Anatomy of the antibody molecule: a continuing analysis based on high-resolution crystallographic structures

Jo Erika T. Narciso, Iris Diana C. Uy, April B. Cabang, Jenina Faye C. Chavez, Juan Lorenzo B. Pablo, Gisela P. Padilla-Concepcion*, Eduardo A. Padlan

The Marine Science Institute, University of the Philippines Diliman, Quezon City 1101, Philippines

Currently available high-resolution crystallographic studies of liganded and unliganded antibody molecules have provided the opportunity to analyze in more detail the structure of the antibody and its interaction with antigen, as well as the interactions between the domains of the molecule and between the framework and the complementarity-determining regions of the variable domains. The structural data now available have also allowed a more detailed analysis of the solvent accessibilities of the residues in the various domains of the molecule. The information resulting from this analysis is useful in the engineering of antibodies for therapeutic and other purposes.

KEYWORDS

antibody structure, high-resolution crystallographic studies, antibody-ligand complexes, solvent accessibilities, inter-residue contacts, antibody engineering

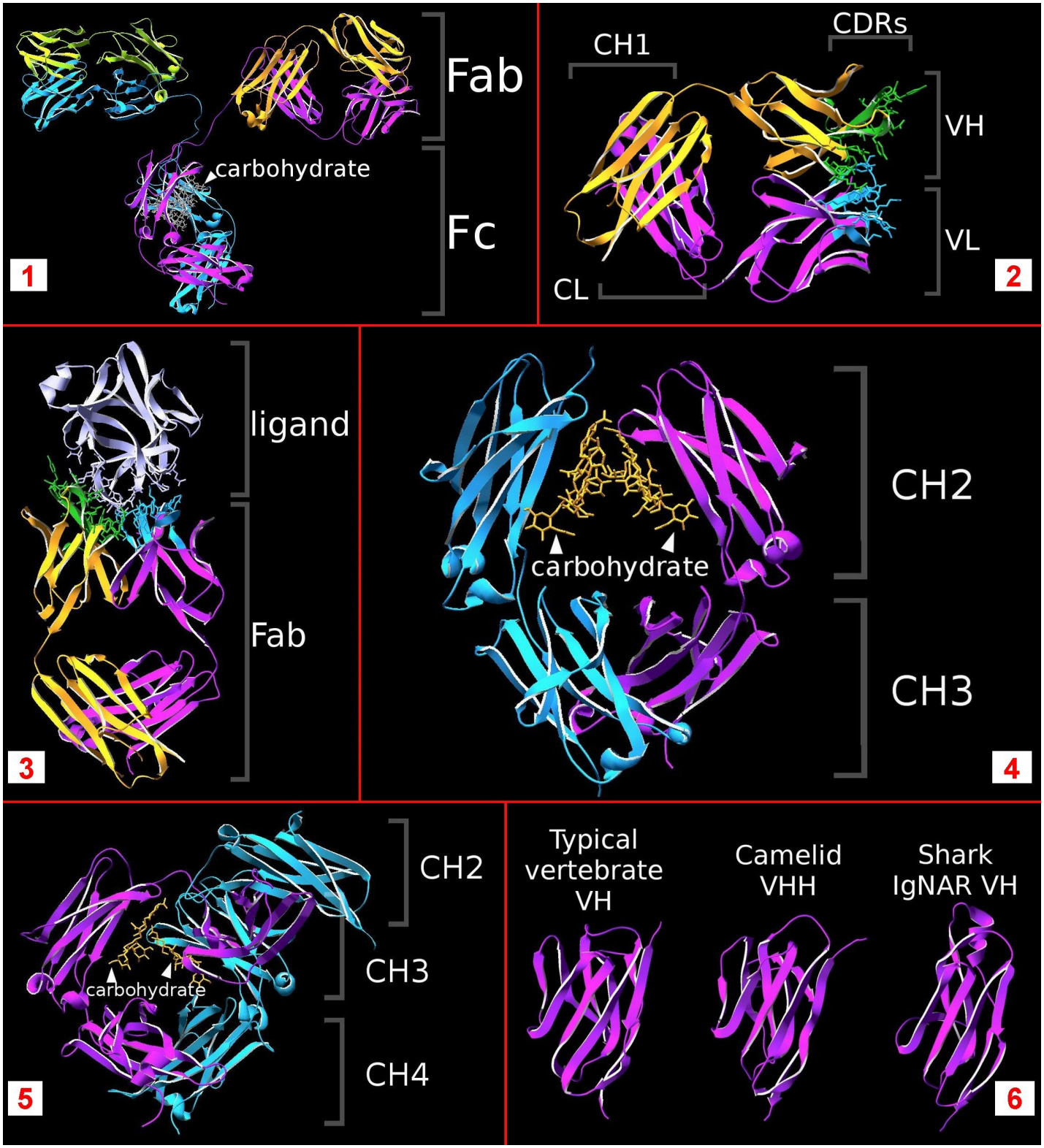
Email Address: gpconcepcion@gmail.com
Submitted: December 14, 2011
Revised: December 30, 2011
Accepted: December 30, 2011
Published: May 29, 2012
Editor-in-charge: Leodevico L. Ilag
Reviewer:
Leodevico L. Ilag

INTRODUCTION

Antibody molecules constitute one of the most important weapons in the arsenal of the immune system and are probably the most extensively studied among the medically significant proteins. Reviews on the structure of antibodies and their interaction with antigen were written soon after the first structures became available (see, for example, Poljak 1973, 1975; Poljak et al. 1976; Davies et al. 1975a,b; Huber 1976; Huber et al. 1976; Capra and Edmundson 1977; Padlan 1977; Amzel and Poljak 1979; Colman 1988; Alzari et al. 1988; Davies et al. 1988, 1990; among others). An extensive analysis of the antibody structure was done by one of us (Padlan 1994). Since then, crystallographic studies have provided higher resolution details of the structure of antibodies and of their interaction with specific ligands. Here, we update that earlier analysis using the structural data now available.

The two most important characteristics of antibodies are exquisite specificity and high binding affinity for the antigens against which they were produced. Those characteristics make antibodies very useful in medicine, research, and industry. Consequently, antibodies have been the subject of extensive engineering. The results of the analysis that we present here will be useful in those engineering endeavors.

Antibodies are produced by all vertebrates and come in a variety of types. In humans, there are five antibody types: IgA (with two subtypes: IgA1 and IgA2), IgD, IgE, IgG (with four subtypes: IgG1, IgG2, IgG3, and IgG4), and IgM, which are often found in specialized locations and all with their specific


Fab
carbohydrate
Fc
1
CH1
CDRs
VH
VL
CL
2
ligand
Fab
3
CH2
carbohydrate
CH3
4
CH2
carbohydrate
CH3
CH4
5
Typical vertebrate VH
Camelid VHH
Shark IgNAR VH
6

functions. For example, IgA is mostly found in the gastrointestinal tract; IgD is found on the surface of the lymphocytes which would eventually produce secreted antibodies; IgE is an important molecule in the fight against parasites and is the antibody type responsible for allergic reactions; IgG is the most prevalent antibody molecule and is primarily responsible for protection against pathogens and their molecules; IgM is the earliest antibody type produced and can also be found on the surface of those lymphocytes which eventually mature to cells that secrete antibodies.

Antibodies come in a variety of sizes. The usual antibody structure is a tetramer of polypeptide chains identical in pairs. One chain is roughly half the size of the other and is called the light (L) chain; the longer chain is called the heavy (H) chain. One light chain is paired with one heavy chain. Both light and heavy chains are built from independently folded structures (domains) of about 110 amino acids. There are two domains in the light chain and four, or five (in the IgE and IgM types), in the heavy chain. The light chain comes in two types: κ (kappa) and λ (lambda). A light chain of either type can associate with the heavy chain of any type.

All antibody domains have a characteristic tertiary structure, that consists mainly of two anti-parallel beta-pleated sheets with loops of varying size and structure connecting the individual strands; this domain structure has been termed the Immunoglobulin Fold (Poljak et al. 1973). In addition to the strong secondary structure that characterizes anti-parallel beta-pleated sheets, the tertiary structure of antibody domains is stabilized by a disulfide bridge connecting the two sheets. Light and heavy chains are usually linked by disulfide bonds and the two heavy chains are linked by one, or more, disulfide bonds.

The N-terminal domains of both light and heavy chains are variable, i.e., they vary from antibody to antibody; those variable domains are referred to as VL and VH, respectively. The other domains of the light and heavy chains are constant, i.e., they are the same for antibodies of the same type. The constant domain of the light chain is called CL and those of the heavy chain are called CH1, CH2, and CH3 (and CH4, in the case of IgE and IgM). There is close association between VL and VH, between CL and CH1, and between the two CH3s in IgA, IgD, and IgG (and between the two CH2s and the two CH4s in IgE and IgM). The structure of a typical IgG molecule is shown in Figure 1.

The close association between domains results in a modular structure for the antibody molecule (Figure 1). The VL:VH module, usually referred to as the Fv (Fragment, variable), is loosely connected to the CL:CH1 module so that there is some degree of freedom in the movement of the Fv relative to CL:CH1. The Fv and the CL:CH1 modules constitute the Fab (Fragment, antigen binding), which has the antigen-binding properties (specificity and affinity) of the antibody. The rest of the constant domains of the two heavy chains (the CH2 and CH3 domains of IgA, IgD, and IgG, and the CH2, CH3, and CH4 domains of IgE and IgM) constitute the Fc (Fragment, crystalline - so named because rabbit Fc, the first one studied, readily formed crystals in distilled water). The "effector functions" of antibodies (for example, recruitment of immune cells, binding to molecules that initiate the destruction of foreign cells, etc) reside in the Fc. The two Fabs and the Fc are connected by a "hinge", that is often unstructured and rather

Figure 1. Ribbon diagram of an intact mouse antibody of the IgG2a type (PDB entry 1IGT) (Harris et al. 1997). Beta strands and helices are shown as wide ribbons, while segments that lack secondary structure are shown as thin strands. The light chains are shown in orange and green, the heavy chains in purple and blue. The 'arms' of the molecule represent the Fabs; the 'stem' represents the Fc. Note that the two Fabs are linked to the Fc by extended strands (the hinge region), the flexibility of which allows essentially unrestricted movement of the Fabs relative to the Fc in this antibody type. The carbohydrates that are normally found between the two chains in the Fc are shown in grey stick representation. This and the other ribbon diagrams (below) were drawn using the modeling software, DeepView v4.0 (Guex and Peitsch 1997), implemented at http://www.expasy.org/spdbv/.

Figure 2. Ribbon diagram of a human Fab (extracted from PDB entry 2VXT). The light chain is shown in purple, the heavy chain in orange. The VL and the VH are at the top of the figure, the CL and CH1 are at the bottom. The CDRs of the light chain are shown in blue; those of the heavy chain are shown in green. The side chains of the CDR residues are shown in stick representation. Note the close association of the VL and VH domains, and of the CL and CH1 domains.

Figure 3. Ribbon diagram of a human Fab complexed to a protein antigen (PDB entry 2VXT). The color scheme for the Fab is the same as in Figure 2; the antigen is shown in lilac. The side chains of the antibody and ligand residues in contact with one another are rendered in stick representation.

Figure 4. Ribbon diagram of a human IgG1 Fc (PDB entry 2DTQ). One chain is shown in blue, the other in purple. The two CH2 domains are at the top; the two CH3 domains are at the bottom. Note the close association of the two CH3 domains. The carbohydrate moieties typically attached to the CH2 domains and located between them in the three-dimensional structure are shown in orange stick representation. The only contact between the two CH2 domains is through these carbohydrates.

Figure 5. Ribbon diagram of a human IgE Fc (PDB entry 2WQR). One chain is shown in blue, the other in purple. The CH2 domains are at the top; the CH3 and CH4 domains are at the bottom. Note the close association of the two CH2 and the two CH4 domains. The carbohydrate moieties attached to the CH3 domains and located between them in the three-dimensional structure are shown in orange stick representation.

Figure 6. Ribbon diagrams of a typical VH (left) [extracted from PDB entry 3D9A], a camelid VHH (middle) [PDB entry 2P49], and a shark IgNAR VH (right) [PDB entry 2I24] shown in approximately the same orientation.

flexible (as in Figure 1), allowing an essentially independent movement of the Fabs relative to the Fc.

The structure of an Fab is shown in Figure 2, where the close association of VL and VH and of CL and CH1 is evident. An isolated Fv has been shown to share similar, but often not identical, antigen-binding properties as the Fab. This is probably because the relative orientation of the VL and the VH in an isolated Fv is not necessarily the same as that found in the Fab - possibly a consequence of the absence in an isolated Fv of the stabilizing effect of the CL:CH1 module.

It was found early on that within both the VL and the VH there are regions that are hypervariable (Wu and Kabat 1970). Those regions have been named "complementarity-determining regions", or CDRs, because they are mainly responsible for the close structural complementarity of the antigen-binding site of the antibody (also called the paratope) and the part of the antigen to which it binds (the epitope). Three CDRs are found in both light and heavy chains, with intervening non-CDR or framework (FR) residues, and they come together at the N-terminal tip of the Fab (see Figure 2). In all the Fab-antigen complexes studied to date, CDR residues predominate in antigen binding, with the occasional involvement of a few neighboring FR residues. The structure of an Fab-antigen complex is shown in Figure 3.

X-ray structures of the Fc from several antibody types have also become available and they have been found to be very similar. The structure of the Fc of a human IgG1 is shown in Figure 4. The structures (not shown) of rabbit Fc and of the two last two domains of avian IgY are very similar to that of human IgG1. The structure of the Fc of a human IgE is shown in Figure 5. Both structures show a close association of the two terminal domains of the heavy chain. The next-to-last domains in the heavy chains (the CH2 domains in IgG and the CH3 domains in IgE) are farther apart, with carbohydrate found in the space between them. The bent structure of an IgE, as predicted from the results of electron-spin resonance studies (Zheng et al. 1991), was confirmed by the crystal structure of the IgE Fc (Figure 5).

Antibody molecules similar in structure to those found in humans are found in other vertebrates. However, unusual antibody types are found in some species. Camelids, in addition to the usual antibody types, have a type that has only heavy chains, i.e., no light chains are found associated with the heavy chains. In that antibody type, the variable domain of the heavy chain (called VHH) is responsible for the entire antigen-binding function of the molecule. Sharks have an unusual antibody type, called IgNAR (Immunoglobulin New Antigen Receptor), which also has no light chains. In shark IgNAR also, the variable domain of the heavy chain constitutes the whole antigen-binding region of the molecule. To compensate for the absence of a VL, the third CDR of both the camelid VHH and the shark IgNAR VH are unusually long and the extra length of this CDR provides enough contact with the antigen to result in significant binding affinity. The VH of a typical antibody, a camelid VHH, and a shark IgNAR VH are compared in Figure 6. Their tertiary structures are seen to be very similar, with the structural differences being mainly in the CDRs.

Antibodies carry out essential functions of the immune system which include: neutralization of toxic antigens, binding to the receptors of pathogens to prevent them from entering cells, immobilization of pathogens to facilitate their ingestion by macrophages and other cells of the immune system, and recruitment of the "complement system" (a cascade system of enzymes and other molecules that is triggered by antibody-antigen complexes and which eventually leads to cell lysis). In view of these functions and, especially, their exquisite specificity, antibodies are widely used in medicine for therapy and diagnostics. Antibodies, again because of their specificity, are also very useful in isolating, tagging with detecting agents for purposes of visualizing, identifying, quantitating, and purifying molecules.

Another unique utilization of antibodies is for the catalysis of chemical reactions for which there are no known natural enzymes. Transition-state analogs can be used to elicit an antibody response, and catalytic antibodies could be obtained and used in place of enzymes (Jencks 1966).

The wide use of antibodies in medicine and in the laboratory is made possible by the ease with which large amounts of antibody molecules of a desired specificity can be generated. One procedure that is in wide use is the hybridoma technology developed by Koehler and Milstein (1973), which involves the immortalization of antibody-secreting cells obtained by usual immunization procedures. Another technology is phage display (Smith 1985), which allows for the generation of many different light and heavy chain combinations (Huse et al. 1989, McCafferty et al. 1990); the random combinations are then screened for molecules that have the desired specificity and affinity.

The specific interaction between antibodies and antigens depends on their respective three-dimensional structures. The elucidation of the three-dimensional structures of antibodies and antigens rapidly progressed in parallel with the development of powerful techniques used in structural biology. Most antibody structures had been generated by x-ray crystallography and some by Nuclear Magnetic Resonance. Several hundred antibody crystal structures have now been elucidated and atomic coordinates for most are available in the Protein Data Bank (Sussman et al. 1998, Berman et al. 2002). The antibodies were from various animal sources (mouse, rat, human, rabbit, chicken, camelids, and shark). Most structures were of natural antibodies, either whole or in fragment form, while others had been modified by protein engineering.

Engineering of the antibody molecule has been encouraged

by the many potential uses of antibodies. For human therapy, antibodies obtained from nonhuman sources (often mice and rats) have to be "humanized", i.e., made less immunogenic to humans. Antibody molecules with greater potency, greater stability, longer half-life, and improved binding properties are desirable for medical purposes, as well as for laboratory and industrial use. Clearly, a detailed knowledge of the structure of antibodies makes it easier to engineer these molecules to achieve the desired characteristics. The aim of this review is to add to that knowledge as more highly refined antibody structures are made available. A preliminary account of this analysis was presented elsewhere (Narciso et al. 2011).

MATERIALS AND METHODS

Structural data

In this review, we analyzed the antibody structures that have been determined by x-ray crystallography and whose coordinates were available in the Protein Data Bank (PDB) as of December 31, 2010 (Table 1). We analyzed representative antibody-antigen interactions and determined the solvent accessibilities of individual residues in the variable domains of the light and heavy chains, and in constant domains. In addition, we analyzed the details of the residue contacts between the variable domain of the L chain (VL) and the variable domain of the H chain (VH), between the constant domain of the L chain (CL) and the first constant domain of the H chain (CH1), and between the constant domains of the heavy chain. We also analyzed the details of the contacts between the residues in the CDRs and the framework regions of the variable domains. We have also analyzed the structures and antigen interactions of the variable domain of the atypical antibodies of camelids (VHH) and of sharks (IgNAR VH).

Although several hundred antibody structures are available in the PDB, we chose to include in our analysis only those which had been determined to high resolution and had been subjected to a high degree of refinement. This is to minimize the uncertainty in the results of our analysis. For purposes of this review, we had arbitrarily designated a structure as "high-resolution/highly refined" if it had been determined at a resolution of 2.00 Angstroms or higher and refined to a crystallographic R-value of 0.200 or better.

We have included in our analysis only naturally occurring molecules and those with native sequences, and have excluded engineered molecules, e.g., those which had been humanized, as well as assemblies of VL and VH domains, whether in the form of paired isolated domains (Fv) or linked together as single chains (scFv). There is no assurance that the antigen-binding site of an Fv will have the same structure and binding properties as the Fab from which the variable domains had been isolated. As mentioned above, a change in the relative orientation of the VL and VH in the Fv versus the Fab was demonstrated very early on (Bhat et al. 1990).

Not infrequently, some segments or atoms are not observed in the electron density maps from the crystallographic studies, even those which had been done at high resolution. The structures with missing parts in the antigen-binding region of the molecule were excluded from our analysis.

Four Fc structures met our criterion for designation as "high-resolution/highly refined" (Table 1). We have included those in our analysis.

Some structures that had been determined at high resolution have shown that water molecules are involved in the interaction between antibody and antigen (e.g., Bhat et al. 1994, Cohen et al. 2005). In the absence of an actual structure determination, the number of water molecules and the nature of their involvement can only be guessed. Obviously, water molecules contribute to the complementarity of the paratope and the epitope. However, no generalizations can be made with the currently available data, so we have decided to forego a discussion of the role of water molecules until more high-resolution structures become available.

Some of the CDRs are mainly loop structures that are exposed to solvent, so that they may be inherently flexible and could be deformed upon binding to ligand - an example of "induced fit" (Koshland et al. 1966). Several examples of "induced fit" in antibody-ligand interactions have been documented (see, for example, Rini et al. 1992, Stanfield et al. 1993). A number of the structures that we have analyzed were unliganded, so that the results for those structures may be different from those which we would have obtained if the structures were liganded.

We wish to remind the reader that our analysis was done on structures that had been determined by x-ray crystallography - a technique that subjects the molecules to non-physiological conditions and damaging radiation. Nonetheless, it is generally accepted that a crystallographically determined structure probably represents one of the more stable conformations of a protein molecule.

Calculations

Surface areas were calculated using the method of Connolly (1983). Fractional solvent accessibilities and interatomic contacts were computed as described earlier (Padlan 1994). Here, two atoms are considered to be in contact if they are within 4.0 Angstroms of each other. A more accurate estimate of interatomic contacts would take into account the error in atomic positions. Error estimates are provided in some, but not all, of the PDB entries and they vary widely. In view of the different resolution and degree of refinement of the structures being analyzed here, only an average value would be appropriate. We chose simply to use a fixed distance of 4.0 Angstroms, which is not unreasonable.

Table 1. High-resolution antibody and antibody-complex structures analyzed in this review (in the PDB as of 12-31-2010)

ANTIBODY	FRAGMENT	PDB CODE	RESOLUTION	R-VALUE	LIGAND	REFERENCE
(Human)						
17B	Fab	2NXY	2.00	0.183	HIV-1 gp120 and CD4	Zhou et al. (2007)
17B	Fab	2NY2	2.00	0.195	HIV-1 gp120 and CD4	Zhou et al. (2007)
7G10	Fab	3D85	1.90	0.172	interleukin-23 subunit	Beyer et al. (2008)
5E1	Fab	3MXW	1.83	0.181	sonic hedgehog protein	Maun et al. (2010)
2F5	Fab	1TJG	2.00	0.198	peptide	Ofek et al. (2004)
3074	Fab(λ)	3MLY	1.70	0.182	peptide	Jiang et al. (2010)
268-D	Fab(λ)	3GO1	1.89	0.192	peptide	Jiang et al. (2010)
ABT-325	Fab	2VXV	1.49	0.155		Argiriadi et al. (2009)
OPG2	Fab	1OPG	2.00	0.160		Kodandapani et al. (1995)
BHA10	Fab	3HC0	1.90	0.182		Jordan et al. (2009)
(Murine)						
125-2H	Fab	2VXT	1.49	0.164	interleukin-18	Argiriadi et al. (2009)
82D6A3	Fab	2ADF	1.90	0.192	von Willebrand factor A3-domain	Staelens et al. (2006)
HyHEL-5	Fab	1YQV	1.70	0.195	lysozyme (hen egg-white)	Cohen et al. (2005)
E8	Fab	1WEJ	1.80	0.200	horse cytochrome c	Mylvaganam et al. (1998)
HyHel-10	Fab	3D9A	1.20	0.191	lysozyme (hen egg-white)	Acchione et al. (2009)
4C3	Fab	3LIZ	1.80	0.178	cockroach Bla g 2 allergen	Li et al. (2011)
MN423	Fab	2V17	1.65	0.160	peptide	Sevcik et al. (2007)
58.2	Fab	1F58	2.00	0.196	peptide	Stanfield et al. (1999)
12A11	Fab	3IFN	1.50	0.188	peptide	Basi et al. (2010)
6A7	Fab	3LEY	1.99	0.180	peptide	Ofek et al. (2010)
101F	Fab	3O41	1.95	0.179	peptide	McLellan et al. (2010)
P20.1	Fab (λ)	2ZPK	1.80	0.173	peptide	Nogi et al. (2008)
F22-4	Fab	3GGW	1.70	0.198	carbohydrate-mimetic peptide	Theillet et al. (2009)
CS-35	Fab	3HNS	2.00	0.172	hexasaccharide	Murase et al. (2009)
F22-4	Fab	3C6S	1.80	0.192	pentasaccharide	Vulliez-Le Normand et al. (2008)
CS-35	Fab	3HNT	1.80	0.199	tetrasaccharide	Murase et al. (2009)
CS-35	Fab	3HNV	2.00	0.177	tetrasaccharide	Murase et al. (2009)
ED10	Fab	2OK0	1.89	0.178	dinucleotide	Sanguineti et al. (2007)
EP2-19G2	Fab	3CFB	1.60	0.187	hapten	Debler et al. (2008)
28B4	Fab	1KEL	1.90	0.199	hapten	Hsieh-Wilson et al. (1996)
M82G2	Fab	1QYG	1.81	0.178	benzoylecgonine	Pozharski et al. (2005)
M82G2	Fab	1RIU	2.00	0.184	norbenzoylecgonine	Pozharski et al. (to be published)
7A1	Fab	2AJV	1.50	0.184	cocaine	Zhu et al. (2006)
4C6	Fab	1NCW	1.30	0.158	benzoic acid	Zhu et al. (2003)
13G5	Fab	1A3L	1.95	0.188	ferrocenyl inhibitor	Heine et al. (1998)
D2.3	Fab	1YEF	2.00	0.199	substrate analog	Gigant et al. (1997)
D2.3	Fab	1YEG	2.00	0.199	reaction product	Gigant et al. (1997)
7A1	Fab	2AJU	1.50	0.184		Zhu et al. (2006)
ACC4	Fab	2W60	1.50	0.171		Uysal et al. (2009)
3A2	Fab	1SBS	2.00	0.180		Fotinou et al. (1998)
17/9	Fab	1HIL	2.00	0.195		Rini et al. (1992)
19D9D6	Fab	1NLB	1.60	0.181		Menez et al. (2003)
F10.6.6	Fab	2Q76	2.00	0.196		Acierno et al. (2007)
J539	Fab	2FBJ	1.95	0.194		Bhat et al. (to be published)
MN423	Fab	3L1O	2.00	0.162		Skrabana et al. (2010)
Fab15	Fab	3NA9	1.70	0.169		Luo et al. (2010)
NC-1	Fab	3OZ9	1.60	0.192		Stanfield et al. (to be published)
(unnamed)	Fab	3I75	1.95	0.181		Riboldi-Tunnicliffe and Isaacs (to be published)
(Camelid)						
HL6	VHH	1OP9	1.86	0.197	lysozyme (human)	Dumoulin et al. (2003)
cAb-Lys3	VHH	1JTP	1.90	0.184	lysozyme (turkey)	Decanniere et al. (2001)
CAB-RN05	VHH	2P49	1.38	0.157	ribonuclease A (cattle)	Koide et al. (2007)
CAB-RN05	VHH	2P4A	1.90	0.186	ribonuclease A (cattle)	Koide et al. (2007)
CABAMD9	VHH	1KXQ	1.60	0.197	pancreatic α-amylase (pig)	Desmyter et al. (2002)
D7	VHH	2XA3	1.50	0.1670		Hinz et al. (2010)
B10	VHH	3LN9	1.80	0.199		Haupt et al. (2011)
(Shark)						
	IgNAR VH	1T6V	1.70	0.195	lysozyme (hen egg-white)	Stanfield et al. (2004)
	IgNAR VH	1SQ2	1.45	0.197	lysozyme (hen egg-white)	Stanfield et al. (2004)
PBLA8	IgNAR VH	2I25	1.80	0.187	lysozyme (hen egg-white)	Stanfield et al. (2007)
PBLA8	IgNAR VH	2I24	1.35	0.174		Stanfield et al. (2007)
(Human)						
	IgG1 Fc	2DTQ	2.00	0.195		Matsumiya et al. (2007)
	IgE Fc	2WQR	1.90	0.1949		Holdom et al. (2011)
(Rabbit)						
	IgG Fc	2VUO	1.95	0.1676		Girardi et al. (2009)
(Avian)						
	IgY C3-C4	2W59	1.75	0.171		Taylor et al. (2009)

Unless otherwise specified, the light chain in the Fab entries is of the kappa (κ) type.

RESULTS

The solvent accessibilities and the identity of the residues involved in the VL:VH contacts in the Fab structures which had been determined at high resolution are presented in Table 2.

The residues involved in the ligand contacts in the antibody-ligand complexes are listed in Table 3.

The contacts between framework and CDR residues in three antibodies are compiled in Tables 4a,b,c. For this analysis, we chose the Fab structures which had been determined at the highest resolution. Details of the contacts between the VL and VH domains in those three Fab structures are presented in Tables 4d,e,f.

The solvent accessibilities and ligand contacts of the residues in a camelid VHH domain are presented in Table 5a. The solvent accessibilities and ligand contacts of the residues in shark IgNAR VH domains are presented in Table 5b. The contacts between the framework and CDR residues in a camelid VHH domain are enumerated in Table 5c. Those contacts between the framework and CDR and hypervariable (H) residues in a shark IgNAR VH domain are listed in Table 5d. Here also, the structures chosen for analysis where those which had been determined at the highest resolution.

The solvent accessibilities of the residues in the CL:CH1 modules of an Fab with a kappa light chain and an Fab with a lambda light chain are compiled in Table 6a; details of the contacts are presented in Tables 6b and 6c, respectively.

The solvent accessibilites of the residues in the Fc of human IgG1, rabbit IgG, and avian IgY, and the residues involved in the contact between the two chains in those molecules are presented in Table 7a. The solvent accessibilities and residue contacts in the human IgE Fc are enumerated separately in Table 7b. Details of the interaction between the two chains in the human IgG1 Fc are presented in Table 7c.

DISCUSSION

The structural information, that we provide here and which we consider to be relevant to the understanding of antibody structure and antigen-binding characteristics, includes the identity of the residues which contact ligand, the exposed and buried residues, the residues involved in the VL:VH interaction, and the framework residues which contact the CDRs. This information is critical in the design of humanization and other engineering protocols while attempting to preserve the antigen-binding properties of the unmodified molecule. We also provide structural information on the Fc part of the molecule since the effector functions, e.g., receptor binding, reside in this part of the molecule and modifications in those functions may be desired. In order to minimize uncertainties, we have limited our analysis to the most accurately determined structures.

The engineering procedure that is the most often done on antibodies is humanization, which aims to reduce the immune response when those molecules are used in medical therapy. There are various techniques used in humanization, including the grafting of the CDRs to a human framework (Jones et al. 1986), grafting only the segments of the CDRs which are involved in the interaction with antigen (abbreviated CDRs), or transfering just the “specificity-determining residues” (SDRs) (Padlan et al. 1995), or by "veneering/resurfacing", i.e., replacing the exposed residues with human counterparts (Padlan 1991, Roguska et al. 1994), or with residues that are expected to be less antigenic based on their physicochemical properties (Padlan 2008, 2010). Humanization by grafting abbreviated CDRs, or by the transfer of only the SDRs, will reduce the likelihood of an anti-idiotypic immune response (i.e., directed at the variable region) against the humanized antibody, compared to that which might be induced by grafting the full CDRs.

A detailed knowledge of the VL:VH contact reveals the identity of the residues which play a major role in the formation of the quaternary structure of the antigen-binding region of an antibody and which should be preserved to maintain its antigen-binding properties. Knowledge of the identity of the framework residues in contact with the CDRs is needed in deciding which framework residues should be preserved when humanizing by CDR-grafting or SDR-transfer (Verhoeyen et al. 1988, Queen et al. 1989, Foote and Winter 1992). Since small differences in binding energy lead to noticeable differences in binding affinity, minor deviations from the original structure of the antigen-binding site could lead to significant changes in the binding properties of an engineered molecule. The changes could be a decrease in binding affinity (see, for example, De Pascalis et al. 2002), or even an increase (see, for example, Brams et al. 2001). Knowing which residues are exposed to solvent is needed when deciding which residues to replace when humanizing by veneering/resurfacing.

We provide similar structural information for the camelid VHH and shark IgNAR VH. It has been shown that those unpaired domains are quite capable of binding to antigen with high affinity and their small size makes them attractive for use in therapy.

To aid in the engineering of the Fc, we also provide the solvent accessibilities of the individual residues of the Fc, as well as details of the interaction between the two chains of the fragment.

In some cases, two or more independently determined sets of high-resolution structural data are available for the same antibody (Table 1). Such duplications provide an opportunity to get a better estimate of the uncertainties in the structure determination and in the structural details. Further, studies of the

Table 2. Residues in contact with the opposite domain and solvent accessibilities in VL and VH domains of high-resolution antibody structures

VL

		CDR-1		CDR-2		CDR-3	
	10 20	27abcdef 30	35 40 49	50 56	60 70 80 88	89 95ab	100 107
(Human)							
		*	* * ** * *	*	*	* * * ***	*
2NXY	DIVMTQSPATLSVSPGERATLSC	RASE------SVSSDLA	WYQQKPGQAPRLLIY	GASTRAT	GVPARFSGSGSGAEFTLTISSLQSEDFAVYYC	QQYNNWPPRYT	FGQGTRLEIK
	ebEBEBeeEebpbee ppBeBeB	*pBpp------eBeepBB*	*BBBBpE ppBpBBBb*	*beebbE*	*BpEbBE e e EeBeBpBEeBpEpBBBeBBB*	*BBBpppEBpBp*	*B e BpBBpp*
		*	* * ** * *	*	*	* * * ***	*
2NY2	DIVMTQSPATLSVSPGERATLSC	RASE------SVSSDLA	WYQQKPGQAPRLLIY	GASTRAT	GVPARFSGSGSGAEFTLTISSLQSEDFAVYYC	QQYNNWPPRYT	FGQGTRLEIK
	ebEBEBEeEebpbee ppBeBeB	*pBpp------eBeepBB*	*BBBBpE ppBpBBBb*	*beebbE*	*BpEbBE E E EpBeBpBeeBpEpBBBeBBB*	*BBBpbpEBpBp*	*B e BpBBpp*
		*	* * ^** * *	*	*	* ** *	*
3D85	DIVMTQSPATLSVTPGDRVSLSC	RASQ------SISDYLH	WYRQKSHESPRLLIK	YASQSIS	GIPSRFSGSGSGSDFTLSINSVEPEDVGVYYC	QNGHSFP--FT	FGSGTKLEIK
	eBEBEBepEepbbee ppBeBeB	*pBpp------eBeppBB*	*BBBBbEEppBpBBBb*	*pbepEbE*	*peEBBE e e eeBeBeBEeBpeeBp eBBB*	*BB pepp--Bp*	*B E BeBpBp*
		* *	* * ** * *	* *	*	* * ** *	*^^
1TJG	ALQLTQSPSSLSASVGDRITITC	RASQ------GVTSALA	WYRQKPGSPPQLLIY	DASSLES	GVPSRFSGSGSGTEFTLTISTLRPEDFATYYC	QQLHFYP--HT	FGGGTRVDVR
	EBeBEBeeEebpBee ppBeBpB	*pBpp------ BeepBB*	*BBBBeE peBeBBBb*	*bbeppbE*	*bEEbBE E e EpBeBbBeeBbEpBbBeBBB*	*BBBpppp--Bb*	*B BpBbep*
			* * ^** * *	^*	*	* * ** *	*
2VXV	EIVMTQSPATLSVSPGERATLSC	RASE------SISSNLA	WYQQKPGQAPRLFIY	TASTRAT	DIPARFSGSGSGTEFTLTISSLQSEDFAVYYC	QQYNNWPS-IT	FGQGTRLEIK
	ebEBeBEeEebpbeE ppBeBeB	*pBep------eBeepBB*	*BBBbpE ppBpBBBb*	*ebeepeE*	*eBeEbBE e E eeBeBbBeeBpEpBBBeBBB*	*BBBpEbpE-Bb*	*B e BpBBpp*
		* *	* * ** *		*	* *^ ** *	*
1OPG	DELLTQSPATLSVTPGDSVSLSC	RASQ------SISNNLH	WYQQKSHESPRLLIK	YASQSIS	GIPSRFSGSGSGTDFTLSINSVETEDFGMYFC	QQSNSWP--LT	FGGGSKLEIK
	eBEBEBEeEEppbEe ppBeBEB	*pBpp------EBEeBBB*	*BBbBbEEppBpBBBp*	*bBepepe*	*beEbBe e e eeBeBeBeEBpEeBb pBBB*	*BBBpepE--Bb*	*B BeBbBp*
		* *	* * ^** * *	* *		* * ** *	*
3HC0	DIQMTQSPSSLSASVGDRVTITC	KASQ------NVGINVA	WYQQKPGKAPKSLIS	SASYRYS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC	QQYDTYP--FT	FGQGTKVEIK
	eBeBEBEeeEbpbee ppBeBEB	*pBpp------eB epBB*	*BBBBpE eeBpbBBb*	*pBeepbE*	*bEEbBE E e eEBpBpBeeBpEpBbBeBBB*	*BBBbepp--bb*	*B p BpBbpp*
		*	* * ^** * *	* *	*	* * * *	*
3MXW	DIVMTQTPKFLLVSAGDKVTITC	KASQ------SVSNDLT	WYQQKPGQSPKLLIY	YASNRYT	GVPDRFTGSGYGTDFTFTISTVQAEDLAVYFC	QQDYGSP--PT	FGGGTKVEIK
	eBEBEBEeeeppbeE ppBeBeB	*eBep------eBeepBB*	*BBBBpe ppBeBBBb*	*pbepbbE*	*BeebBE E e eeBeBpBeEBpEpBbBeBBB*	*BBBp Ep--bb*	*B BeBbBp*
		*	* * ^** * *	*	*	* * ^^* *	*
3G01	SYVLTQPPS-VSVSPGQTARITC	SAEA------LSNQYAY	WYRQRPGQAPLLIIY	KDTKRPS	GIPERFSGSTSGTTVTLTISGVQAEDEADYYC	QSADSSGD-YV	FGGGTKVTVLG
	EBEBEBeEe-ppbeE ppBpBeB	*ebeb------BeeBeBB*	*BBbbpE pEBEBBBb*	*pbepbbE*	*beebBe EeE eeBpBpBe BpEebBBpBBB*	*BBbbEe b-Bp*	*B BeBeBe*
		** *	* * ^** * *		*	^** ^^* * *	*
3MLY	QSVLTQPPS-VSAAPGQKVTISC	SGSSSN----IGNNMVS	WYQQHPGTAPKLLIY	ENSKRPS	GIPDRFSGSRSGTSATLGIIGLQTGDEAEYYC	ATWDGSLR-TV	FGGGTKLTVLS
	eepBEBeEp-pBpEE peBEBEB	*E eEeB----B ebpBB*	*BBBBeE EpBebBBb*	*bbepbpE*	*BeebBE EbE eEBeB BE BpE BpBbBBB*	*BBBB bEb-Bb*	*B BpBBBbE*
(Murine)							
			* * ** * *	* *	*	* * * *	*
1YQV	DIVLTQSPAIMSASPGEKVTMTC	SASS-------SVNYMY	WYQQKSGTSPKRWIY	DTSKLAS	GVPVRFSGSGSGTSYSLTISSMETEDAATYYC	QQWGRN---PT	FGGGTKLEIK
	epEBEBeeEEbpbpE peBeBeB	*EBpe-------eBEpBB*	*BBBbpE epBpBbBb*	*bbeppBE*	*BeEbBE e E eEBeBpBEeBpEpBEBeBBB*	*BBp ep---Bb*	*B BeBbbp*
		*	* * ^** * *	*	*	* * * *	*
3OZ9	QIVLTQSPVIMSASLGEEITLTC	SASS-------SVSYMH	WYQQKSGTSPKLLIY	STSNLAS	GVPSRFSGSGSGTFYSLTISSVEAEDAADYYC	HQWSGF---YT	FGGGTKLEIQ
	ebEBEBEeEebeBee ppbeBeB	*eBee-------eBeeBb*	*BBbBpe ebBebBBb*	*pbeeppE*	*BEEbBE E E eeBpBpBeeBpepBpBpBBB*	*BBpe p---Bb*	*B BeBbBp*
		* *	* * ** * *	* *	*	* * ** * *	*
2FBJ	EIVLTQSPAITAASLGQKVTITC	SASS-------SVSSLH	WYQQKSGTSPKPWIY	EISKLAS	GVPARFSGSGSGTSYSLTINTMEAEDAAIYYC	QQWTYPL--IT	FGAGTKLELK
	eBEBEBeeEEbEBee ppbeBeB	*eBee-------EBpBBB*	*BBBBpE EbBepBBb*	*bpepppE*	*BeEbBE E E eEBeBpBeEBpEeBBBpBBB*	*BBbppEp--Bp*	*B E BpBbBp*
		*	* * ^** * *	* *	*	* * ** *	*^^
2Q76	DIELTQSPATLSVTPGDSVSLSC	RASQ------SISNNLH	WYQQKSHESPRLLIK	YTSQSMS	GIPSRFSGSGSGTDFTLSINSVETEDFGVYFC	QQSGSWP--RT	FGGGTKLDIK
	eBpBEBeeEEbpbee eEBeBeB	*pBpp------eBeebBB*	*BBBBpeEppBpBBBb*	*pbeeEbE*	*beEbBE E E eeBeBpBEeBpepBb pBBB*	*BBB ppB--Bb*	*B BeBbBp*
		*	* * ** * *	* * *	*	* ** * *	*
3LIZ	QIVLTQSPSSMYASLGERVTITC	KASQ------DINNYLS	WFQQKPGKSPKTLIY	RADRLVD	GVPSRVSGSGSGQDYSLTISSLEYEDLGIYYC	LQYDELP--YT	FGGGTKLEIK
	eBEBEBEeEEbpBep ppBeBEB	*eBpp------eBepbBB*	*BBbbpE epBeBbBB*	*bBeppbE*	*bEEbbE e E peBpBpBeEBpbpBp pBBB*	*BBBbppp--bb*	*B BeBbBp*
		*	* * ^ * *		*	* * *	* ^
2VXT	DIQMTQSPSSLSASLGERVSLTC	RASQ------DIGSKLY	WLQQEPDGTFKRLIY	ATSSLDS	GVPKRFSGSRSGSDYSLTISSLESEDFVDYYC	LQYASSP--YT	FGGGTKLAIK
	eBeBEBEeEebpBeE ppBeBeB	*Ebpp------eB EpBb*	*BBBBpEp EBpBBBp*	*EbepbbE*	*BeebBE ebE epBpBbBeEBpEpBBbpBBB*	*BBbeeEp--pp*	*B BpBbEe*
		* *	* * ^^* * *	*	*	* * * *	*
2ADF	DIQMTQSPSSLSASLGGKVTITC	KASQ------DINKYIA	WYQHKPGKGPRLLIH	YTSTLQP	GIPSRFSGSGSGRDYSFSISNLEPEDIATYYC	LQYDNL---RT	FGGGTKLEIK
	eBeBEBeeEEbpeEe eBeBEB	*pBpp------pBeppBB*	*BBBbpE e BpBBBB*	*bBeepbE*	*BpEbBE E e peBeBeBeeBbepBbBeBBB*	*BBBbee---Be*	*B BeBbBp*
	*	*	* * * ^** * *	*	*	* * ** *	*
1WEJ	DIQMTQSPASLSASVGETVTITC	RASG------NIHNYLA	WYQQKQGKSPQLLVY	NAKTLAD	GVPSRFSGSGSGTQYSLKINSLQPEDFGSYYC	QHFWSTP--WT	FGGGTKLEIK
	pbpBEBeeEebpBee pebeBeB	*pBp ------pBpepBB*	*BBBBpp ppBpBBBb*	*pbeebbe*	*BeEbBE E E epBeBpBeeBpEpBB pBBB*	*BBBpeeB--Bb*	*B BeBBep*
			* * ** *		* *	* ** *	* ^
3D9A	DIVLTQSPATLSVTPGNSVSLSC	RASQ------SIGNNLH	WYQQKSHESPRLLIK	YASQSIS	GIPSRFSGSGSGTDFTLSINSVETEDFGMYFC	QQSNSWP--YT	FGGGTKLEIK
	eBEBEBEeEebpBEe eebEBpB	*pBpp------eB ebBb*	*BBBBpeeppBpBBBp*	*pbepEbE*	*beEbBE E E eeBeBeBeeBpEpBp bBBB*	*BBppebb--bp*	*B BpBbbp*
			* * ** * *	*	*	* * ** *	*
2V17	DVQITQSPSYLAASPGETITINC	RASK------SIRKFLA	WYREKPGKTNKLLIY	SGSTLQS	GTPSRFSGSGSGTDFTLTISRLEPEDFAMYYC	QQHNDYP--LT	FGAGTKLELK
	eBeBEBeeEpbbbpe peBeBpB	*pBpp------eBppbBB*	*BBBbbE peBpBBBb*	*b EepbE*	*BeEbBE E E eeBeBpBepBpEeBBBpBBB*	*BBBbbpe--Bp*	*B E BpBBpp*
	*	*	* * ^ * * *	*	*	* * ** *	*
3HNS	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
	ppeBEBEeEEppeee ppBeB B	*pBep------pB epBB*	*BBBbpEe EBpBBBb*	*pbeppbE*	*BpEbBE e p eeBpBpBeeBpbpBp pBBB*	*BBbbEE--Bp*	*B E BEBbBp*
	*	*	* * ^ * * *	*	*	* * ** *	* ^
3HNT	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
	ppeBEBEeEEppeee ppBeB B	*pBpp------pB epBB*	*BBBbpEe EBpBBBb*	*pbeppbE*	*BpEbBE e p eeBpBpBeeBpbpBp pBBB*	*BBbbEe--Bp*	*B E BEBbBp*
	*	*	* * ^ * * *	*	*	* * ** *	* ^
3HNV	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
	ppeBEBEeEEppeee ppBeB B	*pBpp------pB epBB*	*BBBbpEe EBpBBBb*	*pbeppbE*	*BpEbBE e p eeBpBpBeeBpbpBp pBBB*	*BBbbEe--Bp*	*B E BEBbBp*
		*	* * ^ * *	*	*	* ** *	*
3I75	DIQMTQSPSSLSASLGGKVTITC	QSSQ------DINKYIG	WYQHKPGKGPRLLIH	YTSILRP	DIPSRFSGSGSGRDYSFSISNLEPEDTATYYC	LQYDDLL--LT	FGAGTKLELK
	eBeBEBeeEEbeeEE eBeBEB	*eBpe------pBepbB*	*BBBbbE p BpBBBb*	*bBeepbE*	*pbeEbBE E e peBpBpBEEBpppBbBpBBB*	*BBBbbEp--Bp*	*B E BeBbBp*
		*	* * ^^** * *	*	*	* * ** *	*^
3L10	DVQITQSPSYLAASPGETITINC	RASK------SIRKFLA	WYREKPGKTNKLLIY	SGSTLQS	GTPSRFSGSGSGTDFTLTISRLEPEDFAMYYC	QQHNDYP--LT	FGAGTKLELK
	eppBEBepEepeBee peBeBeB	*eBpe------eBeppBB*	*BBbbpE peBpBBBb*	*p eepbE*	*BeEbBE E E eeBeBeBeeBpEpBBBbBBB*	*B8Bpebe--Bp*	*B E BeBBpp*
		^	* * ^** * *	* *	*	* ^ * * *	*
3NA9	DIQMTQSPSSLSASVGDRVTITC	RASQ------SIGLYLA	WYQQKPGKAPKLLIY	AASSLQS	GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC	QQGNTLS--YT	FGQGTKVEIK
	eBpBEBeeEebpbee ppBeBeB	*pBpp------eB epBB*	*BBbBpE eeBpBBBb*	*ebeepbE*	*BpEbBE E e eeBeBeBeEbEpBbBpBBB*	*BB peeE--Bp*	*B p BpBbep*
		^ ^*^	* * ^** * *	* * *	*	** * *	*
1F58	DIVLTQSPASLAVSLGQRATISC	KASQGVDF--DGASFMN	WYQQKPGQPPKLLIF	AASTLES	GIPARFSGRGSGTDFTLNIHPVEEEDAATYYC	QQSHEDP--LT	FGAGTKLELK
	eBEBEBEeEpbeBee ppBpBeB	*eBpp Bpe--e pbpBB*	*BBbBpE ppBeBBBB*	*BbeepbE*	*beEbBe p e eeBpBbBeeBpepBeBpBBB*	*BBbbpbb--bp*	*B E BpBBpp*
		* *	* * ^** * *	* *	*	* * ** * *	*
3O41	DIVLTQSPASLAVSLGQRATIFC	RASQSVDY--NGISYMH	WFQQKPGQPPKLLIY	AASNPES	GIPARFTGSGSGTDFTLNIHPVEEEDAATYYC	QQIIEDP--WT	FGGGTKLEIK
	eBEBEBeeEEpEbep ppBpBeB	*eBpppBbe--e pbbBB*	*BBBbpE ppBeBBBb*	*bbepEbE*	*BeEBBE e e eEBpBpBpEBpppBbBeBBB*	*B8Bppp8--Bp*	*B BEBpBp*
			* * ^** *		*	^* ** * *	*
3GGW	DIVMTQAAFSNPVTLGTSASISC	RSSKSLLHS-DGITYLY	WYLQKPGQSPHLLIY	HLSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGIYYC	AHNVELP--RT	FGGGTKLEIK
	eBEBeBEEebEebEe epbeBeB	*pBeppBeee-b pbbBB*	*BBBBbe pbBebBBp*	*pbeppeE*	*beebBEBE p EeBeBpBeeBpeeBB pBBB*	*BBppbpe--Bp*	*B beBbBp*

Table 2. continued.

	CDR-1			CDR-2			CDR-3		
	10 20	27abcdef 30	35 40 49	50 56	60 70 80 88	89 95ab	100 107		
		* * ** *			*	^* ** *	*		
3C6S	DIVMTQAAFSNPVTLGTSASISC	RSSKSLLHS-DGITYLY	WYLQKPGQSPHLLIY	HLSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGIYYC	AHNVELP--RT	FGGGTKLEIK		
	pBEBeBEEebEebee epbeBEB	*pBeppBepE-b pbbBB*	*BBBBpe pbBpbBBp*	*bbeppeE*	*beeBBEBE p eeBeBpBbpBpEpBB pBBB*	*BBppbpe--Bp*	*B bpBpBp*		
		*	* * ** * *	*	*	* ** *	*		
2OK0	DILMTQTPLSLPVSLGDQASISC	RSSQSIVHS-NGNTYLE	WYLQKPGQSPTLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC	FQGSHIP--LT	FGAGTKLEVK		
	eBEBEBepEpbeBee peBpBpB	*pBpppBepE-p pbpBB*	*BBbBpE ppBEBBBp*	*pbepbpE*	*beeBBE e E eeBpBpBepBpEpBp eBBB*	*BB Bppp--bb*	*B E BpBBee*		
		*	* * ** * *	* *		* ** *	*^^		
3CFB	DIVMTQAAFSNPVTLGTSASISC	RSTKSLLHS-NGITYLY	WYLQKPGQSPQLLIY	QMSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGVYYC	AQNLELP--PT	FGGGTKLEIK		
	pbEBeBEEpBEebEe epBpBEB	*pBEpeBepe-p pbbBB*	*BBbbpE epbeBBBb*	*bbeppBe*	*BeebBEBE E eeBeBpBepBpepBb pBbB*	*BBbbpee--pb*	*b beBbBe*		
		* *	* * ^** * *	*	*	* ^ *	*		
1KEL	DVLMTQTPLSLPVSLGDQASISC	RFSQSIVHS-NGNTYLE	WYLQKSGQSPKLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC	FQGSHVP--RT	FGGGTKLEIK		
	pbEBEBEeEebeBpe ppBeBpB	*pBppeBepe-p pbbBB*	*BBBBbE ppBeBBBb*	*ppepbbE*	*BEebBE e E epBpBpBepBpEeBp pBBB*	*BB bbEE--be*	*B BpBBpp*		
		* *	* * ** * *	*	*	* * ** *	*		
1QYG	DIVMTQAAPSVPVTPGESVSISC	RSSKSLLHS-NGYTYLH	WFLQRPGQSPQLLIY	RVSNLAS	GVPDRFSGSGSGTAFTLRFSRVEAEDVGVYYC	MQHLEYP--FT	FGSGTKLEIK		
	pBEBEBEEEpbEBee pEbeBEB	*eBppeBepE-b pbbBB*	*BBBBbe pbBpBBBb*	*bbeppeE*	*beEbBE e E EEBpBpBEeBpEpBB eBBB*	*BBbbpbe--Bb*	*B E BeBbBp*		
		*	* * ** * *	*	*	* * ** *	*		
1RIU	DIVMTQAAPSVPVTPGESVSISC	RSSKSLLHS-NGYTYLH	WFLQRPGQSPQLLIY	RVSNLAS	GVPDRFSGSGSGTAFTLRFSRVEAEDVGVYYC	MQHLEYP--FT	FGSGTKLEIK		
	eBEBEBEEEepEBee pebeBEB	*eBepeBepE-b pbbBB*	*BBBbbE pbBpBBBb*	*bbeppEE*	*BeebBE e E eEBpBpBEeBpEpBB eBBB*	*BBbbpbe--Bp*	*B E BeBbBp*		
		* * *	* * ^** * *	*	*	* * ** *	* *		
2AJV	DIVITQDELSNPVTSGESVSISC	RSSRSLLYK-DGRTYLN	WFLQRPGQSPQLLIY	LMSTRAS	GVSDRFSGSGSGTDFTLEISRVKAEDVGVYYC	QQFVEYP--FT	FGSGTKLEIK		
	BBEBEBpebEEpbeE ppbeBeB	*pBppeBepe-b pbBBB*	*BBBbpp pbBpBBBb*	*pbeebbE*	*BpeBBE E e eeBeBpBepBeeeBB eBBB*	*BBBbbbb--bp*	*B E BeBbBp*		
		* * *	* * ^** * *	* **	*	* ** *	* ^		
1NCW	DVVMTQSPKTISVTIGQPASISC	KSSQRLLNS-NGKTFLN	WLLQRPGQSPKRLIY	LGTKLDS	GVPDRFTGSGSGTDFTLKISRVEAEDLGVYYC	WQGTHFP--YT	FGGGTKLEIK		
	eBEBEBEeeebebee pEBeBpB	*pBppeBebE-E bpbBB*	*BbbBbe bbBebBBB*	*p EppBe*	*BeeBBE e E eeBpBpBepBpepBB bBBB*	*bB BppB--bb*	*B BpBbBp*		
		* * * *	* * ** * *	*	*	* ** *	*		
1A3L	DIVLTQAAFSNPVTLGASASISC	RSSKSLLNS-NGIIHMY	WYLQKPGQSPQLLIY	QMSKLAS	GAPDRFSGSGSGTDFTLRISRVEAEDVGVYYC	AQNLELP--YT	FGGGTKLEIK		
	pBEBEBEEEpEebEe EEpeBEB	*pBepeBebe-p bbBBB*	*BBBbbE ebBpbBBb*	*bbepppE*	*BeebBE E e epBpBpBpEpBpppBb eBBB*	*pBbbbee--pb*	*B beBpBe*		
		* * *	* * **** *	*	* *	** *	* ^^		
1YEF	DIVMTQSPLTLSVTIGQPASISC	KSSQSLLYS-NGKTYLN	WLLQRPGQSPKRLIH	LVSKLDS	GVPDRITGSGSGTDFTLKISRVEAADLGVYYC	VQGTHFP—YT	FGGGTKLEIL		
	ebEBEBEeEEbpbEe pEBpBEB	*pBppeBepE-b pbbBB*	*BBBbpe epBpBbBB*	*bbepbBE*	*bEebBE e E peBpBeBEpBpEEBB pBBB*	*bB Bppb--bb*	*B BbBBBe*		
		* * * *	* * **** *	*	* *	** *	* ^^		
1YEG	DIVMTQSPLTLSVTIGQPASISC	KSSQSLLYS-NGKTYLN	WLLQRPGQSPKRLIH	LVSKLDS	GVPDRITGSGSGTDFTLKISRVEAADLGVYYC	VQGTHFP--YT	FGGGTKLEIL		
	ebEBEBEeEEbpbee pEBeBEB	*pBppeBepE-b pbbBB*	*BBBbpE ppBpBbBB*	*bbeppBE*	*bEebBE e E peBeBeBEpBpEEBB pBBB*	*bB Bppb--bb*	*B BbBbBe*		
		* *	* * ** * *	*	*	* * ** *	*		
2AJU	DIVITQDELSNPVTSGESVSISC	RSSRSLLYK-DGRTYLN	WFLQRPGQSPQLLIY	LMSTRAS	GVSDRFSGSGSGTDFTLEISRVKAEDVGVYYC	QQFVEYP--FT	FGSGTKLEIK		
	eBEBEBpebEEpbeE ppbeBeB	*pBppeBepe-b pbBBB*	*BBBbpp pbBpBBBb*	*BbeebbE*	*BpeBBE E e eeBeBpBepBeeeBB eBBB*	*BBbbppb--bp*	*B E BeBbBp*		
		* *	* * ** * *	*	^	* ^ ** *	*		
2W60	DVVMTQTPLTLSVTIGQPASISC	KSSQSLLDS-DGKTYLN	WLLQRPGQSPKRLIY	LVSKLDS	GVPDRFTGSGSGTDFTLKISRVEAEDLGVYYC	WQGTHFP--LT	FGAGTKLELK		
	pBEBEBEeEpbpBEe pEBeBpB	*pBppeBebE-b pbbBB*	*BBBbpE epBeBBBb*	*bbepppE*	*peebBE e E eeBeBpBEpBpEpBp eBBB*	*BB Bpep--Bb*	*B E BpBBpe*		
		* *	* * ** * *	*	*	* * ** *	*		
3IFN	DVLMTQTPLSLPVSLGDQASISC	RSSQSIVHS-NGNTYLE	WYLQKPGQSPKLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGIYYC	FQSSHVP--LT	FGAGTKLELK		
	eBEBEBEpeebebee ppBeBeB	*pBppeBEpE-p ppbBB*	*BBBbbE epBpBBBb*	*ppepbpE*	*beebBE E E eeBeBpBeeBpEpBp pBBB*	*BBBBpeB--bb*	*B E BpBBpp*		
		*	* * ** *		* *	* ^ ** *	*		
3LEY	DVVMTQTPLSLSVTLGQPASISC	KSSQSLLDS-DGKTYLN	WLLQRPGQSPKRLIY	LVSKLAS	GVPDRFTGSGSGTDFTLKINRVEAEDLGIYYC	WQGTHFP--WT	FGGGTKLEIK		
	ebEBEBEeEebpbEe eEBEBeB	*eBppeBebE-p pbbBb*	*BBBbpe epBEbbBp*	*pbepppE*	*beebBE e E eeBeBpBppBpeeBb pBBB*	*BB Bbpb--Bp*	*B BpBbBp*		
		*	* * ** * *	* *	*	* * ** *	*		
1SBS	DIVMSQSPSSLAVSVGEKVTMTC	KSSQSLLYSSNQMNYLA	WYQQKPGQSPKLLIY	WASTRES	GVPDRFTGSGSGTDFTLTISSVEAEDLAVYYC	QQYHSYP--FT	FGSGTKLEIK		
	eBEBEBEeEepEbeE peBeBeB	*eBppeBpbEebepBBBB*	*BBbbpp ebBeBBBb*	*bbeebpE*	*BeebBE e e eeBpBeBeEBpEpBbBeBBB*	*BBBbpbe--Bp*	*B E BeBpBp*		
		* * *	* * ** * *	* *	*	* * ** *	*		
1HIL	DIVMTQSPSSLTVTAGEKVTMSC	TSSQSLFNSGKQKNYLT	WYQQKPGQPPKVLIY	WASTRES	GVPDRFTGSGSGTDFTLTISSVQAEDLAVYYC	QNDYSNP—LT	FGGGTKLELK		
	ebEBEBEeEepeBEE ppBeBeB	*eBppeBppE pebbBBB*	*BBbBpE ppBpBBBB*	*bbeebbE*	*beebBE e e eeBeBpBEeBpEpBpBpBBB*	*BBBpepe--bp*	*B BpBbbp*		
		*	* * ** * *	* *	*	* * * *	*		
1NLB	DIVMSQSPSSLAVSAGEKVTMSC	KSSQSLLNSRTRKNYLA	WYQQKPGQSPKVLIY	WASTRES	GVPDRFTGRGSGTDFTLTISSVQAEDQAVYYC	KQAYIPP--LT	FGAGTKLELK		
	eBEBEBEeEebEbeE ppBeBeB	*pBepeBpbEepepBbBB*	*BBbBpE pbBeBBBB*	*bbepbbE*	*bEebBp p e eeBeBpBeeBppeBBBeBBB*	*BBBbeEp--bp*	*B E BpBpBp*		
		* *	* * * ^ *	*	*	* * *	*^^		
2ZPK	ZTVVTQESA-LTTSPGETVTLTC	RSSTGAV---TTSNYAN	WVQEKPDHLFTGLIV	GTNNRVP	GVPPRFSGSLIGDKAALTITGAQTEDEAIYFC	ALWYSNH--WV	FGGGTKLTVLG		
	ppeBeBbEE-ppbee pebeBpB	*pBpe EB---EEEbpBB*	*BBBbpEepeBe BBb*	*BeebbE*	*BeEbBe ebe ebBbBpBE BpEpBbBeBBB*	*BBbbEep--Bp*	*B BbBBBb*		

Table 2. continued.

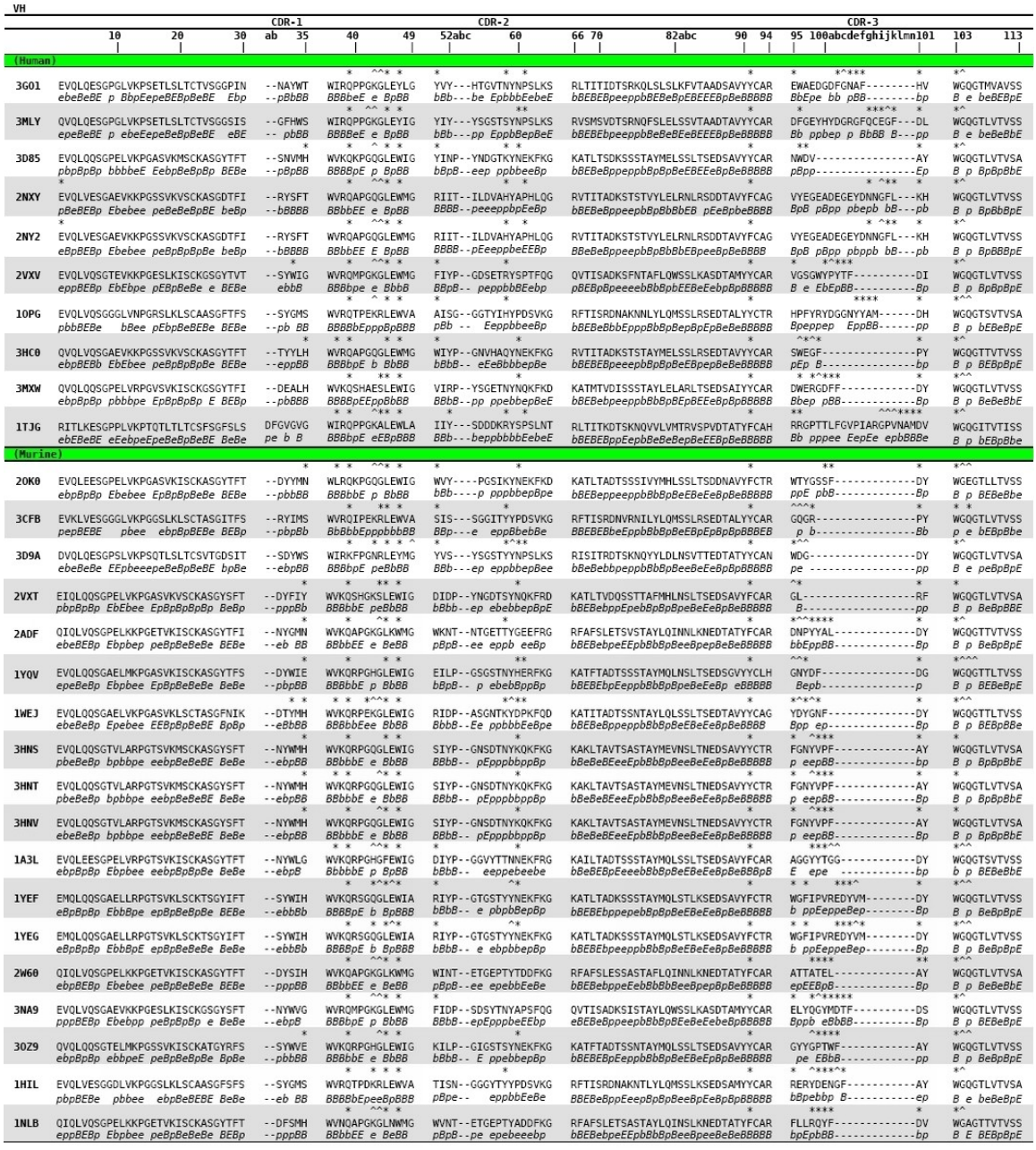
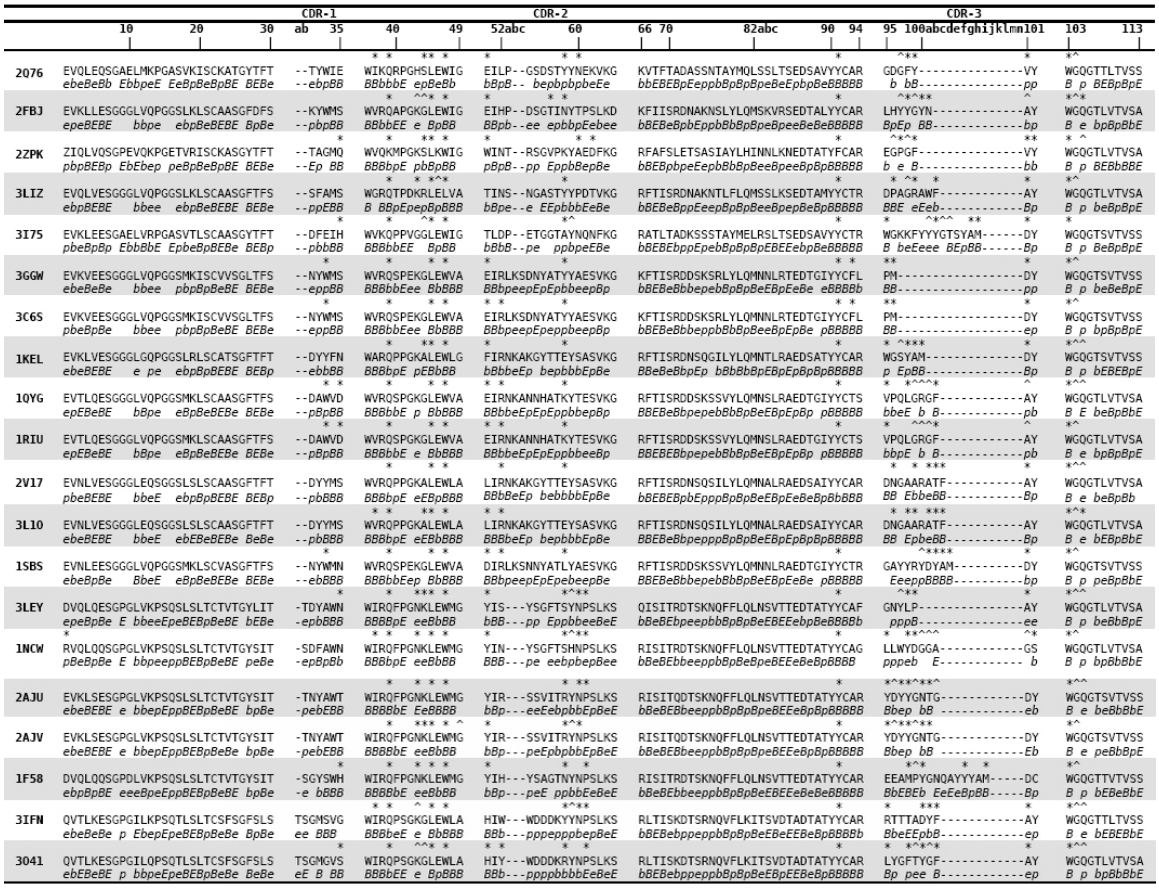


Table 2. continued.



The solvent accessibilities of the side chains are placed under the sequences and are italicized. A residue is designated "*E*" (completely exposed) if the fractional accessibility of its side chain is at least 0.80; "*e*" (mostly exposed) if the accessibility is between 0.60 and 0.80; "*p*" (partly exposed, partly buried) if the accessibility is between 0.40 and 0.60; "*b*" (mostly buried) if the accessibility is between 0.20 and 0.40; and "*B*" (completely buried) if the accessibility is between 0.00 and 0.20 (definitions are from Padlan (1990)). The accessibility of glycine is simply designated as blank, " ". A residue that is in contact by its side chain is indicated by an asterisk, "*", above it; one that is in contact by only its main chain is indicated by a caret, "^". The numbering scheme follows that of Kabat et al. (1991), except in CDR-1 of the heavy chain, where the insertions are placed after residue 30 instead of residue 35 - a structurally more logical placement (Padlan et al. 1995). The molecules are simply identified by their PDB entry codes. The single-letter code is used for the amino acids. The light and heavy chains are listed by subgroup (Kabat et al. 1991). The N-terminal, "Z", in the PDB entry 2ZPK represents polyglutaric acid.

Table 3. Contacts between antibody and ligand in high-resolution antibody-ligand complex structures

VL

		CDR-1		CDR-2		CDR-3	
	10 20	27abcdef 30	35 40 49	50 56	60 70 80 88	89 95ab	100 107
(Human)							
			WYQQK			**	
2NXY	DIVMTQSPATLSVSPGERATLSC	RASE------SVSSDLA	PGQAPRLLIY	GASTRAT	GVPARFSGSGSGAEFTLTISSLQSEDFAVYYC	QQYNNWPPRYT	FGQGTRLEIK
						**	
2NY2	DIVMTQSPATLSVSPGERATLSC	RASE------SVSSDLA	WYQQKPGQAPRLLIY	GASTRAT	GVPARFSGSGSGAEFTLTISSLQSEDFAVYYC	QQYNNWPPRYT	FGQGTRLEIK
		*		*		^*** *	
3D85	DIVMTQSPATLSVTPGDRVSLSC	RASQ------SISDYLH	WYRQKSHESPRLLIK	YASQSIS	GIPSRFSGSGSGSDFTLSINSVEPEDVGVYYC	QNGHSFP--FT	FGSGTKLEIK
		***		* ** *		^*	
3MXW	DIVMTQTPKFLLVSAGDKVTITC	KASQ------SVSNDLT	WYQQKPGQSPKLLIY	YASNRYT	GVPDRFTGSGYGTDFTFTISTVQAEDLAVYFC	QQDYGSP--PT	FGGGTKVEIK
						^*** *	
1TJG	ALQLTQSPSSLSASVGDRITITC	RASQ------GVTSALA	WYRQKPGSPPQLLIY	DASSLES	GVPSRFSGSGSGTEFTLTISTLRPEDFATYYC	QQLHFYP--HT	FGGGTRVDVR
		^^*^*		*	*	^*^^ *	
3G01	SYVLTQPPS-VSVSPGQTARITC	SAEA------LSNQYAY	WYRQRPGQAPLLIIY	KDTKRPS	GIPERFSGSTSGTTVTLTISGVQAEDEADYYC	QSADSSGD-YV	FGGGTKVTVLG
				*	*		
3MLY	QSVLTQPPS-VSAAPGQKVTISC	SGSSSN----IGNNMVS	WYQQHPGTAPKLLIY	ENSKRPS	GIPDRFSGSRSGTSATLGIIGLQTGDEAEYYC	ATWDGSLR-TV	FGGGTKLTVLS
(Murine)							
		* ** *	* *	*		*^ * *	
2VXT	DIQMTQSPSSLSASLGERVSLTC	RASQ------DIGSKLY	WLQQEPDGTFKRLIY	ATSSLDS	GVPKRFSGSRSGSDYSLTISSLESEDFVDYYC	LQYASSP--YT	FGGGTKLAIK
		*	*	* *		* *	
2ADF	DIQMTQSPSSLSASLGGKVTITC	KASQ------DINKYIA	WYQHKPGKGPRLLIH	YTSTLQP	GIPSRFSGSGSGRDYSFSISNLEPEDIATYYC	LQYDNL---RT	FGGGTKLEIK
		** *		*		*^* *	
1YQV	DIVLTQSPAIMSASPGEKVTMTC	SASS------SVNYMY	WYQQKSGTSPKRWIY	DTSKLAS	GVPVRFSGSGSGTSYSLTISSMETEDAATYYC	QQWGRN---PT	FGGGTKLEIK
		* ***		*		^*	
1WEJ	DIQMTQSPASLSASVGETVTITC	RASG------NIHNYLA	WYQQKQGkSPQLLVY	NAKTLAD	GVPSRFSGSGSGTQYSLKINSLQPEDFGSYYC	QHFWSTP--WT	FGGGTKLEIK
		**	*	* *		^* * *	
3D9A	DIVLTQSPATLSVTPGNSVSLSC	RASQ------SIGNNLH	WYQQKSHESPRLLIK	YASQSIS	GIPSRFSGSGSGTDFTLSINSVETEDFGMYFC	QQSNSWP--YT	FGGGTKLEIK
		*				** *	
3LIZ	QIVLTQSPSSMYASLGERVTITC	KASQ------DINNYLS	WFQQKPGKSPKTLIY	RADRLVD	GVPSRVSGSGSGQDYSLTISSLEYEDLGIYYC	LQYDELP--YT	FGGGTKLEIK
		* *				*^ *	
2V17	DVQITQSPSYLAASPGETITINC	RASK------SIRKFLA	WYREKPGKTNKLLIY	SGSTLQS	GTPSRFSGSGSGTDFTLTISRLEPEDFAMYYC	QQHNDYP--LT	FGAGTKLELK
		* *^* *				^*** *	
1F58	DIVLTQSPASLAVSLGQRATISC	KASQGVDF--DGASFMN	WYQQKPGQPPKLLIF	AASTLES	GIPARFSGRGSGTDFTLNIHPVEEEDAATYYC	QQSHEDP--LT	FGAGTKLELK
		* * *				*^^ *	
3GGW	DIVMTQAAFSNPVTLGTSASISC	RSSKSLLHS-DGITYLY	WYLQKPGQSPHLLIY	HLSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGIYYC	AHNVELP--RT	FGGGTKLEIK
		* *		*		* *	
3HNS	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
		* *				*^** *	
3C6S	DIVMTQAAFSNPVTLGTSASISC	RSSKSLLHS-DGITYLY	WYLQKPGQSPHLLIY	HLSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGIYYC	AHNVELP--RT	FGGGTKLEIK
		*	*	*		* *	
3HNT	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
		*	*	*		*	
3HNV	DIQMTQTTSSLSASLGDRVTIGC	RASQ------DIGSYLN	WYQQKPDGAVRLLIY	YTSRLHS	GVPSRFSGSGSGTHFSLTISNLEQEDIGTYFC	HQDTKPP--YT	FGSGTKLEIK
		* * *				^ * *	
2OK0	DILMTQTPLSLPVSLGDQASISC	RSSQSIVHS-NGNTYLE	WYLQKPGQSPTLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC	FQGSHIP--LT	FGAGTKLEVK
		*				*^ * * *	
3CFB	DIVMTQAAFSNPVTLGTSASISC	RSTKSLLHS-NGITYLY	WYLQKPGQSPQLLIY	QMSNLAS	GVPDRFSSSGSGTDFTLRISRVEAEDVGVYYC	AQNLELP--PT	FGGGTKLEIK
		* * *				* ^*	
1KEL	DVLMTQTPLSLPVSLGDQASISC	RFSQSIVHS-NGNTYLE	WYLQKSGQSPKLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC	FQGSHVP--RT	FGGGTKLEIK
		* *				*^ * *	
1QYG	DIVMTQAAPSVPVTPGESVSISC	RSSKSLLHS-NGYTYLH	WFLQRPGQSPQLLIY	RVSNLAS	GVPDRFSGSGSGTAFTLRFSRVEAEDVGVYYC	MQHLEYP--FT	FGSGTKLEIK
		*				* * *	
1RIU	DIVMTQAAPSVPVTPGESVSISC	RSSKSLLHS-NGYTYLH	WFLQRPGQSPQLLIY	RVSNLAS	GVPDRFSGSGSGTAFTLRFSRVEAEDVGVYYC	MQHLEYP--FT	FGSGTKLEIK
		* *				*^^^ *	
2AJV	DIVITQDELSNPVTSGESVSISC	RSSRSLLYK-DGRTYLN	WFLQRPGQSPQLLIY	LMSTRAS	GVSDRFSGSGSGTDFTLEISRVKAEDVGVYYC	QQFVEYP--FT	FGSGTKLEIK
						* ^ *	
1NCW	DVVMTQSPKTISVTIGQPASISC	KSSQRLLNS-NGKTFLN	WLLQRPGQSPKRLIY	LGTKLDS	GVPDRFTGSGSGTDFTLKISRVEAEDLGVYYC	WQGTHFP--YT	FGGGTKLEIK
		*	*			*^* *	
1A3L	DIVLTQAAFSNPVTLGASASISC	RSSKSLLNS-NGIIHMY	WYLQKPGQSPQLLIY	QMSKLAS	GAPDRFSGSGSGTDFTLRISRVEAEDVGVYYC	AQNLELP--YT	FGGGTKLEIK
		* *				* ^^ * *	
1YEF	DIVMTQSPLTLSVTIGQPASISC	KSSQSLLYS-NGKTYLN	WLLQRPGQSPKRLIH	LVSKLDS	GVPDRITGSGSGTDFTLKISRVEAADLGVYYC	VQGTHFP--YT	FGGGTKLEIL
		*				*^^ * *	
1YEG	DIVMTQSPLTLSVTIGQPASISC	KSSQSLLYS-NGKTYLN	WLLQRPGQSPKRLIH	LVSKLDS	GVPDRITGSGSGTDFTLKISRVEAADLGVYYC	VQGTHFP--YT	FGGGTKLEIL
		** * *				*^** *	
3IFN	DVLMTQTPLSLPVSLGDQASISC	RSSQSIVHS-NGNTYLE	WYLQKPGQSPKLLIY	KVSNRFS	GVPDRFSGSGSGTDFTLKISRVEAEDLGIYYC	FQSSHVP--LT	FGAGTKLELK
		* * * *	* *	*		^^ * *	
3LEY	DVVMTQTPLSLSVTLGQPASISC	KSSQSLLDS-DGKTYLN	WLLQRPGQSPKRLIY	LVSKLAS	GVPDRFTGSGSGTDFTLKINRVEAEDLGIYYC	WQGTHFP--WT	FGGGTKLEIK
		*				^***	
3041	DIVLTQSPASLAVSLGQRATIFC	RASQSVDY--NGISYMH	WFQQKPGQPPKLLIY	AASNPES	GIPARFTGSGSGTDFTLNIHPVEEEDAATYYC	QQIIEDP--WT	FGGGTKLEIK
		** *				* *	
2ZPK	ZTVVTQESA-LTTSPGETVTLTC	RSSTGAV---TTSNYAN	WVQEKPDHLFTGLIV	GTNNRVP	GVPPRFSGSLIGDKAALTITGAQTEDEAIYFC	ALWYSNH--WV	FGGGTKLTVLG

Table 3. continued.

VH

		CDR-1		CDR-2		CDR-3	
	10 20 30	ab 35	40 49	52abc 60	66 70 82abc 90 94	95 100abcdefghijklmn101	103 113
(Human)							
2NXY	EVQLVESGAEVKKPGSSVKVSCKASGDTFI	--RYSFT	WVRQAPGQGLEWMG	RIIT--ILDVAHYAPHLQG	RVTITADKSTSTVYLELRNLRSDDTAVYFCAG	VYEGEADEGEYDNNGFL---KH	WGQGTLVTVSS
2NY2	EVQLVESGAEVKKPGSSVKVSCKASGDTFI	--RYSFT	WVRQAPGQGLEWMG	RIIT--ILDVAHYAPHLQG	RVTITADKSTSTVYLELRNLRSDDTAVYFCAG	VYEGEADEGEYDNNGFL---KH	WGQGTLVTVSS
3D85	EVQLQQSGPELVKPGASVKMSCKASGYTFT	--SNVMH	WVKQKPGQGLEWIG	YINP--YNDGTKYNEKFKG	KATLTSDKSSSTAYMELSSLTSEDSAVYYCAR	NWDV----------------AY	WGQGTLVTVSA
3MXW	QVQLQQSGPELVRPGVSVKISCKGSGYTFI	--DEALH	WVKQSHAESLEWIG	VIRP--YSGETNYNQKFKD	KATMTVDISSSTAYLELARLTSEDSAIYYCAR	DWERGDFF------------DY	WGQGTLVTVSS
1TJG	RITLKESGPPLVKPTQTLTLTCSFSGFSLS	DFGVGVG	WIRQPPGKALEWLA	IIY---SDDDKRYSPSLNT	RLTITKDTSKNQVVLVMTRVSPVDTATYFCAH	RRGPTTLFGVPIARGPVNAMDV	WGQGITVTISS
3G01	EVQLQESGPGLVKPSETLSLTCTVSGGPIN	--NAYWT	WIRQPPGKGLEYLG	YVY---HTGVTNYNPSLKS	RLTITIDTSRKQLSLSLKFVTAADSAVYYCAR	EWAEDGDFGNAF--------HV	WGQGTMVAVSS
3MLY	QVQLQESGPGLVKPSETLSLTCTVSGGSIS	--GFHWS	WIRQPPGKGLEYIG	YIY---YSGSTSYNPSLKS	RVSMSVDTSRNQFSLELSSVTAADTAVYYCAR	DFGEYHYDGRGFQCEGF---DL	WGQGTLVTVSS
(Murine)							
2VXT	EIQLQQSGPELVKPGASVKVSCKASGYSFT	--DYFIY	WVKQSHGKSLEWIG	DIDP--YNGDTSYNQKFRD	KATLTVDQSSTTAFMHLNSLTSEDSAVYFCAR	GL------------------RF	WGQGTLVTVSA
2ADF	QIQLVQSGPELKKPGETVKISCKASGYTFI	--NYGMN	WVKQAPGKGLKWMG	WKNT--NTGETTYGEEFRG	RFAFSLETSVSTAYLQINNLKNEDTATYFCAR	DNPYYAL-------------DY	WGQGTTVTVSS
1YQV	EVQLQQSGAELMKPGASVKISCKASGYTFS	--DYWIE	WVKQRPGHGLEWIG	EILP--GSGSTNYHERFKG	KATFTADTSSSTAYMQLNSLTSEDSGVYYCLH	GNYDF---------------DG	WGQGTTLTVSS
1WEJ	EVQLQQSGAELVKPGASVKLSCTASGFNIK	--DTYMH	WVKQRPEKGLEWIG	RIDP--ASGNTKYDPKFQD	KATITADTSSNTAYLQLSSLTSEDTAVYYCAG	YDYGNF--------------DY	WGQGTTLTVSS
3D9A	DVQLQESGPSLVKPSQTLSLTCSVTGDSIT	--SDYWS	WIRKFPGNRLEYMG	YVS---YSGSTYYNPSLKS	RISITRDTSKNQYYLDLNSVTTEDTATYYCAN	WDG-----------------DY	WGQGTLVTVSA
3LIZ	EVQLVESGGGLVQPGGSLKLSCAASGFTFS	--SFAMS	WGRQTPDKRLELVA	TINS--NGASTYYPDTVKG	RFTISRDNAKNTLFLQMSSLKSEDTAMYYCTR	DPAGRAWF------------AY	WGQGTLVTVSA
2V17	EVNLVESGGGLEQSGGSLSLSCAASGFTFT	--DYYMS	WVRQPPGKALEWLA	LIRNKAKGYTTEYSASVKG	RFTISRDNSQSILYLQMNALRAEDSAIYYCAR	DNGAARATF-----------AY	WGQGTLVTVSA
1F58	DVQLQQSGPDLVKPSQSLSLTCTVTGYSIT	-SGYSWH	WIRQFPGNKLEWMG	YIH---YSAGTNYNPSLKS	RISITRDTSKNQFFLQLNSVTTEDTATYYCAR	EEAMPYGNQAYYYAM-----DC	WGQGTTVTVSS
3GGW	EVKVEESGGGLVQPGGSMKISCVVSGLTFS	--NYWMS	WVRQSPEKGLEWVA	EIRLKSDNYATYYAESVKG	KFTISRDDSKSRLYLQMNNLRTEDTGIYYCFL	PM------------------DY	WGQGTSVTVSS
3HNS	EVQLQQSGTVLARPGTSVKMSCKASGYSFT	--NYWMH	WVKQRPGQGLEWIG	SIYP--GNSDTNYKQKFKG	KAKLTAVTSASTAYMEVNSLTNEDSAVYYCTR	FGNYVPF-------------AY	WGQGTLVTVSA
3C6S	EVKVEESGGGLVQPGGSMKISCVVSGLTFS	--NYWMS	WVRQSPEKGLEWVA	EIRLKSDNYATYYAESVKG	KFTISRDDSKSRLYLQMNNLRTEDTGIYYCFL	PM------------------DY	WGQGTSVTVSS
3HNT	EVQLQQSGTVLARPGTSVKMSCKASGYSFT	--NYWMH	WVKQRPGQGLEWIG	SIYP--GNSDTNYKQKFKG	KAKLTAVTSASTAYMEVNSLTNEDSAVYYCTR	FGNYVPF-------------AY	WGQGTLVTVSA
3HNV	EVQLQQSGTVLARPGTSVKMSCKASGYSFT	--NYWMH	WVKQRPGQGLEWIG	SIYP--GNSDTNYKQKFKG	KAKLTAVTSASTAYMEVNSLTNEDSAVYYCTR	FGNYVPF-------------AY	WGQGTLVTVSA
2OK0	EVQLEESGPELVKPGASVKISCKASGYTFT	--DYYMN	WLRQKPGQGLEWIG	WVY----PGSIKYNEKFKD	KATLTADTSSSIVYMHLSSLTSDDNAVYFCTR	WTYGSSF-------------DY	WGEGTLLTVSS
3CFB	EVKLVESGGGLVKPGGSLKLSCTASGITFS	--RYIMS	WVRQIPEKRLEWVA	SIS---SGGITYYPDSVKG	RFTISRDNVRNILYLQMSSLRSEDTALYYCAR	GQGR----------------PY	WGQGTLVTVSS
1KEL	EVKLVESGGGLGQPGGSLRLSCATSGFTFT	--DYYFN	WARQPPGKALEWLG	FIRNKAKGYTTEYSASVKG	RFTISRDNSQGILYLQMNTLRAEDSATYYCAR	WGSYAM--------------DY	WGQGTSVTVSS
1QYG	EVTLQESGGGLVQPGGSMKLSCAASGFTFS	--DAWVD	WVRQSPGKGLEWVA	EIRNKANNHATKYTESVKG	RFTISRDDSKSSVYLQMNSLRAEDTGIYYCTS	VPQLGRGF------------AY	WGQGTLVTVSA
1RIU	EVTLQESGGGLVQPGGSMKLSCAASGFTFS	--DAWVD	WVRQSPGKGLEWVA	EIRNKANNHATKYTESVKG	RFTISRDDSKSSVYLQMNSLRAEDTGIYYCTS	VPQLGRGF------------AY	WGQGTLVTVSA
2AJV	EVKLSESGPGLVKPSQSLSLTCTVTGYSIT	-TNYAWT	WIRQFPGNKLEWMG	YIR---SSVITRYNPSLKS	RISITQDTSKNQFFLQLNSVTTEDTATYYCAR	YDYYGNTG------------DY	WGQGTSVTVSS
1NCW	RVQLQQSGPGLVKPSQSLSLTCTVTGYSIT	-SDFAWN	WIRQFPGNKLEWMG	YIN---YSGFTSHNPSLKS	RISITRDTSKNQFFLQLNSVTTEDTATYYCAG	LLWYDGGA------------GS	WGQGTLVTVSA
1A3L	EVQLEESGPELVRPGTSVKISCKASGYTFT	--NYWLG	WVKQRPGHGFEWIG	DIYP--GGVYTTNNEKFRG	KAILTADTSSSTAYMQLSSLTSEDSAVYFCAR	AGGYYTGG------------DY	WGQGTSVTVSS
1YEF	EMQLQQSGAELLRPGTSVKLSCKTSGYIFT	--SYWIH	WVKQRSGQGLEWIA	RIYP--GTGSTYYNEKFKG	KATLTADKSSSTAYMQLSTLKSEDSAVYFCTR	WGFIPVREDYVM--------DY	WGQGTLVTVSS
1YEG	EMQLQQSGAELLRPGTSVKLSCKTSGYIFT	--SYWIH	WVKQRSGQGLEWIA	RIYP--GTGSTYYNEKFKG	KATLTADKSSSTAYMQLSTLKSEDSAVYFCTR	WGFIPVREDYVM--------DY	WGQGTLVTVSS
3IFN	QVTLKESGPGILKPSQTLSLTCSFSGFSLS	TSGMSVG	WIRQPSGKGLEWLA	HIW---WDDDKYYNPSLKS	RLTISKDTSRNQVFLKITSVDTADTATYYCAR	RTTTADYF------------AY	WGQGTTLTVSS
3LEY	DVQLQESGPGLVKPSQSLSLTCTVTGYLIT	-TDYAWN	WIRQFPGNKLEWMG	YIS---YSGFTSYNPSLKS	QISITRDTSKNQFFLQLNSVTTEDTATYYCAF	GNYLP---------------AY	WGQGTLVTVSA
3O41	QVTLKESGPGILQPSQTLSLTCSFSGFSLS	TSGMGVS	WIRQPSGKGLEWLA	HIY---WDDDKRYNPSLKS	RLTISKDTSRNQVFLKITSVDTADTATYYCAR	LYGFTYGF------------AY	WGQGTLVTVSA
2ZPK	ZIQLVQSGPEVQKPGETVRISCKASGYTFT	--TAGMQ	WVQKMPGKSLKWIG	WINT--RSGVPKYAEDFKG	RFAFSLETSASIAYLHINNLKNEDTATYFCAR	EGPGF---------------VY	WGQGTLVTVSA

See footnote to Table 2.

same antibody in different ligand states provide a better gauge of the flexibility and deformability of the CDR loops, the stability of the quaternary structure, and other structural features of the molecule.

The information provided here could help in the engineering of an antibody, even in the absence of a three-dimensional structure for that antibody. An examination of the structural data presented in the various tables shows that antibodies have structural features that are shared (as well as features that are very different). For example, many of the residues in the framework, which are involved in the contact with the opposite domain (Table 2), are located at the same position and share structural characteristics (size, polarity, etc). That information and the results presented in Table 3 are useful in deciding which framework residues to keep if the quaternary structure of the Fv is to be preserved. However, the residues in the CDRs, which are involved in the VL:VH contact and which also play a role in the

Table 4a. Contacts between Framework and CDR residues in the murine HyHEL-10 Fab [PDB entry 3D9A]

CDRs-L:	CDR1-L									CDR2-L					CDR3-L				
	Arg	Ala	Ser	Gln	*Ser*	Ile	*Gly*	Leu	His	Ala	*Ser*	Gln	Ile	Ser	Gln	Gln	Ser	Pro	Thr
	24	25	26	27	*28*	29	*30*	33	34	51	*52*	53	55	56	89	90	93	95	97
FRs-L:																			
Asp 1																		5	
Ile 2			1	3		1										3	2		
Val 3			1																
Leu 4		2						1							3	1			1
Ser 22	6																		
Tyr 36															1				
Leu 46													1						
Ile 48										2	3	2							
Lys 49									1			2	3						
Ile 58													4						
Ser 67							1												
Thr 69		2		2	2														
Asp 70	5																		
Phe 71						3	2	7		2									

CDRs-H:	CDR1-H				CDR2-H										CDR3-H			
	Ser	Asp	Trp	Ser	*Tyr*	Val	*Ser*	*Tyr*	*Ser*	*Gly*	Tyr	Asn	Ser	Leu	Asp	*Gly*	Asp	Tyr
	31	32	34	35	*50*	51	*52*	*53*	*54*	*55*	59	60	62	63	96	*97*	101	102
FRs-H:																		
Val 2																		7
Leu 4																		1
Val 24			2															
Asp 27	3	7	3															
Ser 28	2																	
Ile 37				1														
Glu 46												1	3	3				
Tyr 47				6														
Met 48														2				
Arg 66														1				
Ile 67											4			3				
Ile 69					1	2					3							
Arg 71						4	2	5	1	5								
Tyr 78			13	7														
Ala 93																		1
Asn 94			1												2	4	9	12

The residues along the horizontal are the complementarity-determining region (CDR) residues; those along the vertical are the framework (FR) residues. The residues, which are in contact by only their main chain, are italicized. In this matrix of contacts, the element *c(i,j)* the number of atom pairs, one from residue *i* the other from residue *j*. Nearest-neighbor contacts and CDR contacts with the intradomain disulfide bridge, with the "invariant tryptophan" (the tryptophan immediately following the CDR-1 in both chains), with the phenylalanine immediately following the CDR-3 in light chains, and with the tryptophan immediately following the CDR-3 in heavy chains, are excluded. The numbering scheme of Table 2 is used.

quaternary structure, often do not share a common location. This is especially true of the contact residues in the CDRs which have different lengths (CDR-1 and CDR-3 in the light chain, and all three CDRs in the heavy chain). A fortuitous similarity in the CDR lengths of an antibody listed in Table 2 with those of the antibody of unknown structure would be fortunate.

The antibody:ligand contacts shown in Table 3 show that the N-terminal segment of CDR-1 of the light chain and the C-terminal segment of the CDR-2 of the heavy chain are not involved in the interaction with ligand. This supports the suggestion that not the entire length of the CDRs need to be preserved when humanizing by CDR-grafting, but only those segments which had been found to be in contact with ligand (Padlan et al. 1995). Likewise, the frequent use of similarly located residues in the antibody-ligand contacts suggests that not all of the CDR residues need to be transferred during humanization, but only those which are deemed to be specificity-determining (Padlan et al. 1995). This information also helps in the humanization of an antibody of unknown structure.

Table 4b. Contacts between Framework and CDR residues in the murine 4C6 Fab [PDB entry 1NCW]

CDRs-L:	CDR1-L									CDR2-L					CDR3-L				
	Lys	Ser	Ser	Gln	Arg	Leu	Thr	Leu	Asn	*Gly*	Lys	Leu	Asp	Ser	*Trp*	Gln	His	Pro	Thr
	24	25	26	27	27a	27b	31	33	34	*51*	53	54	55	56	*89*	90	93	95	97
FRs-L:																			
Asp 1																		5	
Val 2			2	4												1	3		1
Met 4		2						1							1	4			2
Thr 5	1																		
Ser 22	2																		
Arg 46													5	7					
Ile 48										2	4								
Tyr 49								2			9	1	12						
Val 58												2	3	1					
Thr 69		4		2	3														
Asp 70	4																		
Phe 71						1	2	6											

CDRs-H:	CDR1-H					CDR2-H											CDR3-H
	Ser	Asp	Phe	Trp	Asn	Tyr	Ile	*Tyr*	*Ser*	*Gly*	Ser	*His*	*Asn*	Ser	Leu	*Lys*	*Ser*	Ser
	30b	31	32	34	35	50	51	*53*	*54*	*55*	58	*59*	*60*	62	63	*64*	*65*	102
FRs-H:																		
Val 2			1	1														2
Leu 4																		3
Val 24				1														
Tyr 27	3	5	6															
Ile 29				2														
Thr 30								1										
Arg 38															1			
Glu 46														3				
Trp 47					3		2				3	1	2					
Met 48															2			
Arg 66															2	1		
Ile 67															3			
Ile 69						2	3											
Arg 71							3	2	4	7								
Phe 78				12	1		1											
Asn 82A																	1	
Ala 93																		1

See footnote to Table 4a.

The contacts between framework and CDR residues shown in Tables 4a,b,c also reveal similarities that are useful in deciding which framework residues of a nonhuman antibody should be preserved during humanization. The preservation of critical framework residues, i.e., the ones which are involved in the interaction with the CDRs and the ones which dominate in the VL:VH contact (Tables 4d,e,f), is to ensure that the structures of the CDRs and, consequently, of the antigen-binding

Table 4c. Contacts between Framework and CDR residues in the murine 7A1 Fab [PDB entry 2AJU]

CDRs-L:	CDR1-L									CDR2-L					CDR3-L				
	Arg	Ser	Ser	Arg	*Ser*	Leu	Thr	Leu	Asn	Met	*Ser*	Thr	Arg	Ala	Gln	Gln	Val	Glu	Pro
	24	25	26	27	*27a*	27b	31	33	34	51	*52*	53	54	55	89	90	92	93	95
FRs-L:																			
Asp 1																			6
Ile 2		1		1		1										3	1	4	
Val 3			1																
Ile 4	2	3														1			
Thr 5	1																		
Asp 7	7																		
Phe 36									1						3				
Leu 46									1										
Ile 48										2	3	3	1						
Tyr 49									1			5							
Val 58														4					
Asp 60													2						
Thr 69		1			2														
Asp 70	3																		
Phe 71		1				3	1	7		1									

CDRs-H:	CDR1-H				CDR2-H									CDR3-H	
	Thr	Tyr	Trp	Thr	Tyr	Ile	Ser	*Arg*	Tyr	Asn	Ser	Leu		Tyr	Tyr
	30b	32	34	35	50	51	54	*58*	59	60	62	63		95	102
FRs-H:															
Val 2															6
Leu 4															4
Val 24			2												
Tyr 27	12	9	1												
Ser 28	1														
Ile 29			3												
Ile 37				1										2	
Arg 38												1			
Glu 46										2	3				
Trp 47				2	1			2	3	4					
Met 48											2				
Arg 66												4			
Ile 67									6			2			
Ile 69					2	2			3						
Gln 71						1	5								
Phe 78			9	4											
Ala 93															1
Arg 94		16	6												7

See footnote to Table 4a.

Table 4d. Contacts between VL and VH in the murine HyHEL-10 Fab [PDB entry 3D9A]

VH:	Lys	Asn	Leu	Tyr	*Gly*	Tyr	Tyr	Tyr	Asn	Pro	Tyr	Trp	Asp	*Gly*	Trp	*Gly*
	39	43	45	47	*49*	50	58	*59*	60	61	94	98	99	*100*	103	104
VL:																
Tyr 36														4	5	
Gln 38	5										2					
Ser 43															2	2
Pro 44															10	
Leu 46													1	1		
Met 85		3														
Phe 87		2	6													
Gln 89				1												
Trp 94				1	1	3	8	10	1							
Pro 95									1	1						
Tyr 96				6	2							6				
Phe 98			1	5												
Gly100		1														

The residues along the horizontal are from the VH; those along the vertical are from the VL. As in Table 4a, the residues which are in contact by only their main chain are italicized. In this matrix of contacts also, the element *c(i,j)* represents the number of atom pairs, one from residue *i* and the other from residue *j*.

Table 4e. Contacts between VL and VH in the murine 4C6 Fab [PDB entry 1NCW]

VH:	Arg	Ile	Gln	Asn	Leu	Trp	Tyr	Ser	*His*	Asn	Pro	Tyr	Leu	Tyr	Asp	*Gly*	*Gly*	*Ala*	*Gly*	Ser	Trp	*Gly*
	1	37	39	43	45	47	50	58	*59*	60	61	91	95	98	99	*100*	*100a*	*100b*	*101*	102	103	*104*
VL:																						
Lys 30														10								
Phe 32														2	3							
Asn 34													1			1						
Leu 36																					2	
Gln 38			8									1										
Gln 42												1										
Ser 43												4									2	2
Pro 44																					10	
Arg 46	4																	3	7	1		
Tyr 49															4		5					
Leu 50															5							
Asp 55																2		4				
Ser 56	3																					
Val 85				1																		
Tyr 87			1	10	6																	
Trp 89													1									
Phe 94						2		2	2		4											
Pro 95						3				4	1											
Tyr 96						10	1															
Phe 98		1			5																	
Gly100				2																		

See footnote to Table 4d.

site, are maintained.

The solvent exposures and the contacts with specific ligand are presented in Tables 5a and 5b for representative camelid VHH and shark IgNAR VH, respectively. The contacts between FR and hypervariable residues are presented in Tables 5c and 5d. The data presented in those tables will be useful in the engineering of those molecules.

Although the CL and CH1 domains are usually not subjected

Table 4f. Contacts between VL and VH in the murine 7A1 Fab [PDB entry 2AJU]

VH:	Gln 39	Asn 43	Leu 45	Trp 47	Arg 58	Asn 60	Pro 61	Tyr 91	Tyr 95	Asp 96	Tyr 97	Tyr 98	*Gly* *99*	Asn 100	Thr 100a	*Gly* *100b*	Trp 103	*Gly* *104*	Gln 105
VL:																			
Tyr 27D											1								
Tyr 32											5	2							
Asn 34													5						
Phe 36									1								4		
Gln 38	8							1											
Ser 43								4										3	4
Pro 44																	5		
Leu 46														2	3	1			
Tyr 49												1		10					
Leu 50												7							
Tyr 87	1	3	3																
Gln 89									6										
Phe 91									8	2	1		6						
Tyr 94					3		2												
Pro 95				9		3													
Phe 96				7															
Phe 98			4						1										

See footnote to Table 4d.

Table 5a. Solvent accessibilities and residues in contact with ligand in camelid VHH of known structure

```
              10        20         30         40       5052a     60          70         82abc      90          100abcdefghijklmno  103     110
               |         |          |          |        | |       |           |           |          |             |                  |       |
              FR1             [   CDR1  ]        FR2    [       CDR2       ]             FR3                   [         CDR3          ]   FR4
                                  *  *                         ^^ *                                                **  *****    *      *
1JTP  DVQLQASGGGSVQAGGSLRLSCAAS GYTIGPYCMG WFRQAPGKEREGVA AINMGGGITYYADSVKG RFTISQDNAKNTVYLLMNSLEPEDTAIYYCAA DSTIYASYYECGHLSTGGYGYDS WGQGTQVTVSS
      EpeBepE   Ebpp  pbpBeBEBE  BEE BbBB BBBbEE eepp BB BBBB    epbbpepBp  bBEBeBbeEppbBpBeBeeBpepBebeBbBBB BpeebEEpbpB p Bep  e Bpp b e bpBeBeE

                                 * * ^**                                                                       ^^*  * **^            ^
2P49  QVQLVESGGGLVQAGGSLRLSCAAS GYAYTYIYMG WFRQAPGKEREGVA AMDSGGGGTLYADSVKG RFTISRDKGKNTVYLQMDSLKPEDTATYYCAA GGYELRDRTY------------GQ WGQGTQVTVSS
      ebpBEBe   EbeE  pbpBpBEBE bEeeBEBB  BBBpeE eebp BB Bbbp     pebpeeBe  bBEBeBep pBbBbBpBpeBpepBpBeBbBBB   ppppbepB------------ p b e bpBeBeE

                                 * *^^**                                                                       ^^*    **^            ^*
2P4A  QVQLVESGGGLVQAGGSLRLSCAAS GYPWTYIYMG WFRQAPGKEREGVA AMDSGGGGTLYADSVKG RFTISRDKGKNTVYLQMDSLKPEDTATYYCAA GGDALVATRY------------GR WGQGTQVTVSS
      eBeBEBe   ebpE  pbpBpBEBe beeeBEbB  BBBpEE eepp BB BBbp     eebpepBp  BBEBeBpe ebbBbBpBpeBpEpBpBeBbBBB   pEbeeEpB------------ p b e bpBeBeE

      **                        **   *                        **** * *          ^^^^^*                           * **                *
1KXQ  QVQLVESGGGSVQAGGSLSLSCAAS TY---TDTVG WFRQAPGKEREGVA AIYRRTGYTYSADSVKG RFTLSQDNNKNTVYLQMNSLKPEDTGIYYCAT GNSVRLASWEGY----------FY WGQGTQVTVSS
      epeBEBe   EbeE  pbEBpBEBp ep---BbBB  BBBbEE pepb BB BBbBpe epppeeeBp  bBEBebbEEepbBpBpBeEBpEpBp pBbBBB  ebEeBEpbp B----------pp p e bpBeBeE

                                 *    *                       * *                                              ****^    ^^           *
1OP9  QVQLQESGGGSVQAGGSLRLSCSAS GY---TYISG WFRQAPGKEREGVA AIRSSDGTTYYADSVKG RFTISQDNAKNTVYLQMNSLKPEDTAMYYCAA TEVAGWPLDIGIY---------DY WGQGTEVTVSS
      eBpBeBe   EbeE  ebpBeBEpp  e---bbBb BBBpEE eebp BB BBbppp ppbbeepBp  bBEBEbbEEpebpbBpBeeBpepBpBpBbBBB BpEE eBbpe bB---------eB b e bpBeBee

2XA3  AVQLQESGGGLVQAGGSLRLSCTVS ARTSSSHDMG WFRQAPGKEREFVA AISWSGGTTNYVDSVKG RFDISKDNAKNAVYLQMNSLKPEDTAVYYCAA KWRPLRYSDNPSNSDY------NY WGQGTQVTVSS
      EBeBeBE   ebeE  eBpBpBEBp EeBeeeBBB BBBpEE eeppbBB BBBBe   eppbpepBp  BBeBebbpEepeBbBpBeeBpEpBpbpBbBBB bepEEpeEpeBbeeeB------pb b p bpBeBeE

3LN9  EVQLVESGGGLVQPGGSLRLSCTAS GYTFSHRYHR WFRQAPGKEREIVA VISQSGMRTYYADSVKG RFTISRDNAKNTVYLQMNSLKPEDTAMYYCAA GTRKNVWTRQHPF---------DY WGQGTQVTVSS
      pppBeBe   ebpe  ebpBpBeBe  BEBeebBBB BbBbpE peppeBB pBbpe eppbbbepBp  BBEBeBppEeebBbBpBeeBpepBpbpBbBBB  bpeeppBeeeee---------pb p e bpBeBpE
```

The solvent accessibility designations and contact labels follow those in Table 2. The CDR boundaries follow the convention of Desmyter et al. (2001). The numbering scheme of Table 2 is used. PDB entries 2XA3 and 3LN9 are unliganded.

to modification, we have included the results of similar analyses of the CL:CH1 module involving both a kappa- and a lambda-type light chain (Tables 6a,b,c), in the event that the engineering of those modules becomes desirable.

Currently, the Fc is also often being engineered for longer half-life and to control its interaction with receptors and other biologically important ligands. The solvent exposure of the residues in the CH2 and CH3 domains of various IgG-type Fc and in the CH2, CH3 and CH4 domains of IgE Fc, as well as the residues involved in the contact between the two chains in those fragments (Table 7a and 7b, respectively), will be useful in the engineering of the Fc. The identity of the residues that play the more significant role in the contact between the two chains in human IgG1 Fc (Table 7c) will greatly aid in the engineering of this Fc, which is currently the primary subject of attempts to improve the efficacy of therapeutic antibodies.

Knowledge of which residues are exposed on the surface and which could be replaced judiciously without unduly altering tertiary structure is useful in the design of molecules with reduced antigenicity, a procedure called "de-Antigenization" (Padlan 2008), especially in instances when a specific region or function is to be preserved (Padlan 2010), like the antigen-binding site of an antibody.

Table 5b. Solvent accessibilities and residues in contact with ligand in shark IgNAR VH of known structure

```
                10        20        30        40        50        60        70        80        90        100       110
                |         |         |         |         |         |         |         |         |         |         |
                     FR1          [ CDR1 ]     FR2     [  H2  ]    FR3     [H4]         FR4          [      CDR3      ]    FR5
                                   *  ^^                                                ^^^^^^ * **   ***^
1SQ2  RVDQTPRSVTKETGESLTINCVLR DASYALGS TCWYRKKSGE GNEESISK GGRYVETVN SGSK SFSLRINDLTVEDGGTYRCG LGVAGGYCDYALCSSRYAE CGDGTAVTVN
      pBpBEepEpEbpe ppBeBeBeBp eBEpEB p BBBBbbpe p  EBeeBEp   bbEBEee E pb eBpBpBpeBEEeB  eBBB  b pE  epppEbbEbpeBp B e BEBEBe
                                   *  ^^                                                ^^^^^^ * **   ***^
1T6V  RVDQTPRSVTKETGESLTINCVLR DASYALGS TCWYRKKSGE GNEESISK GGRYVETVN SGSK SFSLRINDLTVEDGGTYRCG LGVAGGYCDYALCSSRYAE CGDGTAVTVN
      pBpBEpeEbebpe ppBeBeBeBp ebEpEB p BBBBbpbe p  pBpeBEp   bbEBEee E pb pBeBpBpeBeEeB  eBbB  b pE  ppppEbbEbpeBp B e BEBeBe
                                    *    *                    *                             *^^^^^**      **
2I25  RVDQTPQRITKETGESLTINCVVR DSRCVLST GYWYRKPPGS RNEESISD GGRYVETVN RGSK SFSLRINDLTVKDSGTYRCK PESRYGSYDAVCAALNDQ- YGGGTVVTVN
      pBbBEpepbEbpp ppBeBpBpBp BEpBEBbp  pBbbppE p eEbeEbEp   bbEbEpe p eb pBeBpBpeBeEeBp eBbBb BbBpp EppEebeEBpep- b   BeBEBp

2I24  RVDQTPQRITKETGESLTINCVVR DSRCVLST GYWYRKPPGS RNEESISD GGRYVETVN RGSK SFSLRINDLTVKDSGTYRCK PESRYGSYDAVCAALNDQ- YGGGTVVTVN
      pBpBEeeppebpp ppBeBeBeBp ebpBeBpe  pBbbbeE p eEbeEBEp   bbebEpe p pb eBeBpBpeBeEeBb eBbBb BpBbe EepEEbEEBpep- b   BEBEBp
```

The solvent accessibility designations and contact labels follow those in Table 2. The CDR and hypervariable (H) region boundaries are as defined by Dooley et al. (2006). The numbering is sequential (the first residue is missing in all of these entries). PDB entry 2I24 is unliganded.

Table 5c. Contacts between framework and CDR residues in the camelid CAB-RN05 VHH [2P49]

	CDR1									CDR2									CDR3			
	Gly	*Tyr*	*Ala*	*Tyr*	**Thr**	**Tyr**	*Ile*	*Tyr*	**Met**	*Ala*	**Met**	*Asp*	**Ser**	**Thr**	*Leu*	**Tyr**	**Ser**	**Val**	**Asp**	**Tyr**	*Gly*	**Gln**
	26	*27*	*28*	*29*	30	31	*32*	*33*	34	*50*	51	*52*	52a	57	*58*	59	62	63	100a	100d	*101*	102
FR:																						
Val	2	2				2																3
Leu	*4*																					2
Arg	38																1					
Arg	45																		16			
Ala	49														1	2						
Arg	66																4	3				
Phe	67																	7				
Thr	68															1						
Ile	69									3	2			1		3						
Arg	71					3	2	2	6		3	4	8									
Lys	75		2	1	1																	
Asn	76	1		3	2	6																
Val	78								2													
Ala	93																			3		1
Ala	94					4																
Trp	103																		5	5	1	

The residues along the horizontal are the complementaritydetermining region (CDR) residues; those along the vertical are the framework (FR) residues. The residues, which are in contact by only their main chain, are italicized. In this matrix of contacts also, the element *c(i,j)* represents the number of atom pairs, one from residue *I* and the other from residue *j*. The CDR boundaries are as defined by Desmyter et al. (2001); only the CDR residues which are in contact are listed. The numbering scheme of Table 5a is used.

Table 5d. Contacts between the framework and CDR and hypervariable (H)regions in the shark IgNAR PBLA8 VH [2I24]

	CDR1					H2								H4			CDR3			
	Asp	Ser	Leu	*Ser*	Thr	*Arg*	Asn	Glu	*Glu*	Ser	Ile	*Ser*	Asp	*Gly*	Ser	*Lys*	Pro	*Asn*	*Asp*	Gln
	26	27	31	*32*	33	*44*	45	46	*47*	48	49	*50*	51	*62*	63	*64*	85	*100*	101	102
FR:																				
Arg 2	3																			
Val 3																	1			4
Val 23															1	1				
Val 24		1														2				
Tyr 35					2															
Trp 36											1									
Tyr 37								4	3	3										
Lys 39						4	2													
Pro 40							1													
Ser 43							5													
Tyr 55											3	3								
Glu 57												1	7							
Val 59				1	1															
Asn 60														3	9					
Ser 65															2					
Phe 66			1														4			
Leu 68											1									
Tyr 81											2									
Arg 82								12												
Lys 84																		3		
Tyr104																			2	

The residues along the horizontal are the hypervariable (H) and complementarity-determining region (CDR) residues; those along the vertical are the framework (FR) residues. The residues, which are in contact by only their main chain, are italicized. In this matrix of contacts also, the element *c(i,j)* represents the number of atom pairs, one from residue *i* and the other from residue *j*. The CDR and hypervariable (H) region boundaries are as defined by Dooley et al. (2006); only the residues which form contacts are listed. The numbering scheme of Table 5b is used.

Table 6a. Solvent accessibilities of the residues in the CL:CH1 module

CL:

```
          110       120       130       140       150       160       170       180       190       200       210
           |         |         |         |         |         |         |         |         |         |         |
3MLY      QPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVRAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
          bEpeEBeBeBBeBBepBpeEEeBBBBBBBbpBBB EBEBEBpebEEEbeE pppbEpepBEpebbBBBBBBEBebepBeebppBBBbBeBp BeeppebEpEeE

          110       120       130       140       150       160       170       180       190       200       210
           |         |         |         |         |         |         |         |         |         |         |
1WEJ      RADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC
          bEpeEBEEBbBBpBBEpBpEe  BBBBBBBBpBBepeBEBpBbBp epbeE pBbBebebbEeppBbBBBBBBpBEpebBppbepBpBbBEBppeEEepepebebeeE
```

CH1:

```
          120       130       140       150       160       170       180       190       200       210
           |         |         |         |         |         |         |         |         |         |
3MLY      ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
          EEep BeBbbBbBEEpEEe  EBeB BBBbpBbepEBEBeBbE EbEE bppBeBBebEp pbBBBBbBEBEbEpp EpeBpBbBpBepbEpeeebpBpEeEE

          120       130       140       150       160       170       180       190       200       210
           |         |         |         |         |         |         |         |         |         |
1WEJ      AETTPPSVYPLAPGTAALKSSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSDLYTLTSSVTVPSSTWPSQTVTCNVAHPASSTKVDKKIVPRNC
          beepEBeBbbBpB EEEBeEEpBBB BBBb BBepEBEBeBbE epEE bbpBeBeppEpebbBBBBBebEpeeBbeeEBeBbBEBEEBEbepebpBEppEp
```

The solvent accessibility designations follow those in Table 2. PDB entry 1WEJ is a murine Fab with a kappa light chain; PDB entry 3MLY is a human Fab with a lambda light chain. The numbering schemes are sequential and follow those in the respective PDB entries.

Table 6b. Residue contacts in the CL:CH1 of PDB entry 1WEJ

CH1:	Tyr	Pro	Leu	*Ala*	Pro	Thr	Leu	Lys	His	*Thr*	Phe	Pro	Val	Gln	Thr	Thr	*Ser*	Ser	Lys	Arg	Cys
	126	127	128	*129*	130	141	145	147	168	*169*	170	171	173	175	180	182	*183*	184	212	217	219
CL:																					
Ser 116						3															
Phe 118			6	5	1	2															
Pro 119																				4	
Ser 121	3	1																			
Glu 123		2																	1		
Gln 124	15							1													
Ser 131							1	2													
Phe 135			2								2					2	2	4			
Asn 137									6		1							2			
Asn 138									7												
Val 159														2							
Leu 160													2	5	2						
Asn 161													1								
Ser 162											4	8									
Trp 163												3									
Thr 164										1	2										
Ser 174									7		4										
Met 175											6										
Ser 176											6					4					
Thr 180								3													
Cys 214																					4

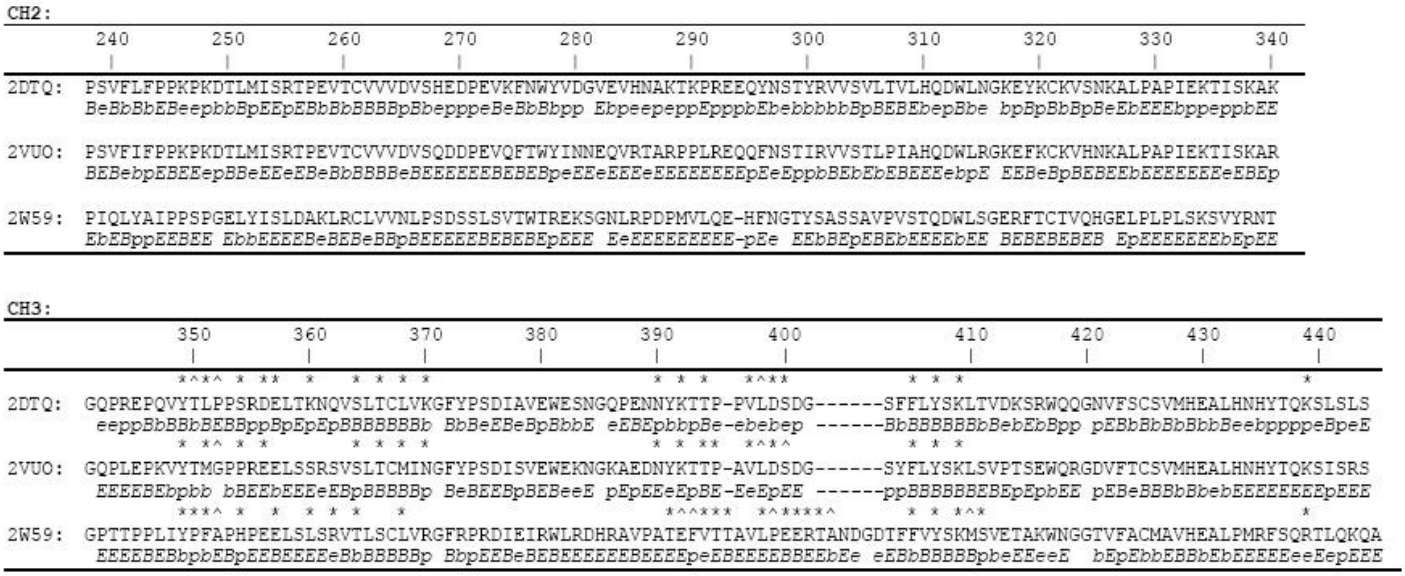
The residues along the horizontal are from the CH1; those along the vertical are from the CL. The residues, which are in contact by only their main chain, are italicized. In this matrix of contacts also, the element *c(i,j)* represents the number of atom pairs, one from residue *i* and the other from residue *j*.

Table 6c. Residue contacts in the CL:CH1 of PDB entry 3MLY

CH1:	Phe	Pro	Leu	*Ala*	*Ala*	*Gly*	Leu	Lys	His	Phe	Pro	*Ala*	Val	*Leu*	Gln	Ser	*Leu*	Ser	Val	Lys	Cys
	122	123	124	*125*	*137*	*139*	141	143	164	166	167	*168*	169	*170*	171	172	*178*	179	181	214	216
CL:																					
Phe 118			5	2	2	1													1		
Ser 121	2	1																		3	
Ser 122																				4	
Glu 123	3	5																		1	
Glu 124	10							1													
Thr 131								2													
Val 133																		1			
Leu 135										1								1	1		
Ile 136										2											
Ser 137									2												
Glu 160													1	1	4	5					
Thr 162												1	3								
Ser 165											3										
Gln 167									3												
Ala 173									2	1											
Ala 174										2											
Ser 175											1										
Tyr 177							1						2					3	6		
Cys 211																					6

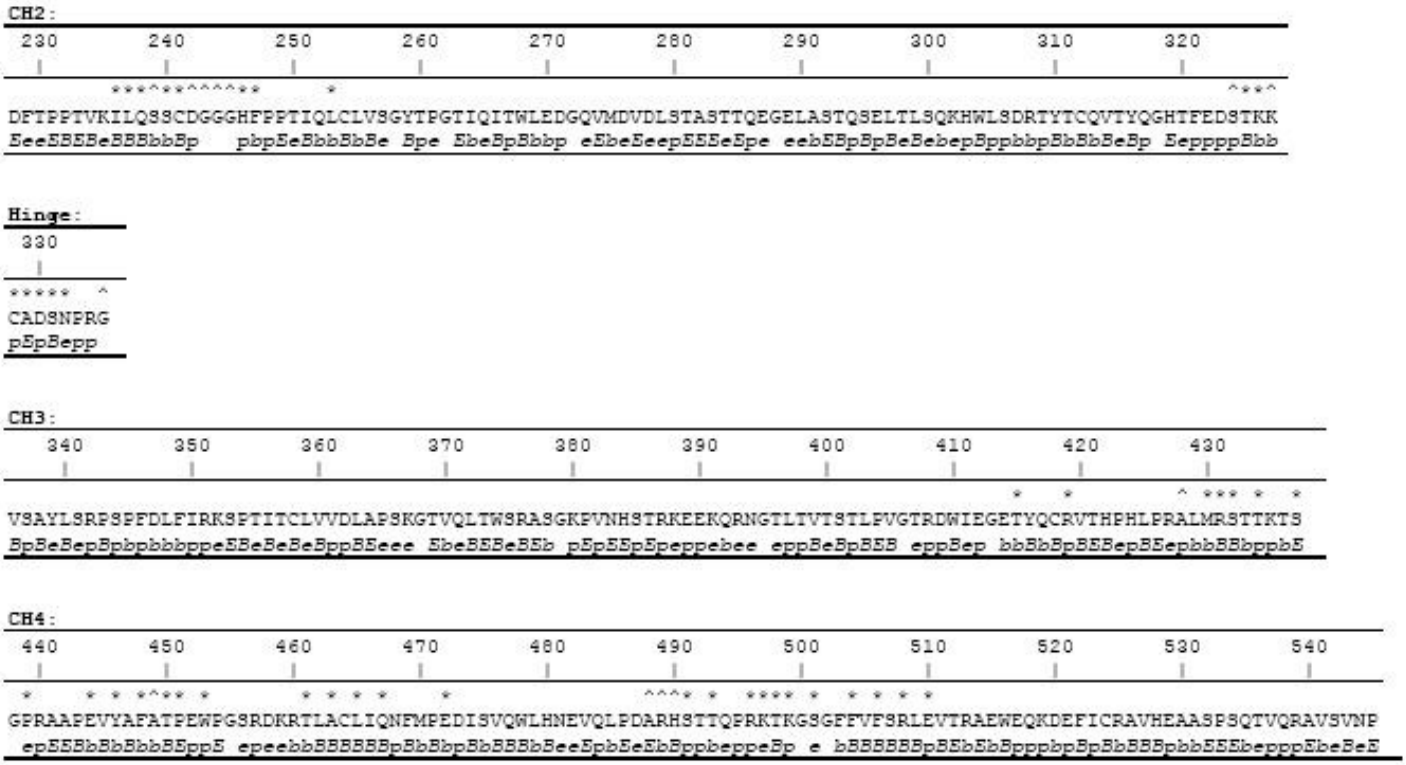
See footnote to Table 6b.

Table 7a. Solvent accessibilities of the CH2 and CH3 domains and contacts with the opposite chain in various Fc



The solvent accessibility designations and contact labels follow those in Table 2. The numbering scheme follows that in PDB entry 2DTQ. The results shown are for the first chain in the Fc.

Table 7b. Solvent accessibilities of the residues in the first chain of human IgE Fc (PDB entry 2WQR) and contacts with the opposite chain



The solvent accessibility designations and contact labels follow those in Table 2. The numbering scheme follows that in PDB entry 2WQR. The results shown are for the first chain in the Fc.

The ability of a single domain to bind with high affinity and specificity to an antigen has clear advantages. The synthesis of a single domain would be simpler in comparison to paired chains. And, when used in medical diagnosis or therapy, the smaller size would allow for greater penetration into tissues for binding to deeply situated targets. When camelid and shark molecules are used in human therapy, it would be desirable to minimize their immunogenicity. If a sufficiently similar human sequence could be found, humanization by CDR-grafting might suffice. Otherwise, veneering/resurfacing may be more appropriate. Indeed, the veneering of a camel VHH has been attempted (Conrath et al. 2005). In this regard, the judicious replacement of the exposed residues with amino acids that are expected not to change the overall structure of the molecule (Padlan 2008, 2010) would be particularly useful, especially in the case of the shark IgNAR VH which is significantly more different from human domains than the camelid VHH.

CONCLUSION

The availability of more high-resolution structures on antibodies allows for a better assessment of antibody structure and function. However, the picture is far from complete. As of this writing, no structure of a whole antibody molecule has been done to high resolution. Further, the various antibody types and subtypes are sufficiently different and have different functions, so that high-resolution structures for all of them would be desirable. A complete analysis should also include the interaction of the Fc fragments of the various antibody types (and subtypes) with their receptors. More high-resolution data on antibody molecules and their interactions with specific ligand and other molecules are eagerly anticipated.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest relating to this work.

CONTRIBUTION OF INDIVIDUAL AUTHORS

All the authors participated in gathering the crystallographic data from the Protein Data Bank; Eduardo A. Padlan performed the calculations; Jo Erika T. Narciso and Eduardo A. Padlan prepared the tables; Iris Diana C. Uy prepared the figures; Jo Erika T. Narciso, Eduardo A. Padlan, and Gisela P. Padilla-Concepcion wrote the manuscript; Gisela P. Padilla-Concepcion provided overall supervision.

Table 7c. Contacts between the two heavy chains in human IgG1 Fc [PDB entry 2DTQ]

Chain1:	Tyr	*Thr*	Leu	*Pro*	Ser	Asp	Glu	Lys	Ser	Thr	Leu	Lys	Asn	Lys	Thr	Val	*Leu*	Asp	Ser	Phe	Tyr	Lys	Lys	BMA	MAN
	349	*350*	351	*352*	354	356	357	360	364	366	368	370	390	392	394	397	*398*	399	400	405	407	409	439	3	4
Chain2:																									
Gln347								4																	
Tyr349					3	2	8	1																	
Thr350					1																				
Leu351			1	1	2					1															
Pro352			1																						
Ser354	6	1	2																						
Asp356	2																						3		
Glu357	9											2													
Lys360	1																								
Ser364											2	2													
Thr366			1																		7				
Leu368									2													1			
Lys370									1																
Asn390																			1						
Lys392																	3	1		3					
Thr394															4	3									
Pro395																1									
Val397															2										
Leu398														2											
Asp399														2								4			
Ser400													2	1											
Phe405														3								3			
Tyr407										8											14	5			
Lys409																		4		3	4				
Lys439						1																			
BMA 3																									1
MAN 4																								1	7

The residues along the horizontal are from the first chain of the Fc; those along the vertical are from the second chain. The residues, which are in contact by only their main chain, are italicized. In this matrix of contacts also, the element *c(i,j)* represents the number of atom pairs, one from residue *I* and the other from residue *j*. The numbering scheme follows that in PDB Entry 2DTQ. BMA is β-D-mannose; MAN is α-D-mannose.

REFERENCES

Acchione M, Lipschultz CA, DeSantis ME, Shanmuganathan A, Li M, Wlodawer A, Tarasov S, Smith-Gill SJ. Light chain somatic mutations change thermodynamics of binding and water coordination in the HyHEL-10 family of antibodies. Mol Immunol 2009; 47:457-464.

Acierno JP, Braden BC, Klinke S, Goldbaum FA, Cauerhff A. Affinity maturation increases the stability and plasticity of the Fv domain of anti-protein antibodies. J Mol Biol 2007; 374:130-146.

Alzari PM, Lascombe M-B, Poljak RJ. Three-dimensional structure of antibodies. Annu Rev Immunol 1988; 6:555-580.

Amzel LM, Poljak RJ. Three-dimensional structure of immunoglobulins. Annu Rev Biochem 1979; 48:961-997.

Argiriadi MA, Xiang T, Wu C, Ghayur T, Borhani DW. Unusual water-mediated antigenic recognition of the proinflammatory cytokine interleukin-18. J Biol Chem 2009; 284:24478-24489.

Basi GS, Feinberg H, Oshidari F, Anderson J, Barbour R, Baker J, Comery TA, Diep L, Gill D, Johnson-Wood K, Goel A, Grantcharova K, Lee M, Li J, Partridge A, Griswold-Prenner I, Piot N, Walker D, Widom A, Pangalos MN, Seubert P, Jacobsen JS, Schenk D, Weis WI. Structural correlates of antibodies associated with acute reversal of amyloid beta-related behavioral deficits in a mouse model of Alzheimer disease. J Biol Chem 2010; 285:3417-3427.

Bhat TN, Bentley GA, Boulot G, Greene MI, Tello D, Dall'Acqua W, Souchon H, Schwarz FP, Mariuzza RA, Poljak RJ. Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci USA 1994; 91:1089-1093.

Bhat TN, Bentley GA, Fischmann TO, Boulot G, Poljak RJ. Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding. Nature 1990; 347:483-485.

Bhat TN, Padlan EA, Davies DR. Refined Crystal Structure of the Galactan-Binding Immunoglobulin Fab J539 at 1.95-Angstroms Resolution. (To be published) cited in PDB Entry 2FBJ.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucl Acids Res 2000; 28:235-242.

Beyer BM, Ingram R, Ramanathan L, Reichert P, Le HV, Madison V, Orth P. Crystal structures of the pro-inflammatory cytokine interleukin-23 and its complex with a high-affinity neutralizing antibody. J Mol Biol 2008; 382:942-955.

Brams P, Black A, Padlan EA, Hariharan K, Leonard J, Chambers-Slater K, Noelle RJ, Newman R. A humanized anti-human CD154 monoclonal antibody blocks CD154-CD40 mediated human B cell activation. Int Immunopharmacol 2001; 1:277-294.

Capra JD, Edmundson AB. The antibody combining site. Sci Am 1977; 236(1):50-59.

Cohen GH, Silverton EW, Padlan EA, Dyda F, Wibbenmeyer JA, Willson RC, Davies DR. Water molecules in the antibody-antigen interface of the structure of the Fab HyHEL-5-lysozyme complex at 1.7 Å resolution: comparison with results from isothermal titration calorimetry. Acta Crystallogr D Biol Crystallogr 2005; 61:628-633.

Colman PM. Structure of antibody-antigen complexes: implications for immune recognition. Adv Immunol 1988; 43:99-132.

Connolly ML. Analytical molecular surface calculation. J Appl Crystallogr 1983; 16:548-558.

Conrath K, Vincke C, Stijlemans B, Schymkowitz J, Decanniere K, Wyns L, Muyldermans S, Loris R. Antigen binding and solubility effects upon the veneering of a camel VHH in framework-2 to mimic a VH. J Mol Biol 2005; 350:112-125.

Davies DR, Padlan EA, Segal DM. Three-dimensional structure of immunoglobulins. Annu Rev Biochem 1975a; 44:639-667.

Davies DR, Padlan EA, Segal DM. Immunoglobulin structures at high resolution. Contemp Top Mol Immunol 1975b; 4:127-153.

Davies DR, Padlan EA, Sheriff S. Antibody-antigen complexes. Annu Rev Biochem 1990; 59:439-73.

Davies DR, Sheriff S, Padlan EA. Antibody-antigen complexes. J Biol Chem 1988; 263:10541-10544.

Debler EW, Kaufmann GF, Meijler MM, Heine A, Mee JM, Pljevaljcic G, Di Bilio AJ, Schultz PG, Millar DP, Janda KD, Wilson IA, Gray HB, Lerner RA. Deeply inverted electron-hole recombination in a luminescent antibody-stilbene complex. Science 2008; 319:1232-1235.

Decanniere K, Transue TR., Desmyter A, Maes D, Muyldermans S, Wyns L. Degenerate interfaces in antigen-antibody complexes. J Mol Biol 2001; 313:473-478.

De Pascalis R, Iwahashi M, Tamura M, Padlan EA, Gonzales NR, Santos AD, Giuliano M, Schuck P, Schlom J, Kashmiri SV. Grafting of "abbreviated" complementarity-determining regions containing specificity-determining residues essential for ligand contact to engineer a less immunogenic humanized monoclonal antibody. J Immunol 2002; 169:3076-3084.

Desmyter A, Decanniere K, Muyldermans S, Wyns L. Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. J Biol Chem 2001; 276:26285-26290.

Desmyter A, Spinelli S, Payan F, Lauwereys M, Wyns L, Muyldermans S, Cambillau C. Three camelid VHH domains in complex with porcine pancreatic α-amylase. J Biol Chem 2002; 277:23645-23650.

Dooley H, Stanfield RL, Brady RA, Flajnik MF. First molecular and biochemical analysis of in vivo affinity maturation in an ectothermic vertebrate. Proc Natl Acad Sci USA 2006; 103:1846-1851.

Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D,

Larsson G, Spencer A, Archer DB, Sasse J, Muyldermans S, Wyns L, Redfield C, Matagne A, Robinson CV, Dobson CM. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. Nature 2003; 424:783-788.

Foote J, Winter G. Antibody framework residues affecting the conformation of the hypervariable loops. J Mol Biol 1992; 224:487-499.

Fotinou C, Beauchamp J, Emsley P, deHaan A, Schielen WJ, Bos E, Isaacs NW. Structure of an Fab fragment against a C-terminal peptide of hCG at 2.0 Å resolution. J Biol Chem 1998; 273:22515-22518.

Gigant B, Charbonnier JB, Eshhar Z, Green BS, Knossow M. X-ray structures of a hydrolytic antibody and of complexes elucidate catalytic pathway from substrate binding and transition state stabilization through water attack and product release. Proc Natl Acad Sci USA 1997; 94:7857-7861.

Girardi E, Holdom MD, Davies AM, Sutton BJ, Beavil AJ. The crystal structure of rabbit IgG-Fc. Biochem J 2009; 417:77-83.

Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 1997; 18:2714-2723.

Harris LJ, Larson SB, Hasel KW, McPherson A. Refined structure of an intact IgG2a monoclonal antibody. Biochemistry 1997; 36:1581-1597.

Haupt C, Morgado I, Kumar ST, Parthier C, Bereza M, Hortschansky P, Stubbs MT, Horn U, Fändrich M. Amyloid fibril recognition with the conformational B10 antibody fragment depends on electrostatic interactions. J Mol Biol 2011; 405:341-348.

Heine A, Stura EA, Yli-Kauhaluoma JT, Gao C, Deng Q, Beno BR, Houk KD, Janda KD, Wilson IA. An antibody exo Diels-Alderase inhibitor complex at 1.95 Angstrom resolution. Science 1998; 279:1934-1940.

Hinz A, Lutje Hulsik D, Forsman A, Koh WW, Belrhali H, Gorlani A, de Haard H, Weiss RA, Verrips T, Weissenhorn W. Crystal structure of the neutralizing Llama V(HH) D7 and its mode of HIV-1 gp120 interaction. PLoS One 2010; 5(5):e10482.

Holdom MD, Davies AM, Nettleship JE, Bagby SC, Dhaliwal B, Girardi E, Hunt J, Gould HJ, Beavil AJ, McDonnell JM, Owens RJ, Sutton BJ. Conformational changes in IgE contribute to its uniquely slow dissociation rate from receptor FcεRI. Nat Struct Mol Biol 2011; 18:571-576.

Hsieh-Wilson LC, Schultz PG, Stevens RC. Insights into antibody catalysis: structure of an oxygenation catalyst at 1.9-Å resolution. Proc Natl Acad Sci USA 1996; 93:5363-5367.

Huber R. Antibody structure. Trends Biochem Sci 1976; 1:174-178.

Huber R, Deisenhofer J, Colman PM, Matsushima M, Palm W. Crystallographic structure studies of an IgG molecule and an Fc fragment. Nature 1976; 264:415-420.

Huse WD, Sastry L, Iverson SA, Kang AS, Alting-Mees M, Burton DR, Benkovic SJ, Lerner RA. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science 1989; 246:1275-1281.

Jencks WP. Catalysis In Chemistry And Enzymology. New York: McGraw-Hill, 1969.

Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature 1986; 321:522-525.

Jordan JL, Arndt JW, Hanf K, Li G, Hall J, Demarest S, Huang F, Wu X, Miller B, Glaser S, Fernandez EJ, Wang D, Lugovskoy A. Structural understanding of stabilization patterns in engineered bispecific Ig-like antibody molecules. Proteins 2009; 77:832-841.

Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. Sequences of Proteins of Immunological Interest. 5th ed. US Department of Health and Human Services, Public Health Service, National Institutes of Health (NIH Publication No. 91-3242) 1991.

Kodandapani R, Veerapandian B, Kunicki TJ, Ely KR. Crystal structure of the OPG2 Fab. An antireceptor antibody that mimics an RGD cell adhesion site. J Biol Chem 1995; 270:2268-2273.

Koide A, Tereshko V, Uysal S, Margalef K, Kossiakoff AA, Koide S. Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar KD of a single-domain antibody with a flat paratope. J Mol Biol 2007; 373:941-953.

Koehler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 1975; 256:495-497.

Koshland DE Jr, Némethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 1966; 5:365-385.

Li M, Gustchina A, Glesner J, Wünschmann S, Vailes LD, Chapman MD, Pomés A, Wlodawer A. Carbohydrates contribute to the interactions between cockroach allergen Bla g 2 and a monoclonal antibody. J Immunol 2011; 186:333-340.

Luo J, Obmolova G, Huang A, Strake B, Teplyakov A, Malia T, Muzammil S, Zhao Y, Gilliland GL, Feng Y. Coevolution of antibody stability and Vκ CDR-L3 canonical structure. J Mol Biol 2010; 402:708-719.

Matsumiya S, Yamaguchi Y, Saito J, Nagano M, Sasakawa H, Otaki S, Satoh M, Shitara K, Kato K. Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. J Mol Biol 2007; 368:767-779.

McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. Nature 1990; 348:552-554.

McLellan JS, Chen M, Chang JS, Yang Y, Kim A, Graham BS, Kwong PD. Structure of a major antigenic site on the

respiratory syncytial virus fusion glycoprotein in complex with neutralizing antibody 101F. J Virol 2010; 84:12236-12244.

Menez R, Bossus M, Muller BH, Sibai G, Dalbon P, Ducancel F, Jolivet-Reynaud C, Stura EA. Crystal structure of a hydrophobic immunodominant antigenic site on hepatitis C virus core protein complexed to monoclonal antibody 19D9D6. J Immunol 2003; 170:1917-1924.

Murase T, Zheng RB, Joe M, Bai Y, Marcus SL, Lowary TL, Ng KK. Structural insights into antibody recognition of mycobacterial polysaccharides. J Mol Biol 2009; 392:381-392.

Mylvaganam SE, Paterson Y, Getzoff ED. Structural basis for the binding of an anti-cytochrome c antibody to its antigen: crystal structures of FabE8-cytochrome c complex to 1.8 A resolution and FabE8 to 2.26 A resolution. J Mol Biol 1998; 281:301-322.

Narciso JET, Uy IDC, Cabang AB, Chavez JFC, Pablo JLB, Padilla-Concepcion GP, Padlan EA. Analysis of the antibody structure based on high-resolution crystallographic studies. New BIOTECHNOLOGY 2011; 28(5):435-447.

Nogi T, Sangawa T, Tabata S, Nagae M, Tamura-Kawakami K, Beppu A, Hattori M, Yasui N, Takagi J. Novel affinity tag system using structurally defined antibody-tag interaction: application to single-step protein purification. Protein Science 2008; 17:2120-2126.

Ofek G, Guenaga FJ, Schief WR, Skinner J, Baker D, Wyatt R, Kwong PD. Elicitation of structure-specific antibodies by epitope scaffolds. Proc Natl Acad Sci USA 2010; 107:17880-17887.

Ofek G, Tang M, Sambor A, Katinger H, Mascola JR, Wyatt R, Kwong PD. Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. J Virol 2004; 78:10724-10737.

Padlan EA. Structural basis for the specificity of antibody-antigen reactions and structural mechanisms for the diversification of antigen-binding specificities. Q Rev Biophys 1977; 10:35-65.

Padlan EA. On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. PROTEINS: Struc Funct Genet 1990; 7:112-124.

Padlan EA. A possible procedure for reducing the immunogenicity of antibody variable domains while preserving their ligand-binding properties. Mol Immunol 1991; 28:489-498.

Padlan EA. Anatomy of the antibody molecule. Mol Immunol 1994; 31:169-217.

Padlan EA. A Novel Method for Designing Vaccines Against Constantly Mutating Pathogens. Phil J Sci 2008; 137:39-51.

Padlan EA. A method for designing molecules for use in directing the antibody response to a chosen region of a protein antigen. Phil Sci Letts 2010; 3(2):36-47.

Padlan EA, Abergel C, Tipper JP. Identification of specificity-determining residues in antibodies. FASEB J 1995; 9:133-139.

Poljak RJ. X-ray crystallographic studies of immunoglobulins. Contemp Top Mol Immunol 1973; 2:1-26.

Poljak RJ. X-ray diffraction studies of immunoglobulins. Adv Immunol 1975; 21:1-33.

Poljak RJ, Amzel LM, Phizackerley RP. Studies on the three-dimensional structure of immunoglobulins. Prog Biophys Mol Biol 1976; 31:67-93.

Pozharski E, Hewagama A, Shanafelt A, Petsko G, Ringe D. Carving a Binding Site: Structural Study of an Anti-Cocaine Antibody in Complex with Three Cocaine Analogs. (To be published) cited in PDB Entry 1RIU.

Pozharski E, Moulin A, Hewagama A, Shanafelt AB, Petsko GA, Ringe D. Diversity in hapten recognition: structural study of an anti-cocaine antibody M82G2. J Mol Biol 2005; 349:570-582.

Queen C, Schneider WP, Selick HE, Payne PW, Landolfi NF, Duncan JF, Avdalovic NM, Levitt M, Junghaus RP, Waldmann TA. A humanized antibody that binds to the interleukin 2 receptor. Proc Natl Acad Sci USA 1989; 86:10029-10033.

Riboldi-Tunnicliffe A, Isaacs NW. Antibody Structure. (To be published) cited in PDB Entry 3I75.

Rini JM, Schulze-Gahmen U, Wilson IA. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. Science 1992; 255:959-965.

Roguska MA, Pedersen JT, Keddy CA, Henry AH, Searle SJ, Lambert JM, Goldmacher VS, Blaettler WA, Rees AR, Guild BC. Humanization of murine monoclonal antibodies through variable domain resurfacing. Proc Natl Acad Sci USA 1994; 91:969-973.

Sanguineti S, Centeno Crowley JM, Lodeiro Merlo MF, Cerutti ML, Wilson IA, Goldbaum FA, Stanfield RL, de Prat-Gay G. Specific recognition of a DNA immunogen by its elicited antibody. J Mol Biol 2007; 370:183-195.

Sevcik J, Skrabana R, Dvorsky R, Csokova N, Iqbal K, Novak M. X-ray structure of the PHF core C-terminus: insight into the folding of the intrinsically disordered protein tau in Alzheimer's disease. FEBS Lett 2007; 581:5872-5878.

Skrabana R, Dvorsky R, Sevcik J, Novak M. Monoclonal antibody MN423 as a stable mold facilitates structure determination of disordered tau protein. J Struct Biol. 2010 Feb 23. (Epub ahead of print) cited in PDB Entry 3L1O.

Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 1985; 228:1315-1317.

Staelens S, Hadders MA, Vauterin S, Platteau C, De Maeyer M, Vanhoorelbeke K, Huizinga EG, Deckmyn H. Paratope determination of the antithrombotic antibody 82D6A3 based on the crystal structure of its complex with the von Willebrand factor A3-domain. J Biol Chem 2006; 281:2225-2231.

Stanfield R, Cabezas E, Satterthwait A, Stura E, Profy A, Wilson

I. Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing fabs. Structure 1999; 7:131-142.

Stanfield RL, Calarese DA, Jiang S, Wilson IA. Crystal Structure of anti-gp41 Fab NC-1. (To be published) cited in PDB Entry 3OZ9.

Stanfield RL, Dooley H, Flajnik MF, Wilson IA. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. Science 2004; 305:1770-1773.

Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. J Mol Biol 2007; 367:358-372.

Stanfield RL, Takimoto-Kamimura M, Rini JM, Profy AT, Wilson IA. Major antigen-induced domain rearrangements in an antibody. Structure 1993; 1:83-93.

Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, Abola EE. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. Acta Crystallogr D Biol Crystallogr 1998; 54(Pt 6 Pt 1):1078-1084.

Taylor AI, Fabiane SM, Sutton BJ, Calvert RA. The crystal structure of an avian IgY-Fc fragment reveals conservation with both mammalian IgG and IgE. Biochemistry 2009; 48:558-562.

Theillet FX, Saul FA, Vulliez-Le Normand B, Hoos S, Felici F, Weintraub A, Mulard LA, Phalipon A, Delepierre M, Bentley GA. Structural mimicry of O-antigen by a peptide revealed in a complex with an antibody raised against *Shigella flexneri* serotype 2a. J Mol Biol 2009; 388:839-850.

Uysal H, Bockermann R, Nandakumar KS, Sehnert B, Bajtner E, Engstrom A, Serre G, Burkhardt H, Thunnissen MMGM, Holmdahl R. Structure and pathogenicity of antibodies specific for citrullinated collagen type II in experimental arthritis. J Exp Med 2009; 206:449-462.

Verhoeyen M, Milstein C, Winter G. Reshaping human antibodies: grafting an antilysozyme activity. Science 1988; 239:1534-1536.

Vulliez-Le Normand B, Saul FA, Phalipon A, Belot F, Guerreiro C, Mulard LA, Bentley GA. Structures of synthetic O-antigen fragments from serotype 2a *Shigella flexneri* in complex with a protective monoclonal antibody. Proc Natl Acad Sci USA 2008; 105:9976-9981.

Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. J Exp Med 1970; 132:211-250.

Zheng Y, Shopes B, Holowka D, Baird B. Conformations of IgE bound to its receptor Fc epsilon RI and in solution. Biochemistry 1991; 30:9125-9132.

Zhou T, Xu L, Dey B, Hessell AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, Burton DR, Dimitrov DS, Sodroski J, Wyatt R, Nabel GJ, Kwong PD. Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature 2007; 445:732-737.

Zhu X, Dickerson TJ, Rogers CJ, Kaufmann GF, Mee JM, McKenzie KM, Janda KD, Wilson IA. Complete reaction cycle of a cocaine catalytic antibody at atomic resolution. Structure 2006; 14:205-216.

Zhu X, Heine A, Monnat F, Houk KN, Janda KD, Wilson IA. Structural basis for antibody catalysis of a cationic cyclization reaction. J Mol Biol 2003; 329:69-83.