Journal of Computer-Aided Molecular Design, 11 (1997) 453–461.
KLUWER/ESCOM

J-CAMD 407

Diverse binding site structures revealed in homology models of polyreactive immunoglobulins

Paul A. Ramsland[a], Luke W. Guddat[b], Allen B. Edmundson[c] and Robert L. Raison[a,*]

[a]*Immunobiology Unit, University of Technology, Sydney, Westbourne Street, Gore Hill, NSW 2065, Australia*
[b]*Centre for Drug Design and Development, The University of Queensland, St. Lucia, QLD 4072, Australia*
[c]*Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, OK 73104, U.S.A.*

Received 15 April 1996
Accepted 22 April 1997

Keywords: Polyreactive IgM; Natural (auto)antibodies; Template-based models; Chronic B lymphocytic leukaemia

Summary

We describe here computer-assisted homology models of the combining site structure of three polyreactive immunoglobulins. Template-based models of Fv (V_L–V_H) fragments were derived for the surface IgM expressed by the malignant CD5 positive B cells from three patients with chronic lymphocytic leukaemia (CLL). The conserved framework regions were constructed using crystal coordinates taken from highly homologous human variable domain structures (Pot and Hil). Complementarity determining regions (CDRs) were predicted by grafting loops, taken from known immunoglobulin structures, onto the Fv framework models. The CDR templates were chosen, where possible, to be of the same length and of high residue identity or similarity. LCDR1, 2 and 3 as well as HCDR1 and 2 for the Fv were constructed using this strategy. For HCDR3 prediction, a database containing the Cartesian coordinates of 30 of these loops was compiled from unliganded antibody X-ray crystallographic structures and an HCDR3 of the same length as that of the B CLL Fv was selected as a template. In one case (Yar), the resulting HCDR3 model gave unfavourable interactions when incorporated into the Fv model. This HCDR3 was therefore modelled using an alternative strategy of construction of the loop stems, using a previously described HCDR3 conformation (Pot), followed by chain closure with a β-turn. The template models were subjected to positional refinement using energy minimisation and molecular dynamics simulations (X-PLOR). An electrostatic surface description (GRASP) did not reveal a common structural feature within the binding sites of the three polyreactive Fv. Thus, polyreactive immunoglobulins may recognise similar and multiple antigens through a diverse array of binding site structures.

Introduction

While the majority of immunoglobulins (Igs) exhibit restricted antigen binding specificity, polyreactive immunoglobulins can recognise multiple structurally dissimilar antigens. Malignant cells from patients with chronic B lymphocytic leukaemia (B CLL) frequently express polyreactive antibodies, thus providing a system for investigations into the structural features important for polyreactive antigen binding. Unlike the situation with monospecific antibodies where structural correlation of the binding site with antigen specificity has been elucidated for a number of antigen/antibody systems [1–5], the structural basis of the polyreactivity of certain antibody combining sites remains unclear. Clarification of this problem requires detailed structural analysis of a large number of polyreactive binding sites, a process that will be greatly assisted by the use of predicted structures in addition to the essential, but limited, number of experimentally determined structures for binding sites of this type.

The antigen binding properties of Ig are largely dictated by six (three on light chains and three on heavy chains) complementarity determining regions (CDRs) in the variable (V) domains which provide a unique molecular surface for specific antigen recognition. The CDRs are supported by a highly conserved molecular scaffolding provided by the framework regions. Various methods have been described to predict the three-dimensional (3D)

*To whom correspondence should be addressed.

Fig. 1. C^{α} representations of the HCDR3 conformations. Left: the 16-residue HCDR3 loops. Bel is coloured blue, Yar 1 is in yellow and Yar 2 in magenta. The crystal structure of the HCDR3 of NC10.14 is in green. Right: the 11-residue HCDR3 loops. Tre is coloured yellow, and the crystal structures of 1FVC, 1HIL and 1MCP are in green, purple and red, respectively.

structure of antibody Fv (V_L–V_H) molecules [6–11]. Usually the structurally conserved framework regions of the V domains are constructed from atomic coordinates taken directly from an Ig crystal structure. The conformations of the six CDR loops are predicted by scanning the sequence for key 'canonical' structure determining residues to select a template loop [12,13] or by randomly sampling conformational space for loops with low potential energy values [14,15]. This approach is usually adequate for five of the six loops, but the third loop on the heavy chain (HCDR3) shows the greatest variability in length and sequence, resulting in an inability to accurately predict its structure. In view of this, there is a need for a combined modelling and crystallographic approach to improve the quality of HCDR3 predictions. In comparison, framework regions and CDR loops other than HCDR3 are well predicted by homology modelling. Homology-based models of variable domain structures, when compared with experimentally determined structures, show root

TABLE 1
COMPARISON OF THE 3D6 CRYSTAL STRUCTURE WITH THE HOMOLOGY Fv MODEL

Region of the variable domain	Rmsd for all C^{α} atoms (Å)[a]
LCDR1	0.67
LCDR2	0.35
LCDR3	0.84
HCDR1	0.23
HCDR2	0.70
HCDR3	1.30
V_L [b]	0.74
V_H [b]	0.49
V_L–V_H [b]	1.0

[a] Rms values shown are for the α carbon atoms used for rigid-body alignments.

[b] V domains and Fv (V_L–V_H) were aligned using the conserved β-sheet framework residues (i.e. framework residues are 1–23, 35–49, 57–88, 98–107 for light chains and 1–30, 36–49, 66–94, 103–113 for heavy chains) [20].

Fig. 2. Side views of the two Yar Fv models. The heavy-chain V domains are represented in magenta and the light chains in yellow. The van der Waals radii are in red for LCDR1 and in green for HCDR3. Top: Yar 1 based on the HCDR3 conformation of NC10.14; bottom: Yar 2 based on the HCDR3 conformation of 1IGM.

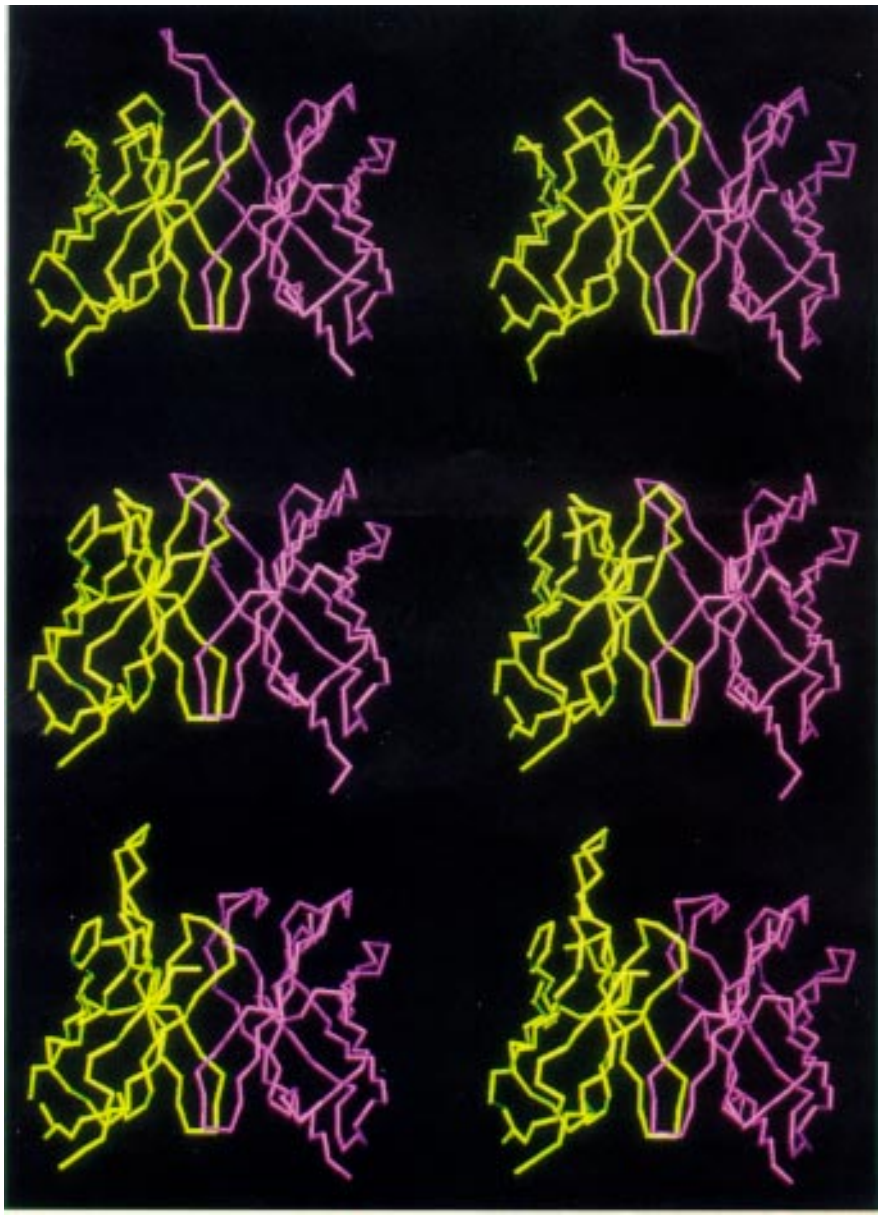

Fig. 3. Stereo C^{α} diagrams of the side views of the polyreactive Fv homology models. The heavy chains are in magenta and the light chains are in yellow. Top: Bel; middle: Tre; bottom: Yar.

mean square (rms) deviations of less than 1.0 Å for framework and canonically predicted CDRs other than HCDR3 [11–13]. Thus, homology modelling can provide a useful structural description of the binding site, comparable to the quality of a low-resolution X-ray structure [16].

In this study, we have applied homology modelling to determine the structures of the variable domains of three polyreactive antibodies. Electrostatic molecular surface models [17] reveal the level of binding site diversity exhibited by this group of immunoglobulins.

TABLE 2
VARIABLE DOMAINS USED TO CONSTRUCT Fv MODEL FRAMEWORK REGIONS

B-CLL Fv model (isotype)	V_L framework residue identity (%) (PDB code)	V_H framework residue identity (%) (PDB code)
Bel (IgM λ)	75 (8FAB)	60 (1IGM)
Tre (IgM κ)	82 (1IGM)	62 (1IGM)
Yar (IgM κ)	69 (1IGM)	87 (1IGM)

Materials and Methods

Immunoglobulins

The complete nucleotide sequences of the variable domains of the Ig described here will be published elsewhere (manuscript in preparation). Here we only show the translated amino acid sequences of the light- and heavy-chain CDR loops (Table 3). The IgM molecules are monoclonal, exhibit polyreactive binding properties and are expressed on the surface of malignant $CD5^+$ B cells from patients with B CLL. The IgM were designated as Bel (IgM λ), Tre (IgM κ) and Yar (IgM κ).

Computer-aided modelling

All computational work was performed on an INDY workstation running IRIX, v. 5.2 (Silicon Graphics Inc.). Template models were constructed using the software TURBO-FRODO, v. 5 (Biographics, France). Minimisation and molecular dynamics were performed using X-PLOR, v. 3.1 [18]. Molecular surfaces and electrostatic potentials were calculated and visualised with GRASP, v. 1.2.5 [17].

Modelling strategy

The B CLL Fv molecules were constructed by starting with known high-resolution crystal structures (1BAF, 1BBD, 1IGM, 2FB4, 8FAB, 1DFB, 1REI, 1HIL, 1MCP) taken from the Brookhaven Protein Databank [19]. Variable domain templates were selected from human Ig fragments on the basis of highest amino acid identity within framework regions. Where necessary, V domain templates were aligned, using rigid-body alignment of conserved Ig β-sheets [20] over the corresponding domain of the Pot (1IGM) structure. Thus, all Fv models have the V_L–V_H domain interface of Pot IgM Fv [21]. CDR templates were selected initially on the basis of an identical number of residues comprising the loop. Sequence homology was used as a secondary criterion. Furthermore, only residues within the CDRs as defined by Kabat et al. [22] were considered. The CDRs were classified, where possible, into canonical classes [12,13]. Loops which were found to be of a different length to the available loops of known conformation were constructed using a template-based approach, where the loop stems were built using CDRs of known conformation. Additional residues were placed onto these stems and chain closure was accomplished by grafting a 'classical' β-turn onto the CDR loop stems. The type of β-turn assigned was based on the criteria detailed by Wilmot and Thornton [23].

Structures were optimised using the CHARMM22 topology and parameter set implemented within X-PLOR, v. 3.1 [18]. The template-based models were subjected to an initial 100 cycles of conjugate gradient energy minimisation [24]. After this initial minimisation step, harmonic coordinate restraints of 20.0 kcal/(mole $Å^2$) were set for all main-chain α carbon atoms. A further 200 cycles of energy minimisation were then carried out. The Fv models were subjected to molecular dynamics and simulated annealing where the structures were initially heated to 1000 K followed by slow cooling to a bath temperature of 300 K. The time step for the simulation was set to 0.5 fs. A final step of 300 cycles of energy minimisation was carried out.

Comparison of structures

The stereochemistry and geometry of Fv models were assessed by calculating the deviations from ideality of bond lengths, bond angles, dihedrals and improper angles using the parameter set of Engh and Huber [25]. All ϕ, ψ torsional angles were analysed by the method of Ramachandran and Sasisekharan [26]. Optimal rigid-body structural alignments and rms deviations between structures were calculated within TURBO-FRODO. Interchain energies were calculated from the van der Waals (vdW) and electrostatic energy terms in X-PLOR [18]. For these calculations, the protein dielectric was set to 2.0 without a nonbonded cutoff using the CHARMM22 force field.

Results

Validation of modelling strategy

To evaluate the template-based modelling strategy, a homology model of the variable domains of the human antibody 3D6 was constructed. The 3D6 structure [27]

TABLE 3
CDR SEQUENCES OF THE POLYREACTIVE Fv MOLECULES AND THE CDR TEMPLATES SELECTED FOR MODELLING

CDR[a]	Bel Fv model (PDB codes)[b,c]	Tre Fv model (PDB codes)[b,c]	Yar 2 Fv model (PDB codes)[b,c]
L1	GGNNIGSDSVH (8FAB)	RASQSISSYLN (1DFB), **2**	KSSQSVLYSSNNKNYLA (1BBD), **3**
L2	YDSDRPS (2FB4)	AASSLQS (1DFB), **1**	WASTRES (1DFB), **1**
L3	QVWDSSSVV (8FAB)	QQSYSTPPYT (1BAF), **1**	QQYYSTPYT (1REI), **1**
H1	NYWIG (8FAB), **1**	SYWIS (1HIL), **1**	SYAMH (1DFB), **1**
H2	IIYPGDSDTRYSPSFQG (2FB4), **2**	RIDPSDSYTNYSPSFQG (1BBD), **2**	AISSNGGTTYYADSVKG (1HIL), **3**
H3	RTGTGDPYYYYYYMDV (NC10.14)	RQWLALGHFDY (1MCP)	TYYDFWSGYSPNWFDP (1IGM, 4PTP)

[a] CDRs were assigned according to Kabat et al. [22].
[b] Residue identities between the template loop (PDB code indicated) and the polyreactive Fv sequences have been underlined.
[c] Where a 'canonical' class could be assigned to a CDR, the class is indicated in bold [12,13].

TABLE 4
INTERCHAIN INTERACTION ENERGIES OF TWO ALTERNATIVE YAR Fv MODELS

Interaction	E-total[a] (kcal/mol)	E-vdW[b] (kcal/mol)	E-ELEC[c] (kcal/mol)
Yar 1			
V_L–V_H	−58.3	−84.1	25.7
V_L–HCDR3	−7.5	−45.9	38.4
Yar 2			
V_L–V_H	−103.4	−86.5	−16.9
V_L–HCDR3	−50.5	−45.8	−4.7

[a] All energy values are calculated for interchain interactions only using X-PLOR [18].
[b] van der Waals energy term.
[c] Electrostatic energy contribution term.

was not in our template database. The coordinates from 1IGM were used to determine the V domain framework regions. Residue identities in these regions of 3D6 and 1IGM were 86% and 90% for V_H and V_L, respectively. Template CDR structures of the same residue length as the 3D6 CDR loops were selected on the basis of maximum sequence homology with the target CDR. LCDR1 and LCDR2 were taken from 1IGM, and HCDR1 and HCDR2 from 2FB4. Template loops of the same length were not available for LCDR3 (7 residues) and HCDR3 (17 residues). CDR3 loops were built using loop stems of known conformation, taken from corresponding loops of 1REI for LCDR3 and NC10.14 (Guddat and Edmundson, unpublished structure) for HCDR3, followed by chain closure with a classical β-turn [23]. The rms deviations (α carbon atoms) between the model and the crystal structure were determined after optimal rigid-body alignments of various regions of the two structures (Table 1). The modelled CDR loops, the individual V domains and the modelled Fv fragment were compared to the published 3D6 structure, with rms deviations between the crystal and model structures of <1.0 Å for all regions except for HCDR3, which has an rms deviation of 1.30 Å.

Template models of polyreactive immunoglobulins

The template V domains used to model the framework regions of the polyreactive Fv are shown in Table 2. In each case, they represent the best identity match from the database with the respective CLL IgM V domains. Residue identities ranged from 87% for the Yar/1IGM V_H pairing to 60% for the V_H of Bel and 1IGM. Residues were replaced within the program TURBO-FRODO (Biographics). The side chains of the mutated residues were manually positioned into a conformation which minimised any nonbonded contacts and the bond geometries were refined against a rotamer library (TURBO-FRODO). The templates used for constructing the CDRs of the three Fv models are summarised in Table 3. Where a canonical class could be assigned to a CDR loop, it is indicated. However, loops were initially chosen as templates on the basis of residue identities and similarities within the CDRs as defined by Kabat et al. [22].

In an attempt to develop a template-based strategy for predicting the structure of HCDR3, we extracted these loop coordinates from the 30 available unliganded Ig crystal structures (29 from the Protein Databank [19] and one from the NC10.14 structure (Guddat and Edmundson, unpublished results)). The length of the HCDR3 loops ranged from 5 to 17 residues but there were no loops of 6 or 13 residues long. Bel and Yar antibodies both have 16-residue HCDR3 loops, the same length as a murine Ig, NC10.14. For the 11-residue HCDR3 of Tre, there were three structures in the database.

The predicted HCDR3 conformations and a comparison with their starting models are presented in Fig. 1. The Bel HCDR3 was constructed from the HCDR3 of the NC10.14 Fab structure. This loop conformation fitted the Bel template model and did not result in unfavourable interactions with the other five CDRs in the Fv model. The Yar HCDR3 was initially constructed using the NC10.14 conformation (Yar 1). However, when this structure was used in the Yar Fv model, the backbone of the LCDR1 loop came into close contact with the HCDR3. An alternative model for the Yar HCDR3 (Yar 2) was then constructed to incorporate the fold seen in the Pot (1IGM) HCDR3 [21]. Additional residues were added and the chain closed using a β-turn taken from the trypsin (4PTP) structure. The two models of Yar Fv are presented (Fig. 2) showing the vdW surfaces of the LCDR1 and HCDR3 loops. It was noted that the vdW radii of the HCDR3 and LCDR1 of Yar 1 exhibited many close contacts, whereas the HCDR3 vdW radii in Yar 2 made fewer close contacts with LCDR1. The interchain energies between the two V domains and between HCDR3 and V_L were calculated for the two proposed structures of Yar Fv using X-PLOR (Table 4). Although the vdW contacts were similar between the two loop structures, unfavourable electrostatic energy was calculated for the HCDR3 structure based on an NC10.14 template (Yar 1). In comparison, modelling based on Pot HCDR3 (Yar 2) resulted in a significantly improved electrostatic energy term. The template for the Tre

TABLE 5
Fv MODEL STEREOCHEMISTRY

Rms deviations[a]	Bel Fv	Tre Fv	Yar 2 Fv
Bond distances (Å)	0.004	0.004	0.004
Bond angles (°)	0.9	1.0	1.0
Dihedral angles (°)	26.0	26.1	26.0
Improper angles (°)	0.8	0.9	0.9

[a] Rms deviations from ideality were calculated for all residues in the given Fv model structure, compared against Engh and Huber parameters [25].

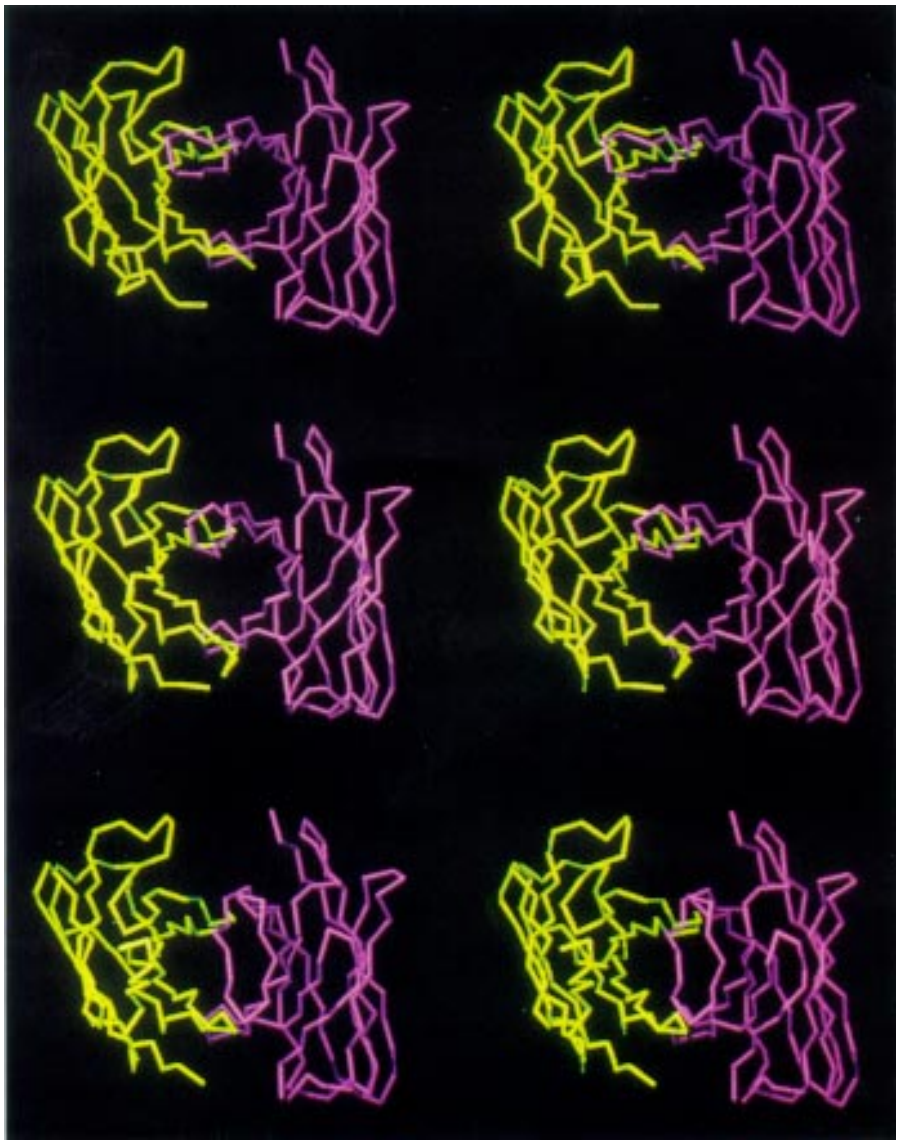

Fig. 4. Stereo C^{α} diagrams of the end-on views of the polyreactive Fv homology models. The heavy chains are in magenta and the light chains are in yellow. Top: Bel; middle: Tre; bottom: Yar.

HCDR3 was the 11-residue loop of 1MCP. This loop showed the greatest similarity in sequence to the Tre loop compared with the other known loops of identical length.

Refined Fv models

The Fv models after energy minimisation and molecular dynamics refinement using the program X-PLOR [18] are presented in Figs. 3 and 4. The rms deviations in bond lengths, bond angles and dihedral angles are presented in Table 5. A Ramachandran plot [26] was calculated and the majority of ϕ,ψ torsion angles around C^{α} atoms are within the allowed regions. These plots are shown for the Fv models after the full positional refinement protocol (Fig. 5) and are a further indication that appropriate stereochemistry and geometry have been achieved in the models.

Electrostatic surface models of polyreactive Fv

Binding site topology is shown as a surface representation [17] in Fig. 6. The solvent-accessible surfaces and the electrostatic potentials are depicted for the polyreactive Fv structures. A deep binding pocket is seen at the V_L–V_H domain interface of the Bel and Tre Fv models. The two alternative Yar Fv models (Yar 1 and Yar 2) are also presented in Fig. 6. The electrostatic surface potential of the Bel binding site is predominantly hydrophobic in nature, with regions of negative potential restricted to the edges of the binding cavity. The Tre binding site has a strong positive electrostatic potential lining the floor of the binding pocket. The binding site of Yar Fv exhibits an essentially hydrophobic electrostatic surface potential which is not significantly altered between the two models presented for this Fv.

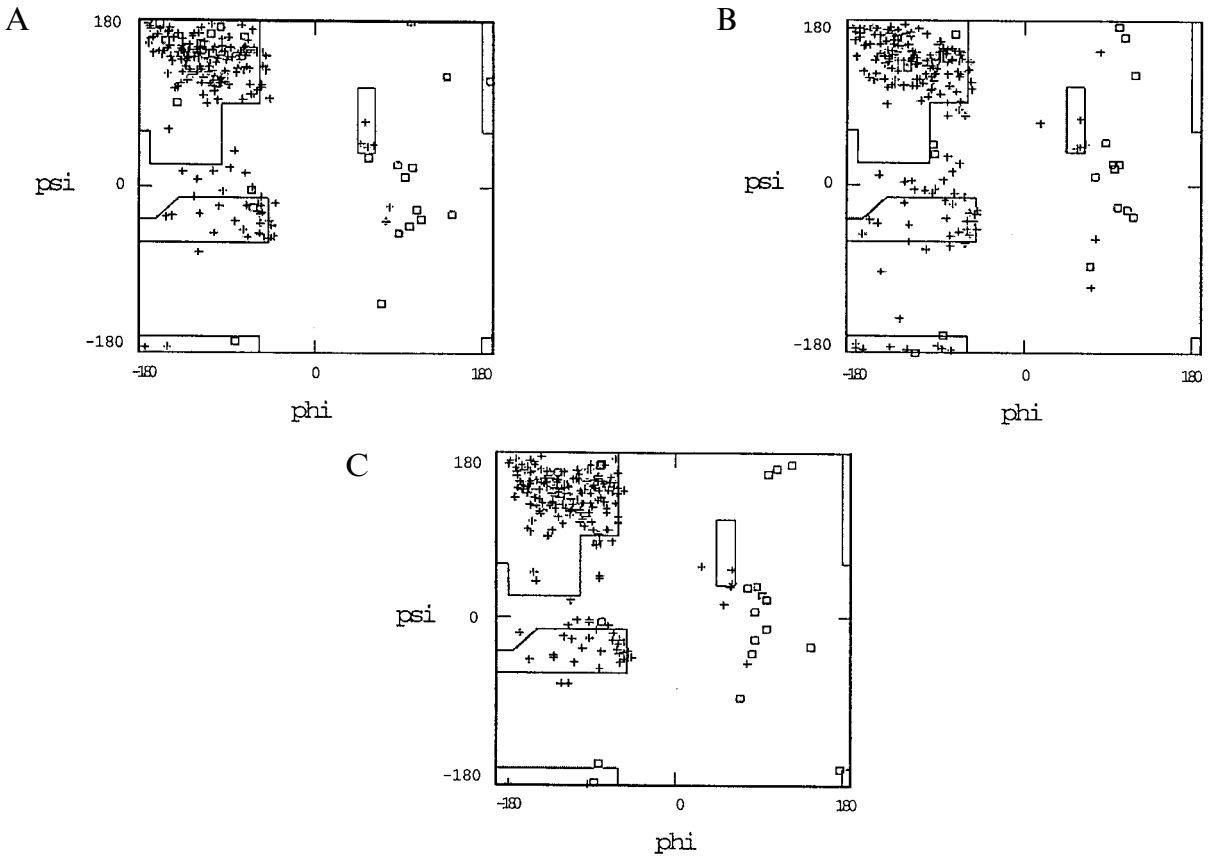

Fig. 5. Ramachandran plots [26] of the ϕ,ψ angles for all α carbon dihedral angles in the variable region fragment homology models. Crosses represent all the residue types excluding glycines which are described by the boxes. Lines delimit the favoured dihedral conformations. Ramachandran plots are labelled: (A) Bel; (B) Tre; (C) Yar 2.

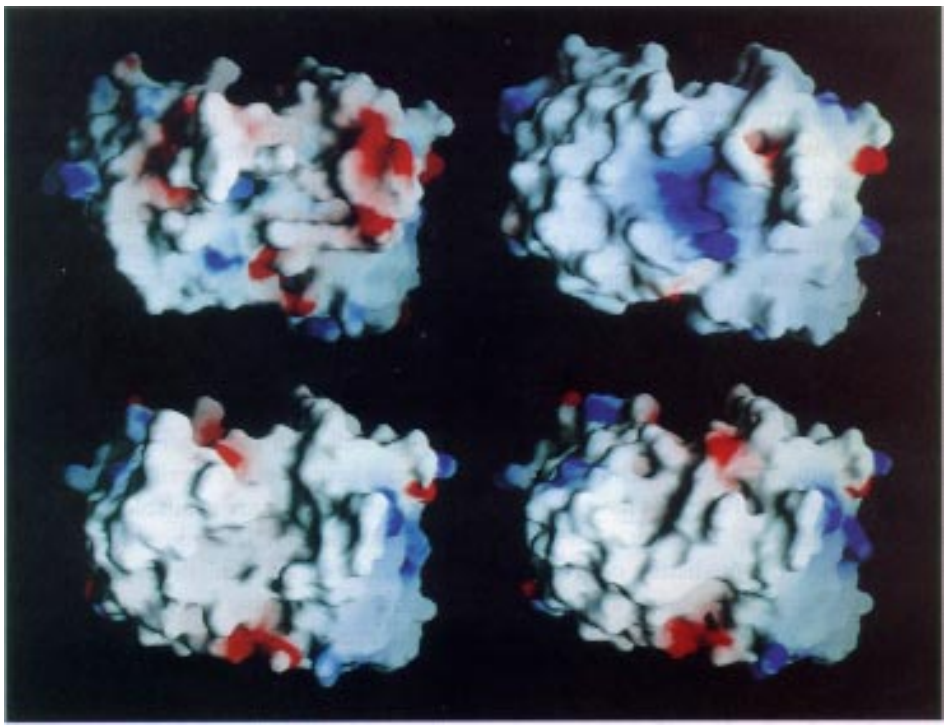

Fig. 6. Electrostatic surface representations of the Fv homology models (end-on views). Top left: Bel; top right: Tre; bottom left: Yar 1; bottom right: Yar 2. The solvent-accessible surface and the electrostatic potential across this surface have been represented using the program GRASP [17]. White represents areas of zero surface potential, blue represents positive surface potential and red represents negative surface potential.

Discussion and Conclusions

Computational techniques for predicting antibody combining site structures have generally approached the problem by utilising the conserved Ig variable domain scaffolding to construct good approximations of antibody variable region fragments [11]. These models provide important information regarding molecular recognition. The homology models of polyreactive Fv molecules presented here were constructed in an attempt to identify structural features of polyreactive Ig antigen binding sites.

The topology of immunoglobulin combining sites is predominantly determined by the primary sequence and folded conformation of the six CDRs. The conformation adopted by CDR loops is largely dependent on the number of residues therein. Additionally, CDR conformation is affected by 'key' structure determining residues both within the loop or in neighbouring framework regions. Chothia et al. [12,13] identified this dependence of loop conformation on length and sequence and produced an elegant theory of canonical CDR loop structures. In our homology models, we did not initially analyse the templates for the presence of 'key' framework residues. Only the residues assigned to CDR loops by Kabat et al. [22] were considered in the choice of the templates. Subsequent analysis has revealed that the CDR loops of the two κ light-chain V domains (Tre and Yar) and all three heavy-chain CDR1 and CDR2 structures fall into previously described canonical classes (Table 3). The antibody Bel expresses a λ light chain. In this case, none of the light-chain CDR loops could be assigned to a canonical class. This is mainly due to a disproportionate number of κ compared to λ light-chain structures available at high atomic resolution. At least five out of the six CDR loops within each of our polyreactive Fv homology models were predicted using template loops of the same length. Previous studies have demonstrated [6,12,13] that these loops are predicted accurately by homology modelling and so provide a good approximation of the basic binding site topology.

No current modelling procedure has consistently predicted the conformation of heavy-chain CDR3 loops with a high level of accuracy. This is mainly due to the heavy-chain CDR3 loop being the most variable with respect to length and sequence. Human HCDR3 loops have lengths varying from 2 to 26 residues [28]. In order to address this difficulty, our approach to the prediction of HCDR3 loops has been to construct these loops using a template-based loop building strategy. We constructed a 3D database of HCDR3 from native X-ray structures. Although the loops range from 5 to 17 residues, not all HCDR3 lengths are represented in the database. In some cases, only one structure is known for a given HCDR3 length. As the number of Ig crystal structures increases, the ability of an empirical approach to predict HCDR3 conformations will become increasingly more accurate. In the present study, Bel and Yar contain 16-residue HCDR3 while Tre Fv contains an 11-residue HCDR3 (Fig. 1). The Bel Fv HCDR3 was constructed from the coordinates of a murine monoclonal antibody NC10.14 which has a well-ordered 16-residue β-loop that protrudes out from the binding pocket into the solvent. The HCDR3 of Yar was constructed using either the conformation of this loop found in the NC10.14 Fab (Yar 1) or that of the 1IGM Fv (Yar 2). A comparison of two alternative conformations for the Yar HCDR3 suggested that the more energetically favourable conformation was for the loop to fold across the binding site, resulting in a relatively flat binding surface (Yar 2). This fold was based on the crystallographic structure for the HCDR3 of Pot Fv described previously by Edmundson's group [21]. While it is a distinct possibility that the HCDR3 of Yar could adopt a conformation somewhere intermediate between that of NC10.14 and Pot HCDR3, the current database of HCDR3 loop folds does not provide a more energetically favourable alternative. It will be of great interest to determine the experimental structure of Yar Fv to examine how influential the relatively long LCDR1 is on the conformation of the HCDR3.

The template for the 11-residue HCDR3 of Tre was 1MCP. The 1MCP loop was selected from the three unique HCDR3 sequences of known conformation (Fig. 1). Perhaps shorter HCDR3 loops are more similar in structure to each other than longer loops, a feature which is at least reflected by the 11-residue HCDR3 loops examined. Although there are differences between the three 11-residue HCDR3 loops of known conformation, these are restricted to small movements towards the top of the loops. It appears that large differences in loop conformation are not common for these shorter loops. In view of the apparent conformational restriction of 11-residue HCDR3, we were encouraged to conclude that the fold of the Tre HCDR3 loop is accurately predicted using 1MCP HCDR3 as a template. Another indication that this loop was an appropriate template for Tre HCDR3 is that the conformation of 1MCP HCDR3 is closely maintained after refinement of the Tre Fv model and so this loop fitted well with the other five CDR loops. Especially in cases where several experimental structures of HCDR3 of the same length are available, a template-based approach to modelling HCDR3 should equal or surpass the predictive capabilities of methods based on systematic conformational searching or the random generation of loop structures [14,15].

One of the strongly held precepts regarding the molecular mechanisms of polyreactive antibody binding is that these immunoglobulins would have a conserved combining site structure [29,30]. Padlan [31] argues in a comprehensive review of antibody structure that the poly-

reactive antibody combining site is relatively large, plastic or 'sticky' (high number of aromatic side chains within CDR) in character. Our current data would appear to contradict this concept and point towards a more diverse pattern of binding site structure for polyreactive immunoglobulins. The polyreactive Fv homology models (Figs. 3 and 4) suggest that at least two basic binding site topologies are involved in polyreactive antigen recognition: (i) a pocket-type binding site is predicted for Bel and Tre antibodies; and (ii) a relatively flat binding surface is described for Yar Ig. Furthermore, the HCDR3 of Bel and the LCDR1 of Yar project out from the compact variable domains into the solvent. These solvent-exposed loops are potentially buried during antigen binding. In particular, the highly aromatic nature of the Bel HCDR3, resulting mainly from a stretch of six tyrosine residues, suggests that this loop is buried at the interface between antigen and antibody. The electrostatic surface representations of the Fv models (Fig. 6) indicate that the binding site of Bel is predominantly hydrophobic in character with one region of negative potential. Yar IgM also has a hydrophobic binding surface. However, the binding 18pocket of Tre immunoglobulin has a strongly positive electrostatic surface potential. Finally, with respect to the electrostatics, an examination of the two Yar models (Fig. 6) shows that the electrostatic nature of the binding site is relatively similar even though the HCDR3 is shifted dramatically between a loop which projects upwards from the V domains (Yar 1) to one which folds across the V_L–V_H domain interface (Yar 2). This indicates that an electrostatic surface representation is a robust technique for investigating the binding sites of antibodies even in cases where the confidence of prediction of the HCDR3 may be low.

Elucidation of the structural basis of polyreactive antigen binding will require analysis and comparison of the binding sites of numerous polyreactive Igs. While X-ray crystallography is arguably the definitive method for achieving this goal, it is time consuming, costly and dependent on the ability to obtain suitable protein crystals. Thus, the development of improved antibody modelling techniques will make a significant contribution towards the solution of this problem. We will attempt to establish X-ray structures of Bel, Tre and Yar Fv from protein obtained through the bacterial expression of the V_L and V_H domains. The model structures (published here in advance) will then be compared with the X-ray structures. Such comparisons will result in refinement of the modelling process and will increase the size of the database of relevant antibody structures.

Acknowledgements

We are grateful to Dr. Leif Hanson for pertinent comments on the manuscript and to Kim Andersen for her expert computing skills. This work was supported by a grant from the Australian National Health and Medical Research Council (NHMRC). P.A.R. is a recipient of an Australian Postgraduate Award (APA).

References

1 Amit, A.G., Mariuzza, R.A., Phillips, S.E.V. and Poljak, R.J., Science, 233 (1986) 747.
2 Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M. and Webster, R.G., Nature, 326 (1987) 358.
3 Herron, J.N., He, X.-M., Mason, M.L., Voss Jr., E.W. and Edmundson, A.B., Proteins Struct. Funct. Genet., 5 (1989) 271.
4 Herron, J.N., He, X.-M., Ballard, D.W., Blier, P.R., Pace, P.E., Bothwell, A.L.M., Voss Jr., E.W. and Edmundson, A.B., Proteins Struct. Funct. Genet., 11 (1991) 159.
5 Arevalo, J.H., Hassig, C.A., Stura, E.A., Sims, M.J., Taussig, M.J. and Wilson, I.A., J. Mol. Biol., 241 (1994) 663.
6 Barry, M.M., Mol, C.D., Anderson, W.F. and Lee, J.S., J. Biol. Chem., 269 (1994) 3623.
7 Bassolino-Klimas, D., Bruccoleri, R.E. and Subramaniam, S., Protein Sci., 1 (1992) 1465.
8 Bajorath, J., Bioconj. Chem., 5 (1994) 213.
9 Chothia, C., Lesk, A.M., Levitt, M., Amit, A.G., Mariuzza, R.A., Phillips, S.E.V. and Poljak, R.J., Science, 233 (1986) 755.
10 De la Paz, P., Sutton, B.J., Darsley, M.J. and Rees, A.R., EMBO J., 5 (1986) 415.
11 Martin, A.C.R., Cheetham, J.C. and Rees, A.R., Methods Enzymol., 203 (1991) 121.
12 Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzari, P.M. and Poljak, R.J., Nature, 342 (1989) 877.
13 Chothia, C., Lesk, A.M., Gherardi, E., Tomlinson, I.M., Walter, G., Marks, J.D., Meirion, B.L. and Winter, G., J. Mol. Biol., 227 (1992) 799.
14 Bruccoleri, R.E. and Karplus, M., Biopolymers, 26 (1987) 137.
15 Bajorath, J. and Fine, R.M., Immunomethods, 1 (1992) 137.
16 Sali, A., Curr. Opin. Biotechnol., 6 (1995) 437.
17 Nicholls, A. and Honig, B., J. Comput. Chem., 12 (1991) 435.
18 Brunger, A.T., X-PLOR v. 3.1. A System for X-ray Crystallography and NMR, Yale University Press, New Haven, CT, U.S.A., 1992.
19 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
20 Guddat, L.W., Shan, L., Anchin, J.M., Linthicum, D.S. and Edmundson, A.B., J. Mol. Biol., 236 (1994) 247.
21 Fan, Z.C., Shan, L., Guddat, L.W., He, X.-M., Gray, W.R., Raison, R.L. and Edmundson, A.B., J. Mol. Biol., 228 (1992) 188.
22 Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. and Foeller, C., Sequences of Proteins of Immunological Interest, 5th ed., National Institutes of Health, Bethesda, MD, U.S.A., 1991.
23 Wilmot, C.M. and Thornton, J.M., J. Mol. Biol., 203 (1988) 221.
24 Powell, M.J.D., Math. Programming, 12 (1977) 241.
25 Engh, R.A. and Huber, R., Acta Crystallogr., A47 (1991) 392.
26 Ramachandran, G.N. and Sasisekharan, V., Adv. Protein Chem., 23 (1968) 283.
27 He, X.-M., Ruker, F., Casale, E. and Carter, D.C., Proc. Natl. Acad. Sci. USA, 89 (1992) 7154.
28 Wu, T.T., Johnson, G. and Kabat, E.A., Proteins Struct. Funct. Genet., 16 (1993) 1.
29 Avrameas, S. and Ternynck, T., Mol. Immunol., 30 (1993) 1133.
30 Cheung, S.C., Takeda, S. and Notkins, A.L., Clin. Exp. Immunol., 101 (1995) 383.
31 Padlan, E.A., Mol. Immunol., 31 (1994) 231.