Structure, Function and Properties of Antibody Binding Sites

I. Saira Mian[1]†, Arthur R. Bradwell[2] and Arthur J. Olson[1]‡

[1]*Department of Molecular Biology*
Research Institute of Scripps Clinic
10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

[2]*Department of Immunology*
Medical School, Birmingham University, B15 2TJ, U.K.

Do antibody combining sites possess general properties that enable them to bind different antigens with varying affinities and to bind novel antigens? Here, we address this question by examining the physical and chemical characteristics most favourable for residues involved in antigen accommodation and binding. Amphipathic amino acids could readily tolerate the change of environment from hydrophilic to hydrophobic that occurs upon antibody–antigen complex formation. Residues that are large and can participate in a wide variety of van der Waals' and electrostatic interactions would permit binding to a range of antigens. Amino acids with flexible side-chains could generate a structurally plastic region, i.e. a binding site possessing the ability to mould itself around the antigen to improve complementarity of the interacting surfaces. Hence, antibodies could bind to an array of novel antigens using a limited set of residues interspersed with more unique residues to which greater binding specificity can be attributed. An individual antibody molecule could thus be cross-reactive and have the capacity to bind structurally similar ligands. The accommodation of variations in antigenic structure by modest combining site flexibility could make an important contribution to immune defence by allowing antibody binding to distinct but closely related pathogens.

Tyr and Trp most readily fulfil these catholic physicochemical requirements and thus would be expected to be common in combining sites on theoretical grounds. Experimental support for this comes from three sources, (1) the high frequency of participation by these amino acids in the antigen binding observed in six crystallographically determined antibody–antigen complexes, (2) their frequent occurrence in the putative binding regions of antibodies as determined from structural and sequence data and (3) the potential for movement of their side-chains in known antibody binding sites and model systems. The six bound antigens comprise two small different haptens, non-overlapping regions of the same large protein and a 19 amino acid residue peptide. Out of a total of 85 complementarity determining region positions, only 37 locations (plus 3 framework) are directly involved in antigen interaction. Of these, light chain residue 91 is utilized by all the complexes examined, whilst light chain 32, light chain 96 and heavy chain 33 are employed by five out of the six. The binding sites in known antibody–antigen complexes as well as the postulated combining sites in free Fab fragments show similar characteristics with regard to the types of amino acids present. The possible role of other amino acids is also assessed. Potential implications for the combining regions of class I major histocompatibility molecules and the rational design of molecules are discussed.

---

† Present address: Sinsheimer Laboratory, Biology Department, University of California Santa Cruz, Santa Cruz, CA 95064, U.S.A.
‡ Author to whom reprint requests should be addressed.
§ Abbreviations used: CDR, complementarity determining region; MHC, major histocompatibility

1. Introduction

Antibodies are powerful recognition and binding molecules that the immune system employs to eliminate foreign molecules. Antibody binding sites are formed by six hypervariable loops or comple-

three from each of the heavy and light chain variable domains, are connected to a relatively invariant β-sheet framework (Alzari *et al.*, 1988; Davies & Metzger, 1983; Capra & Edmundson, 1977; Wu & Kabat, 1970). Early analysis of a data bank of complete and partial sequences of 415 light and 197 heavy chains demonstrated that CDRs are rich in aromatic residues (Kabat *et al.*, 1977). The combining region represents only a small part of the antibody molecule, whose overall three-dimensional structure is highly conserved. Although, the pairing of light and heavy chains can generate some antibody diversity, most of it is generated by the somatic recombination of variable region gene segments (Yancopoulos & Alt, 1986; Wysocki & Gefter, 1989). Such genetic mechanisms yield antibodies exhibiting extensive diversity in hypervariable loop sequences. This potential repertoire is estimated to be approximately 10^9 in mouse (Berek *et al.*, 1985). However, the initial repertoire that confronts an antigenic challenge is smaller than the potential repertoire, since it is restricted to the antibody specificities expressed on existing immunocompetent B cells at a point in time (Holmberg *et al.*, 1986). This available repertoire can yield an apparently unlimited repertoire of antigen binding specificities and affinities.

Although a single antibody has a unique three-dimensional structure, biophysical and biochemical evidence indicates that it is multispecific or cross-reactive (Richards *et al.*, 1975). This capacity to combine both with its inducing antigen and with antigens of similar or disparate structure augments the genetically determined antigen-binding capabilities of antibodies. The extent of molecular complementarity between determinants on the antigen molecule and amino acid residues in the combining site determines the degree of antibody specificity. Increased cross-reactivity, therefore, is at the expense of specificity and affinity.

An improved understanding of both antibody cross-reactivity and binding can be obtained by a study of antibody–antigen interactions at the atomic level. The role of residues in the definition of combining site structure and interaction with antigen can be assessed as a function of the chemical and structural properties of individual amino acids. First, we examine those characteristics that appear to be of general importance in antibody–antigen interactions. This is followed by a detailed study of the binding sites in six antibody–antigen complexes and four free Fab fragments of known three-dimensional structure, and the much larger database of antibody sequences. Padlan (1990) has performed a similar, though not identical, analysis of antibody combining sites in general, and three anti-lysozyme antibody antigen complexes in particular. On the basis of their propensity to occur in the combining sites and their greater exposure relative to those in the framework regions, he has suggested that these amino acids determine specificity. Our results and their interpretation lead us to conclude that Tyr

binding and non-specific antibody–antigen interactions.

2. Physical and Chemical Properties of Amino Acids

Since antibody binding sites are formed by six hypervariable loops supported on a highly conserved β-sheet framework, there is likely to be a bias towards amino acids that are generally found in non-helical regions of proteins. Figure 1 shows the normalized frequencies of occurrence of amino acids in α-helix, β-sheet and reverse turns in 66 globular proteins comprising 31 different conformations (Levitt, 1978). In these structures, the occurrence of Pro, Gly, Tyr, Ser, Thr, Asn, Val, Arg, Ile and Trp in α-helices is less frequent than random. Leu, His, Trp, Thr, Tyr, Phe, Ile and Val have a greater than random probability of occurring in β-sheets; the same is true for Thr, Tyr, Asn, Ser, Asp, Gly and Pro in reverse turns. Arg appears to be equally tolerated in all the secondary structures elements considered. In general structural terms, Tyr and Thr seem to be the most useful non-helix forming residues, since they could be positioned in either the strand or turn regions of the hypervariable loops.

The free energy of interaction between an antibody and its antigen is a function of both enthalpy and entropy. Non-bonded forces between the interacting molecules include hydrophobic, hydrogen bond, van der Waals' and electrostatic interactions (for a review, see Fersht, 1985). In general terms, antibody combining site residues need to be as multifaceted as possible to accommodate the varied stereochemical and electronic features of the antigen. Hence, amino acids with non-polar (for example Leu, Ile and Val) and charged (for example Asp, Glu, Lys and Arg) side-chains would be of more limited usefulness than, for example, His, which is known to be capable of cross-linking sequentially distant but spatially close regions of proteins (Baker & Hubbard, 1984; I.S.M. & A.J.O., unpublished results). Similarly, the amides Asn and Gln would be generally more preferable than Asp and Glu, since the former pair are both hydrogen bond donors and acceptors whereas their charged counterparts are only acceptors.

If a positive charge is required in the antibody combining site, Arg would be more suitable than Lys because of its greater functional versatility; for example, Arg can form a larger number of hydrogen bonds than Lys. As a consequence of its planar nature and π-electron system, the terminal guanidinium group of Arg often exhibits pseudo-aromatic behaviour by participating in most of the interactions previously catalogued for true aromatic–aromatic interactions (I.S.M. & A.J.O., unpublished results). These interactions occur at the intersubunit interfaces of a number of oligomeric proteins, including viral coat proteins and a membrane protein; the photosynthetic reaction centre of *Rhodopseudomonas viridis*. The ability to
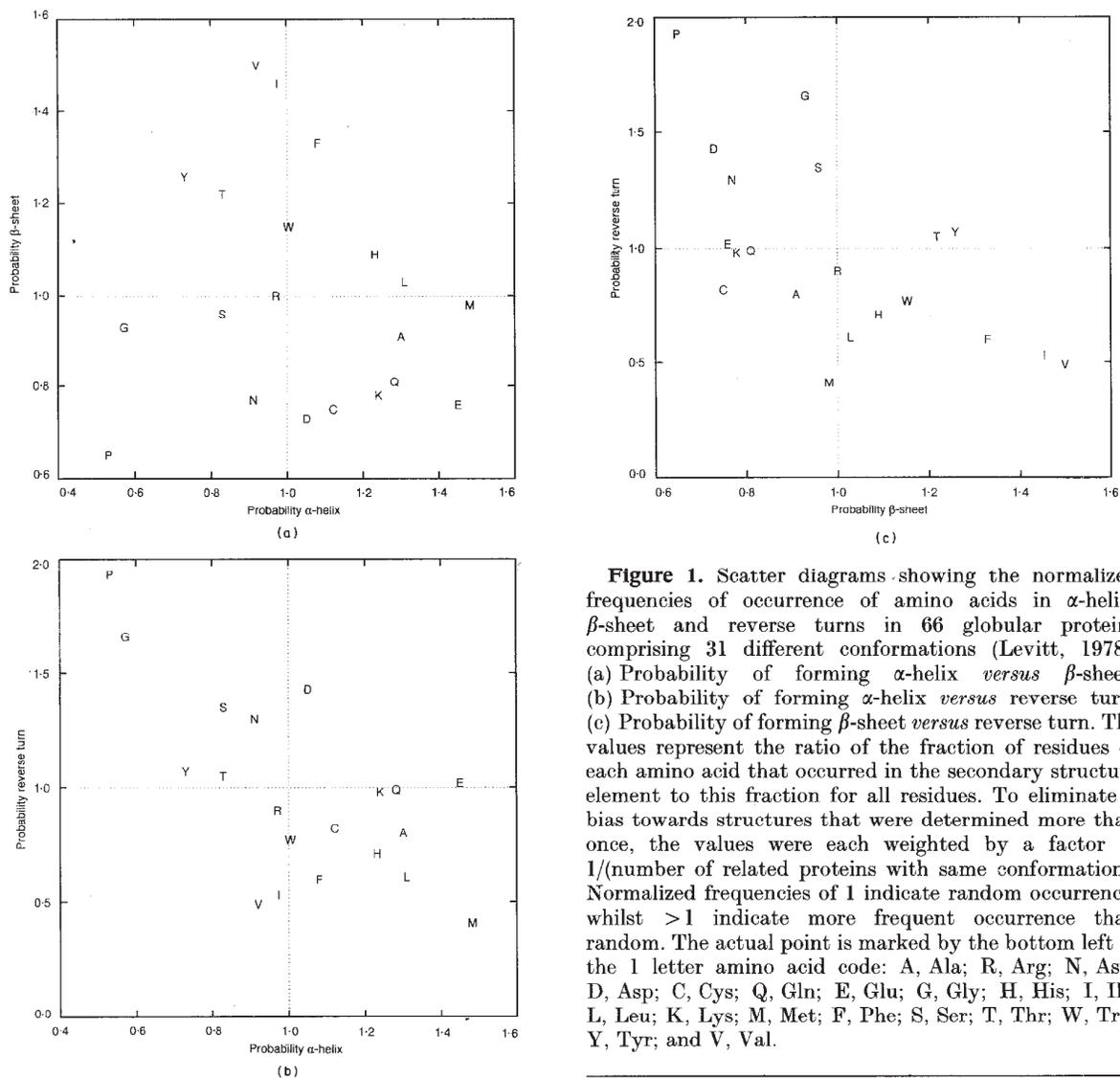
(a)

(b)

(c)

Figure 1. Scatter diagrams showing the normalized frequencies of occurrence of amino acids in α-helix, β-sheet and reverse turns in 66 globular proteins comprising 31 different conformations (Levitt, 1978). (a) Probability of forming α-helix *versus* β-sheet. (b) Probability of forming α-helix *versus* reverse turn. (c) Probability of forming β-sheet *versus* reverse turn. The values represent the ratio of the fraction of residues of each amino acid that occurred in the secondary structure element to this fraction for all residues. To eliminate a bias towards structures that were determined more than once, the values were each weighted by a factor of 1/(number of related proteins with same conformation). Normalized frequencies of 1 indicate random occurrence, whilst >1 indicate more frequent occurrence than random. The actual point is marked by the bottom left of the 1 letter amino acid code: A, Ala; R, Arg; N, Asn; D, Asp; C, Cys; Q, Gln; E, Glu; G, Gly; H, His; I, Ile; L, Leu; K, Lys; M, Met; F, Phe; S, Ser; T, Thr; W, Trp; Y, Tyr; and V, Val.

attractive electrostatic interactions between positively charged groups and aromatic rings permits Tyr and Trp to interact with structurally diverse antigens. Another functional advantage in locating Tyr and Trp in antibody combining sites is that, unlike amino acids having shorter side-chains, such as Asn and Ser, they lack the capacity to interact easily with other groups on the antibody surface but are ideally suited to interact with another molecule.

The accommodation of charged areas on the antigen need not necessitate an antibody combining site possessing amino acids of complementary charge. Analysis of Arg, Lys, Glu and Asp side-chains buried at the intermolecular interfaces of oligomeric systems indicates that oriented dipoles are usually preferred over countercharges in stabilizing these buried residues (I.S.M. & A.J.O., unpublished results). Thus, the peptide backbone and polar side-chains of hypervariable loop residues could be deployed to stabilize both negatively and

may be as effective as employment of formally charged amino acids: in cases of charge–charge interaction, the steric effects of neighbouring regions may prevent the formation of geometrically optimal ion-pairs such that the potentially available energy is not fully realized.

The non-covalent association between antibody and antigen requires the removal of water from surfaces buried by the interacting molecules. Antibody regions involved in this process should be capable of tolerating both the polar and non-polar environments that exist before complex formation and upon antigen binding, respectively. Individual residues exposed on the surface of the free antibody can become completely or partially buried in the complex. In addition to residue amphipathicity, residue size might be a factor. There is a good correlation between the surface area of amino acids and their free energies of transfer from water to an organic phase (Chothia, 1974, 1975; Gelles &
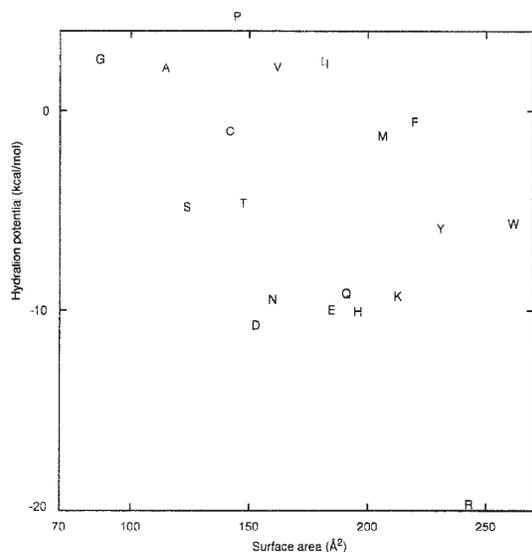
Figure 2. Comparison of the size of amino acids and the affinities of their side-chains for water. The surface area (Rose *et al.*, 1985) comprises the mean solvent accessibility for amino acid X in an ensemble of Gly-X-Gly tripeptides. The hydration potential (Wolfenden *et al.*, 1981) is the effective free energy of transfer from the vapour phase to dilute aqueous solution buffer at pH 7 of molecules having the structure R-H, where R is the side-chain of each amino acid; for P (Pro), only the surface area is indicated, since no hydration potential was evaluated. Side-chains were modelled by the following compounds: A, methane; R, methylguanidine; N, acetamide; D, acetic acid; C, methanethiol; Q, propionamide; E, propionic acid; G, H$_2$; H, 4-methylimidazole; I, isobutane; L, butane; K, *n*-butylamine; M, ethylmethyl sulphide; F, toluene; S, methanol; T, ethanol; W, 3-methylindole; Y, *p*-cresol; and V, propane. As a result of technical difficulties (Wolfenden *et al.*, 1981), methylguanidine (shorter than the side-chain of Arg by 2 methylene groups) was employed to estimate the value for propylguanidine; this probably leads to the hydrophilic and hydrophobic nature of Arg being over- and under-estimated, respectively.

Figure 3. A diagram illustrating specificity and cross-reactivity for a given antibody. The specific binding of the antibodies A and B to antigens A and B is a function of the high degree of complementarity between their molecular surfaces in terms of shape, size and functionality. (a) Cross-reactivity may arise as a result of structural similarity of epitopes between antigens A and A' (Richards *et al.*, 1975). A poor fit in one region may be compensated for by a good fit elsewhere. This could result in a sufficient number of short-range interactions to produce a stable antibody–antigen complex. Another cross-reacting antigen, A'', may fit the antibody combining site in a slightly different way. (b) The antibody B may accommodate the related antigens B and B' if it is able to vary the stereochemical features of the combining site, i.e. if it is intrinsically pliable.

surface area gives a hydrophobic energy of 25 cal/mol (1 cal = 4·184 J: Chothia. 1974). Whilst these van der Waals' energies may be small compared to a hydrogen bond, when summed over the entire combining site they may be important in stabilizing the complex. Figure 2 compares the affinities of amino acid side-chains for water (Wolfenden *et al.*, 1981) with the surface area of the entire amino acid (Rose *et al.*, 1985). The classical groupings into small, large, hydrophobic and hydrophilic amino acids are evident. With respect to amphipathicity, Ser, Thr, Tyr and Trp seem desirable residues to locate in antibody binding sites, since their side-chains are in the midrange of hydrogen potential values. The aromatic residues Tyr and Trp are also two of the largest and are capable of contributing significantly to the total interaction
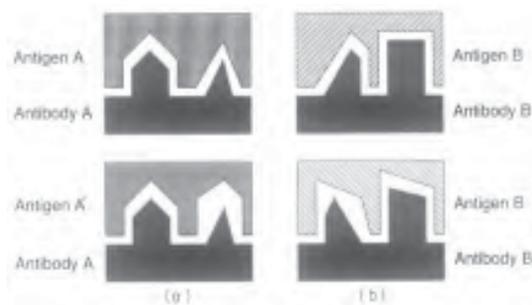
In an antibody–antigen complex, the stabilization energy gained from the various intermolecular forces must more than offset losses due to conformational entropy and conformational strain. The free energy of complex formation could therefore be maximized by minimizing the loss of conformational entropy upon association. It is known that a single antibody is able to combine with a spectrum of different antigens (Richards *et al.*, 1975). Although such cross-reaction may occur either because the antigens share epitopes, or because the epitopes are sufficiently similar in shape to bind the same antibody (Richards *et al.*, 1975), it could arise also if the topography of the combining site could be modulated (Fig. 3). Thus, antibodies could utilize amino acids whose side-chains were sufficiently structurally and functionally flexible to permit them to alter the stereochemical features of the combining site with minimal loss of entropy. The potential importance of side-chain motion has been further highlighted by a recent comparative study of known antibody structures and sequences (Chothia *et al.*, 1989). It has been suggested that the number of main-chain conformations of at least five of the six loops appears to be limited. The adoption of a specific backbone conformation is believed to be a reflection of only a few key conserved residues in the loop or framework of the antibody (Chothia *et al.*, 1989). This small repertoire of canonical structures would represent a reduction in the spectrum of specificity and affinity potentially available to the antibody binding site were the number of conformations proportional to the number of sequences that

Table 1
Preferred conformations of amino acid side-chains as described by their torsion parameters (Cody, 1985)

Residue	ψ	χ^{11}	χ^{12}	χ^{21}	χ^{22}	χ^3	χ^4	χ^5
				Torsion parameters (°)				
A	5; −19							
R	−15	60		180		180	±10	
		−60		180		180	±10	
N	11	60		0				
D	−7; ±35	−60		±5				
C	−30	−60						
		60						
Q	−20	60		180		−15		
		−60		180		25		
E	0; ±35	180		180		±15		
		−60		180		±15		
		60		180		±15		
G	0; 25							
H	−25	−60		60	−120			
		60		60	−120			
I	−15; −45	60	180	180				
		180	−60	180				
		−60	180	180				
L	−18; ±36	180		−60	180			
		−60		−60	180			
		60		60	180			
K	−20	−60		180		180	180	
M	±20	180		180		−60		
		−60		180		180		
		60		180		180		
F	±20	180		90				
		−60		90				
		60		90				
P	±10	±35		±35		±25	±20	±10
S	5	60						
T	−25	−60	180					
		60	−60					
W	−10	−60		90				
		60		90				
Y	±20	−60		90				
		180		90				
		60		90				
V	−11; −35	−60	180					

These were derived from crystallographic studies of amino acids and their derivatives. All atoms are numbered using the Greek letter designations starting with the C^α. $\psi = (O\text{–}C\text{–}C^\alpha\text{–}N)$, $\chi^{11} = (N\text{–}C^\alpha\text{–}C^\beta\text{–}C^{\gamma 1})$, $\chi^{12} = (N\text{–}C^\alpha\text{–}C^\beta\text{–}C^{\gamma 2})$, $\chi^{21} = (C^\alpha\text{–}C^\beta\text{–}C^\gamma\text{–}C^{\delta 1})$, $\chi^{22} = (C^\alpha\text{–}C^\beta\text{–}C^\gamma\text{–}C^{\delta 2})$, $\chi^3 = (C^\beta\text{–}C^\gamma\text{–}C^\delta\text{–}C^\varepsilon)$, $\chi^4 = (C^\gamma\text{–}C^\delta\text{–}C^\varepsilon\text{–}C^\zeta)$. To account for the ring pucker in Pro, $\chi^4 = (C^\gamma\text{–}C^\delta\text{–}N\text{–}C^\alpha)$ and $\chi^5 = (C^\delta\text{–}N\text{–}C^\alpha\text{–}C^\beta)$. The most frequently observed values are given first.

The number of degrees of freedom of each amino acid side-chain can be approximated by examining the range and distribution of the observed conformations (Table 1). Residues with short side-chains such as Ser and Thr lose little entropy when fixed upon antigen association, since they have only one and two variable side-chain torsion angles, respectively. However, they project only a short distance from the surface of the antibody and so could not effect substantial changes in binding site topography. Large residues can elicit great changes in the surface contours of the combining site, since they can sweep out large volumes of space. Residues with the largest surface area (Fig. 2) are Trp, Arg, Tyr and Phe and of these, Arg has twice the number of variable torsion angles of the other three. Additionally, Arg is less suitable than the other

fore requires a more restrictive interaction at the interface.

Tyr represents a balance between the many different, though sometimes conflicting, desirable aspects we believe to be of general importance in antibody–antigen interactions. Thus, it would be expected to be the most common residue in antibody combining sites. Experimental evidence from antibody sequences and structures appears to verify these assumptions.

3. Combining Sites in Known Antibody–Antigen Complexes

Table 2 lists the antibody residues that bind antigen in six complexes whose structures have been determined by X-ray crystallography. The

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming inter-face) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.

fastcase
Smarter legal research.