doi: 10.1016/S0022-2836(02)00038-4 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 318, 135–147

JMB

Crystal Structure of the Anti-His Tag Antibody 3D5 Single-chain Fragment Complexed to its Antigen

Markus Kaufmann, Peter Lindner, Annemarie Honegger, Kerstin Blank Markus Tschopp, Guido Capitani, Andreas Plückthun* and Markus G. Grütter*

Biochemisches Institut Universität Zürich Winterthurerstrasse 190 CH-8057 Zurich, Switzerland

**Corresponding authors*

The crystal structure of a mutant form of the single-chain fragment (scFv), derived from the monoclonal anti-His tag antibody 3D5, in complex with a hexahistidine peptide has been determined at 2.7 Å resolution. The peptide binds to a deep pocket formed at the interface of the variable domains of the light and the heavy chain, mainly through hydrophobic interaction to aromatic residues and hydrogen bonds to acidic residues. The antibody recognizes the C-terminal carboxylate group of the peptide as well as the main chain of the last four residues and the last three imidazole side-chains. The crystals have a solvent content of 77% (v/v) and form 70 Å-wide channels that would allow the diffusion of peptides or even small proteins. The anti-His scFv crystals could thus act as a framework for the crystallization of His-tagged target proteins. Designed mutations in framework regions of the scFv lead to high-level expression of soluble protein in the periplasm of *Escherichia coli*. The recombinant anti-His scFv is a convenient detection tool when fused to alkaline phosphatase. When immobilized on a matrix, the antibody can be used for affinity purification of recombinant proteins carrying a very short tag of just three histidine residues, suitable for crystallization. The experimental structure is now the basis for the design of antibodies with even higher stability and affinity.

Keywords: crystal structure; polyhistidine tag; scFv antigen complex; single-chain antibody; antibody engineering

Introduction

Fusion proteins and tag sequences facilitate the purification and/or detection of recombinant proteins.[1] Many fusion systems have been described, either based on folded domains, such as glutathione-*S*-transferase or maltose-binding protein, which bind to an easily immobilized ligand, or based on short peptide tags (e.g. Flag, VSV, Myc), which are usually recognized by an antibody. As these antibodies first have to be immobilized on a column, this step adds greatly to the expense of large-scale purification. Probably the most widely used peptide tag is the His tag, consisting of 4–6 consecutive histidine residues. The main attraction about this tag is that the affinity matrix for the His tag is a simple metal-chelating resin. In this method, neighboring histidine residues coordinate divalent metal ions such as Ni^{2+}, Co^{2+}, Zn^{2+}, and Cu^{2+} bound to chelating

Present address: K. Blank, nanotype GmbH, Lochhamer Schlag 12, D-82166 Gräfelfing, Germany.

Abbreviations used: CDR, complementarity-determining region; ELISA, enzyme-linked immunosorbent assay; ESI-MS, electron spray ionisation mass spectrometry; ESRF, European Synchrotron Radiation Facility; IDA, iminodiacetic acid; IMAC, immobilized metal-ion affinity chromatography; IPTG, isopropyl-β-D-thiogalactoside; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; Mes, 2-morpholinoethanesulfonic acid; NTA, nitrilotriacetic acid; PDB, Protein Data Bank; PEG, polyethylene glycol; r.m.s.d, root-mean-square deviation; scFv, single-chain antibody fragment; SNBL, Swiss-Norwegian Beam Line; TED, *N*,*N*,*N*′-tris-(carboxymethyl)ethylenediamine; V_L, variable domain of the antibody light chain; V_H, variable domain of the antibody heavy chain; VSV, amino acid residues 497–511 of VSV (vesicular stomatitis virus glycoprotein); wt, wild-type.

E-mail addresses of the corresponding authors: plueckthun@biocfebs.unizh.ch; gruetter@bioc.unizh.ch

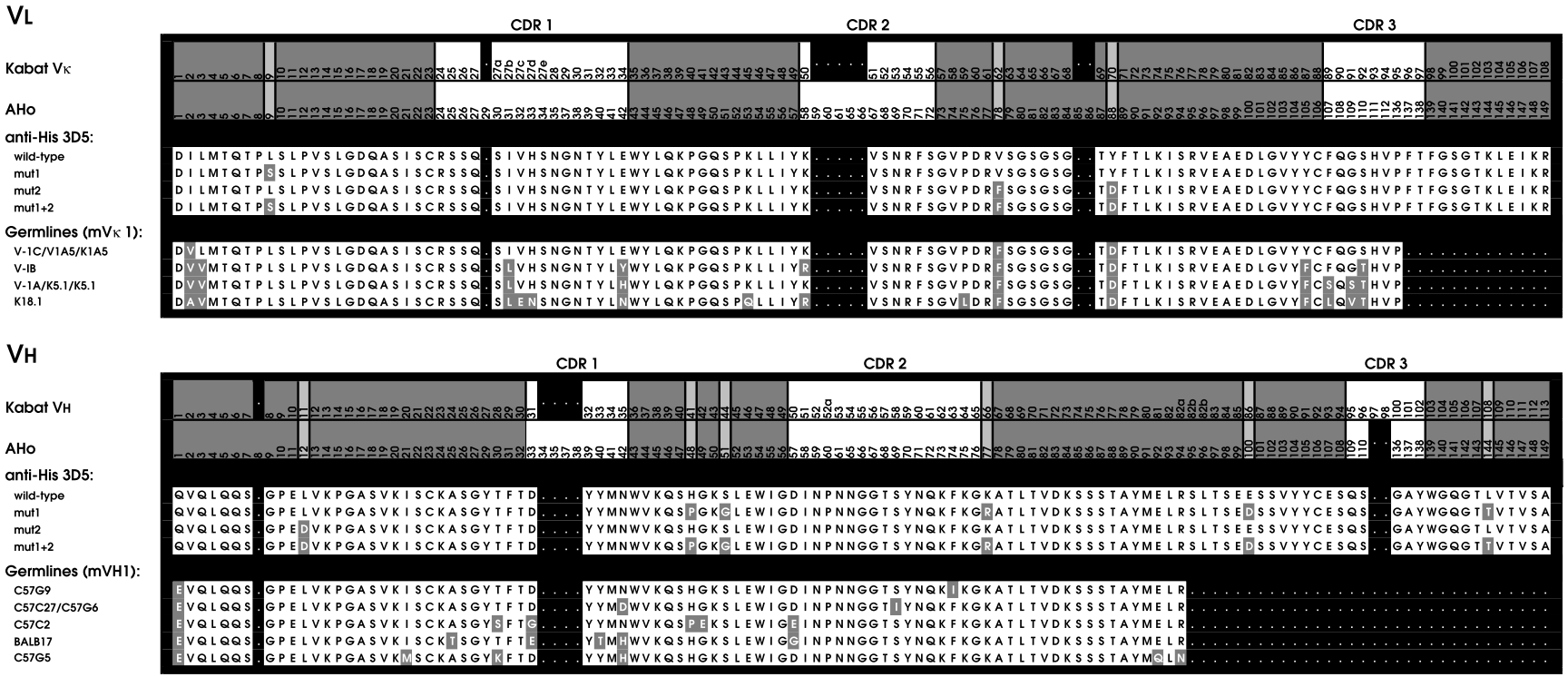

Figure 1. Sequence alignment showing the amino acid sequences of V_L and V_H of the wild-type anti-His antibody and of the constructs mut1, mut2, mut1 + 2, as well as the murine germline sequences (ABG: Germline directories of mouse (http://www.ibt.unam.mx/vir/V_mice.html), most closely related to the murine antibody 3D5, from which the anti-His scFv is derived. Residues of V_L are structurally aligned to the residues of V_H.[19] The construct mut1 carries the mutations Leu L9 to Ser, His H48 to Pro, Ser H51 to Gly, Lys H77 to Arg, Glu H100 to Asp and Leu H144 to Thr, mut2 the mutations Val L78 to Phe, Tyr L88 to Asp and Leu H12 to Asp. mut1+2 combines the mutations of mut1 and mut2. The headers indicate the correspondence between Kabat residue numbering[20] and the AHo residue numbering.[19] The latter was used throughout this paper. Residue numbers with a white background correspond to CDR regions. Residue numbers of mutated residues in the anti-His have a light gray background. The background of non-conserved amino acids is gray.

ligands, e.g. nitrilotriacetic acid (NTA), iminodiacetic acid (IDA) or *N*,*N*,*N*′-tris-(carboxymethyl) ethylenediamine (TED), usually linked to an agarose gel matrix.[2–4] Oligohistidine tags offer the possibility to purify recombinant proteins by immobilized metal-ion affinity chromatography (IMAC) under native or denaturing conditions. The purity of proteins isolated by IMAC is sufficient for most technical applications. Greater purity, however, is required for medical applications and structure determination. His tags can be used in immunoaffinity chromatography, in conjunction with IMAC,[5] to further increase purity by using the same tag twice, but each time removing different impurities. To overcome the prohibitive expense of using monoclonal antibodies on a large scale, a new method was developed in which a recombinant anti-His scFv can be directly immobilized from crude *Escherichia coli* supernatants.[6]

His-tagged proteins have been used successfully in structure determination by X-ray crystallography,[7–9] and by multidimensional NMR spectroscopy,[10] and even as therapeutic proteins in clinical applications.[11,12] It is thus frequently not necessary to remove the tag.

While Ni-NTA derivatized enzymes can be used for the detection of His-tagged proteins (Qiagen, Germany), antibodies are the most versatile reagent class for detection and for further purification. Several antibodies recognizing His tags have been generated and some are available commercially. Certain anti-His tag antibodies, such as His·Tag® Monoclonal Antibody (Novagen, US), Penta·His™ (Qiagen, Germany), and Tetra·His™ (Qiagen, Germany), 13/45/31,[13] recognize histidine residues irrespective of their neighboring sequences, whereas others require a defined environment like RGS·His™ (Qiagen, Germany),[14] HisG (Novagen, US), or the parent antibody 3D5, from which the present single-chain fragment has been derived.[15] The generation of this monoclonal antibody has been described[15] and it is distributed commercially (Anti-His (C-term) Antibody, Invitrogen Life Technologies, Basel, Switzerland).

The anti-His tag antibody 3D5 is of special interest, as it requires a C-terminal histidine residue,[15] thereby offering very high specificity in detection and purification. Only this antibody is available in recombinant form as a scFv fragment,[15] allowing easy access to a wide variety of formats for different applications. The scFv fragment (V_L–$(Gly_4Ser)_4$–V_H) was fused to alkaline phosphatase. *E. coli* extracts from cells expressing this construct can be used directly to detect His-tagged proteins in Western blot experiments or ELISA.[15] By using a novel immobilization strategy based on engineered chitin-binding domains, the anti-His scFv can be immobilized selectively, directly from the crude bacterial extract and used for inexpensive immunoaffinity purification of His-tagged proteins.[6]

We first built a homology model of the Fv fragment of the anti-His tag antibody 3D5 and chose nine mutations, which we predicted to increase the expression level and/or stability of the antibody in *E. coli*. Indeed, this strategy was very successful, increasing the expression level significantly. Here, we describe the crystal structure of this engineered mutant of anti-His scFv 3D5 in complex with a hexahistidine peptide.

Results and Discussion

Rational design of anti-His 3D5 scFv mutants leading to increased expression in *E. coli*

The wild-type scFv, derived from a murine monoclonal antibody,[15] was poorly produced in *E. coli*, especially in the absence of a fusion partner. To obtain sufficient amounts of protein for structural characterization, the stability and folding efficiency of the scFv had to be improved significantly. A homology model[15] of the Fv fragment was built, based on the structures of the anti-cholera toxin antibody TE33 (PDB entry 1TET[16]), the anti-digoxigenin antibody 26-10 (PDB entry 1IGI[17]) and the CDR H3 structure of the liganded anti-HIV-peptide antibody 50.1 (PDB entry 1GGI[18]). On the basis of this model and a comparison with the closest murine germline sequences (Figure 1), nine mutations were proposed that were expected to improve both thermodynamic stability and folding yield. They were all introduced at once as two sets of mutations (mut1 and mut2, Figure 1).†

The construct mut1 comprised six mutations, which had been incorporated successfully in various combinations into other scFvs.[21–24] The mutations Lys H77 to Arg and Glu H100 to Asp were predicted to improve thermodynamic stability significantly. The side-chains of the two residues interact to form a buried salt-bridge. Proba *et al*.[21] showed that in the disulfide-free V_H domain of antibody ABPC48, the replacement of the germline-encoded Lys by Arg led to significant stabilization. Position H100 is Asp in more than 98% of all human and murine V_H germline sequences, and Asp is predominant in this position in all other types of variable domains (V_λ, V_κ, TCR, V_α, V_β, V_δ, V_γ).

The optimal formation of the β-turn at the bottom of the V_H domain (residues H47–H51) was expected to have a strong effect on folding efficiency. Position H48 has a strong preference for Pro (94% of all human and 82% of murine sequences), while in position H51 Gly is particularly prevalent in human sequences (90% of

† Here, we follow a new numbering scheme for antibody and TCR variable domains, which is based entirely on structural superposition.[19] Conversion tables, as well as a complete set of all renumbered and aligned variable domains of the antibody and TCR structure PDB files are available in a new database (http://www.biochem.unizh.ch/antibody). The correspondence with residue numbers according to the Kabat numbering scheme[20] is indicated in Figure 1.

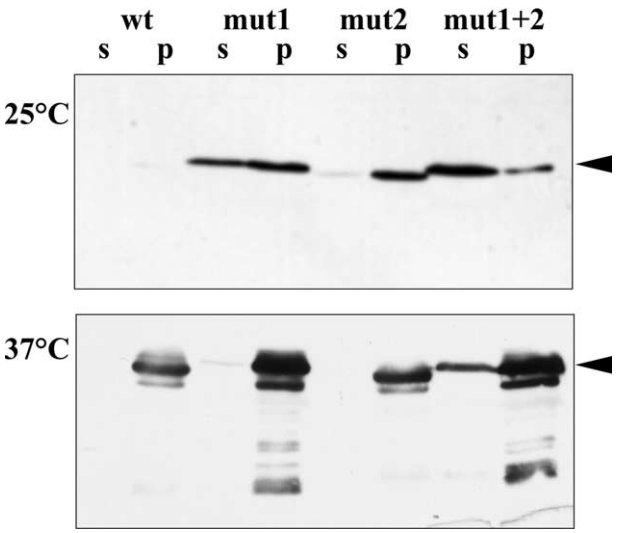

Figure 2. Western blots of *E. coli* cells expressing the different anti-His scFvs (Figure 1) grown at 25°C and at 37 °C. Cells from 20 ml cultures were standardized by resuspending them in 1 ml buffer per final $A_{550\,nm} = 1$ and opened by French Press. Of the resulting lysate, 16 μl soluble supernatant were applied per lane. The insoluble pellet fraction was resuspended in the same corresponding volume. Again, 16 μl were loaded per lane. Introduction of the two sets of mutations lead to higher expression compared to the wild-type scFv. The full synergistic effect, however, shows up in the high level of soluble expression after combining the two sets to mutant "mut1+2" (Figure 1), preferably at 25 °C growth temperature. The scFvs were detected with an antibody against the N-terminal Flag tag. s, supernatant; p, pellet; wt, wild-type anti-His; mut1 and mut2, anti-His carrying the mutations specified in Figure 1; mut1+2, anti-His combining the mutations of mut1 and mut2.

human and 58% of murine sequences). Forsberg *et al*.[22] showed that the combined mutation of His H48 to Pro, Ser H51 to Gly and Val H142 to Gly led to a significant decrease of the toxicity to the bacteria producing the fusion construct derived from the antibody 4T5. While that effect could be explained by the restoration of Gly H142 in the Gly–X–Gly motif forming the β-bulge in framework 4, their results suggested that at least the other two mutations do not have a deleterious effect. In V_L domains, the replacement of Pro L48 by Ser destabilized the domain by 2.4 kJ/mol.[23] We therefore included the mutations of His H48 to Pro and Ser H51 to Gly. Nieba *et al*.[24] showed that scFvs benefit from the replacement of exposed hydrophobic residues of the former V_L/C_L and V_H/C_H interface by more hydrophilic residues. Therefore, we also replaced Leu L9 by Ser and Leu H144 by Thr.

Three new mutations were introduced into the construct mut2 that were expected to have a positive effect on folding and/or stability. The wild-type anti-His antibody contained Val in position L78 instead of Phe found in the germlines of the murine Vκ1 germline subgroup, to which this anti-His antibody belongs. This replacement of a bulky core residue by a smaller side-chain leads to the formation of a cavity in the model. We therefore decided to mutate this Val back to Phe, expecting that this would increase stability. Additional mutations in the mut2 construct were the replacement of the solvent-exposed Tyr L88 in the outer loop by Asp, as found in the Vκ1 germlines, since 96% of all human and 66% of all murine sequences contain an acidic residue in this position. In addition, the hydrophobic surface residue Leu H12 was replaced by Asp. The construct mut1+2 combines all nine mutations. The amino acid sequences of wild-type V_L and V_H are shown in Figure 1, aligned with the sequences of the variants mut1, mut2, mut1+2 and the most closely related germline sequences. For convenience, the variant mut1+2 that was used for structure determination is called anti-His scFv.

Expression of anti-His scFv

Wild-type anti-His scFv is expressed in *E. coli* only poorly. At 25 °C, the construct mut2 led to only a slight improvement in soluble expression, whereas mut1 showed a dramatic increase in soluble periplasmic expression of anti-His scFv. Both sets of mutations in combination (mut1+2) had a synergistic effect and led to an even further increased amount of soluble anti-His scFv, as detected by Western blot analysis (Figure 2). At 37 °C, expression levels were increased dramatically, but still only a fraction of the protein was soluble.

The construct mut1+2 used for crystallization and structure determination contains three additional residues (Asp–Tyr–Lys) of a Flag recognition site at the N terminus. The N-terminal 113 amino acid residues (V_L) and the C-terminal 114 amino acid residues (V_H) of the anti-His scFv (Figure 1) were connected by the 20 residue peptide linker $(Gly_4Ser)_4$.

Purification, characterization and crystallization

About 2 mg of protein was isolated from 3 l of standard shake flask culture from the soluble fraction of *E. coli*. The anti-His scFv was purified by cation-exchange chromatography and affinity chromatography using an immobilized octahistidine peptide, resulting in very pure material. The antibody–ligand complex was formed by mixing anti-His scFv with the hexahistidine peptide in a 1:1.5 molar ratio at pH 6.5. The mass of the protein was confirmed by ESI-MS. The mass of the peptide in the complex was determined by MALDI-TOF MS and, while 60% of the peptide was hexahistidine, 40% of the peptide molecules contained only five histidine residues. This did not have any effect on the complex formation, since anti-His scFv recognizes pentahistidine equally as well as hexahistidine.[15]

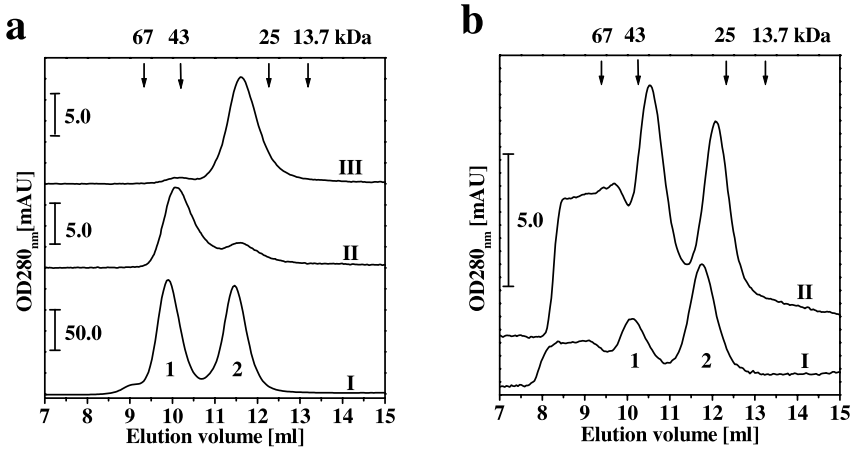

Figure 3. Size exclusion chromatography of the anti-His scFv: hexahistidine peptide complex. Peak 1 is the dimeric form and peak 2 is the monomeric form of the anti-His scFv. The elution volumes of standard protein are indicated with arrows. (a) Analysis of the complex and the stability of the dimeric form: I, complex at the concentration used for crystallization; II, aliquot from peak 1 (dimer) of run I reloaded onto the column after two hours; III, aliquot from peak 1 from run I reloaded after 16 days stored at 4 °C. The sample analyzed in II and III is diluted more then 40-fold, compared to the sample in I. (b) Analysis of the crystallized complex: I, dissolved crystals and mother liquor of one crystallization experiment; II, 100 crystals washed and dissolved in gel filtration buffer.

ScFvs can dimerize,[25,26] depending on the protein concentration, the linker length and structural details in the individual scFv, which are not completely understood. Size-exclusion chromatography of the anti-His scFv–peptide complex at high concentration (3.8 mg/ml) revealed a monomer to dimer ratio of approximately 1:1, with a small fraction of higher aggregates (Figure 3(a), I). Incubation of this preparation at five- and tenfold dilution for 16 days at 4 °C did not change this ratio significantly (not shown). The isolated monomer fraction stored at 40-fold dilution (0.1 mg/ml) for 16 days at 4 °C remained monomeric (not shown), while isolated dimers converted slowly to monomers. After two hours, a small fraction of monomers was detectable (Figure 3(a), II) and after 16 days the dimers had converted to monomers almost completely (Figure 3(a), III). This is consistent with previous results, which showed that at low concentrations scFv monomers are the thermodynamically more stable species.[27]

Hexagonal crystals were grown at pH 6.4 at 4 °C from a mixture of monomers and dimers (Figure 3(a)) using polyethylene glycol and magnesium acetate as precipitant. Crystals belong to space group $P3_221$. There is one molecule of the scFv–peptide complex in the asymmetric unit. In preliminary experiments, an scFv containing an additional Myc tag produced crystals, which were difficult to reproduce, and were found not to be suitable for structure determination. Removal of the flexible Myc tag at the DNA-level improved crystallization significantly.

Size-exclusion chromatography was carried out with proteins from dissolved crystals to assess the oligomeric state of anti-His scFv in the crystals. About 100 crystals were harvested, washed in a stabilizing solution and dissolved in the buffer used for size-exclusion chromatography. The solution from dissolved crystals contained monomers and dimers in a 1:1 ratio (Figure 3(b), II). The protein from the mother liquor was analyzed, too. Monomers and dimers were present in a ratio of 2:1 (Figure 3(b), I). Since it was shown that isolated dimers dissociate at low concentration (Figure 3(a)), these findings are consistent with crystals being built from dimers.

Structure determination

The three-dimensional structure of the anti-His scFv–hexahistidine complex was determined and

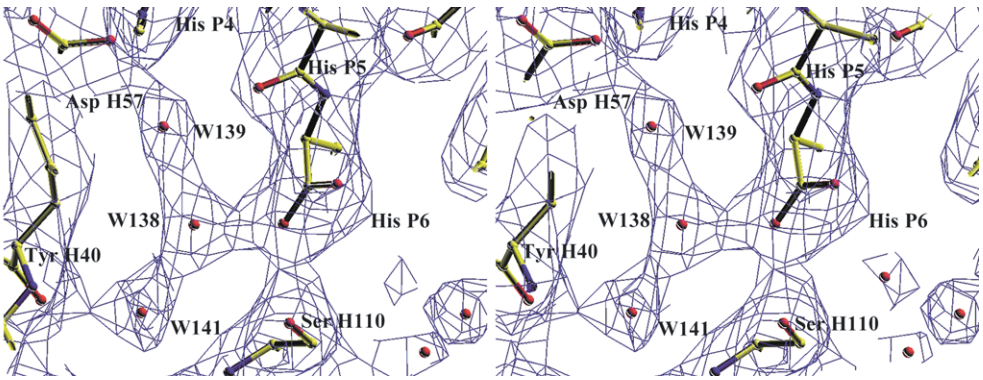

Figure 4. $2F_oF_c$ electron density around the C terminus of the peptide. The map is contoured at 1.0σ. Red spheres represent water molecules. Amino acids are colored by atom type: C (yellow), N (blue), O (red). Picture prepared with O.[43,44]

Table 1. Data collection, refinement statistics and final model

A. Crystal parameters	
Space group	$P3_221$
Unit cell constants	
$a = b$ (Å)	106.51
c (Å)	92.80
Content of asymmetric unit	One scFv–peptide complex
Solvent content (%)	77
B. Data collection	
Beam line at ESRF	SNBL BM1A
Wavelength (Å)	0.873
Temperature (K)	100
No. crystals	1
Resolution range (Å)	30–2.7
Observed reflections	128,480
Unique reflections	16,996
Redundancy	7.6
B-value (Wilson plot) (Å^2)	52.8
Data completeness (%)	99.0
R_{merge} (%)	11.6
C. Refinement and final model	
Resolution range (Å)	30–2.7
No. unique reflections	16,936
R_{work} (%)	19.04
R_{free} (%) (966 test reflections)	22.20
No. non-hydrogen atoms	1946
No. protein residues (atoms)	
Peptide	4 (36)
V_L	113 + 2 linker residues (882)
V_H	114 (872)
No. water molecules	156
B-factor protein (Å^2)	
V_L	33.8
V_H	30.8
Hexahistidine peptide	25.8
r.m.s. deviation from ideal geometry	
Bond (Å)	0.0054
Angles (°)	1.3173
Ramachandran plot statistics (non proline and glycine residues)	
Residues in most favored regions (%)	172 (89.1%)
Residues in additional allowed regions (%)	20 (10.4%)
Residues in general allowed regions (%)	0
Residues in disallowed regions (%)	1 (0.5%)

refined to a final R-factor of 19.0% (R_{free} 22.2%) at 2.7 Å to describe the antigen–antibody interactions at the atomic level and to elucidate why only C-terminal His tags are recognized by the antibody. The final $2F_o - F_c$-electron density map shows well-defined and continuous electron density for all 113 residues of V_L, two linker residues and all 114 residues of V_H. The conformation of Ala H114 as well as the side-chains of a few surface arginine and lysine residues are not well defined. Eighteen linker residues and the three additional N-terminal residues of the Flag tag are disordered. Electron density is visible for main-chain atoms of His P3 to P6 and side-chain atoms of His P4 to P6 of the hexahistidine ligand (Figure 4). For residues His P1 and P2 and the imidazole ring of His P3, there is no electron density. His P3 was therefore modeled as Ala and His P1 and P2 were removed from the model. The structure was validated with the programs PROCHECK[28] and WHAT_CHECK.[29] There was only one non-glycine residue (Val L67), part of CDR 2, in the disallowed region of the Ramachandran plot. However, this combination of ϕ and ψ torsion angles in this position is typical for antibody V_L domain structures.† Crystallographic statistics are given in Table 1.

Crystal packing

Crystals of the anti-His scFv in complex with hexahistidine have an extremely high solvent content of 77% (v/v) (V_M 5.4 Å^3/Da). Large channels can be observed in all unit-cell axis directions (Figure 5). The minimal diameter of the channels in the z-direction (Figure 5) is 70 Å and 38 Å in the x and y-direction (not shown).

Each anti-His scFv is in contact with three symmetry-related molecules. Four regions of molecule 0 are involved in crystal contacts, two in V_L and

† http://www.biochem.unizh.ch/antibody

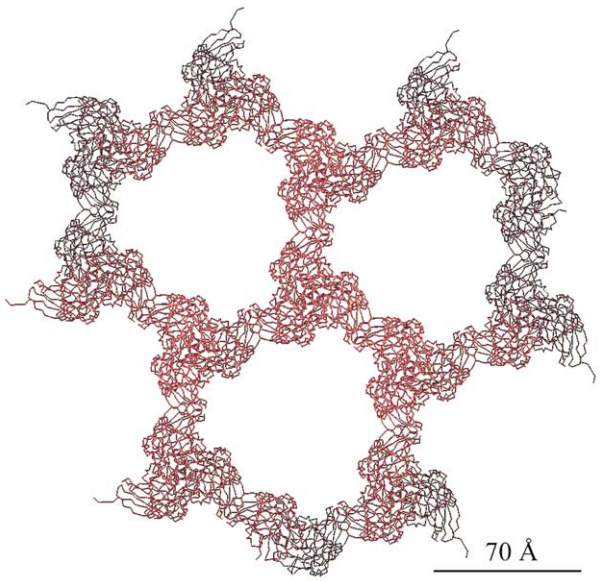

Figure 5. Crystal packing I. scFv-crystals contain large solvent channels. Packing in *xy*-plane, channels of a minimal diameter of 70 Å in the *z*-direction. Pictures prepared with O.[43,44]

two in V_H (Figure 6). There is surface complementarity and hydrophobic interactions to symmetry-related molecules 1–3. Additionally, there are specific contacts. Residues L35/38 and L68–69 of molecule 0 form hydrogen bonds with residues H51 and H82–85 of the symmetry-related molecule 2 (Figure 6). The same symmetry-related interactions are found for contacts between molecules 0 and 3. The size of the contact area of two molecules (0 and 3 or 0 and 2), also called the interface area (B)[30] is 984 Å^2 (calculated using GRASP[31]). It is the difference between the sum of the molecular surface areas of two isolated molecules and the surface of the complex formed by these two molecules. Molecules 0 and 1 share two spatially separated contact areas. The C-terminal V_L residues, L148/L149 of molecule 0, form direct hydrogen bonds to H11 of molecule 2. The first two linker residues, M150 and M151, of molecule 0 interact with H141–H144 of molecule 2. Together with the van der Waals surface complementarity and hydrophobic interactions, this results in a total interface area of 1644 Å^2, which represents a typical value for interfaces formed by subunits in oligomeric proteins.[30] It is even larger than the total interface area of 1253 Å^2 between V_H and V_L within molecule 0. But the V_L/V_H interface has a higher complementarity. The size of the V_L/V_H interface of the anti-His scFv is very small compared with the average V_L/V_H interface[32] of 1622($\pm$184) Å^2 (distribution 1200–2100 Å^2) in antibody structures (unpublished results, based on the analysis of 128 structures taken from the PDB).

It is worth noting that the incorporation of the described mutations (mut1+2) increased solubility and stability of the scFv, and helped in crystallization, since four of the nine mutated residues (Ser L9, Asp H12, Pro H48 and Thr L144) are involved in crystal contacts.

Size-exclusion chromatography showed that the anti-His scFv forms dimers in solution at high concentrations, and that the anti-His scFv from dissolved crystals is partially dimeric. We thus assume that in the crystal, molecule 0 and molecule 1 represent the scFv dimer because of the large buried surface area (Figure 6). Dimers can be formed by a domain swap of two scFvs, where the V_L domain of one scFv (molecule 0) pairs with the V_H of a second scFv (molecule 1) and *vice versa* (as described for a diabody[25,33]). Alternatively, dimers may consist of two scFvs forming intramolecular V_L/V_H interfaces (Figure 6), which are associated through the large crystal interface. Since only the first two residues of the linker are visible in the structure of the anti-His scFv, we cannot distinguish rigorously between the two possibilities.

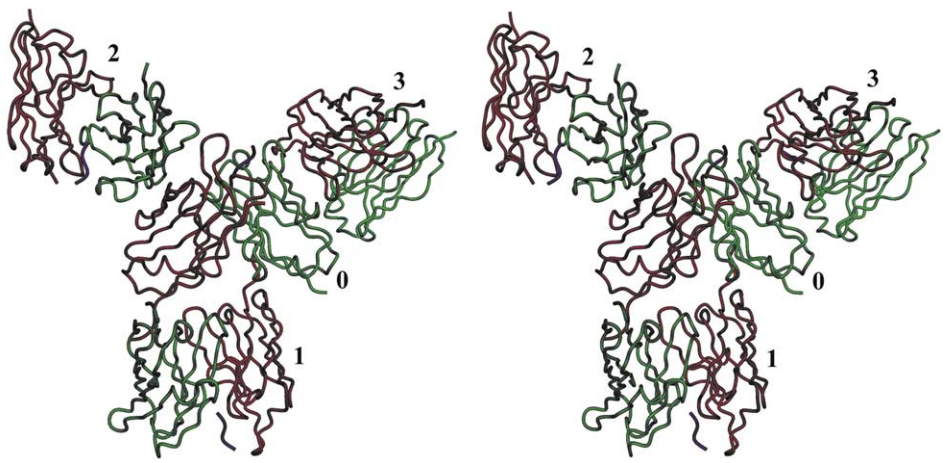

Figure 6. Crystal packing II. One scFv is in contact with three symmetry related molecules. ScFv molecule 0 is in contact with scFv molecules 1–3. Light chains are colored in red, heavy chains in green and the peptide in blue. Picture prepared with DINO. (DINO: Visualizing Structural Biology (2001), http://www.bioz.unibas.ch/~xray/dino)

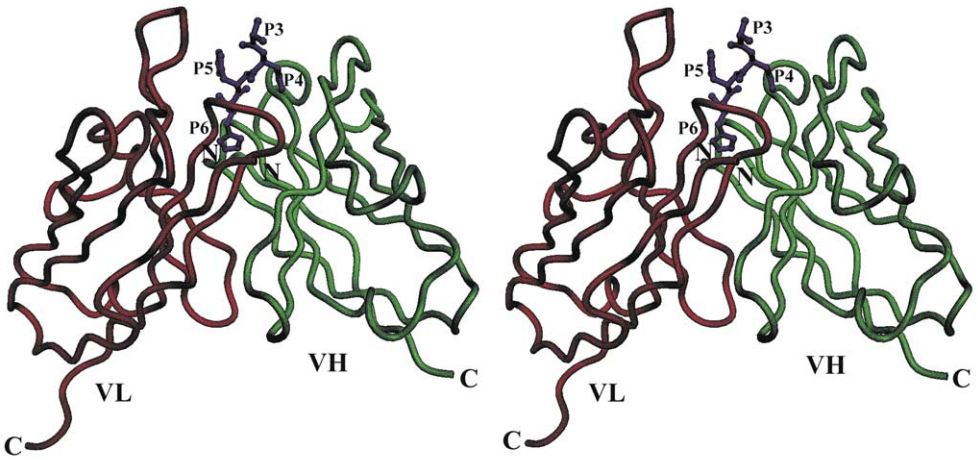

Figure 7. Stereo backbone representation of the anti-His scFv–peptide complex. The oligohistidine peptide (blue) binds in a deep pocket formed by the V_L (red) and V_H (green). Note that the scFv is of the orientation V_L–linker–V_H. The linker is invisible. Picture prepared with DINO.

The flexible 18 linker residues have to span the distance between the second linker residue and the N terminus of the V_H domain. In the structure of the anti-His scFv, the linker would have to span a distance of 18 Å in a diabody arrangement of two scFvs, or 42 Å within one scFv. Since the linker used here is clearly long enough to span the longer distance, the nature of the dimer observed in concentrated solution cannot be deduced from the crystal structure.

Structure description and details of interaction

The hexahistidine binds in a deep pocket formed by the heavy and the light chain of the scFv similar as in hapten-binding antibodies (Figure 7; see footnote on p. 140). Hydrophobic interactions are observed mainly between aromatic side-chains of the antibody and the imidazole ring of the antigen (Figure 8). The imidazole rings of His P4, P5 and P6 interact with the side-chain atoms of Tyr H40, Tyr L40, His L33, Phe L107 and Asn H42. Additional hydrophobic interactions are observed between the aromatic ring of Phe L137 and main-chain atoms of His P5 and P6, and between the imidazole ring of His P6 and main-chain atoms of Gly L109 and Ser H110.

Acidic side-chains in the binding cleft of the scFv provide a strong negative surface potential (Figure 9). They act as hydrogen bond acceptors for the nitrogen-protons of the antigen imidazole side-chains. Hydrogen bonds between the antibody and the antigen and between the antigen and structured water molecules are visualized in Figure 8.

Hydrogen bond acceptors are the carboxyl-group of Asp H57, Glu L42, Glu H107 and the carboxyl group of His P6. There is one hydrogen bond from the main-chain amide group of His P6 to the carboxyl group of Gly L109. A single hydrogen bond from the side-chain imidazole ring of His L33 to the amide nitrogen atom of His P3 stabilizes the antigen main chain.

The positively charged imidazole ring of His P6 is kept firmly in place by the carboxyl groups of Glu H107 and of Glu L42, which accept a split hydrogen bond each from the side-chain protons of His P6 (Figure 8). These interactions could be described as shared salt-bridges. There is no indication that His P4 and His P5 are protonated.

Optimal binding of scFv to the antigen has been observed between pH 6.0 and 7.4.[15] This pH-dependence of antigen binding is consistent with the hydrogen bonding network and with the presence of a salt-bridge. Outside this pH-range, binding is weak. At lower pH-values, aspartate and glutamate side-chains are protonated and cannot act as hydrogen bond acceptors, while at pH values above 7.4, the histidine side-chains are deprotonated, not allowing the stabilizing salt-bridge to be formed. The pH optimum offers the possibility of acid and base elution in immunoaffinity chromatography.

Main-chain and side-chain atoms of His P1 and P2, and the side-chain atoms of His P3 do not interact with the scFv at all. On the basis of the present structure, we deduce that a tetrapeptide containing three C-terminal histidine residues would be sufficient for high-affinity binding to the anti-His scFv. Thus, proteins containing a C-terminal His tag of just three histidine residues should be detectable by the anti-His scFv, which was verified experimentally (see below), or purified using a matrix coupled with the anti-His scFv.

The specificity for C-terminal His tags is explained by the crystal structure. The C-terminal carboxyl group is stabilized by a hydrogen bond donated by Ser H110 and by an indirect hydrogen bond network involving Tyr H40 and the water molecules W141 and W138 as well as Asp H57, and the water molecules W139 and W138 (Figure 4). The structure shows that His P6 is the

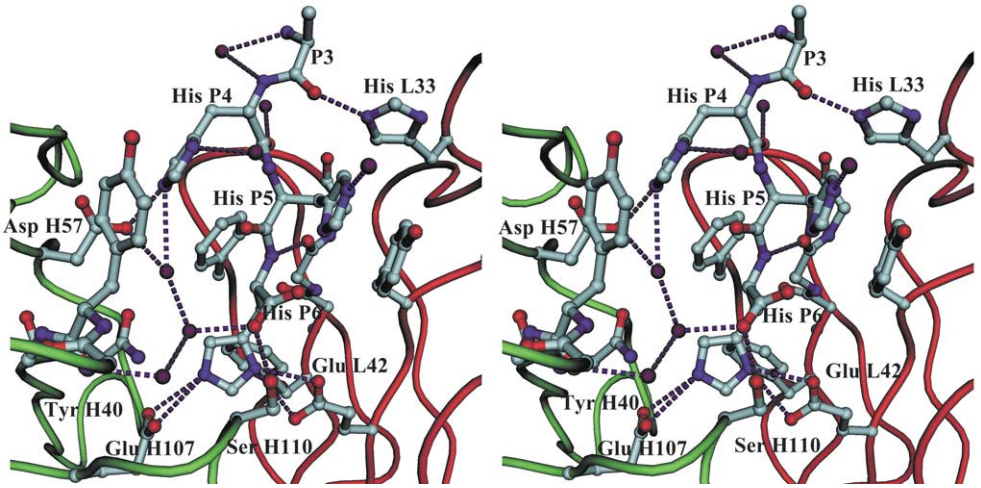

Figure 8. Stereo picture of the antigen-binding site. The backbone of the V_L is shown in red, the backbone of V_H is shown in green. Side chains involved in antigen binding and the oligohistidine peptide are shown and colored by atom type: C (gray), N (blue), O (red). Magenta spheres represent water molecules. Hydrogen bonds calculated with ACT[46] and LIGPLOT[47] are shown as dashed lines. Residues involved in hydrogen bonds are labeled with their residue name. The picture was prepared in DINO.

C-terminal residue, because there is no electron density further down in the cavity, even though there might be space for at least one additional residue in this direction.

Western blot experiments of four *E. coli* crude extracts, each containing an scFv fragment[34,35] with a different His tag as the protein to be detected, support our structural finding. The anti-His scFv fused to alkaline phosphatase[15] was used to test detection of the different His-tagged scFv molecules. Constructs containing three, four or six histidine residues at the C terminus or six histidine residues followed by glycine as the C-terminal residue were tested. Six, four and three histidine residues are recognized equally well by the anti-His scFv. In contrast, no binding was observed for the molecule with a C-terminal glycine residue after the His tag (Figure 10). The anti-His scFv thus recognizes three C-terminal histidine residues of a protein, and the last residue must be histidine.

A homology model of the Fv fragment of wild-type anti-His antibody 3D5 in complex with a tetrahistidine peptide has been published.[15] Some mutations, introduced into mut1+2 anti-His scFv based upon modeling, that had increased stability and solubility, can now be discussed in terms of the experimental structure. His H48 to Pro and probably Ser H51 to Gly caused a conformational change of the loop between H47 and H51, compared to the wild-type. This conclusion can be drawn from a model of the wild-type anti-His scFv and the structure of an anti-melanoma-associated GD2 ganglioside Fab (PDB entry 1PSK[36]), which has the same amino acid sequence in this loop as the wild-type. Mutations Glu H100 to Asp and Lys H77 to Arg were intended to stabilize V_H. A doubly hydrogen bonded salt-bridge between these two residues proves this stabilizing effect. In the wild-type, there was only a single hydrogen bonded Lys–Glu salt-bridge.

Conclusions and Implications

The structure described here shows the details of His tag recognition by anti-His scFv. The peptide binds in a pocket formed by the heavy and the light chain of the variable fragment, which is simi-

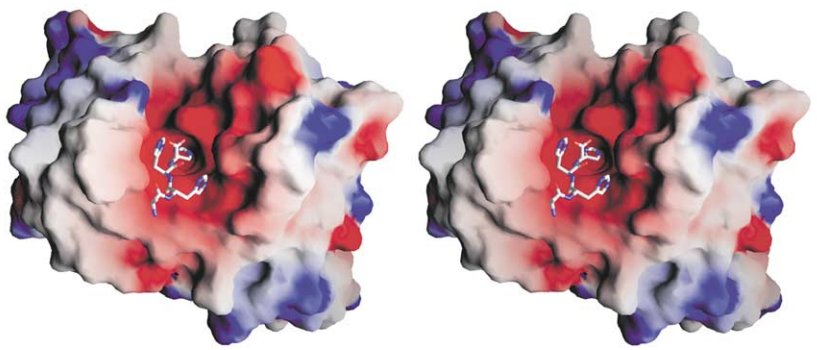

Figure 9. The peptide binds in a deep pocket with a strong negative surface potential. Surface potential of the antigen-binding site calculated and displayed with GRASP.[31] Negative potential is colored in red; positive potential is colored in blue. The peptide is colored by atom type C (white), N (blue), O (red).

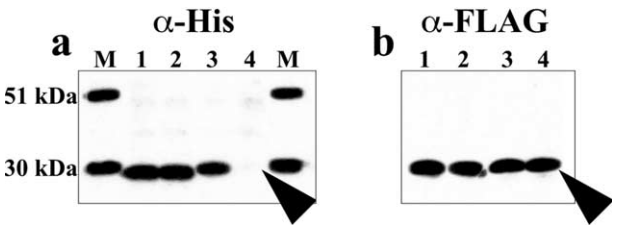

Figure 10. Western blots of four *E. coli* crude extracts each containing a differently his-tagged scFv of antibody 4D5.[35] (M) Two marker proteins for molecular mass reference; (1) scFv–$(His)_3$; (2) scFv–$(His)_4$; (3) scFv–$(His)_6$; (4) scFv–$(His)_6$Gly. Detection was performed using (a) anti-His scFv–phoA fusion or (b) anti-FLAG tag mAb M1 (Sigma). Even three histidine residues are recognized by the 3D5 anti-His tag scFv (lane a1). No binding of the anti-His tag antibody, however, is observed if the histidine residues are followed by a glycine towards the C-terminus (arrow in lane a4). All four constructs also carry the N-terminal sequence DYKD that is recognized by the anti-FLAG antibody M1. This is to demonstrate that comparable protein amounts of scFv were loaded in lanes 1–4.

lar to hapten-binding antibodies. The precise knowledge of the interactions of the anti-His scFv with the His tag opens the possibility to design mutations in the binding site of the anti-His scFv to increase affinity and/or alter specificity. Together, structure determination and Western blot analysis prove that the anti-His scFv exclusively and specifically recognizes three C-terminal histidine residues. Thus, a very short affinity tag of only three histidine residues is sufficient for detection and purification by the anti-His scFv. Proteins with such a "mini-tail" should have a better chance to crystallize, because flexible tails can prevent the formation of crystals.

Anti-His scFv crystals have a very high solvent content and possess channels of up to 70 Å in diameter. This would allow diffusion of peptides and even small proteins into preformed crystals. Crystals of the same form of uncomplexed anti-His scFv might conceivably be used for soaking of small proteins into such crystals. As the antigen-binding site of anti-His scFv is exposed to the channel, proteins with a C-terminal His tag could potentially bind to the anti-His scFv in the crystal. Anti-His scFv crystals could act as framework for the crystallization of His-tagged target proteins that are difficult to crystallize. If all the His-tagged protein molecules adopted the same orientation in the crystal, such protein-scFv crystals could be used for crystal structure determination. For small soaked proteins, initial phases could even be calculated from the anti-His scFv model.

The anti-His scFv can now be produced as a soluble protein in large amounts. This is a requirement for the low-cost production of an immunoaffinity matrix. Conveniently, the anti-His scFv can be linked to a gel matrix directly from a crude *E. coli* extract, for immunoaffinity purification of His-tagged proteins.[6] To purify a protein efficiently, the His tag can be exploited twice in two subsequent and inexpensive purification steps (IMAC and anti-His scFv immunoaffinity chromatography). This can provide highly purified protein that can be used for crystallization, biophysical studies and for the production of therapeutic proteins in clinical applications.

Materials and Methods

Chemicals and enzymes

Except where indicated otherwise, chemicals were purchased from Fluka (Buchs, Switzerland). Restriction enzymes were from New England Biolabs (*via* Bio Concept, Allschwil, Switzerland), Fermentas (*via* Lab Force, Nunningen, Switzerland) or Roche Diagnostics (Rotkreuz, Switzerland).

Site-directed mutagenesis

The mutagenesis was performed with standard protocols using PCR. The two sets of mutations (mut1 and mut2) were combined to mut1+2 by cloning.

Construction of His-tagged scFv 4D5

The scFv of the anti HER2 antibody 4D5[34,35] was PCR-cloned into pAK400[37] and different histidine tails were introduced by PCR-mutagenesis using appropriate primers for testing the specificity of the anti-His tag scFv. Expression was performed at 25 °C in a total volume of 30 ml as described below for the anti-His scFvs.

Expression of scFv constructs and Western blot analyses

The mutant anti-His scFv mut1, mut2 and mut1+2 were expressed from plasmid pAK100His2[15] for periplasmic expression. To compare the expression levels, the resulting plasmids (pAK1Hmut1, pAK1Hmut2 and pAK1Hmut1+2) and the plasmid encoding the wild-type were transformed in the *E. coli* K12 strain SB536 (F^-, WG1, Δ*fhuA* (*tonA*), Δ*hhoAB* (SacII), *shh*).[38] Precultures were inoculated from single colonies and grown at 25 or 37 °C in 5 ml of Miller's LB-medium (Life Technologies) containing 30 μg/ml of chloramphenicol. Main cultures (25 ml) were inoculated to $A_{550} = 0.1$. The cultures were grown at 25 or 37 °C. Expression was induced with 1 mM IPTG (BIOSOLVE, Netherlands) at $A_{550} = 0.5$. Cells from 20 ml cultures were harvested by centrifugation 2.5–3.5 hours after induction. Cell pellets were resuspended in TBS buffer (20 mM Tris-HCl (pH 7.0), 500 mM NaCl, 0.1 mM EDTA), normalized to their end A_{550} using 1.0 ml of TBS buffer per A_{550} unit. Whole-cell extracts were prepared by French press lysis at 10,000 psi (1 psi = 6.9 kPa) and 1 ml of the resulting crude extract was centrifuged in an Eppendorf tube for 30 minutes in a benchtop centrifuge at maximum speed at 4 °C. The supernatants containing the soluble fractions were transferred to a new tube and the pellets were resuspended in 1 ml of TBS buffer. The normalized samples (soluble and insoluble fractions) were mixed

with SDS-PAGE loading buffer and boiled until the pellet dissolved completely. SDS-PAGE was performed under reducing conditions using 12% (w/v) polyacrylamide gels. Western blots with the monoclonal anti-FLAG® antibody M1 (Sigma, Buchs, Switzerland) were carried out as described.[39,40] Anti-His tag Western blots were performed using the anti-His tag scFv–phoA fusion as described.[15]

Large-scale expression and purification of the anti-His scFv

For crystallization of variant mut1+2, the *myc*-tag at the C terminus of the scFv fragment encoded by the pAK100 derivative was removed by PCR, while the N-terminal FLAG-tag was retained for detection of the scFv-fragment during the purification steps. The resulting plasmid (pAK1Hmut1+2_noM) was transformed in the *E. coli* strain SB536. Precultures were grown in 200 ml of SB-medium (20 g/l of Tryptone, 10 g/l of yeast extract, 5 g/l of NaCl, 50 mM K_2HPO_4) containing 30 μg/ml of chloramphenicol at 25 °C overnight. Main cultures (3 × 1 l of SB) were inoculated to $A_{550} = 0.1$ and grown in 5 l baffled shake flasks at 25 °C. Expression was induced with 1 mM IPTG at $A_{550} = 1.5$. Cells were harvested five hours after induction by centrifugation. The cell pellet was resuspended in buffer A (20 mM Mes/NaOH, pH 6.5) using 1 ml of buffer per gram of cells. DNase I (Roche Diagnostics, Rotkreuz, Switzerland) was added and cell disruption was achieved by French press lysis at 10,000 psi. After centrifugation (Sorvall SS34, 20,000 rpm, 4 °C, 30 minutes) the crude extract was dialyzed against buffer A, centrifuged for 60 minutes and then passed through a 0.22 μm filter. After equilibration with buffer A, the crude extract was loaded onto an 8 ml SP-Sepharose column (Amersham Pharmacia Biotech, Dübendorf, Switzerland) at a flow-rate of 0.5 ml/minute. The column was washed (1 ml/minute) with buffer A until the absorption reached the baseline and elution was achieved at a flow-rate of 0.5 ml/minute using the following gradient: NaCl concentration was increased from 0 mM (buffer A) to 500 mM (buffer B consists of 20 mM Mes/NaOH (pH 6.5), 500 mM NaCl) in a volume of 150 ml. During elution, 5 ml fractions were collected. The fractions containing the anti-His scFv were identified by Western blot analysis using the anti-FLAG antibody M1 (Sigma). These fractions were purified further using an affinity column. A Gly–Ser–Gly–Ser–Gly–$(His)_8$–peptide was coupled to an NHS-Sepharose column (Amersham Pharmacia Biotech, Dübendorf, Switzerland) *via* its N-terminal α-amino group exactly as described in the manufacturer's protocol. The column was equilibrated with buffer B and NaCl was added to the pooled fractions to a concentration of 500 mM. The pooled fractions were loaded onto the affinity column at a flow-rate of 0.25 ml/minute. After washing the column with buffer B at a flow-rate of 1 ml/minute, the scFv fragment was eluted at 0.5 ml/minute with the following elution buffer: 50 mM Caps/NaOH (pH 10), 300 mM imidazole, 500 mM NaCl. A combination of pH 10 and imidazole was needed for eluting the protein in a sharp peak. The fractions containing the scFv without any degradation products were identified by Western blot analysis.

Sample preparation for crystallization

The scFv fractions were pooled, dialyzed against buffer A and concentrated (cut-off 10 kDa, Ultrafree 4, Millipore, Volketswil, Switzerland) to a final concentration of 136 μM. The concentration was determined by measuring absorbance at 280 nm using the calculated extinction coefficient of 42,200 $M^{-1}\,cm^{-1}$. The correct mass of the purified scFv fragment was confirmed by mass spectrometry. Before crystallization, a $(His)_6$-peptide was added in a 1.5-fold molar excess. The final solution contained the scFv-peptide complex at a concentration of 3.8 mg/ml.

Size-exclusion chromatography

The anti-His scFv–His_6 complex was analyzed on a Superdex 75 (10/30) gel-filtration column (Amersham Pharmacia Biotech, Dübendorf, Switzerland) using a buffer containing 20 mM Mes (pH 6.4), 150 mM NaCl.

Crystallization

Anti-His scFv was crystallized by the sitting drop, vapor diffusion method at 4 °C. A 1 μl sample of reservoir solution containing 0.1 M Mes (pH 6.4), 0.2 M magnesium acetate, 20% (w/v) PEG 8000, 0.02% NaN_3 (w/v), was mixed with 1–5 μl of protein solution. Crystals of hexagonal shape grew within two weeks to a maximal size of 70 μm. The largest crystals grew when the reservoir to protein ratio was 1:5. Crystals from these last experiments were used for data collection.

Data collection and processing

Data collection was performed at 100 K using a cryoprotectant consisting of 85.5% (v/v) reservoir solution and 12.5% (v/v) ethylene glycol. Diffraction data were collected on a MarResearch imaging plate system at the Swiss Norwegian Beam Line (SNBL) at the European Synchrotron Radiation Facility (ESRF) in Grenoble (France) from a single crystal. Data were processed with DENZO and SCALEPACK.[41] The space group is $P3_221$ with cell dimensions $a = b = 106.51$ Å, $c = 92.80$ Å. The asymmetric unit contains one scFv–peptide complex. The Matthews coefficient V_M is 5.4 $Å^3$/Da. A total of 128,480 measured reflections with $I > 2.5\ \sigma(I)$ were accepted, where I is the intensity and σ the standard deviation. Merging of symmetry-related reflections resulted in 16,996 unique reflections and a multiplicity of 7.6 with $R_{merge} = 11.6\%$. The completeness of the data to 2.7 Å was 99.0%.

Structure solution

The structure was solved by molecular replacement using the program AMoRe.[42] The coordinates of the V_L domain of the anti-cholera toxin peptide 3 Fab (PDB entry 1TET[16]) and the coordinates of the V_H domain of the anti-melanoma-associated GD2 ganglioside Fab (PDB entry 1PSK[36]) served as initial templates to build the search models. A total of 25 non-conserved residues were replaced by alanine or glycine. The templates were selected because of their close similarity to the V_L (94% identity) or V_H (85% identity) of the anti-His scFv, and they are identical in length. A conventional molecular replacement protocol (rotation, translation, rigid body fitting) using an integration radius of 30 Å for the

rotation search was performed for the V_L and V_H independently to deduce the correct relative orientation of V_L and V_H. An unambiguous solution was found in both cases. The solutions determined separately for V_L and V_H were combined to provide the initial model of the anti-His scFv.

Model building and structure refinement

Model building was carried out in O.[43,44] The anti-His scFv model was refined with CNS[45] in the resolution range 30–2.7 Å. To create an R_{free} test set, 5.8% of the reflections were selected randomly. Torsion angle molecular dynamics, simulated annealing using a slow-cooling protocol and a maximum likelihood target function, energy minimization and *B*-factor refinement were used in the refinement. Positions of water molecules were identified with CNS and were checked manually. The *R*-factor of the final model was 19.04% and R_{free}-factor was 22.20%. Structure validation was performed with the programs PROCHECK[28] and WHAT_CHECK.[29] Crystallographic statistics are given in Table 1.

Protein Data Bank accession number

The atomic coordinates and the structure factors have been deposited at the RCSB Protein Data Bank (http://www.rcsb.org/pdb) with entry code 1kTR.

Acknowledgments

We thank Christophe Briand for help in data collection, Peter Gehrig for mass spectrometry analysis and Stefan Klauser for peptide synthesis. We are grateful to the staff of the Swiss-Norwegian Beam Line at the European Synchrotron Radiation Facility for providing their facilities for data collection and for technical assistance. The financial support to M.G.G. from the Baugartenstiftung (CH-8022 Zürich) is gratefully acknowledged.

References

1. Nygren, P. A., Ståhl, S. & Uhlén, M. (1994). Engineering proteins to facilitate bioprocessing. *Trends Biotechnol.* **12**, 184–188.
2. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, **258**, 598–599.
3. Porath, J. & Olin, B. (1983). Immobilized metal ion affinity adsorption and immobilized metal ion affinity chromatography of biomaterials. Serum protein affinities for gel-immobilized iron and nickel ions. *Biochemistry*, **22**, 1621–1630.
4. Hochuli, E., Döbeli, H. & Schacher, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatog.* **411**, 177–184.
5. Müller, K. M., Arndt, K. M., Bauer, K. & Plückthun, A. (1998). Tandem immobilized metal-ion affinity chromatography/immunoaffinity purification of His-tagged proteins: evaluation of two anti-His-tag monoclonal antibodies. *Anal. Biochem.* **259**, 54–61.
6. Blank, K., Lindner, P., Diefenbach, B. & Plückthun, A. (2002). Self-immobilizing recombinant antibody fragments for immunoaffinity chromatography: generic, parallel and scaleable protein purification. *Protein Expt. Purif.*, **24**, 313–322.
7. Briand, C., Kozlov, S. V., Sonderegger, P. & Grütter, M. G. (2001). Crystal structure of neuroserpin: a neuronal serpin involved in a conformational disease. *FEBS Letters*, **505**, 18–22.
8. Liu, S., Neidhardt, E. A., Grossman, T. H., Ocain, T. & Clardy, J. (2000). Structures of human dihydroorotate dehydrogenase in complex with antiproliferative agents. *Struct. Fold. Des.* **8**, 25–33.
9. Knapp, J. E., Carroll, D., Lawson, J. E., Ernst, S. R., Reed, L. J. & Hackert, M. L. (2000). Expression, purification, and structural analysis of the trimeric form of the catalytic domain of the *Escherichia coli* dihydrolipoamide succinyltransferase. *Protein Sci.* **9**, 37–48.
10. Eijkelenboom, A. P., Lutzke, R. A., Boelens, R., Plasterk, R. H., Kaptein, R. & Hard, K. (1995). The DNA-binding domain of HIV-1 integrase has an SH3-like fold. *Nature Struct. Biol.* **2**, 807–810.
11. Casey, J. L., Keep, P. A., Chester, K. A., Robson, L., Hawkins, R. E. & Begent, R. H. (1995). Purification of bacterially expressed single chain Fv antibodies for clinical applications using metal chelate chromatography. *J. Immunol. Methods*, **179**, 105–116.
12. Kaslow, D. C. & Shiloach, J. (1994). Production, purification and immunogenicity of a malaria transmission-blocking vaccine candidate: TBV25H expressed in yeast and purified using nickel-NTA agarose. *BioTechnology*, **12**, 494–499.
13. Zentgraf, H., Frey, M., Schwinn, S., Tessmer, C., Willemann, B., Samstag, Y. & Velhagen, I. (1995). Detection of histidine-tagged fusion proteins by using a high-specific mouse monoclonal anti-histidine tag antibody. *Nucl. Acids Res.* **23**, 3347–3348.
14. Pogge von Strandmann, E., Zoidl, C., Nakhei, H., Holewa, B., Pogge von Strandmann, R., Lorenz, P. *et al.* (1995). A highly specific and sensitive monoclonal antibody detecting histidine-tagged recombinant proteins. *Protein Eng.* **8**, 733–735.
15. Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C. *et al.* (1997). Specific detection of his-tagged proteins with recombinant anti-His tag scFv–phosphatase or scFv–phage fusions. *BioTechniques*, **22**, 140–149.
16. Shoham, M. (1993). Crystal structure of an anticholera toxin peptide complex at 2.3 Å. *J. Mol. Biol.* **232**, 1169–1175.
17. Jeffrey, P. D., Strong, R. K., Sieker, L. C., Chang, C. Y., Campbell, R. L., Petsko, G. A. *et al.* (1993). 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc. Natl Acad. Sci. USA*, **90**, 10310–10314.
18. Rini, J. M., Stanfield, R. L., Stura, E. A., Salinas, P. A., Profy, A. T. & Wilson, I. A. (1993). Crystal structure of a human immunodeficiency virus type 1 neutralizing antibody, 50.1, in complex with its V3 loop peptide antigen. *Proc. Natl Acad. Sci. USA*, **90**, 6325–6339.
19. Honegger, A. & Plückthun, A. (2001). Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J. Mol. Biol.* **309**, 657–670.
20. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesmann, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, NIH Publication No. 91-3242, 5th edit., U.S. Department of Health and Human Services.

21. Proba, K., Wörn, A., Honegger, A. & Plückthun, A. (1998). Antibody scFv fragments without disulfide bonds made by molecular evolution. *J. Mol. Biol.* **275**, 245–253.
22. Forsberg, G., Forsgren, M., Jaki, M., Norin, M., Sterky, C., Enhorning, A. *et al*. (1997). Identification of framework residues in a secreted recombinant antibody fragment that control production level and localization in *Escherichia coli*. *J. Biol. Chem.* **272**, 12430–12436.
23. Chan, W., Helms, L. R., Brooks, I., Lee, G., Ngola, S., McNulty, D. *et al*. (1996). Mutational effects on inclusion body formation in the periplasmic expression of the immunoglobulin VL domain REI. *Fold Des.* **1**, 77–89.
24. Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng.* **10**, 435–444.
25. Hudson, P. J. & Kortt, A. A. (1999). High avidity scFv multimers; diabodies and triabodies. *J. Immunol. Methods*, **231**, 177–189.
26. Kortt, A. A., Malby, R. L., Caldwell, J. B., Gruen, L. C., Ivancic, N., Lawrence, M. C. *et al.* (1994). Recombinant anti-sialidase single-chain variable fragment antibody. Characterization, formation of dimer and higher-molecular-mass multimers and the solution of the crystal structure of the single-chain variable fragment/sialidase complex. *Eur. J. Biochem.* **221**, 151–157.
27. Arndt, K. M., Müller, K. M. & Plückthun, A. (1998). Factors influencing the dimer to monomer transition of an antibody single-chain Fv fragment. *Biochemistry*, **37**, 12918–12926.
28. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
29. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1997). Errors in protein structures. *Nature*, **381**, 272.
30. Janin, J. (1997). Specific *versus* non-specific contacts in protein crystals. *Nature Struct. Biol.* **4**, 973–974.
31. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
32. Stanfield, R. L. & Wilson, I. A. (1994). Antigen-induced conformational changes in antibodies: a problem for structural prediction and design. *Trends Biotechnol.* **12**, 275–279.
33. Plückthun, A. & Pack, P. (1997). New protein engineering approaches to multivalent and bispecific antibody fragments. *Immunotechnology,* **3**, 83–105.
34. Eigenbrot, C., Randal, M., Presta, L., Carter, P. & Kossiakoff, A. A. (1993). X-ray structures of the antigen-binding domains from three variants of humanized anti-p185HER2 antibody 4D5 and comparison with molecular modeling. *J. Mol. Biol.* **229**, 969–995.
35. Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D. *et al*. (1992). High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. *Biotechnology*, **10**, 163–167.
36. Pichla, S. L., Murali, R. & Burnett, R. M. (1997). The crystal structure of a Fab fragment to the melanoma-associated GD2 ganglioside. *J. Struct. Biol.* **119**, 6–16.
37. Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R. & Plückthun, A. (1997). Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods*, **201**, 35–55.
38. Bass, S., Gu, Q. & Christen, A. (1996). Multicopy suppressors of prc mutant *Escherichia coli* include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. *J. Bacteriol.* **178**, 1154–1161.
39. Knappik, A. & Plückthun, A. (1994). An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. *BioTechniques*, **17**, 754–761.
40. Ge, L., Knappik, A., Pack, P., Freund, C. & Plückthun, A. (1995). Expression of antibodies in *Escherichia coli*. In *Antibody Engineering* (Borrebaeck, C. A. K., ed.), pp. 229–266, Oxford University Press, New York.
41. Otwinowski, Z. & Wladek, M. (1996). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
42. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
43. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. (1989). In *O: A Macromolecular Modeling Environment: An Overview. Crystallographic Computing and Modeling Methods in Macromolecular Design* (Buggs, C. & Ealick, S., eds), Springer, New York.
44. Jones, T. A. & Kjeldgaard, M. (1991). *Manual for O.*, 6th edition, Uppsala University, Uppsala, Sweden.
45. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect D*, **54**, 905–921.
46. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
47. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134.

Edited by R. Huber

(Received 10 October 2001; received in revised form 28 January 2002; accepted 31 January 2002)