0 and 200 m*M* imidazole and directly applied onto the anion-exchange column; the contaminants were bound to the column, while the Fab fragment f3p4 was recovered in the flowthrough. The eluate was further passed over a size-exclusion chromatography column (Superose 6 10/300GL, GE Healthcare) in 10 m*M* Tris pH 7.5, 150 m*M* NaCl and 0.05%(*w*/*v*) *n*-dodecyl-β-D-maltopyranoside to yield a pure and homogeneous protein sample. The protein was concen-

Table 1
Scaling statistics.

Space group	R_{merge}† (%)	R_{meas}‡ (%)	$R_{mrgd\text{-}F}$‡ (%)	Completeness (%)	Molecules in the ASU
*P*1	4.5	6.4	9.0	78.2	16
*P*2	5.7	6.9	7.9	94.3	8
*C*2	8.9	11.0	10.3	96.5	8
*P*222	6.5	7.2	6.0	97.4	4
*C*222	9.8	10.9	7.8	98.7	4
*P*4	9.6	10.7	7.6	99.3	4
*P*422	11.1	11.7	5.7	99.2	2

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ For the definitions of R_{meas} and $R_{mrgd\text{-}F}$, see Diederichs & Karplus (1997).

trated to 12 mg ml^{-1} in a centrifugal filter unit (30 kDa molecular-weight cutoff, Amicon Ultra, Millipore) prior to crystallization.

2.2. Crystallization

Crystals of the Fab fragment f3p4 were grown by mixing 1 µl protein solution with 1 µl 23% PEG 4000, 50 m*M* lithium sulfate, 50 m*M* sodium sulfate, 50 m*M* *N*-(2-acetamido)-iminodiacetic acid in the pH range 6.3–6.5. The sitting drops were equilibrated by vapour diffusion at 277 K. Crystals appeared overnight and grew to full size within one week. Needle-like crystals of dimensions 300 × 50 × 50 µm were flash-cooled in liquid propane and 25%(*v*/*v*) ethylene glycol as cryoprotectant.

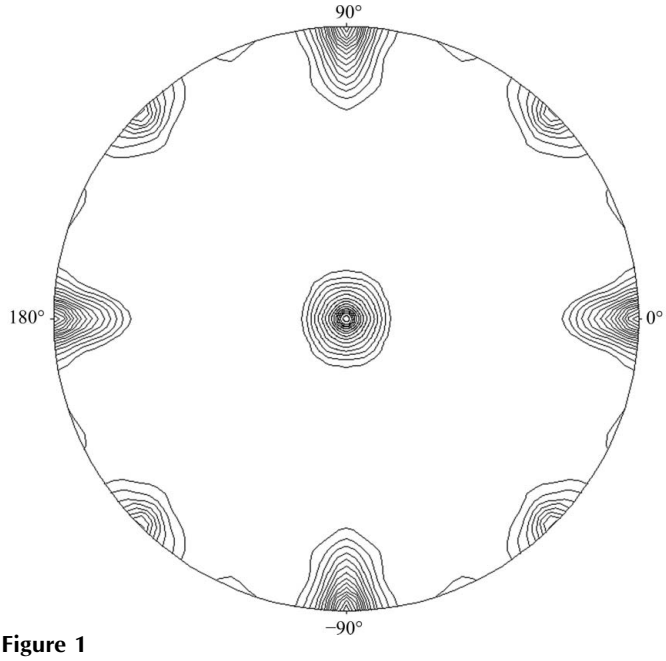


Figure 1
Self-rotation function in space group *P*222. Plot of the $\kappa = 180^\circ$ section. The self-rotation function was calculated in space group *P*222 using the program *POLARRFN* (Collaborative Computational Project, Number 4, 1994). The plot was generated with an integration radius of 25 Å and a contour level >30% of the height of the origin peak using data between 30 and 2 Å resolution. The data reveal *P*222 symmetry with an additional noncrystallographic twofold axis at $\omega = 90^\circ$ and $\varphi = 45^\circ$.

Table 2
Molecular-replacement solutions.

Space group	LLG†/ clashes	Untwinned R/R_{free}‡ (%)	Twin law	Twinned R/R_{free}§ (%)
$P222_1$	−468/5	41/43	$-k, h, l$	27/31
$P2_122_1$	+1465/0	34/38	$-k, h, l$	23/26
$P22_12_1$	−263/0	38/41	$-k, h, l$	23/27
$P2_12_12_1$	−1207/14	41/43	$-k, h, l$	29/34
$P4_1$	−789/0	40/43	$k, h, -l$	28/32
$P4_12_12$	−1200/4	44/48	—	—

† Log-likelihood gain after positional refinement of all molecules. ‡ *R* and R_{free} values reported after refinement with *REFMAC* without twinning. § *R* and R_{free} values with twinning applied during refinement with *SHELX*97.

2.3. Data collection and processing

A data set was collected to 1.9 Å resolution from a single crystal on the X06SA beamline of the Swiss Light Source (SLS) at the Paul Scherrer Institute, Villigen, Switzerland. Diffraction images were indexed and integrated using the *XDS* program package (Kabsch, 1993). Indexing of a total of 3000 spots with $I/\sigma(I) > 5$ revealed unit-cell parameters $a = 102.77$, $b = 102.97$, $c = 185.92$ Å, $\alpha = \beta = \gamma = 90^\circ$. In addition to this primitive cell, centred orthorhombic as well as primitive tetragonal were good candidates for the Bravais lattice of the crystal. Integration and subsequent scaling values of possible lattices are summarized in Table 1.

Systematic absences were analyzed and indicated the presence of a 2_1 or 4_2 screw axis along the long axis *c*. In combination with the self-rotation function (Fig. 1), the space group was assumed to be the orthorhombic space group $P222_1$. An additional peak at $\omega = 90$, $\varphi = 45$, $\kappa = 180^\circ$ with a peak height of 50% of the origin-peak height indicated an imperfect noncrystallographic twofold axis.

2.4. Structure determination and refinement

The structure was solved by molecular replacement using *Phaser* (Collaborative Computational Project, Number 4, 1994; Read, 2001) with a Fab fragment (PDB code 2fb4; Kratzin *et al.*, 1989) as a search model. To accommodate the unknown elbow angle between the constant and variable domains of the Fab fragment, the search model was divided into two parts: the constant and variable domains. One correctly assembled Fab fragment was extracted from the first molecular-replacement attempt and used as a starting model to find the rotation and translation solutions of all four molecules in the asymmetric unit of space group $P222_1$.

Crystal twinning was not initially considered and the structure was refined with *CNS* (Brünger *et al.*, 1998) using the twinned data without any noncrystallographic symmetry (NCS) restraints. After several rounds of model building and refinement, the final values of $R_{cryst} = 45.3\%$ and $R_{free} = 46.7\%$ were unacceptably high. Considering the presence of twinning in the data, the molecular-replacement trials were repeated in the orthorhombic and tetragonal space groups with all possible permutations (Table 2).

Taking twinning into account, the progress of refinement was straightforward and the best results were obtained for

space group $P2_122_1$ (Table 2). Inspection of the (00l) reflections confirmed the presence of the 2_1 screw axis along the long axis c, whereas systematic absences for the second screw axis along the short axes a or b were not observed. Why is the second 2_1 screw axis not observed in the diffraction pattern? When the two axes a and b are related by twinning, the measured intensities of the reflections with indices (h00) are actually a summation of (h00) of the first domain and (0k0) of the second domain of the crystal. This means that the systematic absences for reflections (h00) with the 2_1 screw axis along a are thus masked by the (0k0) reflections, which have no systematic absences owing to the twofold axis along b.

The space group was changed to the conventional setting $P2_12_12$ and all further refinement steps were performed with *SHELX*97 (Sheldrick, 2008) using the twinned intensities obtained from *XDS* (Kabsch, 1993) and *HKLF*4. The TWIN instruction with a twofold twinning around the long axis (h, k, l $\rightarrow$ $-l$, k, h), a starting twin fraction α = 0.37 and local NCS restraints (Usón *et al.*, 1999) were included during the refinement. The test set of reflections for $R_{\rm free}$ calculations was generated in small shells to avoid correlation between twinned reflections belonging to the test and the reference sets (Sheldrick, 2008). The refinement statistics are summarized in Table 3. A similar case of pseudo-merohedral twinning in space group $P2_12_12$ was reported recently in crystals of the enzyme Dicer (MacRae & Doudna, 2007).

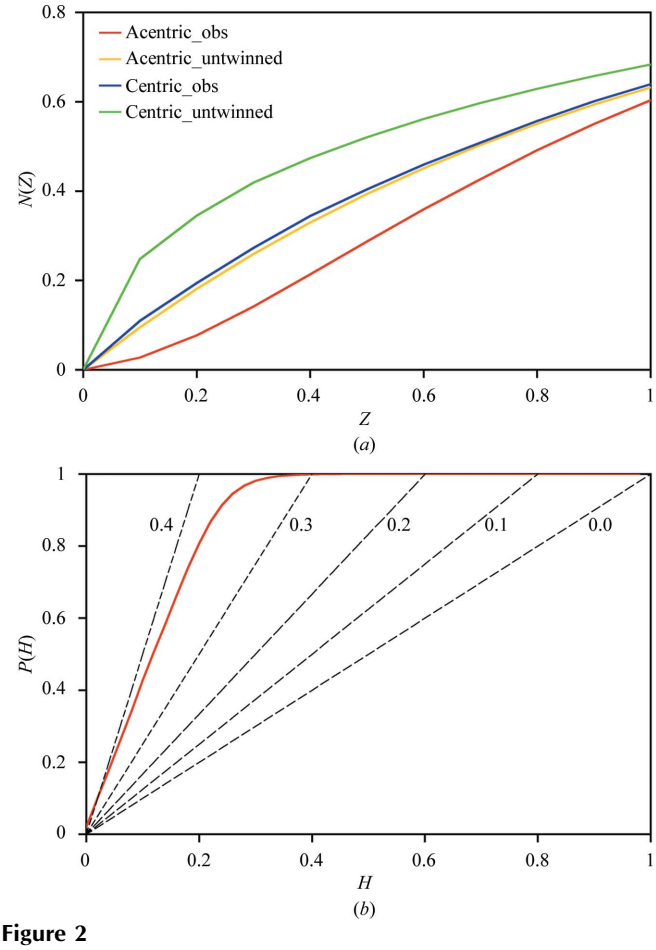


Figure 2
Detection of twinning based on data-set statistics and determination of the twin fraction α. (*a*) Cumulative intensity distribution as implemented in *TRUNCATE*. The fraction of average intensity Z is plotted against the fraction of reflections $N(Z)$ with intensities below the average. Both centric and acentric reflections (blue and red) deviate significantly from the theoretical values shown in green and yellow. (*b*) Determination of the twinning fraction α using the H-test (Yeates, 1997; Yeates & Fam, 1999). $H = |I_1 - I_2|/(I_1 + I_2)$ is plotted against the cumulative distribution $P(H)$, where I_1 and I_2 are the intensities of two reflections related by the twin law. The initial slope of the observed data (red) results in an α of 37%. Theoretical lines for various twinning fractions are shown as dashed lines.

Table 3
Crystallographic data collection and structure refinement.

Values in parentheses are for the highest shell.

Data collection	
Space group	$P2_12_12$
Unit-cell parameters (Å, °)	a = 102.79, b = 185.92, c = 102.99, $\alpha = \beta = \gamma$ = 90
Wavelength (Å)	1.0000
Resolution range (Å)	30–1.90 (1.92–1.90)
Unique reflections	151551 (3952)
Completeness (%)	97.4 (83.4)
$R_{\rm merge}$ (%)	7.7 (33.1)
Multiplicity	4.9 (3.3)
$I/\sigma(I)$	16.9 (4.2)
Refinement	
Resolution range	15–1.92
No. of reflections/test set	143072/4240
$R_{\rm cryst}/R_{\rm free}$	0.155/0.201
$R_{\rm cryst}/R_{\rm free}$ [$F_{\rm o} > 4\sigma(F_{\rm o})$]	0.150/0.196
No. of protein atoms	12877
No. of solvent atoms	540
Twin operator	h, k, $l \rightarrow -l$, k, h
R.m.s. deviation from ideality	
Bonds (Å)	0.006
Angles (Å)	0.023
Ramachandran plot	
Core (%)	89.7
Allowed (%)	9.6
Generously allowed (%)	0.4
Disallowed (%)	0.3
Average B value, protein (Å^2)	24.7

3. Results and discussion

3.1. Analysis of twinning and NCS

The low $R_{\rm meas}$ and $R_{\rm merge}$ during scaling of the data in space groups $P4$ and $P422$ and the unsatisfactorily high R values after initial refinement suggested that the crystals of the Fab fragment f3p4 were twinned. Closer inspection of the data with the program *TRUNCATE* (Collaborative Computational Project, Number 4, 1994) and the L-test (Padilla & Yeates, 2003) as implemented in the *mmtbx.xtriage* package of *PHENIX* (Adams *et al.*, 2004) verified the presence of twinning in the data set.

The cumulative intensity distribution of the measured data from the Fab fragment f3p4 crystals and theoretical statistical data (Rees, 1980) are shown in Fig. 2(*a*). The discrepancy in the distribution of centric and acentric reflections is very pronounced and gives a good indication of the presence of twinning. To further prove the presence of twinning in the data set, an additional test based on local comparison of the

reflection intensities was performed (Padilla & Yeates, 2003). The twinning fraction α was estimated as 37% using the initial slope of the *H*-test (Yeates, 1997; Yeates & Fam, 1999; Fig. 2*b*).

Assignment of the correct space group to the crystals of the Fab fragment f3p4 was hindered by the presence of both twinning and NCS in the crystals. Packing analysis of the crystals revealed fourfold NCS in the direction of the long axis (Fig. 3). This NCS generates a nearly perfect 4_1 screw axis along the long axis, resulting in molecular-replacement solutions for space groups $P4_1$ and $P4_12_12$. The NCS, in combination with the crystallographic 222 symmetry, generates the near-422 symmetry that was observed in the self-rotation function (Fig. 1).

Since NCS and a twinning operator are both present, both or just one might cause the abnormalities in the intensity distribution of the data. Thus, an RvR plot (Lebedev *et al.*, 2006) based on observed and calculated data was generated after refinement to determine the contributions of the operators. $R_{\rm twin}^{\rm obs} \simeq R_{\rm twin}^{\rm calc}$ indicates the absence of twinning, whereas $R_{\rm twin}^{\rm obs} < R_{\rm twin}^{\rm calc}$ would be found when the data show twinning. Lower values of $R_{\rm twin}^{\rm obs}$ correlate with higher twinning fractions α and values close to zero for both $R_{\rm twin}^{\rm obs}$ and $R_{\rm twin}^{\rm calc}$ indicate misspecified crystal symmetry. The $R_{\rm twin}^{\rm obs}$ of 0.11 and $R_{\rm twin}^{\rm calc}$ of 0.52 indicated nearly perfect twinning and furthermore implied that the NCS did not by itself account for the higher pseudo-symmetry observed in the diffraction data.

3.2. Structural analysis of the Fab fragment f3p4

The Fab fragment f3p4 crystallized with four molecules in the asymmetric unit. Analysis of the four individual molecules showed identical elbow angles [defined as the angle between the two, not necessarily intersecting, pseudo-dyad axes relating the light (V_L) and heavy (V_H) chain variable domains and the light (C_L) and heavy (C_H1) chain constant domains; Stanfield *et al.*, 2006], resulting in an r.m.s. deviation of 0.5–1.1 Å^2 for all C$^\alpha$ atoms. The loop between residues Pro162 and Thr171 of the constant domain of C_H1 is highly flexible and the

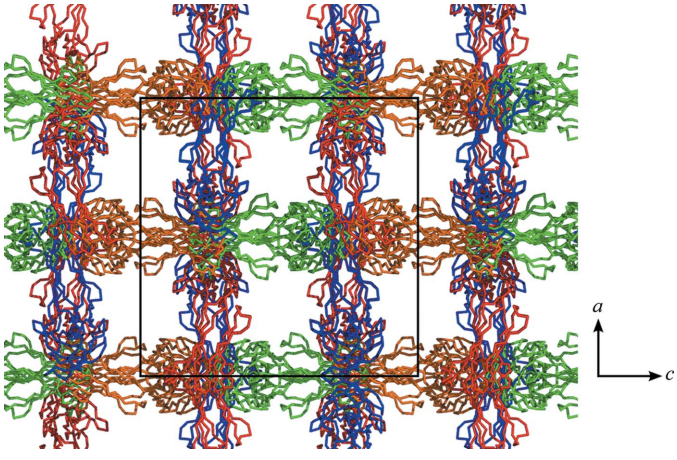


Figure 3
Crystal packing of Fab fragment f3p4. The packing of the molecules generates pseudo-fourfold symmetry parallel to the long axis *b*. The image shows a projection of the *ac* plane with *b* normal to the paper plane. The C$^\alpha$ traces of the light chains of the four molecules are shown in blue, red, green and orange and the unit cell is indicated as a black box.

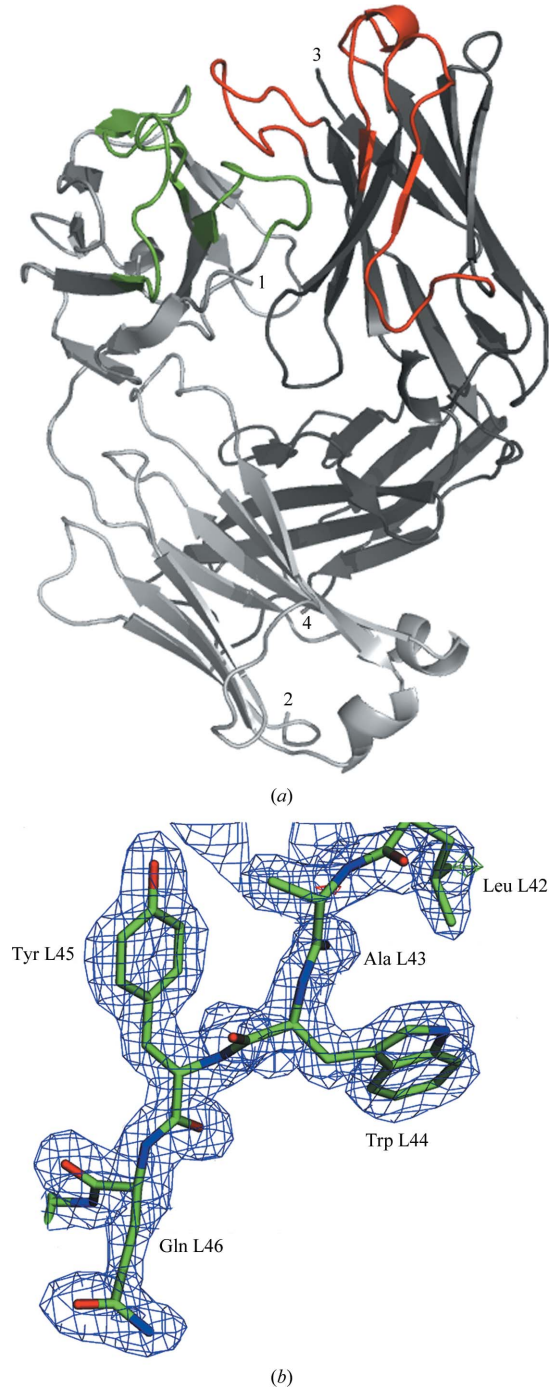


Figure 4
Overall structure of Fab f3p4 and quality of the electron density. (*a*) The overall structure of Fab fragment f3p4 is shown. The light and heavy chains are shown in grey and dark grey, respectively. CDR-L1, CDR-L2 and CDR-L3 of the V_κ chain are coloured green and CDR-H1, CDR-H2 and CDR-H3 of the V_H3 chain are shown in red. Numbers indicate the N- and C-termini of the light (1 + 2) and heavy chains (3 + 4), respectively. (*b*) Amino acids L42–46 (Kabat L34–38) of V_κ are shown. The $2F_o - F_c$ electron-density map at a contour level of 1σ is shown in blue. The $F_o - F_c$ electron density is shown in green (contoured at +3σ) and red (contoured at −3σ).

corresponding electron density is ambiguous; therefore, these residues could not be modelled.

The overall structure of the Fab fragment f3p4 (Fig. 4), which was selected from the HuCAL GOLD library, is very close to the proposed models (PDB codes 1dh5 and 1dhu) of the consensus V_H3 and $V_\kappa3$ frameworks (Knappik *et al.*, 2000). The r.m.s. deviation of all main-chain atoms is 1.31 Å (minimum 0.03 Å, maximum 6.55 Å) for V_H3 and 0.81 Å (minimum 0.08 Å, maximum 2.83 Å) for $V_\kappa3$.

The structure of the Fab fragment f3p4 is the first published crystal structure of a Fab fragment selected against a protein target originating from the synthetic HuCAL GOLD library. It supports the design of the library to contain antibodies corresponding to the natural repertoire. The deviations from the modelled structures (which had been built based on natural antibodies) are minor and are limited to loop regions [L47–51, L83–88, L99–102 and H47–50 in AHo numbering (Honegger & Plückthun, 2001), equivalent to Kabat numbering L39–41, L67–70, L81–84 and H40–43, respectively (Kabat *et al.*, 1991)] outside the central β-sheets, the side-chain orientations of solvent-accessible charged residues and the complementary-determining regions CDR-L1, CDR-L3, CDR-H1 and CDR-H3. These deviations are within the range of the structural variations of antibody structures obtained by hybridoma technology, *i.e.* they are also found in the natural repertoire. Very recently, the structure of a HuCAL Fab fragment binding to a hapten was also reported (Hillig *et al.*, 2008).

The overall conformation of CDR-L1 corresponds well to the typical hairpin structure observed for human $V_\kappa3$. The upper tip of this CDR loop has a high conformational variability in natural antibodies and the detected conformation of this loop is well within the range of human antibody structures (see, for example, PDB codes 1bww, 1dno, 1hez and 1jv5) as well as of the $V_\kappa3$ model. CDR-L1 is also involved in crystal contacts. CDR-L2 hardly varies between antibody V_L domains.

The typical ω shape of the V_κ CDR-L3 is determined by a *cis*-proline at position L136 (Kabat L95), which is carried by more than 90% of the germline V_κ sequences and about 60% of rearranged human V_κ antibody sequences (A. Honegger, unpublished results) as it is the last residue encoded by most of the κ V-segments. In the Fab fragment f3p4, selected from the HuCAL $V_\kappa3/V_H3$ subset (Röthlisberger *et al.*, 2004), a serine is present at this position. Therefore, the CDR-L3 does not adopt an ω-loop conformation. At position L137 (Kabat L96), a *trans*-proline is present and is stabilized by a π–π interaction with the side chain of tryptophan H54 (Kabat H47) of the heavy chain. Residue L137 (Kabat L96) would correspond to the first residue of the κ J_1 segment and would encode a Trp residue (or another aromatic or hydrophobic residue in all J segments), but this residue is frequently mutated in V–J recombination. In contrast to V_κ domains, V_λ domains normally do not have a proline residue at this position in CDR-L3, giving the loop a more hairpin-like shape. The presence of this *trans*-proline in Fab fragment f3p4 prevents a full V_λ CDR-L3-like hairpin conformation, giving it an intermediate shape (Fig. 5).

This amino-acid sequence of CDR-L3 is as expected from the design, as in the HuCAL GOLD library the frequency of serine is 20% at position L136 (Kabat L95) and position L137 (Kabat L96) is fully randomized (except for cysteines). Even though this amino-acid sequence at L136 and L137 (Kabat L95 and L96) is not predominant in natural antibodies, it can be found, for example in the antibody PDB entry 1c1e, which was selected from hybridomas (Fig. 5).

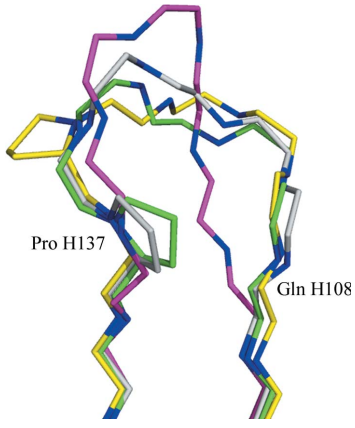


Figure 5
Structure of CDR-L3. A superposition of the main-chain atoms (for clarity, carbonyl O atoms are omitted from the figure) of CDR-L3 of the Fab fragment f3p4 (grey) with typical members of the V_κ and V_λ families is shown. In V_λ members CDR-L3 adopts a hairpin-like structure (magenta; PDB code 8fab). In V_κ members, which predominantly have a proline at position L136 (Kabat L95), the CDR-L3 has an ω-loop conformation (yellow; PDB code 1ay1). The conformation of the CDR-L3 of Fab fragment f3p4 is an intermediate conformation, but is also present in naturally occurring antibodies with a proline residue at L136 (Kabat L95; green; PDB code 1c1e). The compared molecules have an equal number of residues in CDR-L3 and all C^α atoms of L106–L140 (Kabat L88–99) were used for the superposition.

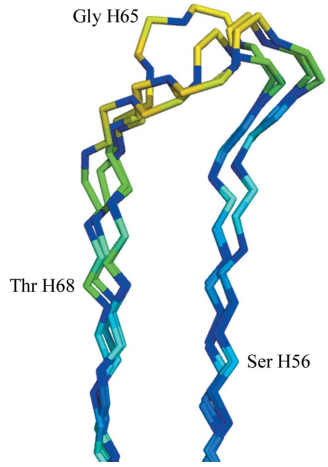


Figure 6
Variability in CDR-H2. Superposition of the four individual CDR-H2 segments of the Fab fragment f3p4 structure. The hairpin loop around Gly H65 (Kabat H54) adopts different conformations in the four molecules and high *B* factors are found in this region. The main-chain trace without carbonyl O atoms is shown. Atoms are coloured light blue, green and yellow according to low, medium and high mobility, respectively. N atoms are coloured dark blue.

In the V_H3 heavy chain of Fab fragment f3p4, CDR-H1 shows a conformation typical of this subclass. In CDR-H2, the glycine at position H65 (Kabat H54) is the centre of a short hairpin loop consisting of amino acids H59–67 (Kabat H52–56). This CDR shows the highest mobility of the structure in all four molecules, which is reflected by multiple conformations and higher *B* values (Fig. 6).

CDR-H3 varies most in sequence and length. Its conformation is therefore difficult to predict. The hydrogen bonding between Arg H108 (Kabat H94) and Asp H137 (Kabat H101) that gives rise to the typical kink of CDR-H3 is also observed in the structure of the Fab fragment f3p4. A relatively open conformation of this CDR is found, allowing residues of the target to penetrate into the groove at the V_L–V_H interface (Fig. 7). CDR-H3 is also involved in crystal contacts.

The side-chain conformations of Phe L101 (Kabat L83) and Arg H82 (Kabat H71) need to be discussed here. Phe L101 (Kabat L83) is rather solvent-exposed in the modelled structure (Knappik *et al.*, 2000) of $V_\kappa3$ (PDB code 1dhu) but is buried in the Fab fragment f3p4 structure (Fig. 8*a*). Both conformations occur in antibody structures and, therefore, its conformation cannot be predicted. Arg H82 (Kabat H71) was previously modelled (Knappik *et al.*, 2000) to be solvent-exposed in V_H3 (PDB code 1dh5). In the Fab fragment f3p4 structure, this residue is fully buried and forms hydrogen bonds to the side chain of Asn H84 (Kabat H73) and to the main chain of Ser H39 (Kabat H34) (Fig. 8*b*). The buried conformation and the hydrogen-bonding pattern is observed in the majority of antibody V_H structures.

The Fab fragment f3p4 precipitates at concentrations above 13 mg ml^{-1} in buffers containing 150 m*M* NaCl or above. This precipitate can be redissolved by simply dialyzing out the salt, indicating an association of native molecules. We have selected four Fab fragments (Röthlisberger *et al.*, 2004) to bind to CitS and two of them show this effect. The framework of all four Fab fragments is the same ($V_\kappa3$/V_H3) and they differ only in the CDRs. CDR-L1 (residues L30, L32, L33, L39 and L40; Kabat L27c, L27e, L27f, L31 and L32), CDR-L2 (L57, L58, L68 and L69; Kabat L49, L50, L52 and L53), together with an exposed tryptophan (L133; Kabat L95*c*) of CDR-H3, are involved in crystal contacts to the constant domain of the light chain (residues L190, L192, L193, L228–232, L234, L236, L244, L245, L247, L249, L251 and L252. This contact consists of 50% nonpolar residues, contains four hydrogen bonds and covers an area of about 450 Å^2. By comparing the CDRs of all four selected Fab fragments, we found that both Fab fragments that do not show precipitation have a CDR-L1 that is shortened by one amino acid and lacks the tyrosine at position L40 (Kabat L32). CDR-L1 might, therefore, not only contribute to

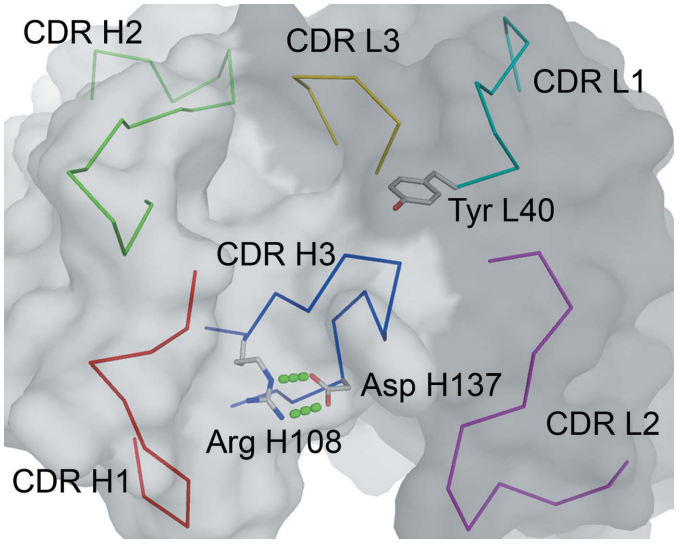
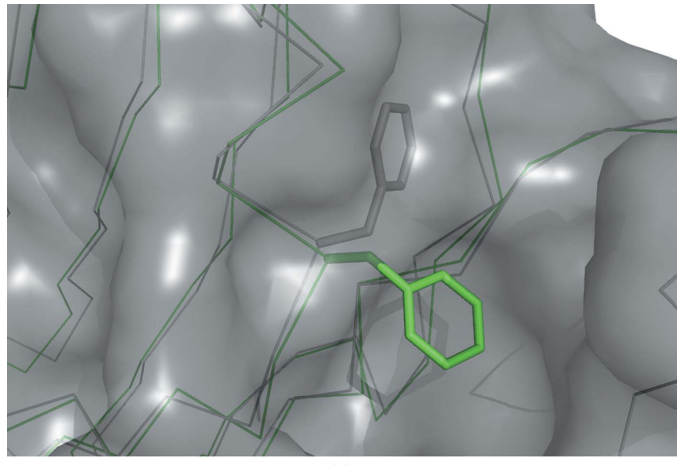
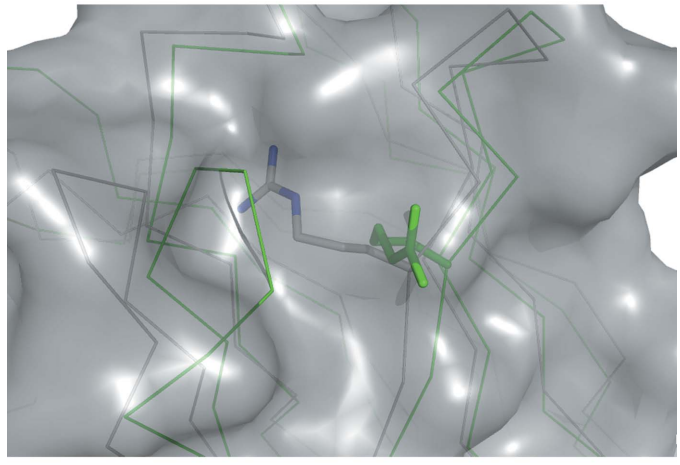


Figure 7
Surface representation of the binding region. C$^\alpha$ traces of CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2 and CDR-L3 are shown and coloured red, green, blue, cyan, magenta and yellow, respectively. Residues Arg H108 (Kabat H94) and Asp H137 (Kabat H101) forming the conserved hydrogen bonds as well as the solvent-exposed residue Tyr L40 (Kabat L32) are shown in stick representation.



(*a*)



(*b*)

Figure 8
Solvent-accessibility of Phe L101 and Arg H82. (*a*) Phe L101 (Kabat L83) is buried in the Fab fragment f3p4 structure (grey) and is solvent-accessible in the model (green; PDB code 1dhu). (*b*) Residue ArgH82 (Kabat H71) is buried in the Fab fragment f3p4 structure (grey) and is more solvent-exposed in the model (green; PDB code 1dh5). The surface was calculated using the coordinates of the Fab fragment f3p4 structure. Models 1dh5 and 1dhu were superimposed with one molecule of the Fab fragment f3p4 structure using the program *Coot* (Emsley & Cowtan, 2004; Krissinel & Henrick, 2004).

the crystal contacts but might also be responsible for the precipitation at high concentration in the presence of salt.

4. Concluding remarks

This work represents the first experimental structure solution of a Fab fragment selected against a protein target derived from the fully synthetic HuCAL GOLD antibody library. Comparison of the structure with theoretical models revealed minor differences in regions that are difficult to predict. These observed discrepancies indicate residues that need careful examination during modelling and thus this structure may help in further improving antibody modelling.

We gratefully acknowledge the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland for providing synchrotron beamtime, Peer Mittl and Annemarie Honegger for helpful discussions and Stefan Dudli and Dominik Hänni for help in protein purification and crystallization. This work was supported by the Swiss National Science Foundation (SNSF), the National Center of Excellence in Research (NCCR) Structural Biology program of the SNSF and SNSF grant 3100-065344.01/1 to AP.

References

Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Pai, R. K., Read, R. J., Romo, T. D., Sacchettini, J. C., Sauter, N. K., Storoni, L. C. & Terwilliger, T. C. (2004). *J. Synchrotron Rad.* **11**, 53–55.

Bass, S., Gu, Q. & Christen, A. (1996). *J. Bacteriol.* **178**, 1154–1161.

Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grütter, M. G. & Plückthun, A. (2004). *Nature Biotechnol.* **22**, 575–582.

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.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Mol. Biol.* **4**, 269–275.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Ewert, S., Huber, T., Honegger, A. & Plückthun, A. (2003). *J. Mol. Biol.* **325**, 531–553.

Hillig, R. C., Urlinger, S., Fanghanel, J., Brocks, B., Haenel, C., Stark, Y., Sulzle, D., Svergun, D. I., Baesler, S., Malawski, G., Moosmayer, D., Menrad, A., Schirner, M. & Licha, K. (2008). *J. Mol. Biol.* **377**, 206–219.

Honegger, A. & Plückthun, A. (2001). *J. Mol. Biol.* **309**, 657–670.

Huber, T., Steiner, D., Röthlisberger, D. & Plückthun, A. (2007). *J. Struct. Biol.* **159**, 206–221.

Hunte, C. & Michel, H. (2002). *Curr. Opin. Struct. Biol.* **12**, 503–508.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesmann, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, 5th ed. Bethesda: NIH Publications.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). *J. Mol. Biol.* **296**, 57–86.

Kohl, A., Amstutz, P., Parizek, P., Binz, H. K., Briand, C., Capitani, G., Forrer, P., Plückthun, A. & Grütter, M. G. (2005). *Structure*, **13**, 1131–1141.

Kratzin, H. D., Palm, W., Stangel, M., Schmidt, W. E., Friedrich, J. & Hilschmann, N. (1989). *Biol. Chem. Hoppe-Seyler*, **370**, 263–272.

Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.

Lebedev, A. A., Vagin, A. A. & Murshudov, G. N. (2006). *Acta Cryst.* D**62**, 83–95.

MacRae, I. J. & Doudna, J. A. (2007). *Acta Cryst.* D**63**, 993–999.

Ostermeier, C. & Michel, H. (1997). *Curr. Opin. Struct. Biol.* **7**, 697–701.

Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.

Pos, K. M. & Dimroth, P. (1996). *Biochemistry*, **35**, 1018–1026.

Rauchenberger, R. *et al.* (2003). *J. Biol. Chem.* **278**, 38194–38205.

Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.

Rees, D. C. (1980). *Acta Cryst.* A**36**, 578–581.

Röthlisberger, D., Pos, K. M. & Plückthun, A. (2004). *FEBS Lett.* **564**, 340–348.

Schweizer, A., Roschitzki-Voser, H., Amstutz, P., Briand, C., Gulotti-Georgieva, M., Prenosil, E., Binz, H. K., Capitani, G., Baici, A., Plückthun, A. & Grütter, M. G. (2007). *Structure*, **15**, 625–636.

Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O. & Grütter, M. G. (2006). *PLoS Biol.* **5**, e7.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Stanfield, R. L., Zemla, A., Wilson, I. A. & Rupp, B. (2006). *J. Mol. Biol.* **357**, 1566–1574.

Usón, I., Pohl, E., Schneider, T. R., Dauter, Z., Schmidt, A., Fritz, H. J. & Sheldrick, G. M. (1999). *Acta Cryst.* D**55**, 1158–1167.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.

Yeates, T. O. & Fam, B. C. (1999). *Structure*, **7**, R25–R29.