IUPAC Technical Report

Douglas M. Templeton* and Kerstin Moehle

Structural aspects of molecular recognition in the immune system. Part I: Acquired immunity

Abstract: Humoral immunity allows the body to mount a defense against pathogens and foreign substances, and to respond with memory to subsequent exposures. The molecular participants may also recognize self-structures, leading to attack on the body and autoimmune disease. The main players in humoral immunity are antibody-producing B lymphocytes, and several classes of T lymphocytes. This review deals with the molecular details of recognition of antigens by soluble antibodies, and of substances presented to receptors on the surfaces of T cells (TCRs). The prototype antibody consists of a dimer of dimers, two heavy (H) chains and two light (L) chains, with antigen recognition capacity lying in variable "head" regions of an H-L pair. Most crystallographic studies are done with this substructure, called a F_{ab} fragment, bound in a soluble antigen complex. Homologous to this arrangement, the prototype TCR consists of two chains (α and β) that complex not soluble antigen, but usually a short peptide or other small molecule presented by proteins of the major histocompatibility complex. In each case a general background on the historical development of understanding the molecular recognition interface is given, followed by a number of examples of crystal structures from the recent literature that have allowed us to refine our understanding of the complex recognition process. Variations on the prototypical structures are also considered. The spectrum of recognition strategies involves interplay of lock-and-key with flexibility, varying degrees of entropic and enthalpic contributions, surface shaping by entrapped water molecules, and combinations of stabilizing hydrogen bonding, electrostatic interactions, salt bridging, and van der Waals forces. Preeminent in the recent literature are details of antibody binding to influenza A and human immunodeficiency viral antigens. Both viral antigens and attempts to understand autoimmunity are prominent in the recent TCR literature.

Keywords: antigens, antigen–antibody recognition, immunity, IUPAC Chemistry and Human Health Division, molecular recognition strategies, T-cell receptors, viral epitopes.

CONTENTS

*Corresponding author: Douglas M. Templeton,** Department of Laboratory Medicine and Pathobiology, University of Toronto, 1 King's College Circle, Toronto M5S 1A8, Canada, e-mail: doug.templeton@utoronto.ca
Kerstin Moehle: Chemistry Department, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

1 General introduction

The immune system provides humans, and a number of other species of higher organisms, with protection against a variety of microorganisms, biological molecules, and other chemical structures that it recognizes as being "foreign" to the self. The human immune system embodies an innate response that offers immediate recognition and defense, as well as acquired immunity that provides defense upon recognition of repeat exposure. Acquired immunity involves both humoral immunity resulting from the production of antibodies by a class of white blood cells called B lymphocytes, as well as cell-mediated immunity achieved by activation of a group of thymocyte-derived cells – the T lymphocytes. Triggers and molecular mechanisms of activation of innate, humoral, and cell-mediated immunity are exceptionally complex, and are discussed in a number of good textbooks of immunology (see [1] for an excellent recent example). The aim of the present paper is not to provide background on mechanisms of immune activation, but rather to discuss structural aspects of molecular recognition within the immune system. Thus, we present principles and examples of recognition of antigens by B lymphocyte-derived antibodies, and of antigens presented in specific molecular complexes to receptors on T lymphocytes. A companion paper deals with interactions that mediate innate immunity.

2 Antigen–antibody (AG-AB) recognition

Information Box 1

Antigens, structures that are specifically recognized by and interact with receptors on the surfaces of cells of the immune system and with secreted antibodies, include protein epitopes, carbohydrates, lipids, nucleic acids, and small-molecule haptens. The recognizing antibody (immunoglobulin) consists of two main regions, called the variable (V) region and the constant (C) region. The variable region is the antigen-recognition site, and its variability from antibody to antibody allows the occurrence of antibodies that each can show specificity against the nearly endless array of foreign and self-antigens that the immune system may encounter.

Information Box 1 (Continued)

The constant region is the effector part of the molecule that links the antigen–antibody complex to common immune system responses such as lymphocyte activation, phagocytosis, and the complement cascade.

Information Box 2

Human immunoglobulins (Ig) consist of two heavy (H) and two light (L) chains connected by disulfide bonds and are traditionally depicted as a planar "Y-shaped" structure (Fig. 1a). Based on the nature of the H chains, human antibodies are classified as IgG, IgA, IgE, IgM, and IgD, with H chains designated γ, α, ε, μ, and δ, respectively. The most widely studied, IgG, has subclasses IgG1–IgG4 based on the occurrence of H chains $\gamma 1$–$\gamma 4$. More detailed structural analysis reveals the L chains have an N-terminal V domain (V_L) and a carboxy-terminal C_L domain. As well as corresponding V_H and $C_H 1$ domains, the H chains have additional $C_H 2$ and $C_H 3$ domains. An extended peptide region between $C_H 2$ and $C_H 3$ is referred to as the hinge region. Each domain is structured with two juxtaposed β sheets consisting of three and four anti-parallel β strands (five and four in the V domains), respectively, protecting interior hydrophobic residues. Domains V_H-V_L, $C_H 1$-C_L, and $C_H 3$-$C_H 3$ are in close pairing contact; a variable carbohydrate chain on each $C_H 2$ domain disrupts $C_H 2$-$C_H 2$ contacts. Antigen recognition resides in the V_L-V_H domain pairing, with three regions of hypervariable amino acid sequence in each domain being the major determinants of recognition. Effector functions reside mainly in the $C_H 2$ domains. When cleaved from each other, the disulfide-bonded V_L-V_H-C_L-$C_H 1$ fragment is referred to as F_{ab} (fragment antigen binding) and the disulfide-bonded $(C_H 2$-$C_H 3)_2$ fragment is denoted F_c (fragment crystallizable).

While the traditional depiction is of a "Y-shaped" molecule, the structure has considerable flexibility in the hinge region (Fig. 1b). Bending (of the F_{ab} at the elbow hinge), wagging (of both the F_{ab} arm and the F_c region), and rotational motions (of the F_{ab} arm) allow the F_{ab}-F_{ab} angle to more closely approximate a "T-shape", the F_c component may rotate out of the F_{ab} plane, and the $C_H 2$ domain may more closely approach one of the $C_H 1$ or C_L domains.

Information Box 3

The variable regions in the F_{ab} fragment binding structures are punctuated by even more variable (thus, hypervariable) regions, three each in the L chain and H chain. The high degree of amino acid sequence heterogeneity in these so-called complementarity-determining regions (CDRs) provides much of the variability necessary to recognize the huge number of possible antigens with high specificity. Each of the CDRs, designated L-chain CDR1-3 and H-chain CDR1-3, can be involved in antigen recognition. However, H-chain CDR3 is special: it is often dominant in binding, and it has an extended finger-like conformation that often allows it to penetrate into grooves and pockets of the antigen, providing multiple favorable contacts. It will be seen below that CDRs also occur in the T-cell receptor. As will be appreciated, the somatic gene rearrangements and sequence shuffling events that give rise to these hypervariable sequences are quite complex, and have been studied in detail, but they are beyond the scope of this discussion of molecular recognition.

2.1 Introduction

In general, recognition of a protein is the responsibility of the F_{ab} end of the antibody and is strongly dependent on the hypervariable regions (complementarity-determining regions or CDRs). A "goodness-of-fit" model applies, with thermodynamic affinity dependent on the number of favorable contacts. Concepts of protein–protein interactions apply, e.g., an increase in entropy when water is excluded from hydrophobic interfacial contacts decreasing the free energy and increasing affinity. Antigenicity refers to the "ability of an antigen
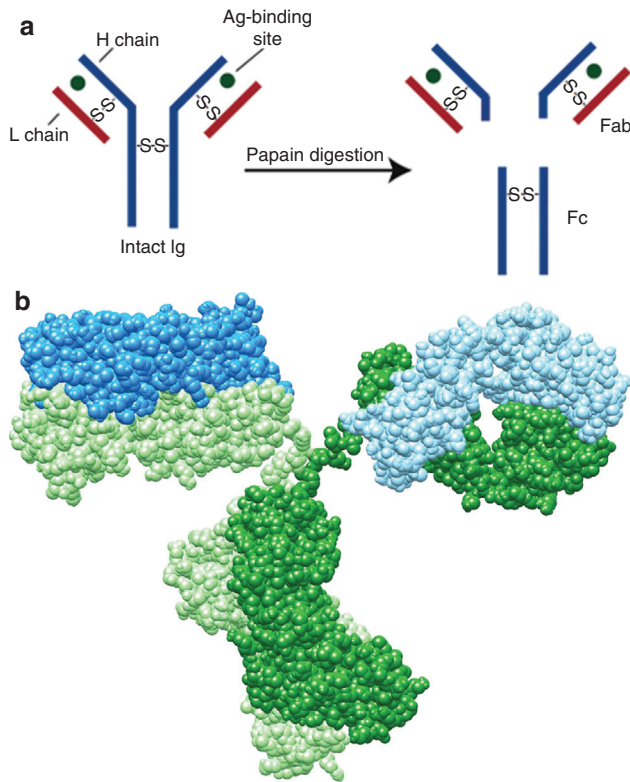
Fig. 1 Depictions of immunoglobulin structure. (a) Conventional schematic of a bivalent immunoglobulin showing disulfide-bonded H chains (blue) and L chains (red), presenting two binding sites for antigen (green). Upon digestion with papain, cleavage of the H chains releases two F_{ab} and one F_c fragments. (b) A more realistic depiction based on the crystallographic data of [2]. The two L chains are shown in blue, the H chains in green.

to bind to a specific antibody...often a measure of its ability to produce immunity" [3], the point being that strength of binding per se does not elicit the immune response, which is better referred to as immunogenicity.

The part of the antigenic molecule recognized by the antibody receptor is referred to as an epitope, and a protein antigen may have multiple epitopes that are recognized by distinct antibodies. Mapping the epitope surface of a protein with different antibodies is called epitope footprinting, and it may be observed that some epitope footprints overlap (for instance, when a particular amino acid is involved in two or more epitopes), while others do not. In antibody recognition of an antigenic protein, it is often noted that amino acid residues from well separated regions of the protein sequence are involved in an epitope, then said to be discontinuous, due to folding of the protein into a tertiary structure that brings these residues into spatial proximity. In special cases, several contiguous amino acids may constitute the epitope, and then the peptide that they define is said to represent a continuous or linear epitope.

A good discussion of the types of protein surface structures "seen" by antibodies is given in Delves et al. [1] (chapter 5). While complementarity of fit (shape) between antigen and antibody is an obvious consideration, flexibility is also important. Flexibility may allow the antibody to conform better to an epitope, or even represent a conformational change induced in the antibody by antigen. Often an antibody will recognize an extended surface, exploiting additivity of thermodynamically favorable contacts. Or, it may have an apparent cavity or groove, especially to accommodate a small molecule (hapten). The hypervariable region H-chain CDR3 is often a major contributor to recognition, and may take on an extended finger-like conformation that may fit into a pocket on a antigenic protein. A common antigenic motif is a loop protruding from the protein surface, which tends to allow more flexibility than the surrounding compact surface. Continuous epitopes are often well-structured, such as an α-helix preserved in the isolated peptide.

We now consider some historical developments that illustrate these points and the various approaches to antigen–antibody structural recognition.

2.2 Historical development

Houghten [4] presented a rapid method of solid phase peptide synthesis that allowed, as a secondary objective, extending existing knowledge that antigenic determinants resided in protein sub-structures, by attempting to identify single amino acids involved in the interaction. The approach was to study monoclonal antibody binding by enzyme-linked immunoabsorbent assay (ELISA) to a series of peptide variants. The antibody was raised to the linear epitope representing amino acid residues 75–110 of influenza hemagglutinin, HA1, and 248 single amino acid variants of residues 98–110 were tested. The results showed that Asp[101] was critical for antibody binding, while Asp[104] and Ala[106] also played significant roles, with the other amino acids having little influence on binding.

An early example of the use of synthetic peptides to probe the antigen–antibody interaction used four monoclonal antibodies recognizing linear epitopes at the 18-residue C-terminus of rhodopsin [5]. It was concluded that ionic interactions can dominate the interaction, but the antigenic determinant was not restricted to highly charged regions. Results were consistent with the view that the free energies of all contact point interactions are additive and a threshold energy must be exceeded to achieve binding. But a deeper understanding of the binding interaction awaited crystallographic data.

At this time, creative use was being made of physical methods for the qualitative understanding of antigen binding. Surface plasmon resonance (SPR) continues to be useful in characterizing antibody-binding affinities, although a great deal of suspect data have probably been generated by over-interpretation of results (see D. G. Myszka's entertaining Foreword in [6] for a thoughtful discussion). An early example of its use was generated in-house at Pharmacia in Uppsala during the development of commercial instrumentation. Fägerstam et al. [7] immobilized an anti-mouse IgG1 on a dextran-derivatized sensor chip and then washed with either a monoclonal antibody to human immunodeficiency virus (HIV) core protein p24, or an unrelated antibody to block unoccupied IgG1 sites. Antigen (p24) was then passed over the surface, followed by a second distinct anti-p24 monoclonal antibody. This protocol allowed a type of epitope mapping by assessing the interactions of the first and second anti-p24 antibodies. Up to five antibodies could be bound sequentially, resulting in a hexamolecular complex of non-overlapping epitopes, and these could be mapped further using competing peptides from known sites on the antigen. Analysis of competition also allowed comment on partial epitope overlap. Although the authors point out the advantage of SPR in not requiring labeling of individual antibodies in these studies, they also recognized the inherently qualitative nature of the results.

Atomic force microscopy was also being brought to bear. In early experiments [8], biotin linked by a spacer arm to albumin served as the antigen and was immobilized on the sensor tip of the atomic force microscope. Antibody (polyclonal anti-biotin IgG) was immobilized on a flat gold substrate with dithiobis(succinimidylundecanoate) crosslinker to minimize antibody-surface contacts and potential attendant conformational changes. After binding, retraction forces in multiples of 60 ± 10 pN were observed, which were absent when the reaction was blocked with streptavidin, when albumin replaced biotinylated albumin, etc. In principle, with monoclonal antibodies and careful antibody orientation, it should be possible to correlate these force measurements with data from thermodynamic experiments.

Notwithstanding the elegant use of these physical techniques, a more detailed appreciation of structural interaction of course awaited crystallographic studies. Not long after the early report of Houghten [4], X-ray structures of two crystal forms of monoclonal F_{ab} HyHEL-5 bound to hen egg white lysozyme were published [9]. While structures of F_{ab} fragments with haptens had been studied previously, and demonstrated the presence of hapten binding pockets [10], the possibility that complexes with protein antigens might more resemble other protein–protein interfaces with exclusion of water and induced conformational changes in the antibody were considered. The results validated earlier epitope-mapping studies [11]. The lysozyme epitope consisted of three sequentially separated regions. A shallow groove accommodated positively charged Arg[45] and Arg[48] of lysozyme that were complemented by two negative Glu residues in the antibody heavy chain (Glu[35] of H1 and Glu[50] of H2) (Fig. 2). Shape and hydrogen bonding are also important, as change of Arg[68] to lysine significantly decreases antibody affinity. Otherwise, the area of contact was relatively flat (about 1.9×2.6 nm), and contained no water in the interface region. The backbone structure of lysozyme was mostly unperturbed, and
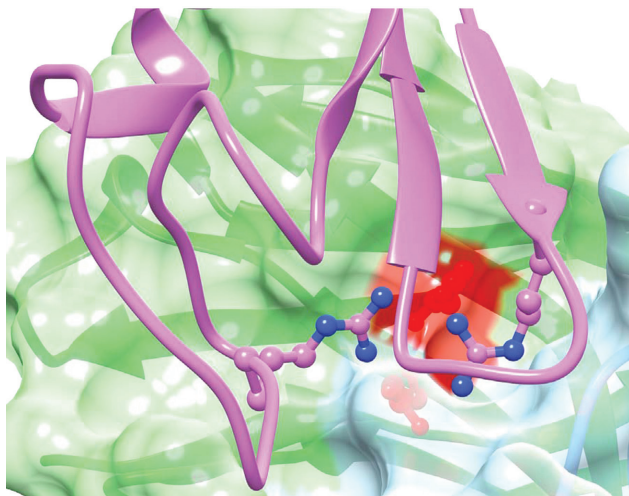
Fig. 2 An epitope in hen egg white lysozyme (pink ribbon) lies in a shallow groove on the antibody surface (green). Two Arg residues (N's blue) from lysozyme contact two Glu residues (red) in the antibody H1 and H2 loops. The figure is based on the crystallographic data of [12].

only 30 % of the residues in the CDRs were in contact with lysozyme. It was concluded [9] that binding of F_{ab} to protein antigen shares with hapten binding charge neutralization at the interface, but results in much greater surface area of contact with exclusion of water. Variable elbow bends in the early structures were indicative of flexibility in this region of the antibody but did not correlate with binding affinity.

When Li et al. published a refined structure of F_{ab} HyHEL-63 bound to lysozyme in 2000 [13] they noted that the three-dimensional structures of about 25 complexes between antibodies and various protein antigens by then had been determined. A significant difference between their experiments and those with the earlier lysozyme-antibody binding study was the ability to crystallize the F_{ab} in the absence of antigen, and refinement of all structures to about 0.20 nm resolution. In addition, they introduced single Ala mutations into the lysozyme epitope to study energetic contributions of individual residues. Their results showed several notable differences from those discussed above. First, all six CDRs of F_{ab} are in contact with the lysozyme epitope, and the interfacial interaction includes 24 hydrogen bonds, a salt bridge, and 170 van der Waals contacts (Fig. 3).
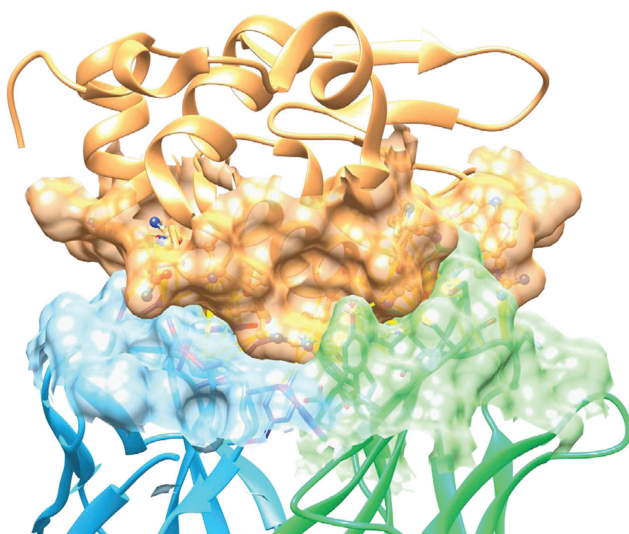

Fig. 3 A more detailed look at the contact of a hen egg white lysozyme antibody with the lysozyme (brown) epitope reveals multiple contacts with all 6 CDRs of the F_{ab} (blue and green) that smooth the surface fit. See text for details. The figure is based on the crystallographic data of [13].

While the antibody side of the interface is predominantly hydrophobic, hydrophilic residues predominate on the lysozyme side, with charged residues tending to be at the periphery pointing away from the center of the interface. Alanine substitutions revealed that residues at the center of the interface contributed most to complex stability. Second, because the F_{ab} antibody was crystallized with and without antigen, a comparison of structures was able to reveal small conformational movements in the CDR loops and flexibility in side chains of amino acids in the interface region that enhanced complementarity of binding. Third, far from being devoid of water, 11 ordered water molecules could be ascertained in the interface region that were not observed at fixed sites in lower resolution. By filling gaps in the interface region, trapped water molecules increased surface complementarity. The authors suggested that in contrast to other protein–protein interactions that have maximized complementarity through coevolution, water molecules may serve to correct imperfections in antigen–antibody interfaces by improving fit and accommodating unpaired hydrogen-bonding residues.

2.3 Current perspectives

The above considerations have briefly set a context for a detailed examination of antigen–antibody interactions. Lock-and-key and induced-fit models are appropriate, and protein recognition by antibodies is firmly established as residing in the F_{ab} domains, strongly involving the CDRs of the hypervariable regions. Thermodynamic principles are useful in explaining binding affinities, with van der Waal's contacts, hydrophobic and π interactions, hydrogen bonding, and electrostatic interactions all contributing in a roughly 2.5×2.5 nm² surface; solvent (water) plays a significant role both entropically with respect to its exclusion, and also by its inclusion to shape the contact surface. Bending and rotating angles in the antibody are somewhat flexible, and such flexibility can contribute significantly to binding energy. Several hundred antigen–antibody crystal structures are available today, and newer structures are focusing not on complexes with model proteins such as lysozyme, albumin, and hemagglutinin, but rather on structures of therapeutic and biomedical significance, such as receptors and viruses. What more can be learned from these to guide future antibody design? Several recent examples are discussed here to illustrate some details of the binding interface, and they demonstrate how the field has quickly matured to studying structures of potential therapeutic significance. But, while each of these recent papers discussed below begins with a discussion of relevance to disease rather than addressing a question in physicochemical determinants of the antigen–antibody interaction, each nevertheless contributes to our growing understanding of the chemical nature of the antigen–antibody interface.

A majority of therapeutic antibodies fall into the category of inhibitors of target protein function. Transmembrane serine proteases are over-expressed on the plasma membrane of some cancer cells, where they may play a role in matrix remodeling and invasion. A phage display library created with a repertoire of F_{ab}s from human B cells was used to identify F_{ab} A11, which bound and inhibited with high specificity the transmembrane type II serine protease, matriptase [14]. The crystal structures of matriptase complexed with A11 and with a previously known potent matriptase inhibitor, F_{ab} S4, were determined at 0.21 and 0.15 nm, respectively, and compared. Both antibodies made multiple contacts with multiple surface loops on the protease and a total buried surface area 10–12 nm².[1] In both cases, H and L chains both contributed to surface coverage, with the H3 loop making the major contribution (63 % of the total for S4, 38 % for A11) (Fig. 4). Most of the A11 contacts are nonpolar, whereas S4 interacts with multiple polar interactions in a network of hydrogen bonds and π-stacking interactions. Binding had little effect on matriptase conformation remote from the binding site. In both cases, the H3 loop resides in the substrate-binding pocket, but the greater potency of the S4 inhibitor may be due to greater interaction of H3 with the active site. While H3 loops of both S4 and A11 form β-hairpin turns in the binding pocket, A11 makes fewer contacts and has less extensive coverage. These observations suggest a means of tailoring the H3 loop to increase inhibitor potency. In contrast, another

---

1 Note that above, a typical footprint of 2.5×2.5 nm = 6.25 nm² was indicated, which would typically bury a surface area of about 12.5 nm², taking both surfaces into account. The distinction between footprint and buried surface area should be kept in mind throughout this text.
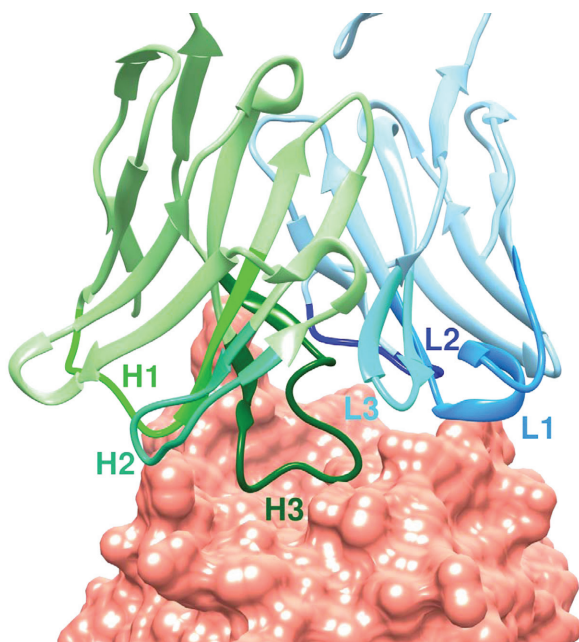
Fig. 4 Complex of the serine protease inhibitor, matriptase, with inhibitory F_{ab} S4. All three CDR loops of the H chain (H1–H3) and the L chain (L1–L3) make contact in the interfacial region, with the H3 loop dominating the interaction. The figure is based on the crystallographic data of [14].

transmembrane type II serine protease, hepsin, is prominent in prostatic cancer. A human monoclonal antibody, hH35, selectively inhibits hepsin, but does not appear to interfere directly with substrate binding, acting rather by allosteric inhibition [15]. F_{ab} hH35 binds to an α-helix well away from the active center, but leads to major structural changes that inactivate the enzyme.

Given the importance of angiogenesis in cancer progression, a good deal of information has been presented on the binding of the vascular endothelial growth factor (VEGF) family members to the VEGF receptors. The F_{ab} portions of two antibodies that block VEGF binding to and/or signaling through VEGF receptor 2/KDR have been crystallized in complex with domain 3 of KDR and solved to 0.22 nm resolution [16]. The antibodies, 1121B and 6.64, had previously been shown to bind exclusively to domain 3, and to do so independently of one another. F_{ab} 6.64 makes contact over an extended region of KDR domain 3 (coverage 9.82 nm², total buried surface area 20.3 nm²), involving a highly discontinuous epitope, and contacts the C-terminus (representing the KDR domain 3/domain 4 interface) (Fig. 5). All six H- and L-chain CDRs make contact, and 45 % of contact is with the L chain. F_{ab} 1121B binds mostly to a more localized region (nearly continuous β-hairpin epitope) at the N-terminus of domain 3, near the domain 3/domain 2 interface, accounting for the lack of competition of the two antibodies (Fig. 6). Binding is dominated by the H chain, with the L chain contributing one charge pair and a few hydrophobic interactions for a coverage of 10.0 nm² (total buried surface 19.2 nm²). There is very slight change in F_{ab} conformation on binding, and also not much change in KDR domain 3 structure, consistent with the localized binding. The results support two different mechanisms for inhibition. While F_{ab} 1121B does not share binding residues with VEGF, it is immediately adjacent to the VEGF binding site and may sterically inhibit binding. On the other hand, F_{ab} 6.64 is not near the VEGF binding site, but may induce conformational changes in the domain 3/domain 4 region that interfere with receptor dimerization.

Potent allergens in house dust mites include the cysteine proteases Der p 1 and Der f 1, which induce a strong response of IgE antibodies [17]. These antigens share 81 % sequence identity and elicit highly cross-reactive responses. A F_{ab} fragment of a monoclonal antibody 4C1, which inhibits IgE binding, was used to identify a common epitope in IgE–Der f 1/p 1 interaction. The epitope was well away from the active center. Buried surface areas of 7.5–7.9 nm² on the two antigens were found in the crystal structure at 0.205 nm resolution. In coverage, 70 % was contributed by contacts with the H chain, mostly arising from the H-chain CDR3
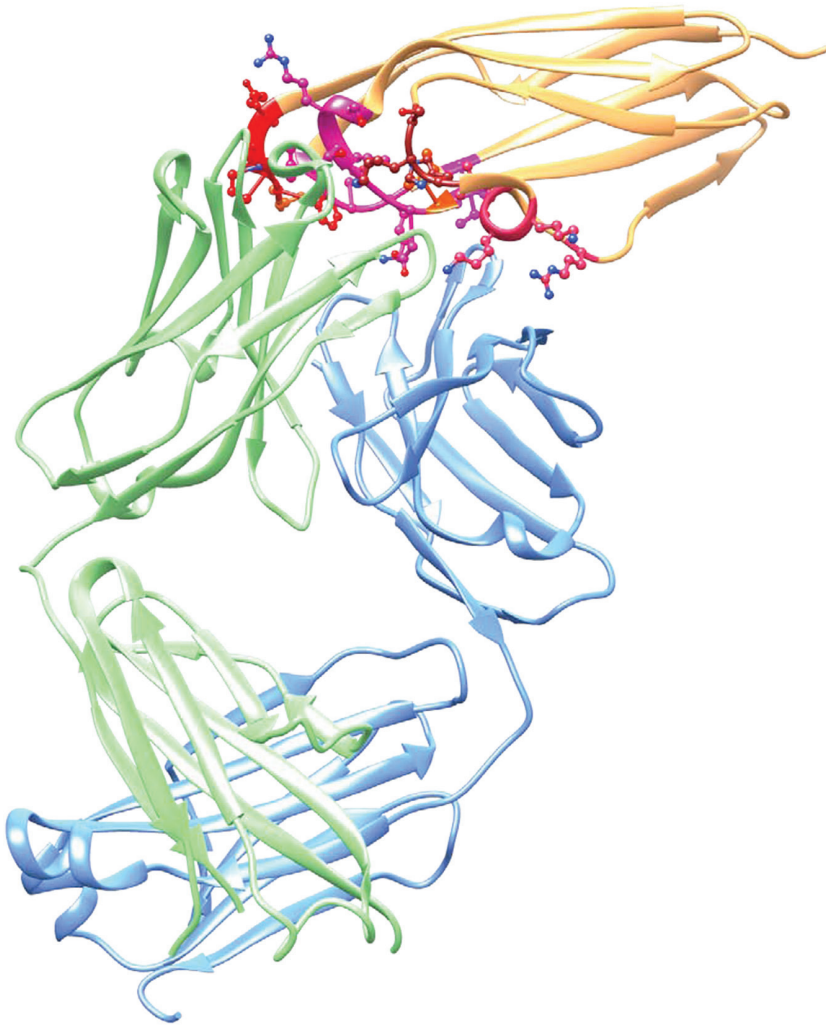
Fig. 5 Structure of the anti-VEGF receptor KDR3 domain bound to F$_{ab}$ 6.64. Note the highly discontinuous epitope, which makes contact with all six H-chain (green) and L-chain (blue) CDRs. See text for more detail. From the crystallographic data of [16].

region being accommodated in a concave region of the antigen surface (Fig. 7). The antigen presented a central rigid surface of discontinuous residues. Overall conformation of the antibody and the conformations of its CDRs were similar in complexed and uncomplexed states, indicating little rearrangement upon binding. Conserved water molecules in the two antigens improved the fit at the interacting surfaces. Elbow angles (C$_L$-C$_H$1 and V$_L$-V$_H$) showed some flexibility (e.g., 146° and 139° in two different crystal forms of 4C1 F$_{ab}$ complexes), consistent with differences of 143° and 167° in two orthorhombic crystal forms of the uncomplexed F$_{ab}$ (Fig. 8). The authors [17] proceeded to analyse 16 available allergen–antibody complexes in the structure database and found a common theme that Tyr and Ser residues in the antibody played a dominant role in interaction with the allergen, followed by Arg, Asn, and Asp. In 4C1 interactions Tyr and Arg were particularly prominent, with potential contributions from hydrogen bonding, π bonding, and hydrophobic interactions. Interestingly, hydrogen-bonding networks in the binding interfaces of the 16 allergens were most frequently formed by Arg-Asp, Arg-Tyr, Asp-Arg, Asp-Tyr, Glu-Arg, Glu-Ser, and Gln-Ser pairs, distinct from a number of lysozyme–antibody complexes where Arg-Glu, Arg-Tyr, Asn-Gln, Asp-Ser, and Lys-Asp predominated [17].

A structure has been reported [18] at 0.28 nm resolution of the interaction of a single-chain variable domain (scF$_v$) antibody, with interferon α (IFNα), of relevance to the study of the autoimmune disease, systemic lupus erythematosus. Single-chain F$_v$ antibodies are fusion proteins of the H- and L-chain variable regions linked with a spacer sequence, isolated following a strategy of cloning variable-domain genes. The
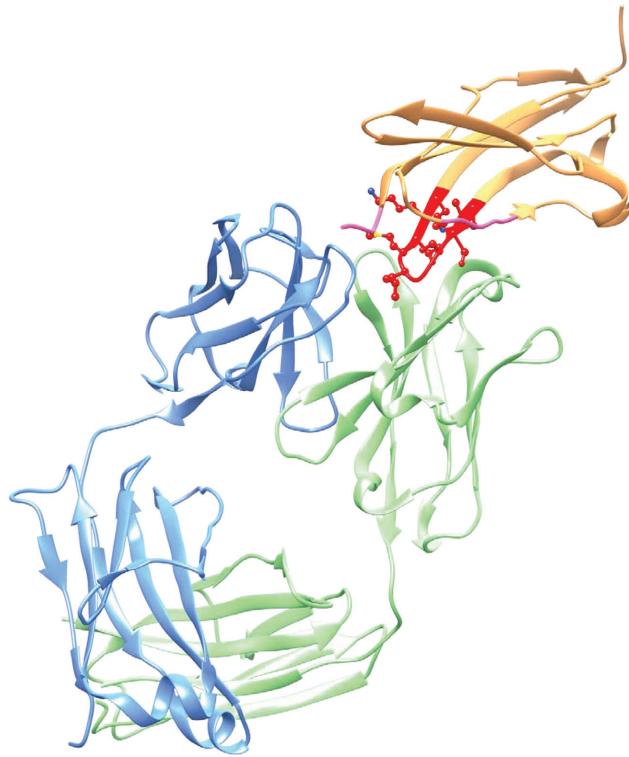
Fig. 6 Structure of the anti-VEGF receptor KDR3 domain bound to F_{ab} 1121B. Note the highly localized β-hairpin epitope in KDR3 that is predominantly bound to the H chain (contrast this with the binding mode described in Fig. 5). From the crystallographic data of [16].

structure of the complex revealed that three distinct regions in the interface contributed to high-affinity binding. In each case, a protruding side chain of one partner fits a cavity in the other – two cavities in IFNα and one in the antibody – allowing the structures to "zip together" (Fig. 9). Shape complementarity was thus a major determinant of binding strength, with an important salt bridge in one contact, numerous van der Waals and π-bonding interactions in another, and structural entrapment of an antibody Phe residue in an IFNα cavity in the third. Sixteen hydrogen bonds are included in the interactions.

The potentially therapeutic antibody ponezumab is directed against the β-amyloid peptide Aβ40, aggregation of which is implicated in the pathogenesis of Alzheimer's disease. La Porte et al. [19] have reported the crystal structure of ponezumab F_{ab} complexed with Aβ40 at 0.25 nm resolution. The antigen-binding site is in a cleft between the H- and L-chain variable domains, includes all six CDRs, and buries 7.59 nm² of the peptide surface. Interaction is dominated by nonpolar amino acid residues, with 10 amino acids from Aβ40 and 24 from ponezumab, 12 of the latter contributed by each chain. The C-terminal carboxyl group plays a key role in binding as demonstrated by truncation studies and SPR, and in the crystal structure it is buried by contact with Lys, Arg, and two Tyr residues of the H-chain variable domains (Fig. 10).

Of relevance to another neurodegenerative disease, the monoclonal antibody 3B5H10 binds to huntingtin (htt), the protein implicated in Huntington's disease. In common with several proteins implicated in neuro-degenerative diseases, htt contains stretches of polyglutamine that are expanded with increasingly severe manifestations of the disease. It is thought that expansion of the polyglutamine region leads to pathological protein self-aggregation with a cross-β-strand structure similar to β-amyloid. A crystal structure of 3B5H10 F_{ab} at 0.19 nm resolution revealed enrichment of solvent-accessible aromatic residues in the H-chain CDRs, facilitating glutamine binding [20]. In fact, 30 % of the H-chain CDRs were aromatic, compared to 9 % in 100 closest homologues. In a second polyglutamine-recognizing antibody, MW1, this number is 39 % [20]. Upon recognition of the polygluatmine epitope, a 3° rotation between the H- and L chains results in a linear groove that accommodates the antigen. An unusual rigid loop structure in L-chain CDR3 forms part of the wall of
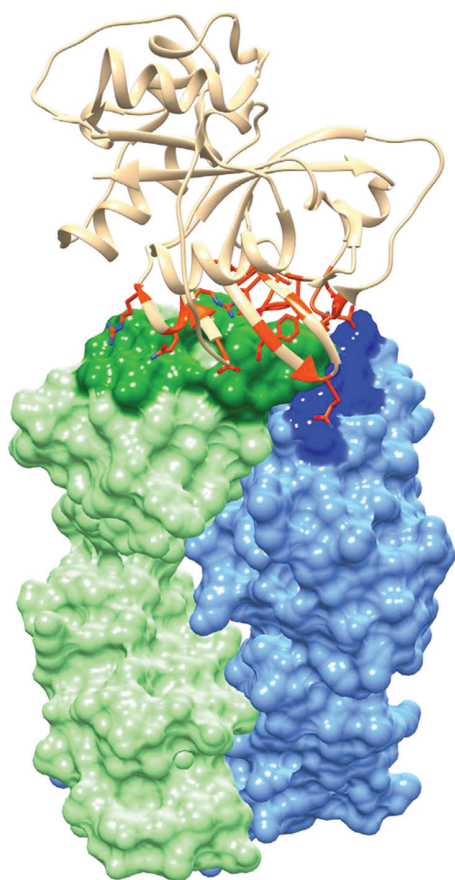
Fig. 7 A F$_{ab}$ fragment of a monoclonal antibody 4C1 was used to identify a common epitope in interaction of IgE with house dust mite antigens Der f 1/p 1 [17]. 70 % of surface contact was contributed by interactions with the H chain, mostly arising from the H-chain CDR3. Antibody contact surfaces within 0.5 nm distances to the antigen are colored dark green (H chain) and dark blue (L chain). Antigen contacts, orange.

the binding groove in both 3B5H10 and MW1. While only a single strand of polyglutamine is accommodated in the MW1 groove (Fig. 11a), in 3B5H10 a widened groove with an extended β-hairpin loop of H-chain CDR2 interacting with H-chain CDR1 and L-chain CDR3 (Fig. 11b) fits a compacted two-stranded hairpin in the polyglutamine peptide. Because recognition by 3B5H10 of htt with expanded polyglutamine sequences is a better predictor of neurodegeneration than recognition by MW1, the compacted polyglutamine loop may be of pathological significance.

The A$_{2A}$ adenosine receptor, A$_{2A}$AR, is a G protein-coupled receptor involved in regulating cardiac blood flow and regulation of release of dopamine and glutamate neurotransmitters in the brain. It is antagonized by caffeine. A monoclonal antibody F$_{ab}$ fragment, Fab2838, blocks binding of receptor agonists, but not the antagonist ZM241385. A$_{2A}$AR with bound ZM241385 has been crystallized in the presence of Fab2838 and solved at 0.27 nm resolution [21]. ZM241385 occupies an extracellular ligand-binding pocket with hydrophobic (Phe and Ile) and hydrogen-bonding interactions (Asn). Fab2838 binds on the intracellular side of the receptor with an unusually long H-chain CDR3 fitting a pocket formed by four receptor helices (Fig. 12) and stabilizes an inactive form of the receptor. Interactions are mainly hydrogen-bonding and van der Waals, with interaction between CDR3 and two of the helices stabilized by a hydrogen-bonding network containing two water molecules. Other CDRs form an additional 14 hydrogen bonds leading to a high-affinity interaction with K_d = 4.4 nM. For comparison, another antibody, Nb80, binds to the β$_2$AR with a similar H-chain CDR3 β-hairpin and recognizes an active form of the receptor, whereas the β-hairpin Fab2838 CDR3 induces a different shape of binding pocket and locks the receptor in an inactive conformation [21].

Fig. 8 Flexibility of elbow angles occurs in two different crystal forms of F_{ab} 4C1 directed against dust mite protease antigen. Superimposed structures are based on the crystallographic data of [17]. Angles are 143° and 167° for $P2_12_12_1$ (blue) and $C222_1$ (orange) forms, respectively.

Venoms that act as blockers of membrane channels have been another target for exploration of potential therapeutic antibodies. The scorpion *Androctonus australis hector* (Aah) produces lethal peptide toxins AahI, II, III, and IV that block mammalian voltage-gated Na$^+$ channels. The peptides share a so-called cystine-stabilized α-helix structure consisting of an α-helix and an antiparallel three-strand β-sheet, stabilized by disulfide bridges. However, variation in the amino acid side chains among the peptides precludes cross-reactivity and prevents generation of a single antiserum against all the toxins. A high-affinity F_{ab} was generated from a monoclonal antibody 4C1 to AahII, and the crystal structure of the F_{ab}4C1–AahII complex was solved at 0.23 nm resolution [22]. Another antibody fragment, F_{ab}9C2 directed against AahI, excluded its antigen from the crystal structure. The F_{ab}4C1–AahII complex was described by the authors as resembling "an egg inserted small-end first in the egg cup". The F_{ab}4C1 "egg cup" consists of a binding pocket 1.3 nm deep and 1.2 nm wide formed by the six CDRs, with an elbow angle between the variable and constant domains of the H and L chains of 140.5° (Fig. 13). An extended L-chain CDR1 and an anionic H-chain CDR2, together with the long H-chain CDR3, form a boundary that "clamps" the cationic toxin. L-chain CDR2 and CDR3, and a hydrophobic H-chain CDR1 serve as additional anchoring points. Strong complementarity buries about 25 % (10 nm^2) of the venom surface at the epitope–paratope interface. A more planar, less anionic, hydrophobic binding surface in F_{ab}9C2 may account for exclusion of antigen in attempts to crystallize its AahI complex. Also, an elbow angle of 171° in F_{ab}9C2 suggests greater flexibility [22].
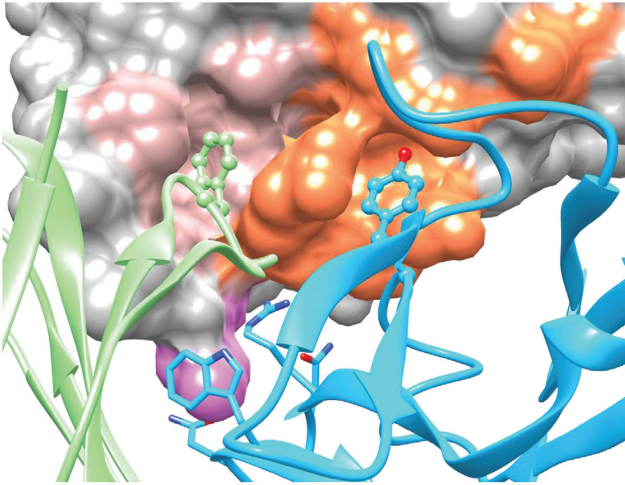
Fig. 9 Interaction of a single-chain variable domain (scF$_v$) antibody with interferon α (IFNα), showing how the structures "zip" together. The antibody is depicted in a ribbon structure contacting IFNα (solid surface). A Met residue (magenta) from IFNα fits into a pocket formed by a Trp, an Arg, and two Asn residues of the antibody, while antibody H-chain (green) Phe and L-chain (blue) Tyr residues are accommodated in distinct cavities of IFNα. After [18].
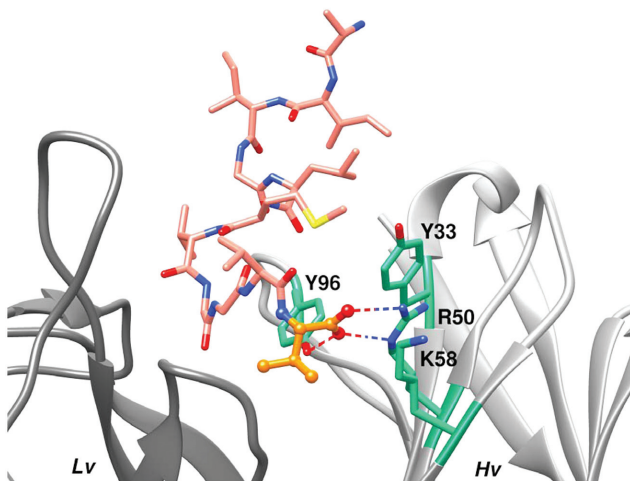

Fig. 10 Structure of the β-amyloid peptide Aβ40 complexed with F$_{ab}$ ponezumab. The antigen binding site is in a cleft between the H- and L-chain variable domains (L$_V$ left, H$_V$ right). Dominating the interactions is that between the peptide carboxyl terminus (orange, O red, N blue) in contact with Lys, Arg, and two Tyr residues of the H-chain variable domains. From [19].

Complement factor D is an enzyme of the complement pathway of innate immunity. Due to a high turnover rate, factor D is present at low concentrations in blood and becomes rate-limiting in the alternative complement pathway, making it a therapeutic target for inappropriate complement activation. Monoclonal antibody 166-32 binds to factor D with high affinity, and a humanized F$_{ab}$ fragment that blocks complement activation has been prepared by grafting the CDRs of 166-32 on human germ line immunoglobulin regions; this construct has been designated AFD. The crystal structures of AFD bound to human and cynomolgus monkey factor D have been solved at 0.24 and 0.23 nm, respectively [23]. The two structures superimposed closely, and no significant changes in factor D structure were induced upon binding. The buried surface of factor D is 9.01 nm^2 (total of epitope plus paratope surface 16.5 nm^2) with 4.85 and 2.63 nm^2 covered on the AFD H chain and L chain, respectively. Eighty percent of the buried surface of factor D involves a protein loop sandwiched between the H chain and L chain and an Arg residue deeply buried in potential hydrogen
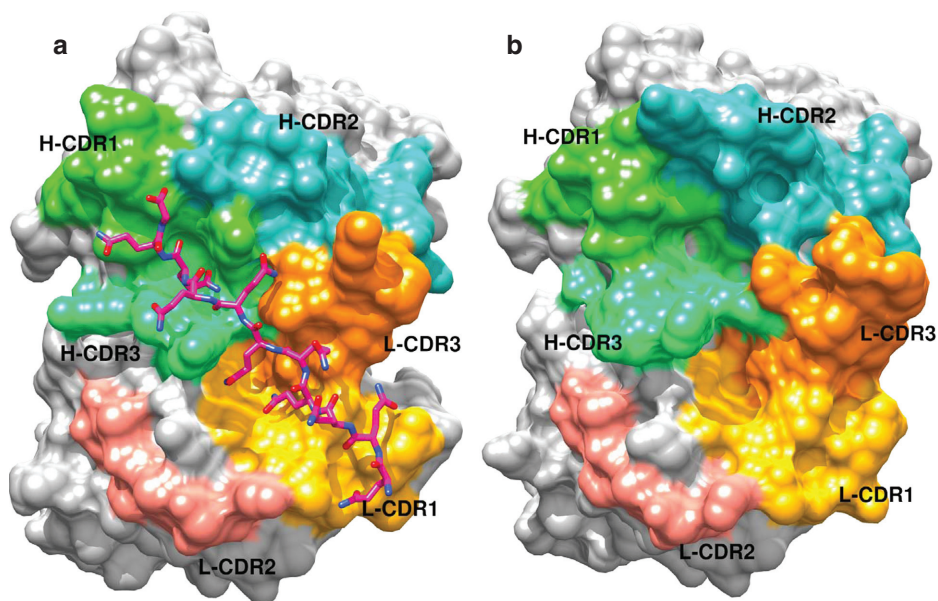
Fig. 11 Huntingtin protein (htt) binding to F_{ab} structures. (a) A polyglutamine sequence of htt lies along a groove formed across the V_H–V_L interface of antibody MW1 [20]. The strong interaction with L-chain CDR3 is somewhat unusual. (b) Note that in another htt-binding Ab, 3B5H10, the polyglutamine groove is widened by an extended H-chain CDR2 that interacts with H-chain CDR1 and L-chain CDR3, potentially accommodating a compact hairpin in the peptide (coordinates with peptide not available).

bonding and electrostatic contact with two Glu residues on the H chain and Tyr and Asn residues on the L chain (Fig. 14). An additional three Asp residues on the L chain form charged interactions with a factor D Lys residue, and contribute to very high affinity binding (K_d < 10 pM). Interactions differ slightly with the
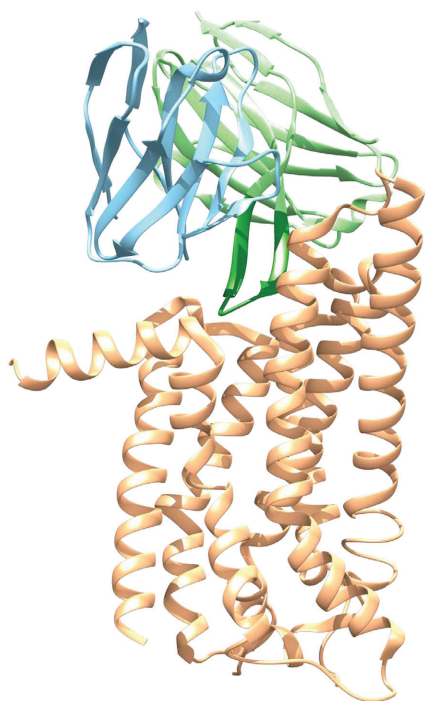

Fig. 12 Monoclonal Fab2838 blocks binding of agonists, but not antagonists, to the adenosine A_{2A} receptor. An unusually long H-chain CDR3 loop (dark green) fits a pocket formed by four adenosine receptor helices (brown). The figure is based on the crystallographic data of [21].
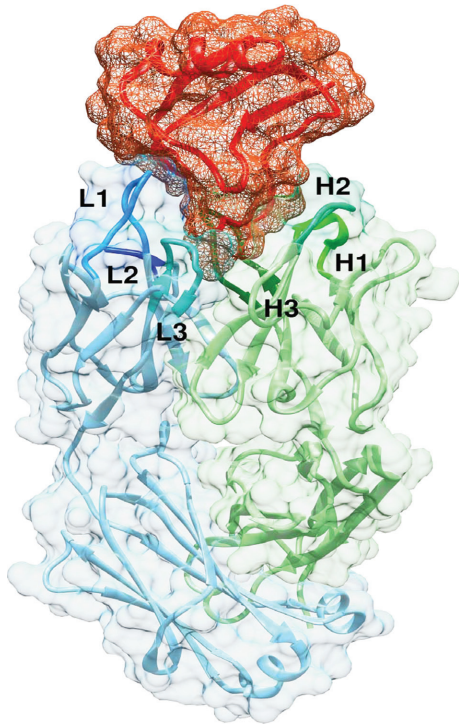
Fig. 13 The fitting of a scorpion toxin (red) into F$_{ab}$4C1 makes contacts with all six CDR loops of the H chain (green) and the L chain (blue). The authors [22] have described this as "an egg inserted small-end first in the egg cup" (see text for details). The basket around the toxin represents the molecular surface of the protein in mesh style, and the surface around F$_{ab}$ antibody is rendered as a transparent solid.

cynomolgus protein, and the affinity is lower. The binding interface is well removed from the catalytic site, and inhibition of complement activation appears to be due to steric hindrance between factor D and its substrate, factor C3bB [23].

"Two-in-one" antibodies are known that can bind with the same site to two different epitopes. An interesting example is D11 [24]. An antibody was generated against the epidermal growth factor receptor (EGFR)/human epidermal growth factor receptor-1 (HER1) from a library with diversity restricted to H-chain CDRs, so
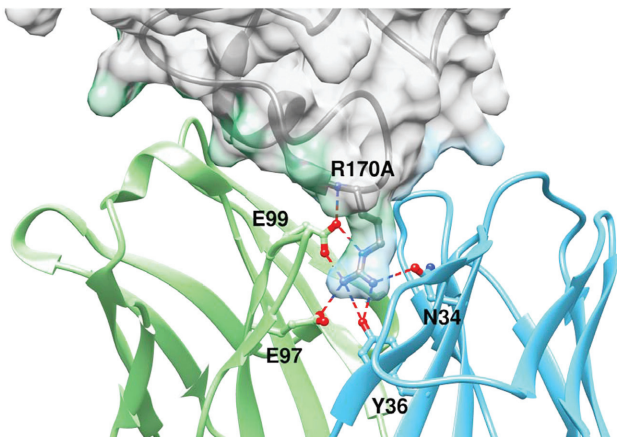

Fig. 14 Crystal structure of complement factor D bound to the engineered antibody AFD (see text for details). A protein loop of factor D (solid surface) is sandwiched between the H chain (green) and L chain (blue) of the antibody, and an Arg residue is in potential contact with two Glu residues on the H chain and Tyr and Asn residues on the L chain. The figure is based on the crystallographic data of [23].
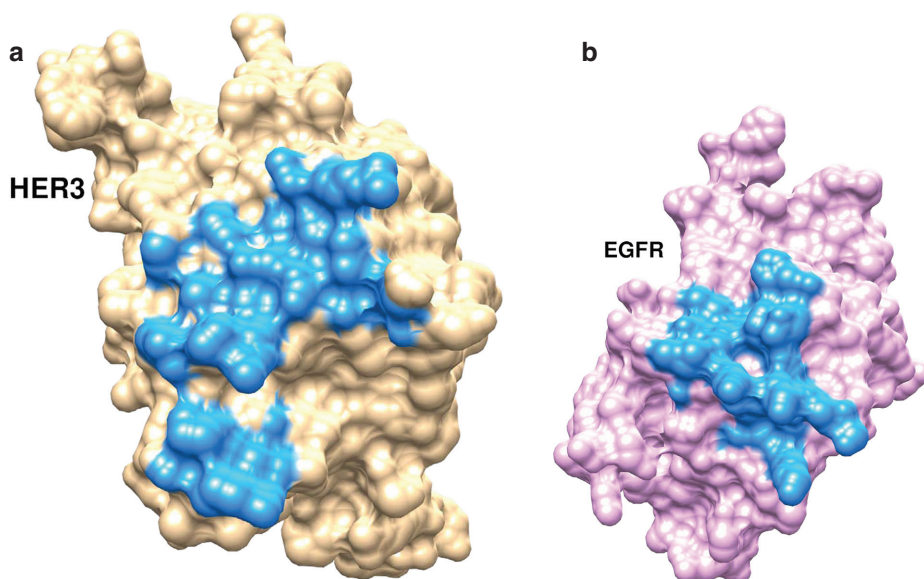
Fig. 15 Comparison of epitope-binding sites of the homologous receptors (a) HER3 and (b) EGFR. The surfaces of the two receptors bound to F_{ab} D11 (antibody not shown) are represented with contact atoms within 0.5 nm of the antibody colored blue. The figure is based on the crystallographic data of [24].

it was assumed EGFR binding resided in the H chain. Mutations were then introduced in L-chain CDRs until a F_{ab} was identified which bound the homologous HER3 receptor, without loss of EGFR binding. It is believed that binding to two distinct members of the HER receptor family will increase anti-cancer activity of the antibody. The epitopes on the two receptors do not exactly correspond (Fig. 15), but binding features are similar.
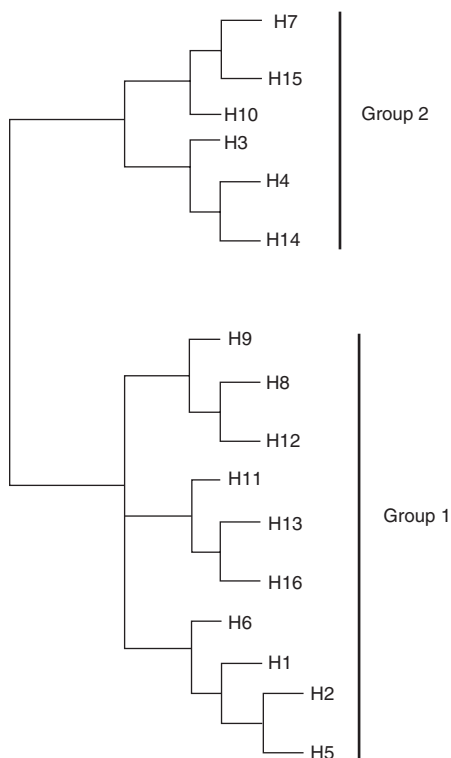

Fig. 16 Phylogenetic relationship among the influenza hemagglutinin subtypes. Adapted from [25].

Surface coverage is somewhat higher in HER3 (8.9 vs. 8.0 nm^2 in EGFR), accounted for by additional involvement of L-chain contacts, as expected.

2.4 Viral epitopes

2.4.1 Influenza A

Among the most widely studied viral antigens are the hemagglutinins of the influenza A virus, an RNA-based orthomyxovirus that causes most human influenza. The two major envelope proteins of influenza A are the glycoproteins hemagglutinin and neuraminidase, and serotypes of these proteins define the various infectious strains, H1N1, H2N2, H5N1, etc., denoting combinations of one of the 16 serotypes of hemagglutinin with one of at least seven neuraminidase variants. Hemagglutinins cluster phylogenetically into two distinct groups on the basis of their primary sequence; group 1 includes subtypes hemagglutinins H1, H2, H5, H6, H8, H9, H11, H12, H13, and H16; group 2 includes the remaining 6 subtypes (Fig. 16). Hemagglutinin is a lectin involved in binding to target cells through sialic acid residues, and so antibodies to hemagglutinin are an attractive therapeutic strategy for immunization against flu. A challenge is to develop a vaccine that neutralizes multiple viral serotypes.

One challenge in developing an effective flu vaccine, then, has been to develop an antibody that reacts against both group 1 and group 2 hemagglutinins. Ekiert et al. [26] developed an antibody, CR6261, that recognizes a highly conserved helical epitope in the hemagglutinin proximal membrane region. However, while CR6261 neutralizes most group 1 viruses, it fails to neutralizes those of group 2. The same laboratory subsequently reported another antibody, CR8020, that neutralizes group 2 [25]. CR6261 and CR8020 recognize distinct epitopes in the hemagglutinin stalk, with two overlapping amino acid residues, and binding is dominated by the F_{ab} H-chain variable V_H1-69 region (typically with extended CDR3) in each case. However, the neutralization patterns are distinct. These studies also suggest the importance of protein glycosylation in antigen escape.

Further insights into viral resistance were achieved with the production of a monoclonal human antibody, CH65, prepared by isolating rearranged H- and L-chains from single plasma cells of an individual who had been immunized with a trivalent influenza vaccine [27]. The H-chain CDR3 region of CH65 inserts into the sialic acid receptor binding pocket of hemagglutinin serotype 1 (H1). CH65 F_{ab} was co-crystallized with a hemagglutinin ectodomain from H1N1 and the structure was solved to 0.32 nm resolution. N-linked glycosylation, an important mechanism of immune evasion, was visualized at all eight potential sites on hemagglutinin. In contrast to CR6261 and CR8020, which bind to the hemagglutinin stem, three CH65 F_{ab}s bind to the globular head region of the hemagglutinin 1 trimer, which encompasses the sialic acid receptor binding pocket (Fig. 17a). All three H-chain CDRs and L-chain CDR1 and CDR3 are involved in coverage (8.58 nm^2 buried on the antibody surface, 7.48 nm^2 on hemagglutinin 1), but as usual H-chain CDR3 dominates: it inserts into the receptor pocket where 7 of its 19 residues contribute 47 % of the complete interface (Fig. 17b). Interestingly, several contacts mimic those of a sialic acid-presenting analog. Affinity maturation of CH65 involved the appearance of Asp^{26} and Arg^{29} in L-chain CDR1 as well as subtle changes in H-chain CDR3. For instance, the backbone amide of antibody Val^{106} H-bonds to the carboxyl oxygen of a hemagglutinin Val residue, and the nonpolar side chain of this Val is in van der Waals contact with hemagglutinin Trp^{153} and Leu^{194}; the amide and methyl moieties of the sialic acid acetamido group interact with the receptor in the same way. And, the sialic acid carboxylate has the same contact as does the side chain of H-chain CDR3 Asp^{107}. CH65 neutralized 30 out of 36 H1N1 strains tested. Resistance appeared to be due to a single-residue insertion of either a basic Lys or Arg residue at a single site near the rim of the sialic-acid pocket [27].

By screening plasma cells isolated from infected or immunized donors, Corti et al. [28] have made a major advance by isolating a monoclonal neutralizing antibody (FI6) that cross-reacts with all 16 serotypes of hemagglutinin. A F_{ab} fragment from this antibody was co-crystallized with both hemagglutinin H1 and H3 proteins, and the structures were solved at 0.34 nm resolution, revealing a common binding groove for the F_{ab} H-chain CDR3 on both hemagglutinins. The CDR3 loop of FI6 crosses a helix (designated helix A) in the side of the groove of the hemagglutinin H2 peptide and makes hydrophobic contacts with Leu, Tyr, Phe, and
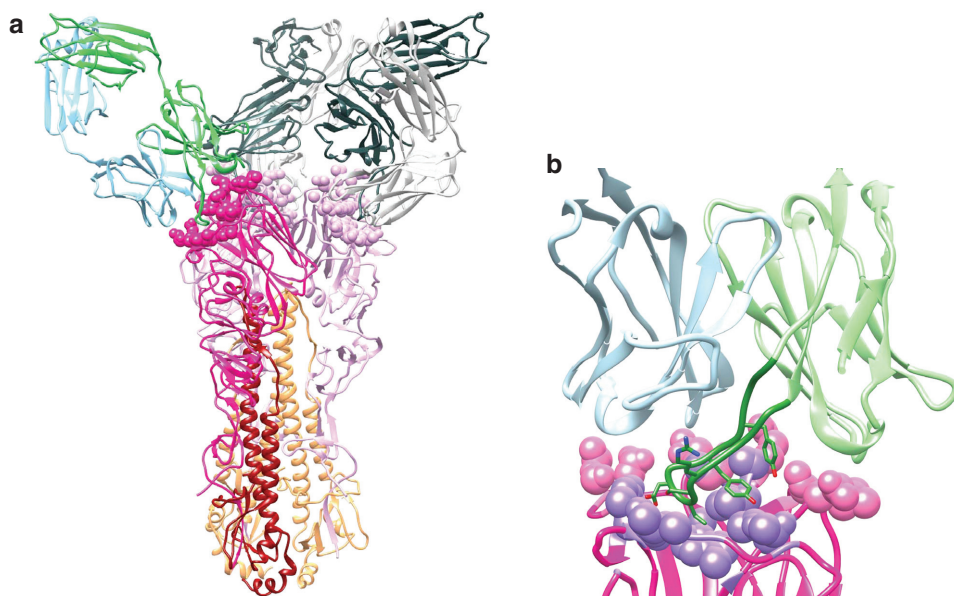
Fig. 17 F$_{ab}$ binding to influenza virus hemagglutinin 1. (a) Three molecules of the CH65 F$_{ab}$ (one copy depicted with L chain in blue and H chain in green) bind to the globular head region of the hemagglutinin 1 trimer. (b) Detail of the binding interaction shows dominance of the H-chain CDR3 (dark green) insertion into the receptor pocket. The spheres indicate the contact atoms with the F$_{ab}$ antibody in both panels; the majority are in close contact (<0.5 nm) with the H-chain CDR3 loop and these are colored violet in (b). Based on the crystallographic data of [27].

Trp residues in the F$_{ab}$ (Fig. 18). Binding is further stabilized with H-bonding contacts with a Thr and main chain carbonyl of hemagglutinin H1, and with polar interactions involving FI6 CDR3 carbonyls and Asn and Thr residues on the hemagglutinin helix. CDR3 buries about 0.75 nm^2 of the surface of both hemagglutinin H1 and hemagglutinin H3 peptides. Hydrophobic and H-bonding contacts of L-chain CDR1 on the opposite surface of the helix are also described, burying an additional 0.19 nm^2 on both peptides. Notably, Asn^{38} is glycosylated in hemagglutinin H3 but not H1; flexibility of this carbohydrate chain allows a rotation that introduces additional contacts with Asp and Asn residues in the H-chain CDR2 region. A difference in orientation
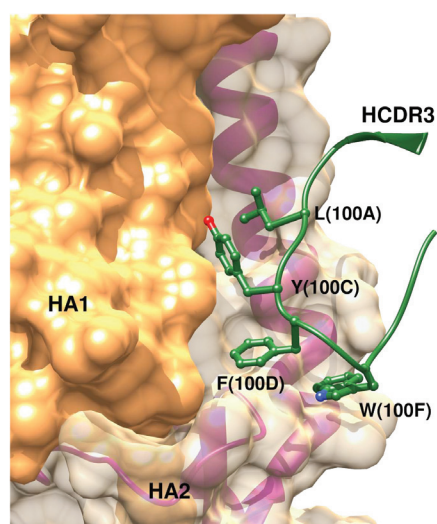

Fig. 18 The H-chain CDR3 loop (green) of the hemagglutinin-neutralizing F$_{ab}$, FI6, crosses a helix (pink) in the hemagglutinin H2 peptide and makes hydrophobic contacts through its Leu, Tyr, Phe, and Trp residues. The figure is based on the crystallographic data of [28]. Residues labeled 100A, B, C... represent insertions between residues 100 and 101 (see [28]).

of about 90° of Trp[21] between hemagglutinin H1 and H3 peptides is accommodated by local rearrangement in FI6 H-chain CDR3 that allows contact with a CDR3 Phe residue to be maintained by burying it about 0.2 nm deeper into the hemagglutinin H1 peptide. Binding to helix A and accommodation of structural differences in the region of Trp[21] are suggested to be important determinants of a pan-influenza vaccine [28].

2.4.2 HIV

Perhaps currently the most avidly studied virus from the viewpoint of structural aspects of antibody design and vaccine development is the human immunodeficiency virus, HIV. This lentivirus member of the *Retroviridae* family is the cause of acquired immunodeficiency syndrome, AIDS, and infects and kills cells of the immune system, notably CD4+ T cells, dendritic cells, and macrophages. The viral envelope consists of phospholipid derived from the plasma membrane of the host cell. It presents two viral surface proteins, gp41, a 41 kDa glycoprotein stem that anchors to the envelope, and gp120, a cap glycoprotein of 120 kDa (Fig. 19). These glycoproteins are involved in attachment and fusion with target cells; in particular, gp120 contacts the CD4 receptor. Variable loops occur in the gp120 structure, and the third of these, V3, is especially involved in the processes of attachment and fusion. Because of their exposure on the surface of the viral particle, and their role in attachment, gp120, and to a lesser extent gp41, have been attractive targets for anti-HIV antibody and vaccine development. However, these proteins have efficient mechanisms of immune evasion, especially by expressing an extensive glycan coat. Of two subtypes of HIV, HIV-1 is most virulent, most widespread, and by far the most intensively studied from an immunological perspective.

A number of neutralizing antibodies to HIV-1 gp120 have been described (reviewed in [30]) including the glycan-reactive antibodies 2G12 and PGT121-137; those preferring quaternary structure – showing a higher affinity for the entire stem-plus-cap spike structure than for the gp120 monomer – including PG9, PG16, and CH01-04; and antibodies b12, HJ16, and VRC01-03, which are directed against the region of gp120 involved in initial contact with the CD4 receptor. Wu et al. [30] note an unusually high level of somatic mutation in the variable regions of the H chain during affinity maturation of these antibodies, especially in the CD4-contact region antibodies (40–46 %). Notably, VRC01, which neutralizes about 90 % of HIV-1 isolates, shows about 70 amino acid sequence changes during the maturation process. Wu et al. then isolated thousands of VRC01-like antibodies from dozens of donors to understand how extensive maturation leads to convergent recognition of an invariant gp120 domain. The crystal structure of a new donor antibody F_{ab}, VCR-PG04, with gp120 was
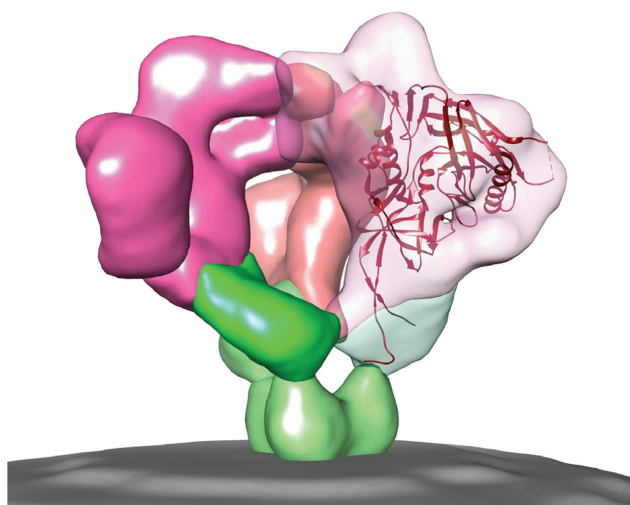

Fig. 19   Density map for the HIV Env trimer. The location of the trimeric gp120 in the density map is shown in the red colors, gp41 in green colors, and the viral membrane in gray. The CD4-bound gp120 core structure was fitted into the Env-trimer density and is shown as a dark red ribbon. The gp120 core is missing the V1, V2, and V3 variable regions. Based on data from [29].

solved at 0.20 nm resolution and compared with that of the original VCR01. A further comparison was made with another donor-derived VCR03 F_{ab}-gp120 complex at 0.19 nm. Despite a 50 % change in amino acids in the H-chain variable region between VCR-PG04 and VCR01, and a 49 % difference between VCR-PG04 and VCR03, a remarkably similar interface was observed for all three antibodies with similar pair-wise interactions and, for example, H-chain CDR2 and L-chain CDR3 Cα carbons showing root mean square deviations of only 0.5–0.14 nm after superposition of VCR-PG04 and VCR01 on gp120. Binding energy during antibody convergence correlated with the average energy of hydrophobic interactions. It can be inferred that the multiple mechanisms of immune evasion protecting the CD4 binding site necessitate an unusual degree of affinity maturation to maintain critical hydrophobic contacts in a conserved site of gp120 vulnerability.

A more potent variant of VRC01 was isolated from the same donor and designated NIH45-46. It contains a four-amino-acid insertion in the H-chain CDR3. The crystal structure of NIH45-46 F_{ab} alone at 0.26 nm and complexed with gp120 [31] revealed a similar structure to VRC01 and little change in antibody conformation upon binding (with minor changes noted in L-chain CDR1 and H-chain CDR3). The primary target is the CD4 receptor binding site and outer domain loops, with H-chain CDR3 reaching into an inner domain, and the main interactions of VRC01 preserved. A main difference is in the H-chain CDR3 insertion, with three of the four residues making contact with gp120 (Fig. 20). First, a Tyr residue in the insertion hydrogen bonds with a main chain carbonyl, decreasing chances for evasion by gp120 mutation. Electrostatic Asp–Lys and hydrogen-bonding Arg–Asn interactions are also introduced. The insertion is additionally stabilized by two intramolecular hydrogen bonds, one with H-chain CDR2. Together, these interactions lead to a larger buried surface area on gp120 that more closely resembles the CD4 receptor footprint [31], and presumably account for the increased potency of NIH45-46 compared to VRC01. Importantly, Diskin et al. [31] then noted that a Phe residue on CD4 provides critical hydrophobic interaction in "welding" gp120 to CD4. By superimposing the structures of CD4-gp120 and NIH45-46-gp120, they identified an NIH45-46 Gly residue in proximity to the critical CD4 Phe residue. By substituting the NIH45-46 Gly with Trp, they achieved an increase in neutralizing potency of 10-fold, with neutralization of a greater number of strains.
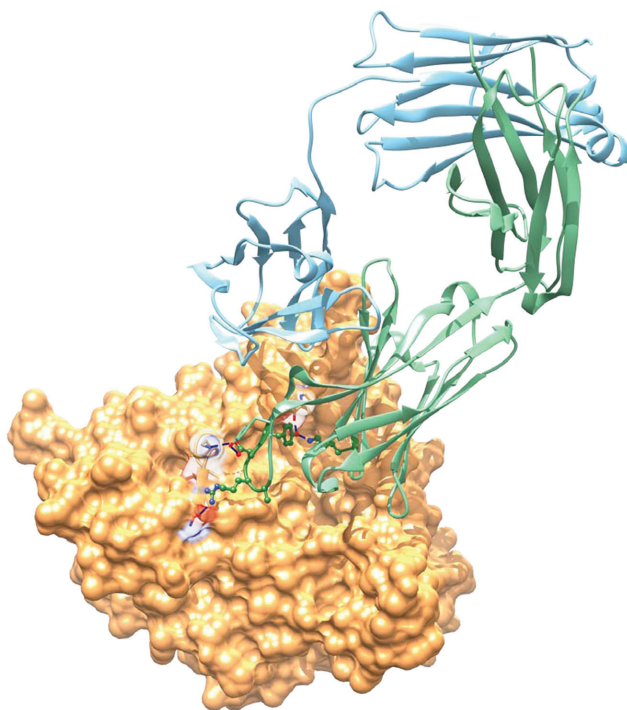

Fig. 20 The anti-HIV antibody fragment NIH45-46 F_{ab}, derived from donor serum, binds to the gp120 viral domain (surface depiction). H-chain (green) CDR3 insertion is prominent. CDR3 Tyr, Asp, and Arg residues are shown at atomic detail. They contact Ala carbonyl, Lys, and Asn residues in gp120, respectively. The figure is based on the crystallographic data of [31].

While involvement of the gp120 V3 region in the CD4 binding domain has made it a popular target, its inherent variability would seem to implicate it in viral evasion and mitigate against development of broadly specific antibodies directed to this site. However, an interesting observation was made that a single immunoglobulin gene locus (VH5-51) was responsible for encoding 18 of 51 monoclonal anti-V3 antibodies occurring naturally in patients' serum [32]. This would suggest that preferential gene usage occurs in antibody production, but the corollary – important for development of anti-V3 gp120 antibodies – is that specific, conserved epitopes must exist within this variable region that guide gene usage. Gorny et al. [32] constructed an epitopic peptide mimetic (mimotope) that was recognized by several anti-V3 antibodies coded by VH5-51, but not by anti-V3 antibodies coded by other genes. The crystal structure of this mimotope complexed with the F_{ab} of a monoclonal antibody encoded by VH5-51 showed similar interactions occurring with four different V3 peptides. The crystal structures of five different VH5-51-derived F_{ab}s with V3 peptides likewise showed good superimposition of H-chain CDR1 and CDR2, and L-chain CDR1 and CDR2. However, the often-important H-chain CDR3 did not superimpose, indicating its lack of relevance here for binding recognition, although it may contribute to the degree of neutralizing activity. Identical lengths of the H-chain CDR1, H-chain CDR2, and L-chain CDR2 point to a strong structural fit of antibody to antigen along the stable backbone conformation of VH5-51 – derived immunoglobulins. Indeed, when the various VH5-51 – derived monoclonal antibodies were complexed with different V3 peptides, binding did not produce a significant shift in stable backbone conformation.

Another strategy to overcome immune evasion is to generate antibodies to the glycan coat of gp120 itself. These exist and provide additional insight into structural interactions available at the epitope interface. A class of antibodies designated PGT do just this. PGT128 F_{ab} was crystallized first with a synthetic high-mannose glycan (Man_9), and then with a fully glycosylated gp120 engineered with a truncated V3 loop [33]. Terminal mannose residues on the arms of the synthetic construct were extensively contacted with hydrogen-bonding interactions (11 of 16 total antibody hydrogen bond contacts). One of the arms of Man_9 was bound by consecutive Asn, Trp, and Asp residues of the L-chain CDR3, while a six-amino-acid insertion in H-chain CDR2 contacted another arm (Fig. 21a), for a total of 0.39 nm^2 buried surface on the carbohydrate. Several ordered water molecules were present in the interface. In addition to binding to two mannose glycans in the engineered gp120, the antibody penetrates the glycan shield to contact a short β-strand on gp120 V3.
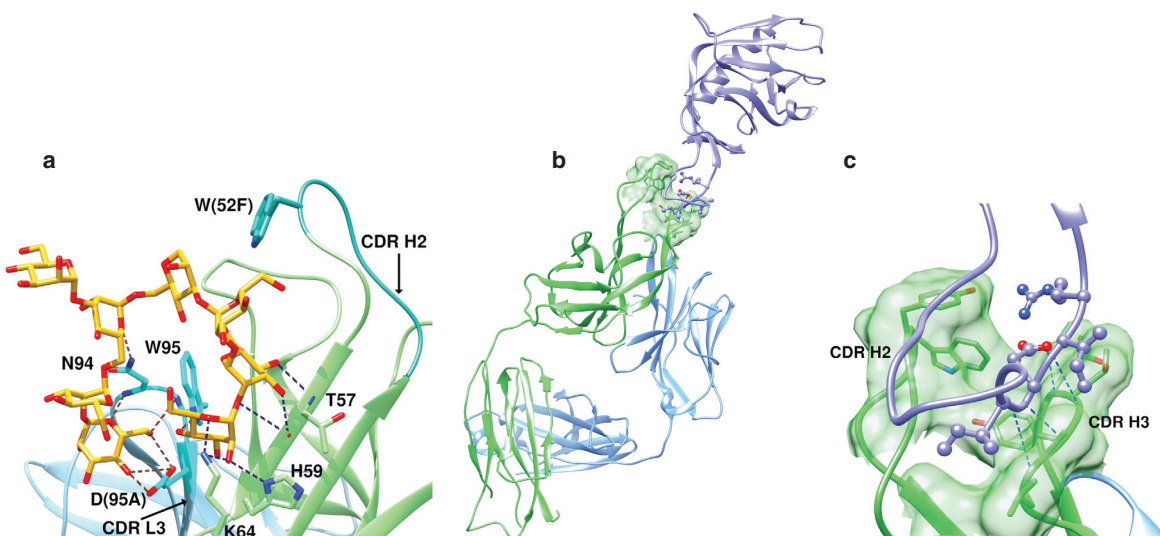

Fig. 21 A class of antibodies designated PGT have been designed to bind to the glycan coat of HIV gp120. (a) Contact of L-chain (blue) CDR3 and H-chain (green) CDR2 residues with a mannose arm (yellow, O red). (b) The antibody also penetrates the glycan shield to contact a short β-strand on gp120 region V3. C-terminal residues of V3 lie in a groove between H-chain CDR2 and CDR3, and the H-chain CDR3 contacts the apex of the V3 loop. (c) Close-up of the binding region in (b). The figure is based on the crystallographic data of [33].

C-terminal residues of V3 lie in a groove between H-chain CDR2 and CDR3, and the H-chain CDR3 contacts the apex of the V3 loop (Fig. 21b, c).

Variable regions V1 and V2 in gp120 show extensive sequence diversity that contributes to different patterns of N-linked glycosylation. Although not essential for viral entry, this 50–90 amino acid region, the most highly variable in gp120, is glycosylated on about 10 % of its residues, and is critically linked to viral evasion. Nevertheless, antibodies of the PG series (mentioned above) targeting the V1/V2 region quaternary structure have been shown to be effective, and PG9 has been found to have broad capacity, neutralizing 80 % of HIV-1 isolates. Crystal structures of PG9 F_{ab} were solved in complex with the gp120 V1/V2 domains of two HIV-1 strains, at better than 0.22 nm resolution [34]. To maintain the V1/V2 fold in native conformation, it was linked to small scaffold proteins, and chimeras that retained PG9 binding were selected for study. The V1/V2 structure exhibits a 4-β-strand structure with relatively conserved sequences in the strands, which are held together by intramolecular disulfide and hydrogen bonds. This allows a common structural fold despite overall sequence variation and variability of glycosylation. PG9 wraps a glycan moiety; its H-chain CDR3 hydrogen bonds with terminal mannose residues and extends into the protein. The L chain contributes further to a total of 11 hydrogen bonds and total surface coverage of 1.15 nm² in the PG9-glycan interface; affinity maturation appears important for glycan recognition. However, the H-chain CDR3 also extends past the glycan barrier to contact the protein β-strand structure, where specific electrostatic interactions occur between cationic residues on V1/V2 and acidic residues, notably several sulfated Tyr's on H-chain CDR3 that provide a closer length match to the side chains of V1/V2 Lys and Arg residues [34] (Fig. 22). H-chain CDR3 sequences from V1/V2-directed neutralizing antibodies isolated from several individuals showed common features of hairpin structures capable of penetrating the glycan shield and then presenting anionic moieties, including sulfated Tyr residues, for electrostatic interaction with the cationic protein surface.

2.4.3 Other viruses

Although influenza and HIV have obvious public health issues driving their study, a number of other viruses also present pressing concerns, and investigations into their structures also contribute insight into the nature of antigen–antibody interactions. The dengue virus (DENV) of the *Flaviviridae* genus is coated in a lipid bilayer that presents 180 copies of an envelope glycoprotein, E, required for receptor-mediated endocytosis of the
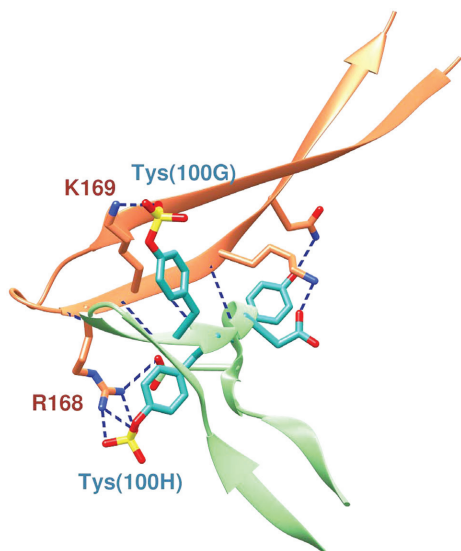

Fig. 22 PG9 F_{ab} H-chain CDR3 (green) uses sulfated Tyr residues (Tys; sulfur yellow) to length-match residues in the gp120 V1/V2 region. Contacts with Lys and Arg residues in the gp120 protein (brown) are shown. Hydrogen bonds, dashed lines. The figure is based on the crystallographic data of [34].
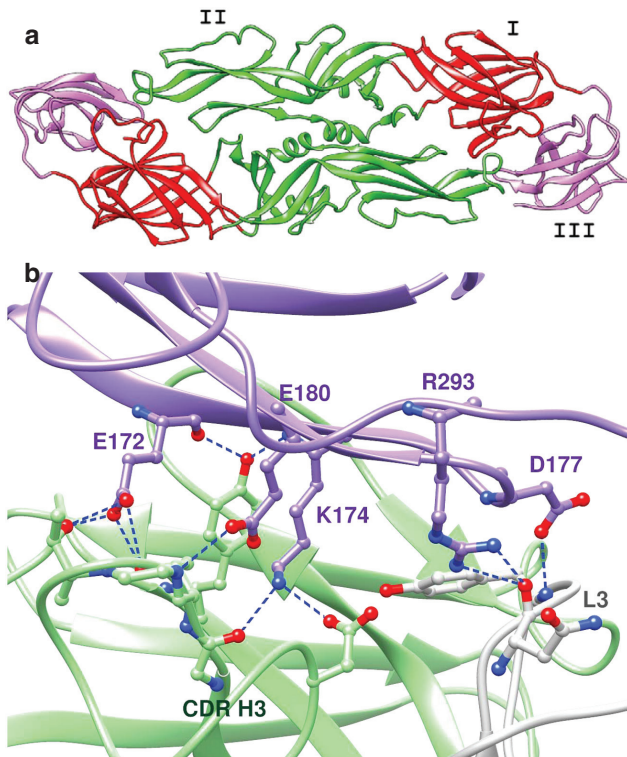
Fig. 23 The dengue virus E protein has three domains, and antigenicity appears to reside in domain I. (a) Domain structure of the DENV E protein crystallized as a dimer. Note the β-barrel in domain I. (b) Interaction surface of the 5H2 antibody in complex with an engineered DENV E domain I (purple), showing potential hydrogen-bonding interactions of H-chain (green) CDR3 with protein Glu, Arg, and Asp residues, and a salt bridge involving peptide Lys[174]. See text for details. The figure is based on the crystallographic data of [35].

virus and fusion with the endosomal membrane. Immunity against one of the four serotypes of DENV appears to increase the severity of infection with other serotypes, and a safe vaccine should therefore protect against all four serotypes. Domain I (DI) of the DENV E protein contains a 9-stranded β-barrel (Fig. 23a). A monoclonal antibody, 5H2, against DI of serotype DENV-4 has been isolated from the chimpanzee, and a 5H2 F_{ab} has been crystallized in complex with DENV-4 soluble E protein and analyzed at 0.32 nm resolution [35]. Conformational changes in E that are necessary for fusion are prevented by 5H2. As the epitope was found on DI, an engineered recombinant DI was then crystallized with 5H2 F_{ab} at 0.27 nm resolution. DI Asp, Glu, Lys, and Arg residues form hydrogen bonds with the F_{ab}, and a Lys[174] residue forms a salt bridge with an Asp residue in the H-chain CDR3 (Fig. 23b). These residues are variable across DENV serotypes, except for conserved Glu[172]. In particular, Lys[174] is replaced by Gln, Glu, or Ile in the other three DENV serotypes, and mutation to Glu in DENV-4 renders it non-binding to 5H2 [35]. However, subsequent study of a murine monoclonal antibody, 4E11, capable of neutralizing all four serotypes, gave additional insight. A single-chain variable fragment, scF_v 4E11, was crystallized with the DIII domain of all four DENV E proteins at (0.16–0.21) nm [36]. A common hydrophobic core in all E proteins was recognized by the antibody, with critical Leu and Trp residues conserved in all four E proteins, although serotype-specific contacts were of course also present in all structures. It is concluded that the murine germ line contains gene segments coding for inherent high-affinity binding to DIII.

Another recent example shows how extreme conformational flexibility in a viral structure can work in favor of antibody recognition. Human noroviruses are common causes of outbreaks of gastroenteritis, and are genetically classified into two main groups, GI and GII, together representing about 25 known genotypes [37]. For example, the well-known prototype Norwalk virus is norovirus GI.1. The single capsid protein self-assembles into a virus-like particle thought to be structurally and antigenically similar to the intact virus, and consists of shell (S) and protruding (P) domains. The P domain is responsible for attachment and includes the

determinants of strain diversity. Several structural studies have shown that it may lie against the S domain, or be raised up to about 1.5 nm off the surface, and may show rotation when raised by up to 40°. It contains two subdomains, P1 consisting of an α-helix, and P2 with six anti-parallel β-strands. Hansman et al. [37] solved the crystal structure of strain GII.10 P domain complexed with a F_{ab} of monoclonal antibody 5B18 (Fig. 24a). They note that as the resolution was 0.33 nm, water molecules were not added to the structure. The interface buried 15.0 nm^2 of protein surface (7.7 nm^2 on the P domain and 7.3 nm^2 on the F_{ab}). Interestingly, the interaction was dominated by interaction of the P1 subdomain with the κ L-chain (Fig. 24b), which contributed eight of nine hydrogen bonds, and binding was directed to two regions of negative charge on the V_L surface. 5B18 binds numerous GII genotypes, but not GI, whereas several other cross-reactive antibodies bind only GI genotypes. However, these structural studies of 5B18 F_{ab}–GII.10 P1 demonstrate an epitope in close proximity to those deduced from epitope mapping of other cross-reactive antibodies in the P1 subdomain of either GI or GII viruses [37]. While the epitope lies in an occluded region against the S domain surface, the conformational flexibility of the P1 domain with respect to orientation to the S surface may render a highly conserved epitope accessible to broad antigenic recognition.

Another example of viral conformational flexibility that fails in immune evasion is provided by one of the pneumoviruses. The human metapneumovirus (HMPV) is a member of the pneumovirus subfamily of *Paramyxoviridae* that causes lower respiratory tract infections. The paramyxovirus fusion protein, F, exists in a pre-fusion conformation and undergoes extensive refolding during fusion; the pre-fusion conformation is a trimeric head-and-stalk structure (Fig. 25a) that becomes an asymmetric 'T' shape upon fusion with the host cell (Fig. 25b). The structure of anti-HMPV F protein monoclonal antibody DS7 F_{ab} complexed with HMPV F was solved at 0.339 nm resolution [38]. The interface buries 16 nm^2 of surface and involves 22 DS7 contacts with 27 residues in HMPV F. The DS7 H-chain CDR3 inserts into a hydrophobic pocket on the F-protein, while L-chain CDR1 spans a surface region. Importantly, the epitope lies in a structural feature that is conserved in pre- and post-fusion conformations of the F-protein.

Finally, another recent study shows the importance that one critical amino acid residue can take on in viral neutralization. Ebola, an RNA virus of the *Filoviridae* family, causes a highly lethal hemorrhagic fever.
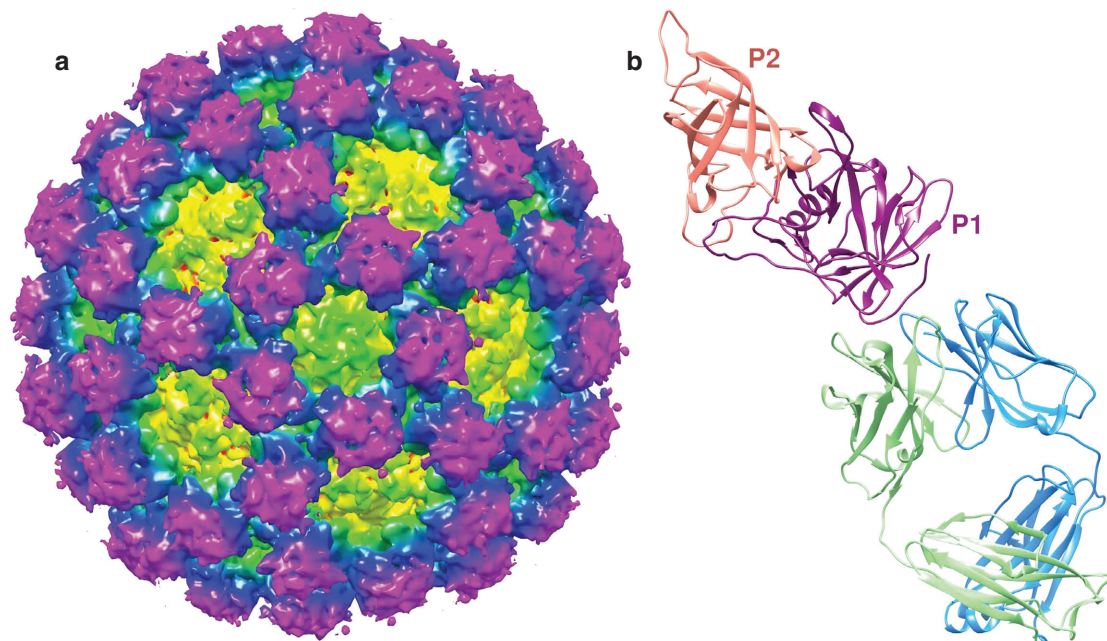

Fig. 24 Norovirus GII structure and antibody binding. (a) Cryo-EM structure of GII-10, showing the S domain structure (yellow, green) surrounded by 90 P1/P2 domain dimers (blue/pink). (b) Binding to antibody 5B18 F_{ab} is dominated by interaction of the antigen P1 subdomain (magenta) with the κ L chain (blue). See text for details. Structures are derived from [37].
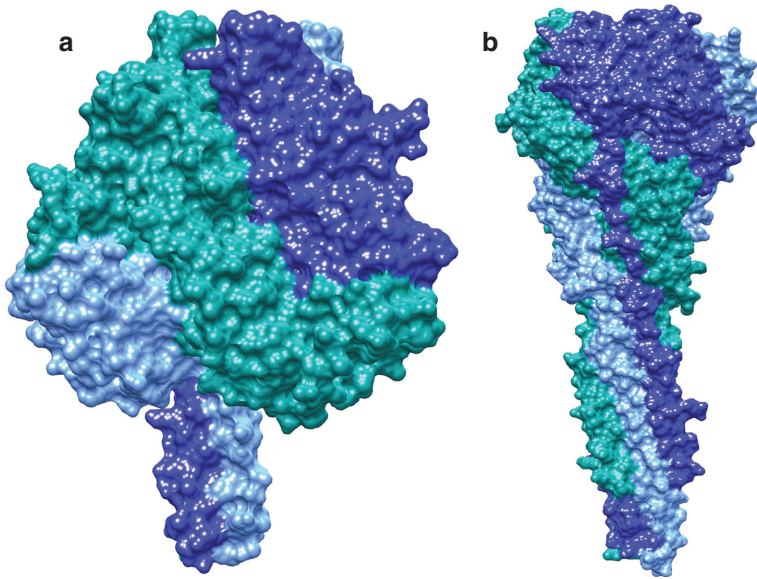
Fig. 25 Structure of the human metapneumovirus protein. The viral fusion protein F exists in a pre-fusion conformation consisting of a trimeric head-and-stalk structure (a) that undergoes extensive refolding (b) during fusion with the host cell. A F_{ab} epitope on F protein is conserved between the two conformations. Structures are based on data from [38].

Its genome codes a trimeric glycoprotein called GP that is anchored in the viral membrane and is responsible for viral attachment and entry. However, due to transcriptional editing, only about 20 % of *GP* gene product ends up as GP, the remaining 80 % representing a secreted form, sGP. A consequence is that antibodies occurring naturally during infection preferentially react with sGP, and even those that cross-react with GP are absorbed by sGP [39]. Thus, a therapeutic challenge is to find antibodies specific for viral surface GP, and a main strategy has been to study antibodies that target unique 150-amino acid mucin-like sequences in GP that extend from its top and sides and are heavily covered in both N-linked and O-linked glycans. Olal et al. [39] have solved at 0.28 nm resolution the structure of a neutralizing monoclonal antibody 14G7 F_{ab} bound to a linear epitope (residues 477–493) in this region of GP. The GP peptide adopts a tandem β-hairpin structure that extends to a depth of 1.4 nm into a pocket in the F_{ab} V_h–V_L interface (Fig. 26), and covers a total of 17 nm^2 of protein surface (approximately equal areas from peptide and F_{ab}). Seven hydrogen-bonding and 12 van der
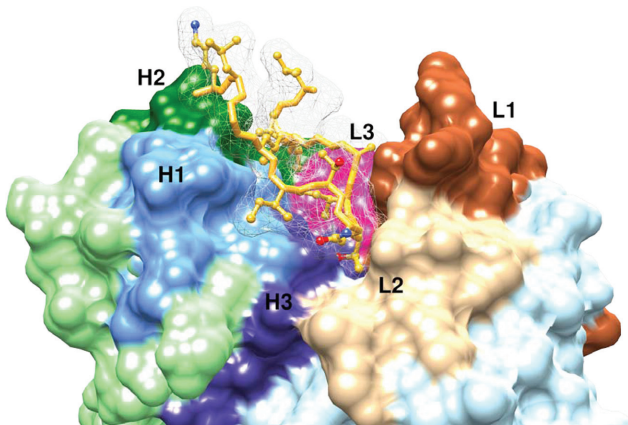

Fig. 26 Glycoprotein GP from *Filoviridae* binds a neutralizing monoclonal antibody 14G7 F_{ab} through a linear epitope (yellow chain). A tandem β-hairpin structure extends to a depth of 1.4 nm into a pocket in the F_{ab} V_h–V_L interface. Multiple contacts involve H-chain CDR1 (blue), CDR2 (green), and CDR3 (violet), as well as L-chain CDR1 (brown) and CDR3 (pink). The figure is based on the crystallographic data of [39].

Waals contacts are documented. Five of the hydrogen bonds are with H-chain residues, and all three H-chain CDRs as well as L-chain CDR1 and CDR3 are in contact with the peptide. A critical Thr residue at the apex of the hairpin hydrogen bonds to Ala and Asp residues of H-chain CDR3 and to an imidazole group of L-chain CDR1 (Fig. 26). This residue is critical for binding, and is conserved among all Ebola viruses [39].

2.5 Conclusions

Initial models of antigen–antibody interactions as lock-and-key fits have proven to be largely correct. Many detailed crystallographic studies, now generally conducted with F_{ab} complexes of antigens of clinical significance at resolution of about 0.2–0.3 nm, have borne out many of the early predictions about the chemistry of the antigen–antibody interface. Contact surfaces and buried areas are strikingly uniform, of the nature of 5–10 nm^2 of surface on both antigen peptide and F_{ab}, and often the extended, comparatively flexible H-chain CDR3 makes the major contribution. But this is not a safe generalization; studies with anti-HIV-1 gp120 in particular have highlighted the importance of other CDRs, and an example of a norovirus structure is cited above where interaction is dominated by the V_L surface. As structures are now routinely presented to at least 0.2–0.3 nm resolution, water is observed in the interface, diminishing early predictions of an important entropic contribution of solvent exclusion to binding stability; indeed, entrapped solvent may help to improve the fit of the contact surface. On the other hand, hydrophobic interactions, hydrogen bonding, salt bridges, electrostatic contacts, and van der Waals contacts are generally all involved in the binding interaction, although depending on the nature of the antigen target and, for example, strategies that viruses may have used for immune evasion, specific types of interactions may assume greater importance. Antibody flexibility can be a significant contributor to accommodation of variations in antigen peptide sequence, but often the structures of bound and uncomplexed F_{ab}s are very similar. In these cases, antibody maturation may play a more significant role in fine-tuning the binding interaction.

3 T-cell receptor (TCR) recognition

Information Box 4

In contrast to recognition of soluble antigens by antibodies in circulation or expressed on the surface of B cells, T cells only "see" the antigen when it is presented by another cell, a so-called antigen-presenting cell (APC). In this case, the antigen is generally a peptide of <20 (often 10) amino acids that has been derived from either self- or non-self-proteins, and this imparts a continuous surveillance function to the immune system. The recognition process involves a receptor on the T cell [the T-cell receptor (TCR)] interacting with the foreign peptide bound to a molecule on the surface of the APC, called a major histocompatability complex (MHC) molecule, named after a gene complex originally identified as encoding proteins that determine acceptance or rejection of foreign (transplanted) tissues or organs. [Human MHC molecules are also referred to as human leukocyte antigens (HLA).] Just as binding of a soluble antigen to an antibody on a B cell activates that cell for clonal expansion, so the binding interaction of a TCR with an MHC molecule–peptide complex activates the T cell.

The MHC molecules occur in two major varieties, called MHC class I and class II molecules. If a cell is making foreign (non-self) proteins, typically derived from an infecting bacterium or virus, proteolytic degradation products will be expressed bound to a MHC class I molecule expressed on most cells. Expression of MHC class II molecules, on the other hand, is restricted to APC cells of the immune system, such as dendritic cells and macrophages, that internalize and process peptides of extracellular origin. Just as binding of a soluble antigen to an antibody on a B cell activates that cell for clonal expansion, so too does the binding interaction of a TCR and an MHC molecule–peptide complex activate the T cell. The binding site for antigenic peptide on the MHC molecule is described structurally as a groove or cleft.

Information Box 5

The TCR bears a structural homology to the immunoglobulins. It typically consists of two disulfide-linked chains, α and β, each of which folds into two immunoglobulin-like regions, one described as constant and the other as variable. Thus, the receptor resembles an antibody F_{ab} fragment. Each variable domain contains three hypervariable regions (the complemantarity-determining regions CDR1, CDR2, and CDR3). As with the immunoglobulins, the CDRs are usually the main site of antigen contact and show sufficient structural flexibility to conform to the antigen structure. The αβTCR interaction with the MHC molecule-peptide complex also involves the presence of a co-receptor molecule; CD8 in the case of MHC class I molecules and CD4 in the case of MHC class II molecules. These are transmembrane proteins that are involved in antigen recognition and determine whether the activated T cell will respond to MHC class I presentation and become a cytotoxic (CD8+) T cell (CTL), or detect MHC class II complex and become a T-helper (CD4+) cell (Th). The CD8 and CD4 co-receptors both recruit the protein tyrosine kinase Lck to their cytoplasmic domains upon antigen recognition.

3.1 Introduction

As recently noted by Gras et al. [40], "Defining what determines an optimal peptide-HLA I determinant and TCR response is essential to improve immunotherapeutic approaches for dealing with both cancer and persistent infections that have seemingly escaped from immune surveillance". Viral infections and autoimmune diseases have been areas of intensive study of these interactions. As will be seen, binding of the α and β chains of the TCR to presented antigen follows many of the same principles as binding of the H and L chains of antibodies, with prominent involvement of the CDR loops. However, there is added complexity in that the TCR must also interact with residues in the MHC molecule. Structures for many TCR–antigen peptide/MHC class I complexes are now known; less is understood about the corresponding complexes involving MHC class II molecules, although this situation is changing rapidly.

3.2 Historical development

In contrast to antigen–antibody complexes, TCR recognizes not only a linear epitope of the peptide in the groove, but also MHC haplotype restriction demonstrates a direct requirement for MHC recognition, in addition to CD4 and CD8 co-receptor participation. However, it can be anticipated that similar considerations will apply regarding the underlying aspects of molecular recognition. Thus, while the relative simplicity of F_{ab}-peptide interactions began yielding structures in the 1980s, high-resolution crystal structures of TCR complexes were soon to follow. In 1995, Bentley et al. [41] determined the first structure of an extracellular domain of a murine TCR (the β chain) at 0.17 nm resolution after using site-directed mutagenesis to prevent glycosylation that interfered with crystallization. They showed structural homology to immunoglobulins. However, more restricted flexibility, particularly more restricted sets of conformations in the first and second hypervariable loops and close contacts between variable and constant domains, was postulated to play a role in signal transduction. Structure of the α chain followed quickly [42]. The variable (Vα) domain exists as a homodimer with the three CDR loops of each monomer directed to form a combined antigen-binding site identical to the loops of the CDRs of the antibody L chain–H chain heterodimers. And, multiple van der Waals and hydrogen-bonding contacts across the homodimer interface were analogous to those in the antibody heterodimers. A model was postulated in which a $(αβ)_2$TCR tetramer is positioned with Vα CDR3 over the N-terminal of the MHC-II-presented peptide, and Vβ CDR3 over the C-terminus. Dimerization mediated by the Vα chain was postulated to contribute to the signal-transducing activity of the TCR [42]. Subsequent studies essentially confirmed these predictions almost immediately, when Garcia et al. [43] crystallized the complete extracellular portion of a glycosylated αβTCR and deduced its structure bound to an MHC class I–peptide complex. They concluded that the TCR resembles an antibody in the variable α and β domains, although differences were found in the
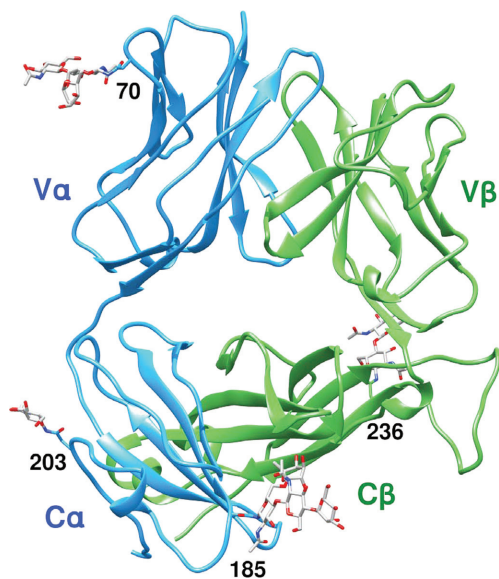
Fig. 27 The crystal structure of the complete extracellular portion of a glycosylated αβTCR [43] resembles that of an antibody in the variable α and β domains, although differences are noted in the constant domain pairing regions. In particular, of four glycosylated Asn residues (Asn[70], Asn[185], Asn[203], and Asn[236]; sugars shown as stick models), one (Asn[185]) lies in the Cα–Cβ interface.

constant domain-pairing regions, with a carbohydrate chain of one of four glycosylated Asn residues (Asn[185]) lying in the Cα–Cβ interface (Fig. 27). In contrast to the model postulated by Fields et al. [42], however, the MHC-bound peptide lay in a hydrophobic cleft between the α and β CDR3s, with Vα CDR1 and CDR2 positioned over the amino terminal of the peptide, and Vβ CDR1 and CDR2 over the carboxyl terminal.

Both MHC-peptide complex and TCR present relatively flat interfacial surfaces. While it had been postulated that the contacts might be either congruent or orthogonal, in fact, a closer look at 0.26 nm resolution at a soluble human αβTCR with the Tax peptide (LLFGYPVYV) of the human T-cell lymphotropic virus HTLV-1 bound to an MHC class I molecule (HLA-A2) [44] shows that the TCR CDR loops lie diagonally across β strands of the MHC binding groove (Fig. 28). TCR covers 9.98 nm^2 of solvent-accessible surface on the HLA-A2–Tax complex. While the β-chain CDR2 does not contact MHC in this structure, it forms a principle contact
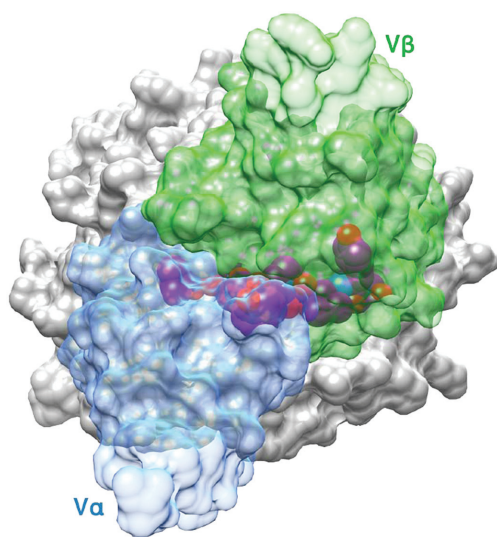

Fig. 28 Structure of the αβTCR–Tax/HLA-A2 complex. The CDR loops of the TCR (blue, green) lie diagonally across the binding groove of the MHC molecule (gray), indicated by the presence of the Tax peptide (magenta). Based on the crystallographic data of [44].

with MHC class II in a TCR–staphylococcal enterotoxin/MHC complex [45]. Although MHC class I and class II molecules are very similar in structure, this difference probably contributes to the ability of subsets of T cells to distinguish presentation by the two MHC classes of molecules [46]. More recently, it has been shown that TCR binding to peptide–MHC class I complexes can be quite flexible, with CDR loop movements >0.5 nm [47], and dominated by enthalpic (electrostatic) interactions despite unfavorable entropic changes [48], although certainly other binding strategies are used.

3.3 Current perspectives

3.3.1 General aspects

More than a decade since it was first tackled, understanding of TCR interaction with presented antigens can be understood based on many more structures. A consensus has emerged that a diagonal TCR orientation of about 60° to MHC-bound peptide axis is prototypical, in which the TCR α- and β-chain V domains lie central to the peptide, over the α2 and α1 helices of the MHC molecule, respectively [49]. CDR1 and CDR2 loops of both α and β TCR chains are usually presented to the MHC molecule, with CDR3 loops making extensive contact with the bound peptide, but also available for MHC binding. CDR1α loops tend to orient to the peptide N-terminus, and CDR1β loops to the C-terminus. A selection of structures is discussed here to give an overview of the current status of the field.

Scott et al. [50] have shown the importance of flexibility in the CDR3 loops of both the α and β chains in determining specificity and binding. TCR-A6 was studied in complex with modified Tax [LLFG(F/Y)PVYV] bound to MHC class I HLA-A2, solved at 0.23 nm resolution. The authors discuss their results in the context of competing, but not mutually exclusive, selection and induced fit models, wherein the right fit is selected from a repertoire of conformations, or conformational adjustment maximizes the fit after encounter, respectively. They note that αβTCR A6 forms one of the best TCR structural databases, with crystal structures known in multiple peptide complexes (including six Tax variants) presented by HLA-A2, and previous data had shown that significant conformational flexibility in the TCRβ CDR3 loop, whereas other CDRs, including CDR3α, position themselves independently of the presented peptide. The structures solved by Scott et al., together with molecular dynamics simulations and thermodynamic data, revealed a high degree of conformational flexibility in the CDR3β loop that allows it to sample rapidly a variety of ligand pairings across the peptide, whereas greater rigidity of the CDR3α confines it to a restricted set of peptides, and probably only those present in HLA-A2 itself [50]. One pairing is shown in Fig. 29. Thus, TCR A6 restriction to HLA-A2 is strongly influenced by CDR3α.

As well as peptide sequence, fragment length has been shown to be a determinant of successful CD8+ T-cell recognition [51]. TCR cross-reactivity is necessary to allow a large, but finite, TCR repertoire to recognize the potentially huge number of peptides that can be presented by MHC molecules. In part, this cross-reactivity is manifested by the ability of a single TCR to recognize peptide antigens bound to both MHC class I and class II molecules [52], and to enforce conformational changes in the antigenic peptide itself [53]. Differential pairing of TCRα and TCRβ chain V regions may also play a role in determining specificity, cross-reactivity and host restriction for MHC class I and MHC class II antigens [54]. While the majority of TCR diversity occurs in the V(D)J gene region of the CDR3 loop that contacts the bound peptide, contact of MHC by CDR1 and CDR2 represents regions coded by V gene segments of relatively limited diversity. Stadinsky et al. [54] isolated a number of TCRs that had an identical TCRβ chain but differed in TCRα sequence. They recognized a single peptide, but showed different degrees of MHC self-tolerance and cross-reactivity. Pairing with a different TCRα variable region changed the structure of contacts not only of TCRβ CDR1 and CDR2 with MHC, but also the conformation of TCRβ CDR3 in the peptide–MHC complex. In one example, solved at 0.27 nm resolution, changing the TCRα chain affected TCRβ CDR3 conformation by introducing a new hydrogen bond between a CDR3β Asp residue and both CDR1α Tyr and CDR2α Arg, and a new van der Waals contact between a CDR3β Trp residue and a CDR3α Ala (Fig. 30).

TCR engagement has also been explored in the context of superantigens. These are bacterial or viral antigens that bind TCR as unprocessed molecules, i.e., independently of antigen presentation, and thus activate
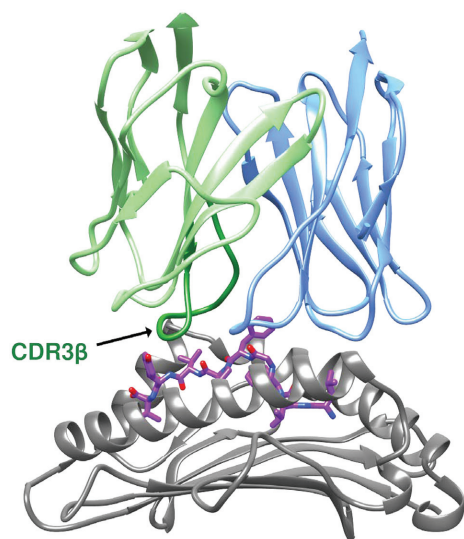
Fig. 29 Flexibilty in the CDR3β loop in a TCR–Tax/MHC complex. The crystal structure of TCR-A6 in complex with a modified Tax bound to HLA-A2 is based on the data of [50]. TCR blue/green, Tax-Y5F magenta, MHC gray. In superposition upon another structure TCR–Tax/MHC complex (yellow with peptide in orange – from O.Y. Borbulevytch and B.M. Baker, unpublished, Protein Data Bank ID 4FTV), note the flexibility of the CDR3β loop (orange arrow) in comparison to the rigidity of the adjacent CDR3α.

large numbers of T cells with resultant massive release of inflammatory cytokines. An example of the binding of a superantigen to a TCR in the absence of an MHC molecule has been provided by Fernández et al. [55], who studied the structure of the complex between a mouse TCR β chain (mVβ8.2) and Staphylococcal enterotoxin G (SEG) at 0.26 nm resolution. The buried surface in the complex was 12.9 nm^2, with equal contribution from both molecules. The major contact was with the CDR2 loop of the β chain and its flanking framework regions, binding in a cleft between two structural domains of the antigen (Fig. 31a). At least 8 hydrogen-bonding and 15 van der Waals contacts were documented involving TCR residues between His^{47} and Ser^{67}, which encompass CDR2β (Fig. 31b). Superantigens may also bridge αβTCR–MCH class II molecules with unconventional protein–protein interactions [56].
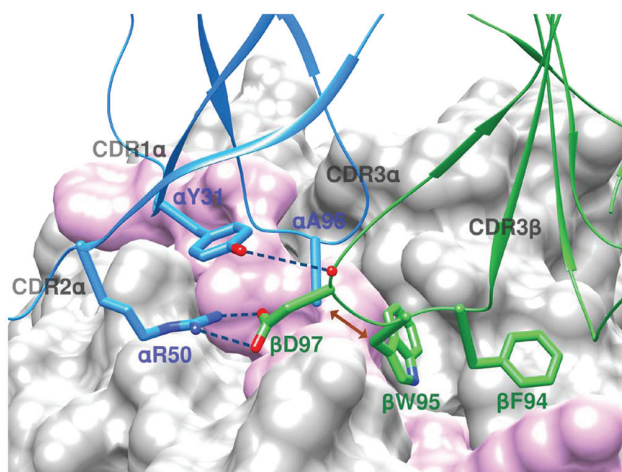

Fig. 30 Interchain α and β contacts in a TCR. Stadinsky et al. [54] isolated a number of TCRs that had an identical TCRβ chain but differed in TCRα sequence. Pairing with a different TCRα variable region changed the structure of contacts not only of TCRβ CDR1 and CDR2 with MHC, but also the conformation of TCRβ CDR3 in the peptide/MHC complex (see text for details). In the example shown here, changing the TCRα chain rotated TCRβ CDR3, introducing new hydrogen bonds between CDR3β Asp^{97} and both CDR1α Tyr^{31} and CDR1α Arg^{50} (dashed lines), and a new van der Waals contact between CDR3β Trp^{95} residue and CDR3α Ala^{96}.
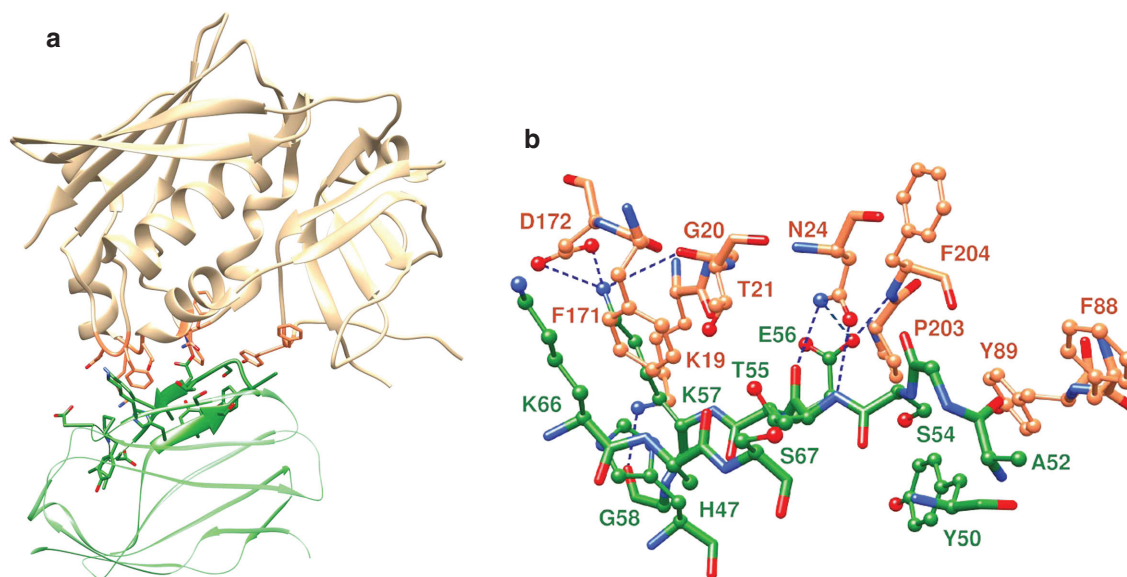
Fig. 31 Direct binding of a TCR β chain to a superantigen, staphylococcal enterotoxin G (SEG). (a) The major contact is with the CDR2 loop region of the β chain (green), binding in a cleft between two structural domains of SEG (orange). Residues involved in the interaction are shown as stick models. (b) Close-up of the interface region showing at least eight hydrogen-bonding interactions (dashed lines) and multiple van der Waals contacts involving TCR residues between His[47] and Ser[67], that encompass CDR2β. Based on the crystallographic data of Fernández et al. [55].

3.3.2 Viral antigens

Although there is a huge repertoire of TCR molecules, certain common, or "public", TCRs are linked to auto-immune diseases and viral infections in individuals that share common MHC types. One such public TCR (TK3) recognizes an 11-peptide sequence (HPVGEADYFEY, referred to as HPVG) from an Epstein–Barr virus (EBV) protein presented in an HLA-B (MHC class 1)-restricted manner. Gras et al. [57] studied allelic poly-morphism in one of the TCRβ variable gene loci as it effects immune response to EBV in individuals showing the appropriate HLA-restriction. Specifically, they determined the effects of a naturally occurring Gln-to-His mutation in the TK3 CDR2β that diminishes functional epitope recognition and immune response, as well as a non-naturally occurring Gln-to-Ala mutation. Mutants were compared to the native structure solved at 0.20 nm while the overall structure of the appropriate HLA-B molecule remained relatively rigid on complexa-tion with TK3 TCR, some residues changed conformation to accommodate CDR loops. Both mutants of TK3 were in similar orientation in the HLA-B/MHC–HPVG complex, and buried comparable surface areas, but within the contact surface the number of contacts differed significantly, and contacts with the peptide dif-fered markedly (Fig. 32). Thus, in the series TK3-Gln, TK3-His, and TK3-Ala, the number of van der Waals con-tacts decreased respectively from 172, to 155, to 129. In the same respective order, hydrogen bonds decreased as 18, 15, and 13; while salt bridges were, respectively, 3, 2, and 1. Nevertheless, the Gln-to-Ala mutant had a similar binding affinity to wild-type TK3, and the greatly diminished binding of the naturally occurring Gln-to-His mutant was attributable to a lesser accommodation by the latter of the Asp residue in the peptide. Retention of the Gln-to-His mutant as a public TCR may be due to its higher affinity for an HLA-A2-resticted HIV peptide epitope [58].

In a complement to these studies, another EBV epitope (FLRGRAYGL) bound to an HLA-B8 molecule was studied by the same group in complex with three different TCRs (designated LC13, CF34, and RL42) from various individuals showing biased CTL response to the epitope [40]. This allowed evaluation of the relative importance of TCR structure and viral epitope in determining the immune response. Energetically, interactions of TCR with the FLRGRAYGL peptide dominated binding. For instance, CDR1α, CDR2α, CDR3α, and CDR3β loops from RL42 contacted the peptide and contributed 20 % of the total buried surface area. Side chains from Phe, Ala, and Tyr in the peptide were exposed for interaction. (Similar loop involvement,
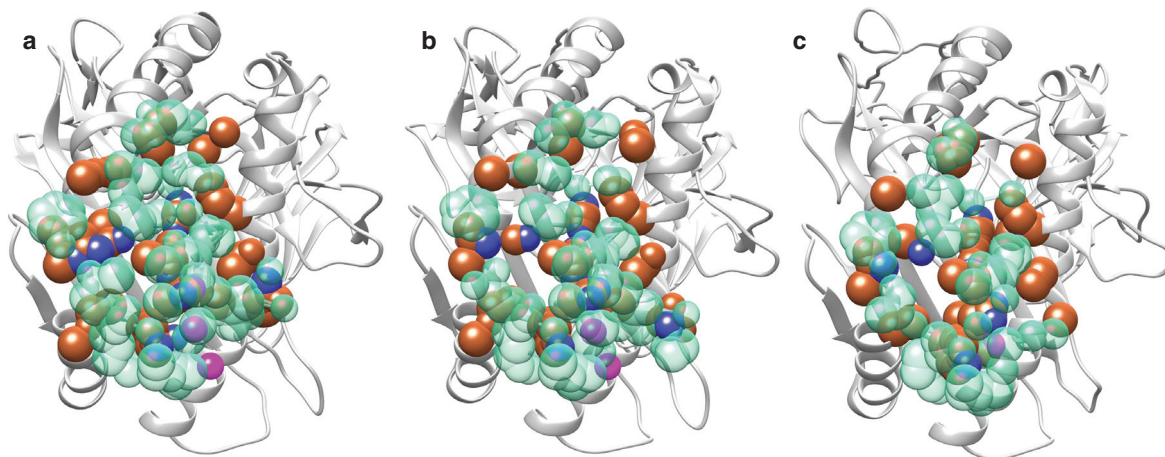
Fig. 32 Natural variations in TCR CDR2β sequence influence binding surface interactions. The TCR designated TK3 recognizes an 11-peptide sequence (HPVGEADYFEY) from an Epstein–Barr virus protein presented in an HLA-B (MHC class 1)-restricted manner. Wild-type TK3 (a) shows allelic variants with Gln[55]-to-His[55] (b) and Gln[55]-to-Ala[55] (c) substitutions. Although these bind the peptide/MHC complex in the same orientation with a similar interface, the binding contacts decrease from wild type to His[55] to Ala[55]. The peptide/MHC complex is shown in gray, and its atoms involved in contacts are coded as brown (van der Waals), blue (hydrogen bonding), and pink (salt bridges). TK3 contact atoms are shown semi-transparent in front, colored cyan. Based on the crystallographic data of Gras et al. [57].

coverage, and binding strength were observed with LC13 and CF34, which nevertheless used other residue patterns in the peptide.) By using alanine-scanning mutagenesis across the binding interface, hot spots for interaction of the TCRs with HLA-B8 were then identified, and the interactions of CDR loops of the TCRs with these hot spots were then compared in the crystal structures solved at 0.18 nm, where quite different interactions with the CDR3 loops of each of the TCRs were observed. The authors concluded that the viral antigen peptide acts as the "molecular glue" for binding, whereas nearby interactions of TCR with the MHC molecule stabilize binding and provide a fine-tuning on interactions that are responsible for discrimination of T cell repertoire selection, perhaps through affinity and/or resulting signal strength of the TCR-MHC class I interaction.

While the presented peptide is usually 10 amino acids or less, such as the EBV and Tax fragments discussed above, 5–10 % of MHC class I-bound peptides can be longer, and because of restrictions on the dimensions of the peptide-binding cleft, this can cause the peptide to bulge out of the binding pocket [59]. If the peptide is flexible, it may be flattened upon TCR engagement, but if it is rigid the bulge presents a new challenge for tight TCR docking. An example of the latter is provided by another EBV epitope, LPEPLPQGQLTAY. This epitope is recognized by a biased TCR (designated SB27) and is restricted to a particular HLA-B (HLA-B*3508). The overall structure of the SB27-HLA-B*3508–peptide complex has been solved [60] and shows engagement of the peptide by TCR with limited MHC class I contacts (Fig. 33a)]. CDR1α and CRR2α bind HLA, while CDR3α contacts both HLA-B and peptide, and CDR1β, CDR2β, and CDR3β bind to the peptide bulge (Fig. 33b). Alanine-scanning mutagenesis was used to map the energetics of the contacts [59]. Four of the six peptide residues in contact with CDR loops are critical to the interaction and strongly dominate the energetic profile. Interestingly, despite limited contacts and small energetic contributions from HLA-B*3508, TCR SB27 does not cross-react with the closely related allomorph HLA-B*3501, indicating a high level of fine-tuning of specificity.

MHC class II binding determinants are generally less well understood than those of MCH class I. HLA-D (MHC class 2)-restricted molecules have been used to study structures of complexes of TCR with influenza hemagglutinin antigen. The structure of the αβTCR HA1.7 has been solved unligated and with a 13-amino acid hemagglutinin epitope presented by HLA-DR1 and HLA-DR4 serotypes [61]. HA1.7 binds promiscuously to MHC class II molecules and recognizes a range of peptide epitope structures. In the study of Holland et al.
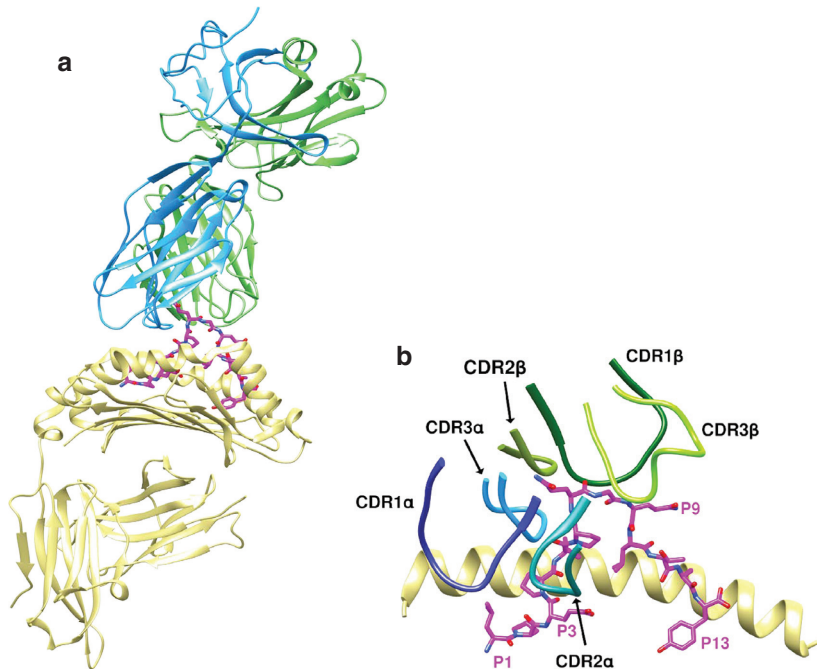
Fig. 33 Accommodation of longer peptides. While peptides presented to the TCR are often <10 amino acids, longer peptides presenting bulges, such as the EBV epitope, LPEPLPQGQLTAY, may be recognized, in this case by the TCR designated SB27. (a) The overall structure of the SB27–peptide/HLA-B*3508 complex shows engagement of the peptide by TCR with limited MHC contacts. SB27 blue/green, HLA-B*3508 yellow, peptide magenta. (b) Close-up of the binding region. CDR1α and CDR2α bind HLA, while CDR3α contacts both HLA-B and peptide, and CDR1β, CDR2β, and CDR3β bind to the peptide bulge. Based on the crystallographic data of Liu et al. [60].

[61], the bound structures were fairly rigid, with average CDR loop movements of only 0.128 and 0.152 nm when peptide was presented by HLA-DR1 and HLA-DR4, respectively (Fig. 34). Greatest flexibility occurred in the CDR1β loop, allowing two salt bridges to form between Asp^{28} of CDR1β and a Lys residue in the peptide (Fig. 35). Minimal conformational changes upon binding were confirmed with B factor heat plots. Thermodynamic studies revealed that, despite the Asp-Lys salt bridging, extremely favorable entropic contributions dominated binding, overcoming negative enthalpic contributions, in contrast to most other studies with MHC
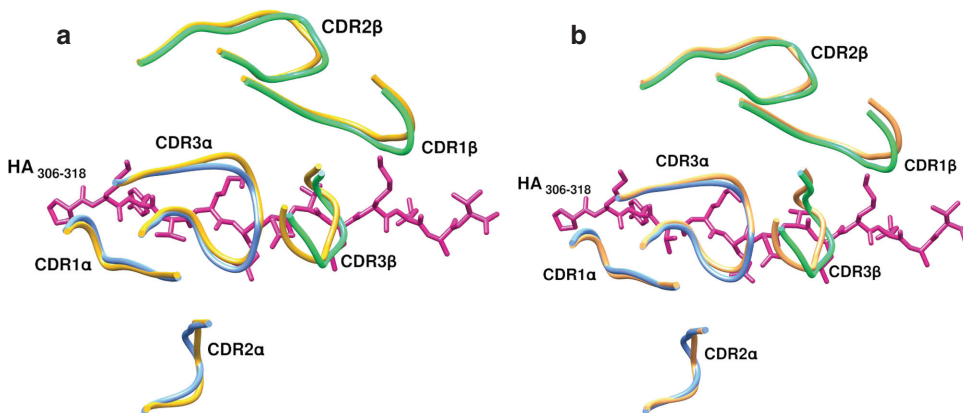

Fig. 34 Rigid binding of an αβTCR to a hemagglutinin antigen peptide. The binding loops of the unligated TCR are shown (α chain blue, β chain green) superimposed upon their structures (yellow) when bound to the peptide/MHC complex. Peptide is magenta, and MHC residues are not shown. (a) Presentation of peptide by HLA-DR1. (b) Presentation of peptide by HLA-DR4. The TCR structure is rigid, changing little upon binding and relatively independent of the MHC variant. Based on the crystallographic data of Holland et al. [61].
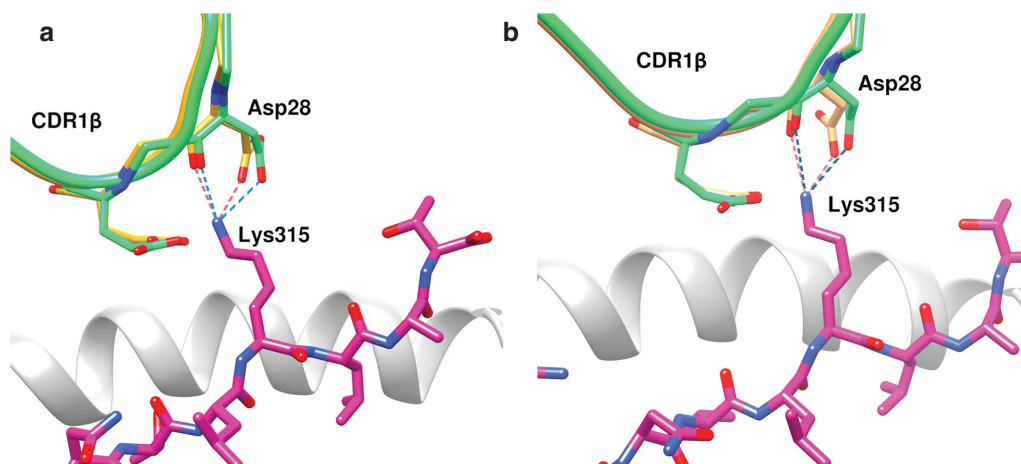
Fig. 35 Rigid binding of an αβTCR to a hemagglutinin antigen peptide. The same structures as shown in Fig. 34 are represented, showing the contact region of HLA region 306–318 with the CDR1β loop. Greatest flexibility occurs in the CDR1β loop, allowing two salt bridges to form between Asp28 of CDR1β and a Lys residue in the peptide. (a) Presentation of peptide by HLA-DR1. (b) Presentation of peptide by HLA-DR4. After [61].

class I systems. This perhaps accounts for an ability of CD4+ T-cell activation to tolerate mutations in key contact residues of the peptide.

3.3.3 Autoimmunity

From the above discussion, it will be evident that subtle interactions between TCR and MHC molecule with bound antigenic peptide determine the degree of HLA restriction, cross-reactivity, and biased TCR expression. Structural details of the interactions are important for understanding the principles underlying these phenomena, and by extension for developing therapeutic approaches when the underlying processes contribute to disease. Autoimmune diseases occur when a self-peptide elicits an immune response, and a key contributing factor may be mimicry of a foreign antigen that results in TCR cross-reactivity. In many cases, as in the examples of celiac disease and multiple sclerosis discussed below, this involves presentation by MHC class II molecules. Thus, there is an intense interest in understanding the structural dimensions of TCR recognition of self-antigens, in the hope of understanding how some autoreactive TCRs escape negative selection in the thymus. Negative selection refers to elimination of self-directed T-cell clones by apoptosis initiated through strong binding, and one goal of structural studies is to determine whether escape is due to unusual binding modes and/or weak/unstable complexes.

As an example of the approaches being taken, consider Borbulevych et al. [62], who noted that mimicry of the Tax HTLV antigen presented by MHC class I HLA-A2 to the αβTCR A6 described above could shed light on the autoimmune disease known as HTLV-associated myelopathy/tropical spastic paraparesis, wherein TCR-A6 also seems to recognize a self-peptide from the neuronal protein HuD in the context of HLA-A2. The HuD epitope LGYGFVNYI is an obvious molecular mimic of the Tax peptide LLFGYPVYV. The HuD–HLA-A2 complex at 0.165 nm resolution was a close but not perfect mimic of Tax–HLA-A2. While the backbones superimpose almost exactly, Tyr^3 of HuD was rotated 110° compared to Phe^3 of Tax, and HuD Phe^5 is rotated 120° compared to Tyr^5 of LLFGYPVYV (Fig. 36a). Thus, while the presentation of these critical residues to the TCR is different, the incoming TCR positions itself in a very similar way in the ternary complex of each structure (Fig. 36b), albeit with a 70-fold lower affinity (~140 μM). The authors suggest that this weakened binding to the peptide is not sufficiently weak to allow T-cell escape from negative selection; relaxed stringency with respect to the peptide still allows complex formation, signaling, and T-cell survival.

Perhaps one of the most important examples of autoimmunity is self-destruction of pancreatic β cells in type I diabetes. Bulek et al. [49] determined the structure at 0.26 nm resolution of an HLA-A-restricted
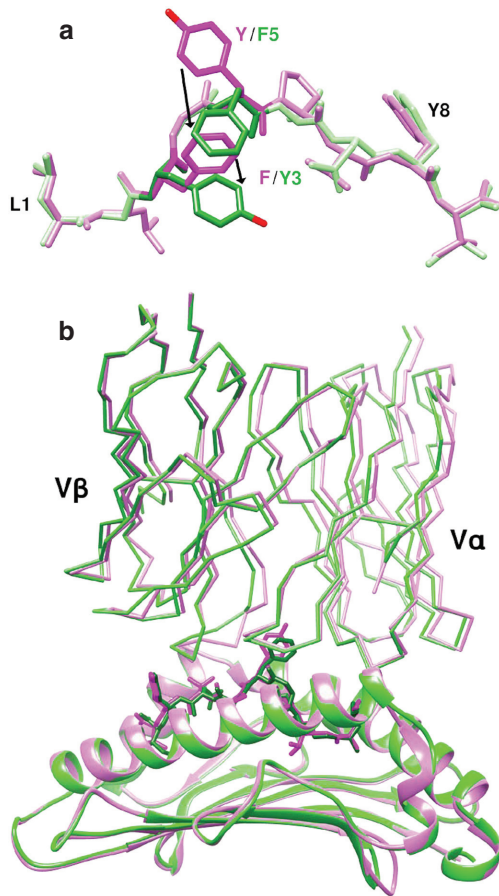
Fig. 36 Mimicry in autoimmunity. The αβTCR A6 recognizes the HTLV Tax peptide (LLFGYPVYV), but in the autoimmune disease known as HTLV-associated myelopathy/tropical spastic paraparesis, it may also recognize a self-peptide from the neuronal protein HuD (LGYGFVNYI). (a) When bound to the HLA-A2 presenting molecule, the backbones of the Tax (magenta) and HuD (green) superimpose almost exactly, except Tyr³ of HuD is rotated 110° compared to Phe³ of Tax, and Phe⁵ of HuD is rotated 120° compared to Tyr⁵ of Tax. (b) Superposition of TCR–HuD/HLA-A2 complex (green) with TCR–Tax/HLA-A2 (magenta). Although the presentation of these critical residues to the TCR is different, the incoming TCR positions itself in a very similar way in the ternary complex of each structure. From the crystallographic data of Borbulevych et al. [62].

preproinsulin peptide, ALWGPDPAAA, in ternary complex with the autoreactive CD8⁺ T cell-derived TCR denoted 1E6. Compared to high-resolution structures of 1E6 alone and HLA-A-bound peptide, neither component showed conformational changes in the ternary complex, indicating rigid lock-and-key binding. Binding orientation was typical, with the TCR α-chain positioned over the α2 helix of the MHC class I molecule and the TCR β-chain over the MHC class I α1 helix. The receptor binding angle of 58.4° with respect to the peptide-binding groove was noted to be within the range observed for other human TCR–peptide/MHC complexes (cited as 32°–80°), and the total buried surface area of the complex, 16.4 nm², was also within the known TCR-peptide-MHC range cited as 14.7–24.5 nm². CDR3α and CDR3β loops dominated binding, contacting a solvent-exposed Gly-Pro-Asp bulge of the peptide (Fig. 37). Binding was weakened due to limited MHC contacts, and was highly peptide-centered. One conclusion is that whereas autoreactive TCRs are sometimes suggested to have atypical docking modes, binding here was typical of canonical MHC class I-TCR docking.

Celiac disease is a T-cell-mediated inflammatory autoimmune disease in which consumption of wheat gluten leads to intestinal inflammation and malabsorption. It is an autoimmune disorder associated with an MHC class II phenotype expressing HLA-DQ2 and/or HLA-DQ8. T cells specific for the HLA-DQ8-restricted α-I-gliadin peptide EGSFQPSQE often occur in celiac patients. Several patient-derived DQ8-restricted α-I-gliadin T cell clones were examined for CDR3α and CDR3β sequences, and the structure of a derived αβTCR with α-I-gliadin peptide/HLA-DQ8 was determined at 0.32 nm resolution [63]. The TCR binds diagonally at 70° across
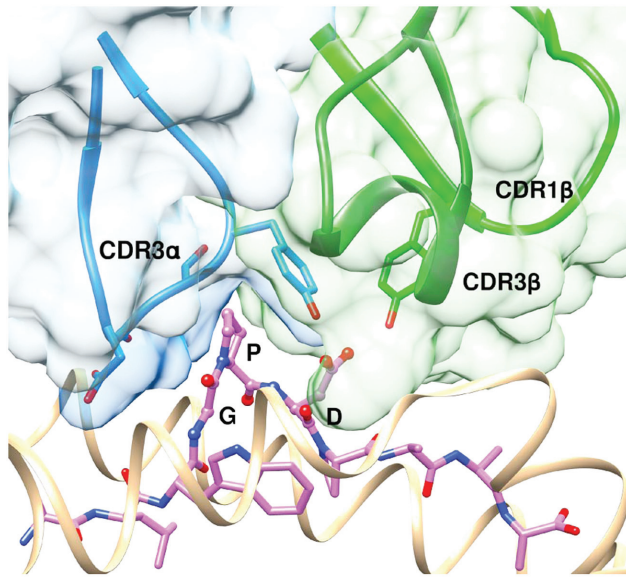
Fig. 37 Binding of a preproinsulin peptide (ALWGPDPAAA)/MHC complex to the autoreactive CD8+ T cell-derived TCR denoted 1E6. The peptide (magenta, N blue, O red) is presented by MHC (yellow). Peptide binding is dominated by the α chain CDR3 loop (blue) with two hydrogen bonds and 32 van der Waals contacts, and β chain CDR3 and CDR1 loops (green) with a total of four hydrogen bonds and 21 van der Waals contacts mainly contributed by CDR3β. Unlike some autoreactive TCRs suggested to have atypical docking modes, binding here is typical of canonical MHC class I-TCR binding. From Bulek et al. [49].

the axis of the peptide-binding cleft, as is typical, and covers a rather low interface of 8.80 nm^2 (Fig. 38a). All CDR loops contribute, but the interaction is rather rigid; the CDRβ loops do not move at all upon engagement, the CDR1α and CDR3α loops move only slightly, and only two side chains of the HLA-DQ8 reorient to avoid steric clashes with a Tyr residue in the TCR β chain. Together, CDR3α and CDR3β contribute just over 50 % to the binding surface area, in roughly equal proportions. Binding contributions from HLA-DQ8 are relatively small and dominated by van der Waals interactions. Interactions with the peptide are more extensive with contacts to residues P1-Glu, P2-Gly, P3-Ser, P5-Gln, and P8-Gln of the peptide involving contacts with CDR3α and all three CDRβ loops. Eight of ten hydrogen bonds in the interface involve the peptide, four of them
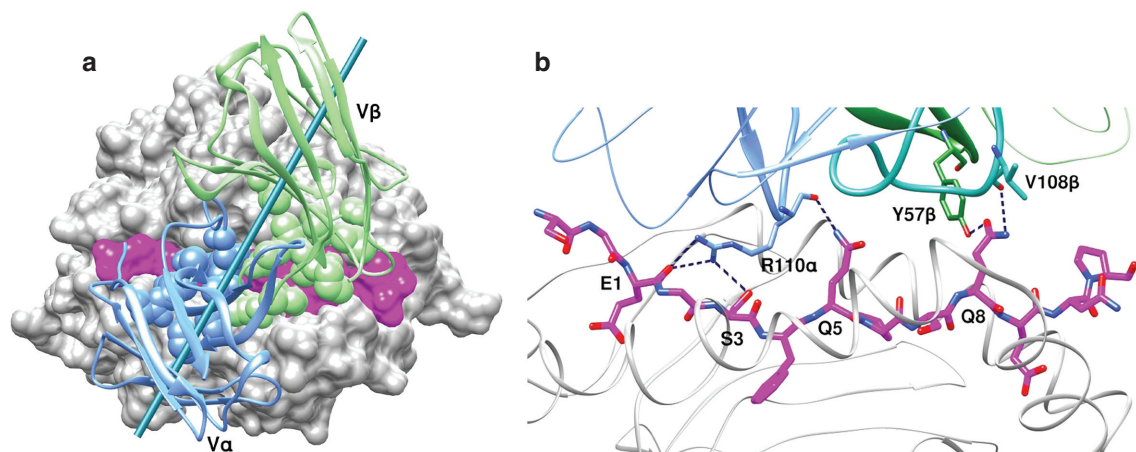

Fig. 38 Binding of αβTCR to a self-peptide derived from α-I-gliadin, presented by HLA-DQ8. (a) The TCR (Vα chain blue, Vβ chain green) binds the peptide (magenta) diagonally at 70° across the axis of the peptide-binding cleft of the HLA molecule (gray). The rod shows the major axis of inertia of the TCR variable region calculated by Chimera software. The spheres on the Vα and Vβ chains show atoms in 0.5 nm contact distance. (b) Six of eight hydrogen bonds (dashed lines) in the interface involve the peptide, four of them contributed by a critical residue, CDR3α Arg110. Based on the crystallographic data of [63].

contributed by a critical residue, CDR3α Arg[110] (Fig. 38b). The dominance of peptide recognition suggests a target for therapeutic intervention in this subset of patients.

Another MHC class II-mediated autoimmunity is multiple sclerosis, where myelin-derived peptides are presented, often with HLA-DR restriction. An autoreactive TCR repeatedly isolated from a multiple sclerosis patient, designated MS2-3C8, was crystallized with the myelin basic protein peptide FSWGAEGQRPGFG bound to the MHC class II molecule HLA-DR4 [64]. To overcome problems of very weak peptide–MHC interaction, the peptide was covalently linked to the N-terminus of the HLA chain with a 16-residue peptide linker. Under these circumstances, MS2-3C8 binds the peptide–MHC complex tightly (K_D 5.5 μM, comparable to strongly bound antimicrobial MHC class I complexes). MS2-3C8 docks centrally over the peptide at a canonical angle of 65°. Thus, weakness of binding in the native complex is due to the low peptide–MHC affinity, and not to the TCR–peptide complex interaction per se. The interpretation here is that peptide–MHC interaction weakens the interaction sufficiently to escape negative selection, but the TCR–peptide/MHC complex is strong enough due to affinity of the TCR to mount a peripheral immune response.

A good example where an unusual, non-canonical binding mode may underlie autoimmunity is provided by a TCR from another multiple sclerosis patient where crystallographic resolution at 0.255 nm showed a tilted binding orientation that limits TCR-MHC class II protein contact to a single CDR loop; the CDR3β loop is unable to contact the myelin basic protein peptide (ENPVVHFFKNIVTPR) presented by HLA-DQ [65]. A strong tilt of 14.5° compared to a TCR–hemagglutinin/MHC complex (Fig. 39a, b), together with a low crossing angle of only 40° with respect to the peptide, severely limits TCR-MHC contacts. Buried surface area was 16.5 nm², with approximately equal contributions from MHC-peptide and TCR. Nevertheless, binding is
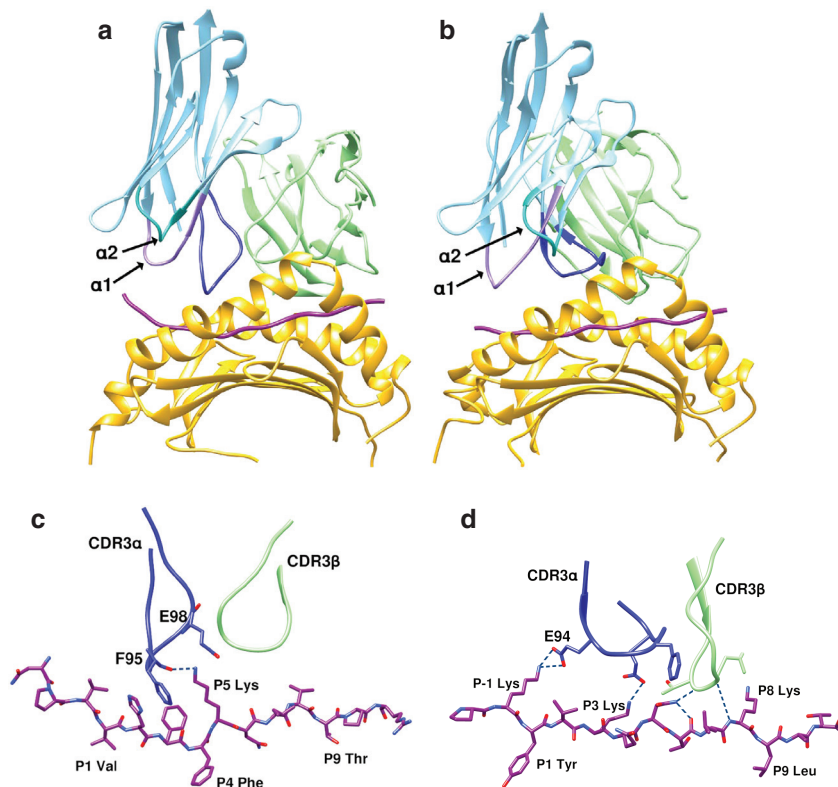

Fig. 39 An unusual binding mode in a myelin basic protein (MBP)-derived peptide complex. (a,c) Binding of TCR to the peptide (ENPVVHFFKNIVTPR) presented by HLA-DQ is compared with (b,d) a hemagglutinin complex. TCR Vα is blue, TCR Vβ green, peptide magenta, and the MHC molecule yellow. A tilt in the MBP-derived peptide complex prevents contact of the CDR3β loop with the peptide and limits TCR contacts with the MHC molecule (a) vs. (b). Binding is stabilized by contact between the MBP peptide and CDR3α, CDR3β being tilted away (c) vs. (d).

stabilized by contact between the peptide and CDR3α, CDR3β being tilted away. Again, comparison with an anti-influenza peptide is instructive (Fig. 39c, d).

The descriptions above have focused on TCR-antigen peptide-MHC molecule complexes while ignoring the presence of co-receptors that associate with the complex and are necessary for T-cell activation and signal transduction. An elegant structure that includes CD4 in a TCR–myelin basic protein-derived peptide/HLA-DR4 quaternary complex has been presented [66]. The weak binding of CD4 in the complex presented a challenge that was overcome by site-directed mutagenesis and in vitro evolution by yeast surface display to engineer higher-affinity binding. The resultant mutant bound HLA-DR with a $K_D = 8.8\ \mu M$, and spontaneously crystallized from an equimolar solution of the MS2-3C8 TCR mentioned above and the peptide-MHC complex; data were collected at 0.40 nm resolution. Despite this low resolution, electron density could be assigned throughout the complete structure and unambiguous assignment of the complete structure was possible; carbohydrates attached to the two N-linked Asn glycosylation sites in CD4 could also be assigned. The solved structure accommodates glycosylation on the CD4 Asn residues (Fig. 40). CD4 and TCR are not in contact, but form an arched 'V' in which both contact MHC–peptide complex, TCR at the peptide-binding groove and CD4 through its distal D1 domain reaching the MHCβ chain. This explains how CD4 and TCR can simultaneously yet independently bind MHC. TCR and CD4 are tilted about 60° and 70°, respectively, from the T-cell surface, and the MHC–CD4 arch angle is about 50° at the apex; glycosylation outside the arch may facilitate the tilt.

3.3.4 Natural killer T (NKT) cell receptors

Natural killer T (NKT) cells recognize lipid antigens presented to the receptor by an MHC class I-like molecule, CD1d. In contrast to the classical MHC peptide binding molecules, they come in two varieties, type I that prototypically recognize α-galactosylceramide (α-GalCer), and type II that recognize a broader range of glycolipids and sulfatides. Type I NKT cells express an invariant α-chain variable region–joining region ($V_\alpha 14$-$J_\alpha 18$ in mice, $V_\alpha 24$-$J_\alpha 18$ in humans); $J_\alpha 18^{-/-}$ mice have been used as a model of NKT cell deficiency. Type II NKT cells express a more diverse TCR repertoire. In comparison to CD4+/CD8+ cell TCRs binding to MHC class I/class
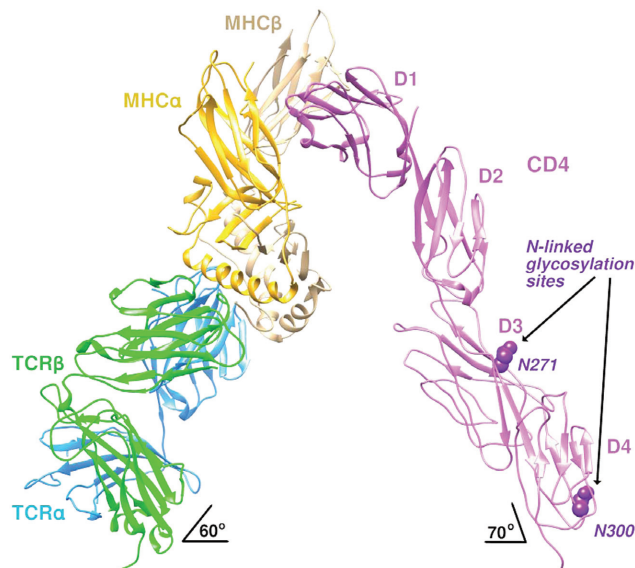

Fig. 40 Structure of a CD4–TCR–MBP peptide/HLA-DR4 quaternary complex. CD4 and TCR are not in contact, but form an arched 'V' in which both contact the peptide/MHC, TCR at the peptide-binding groove and CD4 through its distal D1 domain reaching the MHCβ chain. Glycosylation sites are present on Asn^{271} and Asn^{300} of CD4, shown as spheres. Coloring is TCR α chain blue, TCR β chain green, MHC α chain yellow, MHC β chain brown, CDR magenta. The peptide can be seen as a brown string in MHC binding groove at the MHC-TCR interface. From [66].
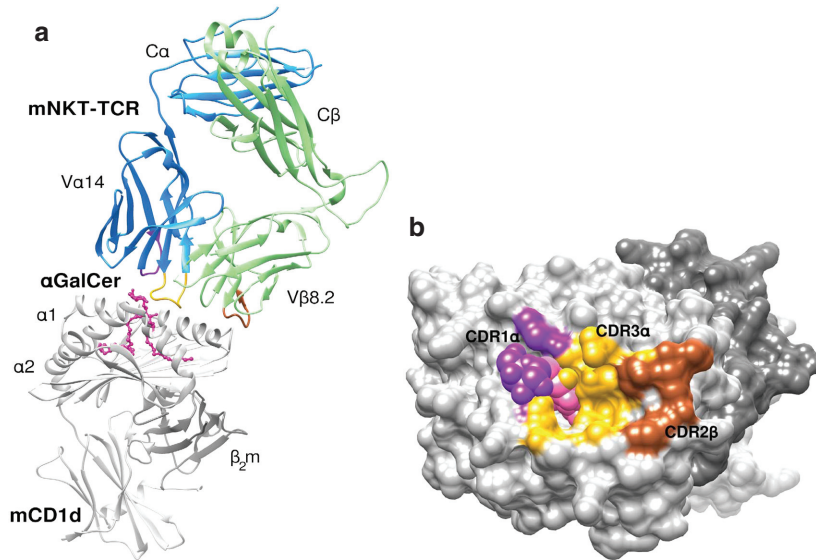
Fig. 41 Structure of a mouse αβTCR with an α-GalCer/CD1d complex. (a) CD1d is gray and the ceramide is shown as a pink chain. The TCR α and β chains are in blue and green, respectively, with the loops color-coded as CDRα1 purple, CDRα3 yellow, and CDRβ2 brown. (b) Footprint of the NKT TCR on the surface of α-GalCer/CD1d, with contact points colored as in (a). Based on crystallographic data from Pellicci et al. [69].

II molecules, NKT cell type I TCR–antigen/CD1d structures reveal a conserved docking mode, binding being nearly parallel to, and tilted with respect to, the antigen-binding cleft. In this common structure, the invariant NKT cell TCR α-chain, and in particular the CDR3α loop, dominates the interaction with CD1d [67], and the receptor binds in a fairly rigid pattern-recognition mode to CD1d [68]. A mouse NKT V_α14-TCR in a complex with α-GalCer/CD1d (Fig. 41) is typical [69].

The role played by variability in the NKT cell CDR3β loop is illustrated by recognition of the CD1d-bound self-antigen, lysophsphatidyl choline (LPC), by a specific NKT cell. LPC binding to human CD1d is relatively weak ($K_D \sim 2.1$ μM, compared to ~0.11 μM observed for α-GalCer) [70]. Engagement of the NKT TCR by the CD1d-bound LPC, observed at 0.305 nm resolution, involves significant movement of the CD1d α1 helix and a 0.75 nm displacement of the LPC head group (Fig. 42). The binding interface of 8.5 nm² is typical of NKT
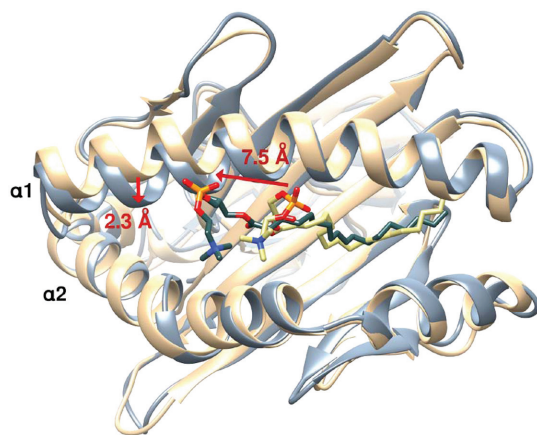

Fig. 42 Structural changes in the lysophosphatidyl choline/CD1d complex upon engagement of NKT TCR. In the unligated structure, the phosphatide (yellow main chain) lies in the α1/α2 groove of the CD1d (brown). Upon engagement of the NKT TCR (TCR residues not shown), there is inward movement of the CD1d α1 helix (gray) to resemble more its structure in α-GalCer/CD1d, and a 0.75 nm displacement of the LPC head group green main chain. Based on [70].

TCR–CD1d binding. The positioning of the CDR3α loop is conserved compared to that in the same NKT TCR bound to α-GalCer–CD1d; of the 31 contacts the loop makes with CD1d–LPC complex, 20 are identical to those in the CD1d/α-GalCer complex [70]. However, unlike the α-GalCer complex, CDR3α does not contact the ligand (LPC), essentially being blocked from doing so by CDR1α and CDR2α (Fig. 43a). A CDR1α Ser residue makes both van der Waals and hydrogen-bonding contacts with the LPC phosphate group with Pro and Phe residues also contacting the LPC head, and a van der Waals interaction of a CDR2α Phe occurs with a methyl group of the choline quaternary ammonium. CDR3β is locked in an extensive water-mediated hydrogen bonding structure with CDR3α and CD1d (Fig. 43b). One might suggest that the disposition of the CDR3 loops, and particularly exclusion of CDR3α from LPC contact, is a major factor determining specificity of the receptor for LPC.

A V-J variant NKT type I TCR α chain (V_α10-TCR) has been described that forms type I TCR–α-GalCer/CD1d complexes but has an expanded recognition capability that includes α-glucuronic acid–containing glycolipids and α-glucosylceramide. The structures of the V_α10 TCR–α-GlcCer/CD1d and type I NKT cell-TCR–α-GalCer/CD1d ternary complexes were compared at 0.22 nm resolution [67]. V_α10 TCR docked to α-GlcCer/CD1d in a similar mode to type I NKT cell-TCR–α-GalCer/CD1d binding, spanning both α1 and α2 helices of CD1d (analogous to the MHC helices) (Fig. 44). V_α10 TCR buries 9.10 nm² of the interface, slightly more than the typical 7.6–8.6 nm² of other NKT TCRs. However, whereas interfacial contact of type I NKT cell TCR with presented α-GalCer typically involves only CDR1α, CDR3α, and CDR2β, contact in the V_α10 TCR–α-GlcCer/CD1d complex involves all six CDRs (Fig. 45). In the V_α10 TCR complex, Thr and Asn residues of CDRα1 made van der Waals contacts with Val and Asp residues in CD1d, whereas in type I prototype complexes CDRα1 typically contacts only α-GalCer. A bulky Phe side chain in V_α10 TCR CDR3α inserts into the cleft between the CD1d α1 and α2 helices, making van der Waals contacts with two residues of each, replacing a network of polar interactions used by CDR3α in the prototype NKT TCR–CD1d interface. Notable differences are also seen in other loops. Interaction with the α-GlcCer moiety was exclusively through the V_α10 TCR α chain, with CDR3α lying directly over the sugar (Fig. 46).

Binding of the more diversified type II NKT TCR to antigen is typified with CD1d presentation of the sulfatide, sulfated β-GalCer from neuronal tissue. In general, type I NKT cells do not bind this sulfatide, while type II NKT cells do not recognize α-GlcCer. To attempt to understand this specificity, the TCR from one of the first characterized type II NKT cell lines, XV19, was studied in complex with β-GalCer sulfatide/CD1d at 0.31 nm resolution [68]. While the overall structure of the XV19–sulfatide/CD1d complex (Fig. 47a) resembles that of a type I NKT TCR in complex with α-GalCer/CD1d, significant differences occur in the interfacial region. First, XV19 binds nearly orthogonally to the antigen-binding cleft (100°), whereas the type I TCR binds nearly
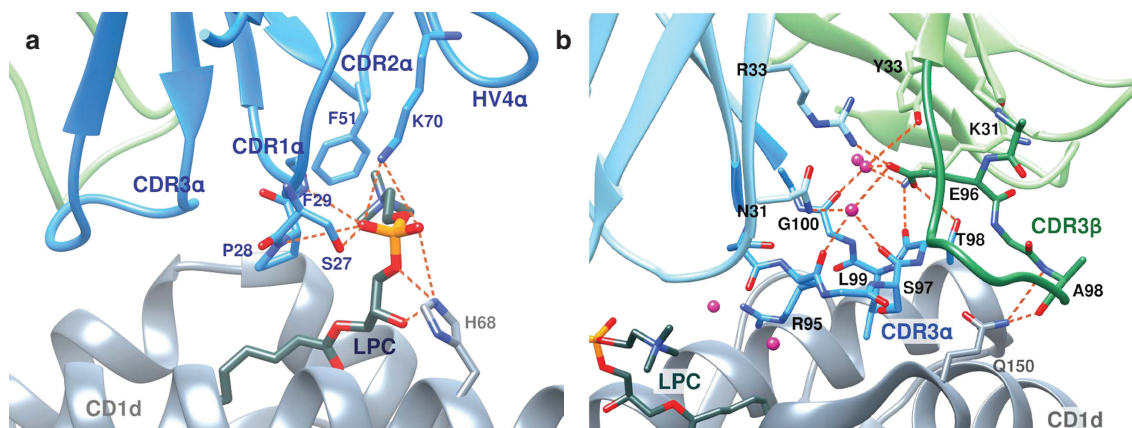

Fig. 43 Details of the binding interface of the NKT TCR–LPC/CD1d complex shown in Fig. 42. (a) TCR CDR3α does not contact the ligand (LPC), essentially being blocked from doing so by CDR1α and CDR2α. (b) CDR3β is locked in an extensive water-mediated hydrogen-bonding structure with CDR3α and CD1d. Dashed lines show hydrogen bonds. TCR α chain blue, TCR β chain green, CD1d gray. Water molecules in (b) are pink. Based on [70].
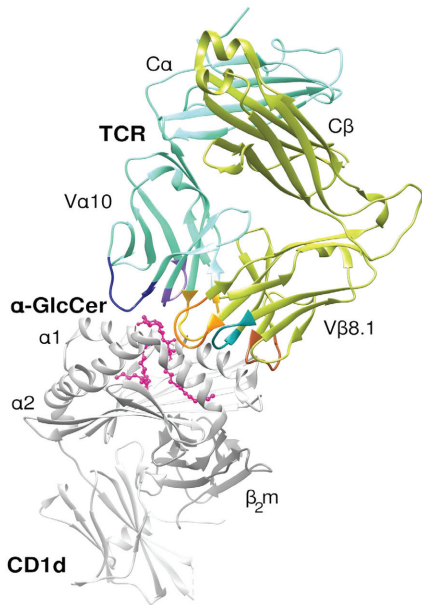
Fig. 44 Structure of an NKT $V_\alpha 10$ TCR docked to α-GlcCer/CD1d. The receptor binds across both helices of CD1d in a manner very similar to binding in the α-GalCer/CD1d shown in Fig. 41a. The ceramide skeleton is shown in pink, and the CDR loops are color-coded as follows: CDR1α purple, CDR2α blue, CDR3α yellow, CDR1β cyan, CDR2β brown, CDR3β orange. Based on the crystallographic data of Uldrich et al. [67].

parallel (20°). There were almost no common interacting CD1d residues in the two complexes, and whereas type I binding used mainly CDR3α in binding, with contribution from CDR1α and CDR2β, all six loops made contact in the XV19 complex (Fig. 47b) [68], as noted above to occur in the type I $V_\alpha 10$ TCR–α-GlcCer/CD1d complex [67]. The sulfated head group showed little movement on engagement of XV19, in contrast to flattening of the glycosyl head groups of β-linked TCR type I ligands, and the sulfatide was contacted only by the CDR3β loop. Thus, quite different binding geometries of type I and II NKT cell TCRs with CD1d may underlie the different antigen specificities of these classes, but more structures are needed to establish details.
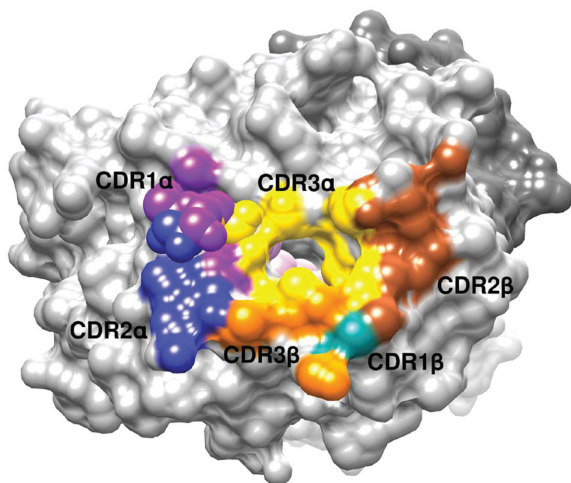

Fig. 45 Contact in the $V_\alpha 10$ TCR–α-GlcCer/CD1d complex involves all six CDRs. The footprint of the NKT TCR on the surface of α-GlcCer/CD1d is shown, with contact points (colored) with the TCR CDR loops indicated. Compare with Fig. 41b wherein the corresponding α-GalCer/CD1d complex only loops CDR1α, CDR3α, and CDR2β contact the surface. Based on the crystallographic data of Uldrich et al. [67].
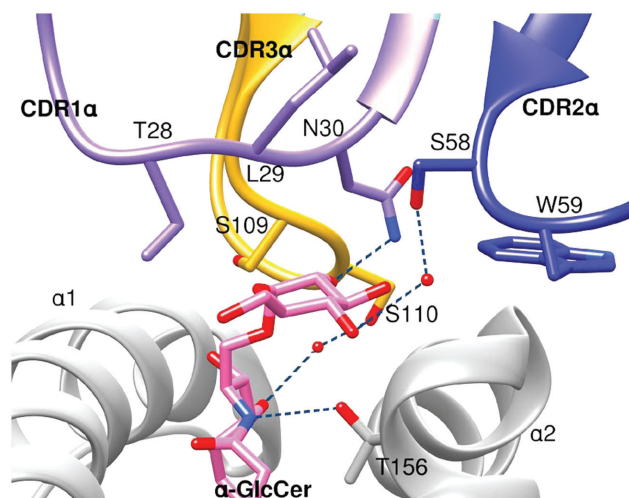
Fig. 46 Contact with the ceramide in the V$_\alpha$10 TCR–α-GlcCer/CD1d complex. Contact is exclusively through the TCR α chain, with CDR3α lying directly over the sugar. Based on the crystallographic data of Uldrich et al. [67]. The TCR α chain CDRs (purple, blue, yellow), α-GlcCer (pink with O red and N blue), and the CD1d α helices (gray) are indicated. From Uldrich et al. [67].

3.3.5 $\gamma\delta$TCR structures

The $\gamma\delta$TCR, composed of γ and δ chains, is much less common than the $\alpha\beta$TCR and occurs mainly in the gut mucosa. Although these receptors are important in defense against intestinal pathogens and malignant-cell surveillance, little is known about their modes of antigen recognition. Few structural studies have addressed $\gamma\delta$TCR antigen binding, and although these receptors may recognize antigen without MHC presentation, identification of $\gamma\delta$TCR-ligand pairs have been difficult to identify. Indeed, the major ligands for $\gamma\delta$TCR so far identified are MHC class I homologous proteins and CD1d without bound glycolipid. The crystal structure of a mouse $\gamma\delta$TCR, designated G8, bound to a non-classical MHC class I ligand, H-2T22, [71] has shown structural homology between the $\gamma\delta$TCR and $\alpha\beta$TCR, but with a different binding geometry, and strong domination of the CDR3δ loop in binding (the loop protrudes more than 1.0 nm from the receptor surface), but more structures are needed before generalizations can be made. Xu et al. [72] have studied the interaction at 0.30 nm resolution of a MHC class I-homologous (MIC) self-antigen that does not associate with antigenic peptides or other ligands, with a single-chain variable fragment (scFv) constructed from the γ and δ chains of a $\gamma\delta$TCR designated Vδ1. Interestingly, in marked distinction from the G8–H-2T22 with CDR3δ-dominated binding, the Vδ1 scFv–MIC uses CDR1δ and CDR2δ, but appears not to involve CDR3δ significantly. Rather than extending, the CDR3δ loop folds over to engage in several non-specific (mainly hydrophobic) contacts. The relatively low resolution and some regions of poor electron density precluded a detailed depiction, but the orientation of the CDR loops of the Vδ1 construct can be seen (Fig. 48).

3.3.6 Metal ion recognition

A number of metal ions cause immune sensitization in humans [73], and nickel allergy is a common and classic example. It is generally thought that the metal ion is bound to a self-peptide presented by an MHC molecule. The CD4$^+$ T cell, ANi2.3, contains a Vβ17 region that is over-represented in patients suffering from severe Ni contact hypersensitivity [74]. ANi2.3 $\alpha\beta$TCR reacts with Ni^{2+} in the context of an unknown self-peptide bound to an MHC class II molecule, HLA-DR52c. Random screening of nine-amino acid-long peptide libraries identified sequences that caused the ANi2.3 TCR to bind to HLA-DR52c in the absence of Ni$^2\beta$, and these peptides had in common a Lys residue at position 7, with an orientation that exposed the Lys ε-amino group [74]. It is proposed that the positive charge of the ε-amino group serves as a mimotope of the Ni^{2+}-bound unknown native present-
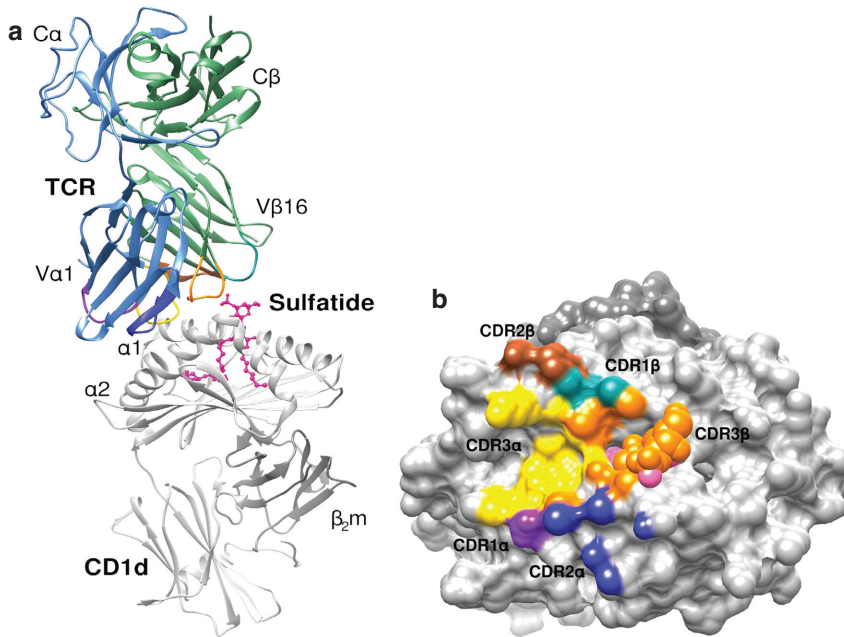
Fig. 47 Structure of a type II NKT TCR in complex with β-GalCer sulfatide/CD1d. The TCR is derived from the cell line XV19. (a) While the overall structure of the XV19 TCR−sulfatide/CD1d complex resembles that of a type I NKT TCR in complex with α-GalCer/CD1d (Fig. 41a), notable differences occur in the interfacial region. In particular, (b) all six TCR CDR loops make contact with the sulfatide/CD1d complex (compare Fig. 41b). Structures are based on the crystallographic data of Patel et al. [68].

ing peptide. The structures of two of these peptides in an ANi2.3 TCR−HLA-DR52c complex solved at 0.33 nm resolution showed conventional interaction with the CDR1 and CDR2 loops docking to the MHC helices in the conventional diagonal binding mode and the CDR3 loops centered over the peptide. The CDR3β loop extends into the peptide binding groove and contacts three peptide amino acids, while a critical salt bridge is formed between an Asp carboxylate at the tip of CDR3β and $Lys^7 \varepsilon$-NH_2 (Fig. 49). The peptide Lys^7 also forms a hydrogen bond with an HLA-DR52c helix, and it is proposed that contact between metal ion bound to a native peptide and this HLA-DR52c β1 helix may be a common site of metal cation binding in metal allergies.

3.4 Conclusions

TCR binding to presented antigen with the common α and β chains can be visualized in analogy to the H chain and L chain of immunoglobulin F_{ab}−antigen binding, with the added complexity that now the recognition
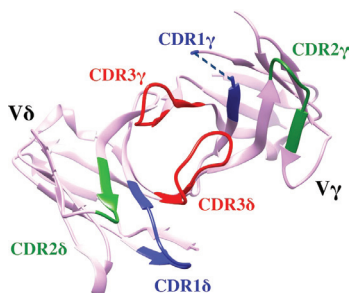

Fig. 48 Structure of a γδTCR bound to an MHC class I-homologous (MIC) self-antigen. Due to low resolution of the structure, only the orientation of the variable CDR loops is shown. CDR1δ and CDR2δ appear to dominate binding while CDR3δ appears to fold back and engage in hydrophobic interactions [72]. Few structures of γδTCR complexes are known.
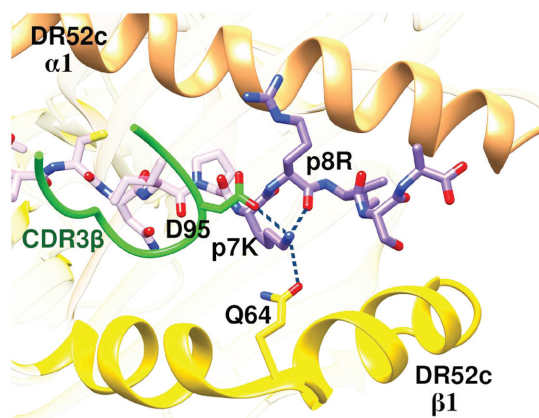
Fig. 49 Structure of an αβTCR bound to a HLA-DR52c complex of a potential Ni^{2+}-binding peptide mimotope. The peptide-binding groove of HLA-DR52c (α chain brown, β chain yellow) accommodates the peptide (purple chain), and TCR CDR3β (green) extends into the groove. The peptide has a Lys residue in position 7 (p7K) whose ε-NH_2 group forms a salt bridge with an Asp-residue carboxylate (D95) at the tip of CDR3β. In the unidentified endogenous peptide associated with nickel allergy, a bound Ni^{2+} ion may replace the ε-NH_2 group. From Yin et al. [74].

surface includes the antigen-presenting (typically MHC class I or II) protein as well as a small peptide fragment. As in antigen–antibody binding, the six CDR loops are generally involved in TCR binding, to varying degrees; the CDR1 and CDR2 loops frequently interact with the MHC partner and determine HLA class restriction, and the CDR3 loops often contribute antigen peptide specificity. Another generalization is that the αβTCR lies at a diagonal angle over the antigen peptide-binding cleft, in contact with the conserved helices of MHC class I and II molecules. Related TCR-recognition profiles include glycolipids presented to αβTCR by CD1d, bacterial superantigens that elicit broad T-cell responses, metal ion presentation, and involvement of γδTCRs in direct recognition of self-MHC-like molecules. Detailed structures of more of these latter, less ubiquitous binding interactions will no doubt yield further insight into the structural basis of antigen specificity and T-cell restriction.

4 Abbreviations

APC	antigen-presenting cell
C region (domain)	constant region (domain)
CD	cluster of differentiation antigen
CDR	complementarity-determining region
DENV	Dengue virus
EBV	Epstein–Barr virus
EGFR	epidermal growth factor receptor
F_{ab}	antibody fragment (substructure; see Fig. 1a)
F_c	crystallizable fragment (substructure; see Fig. 1a)
F_v	variable fragment
H chain	heavy chain
HEL	hen egg white lysozyme (often used in naming antibodies against)
HER	human epidermal growth factor receptor
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
HPMV	human metapneumovirus
HTLV	human T-cell lymphotropic virus

IFN	interferon
Ig	immunoglobulin
L chain	light chain
LPC	lysophsphatidyl choline
MHC	major histocompatability complex
MIC	MHC class I-related chain
NKT	natural killer T cell
scFv	single-chain variable fragment
SPR	surface plasmon resonance
TCR	T-cell receptor
V region (domain)	variable region (domain)
VEGF	vascular endothelial growth factor

5 Membership of sponsoring body

Membership of the IUPAC Chemistry and Human Health Division Committee for the period 2014–2015 is as follows:

President: T. J. Perun (USA); *Vice President:* R. Cornelis (Belgium); *Secretary:* M. Schwenk (Germany); *Titular Members:* E. Differding (Belgium); J. Fischer (Hungary); V. Gubala (Slovakia); P. Illing (UK); L. Johnston (Canada); H. Møller Johannessen (Denmark); W. A. Temple (New Zealand); *Associate Members:* Vincenzo Abbate (UK); M. Kiilunen (Finland); Y. C. Martin (USA); S. Mignani (France); D. Rotella (USA); *National Representatives:* N. Nahar (Bangladesh); M.-X. Wang (China/Beijing); R. Jih-Ru Hwu (China/Taipei); S. Alihodžić (Croatia); A. Rahatgoanker (India); G. Bee Teh (Malaysia); R. Leurs (Netherlands); B. Haug (Norway); S. Bachurin (Russia); P. Ploypradith (Thailand).

References

[1] P. J. Delves, S. J. Martin, D. R. Burton, I. M. Roitt. *Roitt's Essential Immunology*, 12th ed., Wiley-Blackwell, Chichester, UK (2011).

[2] L. J. Harris, S. B. Larson, K. W. Hasel, A. McPherson. *Biochemistry* **36**, 1581 (1997).

[3] D. M. Templeton, M. Schwenk, R. Klein, J. H. Duffus. *Pure Appl. Chem.* **84**, 1113 (2012).

[4] R. A. Houghten. *Proc. Natl. Acad. Sci. USA* **82**, 5131 (1985).

[5] R. S. Hodges, R. J. Heaton, J. M. R. Parker, L. Molday, R. S. Molday. *J. Biol. Chem.* **263**, 11768 (1988).

[6] R. B. M. Schasfoort, A. J. Tudos (Eds.). *Handbook of Surface Plasmon Resonance*, Royal Society of Chemistry, Cambridge, UK (2008).

[7] L. G. Fägerstam, Å. Frostell, R. Karlsson, M. Kullman, A. Lawn, M. Malmqvist, H. Butt. *J. Mol. Recognit.* **3**, 208 (1990).

[8] U. Dammer, M. Hegner, D. Anselmetti, P. Wagner, M. Dreier, W. Huber, H.-J. Güntherodt. *Biophys. J.* **70**, 2437 (1996).

[9] S. Sheriff, S. W. Silverton, E. A. Padlan, G. H. Cohen, S. J. Smith-Gill, B. C. Finzel, D. R. Davies. *Proc. Natl. Acad. Sci. USA* **84**, 8075 (1987).

[10] Y. Satow, G. H. Cohen, E. A. Padlan, D. R. Davies. *J. Mol. Biol.* **190**, 593 (1986).

[11] S. J. Smith-Gill, A. C. Wilson, M. Potter, E. M. Prager, R. J. Feldmann, C. R. Mainhart. *J. Immunol.* **128**, 314 (1982).

[12] G. H. Cohen, E. W. Silverton, E. A. Padlan, F. Dyda, J. A. Wibbenmeyer, R. C. Willson, D. R. Davies. *Acta Crystallogr., Sect. D* **61**, 628 (2005).

[13] Y. Li, H. Li, S. J. Smith-Gill, R. A. Mariuzza. *Biochemistry* **39**, 6296 (2000).

[14] E. L. Schneider, M. S. Lee, A. Baharuddin, D. H. Goetz, C. J. Farady, M. Ward, C. I. Wang, C. S. Craik. *J. Mol. Biol.* **415**, 699 (2012).

[15] T. Koschubs, S. Dengl, H. Durr, K. Kaluza, G. Georges, C. Hartl, S. Jennewein, M. Lanzendorfer, J. Auer, A. Stern, K. S. Huang, K. Packman, U. Gubler, D. Kostrewa, S. Ries, S. Hansen, U. Kohnert, P. Cramer, O. Mundigl. *Biochem. J.* **442**, 483 (2012).

[16] M. C. Franklin, E. C. Navarro, Y. Wang, S. Patel, P. Singh, Y. Zhang, K. Persaud, A. Bari, H. Griffith, L. Shen, P. Balderes, P. Kussie. *Structure* **19**, 1097 (2011).

[17] M. Chruszcz, A. Pomés, J. Glesner, L. D. Vailes, T. Osinski, P. J. Porebski, K. A. Majorek, P. W. Heymann, T. A. E. Platts-Mills, W. Minor, M. D. Chapman. *J. Biol. Chem.* **287**, 7388 (2012).

[18] S. Ouyang, B. Gong, J.-Z. Li, L.-X. Zhao, W. Wu, F.-U. Zhang, L. Sun, S.-J. Wang, M. Pan, C. Li, W. Liang, N. Shaw, J. Zheng, G.-P. Zhao, Y. Wang, Z.-J. Liu, M. Liang. *J. Mol. Med.* **90**, 837 (2012).

[19] S. L. La Porte, S. S. Bollini, T. A. Lanz, Y. N. Abdiche, A. S. Rusnak, W.-H. Ho, D. Kobayashi, O. Harrabi, D. Pappas, E. W. Mina, A. J. Milici, T. T. Kawabe, K. Bales, J. C. Li, J. Pons. *J. Mol. Biol.* **421**, 525 (2012).

[20] C. Peters-Libeu, J. Miller, E. Rutenber, Y. Newhouse, P. Krishnan, K. Cheung, D. Hatters, E. Brooks, K. Widjaja, T. Tran, S. Mitra, M. Arrasate, L. A. Mosquera, D. Taylor, K. H. Weisgraber, S. Finkbeiner. *J. Mol. Biol.* **421**, 587 (2012).

[21] T. Hino, T. Arakawa, H. Iwanari, T. Yurugi-Kobayashi, C. Ikeda-Suno, Y. Nakada-Nakura, O. Kusano-Arai, S. Weyand, T. Shimamura, N. Nomura, A. D. Cameron, T. Kobayashi, T. Hamakubo, S. Iwata, T. Murata. *Nature* **482**, 237 (2012).

[22] I. P. Fabrichny, G. Mondielli, S. Conrod, M.-F. Martin-Eauclaire, Y. Bourne, P. Marchot. *J. Biol. Chem.* **287**, 14136 (2012).

[23] K. J. Katschke Jr., P. Wu, R. Ganesan, R. F. Kelley, M. A. Mathieu, P. E. Hass, J. Murray, D. Kirchhofer, C. Wiesmann, M. van Lookeren Campagne. *J. Biol. Chem.* **287**, 12886 (2012).

[24] G. Schaefer, L. Haber, L. M. Crocker, S. Shia, L. Shao, D. Dowbenko, K. Totpal, A. Wong, C. V. Lee, S. Stawicki, R. Clark, C. Fields, G. D. Lewis Phillips, R. A. Prell, D. M. Danilenko, Y. Franke, J. P. Stephan, J. Hwang, Y. Wu, J. Bostrom, M. X. Sliwkowski, G. Fuh, C. Eigenbrot. *Cancer Cell* **20**, 472 (2011).

[25] D. C. Ekiert, R. H. E. Friesen, G. Bhabha, T. Kwaks, M. Jongeneelen, W. Yu, C. Ophorst, F. Cox, H. J. W. M. Korse, B. Brandenburg, R. Vogels, J. P. J. Brakenhoff, R. Kompier, M. H. Koldijk, L. A. H. M. Cornelissen, L. L. M. Poon, M. Peiris, W. Koudstaal, I. A. Wilson, J. Goudsmit. *Science* **333**, 843 (2011).

[26] D. C. Ekiert, G. Bhabha, M.-A. Elsliger, R. H. E. Friesen, M. Jongeneelen, M. Throsby, J. Goudsmit, I. A. Wilson. *Science* **324**, 246 (2009).

[27] J. R. R. Whittle, R. Zhang, S. Khurana, L. R. King, J. Manischewitz, H. Golding, P. R. Dormitzer, B. F. Haynes, E. B. Walter, M. A. Moody, T. B. Keple, H.-X. Liao, S. C. Harrison. *Proc. Natl. Acad. Sci. USA* **108**, 14216 (2011).

[28] D. Corti, J. Voss, S. J. Gamblin, G. Codoni, A. Macagno, D. Jarrossay, S. G. Vachieri, D. Pinna, A. Minola, F. Vanzetta, C. Silacci, B. M. Fernandez-Rodriguez, G. Agatic, S. Bianchi, I. Giacchetto-Sasselli, L. Calder, F. Sallusto, P. Collins, L. F. Haire, N. Temperton, L. P. M. Langedijk, J. J. Skehel, A. Lanzavecchia. *Science* **333**, 850 (2011).

[29] M. Pancera, S. Majeed, Y. E. Ban, L. Chen, C. C. Huang, L. Kong, Y. D. Kwon, J. Stuckey, T. Zhou, J. E. Robinson, W. R. Schief, J. Sodroski, R. Wyatt, P. D. Kwong. *Proc. Natl. Acad. Sci. USA* **107**, 1166 (2010).

[30] X. Wu, T. Zhou, J. Zhu, B. Zhang, I. Georgiev, C. Wang, X. Chen, N. S. Longo, M. Louder, K. McKee, S. O'Dell, S. Perfetto, S. D. Schmidt, W. Shi, L. Wu, Y. Yang, Z. Y. Yang, Z. Yang, Z. Zhang, M. Bonsignori, J. A. Crump, S. H. Kapiga, N. E. Sam, B. F. Haynes, M. Simek, D. R. Burton, W. C. Koff, N. A. Doria-Rose, M. Connors, J. C. Mullikin, G. J. Nabel, M. Roederer, L. Shapiro, P. D. Kwong, J. R. Mascola. *Science* **333**, 1593 (2011).

[31] R. Diskin, J. F. Scheid, P. M. Marcovecchio, A. P. West Jr., F. Klein, H. Gao, P. N. P. Gnanapragasam, A. Abadir, M. S. Seaman, M. C. Nussenzweig, P. J. Bjorkman. *Science* **334**, 1289 (2011).

[32] M. K. Gorny, J. Sampson, H. Li, X. Jiang, M. Totrov, X.-H. Wang, C. Williams, T. O'Neal, B. Volsky, L. Li, T. Cardozo, P. Nyambi, S. Suan Zolla-Pazner, X.-P. Kong. *PLoS One* **6**, e27780 (2011).

[33] R. Pejchal, K. J. Doores, L. M. Walker, R. Khayat, P. S. Huang, S. K. Wang, R. L. Stanfield, J. P. Julien, A. Ramos, M. Crispin, R. Depetris, U. Katpally, A. Marozsan, A. Cupo, S. Maloveste, Y. Liu, R. McBride, Y. Ito, R. W. Sanders, C. Ogohara, J. C. Paulson, T. Feizi, C. N. Scanlan, C. H. Wong, J. P. Moore, W. C. Olson, A. B. Ward, P. Poignard, W. R. Schief, D. R. Burton, I. A. Wilson. *Science* **334**, 1097 (2011).

[34] J. S. McLellan, M. C. Pancera Carrico, J. Gorman, J.-P. Julien, R. Khayat, R. Louder, R. Pejchal, M. Sastry, K. Dai, S. O'Dell, N. Patel, S. Shahzad-ul-Hussan, Y. Yang, B. Zhang, T. Zhou, J. Zhu, J. C. Boyington, G.-Y. Chuang, D. Diwanji, I. Georgiev, Y. D. Kwon, D. Lee, M. K. Louder, S. Moquin, S. D. Schmidt, Z.-Y. Yang, M. Bonsignori, J. A. Crump, S. H. Kapiga, N. E. Sam, B. F. Haynes, D. R. Burton, W. C. Koff, L. M. Walker, S. Phogat, R. Wyatt, J. Orwenyo, L.-X. Wang, J. Arthos, C. A. Bewley, J. R. Mascola, G. J. Nabel, W. R. Schief, A. B. Ward, I. A. Wilson, P. D. Kwong. *Nature* **480**, 336 (2011).

[35] J. J. B. Cockburn, M. E. N. Sanchez, A. P. Goncalvez, E. Zaitseva, E. A. Stura, C. M. Kikuti, S. Duquerroy, P. Dussart, L. V. C. Chernomordik, C.-J. Lai, F. A. Rey. *EMBO J.* **31**, 767 (2012).

[36] J. J. Cockburn, M. E. Navarro Sanchez, N. Fretes, A. Urvoas, I. Staropoli, C. M. Kikuti, L. L. Coffey, F. Arenzana Seisdedos, H. Bedouelle, F. A. Rey. *Structure* **20**, 303 (2012).

[37] G. S. Hansman, D. W. Taylor, J. S. McLellan, T. J. Smith, I. Georgiev, J. R. Tame, S. Y. Park, M. Yamazaki, F. Gondaira, M. Miki, K. Katayama, K. Murata, P. D. Kwong. *J. Virol.* **86**, 3635 (2012).

[38] X. Wen, J. C. Krause, G. P. Leser, R. G. Cox, R. A. Lamb, J. V. Williams, J. E. Crowe, T. S. Jardetzky. *Nat. Struct. Mol. Biol.* **19**, 461 (2012).

[39] D. O. Olal, A. Kuehne, S. Bale, P. Halfmann, T. Hashiguchi, M. L. Fusco, J. E. Lee, L. B. King, Y. Kawaoka, J. M. Dye, E. O. Saphire. *J. Virol.* **86**, 2809 (2012).

[40] S. Gras, P. G. Wilmann, Z. Chen, H. Halim, Y. C. Liu, L. Kjer-Nielsen, A. W. Purcell, S. R. Burrows, J. McCluskey, J. Rossjohn. *J. Immnol.* **188**, 311 (2012).

[41] G. A. Bentley, G. Boulot, K. Karjalainen, R. A. Mariuzza. *Science* **267**, 1984 (1995).

[42] B. A. Fields, B. Ober, E. L. Malchiodi, M. I. Lebedeva, B. C. Braden, X. Ysern, J.-K. Kim, X. Shao, E. S. Ward, R. A. Mariuzza. *Science* **270**, 1821 (1995).

[43] K. C. Garcia, M. Degano, R. M. Stanfield, A. Brunmark, M. R. Jackson, P. A. Peterson, L. Teyton, I. A. Wilson. *Science* **274**, 209 (1996).

[44] D. N. Garboczi, P. Ghosh, U. Utz, Q. R. Fan, W. E. Biddison, D. C. Wiley. *Nature* **384**, 134 (1996).

[45] B. A. Fields, E. L. Malchiodi, H. Li, X. Ysern, C. V. Stauffacher, P. M. Schlievert, K. Karjalainen, R. A. Mariuzza. *Nature* **384**, 188 (1996).

[46] P. Parham. *Nature* **384**, 109 (1996).

[47] K. M. Armstrong, K. H. Piepenbrink, B. M. Baker. *Biochem. J.* **415**, 183 (2008).

[48] K. M. Armstrong, F. K. Insaidoo, B. M. Baker. *J. Mol. Recognit.* **21**, 275 (2008).

[49] A. M. Bulek, D. K. Cole, A. Skowera, G. Dolton, S. Gras, F. Madura, A. Fuller, J. J. Miles, E. Gostick, D. A. Price, J. W. Drijfhout, R. R. Knight, G. C. Huang, N. Lissin, P. E. E. Molloy, L. Wooldridge, B. K. Jakobsen, J. Rossjohn, M. Peakman, P. J. Rizkallah, A. K. Sewell. *Nat. Immunol.* **13**, 283 (2012).

[50] D. R. Scott, O. Y. Borbulevych, K. H. Piepenbrink, S. A. Corcelli, B. M. Baker. *J. Mol. Biol.* **414**, 385 (2011).

[51] J. Ekeruche-Makinde, J. J. Miles, H. A. van den Berg, A. Skowera, D. K. Cole, G. Dolton, A. J. Schauenburg, M. P. Tan, J. M. Pentier, S. Llewellyn-Lacey, K. M. Miles, A. M. Bulek, M. Clement, T. Williams, A. Trimby, M. Bailey, P. Rizkallah, J. Rossjohn, M. Peakman, D. A. Price, S. R. Burrows, A. K. Sewell, L. Wooldridge. *Blood* **121**, 1112 (2013).

[52] L. Yin, E. Huseby, J. Scott-Browne, K. Rubtsova, C. Pinilla, F. Crawford, P. Marrack, S. Dai, J. W. Kappler. *Immunity* **35**, 23 (2011).

[53] W. A. Macdonald, Z. Chen, S. Gras, J. K. Archbold, F. E. Tynan, C. S. Clements, M. Bharadwaj, L. Kjer-Nielsen, P. M. Saunders, M. C. J. Wilce, F. Crawford, B. Stadinsky, D. Jackson, A. G. Brooks, A. W. Purcell, J. W. Kappler, S. R. Burrows, J. Rossjohn, J. McCluskey. *Immunity* **31**, 897 (2009).

[54] B. D. Stadinski, P. Trenh, R. L. Smith, B. Bautista, P. G. Huseby, G. Li, L. J. Stern, E. S. Huseby. *Immunity* **35**, 694 (2011).

[55] M. M. Fernández, S. Cho, M. C. De Marzi, M. C. Kerzic, H. Robinson, R. A. Mariuzza, E. L. Malchiodi. *J. Biol. Chem.* **286**, 1189 (2011).

[56] L. Wang, Y. Zhao, Z. Li, Y. Guo, L. L. Jones, D. M. Kranz, W. Mourad, H. Li. *Nat. Struct. Mol. Biol.* **14**, 169 (2007).

[57] S. Gras, Z. Chen, J. J. Miles, Y. C. Liu, M. J. Bell, L. C. Sullivan, L. Kjer-Nielsen, R. M. Brennan, J. M. Burrows, M. A. Neller, R. Khanna, A. W. Purcell, A. G. Brooks, J. McCluskey, J. Rossjohn, S. R. Burrows. *J. Exp. Med.* **207**, 1555 (2010).

[58] S. J. Vessey, J. I. Bell, B. K. Jakobsen. *Eur. J. Immunol.* **26**, 1660 (1996).

[59] Y. C. Liu, Z. Chen, R. S. Burrows, A. W. Purcell, J. McCluskey, J. Rossjohn, S. Gras. *J. Biol. Chem.* **287**, 12267 (2012).

[60] F. E. Tynan, S. R. Burrows, A. M. Buckle, C. S. Clements, N. A. Borg, J. J. Miles, T. Beddoe, J. C. Whisstock, M. C. Wilce, S. L. Silins, J. M. Burrows, L. Kjer-Nielsen, L. Kostenko, A. W. Purcell, J. McCluskey, J. Rossjohn. *Nat. Immunol.* **6**, 1114 (2005).

[61] C. J. Holland, P. J. Rizkallah, S. Vollers, J. M. Calvo-Calle, F. Madura, A. Fuller, A. K. Sewell, L. J. Stern, A. Godkin, D. K. Cole. *Sci. Rep.* **2**, 629 (2012).

[62] O. Y. Borbulevych, K. H. Piepenbrink, B. M. Baker. *J. Immunol.* **186**, 2950 (2011).

[63] S. E. Broughton, J. Petersen, A. Theodossis, S. W. Scally, K. L. Loh, A. Thompson, J. van Bergen, Y. Kooy-Winkelaar, K. N. Henderson, T. Beddoe, J. A. ye-Din, S. I. Mannering, A. W. Purcell, J. McCluskey, R. P. Anderson, F. Koning, H. H. Reid, J. Rossjohn. *Immunity* **37**, 611 (2012).

[64] Y. Yin, Y. Li, M. C. Kerzic, R. Martin, R. A. Mariuzza. *EMBO J.* **30**, 1137 (2011).

[65] D. K. Sethi, D. A. Schubert, A.-K. Anders, A. Heroux, D. A. Bonsor, C. P. Thomas, E. J. Sundberg, J. Pyrdol, K. W. Wucherpfen-nig. *J. Exp. Med.* **208**, 91 (2011).

[66] Y. Yin, X. X. Wang, R. A. Mariuzza. *Proc. Natl. Acad. Sci. USA* **109**, 5405 (2012).

[67] A. P. Uldrich, O. Patel, G. Cameron, D. G. Pellicci, E. B. Day, L. C. Sullivan, K. Kyparissoudis, L. Kjer-Nielsen, J. P. Vivian, B. Cao, A. G. Brooks, S. J. Williams, P. Illarionov, G. S. Besra, S. J. Turner, S. A. Porcelli, J. McCluskey, M. J. Smyth, J. Rossjohn, D. I. Godfrey. *Nat. Immunol.* **12**, 616 (2011).

[68] O. Patel, D. G. Pellicci, S. Gras, M. L. Sandoval-Romero, A. P. Uldrich, T. Mallevaey, A. J. Clarke, J. Le Nours, A. Theodossis, S. L. Cardell, L. Gapin, D. I. Godfrey, J. Rossjohn. *Nat. Immunol.* **13**, 857 (2012).

[69] D. G. Pellicci, O. Patel, L. Kjer-Nielsen, S. S. Pang, L. C. Sullivan, K. Kyparissoudis, A. G. Brooks, H. H. Reid, S. Gras, I. S. Lucet, R. Koh, M. J. Smyth, T. Mallevaey, J. L. Matsuda, L. Gapin, J. McCluskey, D. I. Godfrey, J. Rossjohn. *Immunity* **31**, 47 (2009).

[70] J. López-Sagaseta, L. V. Sibener, J. E. Kung, J. Gumperz, E. J. Adams. *EMBO J.* **31**, 2047 (2012).

[71] E. J. Adams, Y. H. Chien, K. C. Garcia. *Science* **308**, 227 (2005).

[72] B. Xu, J. C. Pizarro, M. A. Holmes, C. McBeth, V. Groh, T. Spies, R. K. Strong. *Proc. Natl. Acad. Sci. USA* **108**, 2414 (2011).

[73] D. M. Templeton. *Pure Appl. Chem.* **76**, 1255 (2004).

[74] L. Yin, F. Crawford, P. Marrack, J. W. Kappler, S. Dai. *Proc. Natl. Acad. Sci. USA* **109**, 18517 (2012).

[75] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin. *J. Comput. Chem.* **25**, 1506 (2004).