Protein Engineering vol.13 no.5 pp.353–360, 2000

Design and synthesis of germline-based hemi-humanized single-chain Fv against the CD18 surface antigen

Cristina Caldas[1], Verônica P.C.V.Coelho[2], Daniel J.Rigden[3], Goran Neschich[3], Ana Maria Moro[4] and Marcelo M.Brígido[1,5]

[1]Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, 70910-900, [2]Laboratório de Imunologia de Transplantes, INCOR/ Universidade de São Paulo, São Paulo, SP, 05403-000, [3]EMBRAPA-CENARGEN, Brasília, DF, 70770-900 and [4]Instituto Butantan, São Paulo, SP, 05503-900, Brazil

[5]To whom correspondence should be addressed

The 6.7 murine monoclonal antibody (mAb) recognizes the human CD18 antigen and is therefore of interest as an anti-inflammatory agent. The 6.7 heavy variable chain (VH) was humanized using the closest human germline sequence as the template on to which to graft the murine complementary determining regions (CDRs). Two versions were proposed, one in which the residue proline 45 of the murine form was maintained and another in which this framework residue was changed to the leucine found in the human sequence. These VH humanized versions were expressed in the yeast *Pichia pastoris* as hemi-humanized single-chain Fv (scFvs), with the VL from the murine antibody. The scFv from the murine antibody was also expressed. The binding activities of the murine and both hemi-humanized scFvs were determined by flow cytometry analysis. All the constructions were able to recognize human lymphocytes harboring CD18, indicating successful humanization with transfer of the original binding capability. Some differences between the two hemi-humanized versions were observed. The method used was simple and straightforward, with no need for refined structural analyses and could be used for the humanization of other antibodies.

Keywords: CD18/humanization/monoclonal antibody/*Pichia pastoris*/scFv

Introduction

Owing to the capacity that antibody molecules possess to recognize specifically a vast number of different antigens, they have long been proposed and used as therapeutic agents (reviewed by Gronski *et al.*, 1991). With the development of hybridoma technology (Köhler and Milstein, 1975), many monoclonal antibodies (mAbs) relevant to therapy were generated (Mountain and Adair, 1992). Because the majority of mAbs produced by this technology are of murine origin, there are limitations in their clinical application, mainly due to the HAMA (human anti-mouse antibody) response (Benjamin *et al.*, 1986; Brüggemann *et al.*, 1989). In this context, antibody humanization technology was developed (Jones *et al.*, 1986; Riechmann *et al.*, 1988; Verhoeyen *et al.*, 1988) for the production of a new molecule that maintains the capacity to recognize the antigen but looks more like a human antibody. This is done by the transfer of the murine CDRs (complementary determining regions) to a human framework. Many recent studies have shown the therapeutic effectiveness of humanized antibodies such as anti-TNF-α, CAMPATH-1 and anti-IL-2 receptor in patients (reviewed by Vaswani and Hamilton, 1998).

There are different strategies to humanize an antibody. The first humanized antibodies were constructed based on human sequences with known crystal structure (Riechmann *et al.*, 1988; Foote and Winter, 1992; Graziano *et al.*, 1995), which permits the identification of residues contributing to the antigen binding. In the best fit strategy, the closest human sequence is used as a framework to receive the murine CDRs (Queen *et al.*, 1989; Co *et al.*, 1991; Co *et al.*, 1996) and usually a rearranged sequence is chosen. By choosing a framework as similar as possible to the original antibody, the chances that the structural features of the framework will be preserved in the humanized antibody are maximized, leading to a correct spatial orientation of the CDRs. Another strategy to humanize antibodies is to choose the closest human germline sequence (Tomlinson *et al.*, 1992) as the framework to receive the murine CDRs. This germline approach is based on the same rationale as the best fit strategy, but only germline sequences are searched in the databases. Only a few antibodies have been humanized by this method (Rosok *et al.*, 1996; Johnson *et al.*, 1997), but it is a very promising strategy because a human germline sequence does not present somatic hypermutation that is potentially immunogenic. CDRs can also be grafted using the consensus strategy, in which one of the human subgroups is used as the framework (Presta *et al.*, 1993; Couto *et al.*, 1994, 1995; Werther *et al.*, 1996; O'Connor *et al.*, 1998). The resurfacing strategy was proposed by Padlan (1991) and involves the replacement of solvent exposed murine framework residues in the variable regions with human residues (Roguska *et al.*, 1994, 1996; Studnicka *et al.*, 1994).

The final purpose of this work is to humanize the 6.7 murine anti-human CD18 antibody (David *et al.*, 1991), which is of interest as an anti-inflammatory agent. In this stage, the objectives were to clone and sequence the heavy (VH) and light (VL) variable chains of the original (murine) antibody, followed by the chemical synthesis of two humanized VH versions using the closest human germline sequence. The murine and both hemi-humanized versions were cloned in a single-chain variable fragment (scFv) vector and expressed in the yeast *Pichia pastoris.* We were able to purify these scFv molecules and test them on $CD18^+$ cells by flow cytometry analysis. The results were compared with a molecular model generated for the hemi-humanized version.

Materials and methods

Cloning and sequencing of 6.7 variable regions

Hybridoma 6.7 was grown in DME medium (Gibco) containing 10% fetal calf serum (FCS) (Cultilab, Brazil). About 10^6 cells were spun down and washed with PBS (150 mM NaCl, 10 mM NaH_2PO_4, pH 7.2 and NaN_3, 0.002%). The mRNAs were isolated and the cDNAs that code for the VH and VL were synthesized using the Lysate mRNA Capture Kit for RT-PCR

© Oxford University Press

(Amersham), according to the manufacturer's instructions. For the cDNA synthesis, the primers used were κ18 (5′TACAGTTGGTGCAGCATC3′) for Cκ and γ1 (5′TGGACAGGGATCCAGAGTTCCAGGTCACT3′) for Cγ, 1.25 μM of each. For amplification of heavy and light chain V region cDNAs we used a library of sense primers previously described by Coloma *et al.* (1991) and Zhou *et al.* (1994) and the same anti-sense primers used for cDNA synthesis. Samples were subjected to 25 thermal cycles of 94°C for 1 min, 48 or 55°C (for VH and VL, respectively) for 1 min and 72°C for 1 min. Amplified VH and VL cDNAs were cloned in the pGEM-T vector (Promega) following the manufacturer's instructions. Three clones of each variable region were sequenced in both directions using a T7 Sequencing Kit (Pharmacia).

Construction of the VH humanized versions

For the design of the humanized VH, a search was carried out for the closest human germline sequence in the Swissprot, GenBank and EMBL databanks using the FASTA program (Wisconsin Package Version 9.1, Genetics Computer Group, GCG, Madison, WI). This sequence was used as a framework to accept the murine CDRs. Two versions were synthesized using eight overlapping oligonucleotides ranging from 42 to 69 bp, with 15 bp of complementarity. In each case, the eight oligonucleotides were assembled using a PCR-based protocol. Aliquots of each pair of complementary oligonucleotides (10 pmol) were annealed separately in a 50 μl reaction containing 9 mM Tris–HCl (pH 7.6), 13 mM $MgCl_2$, 21 mM DTT and 200 μM dNTPs. The samples were incubated in 400 ml of boiling water for 5 min and left standing until the water reached room temperature. Each pair of primers was elongated by the addition of 24 U DNA Polymerase I (Klenow fragment, Biolabs) for 30 min at room temperature. The two pairs of primers coding for the N-terminus were mixed and amplified by PCR. The two pairs for the carboxy terminus were also mixed in a separate tube and amplified under the same conditions. Samples were subjected to 20 thermal cycles of 94°C for 30 s, 60°C for 40 s and 72°C for 2 min and analyzed in an agarose gel. Finally, the full-length DNA fragment was amplified using as DNA templates the amplified fragments from the first PCR extracted directly from the agarose gel with the aid of a pipet tip. The 5′ and 3′ external primers were used in this second PCR and after 25 thermal cycles of 94°C for 30 s, 60°C for 40 s and 72°C for 2 min, the amplified fragments were cloned in the pGEM-T vector (Promega) and sequenced as above.

Expression of the murine and hemi-humanized scFvs in Pichia pastoris

The murine and hemi-humanized scFvs genes were assembled based on the bacterial expression vector pIg17Z22, previously described by Brígido (1992). This vector possesses the gene that codes for the Z22 scFv fusioned to a staphylococcal protein A domain. For the assembly of the murine anti-CD18 scFv cassette, the Z22 VH and VL genes were replaced by the anti-CD18 VH and VL genes after digestion with the restriction enzymes *Nde*I/*Xba*I and *Bgl*II (*Bcl*I)/*Nco*I, respectively, producing the vector pIgCD18. Restriction sites were created by PCR of the original clones with especially designed primers. In the light chain a *Bcl*I site was introduced instead of *Bgl*II due to an internal *Bgl*II site. VL was digested with *Bcl*I and *Nco*I and ligated to *Bgl*II/*Nco*I digested pIg17. For the assembly of the hemi-humanized scFv vectors, the murine VH of the pIgCD18 was replaced by the humanized versions L and P, originating the vector pIgCD18VHL and pIgCD18VHP. All the constructions maintained the original murine VL. After the construction of the three expression cassettes, these were transferred to the *Pichia pastoris* expression vector pPIg16, which is a pPIC9 derived vector (Invitrogen), into which the Z22 scFv gene has been cloned (Santos-Silva,M., Andrade,E.V., Albuquerque,F.C., Moraes,L.P. and Brígido,M.M., in preparation). The vectors pIgCD18, pIgCD18VHL and pIgCD18VHP were digested with the restriction enzymes *Xma*I and *Eco*RI and the desired fragments were purified and ligated to the linearized pPIg16 vector, producing the plasmids pPIgCD18, pPIgCD18VHL and pPIgCD18VHP, respectively.

Pichia pastoris GS115 cells (Invitrogen) were grown in liquid medium and made competent by resuspension in 1 M sorbitol. The cells were electroporated by pulse discharge (1500 V, 25 μF, 400 Ω; Bio-Rad Gene Pulser) for 5 ms in the presence of 5–10 μg of plasmid DNA linearized with *Sal*I. This enzyme cuts within the plasmid-encoded *HIS4* gene and favors homologous recombination with the endogenous, non-funcional *his4* gene of GS115 cells. Therefore, transformants (His+) were screened by their capacity to grow in the absence of histidine as described (Invitrogen). Several His+ transformants per construct were screened using the method described by Wung and Gascoigne (1996) for protein-secreting clones. Protein expression kinetics were determined by growing clones expressing the three scFvs in 25 ml of BMGY medium (1% yeast extract, 2% peptone, 10 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 4×10^{-5}% biotin, 1% glycerol) at 30°C in a shaking incubator (250 r.p.m.) until the culture reached A_{600} = 2.0–6.0. Cells were then centrifuged and resuspended in 100–200 ml of BMMY medium (1% yeast extract, 2% peptone, 10 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 4×10^{-5}% biotin, 0.5% methanol) to induce protein expression. Cells were incubated for 4 days at 30°C. Aliquots of culture supernatants were taken daily and examined by SDS–PAGE and Western blotting. For large-scale expression, the clones were grown in exactly the same way as above, for 80 h at 30°C under agitation. The supernatants were harvested following centrifugation and filtration through a 0.45 μm cellulose acetate filter. After the addition of 80 μg of Pepstatin A and 14 μg of PMSF to the supernatants, these were concentrated to about 5 ml using Centripep-10 (Amicon) according to manufacturer's instructions.

To purify the scFvs, the concentrated supernatants were run through an IgG Sepharose 6B Fast Flow column (Pharmacia) previously activated by three alternating washes with 0.5 M acetic acid, pH 3.4, and PBST (PBS and Tween 20, 0.1%) and finally equilibrated with PBS. scFv fragments were eluted with 0.5 M acetic acid, pH 3.4, and immediately neutralized with 1.5 M Tris–HCl, pH 8.8. The purified proteins were dialyzed against PBS and quantified using the BCA Protein Assay Kit (Pierce). The recombinant proteins produced were called Fv-Mu (murine anti-CD18 scFv), Fv-VH(L) (hemi-humanized scFv, with the humanized VH version L and murine VL) and Fv-VH(P) (hemi-humanized scFv, with the humanized VH version P and murine VL).

Binding to surface CD18 molecules: immunofluorescence and flow cytometry analysis

Immunofluorescence assays and flow cytometry analysis were performed to evaluate the binding capacity of the different recombinant anti-CD18 scFvs to surface CD18 molecules

on human lymphocytes. Peripheral blood mononuclear cells (PBMC) obtained from a normal individual by gradient centrifugation were used for immunofluorescence assays. Antibodies utilized were recombinant anti-CD18 scFvs [Fv-MU, Fv-VH(L) and Fv-VH(P)], recombinant Z22 scFv (negative control), rabbit anti-human IgG FITC (Dakopatts, Denmark; used in the second step of the indirect immunofluorescence to bind to the protein A domain of the recombinant anti-CD18 scFvs through the Fc fragment), rabbit anti-mouse IgG (Sigma), 6.7 FITC (Instituto Butantan-INCOR, Brazil), anti-CD19PE and anti-CD3PE (Dakopatts, Denmark). The sample incubated with both anti-CD19PE and rabbit anti-human IgG FITC was used to evaluate binding of the rabbit antibody to IgG expressed on B cells and exclude any other unspecific binding from the tests with the scFvs. Anti-CD3 was used as positive control of the assay. 2×10^5 cells were incubated with the different antibodies for 30 min at 4°C and washed three times. For the samples with the recombinant anti-CD18 scFvs, a second incubation was performed with rabbit anti-human IgG FITC. All samples were resuspended in 400 µl of FACS buffer (PBS, 2% FCS and 0.01% sodium azide) and analyzed using a FACScan flow cytometer (Becton Dickinson, CA, USA). Ten thousand events were analyzed for each sample, inside the gate of lymphocytes, according to FSC (forward scatter – cell size) and SSC (side scatter – cell granularity) parameters. Recombinant proteins were added in an equimolar quantity. Results are expressed as the percentage of stained cells. To analyze the binding specificity of the hemi-humanized scFvs, a second set of experiments was performed in which the capacity of the scFvs to block the binding of the original 6.7 FITC to surface CD18 molecules was tested. Cells were initially incubated with the different recombinant fragments, washed, incubated with rabbit anti-mouse IgG (to block the protein A domain) and then incubated with 6.7 FITC. The percentage of positive cells and the intensity of immunofluorescence (IF) were compared in samples with 6.7 FITC alone and samples with the different scFvs plus 6.7 FITC. The percentage of inhibition was calculated considering these differences, that is, (i) % positive cells with anti-CD18 FITC versus % positive cells with the different scFvs + anti-CD18 FITC) and (ii) median IF with anti-CD18 FITC versus median IF with the different scFvs + anti-CD18 FITC.

Structural modeling of the hemi-humanized scFv

Protein models were constructed using the MODELLER-4 package (Sali and Blundell, 1993). Given the high quality of the available template structures, a lower than default coordinate randomization value of 2 Å was used, prior to model refinement against derived restraints. Ten models were made and evaluated for each tested alignment. Several programs were used for the rigorous evaluation of the protein models. PROCHECK (Laskowski *et al.*, 1993) was used to monitor stereochemical quality while PROFILER_3D (Lüthy *et al.*, 1992) and PROSA II programs (Sippl, 1993) were used to measure overall protein quality in terms of packing and solvent exposure. The PROSA II output was the primary information used to guide the process of alignment adjustment. The best alignment was taken to be that which yielded the model having the lowest PROSA II *z*-score. However, care was taken to ensure that better overall scores resulted from improvements of regions in which an alteration to modeling strategy had been made, rather than by chance. The program O (Jones *et al.*, 1991) was used for inspection and manipulation of models. Three-dimensional superposition of proteins was carried out using LSQMAN (Kleywegt, 1996). CLUSTAL W (Thompson *et al.*, 1994) was used to make the initial sequence alignment. PSI-BLAST (Altschul *et al.*, 1997) was run at http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-psi_blast.

The molecular surface and interface between L and H chain was calculated and displayed using the GRASP package (Nicholls *et al.* 1991). The STING package was used to define Interface Forming Residues (IFR) (Neshich *et al.*, 1998).

Results

V region sequence analysis of murine antibody

The VH and VL cDNAs of the 6.7 anti-CD18 antibody were synthesized by PCR using primers specific for heavy and light chain V regions. The clones were sequenced and the sequences were deposited in GenBank with accession numbers AF134808 for VH and AF135165 for VL. Analysis of the VH nucleotide sequence showed that the murine VH of the 6.7 anti-CD18 belongs to the J558 family and uses the JH4 genic fragment. Two types of VL sequences were isolated. One of the VL sequences had the Cys 23 residue replaced with Tyr and also had a reading frameshift within CDR3. The same fact has been observed by others (Jang and Stollar, 1992; Tempest *et al.,* 1994) and originates from the fusion partner used in the production of the murine hybridomas. The other VL had a sequence compatible to productive VL chain and is a member of the Vκ1 family using the Jκ1 genic fragment. The deduced amino acid sequences of 6.7 VH and VL are shown in Figure 1.

Design of the two humanized VH versions

The closest human germline sequence found after the data bank search was HG3 (Rechavi *et al.*, 1983). We opted for the use of a germline sequence in this humanization process to avoid somatic hypermutation of rearranged V regions, that can be unique for that specific antibody and thus be seen as immunogenic to patients. The HG3 sequence was used as the human framework to accept the murine CDRs. We used an expanded CDR1, which includes both the Kabat (Kabat *et al.*, 1991) and the Chothia (Chothia *et al.*, 1989) CDR definitions. This means that the murine CDR1 that was grafted to the human framework included the H1 loop and neighboring residues (residues 32–35). The CDR3, which is derived from DJ recombination, was introduced and joined to the human germline JH4 genic fragment. Besides this grafting of CDRs, the framework residue Ala71 was maintained as in the original murine antibody. The murine residue Ala was maintained for the assembly of a correct H2 loop in the humanized antibody (Tramontano *et al.*, 1990). Analysis of the crystallized Fab with sequence most similar to HG3 (1AD9) showed that residue 45 is located in a position of low flexibility near to the H CDR2 and L CDR3. We therefore thought that this residue could exert an important role in the maintenance of the H2 and L3 loop structure. To test this hypothesis, two versions for humanized VH were constructed, one in which the proline residue 45 of the murine form was maintained (version P) and another in which this framework position was changed to the leucine found in the closest germline human sequence (version L). These two versions were synthesized using eight complementary oligonucleotides, which were annealed, elongated and amplified by PCR. The 5′ and 3′ fragments were recovered from the agarose gel with the aid of a pipet tip, then they were mixed and amplified with two external primers. About 12 clones obtained from the reaction

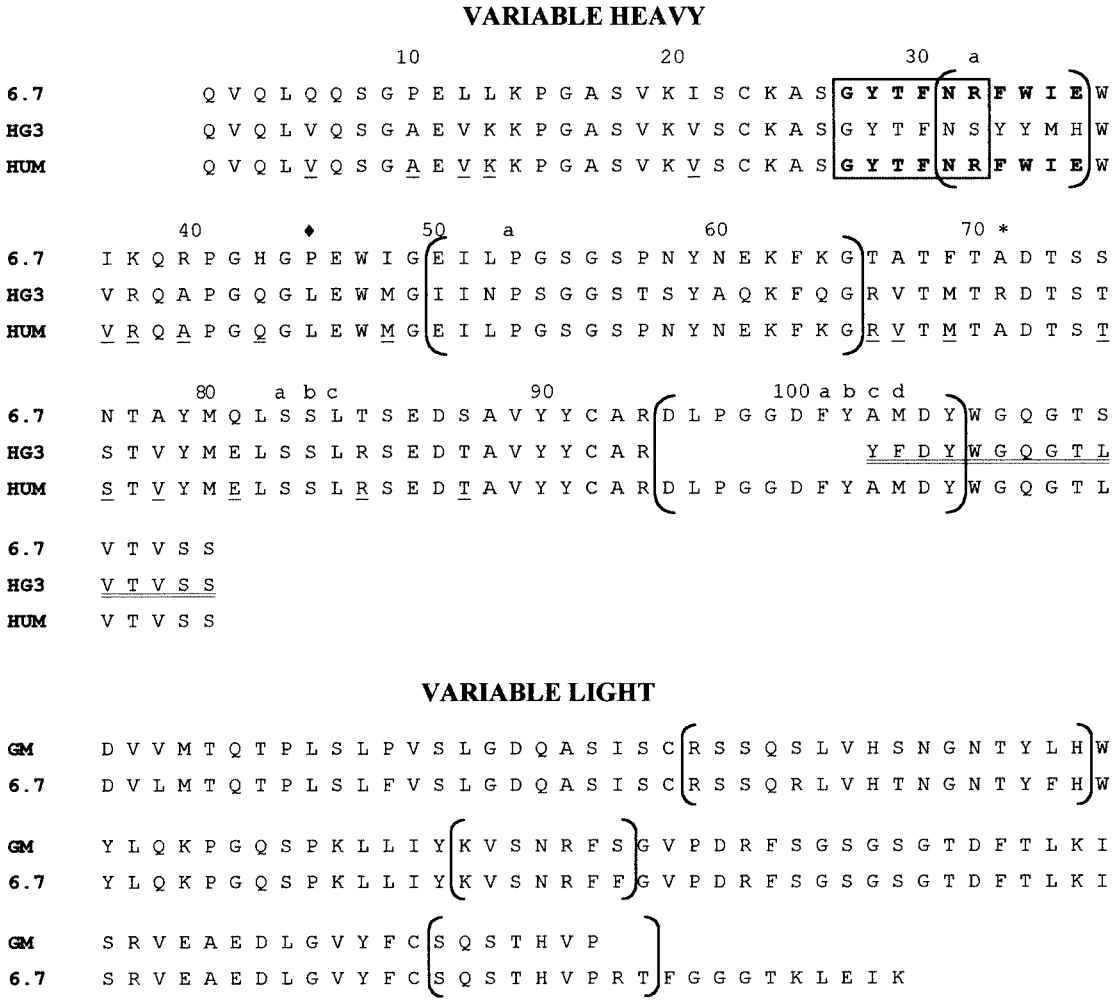

Fig. 1. Comparison of amino acid sequences of V regions of murine (6.7), closest human germline VH (HG3), humanized VH versions (HUM) and the putative murine germline VL (GM, Gen Bank D00080). CDR residues are bracketed. Definition of CDRs according to Kabat *et al.* (1991). Insertions are shown as a, b, c. The rectangle labels the H1 loop (Chothia *et al.*, 1989). The residues which differ in the humanized VH compared with the original antibody are underlined. The mutated residue 45 is labelled (◆). The residue alanine 71 was maintained in the humanized VH as in the murine antibody (*). The extended CDR1 used in the VH humanization is shown emboldened. Human genic fragment JH4 is double underlined.

had to be characterized since there were many partial products including basically N-1 products but also products missing long stretches of sequences. This was probably due to the use of non-purified oligonucleotides. The alignment of the amino acid sequences of the murine, HG3 and humanized VH versions is shown in Figure 1. In the proposed reshaped version, we end up with 19 residues which differ from the original murine sequence.

Expression and purification of the scFvs

The scFvs were expressed in the methylotrophic yeast *Pichia pastoris*, which is a methanol-inducible expression system (Cregg *et al.*, 1985). After the selection of the protein expressing clones, a large-scale preparation was done using the time of 80 h of cultivation determined by the kinetic induction. The kinetic experiment showed that in 96 h of induction more protein was expressed, but a greater proteolytic activity was observed. For this reason, we chose a shorter time in which less proteolytic degradation was observed. The supernatants could be affinity purified on an IgG Sepharose column owing to the presence of a protein A domain in all constructions. After the elution, the scFvs were quantified and analyzed by SDS–PAGE and Western blotting (data not shown). The yield of purified Fv-MU, Fv-VH(L) and Fv-VH(P) was 1.00, 0.98 and 0.91 mg/l, respectively.

Binding of scFvs to CD18$^+$ cells

The FACS analyses showed that 65, 84 and 79% of the cells were stained with Fv-MU, Fv-VH(L) and Fv-VH(P), respectively (Figure 2). In samples in which the cells were incubated only with the original monoclonal 6.7 FITC, 92% of the cells were stained. As shown in Figure 2, two sub-populations were observed in the cells incubated with the monoclonal 6.7. A third, smaller, sub-population with higher IF was observed in the cells stained with the hemi-humanized Fv-VH(L) and Fv-VH(P). This sub-population may be related to the cells stained with the rabbit anti-human IgG FITC binding directly to the IgG expressed on B cells and also to the protein A domain of the recombinant scFvs bound to the CD18. The binding specificity of the scFvs was confirmed in the blocking experiment. The Fv-VH(L) was able to block 76.3% of the CD18 and the Fv-VH(P) 69% (Figure 3). The loss of fluorescence observed in the samples incubated with the scFvs (Table I) confirmed that the hemi-humanized scFvs were able to bind specifically to the CD18, preventing the binding of the original monoclonal 6.7 FITC.

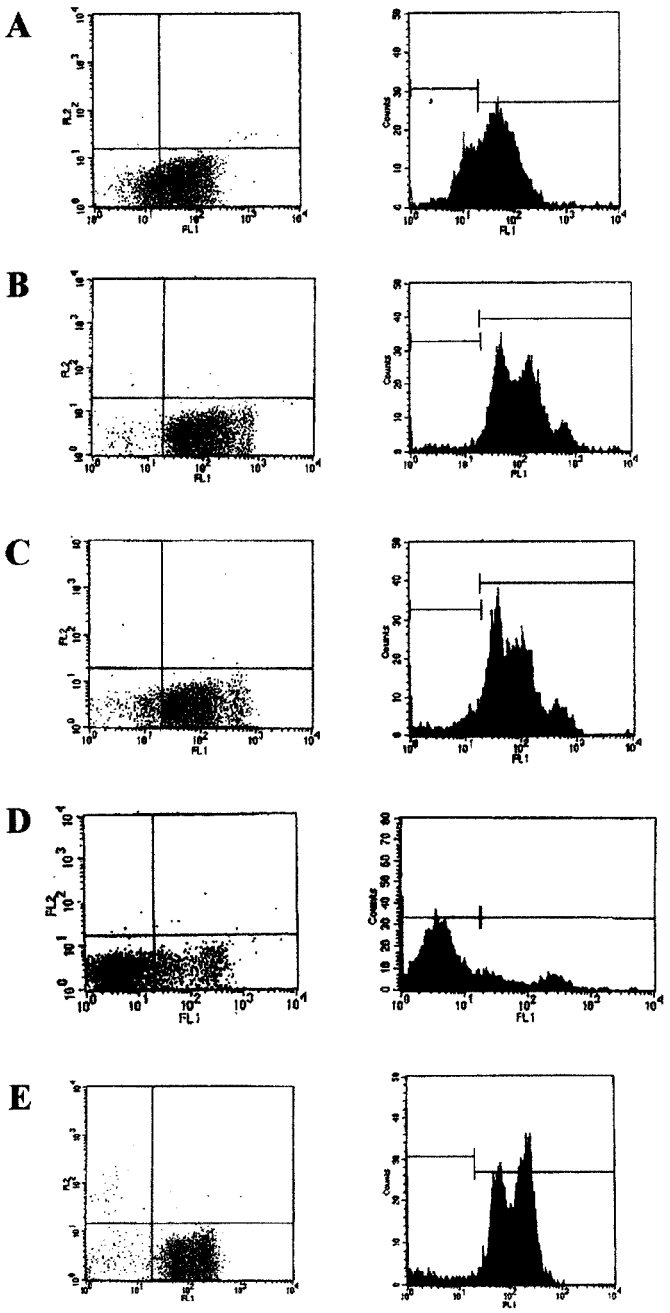

Fig. 2. Analysis of the binding of the scFvs to human lymphocyte surface. 65, 84 and 79% of the cells incubated with the Fv-MU (**A**), Fv-VH(L) (**B**) and Fv-VH(P) (**C**), respectively, were stained with the anti-human IgG FITC that bound to the protein A domain present in these recombinant proteins. The Z22 scFv was used as negative control (**D**). The FITC-conjugated original monoclonal antibody was used as positive control (**E**) and was able to stain 92% of the cells. FL1 refers to FITC fluorescence and FL2 was not analyzed in this experiment. Counts refers to cell number.

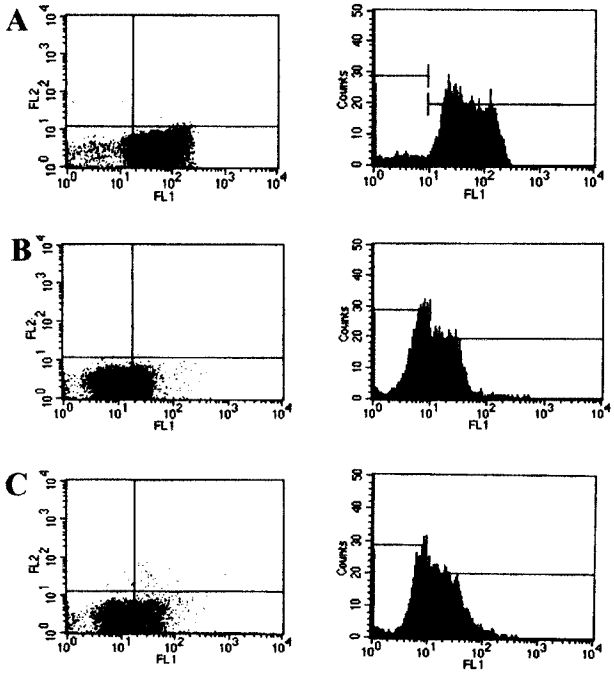

Fig. 3. Hemi-humanized scFvs fragments block the binding of monoclonal 6.7 FITC to $CD18^{+}$ cells. Cells incubated only with the 6.7 FITC (**A**). scFvs were added to the cells prior the addition of the 6.7 FITC. Fv-VH(L) (**B**) and FV-VH(P) (**C**) were able to block 76.3 and 69% of the CD18, respectively. FL1 refers to FITC fluorescence and FL2 was not analyzed in this experiment. Counts refers to cell number.

Table I. Blocking of 6.7 FITC binding to human lymphocytes by the hemi-humanized scFvs: median IF of lymphocytes pre-incubated with the scFvs followed by the addition of 6.7 FITC

	Median IF	Blocking (%)[a]
6.7 FITC	40.32	0
Fv-VH(L)	9.56	76.3
Fv-VH(P)	12.52	69.0

[a]The blocking percentage was calculated considering the median IF value for the 6.7.

Model construction and validation

For the choice of template we used the overall PSI-BLAST scores as a guide to the crystal structures with sequences most similar to the humanized antibody sequence. The top five scoring structures were used for modeling of each of the VH and VL. For VH 1AD9, 1AXS, 1MLB, 1CGS and 1VGE were used, with 66–76% sequence identity to the humanized antibody domain. In the case of VL we used 1MRC, 1MPA, 1AD9, 1A4J and 1RMF with 68–71% sequence identity. The 1AD9 structure was used to define the orientation between VH and VL. Neglecting the clearly structurally different CDR3 of the VH domain which required special treatment (see below), the templates show very similar structures. Using 1AD9 as a reference, the VL templates show Cα r.m.s. differences in the range 0.8–1.0 Å and the VH templates r.m.s. differences of 0.6–0.8 Å. These values compare with estimated coordinate errors, quoted in some of the template pdb files, of 0.3–0.5 Å. We only modeled the humanized antibody in regions for which structural templates existed; no attempt was made to model the Gly/Ala-rich linker region between the two domains.

In initial models CDR3 of the VH domain scored poorly by PROCHECK and PROSA II analysis owing to the radically different conformations adopted by CDR3 loops in the five templates. We therefore reduced the number of template structures used to construct the CDR3 region of the model. Examination of the alignment of humanized VH with the templates showed 1VGE and 1AD9 as equally favorable templates for this region. We therefore constructed two new sets of VH models using either 1AD9 or 1VGE as sole CDR3 template and all five structures as templates for the remainder

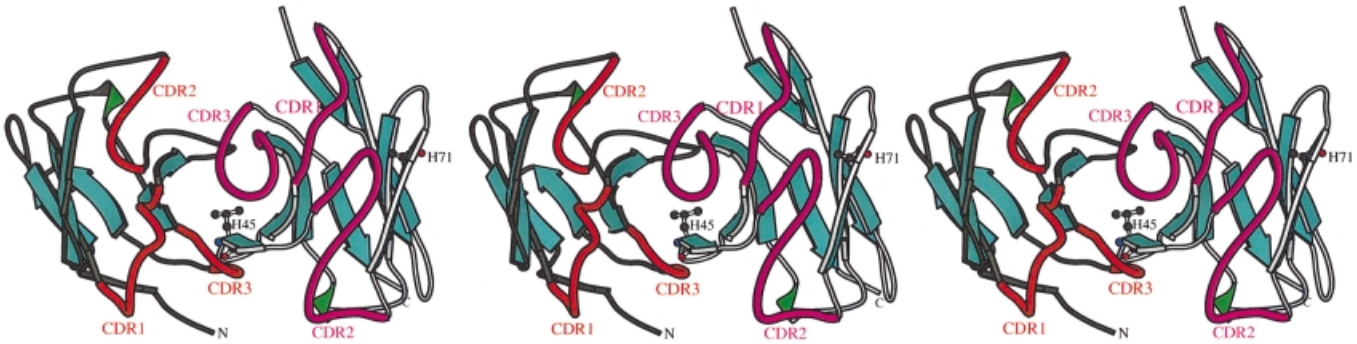

Fig. 4. A stereo Molscript (Kraulis, 1991) diagram of the final structure. The VH domain is on the right with CDRs drawn and labeled magenta. VL CDRs are drawn and labeled red. LeuH45 and Ala71 are shown in a ball-and-stick representation and labeled. The C-terminus of the VH domain and the N-terminus of the VL domain, joined by the unmodeled linker region, are labeled C and N respectively.

of the molecule. Both new sets of models improved on the original set by overall PROSA II, PROFILER_3D and PROCHECK analysis but the 1VGE-derived models had a greater number of residues in disallowed areas of the Ramachandran plot owing to the presence of a disallowed glycine in the CDR3 region of 1VGE.

The best scoring 1AD9-derived model yielded an overall PROFILER_3D score of 115.6 and a PROSA II score of –10.1. These figures are above the averages for correct protein structures of this size (Lüthy *et al.*, 1992; Sippl, 1993) and are comparable to the templates (data not shown). This model had 92% of residues in core areas of the Ramachandran plot. This final model had two unusual characteristics but both were shared with the templates. The first was the position of AlaL84 in a 'generously allowed area' of the Ramachandran plot which recent data have shown to be very sparsely populated (Kleywegt and Jones, 1996). The backbone angles in the model structure were within 10° of those seen in 1AD9, the only template to share an alanine at this position so that these slightly unusual backbone angles are probably a genuine feature of the structures. The second feature was a region of positive PROSA II energy near the CDR3 region of the VL model which is well conserved among the templates both structurally and in sequence. Examination of the 1MRC and 1MPA templates, which share identical VL CDR3 sequences with the humanized antibody, revealed several unusual features. The hydrophilic residues Ser89 and Arg96 are both unfavorably buried whereas the hydrophobic Val94 is partly exposed. Thus, the unusual packing characteristics of the area, leading to positive PROSA II energy, are a general feature of the molecules. This is consistent with the finding that regions of positive PROSA II energy are rare, but by no means unknown, in correct protein structures (Sippl, 1993). A stereo drawing of the final model, produced with MOLSCRIPT and highlighting the positions of CDRs and residues H71 and H45, is shown in Figure 4.

Model analysis

The final model was used to rationalize the finding that the humanized antibody variant containing a leucine residue at position 45 of VH had greater activity than the proline containing variant. This difference should be due solely to side chain effects since the main chain ϕ, φ angles of the leucine are —82, 152° within a highly favored Ramachandran plot area for proline residues (MacArthur and Thornton, 1991). Inspection of the final model and modeling of the Leu→Pro mutation highlight two consequences of this change. The first would be a loss of hydrophobic interactions with neighboring residues ValH37, GlnH39, TyrH91, TrpH103, ProL44 and GlnL38 (Figures 4 and 5). Some of these residues lie near CDRs so that changes in their packing could have consequences for CDR conformation and hence for antibody activity. Additionally, any change at the interface leading to an alteration in relative orientation of VH and VL domains would have consequences for activity. The second consequence of the Leu→Pro mutation would be steric clashes with either PheL98 or TyrL86, depending on the side chain conformation adopted by the proline. Both of these residues are neighbors of L CDR3 so that the effect of the mutation could be transmitted through them to the antibody binding site. Thus, the location of LeuH45 as revealed by the model, packing against CDR neighbors of both domains, can readily explain the antibody's sensitivity even to its relatively conservative mutation to proline.

Discussion

The major goal of the humanization process is to produce more convenient products for clinical use. This is achieved through the correct choice of the constant region followed by the modification of the variable region through replacement of the original murine residues with homologous human residues. The latter process is more time consuming and involves the art of reducing the immunogenicity of the variable domain while preserving the appropriate binding site. Many different procedures have been proposed.

The first humanized antibodies were based on known human antibodies, normally myeloma proteins, that had been crystallized and whose three-dimensional structure had been described (Riechmann *et al.*, 1988). This procedure was convenient owing to the ease of predicting structural incompatibilities between the original murine and the proposed humanized version of the variable region. The major problem in this process was the limited repertoire of available high-resolution human antibody structures, so that often only limited sequence similarity between known structures and the sequence of interest was observed. The consequence was that a great number of murine residues had to be introduced into the proposed humanized version to mold the original paratope, so that the humanized version preserved a strong immunogenic potential.

More recent attempts have used the closest human sequence (Queen *et al.*, 1989; Co *et al.*, 1996). The choice of a sequence is mostly made from rearranged immunoglobulin genes derived

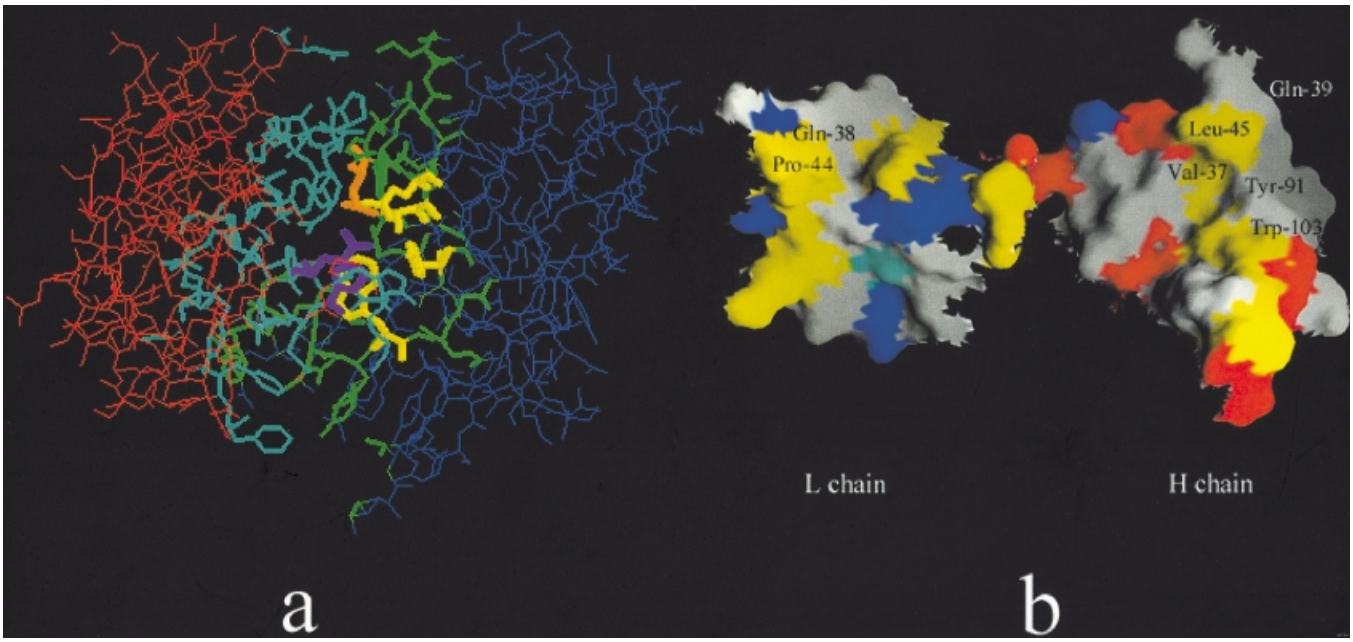

Fig. 5. (**a**) Sting produced image of the L chain (red wireframe) and its FRs (cyan sticks) and H chain (blue wireframe) with its IFRs (green sticks). Orange stick is Leu45. Gln39, Val37, Tyr91 and Trp103 are yellow sticks. Pro44 and Gln38 are from L chain and are purple. (**b**) Open book view of the interface between L and H chain. This presentation was created by rotating two surfaces of the L and H chain in opposing directions (by 90°) and around the axis (within the plane of the figure) that crosses the two molecules at their intersection. Appropriate coloring of residues at the surface indicates elements of charge complementarity as well as hydrophobic interactions. Color coded is dione with respect to the type of residue that makes visible surface: positively charged residues are blue, histidine is cyan, negatively charged residues are red, hydrophobic residues are yellow and polar residues are white.

from splenic B-lymphocytes, which are the commonest sequences in databanks. Since the hypermutation process is unique to each B cell clone under the pressure of a particular individual, there is concern as to whether hypermutation-derived residues could be immunogenic to other individuals. To avoid this potential problem, we chose the best fitting human Ig sequence that was not modified by the clonal development of the lymphocyte.

The 6.7 monoclonal antibody recognizes the CD18 antigen on the surface of leukocytes and it has been shown to inhibit adhesion of $CD18^{+}$ cells (David *et al.*, 1991). The nucleotide sequence that codes for the variable portion of the 6.7 mAb reveals a hypermutated VH. VDJ recombination of a JH4 gene fragment was found. VL is formed by the recombination of a Vk1 to Jk1.

In order to design the humanized version of the heavy chain, we chose a human germline VH sequence and grafted on to it the murine CDR2 and an extended CDR1, while the murine CDR3 was joined to the human germline JH4 genic fragment. We opted for grafting a long CDR1 (Kabat residues 26–35) including both the archetypal Kabat CDR1 (Kabat *et al.*, 1991) and the loop H1 (Chothia *et al.*, 1989) for the following reason. Unlike CDR2, which includes H2, CDR1 is offset from H1. This means that the loop H1 is less variable than the nearby region classified by Kabat *et al.* (1991) as the first CDR, suggested to have a role in determining binding specificity. Despite this observation, CDR1 has been implicated in binding specificity probably by affecting the H1 conformation. In addition to the CDRs, the framework residue alanine 71 of the original murine antibody was maintained in the humanized versions since the size of the residue in this position was shown to be the major determinant of the position and conformation of the H2 loop (Tramontano *et al.*, 1990). In this approach we included only one murine residue in the human framework. Other methods almost always require the introduction of many murine residues in the reshaped antibody, for example, four (Presta *et al.*, 1993), seven (Queen *et al.*, 1989), 12 (Hakimi *et al.*, 1993) or even 20 (Woodle *et al.*, 1992) amino acid substitutions in the humanized VH. A smaller number of murine residues in the humanized antibody favors a lower immunogenic potential in patients.

Examination of the Fv structure 1AD9 of the PDB database, an antibody with high sequence similarity to the HG3, was used to highlight differences between the human germline and the mouse VH sequences that could be critical for the formation of a compatible paratope. Most differences represented solvent-exposed residues and these positions were maintained as in the human sequence. Among the different buried residues only position 45 (Pro in murine, Leu in human) was seen as potentially significant since it was in close contact to residues of H CDR2. Although both of these residues are hydrophobic, proline is known to have specific effects on protein backbone structure (MacArthur and Thornton, 1991). We therefore prepared two versions for the VH, Fv-VH(P) and Fv-VH(L).

These humanized VH versions were synthesized completely by PCR. Selected clones containing the expected VH sequences were transferred to the scFv vector, carrying the murine VL, for expression in *P.pastoris*. Purified scFvs were tested by FACS and compared with an unrelated anti-Z-DNA scFv (Brígido *et al.*, 1993). The binding of scFvs was detected through its C-terminal protein A domain by the use of a rabbit immunoglobulin conjugated to FITC. In Figure 2 we note that all anti-CD18 scFvs stained lymphocytes. The blocking experiments (Figure 3) confirm that the scFv binding to lymphocytes was dependent on CD18 antigen, suggesting that the scFvs were recognizing the same epitope as the original mAb. Moreover, in this case, the Fv-VH(L) construction was more efficient in blocking the mAb, consistent with its apparent

higher affinity. We cannot discard the possibility that there are different fractions of folded molecules in each of the hemi-humanized preparations, although the use of an efficient host such as *Pichia pastoris* for the expression of eukaryotic proteins should minimize folding problems.

A molecular model was built in order to rationalize the different behavior of the Fv-VH(L) and Fv-VH(P) humanized antibodies, with the presence of Leu at position 45 improving the binding activity towards CD18. The explanation of this seems to lie with perturbations to intra- and inter-chain packing caused by the presence of the cyclic proline side chain, rather than with the unusual main chain conformational preferences of proline residues. Surprisingly this position is occupied by a Leu residue in both a putative murine germline (Gen Bank M13787) and the chosen human germline (HG3). It is plausible that the Pro residue may have appeared spuriously as the result of genetic drift in the continuously cultured hybridoma clones. If the proline was introduced during affinity maturation, the apparent lower affinity of this version compared with the Fv-VH(L) may be due to differences between binding of whole antibodies and scFv molecules.

The results presented here support the proposed strategy for the quick design of humanized antibody variable sequences. The approach is based on the search for the closest human germline sequences and the grafting on to it of an extended CDR1 and the Kabat defined CDR2 and CDR3. We can detect binding differences between the two humanized versions, but both were able to bind confirming the success of the binding specificity transfer. Remarkably only one murine residue, Ala71, was introduced in the human framework. This finding was probably due to the fact that germline sequences may possess the basic features necessary for the correct assembly of the paratope. The simplicity of this method may be substantiated by evidence that CDRs affect the overall Fv structure (Holmes and Foote, 1997). The determinism of such rules will dictate the success of this strategy, allowing more predictive protocols to be made for humanization and other structural manipulation of antibodies.

Acknowledgements

We thank Sandra Maria Monteiro for the FACS analyses, Lidia Maria Pepe de Moraes for help with the protein expression experiments, Linda Styer Caldas for English corrections to the manuscript, Andrea Maranhão and Edmar Vaz de Andrade for helpful discussions and technical advice and Cynthia Kiaw for the final format of the figures. This work was supported by CNPq and CAPES.

References

Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) *Nucleic Acids Res*., **25**, 3389–3402.

Benjamin,R.J., Cobbold,S.P., Clark,M.R. and Waldmann,H. (1986) *J. Exp. Med.*, **163**, 1539–1552.

Brígido,M.M. (1992) PhD Thesis, University of Sao Paulo.

Brígido,M.M., Polimenis,M. and Stollar,B.D. (1993) *J. Immunol*., **150**, 469–475.

Brüggemann,M., Winter,G., Waldmann,H. and Neuberger,M.S. (1989) *J. Exp. Med*., **170**, 2153–2157.

Chothia,C. *et al.* (1989) *Nature*, **342**, 877–883.

Co,M.S., Deschamps,M., Whitley,R.J. and Queen,C. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2869–2873.

Co,M.S. *et al.* (1996) *Cancer Res*., **56**, 1118–1125.

Coloma,M.J., Larrick,J.W., Ayala,M. and Gavilondo-Cowley,J.V. (1991) *BioTechniques*, **11**, 152–156.

Couto,J.R., Padlan,E.A., Blank,E.W., Peterson,J.A. and Ceriani,R.L. (1994) *Hybridoma*, **13**, 215–219.

Couto,J.R., Christian,R.B., Peterson,J.A. and Ceriani,R.L. (1995) *Cancer Res. (Suppl.)*, **55**, 5973s–5977s.

Cregg,J.M., Barringer,K.J., Hessler,A.Y. and Madden,K.R. (1985) *Mol. Cell. Biol*., **5**, 3376–3385.

David,V., Leca,G., Corvaia,N., Deist,F.L., Boumsell,L. and Bensussan,A. (1991) *Cel. Immunol*., **136**, 519–524.

Foote,J. and Winter,G. (1992) *J. Mol. Biol*., **224**, 487–499.

Graziano,R.F. *et al.* (1995) *J. Immunol.*, **55**, 4996–5002.

Gronski,P., Seiler,F.R. and Schwick,H.G. (1991) *Mol. Immunol*., **28**, 1321–1332.

Hakimi,J. *et al.* (1993) *J. Immunol.*, **151**, 1075–1085.

Holmes,M.A. and Foote,J. (1997) *J. Immunol*., **158**, 2192–2201.

Jang,Y. and Stollar,B.D. (1992) *Anal. Biochem*., **204**, 407–408.

Johnson,S. *et al.* (1997) *J. Infect. Dis*., **176**, 1215–1224.

Jones,P.T., Dear,P.H., Foote,J., Neuberger,M.S. and Winter,G. (1986) *Nature*, **321**, 522–525.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta. Crystallogr.*, **A47**, 110–119.

Kabat,E.A., Wu,T.T., Perry,H.M., Gottesmann,K.S. and Foeller,C. (1991) *Sequences of Proteins of Imunological Interest.* 5th edn. US Department of Health and Human Services, Bethesda, MD.

Kleywegt,G.J. (1996) *Acta Crystallogr.*, **D52**, 842–857.

Kleywegt,G.J. and Jones,T.A. (1996) *Structure*, **4**, 1395–1400.

Köhler,G. and Milstein,C. (1975) *Nature*, **256**, 495–497.

Kraulis,J. (1991) *J. Appl. Cryst.*, **24**, 946–950.

Laskowski, R., MacArthur,M., Moss,D. and Thornton,J. (1993) *J. Appl. Crystallogr.*, **26**, 283–290.

Lüthy,R., Bowie,J.U. and Eisenberg,D. (1992) *Nature*, **356**, 83–85.

MacArthur,M.W. and Thornton,J.M. (1991) *J Mol Biol.,* **218**, 397–412.

Mountain,A. and Adair,J.R. (1992) *Biotechnol. Gen. Engng Rev.*, **10**, 1–141.

Neshich,G., Togawa,R., Vilella,W. and Honig,B. (1998) *Brookhaven Protein Databank Q. News*, **85**, 6–7.

Nicholls,A., Sharp,K. and Honig,B. (1991) *Proteins: Struct. Funct. Genet.*, 11, 281–286.

O'Connor,S.J., Meng,Y.G., Rezaie,A.R. and Presta,L.G. (1998) *Protein Engng* **11**, 321–328.

Padlan,E.A. (1991) *Mol. Immunol*., **28**, 489–498.

Presta,L.G., Lahr,S.J., Shields,R.L., Porter,J.P., Gorman,C.M., Fendly,B.M. and Jardieu,P.M. (1993) *J. Immunol.*, **151**, 2623–2632.

Queen,C. *et al.* (1989) *Proc. Natl Acad. Sci. USA*, **86**, 10029–10033.

Rechavi,G., Ram,D., Glazer,L., Zakut,R. and Givol,D. (1983) *Proc. Natl Acad. Sci USA*, **80**, 855–859.

Reichmann,L., Clark,M., Waldmann,H. and Winter,G. (1988) *Nature*, **332**, 323–327.

Roguska,M.A. *et al.* (1994) *Proc. Natl Acad. Sci. USA*, **91**, 969–973.

Roguska,M.A. *et al.* (1996) *Protein Engng*, **9**, 895–904.

Rosok,M.J. *et al.* (1996) *J. Biol. Chem*., **271**, 22611–22618.

Sali,A. and Blundell,T.L. (1993) *J. Mol. Biol.*, **234**, 779–815.

Sippl,M.J. (1993) *Proteins*, **17**, 355–362.

Studnicka,G.M., Soares,S., Better,M., Williams,R.E., Nadell,R. and Horwitz,A.H. (1994) *Protein Engng*, **7**, 805–814.

Tempest,P.R. *et al.* (1994) *Protein Engng*, **7**, 1501–1507.

Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res*., **11**, 4673–4680.

Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) *J. Mol. Biol*., **227**, 776–798.

Tramontano,A., Chothia,C. and Lesk,A.M. (1990) *J. Mol. Biol.*, **215**, 175–182.

Vaswani,S.K. and Hamilton,R.G. (1998) *Ann. Allerg. Asth. Immunol*., **81**, 105–115.

Verhoeyen,M., Milstein,C. and Winter,G. (1988) *Science*, **239**, 1534–1536.

Werther,W.A. *et al.* (1996) *J. Immunol*., **157**, 4986–4995.

Woodle,E.S. *et al.* (1992) *J. Immunol*., **148**, 2756–2763.

Wung,J.L. and Gascoigne,R.J. (1996) *BioTechniques*, **21**, 811–812.

Zhou,H., Fisher,R.J. and Papas,T.S. (1994) *Nucleic Acids Res*., **22**, 888–889.

Received September 30, 1999; revised December 21, 1999; accepted February 29, 2000