Structural and Functional Definitions of Epitopes Reacting with Mouse Monoclonal Antibodies

Rene J. Duquesnoy, Ph.D.
Division of Transplantation Pathology
Thomas E. Starzl Transplantation Institute
University of Pittsburgh Medical Center

ABSTRACT

This report represents a supplement to two recent publications in Human Immunology on the design of the eplet version of HLAMatchmaker [1, 2]. It is addresses stereochemical modeling of crystallized complexes of antibodies and different protein antigens and contributions of amino acid residues to antigen-antibody binding energy.

STRUCTURAL ANALYSIS TOOLS

Studies on complexes of protein antigens and antibody domains (Fab and Fv) have provided detailed stereochemical descriptions of antigen-antibody recognition, interactions and shape complementarity. The Entrez Molecular Modeling Database (MMDB) of the National Center for Biotechnology Information (NCBI) stores on its website (http://www.ncbi.nlm.nih.gov/Structure) an extensive collection of crystallographic structures of antibody-antigen complexes that can be viewed with the Cn3D structure and sequence alignment software program [3]. The atomic coordinates of these molecular complexes are stored as specific PDB codes in the Protein Data Bank. The Cn3D molecular viewer identifies the locations of selected residues and their exposure on the molecular surface. This determines the shapes of epitopes defined by clusters of residues in linear and discontinuous sequences. The Cn3D program has also a “select by distance” (in Ångstroms) command that permits an assessment of the sizes of epitopes and paratopes and the intermolecular distances between them.

Sequence differences between antigenic proteins and corresponding self-proteins of the antibody producer were determined on the website (http://www.ncbi.nlm.nih.gov/BLAST) with the Basic Local Alignment Search Tool (BLAST) [4]. The “space fill” command of the Cn3D molecular viewer was used to identify on antigenic proteins, surface-exposed residues as potential contact sites for antibody.

Determinations of epitope structures were based on experimental findings reported in the literature about binding energy in antigen-antibody complexes selected for this analysis. This information was then applied to develop structural models of functional epitopes with the Cn3D molecular viewer.

STRUCTURAL ASPECTS OF THE ANTIGEN-ANTIBODY INTERFACE

The specific reactivity of antibody is determined by about 50 hypervariable amino acid residues in two variable domains (VH and VL) of heavy and light chains. Both VH and VL domains display high sequence diversity in three complementarity-determining regions (CDRs) separated by relatively conserved intervals termed framework regions [5, 6]. These CDRs form loops that constitute the antigen-binding site: CDR-H3 and CDR-L3 lie generally in the center of the traditional antigen binding site and CDR-H1, CDR-H2, CDR-L1, and CDR-L2 form the outside border [7-13]. CDR-H3 exhibits by far the greatest sequence diversity and plays a dominant role in determining the specificity and affinity of antibody. The other CDRs adopt limited sets of main-chain conformations referred to as canonical structures [8, 14].

Crystallographic models of antigen-antibody complexes have shown that small numbers of residues in the CDR loops make contact with a protein antigen [12, 15, 16]. Conversely, protein antigens have small clusters of amino acid residues that constitute the contact sites for the CDR loops. The resulting interface on average involves about 15-22 amino acid residues of the protein antigen and a similar number of antibody residues. The overall contact area on antigen ranges from 650 to 900 $Å^2$. In comparison, an HLA molecule seen from the top (i.e. binding groove) has a surface area of about 750 $Å^2$.

STRUCTURAL AND FUNCTIONAL DEFINITION OF AN EPITOPE

Considering this information about the structure of the antigen-antibody interface we must address the question what constitutes an epitope. In 1960, Niels Jerne coined the term epitope when he proposed that an antigen particle carries several epitopes [17]. Many epitopes are antigenic determinants expressed on the molecular surface of antigen [18-21]. Others are hidden epitopes (or cryptotopes) that become immunologically available after fragmentation or denaturation of antigen. Processed antigenic peptides presented by major histocompatibility complex molecules to T-cells belong to this group and should be considered cryptotopes rather than T-cell epitopes [22]. Two groups of protein epitopes have been proposed: continuous (or linear) epitopes involving a single continuous amino acid sequence and, discontinuous epitopes that comprise amino acids separated in the primary sequence but clustered together on the molecular surface by folding of the native protein [23, 24]. Mapping studies of antibody reactivity patterns with natural variants and mutated protein antigens have generated information about the location of epitopes and have also suggested that epitopes can generally be defined by small numbers of amino acid residues.

Stereochemical analyses of crystals of antigen-antibody complexes have led to a structural definition of an epitope as that part of the antigen that is contacted by the CDR loops of antibody [15]. This means that with about 15-22 contact residues, a structural epitope comprises a rather large area on the antigen surface and involves many amino acid residues that make contact with a large group of residues on CDRs collectively referred to as the paratope of antibody. Direct contact between epitope

and paratope residues is established through electrostatic forces such as hydrogen bonds, salt bridges, van der Waals forces of hydrophobic surfaces and shape complementarity [18-21]. The interface has also bound water molecules or other co-factors that contribute to the specificity and affinity of antigen-antibody interactions [25].

The binding energy of an antigen-antibody complex is primarily mediated by a small subset of contact residues in the epitope-paratope interface [20, 26, 27]. Substitutions of such "energetic" residues [28] as seen in naturally occurring antigenic variants or induced by site-directed or alanine scanning mutagenesis lead often to dramatic decreases in the binding of antigen to antibody [29-31]. Mapping studies have located energetic residues in "hot spots" of epitopes and paratopes, i.e. regions made up of small numbers of residues that contribute most of the binding energy [32]. Energetic residues are often located in the center of the epitope-paratope interface [33]. Contact residues in periphery of the interface make generally minor contributions to the binding energy; their replacements have frequently little effect on the binding with antigen. The considerable flexibility of CDR3 loops allows a mutual adaptation of both epitopes and paratopes, making it possible for a single antibody molecule to interact with a large number of related antigens [34]. This concept helps to understand the structural basis of serological cross-reactivity.

Thus, the binding or functional activity of an epitope involves a small subset of energetic residues centrally located in the structural epitope and contacted by the specificity-determining CDRs [35]. The assignment of a functional epitope on an antigenic protein should consider two criteria. In order to be immunogenic, a functional epitope must have at least one non-self residue, i.e. the antibody producer's homologous proteins must have a different residue in the corresponding sequence position. Such residue must be on the molecular surface so it can make contact with the specificity-determining CDR

In the epitope-paratope interface, the energetic residues of a functional epitope are often in close contact with the energetic residues of the specificity-determining CDRs. The latter can be identified by site directed mutagenesis of CDR loops [32, 35, 36] and constitute what might be called the functional paratope. Although CDR-H3 plays often a dominant role, other CDRs with energetic residues may provide important contacts with epitope. [12]

The question must be raised how many and which residues might define a functional epitope. Such information would be useful in the design of a model for structurally based HLA compatibility.

IDENTIFICATION OF FUNCTIONAL EPITOPES ON PROTEIN ANTIGENS

There is an extensive literature on the structural analysis of crystallized antigen-antibody complexes. This report addresses only anti-protein antibodies that have well-characterized reactivity patterns with natural antigenic variants and mutated antigens. In addition, sequence information about the antibody

producer's own structurally similar proteins will permit an assessment addressing which residues of the antigenic protein are non-self.

The murine antibody response to hen egg white lysozyme (HEL) represents an excellent model to address the assignment of functional epitopes. A great variety of HEL-specific monoclonal antibodies (mAbs) have been tested with structurally defined, naturally occurring antigenic variants (from different avian species) and HEL variants produced by site-directed mutagenesis. These findings have generated information about serological cross-reactivity patterns between structurally related lysozymes and residues associated with serological determinants. Extensive studies of a few dozen crystallized HEL-antibody complexes have provided detailed information about the epitope-paratope interface and how residue substitutions can affect the interactions between antibody and antigen.

HEL is a 129-residue, antiparallel α-helical protein internally cross-linked by four disulfide bonds [37]. A BLAST analysis of HEL and two mouse lysozymes, type P intestinal [38] and type M milk [39], identifies 49 non-self residues on HEL (they are indicated with the superscript ^{NS}), 30 of them are visible on the molecular surface and would therefore be antibody-accessible (Table 1). This rather large number of exposed non-self residues is consistent with the experimental evidence for multiple epitopes on HEL antigen [40-43].

Table 1 Amino Acid Sequences of Mouse and Avian Lysozymes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
intestinal	K	V	Y	N	R	C	E	L	A	R	I	L	K	R	N	G	M	D	G	Y	R	G	V	K	L	A	D	W	V	C	L	A	Q	H	E	S	N	Y	N	T	R	A	T
milk	K	V	Y	E	R	C	E	F	A	R	T	L	K	R	N	G	M	A	G	Y	Y	G	V	S	L	A	D	W	V	C	L	A	Q	H	E	S	N	Y	N	T	R	A	T
HEL	K	V	F	G	R	C	E	L	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	W	V	C	A	A	K	F	E	S	N	F	N	T	Q	A	T
HEL diff	-	-	F	G	-	-	-	-	-	A	A	M	-	-	H	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	-	Q	-	-
Cal Quail	-	-	F	G	-	-	-	-	-	A	A	M	-	-	H	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	S	Q	-	-
BW Quail	-	-	F	G	-	-	-	-	-	A	A	M	-	-	H	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	S	Q	-	-
Jap Quail	-	-	-	G	-	-	-	-	-	A	A	M	-	-	H	-	L	-	K	-	Q	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	-	Q	-	-
Turkey	-	-	-	G	-	-	-	-	-	A	A	M	-	-	L	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	-	H	-	-
RN Pheasant	G	-	-	G	-	-	-	-	-	A	A	M	-	-	M	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	-	G	-	-
Chachalaca	-	I	-	K	-	-	-	-	-	A	A	M	-	-	Y	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	R	Y	F	-	-	-	-	-	-	Q	-	-
Duck A	-	-	-	-	-	-	-	-	-	A	A	M	-	-	L	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	N	Y	F	-	-	S	F	-	S	Q	-	-
Guinea Fowl	-	-	F	G	-	-	-	-	-	A	A	M	-	-	H	-	L	-	N	-	-	-	Y	S	-	G	N	-	-	-	A	-	K	F	-	-	-	F	-	-	Q	-	-

	44	45	46	47		48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
intestinal	N	Y	N	R	G	D	R	S	T	D	Y	G	I	F	Q	I	N	S	R	Y	W	C	N	D	G	K	T	P	R	S	K	N	A	C	G	I	N	C	S	A	L	L	Q
milk	N	Y	N	R	G	D	Q	S	T	D	Y	G	I	F	Q	I	N	S	R	Y	W	C	N	D	G	K	T	P	R	A	V	N	A	C	G	I	N	C	S	A	L	L	Q
HEL	N	R	N	T	*	D	G	S	T	D	Y	G	I	L	Q	I	N	S	R	W	W	C	N	D	G	R	T	P	G	S	R	N	L	C	N	I	P	C	S	A	L	L	S
HEL diff	-	R	-	T	*	-	G	-	-	-	-	-	-	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	R	-	L	-	N	-	P	-	-	-	-	-	S
Cal Quail	-	R	-	T	*	-	G	-	-	-	-	-	V	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	R	-	L	-	N	-	P	-	-	-	-	-	S
BW Quail	-	R	-	T	*	-	G	-	-	-	-	-	V	L	-	-	-	-	-	W	-	-	-	-	-	-	-	-	G	-	R	-	L	-	N	-	P	-	-	-	-	-	S
Jap Quail	-	R	-	T	*	-	G	-	-	-	-	-	-	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	R	-	L	-	N	-	P	-	-	-	-	-	S
Turkey	-	R	-	T	*	-	G	-	-	-	-	-	-	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	-	-	L	-	N	-	P	-	-	-	-	-	S
RN Pheasant	-	R	-	T	*	-	G	-	-	-	-	-	-	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	-	-	L	-	H	-	P	-	-	-	-	-	S
Chachalaca	-	R	S	N	*	-	G	-	-	-	-	-	-	E	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	T	-	-	L	-	H	-	S	-	-	-	-	M	G
Duck A	-	R	-	T	*	-	G	-	-	-	-	-	V	L	-	-	-	-	-	W	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	-	P	-	-	V	-	-	R
Guinea Fowl	-	R	-	T	*	-	G	-	-	-	-	-	-	L	-	-	-	-	-	W	-	-	-	-	-	R	-	-	G	-	R	-	L	-	N	-	P	-	-	-	-	Q	S

	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129
intestinal	D	D	I	T	A	A	I	Q	C	A	K	R	V	V	R	D	P	Q	G	I	R	A	W	V	A	W	R	T	Q	C	Q	N	R	D	L	S	Q	Y	I	R	N	C	R	L
milk	D	D	I	T	A	A	I	Q	C	A	K	R	V	V	R	D	P	Q	G	I	R	A	W	V	A	W	R	A	H	C	Q	N	R	D	L	S	Q	Y	I	R	N	C	R	L
HEL	S	D	I	T	A	S	V	N	C	A	K	K	I	V	S	D	G	N	G	M	N	A	W	V	A	W	R	N	R	C	K	G	T	D	V	Q	A	W	I	R	G	C	R	L
HEL diff	S	-	-	-	-	S	V	N	-	-	-	K	I	-	S	-	G	N	-	M	N	-	-	-	-	-	-	N	R	-	K	G	T	-	V	Q	A	W	-	-	G	-	-	-
Cal Quail	S	-	-	-	-	T	V	N	-	-	-	K	I	-	S	-	G	N	-	M	N	-	-	-	-	-	-	N	R	-	K	G	T	-	V	H	A	W	-	-	G	-	-	-
BW Quail	S	-	-	-	-	T	V	N	-	-	-	K	I	-	S	-	G	N	-	M	N	-	-	-	-	-	-	N	R	-	K	G	T	-	V	Q	A	W	-	-	G	-	-	-
Jap Quail	S	-	-	-	-	S	V	N	-	-	-	K	I	-	S	-	V	H	-	M	N	-	-	-	-	-	-	N	R	-	K	G	T	-	V	N	A	W	-	-	G	-	-	-
Turkey	S	-	-	-	-	S	V	N	-	-	-	K	I	A	S	G	G	N	-	M	N	-	-	-	-	-	-	N	R	-	K	G	T	-	V	H	A	W	-	-	G	-	-	-
RN Pheasant	S	-	-	-	-	S	V	N	-	-	-	K	I	-	S	-	G	N	-	M	N	-	-	-	-	-	K	H	R	-	K	G	T	-	V	N	V	W	-	-	G	-	-	-
Chachalaca	A	-	A	P	-	S	V	R	-	-	-	-	I	-	S	-	G	N	-	M	N	-	-	-	-	-	K	H	R	-	K	G	T	-	V	-	T	W	-	K	D	-	K	-
Duck A	S	-	-	E	-	-	V	R	-	-	-	-	I	-	S	-	G	N	-	M	N	-	-	-	-	-	-	N	R	-	R	G	T	-	V	-	K	W	-	-	G	-	-	-
Guinea Fowl	S	-	-	-	T	T	A	N	-	-	-	K	I	-	S	-	G	N	-	M	N	-	-	-	-	-	K	H	R	-	K	G	T	-	V	R	V	W	-	K	G	-	-	-

Reactivity patterns with various avian lysozymes and mutated HEL antigens have shown that anti-HEL mAbs react with serological determinants in one of four distinct antigenic regions that cover most of the lysozyme surface [40, 43]. Detailed structural information has become available on antigen-

antibody complexes with mAbs specific for epitopes representative of three regions. This analysis addresses the reactivity of five HEL-specific mAbs (Table 2).

HEL-D1.3 The reactivity patterns of this mAb with natural lysozyme variants and mutated HEL have indicated that the D1.3 reacts with an epitope that includes the 117-121 sequence in one of the loops of HEL [43]. In the crystalline antigen-antibody complex, the CDRs of D1.3 make contact with 16 residues in the 18-27 and 117-125 sequence segments of HEL [44, 45]. The paratope of D1.3 has 9 VH and 8 VL contact residues. Alanine scanning has indicated that HEL residues 121Q ^{NS}, 125R, 124I and 119A dominate the energetics of binding with D1.3 [33, 46, 47]. Turkey egg lysozyme differs from HEL by one amino acid in the interface (121H versus 121Q^{NS}) and has a 400-fold lower affinity with D1.3 than HEL [42, 48]. Other lysozymes that differ at position 121 (Glutamine to predominantly histidine) are non-reactive suggesting that the non-self 121Q^{NS} residue plays a critical role in the binding with D1.3.

Thus, the functional D1.3 epitope can be defined by the sequence 119A, 121Q^{NS}, 124I and 125R (Table 2). The functional paratope of D1.3 has five energetic residues; three of them, H100D, H101Y and H102R are on CDR-H3.

Table 2 Residue descriptions of interfaces of eight crystallized protein antigen-antibody complexes

PDB	Antigen	Monoclonal Antibody	Number of Residues Structural Epitope	Number of Contact Residues in H and L Chains	Functional Paratope Residues [1]	Residues in Functional Epitope [2]
1VFB	Hen Egg Lysozyme	D1.3	16	H: 9 L: 8	H: 100D, 101Y, 102R L: 32Y, 92W:	119A, 121Q^{NS}, 124I, 125R
3HFL	Hen Egg Lysozyme	HyHEL-5	13	H: 12 L: 6	H: 35E, 50E L: 90W, 91G, 92R	45R^{NS}, 68R^{NS}
3HFM	Hen Egg Lysozyme	HyHEL-10	20	H: 13 L: 9	H: 32D, 50Y, 98W L: 49K, 53Y	20Y, 21R, 93N^{NS}, 97K^{NS}
1FBI	Hen Egg Lysozyme	F9.13.7	15	H:14 L:4	H: 52D, 54S, 100S L: 32Y, 92Y	20Y, 73R^{NS}, 96K^{NS}
1DQJ	Hen Egg Lysozyme	HyHEL-63	21	H: 10 L: 9	H: 33Y, 98W L: 32N, 50Y	20Y, 96K^{NS}, 97K^{NS}
1WEJ	Horse Cytochrome C	E8	10	H: 5 L: 6	H: 33Y, 50R L: 31N, 32Y, 50N	60K^{NS}, 62E^{NS}, 103N, 104E
1JRH	Interferonγ Receptor	A6	14	H: 13 L: 9	H: 99Y, 102H L: 30Y, 92W, 96W	49Y, 50G^{NS}, 51V^{NS}, 52K^{NS}, 53N^{NS}

1 The numbering of CDR residues is according to the Kabat system [123]. H1: 31-35, H2: 50-65, H3: 95-102, L1: 24-34, L2: 50-56.
2 Non-self residues are marked as ^{NS}.

HyHEL-5 The reactivity pattern with different avian lysozymes suggests that this mAb is specific for a serological determinant associated with residue 68R^{NS} and residues 45 to 48 [40, 49]. In the crystalline complex, the structural epitope recognized by HyHEL-5 comprises 13 contact residues in three sequences 41-53, 67-70 and 79-84 [50]. The surface of this epitope is relatively flat except for a protruding ridge made up of 45R^{NS} and 68R^{NS}; these residues interact with H50E of CDR-H2 and H35E of CDR-H1. Studies on mutated HEL and natural variants have shown that 45R^{NS} and 68R^{NS}

make the highest energetic contacts with HyHEL-5 [36, 51, 52]. These findings suggest that the functional epitope recognized by HyHEL-5 involves 45R^{NS} and 68R^{NS} (Table 1). None of the five energetic contact residues are on CDR-H3 [36, 51, 52] suggesting that this CDR plays a minor role in the specific reactivity of HyHEL-5.

HyHEL-10 Studies with avian variants of lysozyme have suggested that HEL residues 19N, 21R, 101D, 102G and/or 103Q contribute to the binding with this mAb [40]. Two crystal structures (PDB 3HFM and PDB 1C08) have shown that these residues are included in the 20-residue structural epitope that contacts the 22 contact residues of HyHEL-10 [53, 54]. Site-directed mutations of 20Y, 21R, 93N^{NS} and 97K^{NS} caused markedly reduced affinity with HyHEL-10 [55-58]. These residues appear to define the functional epitope recognized by HyHEL-10 (Table 1). The most energetic 97K^{NS} interacts with H32D [59].

HEL-F9.13.7 Although HEL-F9.13.7 reacts with a similar structural epitope as HyHEL-10, this mAb has little sequence similarity at any of their CDRs and its binding energetics is also different [60]. HEL-F9.13.7 has 18 CDR contact residues that interact with 15 residues of the structural HEL epitope. The CDR-H3 of both mAb reacts with the HEL hot spots 20Y and 96K^{NS} but F9.13.7 reacts also with a third HEL hot spot 73R^{NS} through its CDR-H1. Alanine scanning mutagenesis identified three highly energetic residues that bind to HEL-F9.13.7, namely 20Y, 73R^{NS} and 96K^{NS} [60]. The crystal complex shows that 20Y and 96K^{NS} interact with H52D, H53S and H100S whereas 73R^{NS} contacts L32Y and L92Y [57]. These findings suggest that the functional HEL-F9.13.7 epitope can be defined by 20Y, 73R^{NS} and 96K^{NS}.

HyHEL-63 This mAb reacts with an epitope that also overlaps with the HyHEL-10 epitope [61]. Alanine substitutions of HEL residues 96K^{NS}, 97K^{NS} and 20Y caused a considerable reduction of binding energy [56]. These residues are located at the center of the interface. In contrast, alanine substitutions of HEL residues 21R and 101D had much less effect on the binding energy of the complex [62]. Thus, the functional epitope recognized by HyHEL-63 includes 20Y, 96K^{NS} and 97K^{NS}.

The following examples involve different protein antigens:

Horse Cytochrome C A mouse mAb E8 was raised against horse cytochrome C (hCytC) [63]. Sequence comparison with mouse cytochrome C shows that this 105-residue antigenic protein has six non-self residues (Table 3); all of them are exposed on the molecular surface and can elicit specific antibodies [18] [64]. Crystallographic analysis of the hCytC-FabE8 complex has shown that the structural E8 epitope constitutes a patch of ten residues located in three sequence segments of hCytC, namely 36F+37G, 60K^{NS}+61E+62E^{NS} and 96A+99K+100K+103D [65]. The two non-self residues 60K^{NS} and 62E^{NS} and are located in the center of the interface. Rat cytochrome C has an identical sequence as mouse cytochrome C and does not bind E8 [66]. Highly homologous cytochrome C from cow, rabbit, guanaco and dog, which have different residues in positions 60, 62 and/or 103, react poorly with E8 [66, 67]. A fourth and evolutionarily conserved residue 104E makes

also a significantl contribution to the binding with E8 [65]. Thus, the functional E8 epitope consists of $60K^{NS}$ and $62E^{NS}$ and two self-residues 103K and 104E. The E8 paratope has six VL and five VH residues in the CDR loops, five of which make contact with the functional epitope.[65].

Table 3 Amino Acid Sequences of Horse and Mouse Cytochrome C

```
                  1          2          3          4          5
                  0          0          0          0          0
Horse    GDVEKGKKI FVQKCAQCHT VEKGGKHKTG PNLHGLFGRK TGQAPGFTYT
Mouse    GDVEKGKKI FVQKCAQCHT VEKGGKHKTG PNLHGLFGRK TGQAAGFSYT
                                                            1
                   6          7          8          9          0
                   0          0          0          0          0
Horse    DANKNKGITW KEETLMEYLE NPKKYIPGTK MIFAGIKKKT EREDLIAYLK KATNE
Mouse    DANKNKGITW GEDTLMEYLE NPKKYIPGTK MIFAGIKKKG ERADLIAYLK KATNE
```

Human Interferon-γ Receptor The murine mAb A6 reacts with the first domain of human IFNγR [68]. BLAST analysis of this domain has shown that human and mouse IFNγR [69] have 42 of 92 residue differences, 35 of them involve positions on the molecular surface. Simultaneous mutations of five residues (48, 53, 54, 55, and 57) from non-self to self and a deletion in positions 50-53 of three residues absent in mouse IFNγR, led to a drastically reduced binding with A6 [70]. In the crystallized complex of FabA6 with the first domain of IFNγR shows 14 residues in the structural epitope that make contact with 13 VH and 9 VL residues [70]. CDR-H3 and CDR-L3 have tight interactions with IFNγR residues 49Y, $50G^{NS}$, $51V^{NS}$, $52K^{NS}$ and $53N^{NS}$ in the center of the interface. The functional epitope for A6 appears to comprise these five residues; all of them except 49Y, are non-self.

Table 4 Amino Acid Sequences of Human and Mouse First Interferon Gamma Receptor Domain

```
             1          2          3          4          5          6
             0          0          0          0          0          0
Human  TADLGPS SVPTPTNVTI ESYNMNPIVY WEYQIMPQVP VFTVEVKNYG VKNSEWIDAC
Mouse  TEDPEPP SVPVPTNVLI KSYNLNPVVC WEYQNMSQTP IFTVQVKVY- --SGSWTDSC

                                            1
                 7          8          9          0
                 0          0          0          0
Human   INISHHYCNI SDHVGDPSNS LWVRVKARVG QKESAYAKSE EFAVSRDG
Mouse   TNISDHCCNI YGQIMYPDVS AWARVKAKVG QKESDYARSK EFLMCLKG
```

TOPOGRAPHY OF FUNCTIONAL EPITOPES ON PROTEIN ANTIGENS

In these seven crystal models, the structural epitopes consist of 10 to 21 amino acid residues in contiguous patches on the protein surface (Table 1). Small subsets ranging from 2 to 5 highly energetic contact residues in mostly central locations are considered to constitute the functional epitopes. These residues are largely in discontinuous sequences and at least one of them is non-self. This analysis has also shown that about 4-5 highly energetic antibody residues might define the functional paratopes. As expected, the CDR-H3 loops play frequently but not always a major role and other CDRs are involved in the binding with functional epitopes.

The next step of this analysis was an assessment of the shapes and sizes of functional epitopes and the locations of contact residues in the interface. Paratopes comprise energetic residues listed in Table 1 and CDR residues within 3.5 Å of the functional epitope. Figure 1 illustrates the Cn3D modeling results for functional epitopes (left column), the corresponding paratope (middle column) and the residues in the interface (right column). Functional epitope residues seem to cluster in two distinct shapes, namely as a single patch or a two-patch shape.

In four cases, the functional epitope has the shape of a single patch (Figure 1).

HEL-D1.3 The four residues of the functional HEL-D1.3 epitope form a single patch that is well exposed on the HEL surface (Figure 1a). The centrally located $121Q^{NS}$ protrudes from the HEL surface. The other residues 119D, 124I and 125R are within 3.0 Å of $121Q^{NS}$. The functional paratope has three residues H101Y, L92W and L32Y that are less than 3.5 Å away from the functional HEL-D1.3 epitope.

HyHEL-5 The two residues $45R^{NS}$ and $68R^{NS}$ of this functional epitope form a single patch (Figure 1b). Although their sequence positions seem quite distant, these residues are only 3.5 Å apart from each other. In the functional paratope, CDR-L3 makes the closest contact (<3.5 Å) with both $45R^{NS}$ and $68R^{NS}$ and the VH residue H50E interacts primarily with $68R^{NS}$.

HyHEL-63 This functional epitope is defined by three residues 20Y, $96K^{NS}$ and $97K^{NS}$ that form a single patch (Figure 1c). Although 20Y and the other two residues seem far removed in sequence, they are only 3.5 Å apart. Three residues on the functional paratope, L32N, L50Y and H33Y are less than 3.5 Å away from the functional epitope. The CDR-H3 residue H98W is 4.0 Å from 97K.

IFNγR-A6 The five residues of this functional epitope cluster as a patch in which $52K^{NS}$ has a central position and the other residues are within 3.4 Å (Figure 1d). In the functional paratope, CDR-H3 interacts with $49Y^{NS}$, $50G^{NS}$, $51V^{NS}$ and $52K^{NS}$ whereas CDR-L3 binds to 49Y and $54N^{NS}$.

These four patches are centrally located within the structural epitope (not shown) and their residues are 3-3.5 A apart from each other.

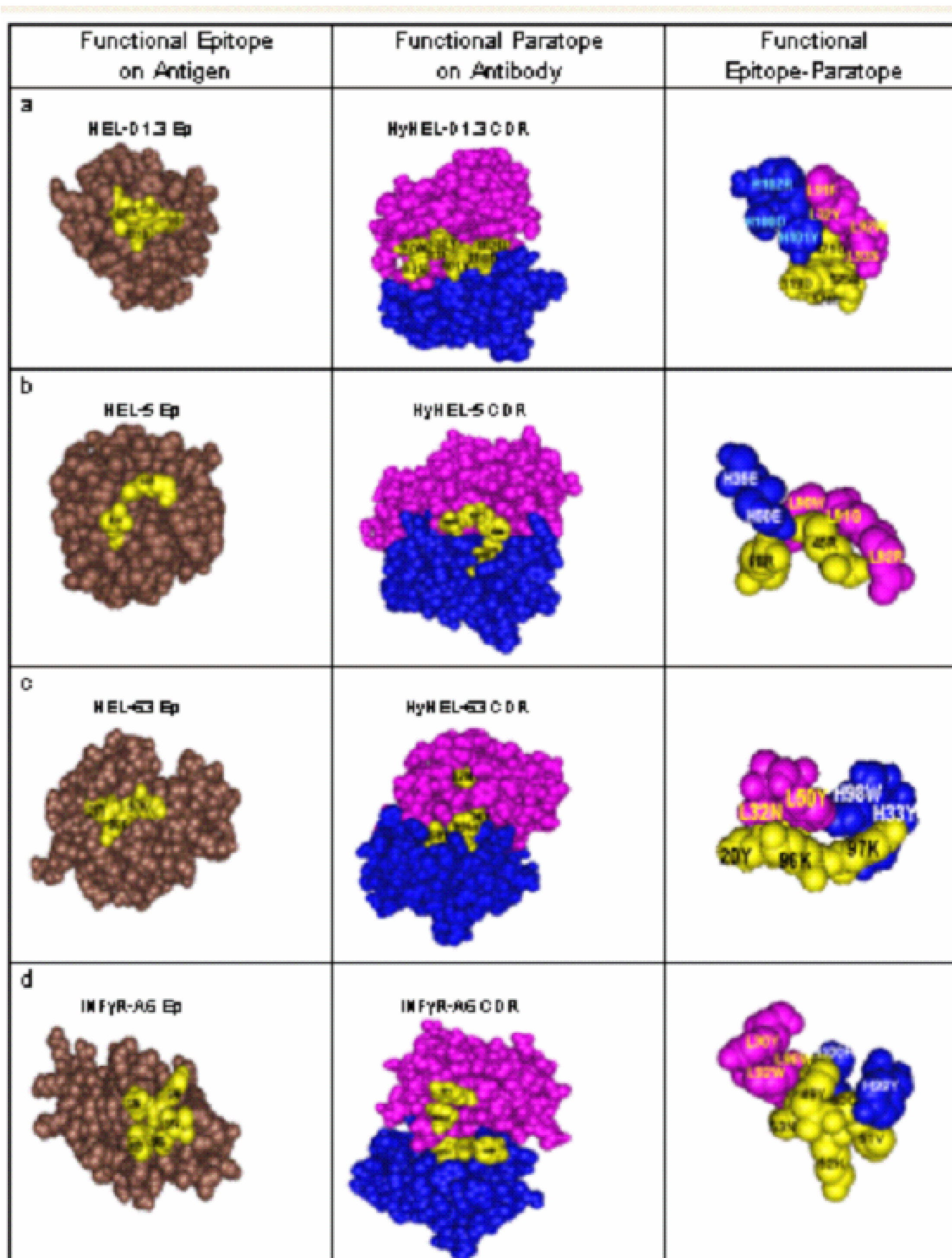

Figure 1 Locations of energetic residues on antigen surface (functional epitope), antibody surface (functional paratope) and in the epitope-paratope interface

The three remaining functional epitopes in Table 2 have a two-patch shape (Figure 2).

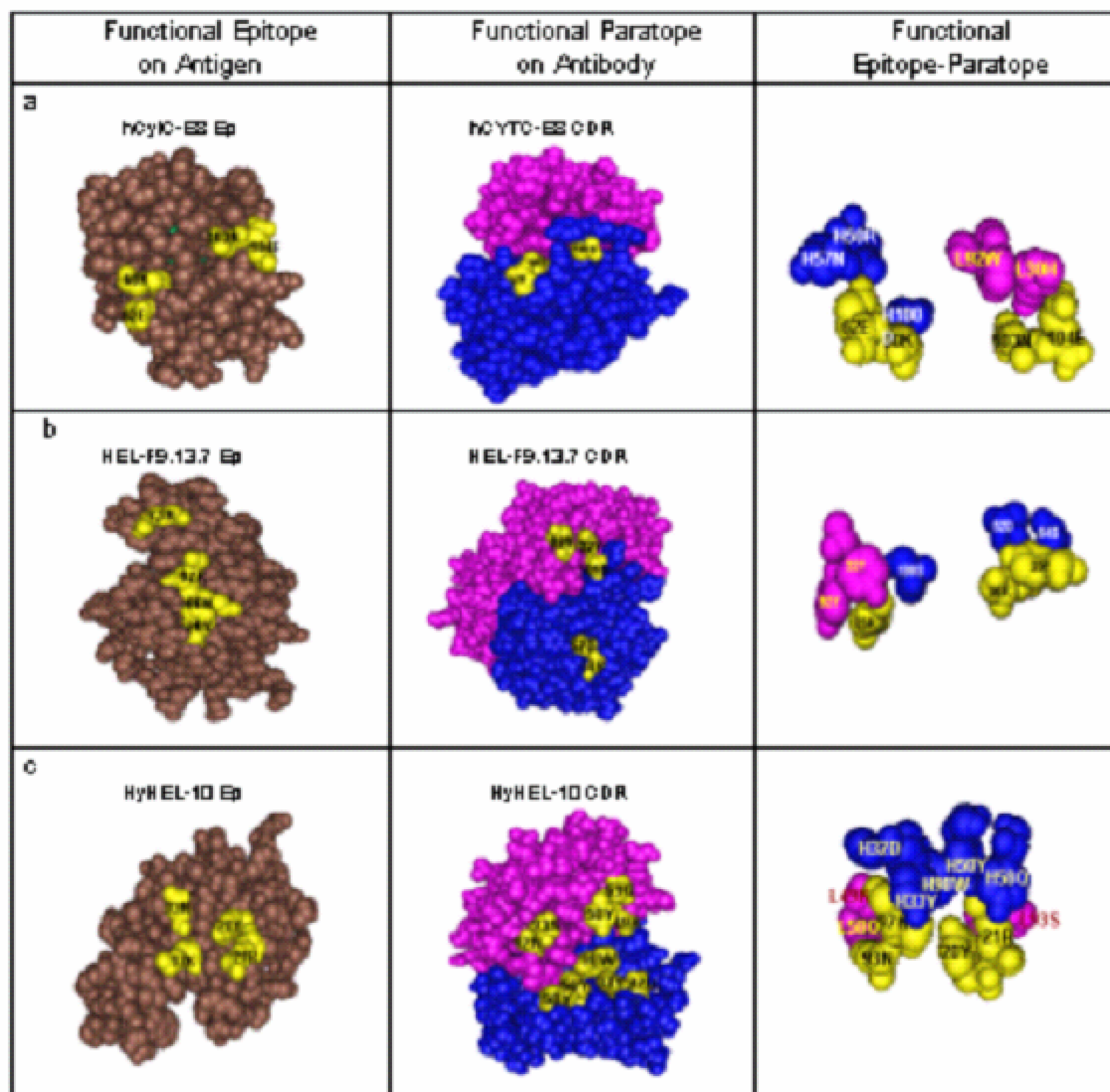

Figure 2 Two-patch shapes energetic residues on antigen surface (functional epitopes), antibody surface (functional paratopes) and in the epitope-paratope interface

<u>hCytC-E8</u> This functional epitope has two distinct patches (Figure 2a). One patch comprises $60K^{NS}$ and $62E^{NS}$ which are 3.3 Å apart and which make contact with VH residues H50N, H57N and H100D. The second pair consists of two self-residues 103N and 104E that interact with VL residues L30H and L92W. These two patches are well separated as shown by the 11.5 Å distance between $60K^{NS}$ and 103N. The closest distance between the corresponding VH and VL contact residues is about 9.5 Å.

<u>HEL-F9.13.7</u> This functional epitope has two well-separated patches, each of which has a non-self residue (Figure 2b). Residues 20Y and $96K^{NS}$ are 3.1 Å apart and form one patch that makes contact with H52D and H54S. The second patch has the highly energetic $73R^{NS}$ that protrudes from the molecular surface. This residue makes contact with H100S, L92Y and L32Y. The distance between

$73K^{NS}$ and $96K^{NS}$ is about 13 Å, whereas the CDR-H2 residues H52D and H54S are about 15 Å away from H100S in CDR-H3.

HyHEL-10 The four residues of this functional epitope separate into two patches: $93N^{NS}$+$97K^{NS}$ and 20Y+21R (Figure 2c). These patches are about 7.5 Å apart. The first is a 3.3 Å patch comprised of discontinuous sequence residues $93N^{NS}$ and $97K^{NS}$. CDR-H1 and CDR-H3 makes contact with $97K^{NS}$ and CDR-L2 residues make contact with $93N^{NS}$. The 20Y+21R patch has only self-residues and their contact involves primarily the VL CDR loops.

This analysis shows patches of highly energetic residues that are 3-3.5 Å apart from each other. There are three configurations: 1) most common are single patches with one or more non-self residues (Figure 1). Other study models have also shown an abundance of single patches [71, 72]. 2) Two patches and both of them have at least one non-self residue (Figure 2a). 3) A combination of a patch with one (or more) non-self residues and a patch consisting of only self residues (Figures 2b and 2c). Although a single patch consists of a short linear or discontinuous sequence of residues clustered closely together, it can make contact with several different CDR residues (Table 1 and Figure 1). These CDR residues might constitute the "functional" paratope that plays a determining role in the specific interaction with the functional epitope represented by a patch.

DISCUSSION

The above models of antigen-antibody complexes illustrate how structural epitopes contain patches of surface residues that play a dominant role in determining recognition by specific antibody. These residues cluster about 3-3.5 Å from each other. Non-self residues in these patches could be considered the driving force of functional epitopes in terms of immunogenicity (i.e. the ability to elicit antibody formation) and antigenicity (i.e. the ability to react with antibodies). The redesign of HLAMatchmaker incorporates these concepts. Instead of using triplets (i.e. linear sequences of three amino acid residues), the new algorithm considers patches of residues in linear and discontinuous sequences. Details are presented elsewhere [1].

Acknowledgement: This study is supported by grant AI-55933 from the National Institutes of Health

REFERENCES

1. Duquesnoy, R., *A Structurally Based Approach to Determine HLA Compatibility at the Humoral Immune Level.* Human Immunol.. 2006. 67: p. 847-862.
2. Duquesnoy, R.J. and M. Askar, *HLAMatchmaker: A Molecularly Based Algorithm for Histocompatibility Determination V. Eplet Matching for HLA-DR, HLA-DQ and HLA-DP.* Human Immunol.. 2007. 68: p. 12-25.
3. Hogue, C., *Cn3D: a new generation of three-dimensional molecular structure viewer.* Trends Biochem Sci. , 1997. 22: p. 314-316.
4. Altschul, S., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res. , 1997. 25: p. 3389-3402.

5. Kabat, E.A., T. Wu, and H. Bilofsky, *Unusual distribution of amino acids in complementarity-determining (hypervarible) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites.* J. Biol. Chem., 1977. 252: p. 6609-6617.
6. Kabat, E.A., *The structural basis of antibody complementarity.* Adv. Protein Chem., 1978. 32: p. 1-75.
7. Tonegawa, S., *Somatic generation of antibody diversity.* Nature, 1983. 302: p. 575-581.
8. Chothia, C., et al., *Conformation of immunoglobulin hypervariable regions.* Nature 1989. 342: p. 877–883.
9. Wilson, I.A. and R.L. Stanfield, *Antibody–antigen interactions: new structures and new conformational changes.* Curr. Opin. Struct. Biol., 1994. 4: p. 857-867.
10. Padlan, E.A., *Anatomy of the antibody molecule.* Mol. Immunol., 1994. 31: p. 169-217.
11. Kirkham, P.M. and H.W. Schroeder Jr, *Antibody structure and the evolution of immunoglobulin V gene segments.* Seminars in Immunology, 1994. 6: p. 347-360.
12. MacCallum, R.M., A.C.R. Martin, and J.M. Thornton, *Antibody-Antigen Interactions: Contact Analysis and Binding Site Topography.* J. Mol. Biol., 1996. 262: p. 732-745.
13. Zemlin, M., et al., *Expressed Murine and Human CDR-H3 Intervals of Equal Length Exhibit Distinct Repertoires that Differ in their Amino Acid Composition and Predicted Range of Structures.* J. Mol. Biol., 2003. 334: p. 733–749.
14. Al-Lazikani, B., A. Lesk, and C. Chothia, *Standard conformations for the canonical structures of immunoglobulins.* J Mol Biol, 1997. 273 p. 927–948.
15. Davies, D., E. Padlan, and S. Sheriff, *Antibidy-antigen complexes.* Annu. Rev. Biochem., 1990. 59: p. 439-473.
16. Almagro, J., *Identification of differences in the specificity-determining residues of antibodies that recognize antigens of different size: implications for the rational design of antibody repertoires.* J. Mol. Recognit., 2004. 17: p. 132-143.
17. Jerne, N., Ann. Rev. Microbiol., 1960. 14: p. 341-358.
18. Benjamin, D.C., et al., *The antigenic structure of proteins: A reappraisal.* Ann. Rev. Immunol., 1984. 2: p. 67-101.
19. Colman, P.M., *Structure of antibody-antigen complexes: implications for immune recognition.* Adv. Immunol., 1988. 43: p. 99-132.
20. Getzoff, E.D., et al., *The chemistry and mechanisnm of antibody binding to protein antigens.* Adv. Immunol., 1988. 43: p. 1-98.
21. Van Regenmortel, M.H.V., *Structural and functional approaches to the study of protein antigenicity.* Immunol. Today 1989. 10: p. 266-272.
22. Laver, W.G., et al., *Epitopes on protein antigens: misconception and realities.* Cell, 1990. 61: p. 553-559.
23. Sela, M., et al., *Antibodies to sequential and conformational determinants.* Cold Spring Harbor Symp. Quant. Biol., 1967. 32: p. 537-545.
24. Atassi, M. and J. Smith, *A proposal for the nomenclature of antigenic sites in peptides and proteins.* Immunochemistry, 1978. 15: p. 609-610.
25. Braden, B., et al., *Anatomy of an antibody molecule: structure, kinetics, thermodynamics and mutational studies of the antilysozyme antibody D.* Immunol. Rev., 1998. 163: p. 45-57.
26. Laune, D., et al., *Systematic Exploration of the Antigen Binding Activity of Synthetic Peptides Isolated from the Variable Regions of Immunoglobulins.* J. Biol. Chem., 1997. 272: p. 30937-30944.
27. Van Regenmortel, M., *Reductionism and the search for structure-function relationships in antibody molecules.* J. Mol. Recognit. , 2002. 15: p. 240-247.
28. Novotny, J., *Protein antigenicity: a thermodynamic approach.* Mol. Immunol., 1991. 28: p. 201-208.
29. Cunningham, B. and J. Wells, *Comparison of a structural and a functional epitope.* J. Mol. Biol., 1993. 234: p. 554-563.
30. Wells, J., *Binding in the growth hormone receptor complex.* Proc. Nat. Acad. Sci., U.S.A., 1996. 93: p. 1-6.
31. Davies, D. and G. Cohen, *Interactions of protein antigens with antibodies.* Proc. Natl. Acad. Sci. USA, 1996. 93: p. 7-12.

32. DeLano, W., *Unraveling hot spots in binding interfaces: progress and challenges.* Curr. Opin. Struct. Biol., 2002. 12: p. 14-20.
33. Hawkins, R., et al., *The contribution of contact and non-contact residues of antibody in the affinity of binding to antigen. The interaction of mutant D1.3 antibodies with lysozyme.* J. Mol. Biol., 1993. 234: p. 958-964.
34. Roberts, V., E. Getzoff, and J. Tainer, *Structural basis of antigenic cross-reactivity*, in *Structure of Antignes*, M. Van Regenmortel, Editor. 1993, CRC press: Boca Raon, FL. p. 31-53.
35. Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces.* J. Mol. Biol. , 1998. 280: p. 1–9.
36. Lavoie, T.B., W.N. Drohan, and S.J. Smith-Gill, *Experimental analysis by site-directed mutagenesis of somatic mutation effects on affinity and fine specificity in antibodies specific for lysozyme.* J. Immunol. , 1992. 148: p. 503–513.
37. Tainer, J.T., et al., *The Atomic Mobility Component of Protein Antigenicity* Ann. Rev. Immunol., 1985. 3: p. 501-35.
38. Cortopassi, G.A. and A.C. Wilson, *Recent origin of the P lysozyme gene in mice.* Nucleic Acids Res., 1990. 18: p. 1911-1920.
39. Obita, T., T. Ueda, and T. Imoto, *Solution structure and activity of mouse lysozyme M.* Cell. Mol. Life Sci., 2003. 60: p. 176-184.
40. Smith-Gill, S.J., T.B. Lavoie, and C. Mainhart, *Antigenic regions defined by monoclonal antibodies correspond to structural domains of avian lysozymes.* J. Immunol. , 1984. 133: p. 384-393.
41. Metzger, D.W., et al., *The expressed lysozyme-specific B cell repertoire: 1. Heterogeneity in the monoclonal anti-hen egg white lysozyme specificity repertoire. and its difference from the in situ repertoire.* Eur. J. Immunol. , 1984. 14: p. 87-96.
42. Harper, M., et al., *Antigen specificity and cross-reactivity of monoclonal anti-lysozyme antibodies.* Mol. Immunol. , 1987. 2: p. 97-108.
43. Newman, M., et al., *Patterns of antibody specificity during the balb/c immune response to hen eggwhite lysozyme'.* J . Immunol., 1992. 149: p. 3260-3272.
44. Amit, A., et al., *Three-dimensional structure of an antigen-antibody complex at 2.8 Angstrom resolution.* Science, 1986. 233: p. 747-752.
45. Fischmann, T.O., et al., *Crystallographic refinement of the three-dimensional structure of the FabD1.3-lysozyme complex at 2.5-A° resolution.* J. Biol. Chem. , 1991. 266: p. 12915–12920.
46. Dall'Acqua, W., et al., *A Mutational Analysis of Binding Interactions in an Antigen-Antibody Protein-Protein Complex.* Biochemistry 1998. 37: p. 7981-7991.
47. Bhat, T.N., et al., *Bound water molecules and conformational stabilization help mediate an antigen-antibody association. .* Proc. Natl Acad. Sci. USA,. 1994. 91: p. 1089–1093.
48. Fields, B., A, et al., *Hydrogen Bonding and Solvent Structure in an Antigen-Antibody Interface. Crystal Structures and Thermodynamic Characterization of Three Fv Mutants Complexed with Lysozyme.* Biochemistry, 1996. 35: p. 15494-15503.
49. Smith-Gill, S., et al., *Mapping the antigenic epitope for a monoclonal antibody against lysozyme.* J. immunol., 1982. 128: p. 314-320.
50. Sheriff, S., et al., *Three-dimensional structure of an antibody-antigen complex* Proc. Natl Acad. Sci. USA, 1987. 84: p. 8075-8079.
51. Shick, K., et al., *Association of the anti-hen egg lysozyme antibody HyHEL-5 with avian species variant and mutant lysozymes.* Biocm. Biophys. Acta, 1997. 1340: p. 205-214.
52. Chacko, S., et al., *Structure of an Antibody-Lysozyme Complex Unexpected Effect of a Conservative Mutation.* J.Mol. Biol., 1995. 245: p. 261-274.
53. Padlan, E.A., et al., *Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab-lysozyme complex. Proc. .* Nat. Acad. Sci., U.S.A. , 1989. 86: p. 5938–5942.
54. Kondo, H., et al., *Crystal Structure of Anti-Hen Egg White Lysozyme Antibody (HyHEL-10) Fv-Antigen Complex.* J. Biol. Chem., 1999. 274: p. 27623-2731.
55. Kam-Morgan, L.N.W., et al., *High resolution mapping of the HyHEL-10 epitope of chicken lysozyme by site-directed mutagenesis. .* Proc. Nat. Acad. Sci., U.S.A. , 1993. 90: p. 3958–3962.
56. Li, Y.C., et al., *Mutations of an epitope hot-spot residue alter rate limiting steps of antigen-antibody associations.* Biochemistry, 2001. 40: p. 2011-2022.

57. Pons, J., J. Stratton, and J. Kirsch, *How do two unrelated antibodies, HyHEL-10 and F9.13.7, recognize the same epitope of hen egg-white lysozyme?* Protein Science., 2002. 11: p. 2308–2315.
58. Sinha, N., et al., *Differences in Electrostatic Properties at Antibody–Antigen Binding Sites: Implications for Specificity and Cross-Reactivity.* Biophys. J., 2002. 83: p. 2946-2968.
59. Pons, J., A. Rajpal, and J.F. Kirsch, *Energetic analysis of an antigen/antibody interface: alanine scanning mutagenesis and double mutant cycles on the HyHEL-10/lysozyme interaction.* Protein Sci., 1999. 8: p. 958-968.
60. Lescar, J., et al., *Crystal structure of a cross-reaction complex between Fab F9.13.7 and guinea fowl lysozyme. .* J. Biol. Chem. 1995. 270 p. 18067–18076.
61. Li, Y., et al., *Three-Dimensional Structures of the Free and Antigen-Bound Fab from Monoclonal Antilysozyme Antibody HyHEL-63.* Biochemistry, 2000. 39: p. 6296-6309.
62. Li, Y., et al., *Dissection of binding interactions in the complex between the anti-lysozyme antibody HyHEL-63 and its antigen. .* Biochemistry 2003. 42: p. 11-22.
63. Carbone, F.R. and Y. Paterson, *Monoclonal antibodies to horse cytochrome c expressing four distinct idiotypes distribute among two sites on the native protein. .* J. Immunol. , 1985. 135: p. 2609-2616.
64. Bushnell, G.W., G.V. Louie, and G.D. Brayer, *High-resolution three-dimensional structure of horse heart Cytochrome c.* J. Mol. Biol. , 1990. 214: p. 585-595.
65. Mylvaganam, S., Y. Paterson, and E. Getzoff, *Structural Basis for the Binding of an Anti-cytochrome c Antibody to its Antigen: Crystal Structures of FabE8-Cytochrome c Complex to 1.8 AÊ Resolution and FabE8 to 2.26 AÊ Resolution.* J Mol. Biol., 1998. 281: p. 301-322.
66. Jemmerson, R. and Y. Paterson, *Mapping antigenic sites on proteins: implications for the design of synthetic vaccines. .* BioTechniques, 1986. 4: p. 18-31.
67. Cooper, H.M., et al., *Site-directed chemical modification of horse cytochrome c results in changes in antigenicity due to local and long-range conformational perturbations.* J. Biol. Chem., 1987. 262: p. 11591-11597.
68. Aguet, M. and G. Merlin, *Purification of human γ interferon receptors by sequential affinity chromatography on immobilized monoclonal antireceptor antibodies and human g interferon. .* J. Expt. Med., 1987. 165,. p. 988-999.
69. Gray, P.W., et al., *Cloning and expression of the cDNA for the murine interferon gamma receptor.* Proc. Natl. Acad. Sci. U.S.A., 1989. 86: p. 8497-8501.
70. Sogabe, S., et al., *Neutralizing Epitopes on the Extracellular Interferon γ Receptor (IFNγR) α-Chain Characterized Homolog Scanning Mutagenesis and X-ray Crystal Structure of the A6 Fab-IFNγR1-108 Complex.* J. Mol. Biol. , 1997. 273: p. 882-897.
71. Smith, A.M. and D.C. Benjamin, *The antigenic surface of staphylococcal nuclease.* J. Immunol. , 1991. 146: p. 1259–1264.
72. Jin, L., B.M. Fendly, and J.A. Wells, *High resolution functional analysis of antibody–antigen interactions.* J. Mol. Biol. , 1992. 226: p. 851-865.