Analysis and prediction of V_H/V_L packing in antibodies

K.R.Abhinandan[1,2] and Andrew C.R.Martin[1,3]

[1]Institute of Structural and Molecular Biology, Darwin Building, University
College London, Gower Street, London WC1E 6BT, UK and
[2]Present address: NIBR Biologics Center (NBC), Novartis Institutes for
BioMedical Research (NIBR), 4056 Basel, Switzerland

[3]To whom correspondence should be addressed.
E-mail: andrew@bioinf.org.uk, a.martin@biochem.ucl.ac.uk

**The packing of V_H and V_L domains in antibodies can
vary, influencing the topography of the antigen-combining
site. However, until recently, this has largely been ignored
in modelling antibody structure. We present an analysis of
the degree of variability observed in known structures
together with a machine-learning approach to predict the
packing angle. A neural network was trained on sets of
interface residues and a genetic algorithm designed to
perform 'feature selection' to define which sets of interface
residues could be used most successfully to perform
the prediction. While this training procedure was very
computationally intensive, prediction is performed in a
matter of seconds. Thus, not only do we provide a rapid
method for predicting the packing angle, but also we
define a set of residues that may be important in antibody
humanization in order to obtain the correct binding
site topography.**
Keywords: antibody modelling/antibody structure/feature
selection/humanization/machine learning

Introduction

The variability of antibodies is encoded in the variable fragment (or *Fv* region), which consists of two protein domains (V_H and V_L) from the heavy and light chains, respectively. The V_H/V_L interface, which influences the stability of the *Fv* region, has been shown to affect the binding kinetics of a peptide (Chatellier *et al*., 1996). The framework region at the V_H/V_L interface consists of two β sheets (Poljak *et al*., 1973), the structures of which are conserved across *Fv*, *Fab* and light chain dimers (Chothia *et al*., 1985; Novotný and Haber, 1985). Packing of the V_H and V_L domains was analysed in detail by Chothia *et al*. (1985). They recognized that V_H/V_L packing involved a 'three-layer packing' with primarily aromatic side chains being involved in the interface. However, this analysis was based on just three antibody structures, a data set too small to analyse variability in the packing angle. They simply stated that the angle is $\sim -50°$, although they did not state how this value was calculated.

The contribution of residues in the framework regions to interactions with the antigen remains poorly understood. It has been demonstrated that modification of residues relatively distant from the antigen-combining site of the antibody can have a significant effect on the binding affinity for the antigen (Chatellier *et al*., 1996; Roguska *et al*., 1996; Adair *et al*., 1999). For example, Adair *et al*. (1999) demonstrated that modification of residue H23 could significantly affect binding of antibody and antigen.

Earlier work on antibody modelling has ignored variation in V_H/V_L packing (Martin *et al*., 1989; Martin *et al*., 1991; Whitelegg and Rees, 2000), though very recent work has started to consider this as an important factor (Narayanan *et al*., 2009; Sircar *et al*., 2009; Sivasubramanian *et al*., 2009). However, the work that has been performed has not included a thorough analysis of the distribution of packing angles and predictions of the interaction have used computationally intensive energy calculations.

Here we present an analysis of the distribution of the V_H/V_L packing angle and a method to predict the interface angle using machine learning. The trained machine learning method is able to provide a very rapid prediction of the packing angle. Knowing the packing angle prior to modelling the variable region light and heavy chain may help in choosing more appropriate template structures upon which models may be based.

The process of humanization involves grafting of murine CDRs onto human framework regions (Jones *et al*., 1986). Further modification of framework residues may be required to restore the binding affinity of the mouse antibody (Riechmann *et al*., 1988). The Adair patent (Adair *et al*., 1999) includes V_H/V_L interface residues as one of the classes of residues that may need to match their murine counterparts in order to preserve the topography of the paratope. However, guidance on precisely which residues are likely to have the greatest influence is limited.

Thus this work has two main applications: (i) in antibody modelling in order to predict the correct packing angle and (ii) to identify the key interface residues that are important in determining the packing angle in order to improve antibody humanization.

Materials and methods

In summary, we built a data set of antibody *Fv* regions and analysed the distribution of V_H/V_L packing angles. We then used artificial neural networks to predict the packing angle from the interface residues and performed feature selection using genetic algorithms (GA) to select the most informative sets of interface residues.

Preparation of the data set

A list of *Fv* and *Fab* structures was extracted from the SACS XML file (Allcorn and Martin, 2002). This yielded a set of 561 antibody structures including 6 anti-idiotype antibodies (PDB codes: 1cic, 1dvf, 1iai, 1pg7, 1qfw and 2dtg), which each consist of two interacting antibodies, and thus were split into two files. The final data set consisted of 567 antibody structures. This set comprised 314 structures for which the sequences of the light chain and heavy chain were distinct. Conformational changes in the antibody CDRs upon binding
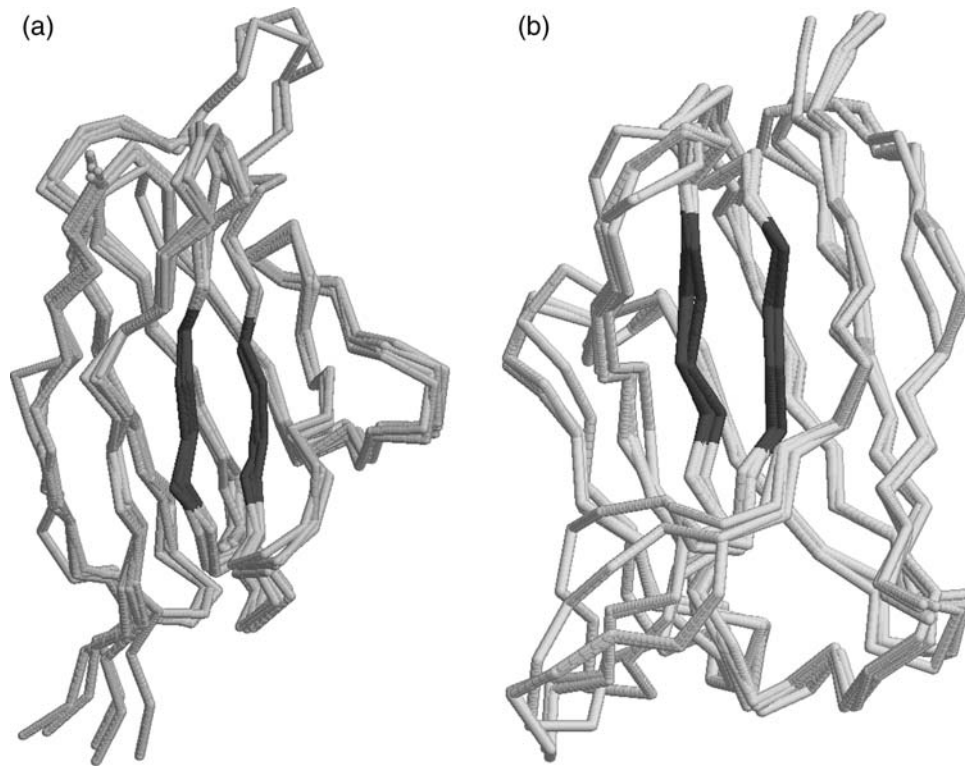
Fig. 1. Rigid body superposition of the Cα atoms in five antibody *Fv* structures. (**a**) Light chains highlighting residues L35–L38 and L85–L88 (PDB codes: 12e8, 15c8, 1a0q, 1a3l, 1a3r). (**b**) Heavy chains highlighting H36–H39 and H89–H92 (PDB codes 1oax, 1yec, 1yef, 2ddq, 8fab).

with the antigen have been observed in several studies (Colman *et al*., 1987; Bhat *et al*., 1990; Herron *et al*., 1991; Rini *et al*., 1992; Wilson and Stanfield, 1994; Mylvaganam *et al*., 1998). Redundancy was retained in the data set to allow for variability in a given structure. Thus, identical interface residues may result in different packing angles and it is important that we allow for this. Structural fitting of antibodies was performed using ProFit (http://www.bioinf.org. uk/software/profit/), which implements the McLachlan algorithm (McLachlan, 1982). AbNum (Abhinandan and Martin, 2008) was used to apply Chothia numbering (Chothia and Lesk, 1987; Al-Lazikani *et al*., 1997) to the PDB files of antibody structures. All structures were reduced to just the *Fv* region.

Potential interface residues were defined as positions for which there is any change in solvent accessibility as a result of V_H/V_L interaction. Accessibility was calculated using NACCESS (Simon Hubbard, unpublished), which implements the algorithm of Lee and Richards (1971).

Programs for analysis were written in C and Perl. Graphs were created using GNUPLOT (http://www.gnuplot.info/) and GRACE (http://plasma-gate.weizmann.ac.il/Grace/). The GRASS library (Team, 2006) was used for calculation of Eigen vectors and values. The Sun Gridengine was used to distribute jobs across two compute farms consisting of 96 IBM Series 335 nodes and 120 AMD Opteron compute cores, respectively.

Calculation of the packing angle

The packing angle was defined as a torsion angle at the V_H/V_L interface calculated as follows:

(i) Sets of eight structurally conserved residues at the V_H/V_L interface were identified by fitting five antibody light and heavy chains on all residues in the variable region using ProFit. Figure 1 shows the fitted structures and highlights the highly conserved residues, which form part of the β-sheet core of the interface (L35–L38, L85–L88, H36–H39, H89–H92).
(ii) For a given structure, the Cα coordinates for the sets of residues (S_L and S_H) were extracted.
(iii) The centroid for each set was identified (C_L and C_H).
(iv) The best-fit line for each set (S_L or S_H), passing through the respective centroid (C_L and C_H) was calculated using principal components analysis.
(v) A point on each line, P_L and P_H, on the same side relative to the respective centroid, was identified
(vi) The packing angle was calculated as the torsion angle between the points P_L, C_L, C_H, and P_H.

Neural networks

The total number of variables N_v in a fully connected artificial neural network is:

$$N_v = (S_i \times S_h) + (S_h \times S_o), \qquad (1)$$

where S_h is the number of nodes in the hidden layer, S_o is the number of nodes in the output layer and S_i is the number of nodes in the input layer and is calculated as:

$$S_i = N_i \times S_e, \qquad (2)$$

where N_i is the number of inputs and S_e is the size of the encoding vector. As a general rule of thumb, one should aim

for a training set size of $3N_v$ although, in practice, smaller training sets are often used. If we were to use the total of 124 potential interface positions identified (see Results) and a standard encoding vector size of 20 to represent each amino acid, then $S_i = 2480$. If we use 10 hidden nodes and a single output node to represent the packing angle, then $N_v = 24\,810$. This far exceeds the number of structures available for training the network (567).

To reduce the size of the input space, the normal approach of using a 20-dimensional encoding vector (where each dimension is a value from a substitution matrix or is zero for 19 of the values and one for the other value to represent a particular amino acid) was replaced by a four-dimensional vector describing physico-chemical properties of the amino acids: (i) the total number of side-chain atoms; (ii) the number of side-chain atoms in the shortest path from the Cα atom to the most distal atom; (iii) the Eisenberg consensus hydrophobicity (Eisenberg et al., 1982); (iv) the charge (histidine was assigned a charge of $+0.5$).

A fully connected artificial neural network using an input layer of 20×4 nodes, a single hidden layer of 10 nodes and an output layer consisting of a single real-valued node was constructed using the Stuggart Neural Network Simulator (SNNS, http://www.ra.cs.uni-tuebingen.de/SNNS/). The Resilient Backpropagation (RProp) learning function (Riedmiller and Braun, 1993) was used for training the network. RProp is a modification to the standard back-propagation algorithm, which implements dynamic learning-rate constants, and has been shown to be superior to other learning algorithms in terms of both speed and quality of learning (Schiffmann et al., 1993). Training used early stopping after 150 cycles, or a sum-of-square error ≤ 1.5. Magnitude pruning and input shuffling were also found to be beneficial.

Output values (packing angles) were scaled to the range of $0-1$ using:

$$\theta_f = 1 - \left(\frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}} \right), \quad (3)$$

where θ_f is the interface angle fraction, θ is the interface angle, $\theta_{\max}$ and $\theta_{\min}$ are the maximum and minimum observed interface angles.

Genetic algorithms

GAs use a population of individuals and iteratively repeat three steps: (i) evaluation and selection of the fittest individuals as parents for the next generation, (ii) crossover of two selected parents, (iii) mutation to make random changes to alleles in the offspring.

Individuals in the population consisted of binary vectors of length 124 or 64 (representing the total number of potential interface positions or the framework positions, respectively). Each allele was either 1 or 0 to indicate whether the interface position is included in training the neural network. Any individual was only allowed to have a maximum of 20 alleles set to 1 at a time.

Initially a random population of individuals was created and the quality of each individual was assessed using the result of training and validating a neural network using 5-fold cross-validation. A new population was generated by cross-over of high-scoring individuals followed by random mutation at a specified rate (μ) with a default of $\mu = 0.0001$. Evaluation and generation of new populations were repeated as required.

A 'generational replacement' strategy was used in which the entire population of parents was replaced by children to allow rapid exploration of the interface position space. The best individual from every generation was recorded in case the fitness of the best individual decreased in future generations.

Two of the most common strategies for selecting parents were evaluated: Roulette-wheel selection and Rank-based selection. Roulette wheel selection is a fitness-proportionate selection method where the likelihood of a particular parent being selected is given by the fitness of the parent divided by the average fitness of the population. In rank-based selection (Baker, 1985) the population is ranked by fitness, with selection performed in a similar way to the Roulette-wheel scheme, but with absolute scores replaced by ranks. In either case, a crossover point is chosen randomly within the selected parents and the two parts of the parents are combined to yield offspring.

One problem with these schemes is premature convergence of the population. Initially, the population is quite diverse, but parents that score significantly better than others are selected more frequently and can therefore result in identical children. When the number of redundant individuals in the population increases, the chances of choosing two identical individuals randomly for crossover also increases. Crossover of identical individuals would clearly yield a child identical to the parents. Since the mutation rate applied to the offspring individual is very low ($\mu = 0.0001$), the final offspring are likely to be unchanged. However, a higher mutation rate ($\mu = 0.001$) did not help curb the exponential rise in the number of redundant individuals (see results, Fig. 5).

We therefore developed a method in which parent individuals were selected using rank-based selection, but the mutation rate was varied dynamically, depending on the similarity of the parents selected for crossover. The method, which we term intelligent selection, is described below:

(i) For every child individual to be created, select 2 parents P$_1$ and P$_2$ based on rank-based selection.
(ii) Randomly choose a crossover point and splice P$_1$ and P$_2$ to create a child O_i.
(iii) Calculate the degree of similarity $S_{(P1,P2)}$ between parents P$_1$ and P$_2$:

$$S_{(P1,P2)} = 2 \times \frac{|P_1 \cap P_2|}{|P_1| + |P_2|}, \quad (4)$$

where $|P_1 \cap P_2|$ is the number of active alleles (i.e. with value = 1) shared between P$_1$ and P$_2$ while $|P_i|$ is the number of active alleles in P$_1$ or P$_2$. When the two parents are completely identical, the similarity is 1.0 whereas when they have no common alleles, the similarity is 0.

(iv) if ($0.9 \leq S_{(P1,P2)} \leq 1.0$), then swap five 0s and 1s in O_i
(v) if ($0.7 \leq S_{(P1,P2)} < 0.9$), then $\sigma = 0.01$
(vi) if ($0.5 \leq S_{(P1,P2)} < 0.7$), then $\sigma = 0.008$
(vii) if ($0.3 \leq S_{(P1,P2)} < 0.5$), then $\sigma = 0.005$
(viii) if ($0 \leq S_{(P1,P2)} < 0.15$), then $\sigma = 0.001$

Evaluation of prediction performance

During development and training of the GA, the performance of the neural network was assessed using 5-fold cross validation and the overall performance was averaged over the five folds. The final performance was evaluated using full leave-one-out jack-knifing.

Initially the Pearson's correlation coefficient (r) was used to compare the output of the neural network and the actual scaled packing angle (between 0 and 1):

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y},$$ (5)

where r_{xy} is the Pearson's correlation coefficient between two sets of variables x and y, n is the number of data points and s_x and s_y are the standard deviations of the two distributions x and y.
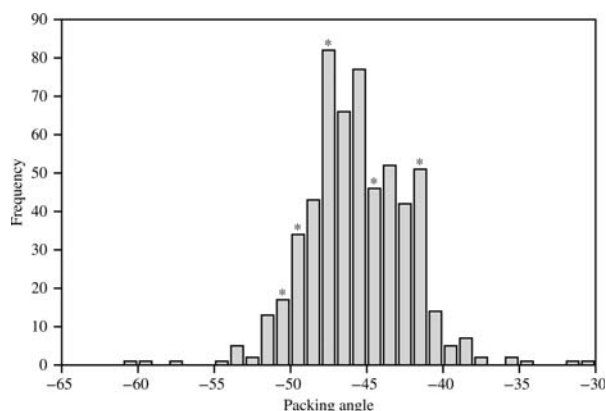

Fig. 2. Frequency distribution of the packing angle. Bars with an asterisk at the top indicate the presence of one or more of the eight single-chain Fv structures in our data set.

However, Pearson's r is not a very intuitive measure of the actual performance of the neural network in terms of prediction accuracy and does not reflect the presence of outliers very well. The relative RMS error (RELRMSE) (Masters, 1993) calculates the RMS value of the error and takes the ratio of this value with respect to the sum of the squares of the actual values:

$$RELRMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - p_i)^2}{\sum_{i=1}^{n} x_i^2}},$$ (6)

where x_i is the actual interface angle fraction and p_i is the predicted interface angle fraction. Since RELRMSE is a ratio, it is a dimensionless value. RELRMSE is calculated over 5-folds for every individual and the score for an individual is calculated as:

$$SCORE = 1 - RELRMSE.$$ (7)

From initial performance statistics (data not shown), it appeared that RELRMSE was much more sensitive to errors in predictions of small and large packing angles than the Pearson's r or other measures such as RMSE:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(x_i - p_i)^2}{n}}.$$ (8)

RELRMSE was therefore used to assess the quality of prediction in the main GA runs.

Results

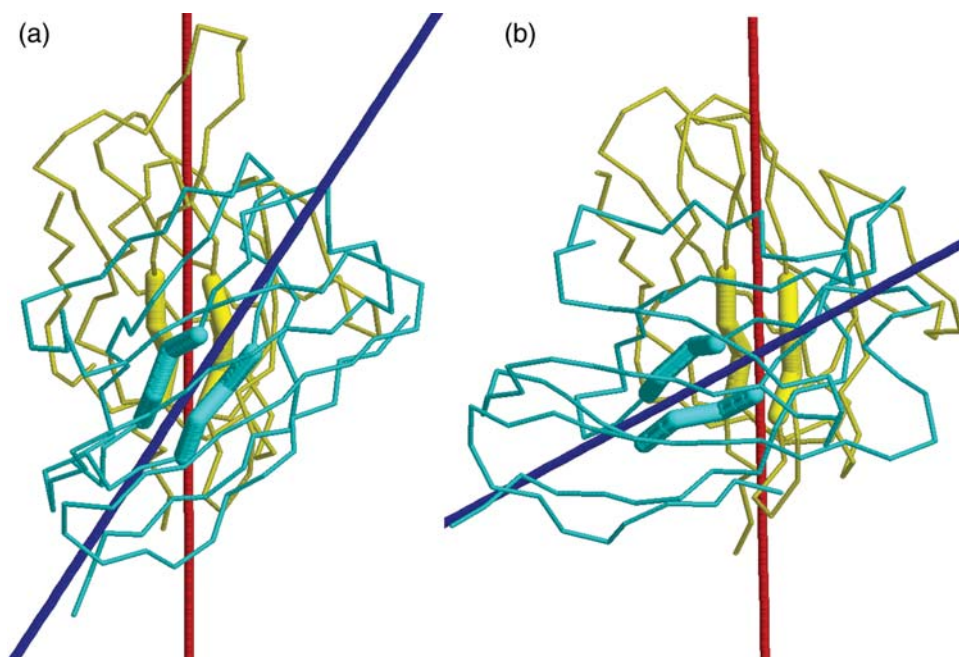A data set of 567 PDB files of antibody Fv regions was prepared and prenumbered using the Chothia numbering

Fig. 3. Extreme packing angles in (**a**) 1fl3, $-31.0°$ and (**b**) 1bgx, $-60.8°$. The images show the light chain (in yellow) in approximately the same orientation, with the heavy chain shown in cyan. The conserved residues used to define the packing angle are shown with thicker lines. The regression lines, fitted though these coordinates, are shown in red for the light chain and blue for the heavy chain.
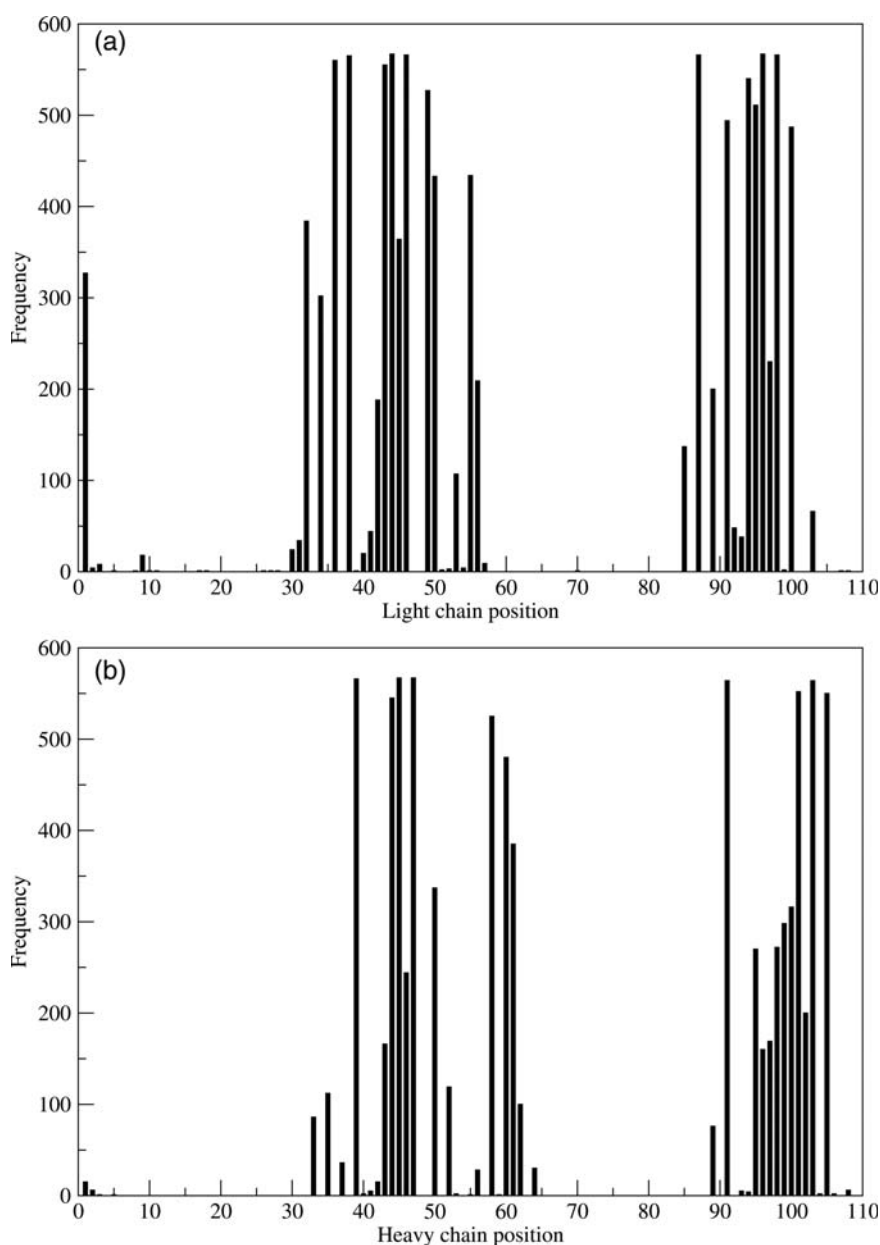
Fig. 4. Frequency of occurrence of residues in the interface (**a**) light chain and (**b**) heavy chain.

scheme. The packing angle was defined and calculated for each structure as described in the Materials and Methods. Potential interface residue amino acid types and packing angles were extracted and tabulated.

Distribution of packing angles

The distribution of packing angles was plotted and was found to follow an approximately normal distribution with a mean of $-45.6°$ and a standard deviation of $3.36°$ (Fig. 2). The observed packing angle varies quite considerably across different structures ranging from $-31.0°$ to $-60.8°$ in 1fl3 (Simeonov *et al.*, 2000) and 1bgx (Murali *et al.*, 1998), respectively, as shown in Fig. 3. We had expected that one or both of these might come from single-chain *Fv* fragments (which may have different constraints on packing), but this proved not to be the case and Fig. 2 shows that the eight

single-chain *Fv* fragments in our data set have evenly distributed packing angles.

Initial prediction of packing angle from interface residues

On the basis of a change in solvent accessibility, a total of 124 Chothia-numbered amino acid positions (63 in the light chain and 61 in the heavy chain) were identified as contributing to the interface in at least one of the 567 structures.

As described in the Materials and methods, a process of 'feature selection' was required to choose a subset of these 124 potential interface positions in order to train the neural network. Initially, manual selections of 20 interface residues (10 light and 10 heavy) most likely to influence packing angle were made based on an analysis of the change in solvent accessible surface area and the frequency of occurrence in the interface (Fig. 4). Owing to the variability in the V_H/V_L packing angle, the interface residues in any given

Table I. Manually chosen interface positions and performance of neural nets trained on those positions

Method[a]	Interface positions	C_p
Method I	L34, L36, L44, L46, L50, L87, L89, L91, L96, L98, H35, H47, H91, H100B, H100C, H100D, H100I, H100G, H100M, H103	0.32
Method II	L34, L36, L43, L44, L46, L86, L87, L89, L91, L98, H35, H47, H91, H100B, H100C, H100D, H100G, H100I, H100M, H103	0.38
Method III	L32, L34, L36, L44, L46, L50, L87, L91, L96, L98, H45, H47, H50, H91, H99, H100, H100A, H100B, H101, H103	0.40
Method IV	L34, L36, L38, L43, L44, L46, L87, L91, L96, L98, H39, H45, H47, H91, H99, H100, H100A, H100B, H101, H103	0.30

[a]Interface residues were chosen as the top 10 light and heavy chain residues using four methods, see text. C_p is the averaged Pearson's correlation coefficient (*r*) over 5-fold cross-validation.

structure will be a subset of the total set identified in all the observed structures. Table I shows the four manual selections of interface residues on the basis of: (i) highest change in solvent accessible surface area (ASA), (ii) highest average change in ASA, (iii) most frequently occurring positions with highest change in ASA and (iv) most frequently occurring positions with highest average change in ASA.

Table I also shows the result of training and validating the neural network based on the manual selection of interface positions. None of the methods to select interface residues manually worked particularly well as the 5-fold cross-validated Pearson's correlation coefficient (*r*) for all methods was low. However, from examination of individual folds, correlation coefficients as high as 0.6 were observed giving us confidence that the data did have useful predictive power. Thus, a new strategy was devised in order to improve the training feature selection.

Using a GA for feature selection

Rather than selecting interface residues for use in predicting the packing angle based on accessibility and occurrence in the interface, it was decided to use a GA to perform feature selection. The GA was designed to select a maximum of 20 interface positions that were optimal in training the neural network.

In generating an offspring population from parents, selection is biased towards parents with high scores. There are many selection methods for choosing parents, the aim of the selection procedure being to keep the population diverse in order to avoid local minima, while achieving progression towards a global minimum in a reasonable number of generations. We used small test populations over limited iterations to compare two approaches: rank-based and Roulette-wheel based selection (see Materials and methods). For this application, rank-based selection was found to out-perform Roulette-wheel selection (results not shown).

However, we found that the population tended to converge quite rapidly even with higher mutation rates. Figure 5 shows a high degree of convergence after around 40 generations with a mutation rate (σ) of 0.0001 (population of 5000) where the best individual had a Pearson's *r* of 0.638. With a mutation rate of 0.001 (population of 1000), the population converged after around 60 generations.

We implemented a modification of rank-based selection which we term 'intelligent selection' (see Materials and methods), which was used for all future GA runs. This alters the mutation rate dynamically during crossover.

By varying the mutation rate using our intelligent selection procedure, it became possible to keep the population diverse avoiding local minima and therefore sampling many different
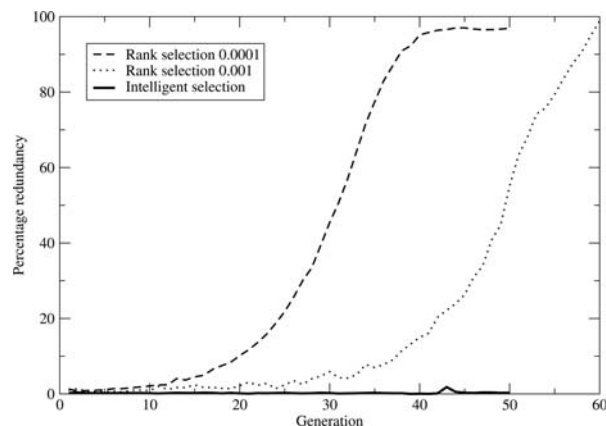

Fig. 5. Percentage of redundant individuals in GA runs using rank-based and intelligent selection. The dashed line shows rank-based selection with 5000 individuals and σ = 0.0001; the dotted line shows rank-based selection with 1000 individuals and σ = 0.001; the solid line shows intelligent selection with 5000 individuals. Note that the intelligent selection strategy results in virtually no redundancy in the population.

combinations in the possible 'interface position space'. Figure 5 also shows the results for intelligent selection for the GA using a population of 5000 individuals over a limited run of 50 generations. At the end of this period, while rank-based selection showed a very high degree of redundancy, intelligent selection showed almost no convergence. The best individuals had a Pearson's *r* of 0.638 and 0.630 in rank-based and intelligent selection, respectively. Thus, in this limited number of generations, intelligent selection was able to find as good a best solution as rank-based selection while maintaining a diverse population to avoid local minima. Intelligent selection was used for all further GA runs.

Results of the main GA runs

After the above preliminary experiments designed to optimize parameters for the neural network and the GA, large-scale GA runs were performed using thousands of individuals over several thousand generations. Scoring of each individual in the GA involves training and 5-fold cross-validation of a neural network and typically took about 25 s per individual. The runs were performed on large compute farms over a period of several months.

Two types of runs were performed. In the first, all 124 potential interface positions were considered. This included a number of CDR residues as it was initially felt that these could influence the V_H/V_L packing angle. In the second, only the 64 framework residues were considered. For this purpose, the CDRs were defined as the structurally variable regions as used by Chothia (Al-Lazikani *et al*., 1997).
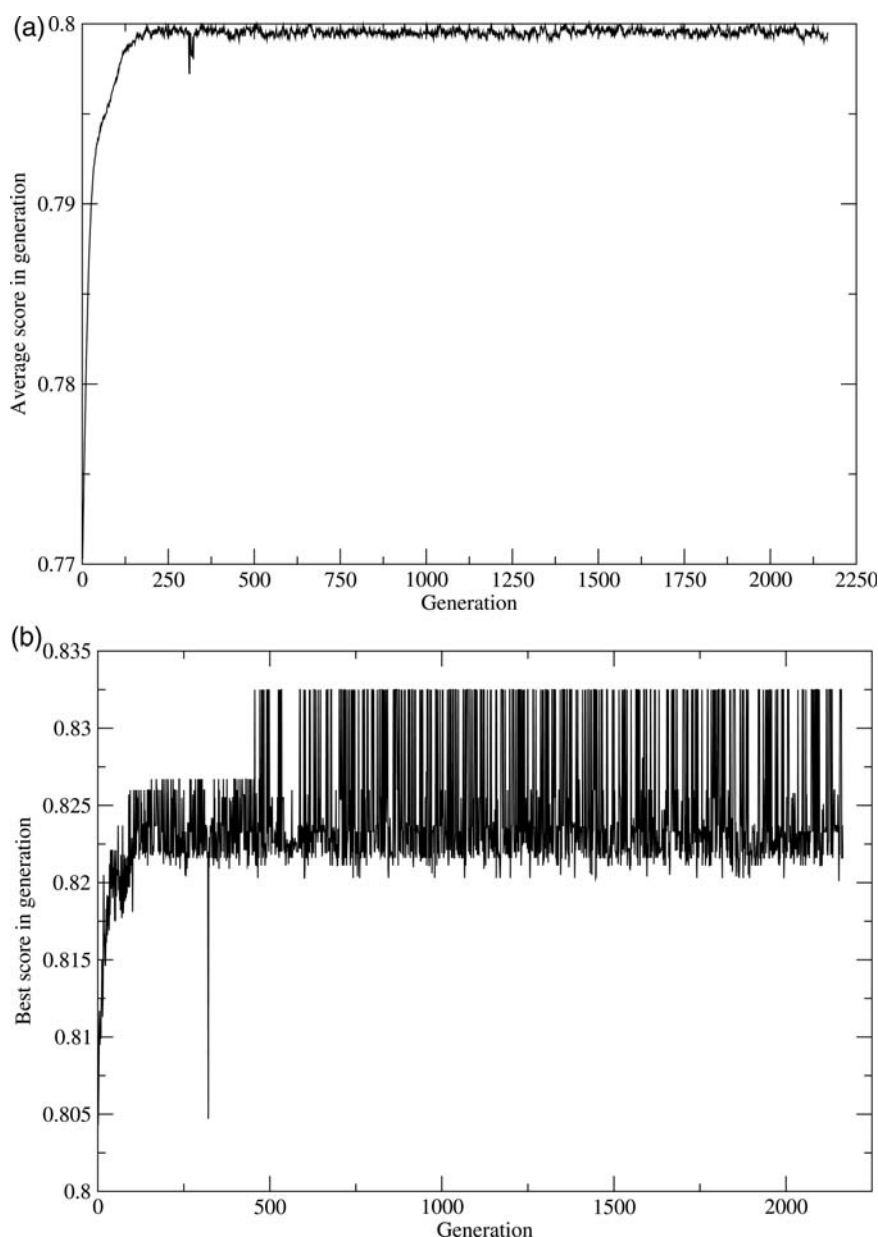
Fig. 6. GA runs involving non-CDR interface positions. (**a**) Average score in each generation and (**b**) best score in each generation.

Overall performance was better when only framework residues were considered suggesting that CDR residues have only a small influence on V_H/V_L packing (data not shown). Consequently only results using the framework residues alone are reported here.

A GA run was performed using a population of 15 000 individuals for a total of 2166 generations over a wall-clock period of approximately 4 months. Results were assessed on the basis of (i) the score of the best individual at the end of each generation and (ii) the average score of individuals in each generation.

Results of the run are shown in Fig. 6. The average and best scores increase sharply for the first 150 generations and then stabilize for the remaining generations. The best score of 0.833 (i.e. RELRMSE = 0.167, see Materials and methods) was first seen after 146 generations. The interface positions defined by the best individual were: L38, L40, L41, L44, L46, L87, H33, H42, H45, H60, H62, H91 and H105.

Jackknifing and analysis of errors of the best individuals

Having identified the optimum set of interface residues using the GA, a full jackknifed (leave-one-out) evaluation of the neural network was performed. Results are shown in Fig. 7 where the predicted packing angle is plotted against the actual packing angle. Perfect predictions would lie on the diagonal. Figure 8 shows the distribution of the errors and shows an approximately normal distribution with a strong peak at an error of $0°$. Figure 9 shows the squared error plotted against the actual packing angle indicating that the most significant errors are for the outlying structures, which have unusually large (or small) packing angles and are under-represented in the training data (see Fig. 2).

Discussion

In this paper, we have defined and analysed the V_H/V_L packing angle across a panel of 567 antibody structures. The
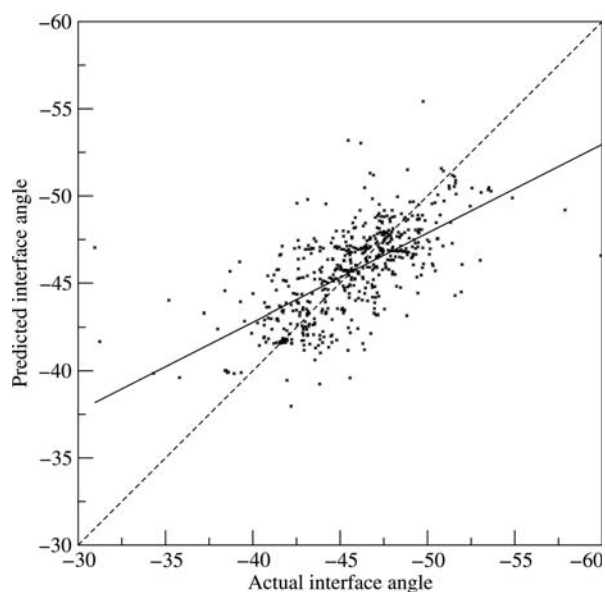
Fig. 7. Predicted vs. actual packing angle results for jackknifing of the best individual from the GA run. Perfect predictions would lie on the dashed line. The solid line shows the best-fit regression line for the data points, although it should be noted that the errors are not evenly distributed such that regression is not very accurate.
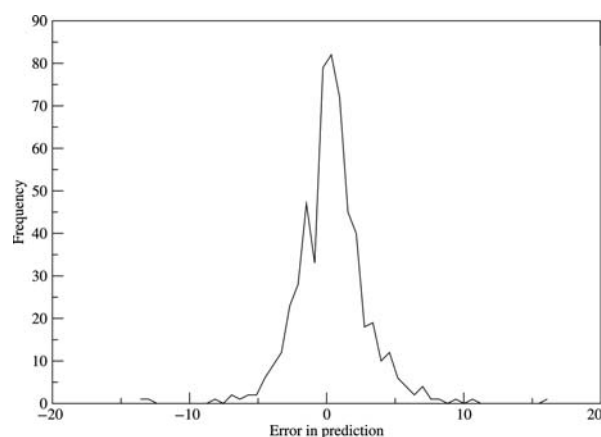

Fig. 8. Distribution of errors calculated as the difference between the predicted and actual interface angle for the best individual from the GA run.
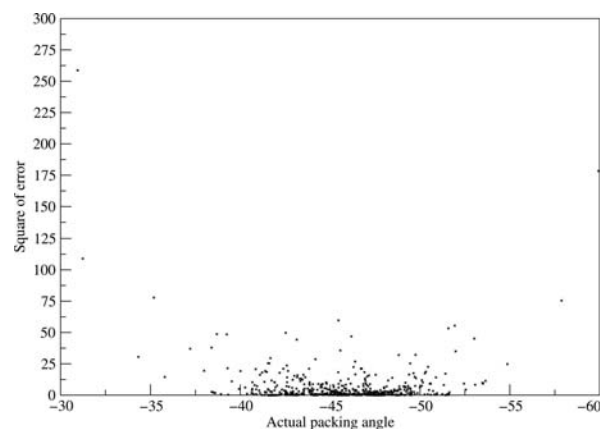

Fig. 9. Squared error in predicted packing angle against actual packing angle.

packing angle is approximately normally distributed with a mean of $-45.6°$, but ranges from $-60.8°$ to $-31.0°$. Such a variation can have a significant effect on the topography of the combining site.

Using a GA for feature selection, we have identified a set of interface residues, which appear to have the largest predictive power and are therefore most likely to influence the packing angle. In antibody humanization, where a set of non-human (typically murine) CDRs are grafted onto a human framework region, it is generally necessary to ensure that some additional human framework residues match the donor mouse residues in order to recover good binding. It had previously been suggested that these residues may include interface residues (Adair *et al.*, 1999) and this work suggests a key set of such positions, which may prove important.

Using the identified set of interface residues, we have performed a full jack-knifed analysis of the predictive ability of a neural network trained using these locations and a physico-chemical representation of the amino acids at those locations. The results show an approximately normal distribution of errors centred around $0°$. The RELRMSE on the packing angle is just 0.056. It is well known that neural networks make poorer predictions on data that are sparsely represented. This appears to be the case for predicting packing angles that are less than $-50°$ or greater than $-43°$ (Fig. 9). Application of boosting techniques (Haykin, 1994) or simple multiple presentation of outliers to the neural network may improve performance.

A web server, which takes a light and heavy chain sequence and implements the prediction method described, has been made available at http://www.bioinf.org.uk/abs/paps/. The neural network used in the final server was trained on the complete data set and results are similar to those obtained from the full jack-knifed evaluation.

It is hard to make a direct comparison of the performance of our method with the energy-based methods described by Sivasubramanian *et al.* (2009) or Narayanan *et al.* (2009) since these report performance in terms of the RMSD of one domain when fitting is performed on the other domain. We feel that our method gives a more direct measure of the results, which is not dependent on the similarity of the domains and hence the quality of fitting. Because predictions made using our trained machine learning method are very fast (taking a matter of seconds), the evaluation of our method is much more extensive than the work presented in these earlier papers.

The applications of this work are 2-fold. First, the set of interface residues may be useful in humanization work. Second, our ability to predict the packing angle can be used in improving antibody modelling: either by improving the selection of a framework consisting of both light and heavy chains or in imposing a packing angle on individually selected light and heavy chains.

References

Abhinandan,K.R. and Martin,A.C.R. (2008) *Mol. Immunol.*, **45**, 3832–3839.

Adair,J.R., Athwal,D.S. and Emtage,J.S. (1999) *Humanised Antibodies*. US patent 5,859,205.

Al-Lazikani,B., Lesk,A.M. and Chothia,C. (1997) *J. Mol. Biol.*, **273**, 927–948.

Allcorn,L.C. and Martin,A.C.R. (2002) *Bioinformatics*, **18**, 175–181.

Baker,J.E. (1985) Adaptive Selection Methods for Genetic Algorithms. In Grefenstette,J.J. (ed.), *Proceedings of the First International Conference on Genetic Algorithms*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., pp. 101–111.

Bhat,T.N., Bentley,G.A., Fischmann,T.O., Boulot,G. and Poljak,R.J. (1990) *Nature*, **347**, 483–485.

Chatellier,J., Van Regenmortel,M.H., Vernet,T. and Altschuh,D. (1996) *J. Mol. Biol.*, **264**, 1–6.

Chothia,C. and Lesk,A.M. (1987) *J. Mol. Biol.*, **196**, 901–917.

Chothia,C., Novotný,J., Bruccoleri,R. and Karplus,M. (1985) *J. Mol. Biol.*, **186**, 651–663.

Colman,P.M., Laver,W.G., Varghese,J.N., Baker,A.T., Tulloch,P.A., Air,G.M. and Webster,R.G. (1987) *Nature (London)*, **326**, 358–363.

Eisenberg,D., Weiss,R., Terwilliger,T. and Wilcox,W. (1982) *Faraday Symposia of the Chemical Society*, **17**, 109–120.

Haykin,S. (1994) *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR Upper Saddle River, NJ, USA.

Herron,J.N., He,X.M., Ballard,D.W., Blier,P.R., Pace,P.E., Bothwell,A.L., Voss,E.W. and Edmundson,A.B. (1991) *Proteins*, **11**, 159–175.

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

Lee,B. and Richards,F.M. (1971) *J. Mol. Biol.*, **55**, 379–400.

Martin,A.C., Cheetham,J.C. and Rees,A.R. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9268–9272.

Martin,A.C., Cheetham,J.C. and Rees,A.R. (1991) *Meth. Enzymol.*, **203**, 121–153.

Masters,T. (1993) *Practical Neural Network Recipes in C++*. San Francisco: Morgan Kaufmann.

McLachlan,A. (1982) *Acta Crystallogr. A*, **38**, 871–873.

Murali,R., Sharkey,D.J., Daiss,J.L. and Murthy,H.M. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 12562–12567.

Mylvaganam,S.E., Paterson,Y. and Getzoff,E.D. (1998) *J. Mol. Biol.*, **281**, 301–322.

Narayanan,A., Sellers,B.D. and Jacobson,M.P. (2009) *J. Mol. Biol.*, **388**, 941–953.

Novotný,J. and Haber,E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4592–4596.

Poljak,R.J., Amzel,L.M., Avey,H.P., Chen,B.L., Phizackerley,R.P. and Saul,F. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 3305–3310.

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

Riedmiller,M. and Braun,H. (1993) *A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm*. Proceedings of the ICNN, San Francisco, http://citeseer.comp.nus.edu.sg/186821.html.

Rini,J.M., Schulze-Gahmen,U. and Wilson,I.A. (1992) *Science*, **255**, 959–965.

Roguska,M.A., *et al*. (1996) *Prot. Eng.*, **9**, 895–904.

Schiffmann,W., Joost,M. and Werner,R. (1993) *Proceedings of the European Symposium on Artificial Neural Networks, ESANN*, **93**, 97–104.

Simeonov,A., *et al*. (2000) *Science*, **290**, 307–313.

Sircar,A., Kim,E.T. and Gray,J.J. (2009) *Nucleic Acids Res.*, **37**, W474–W479.

Sivasubramanian,A., Sircar,A., Chaudhury,S. and Gray,J.J. (2009) *Proteins*, **74**, 497–514.

Team,G.D. (2006) *Geographic Resources Analysis Support System (GRASS) Software*. ITC-irst, Trento, Italy. http://grass.itc.it/.

Whitelegg,N. and Rees,A. (2000) *Prot. Eng. Des. Sel.*, **13**, 819–824.

Wilson,I.A. and Stanfield,R.L. (1994) *Curr. Opin. Struct. Biol.*, **4**, 857–867.