The three-dimensional structure of antibodies

Markus Marquart and Johann Deisenhofer

Max-Planck Institut für Biochemie, Abteilung Strukturforschung II, D-8033 Martinsried, F.R.G.

Antibody molecules are glycoproteins which occur in vertebrate species. They recognize and bind an enormous variety of foreign substances (antigens) and subsequently trigger further defense mechanisms at the molecular or cellular level. Specific recognition requires surface structures complementary to the antigen and hence a huge variety of antibody molecules. In contrast the effector functions need identical interaction sites in all antibody molecules.

The determination of the primary structure of immunoglobulins[1-3] and the X-ray crystallographic studies of several antibody molecules and fragments[4,5,7,10,12-15] led to an advanced understanding of the way in which antibodies meet these opposing requirements.

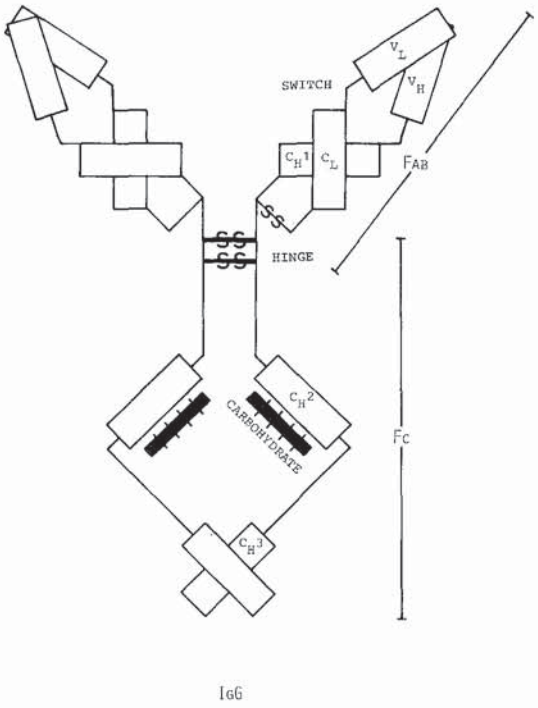

Fig. 1 Schematic representation of an IgG1 immunoglobulin molecule.
The arms of the Y-shaped molecule are formed by the Fab parts, the stem is made up by the Fc part. The light chains are linked to the heavy chains by a disulphide bridge close to the C-terminus. The two heavy chains are connected via two disulphide linkages in the hinge region.

0167-4919/82/0000-0000/$2.75

Fig. 1 is a schematic drawing of an antibody molecule of class IgG1. It is composed of two identical heavy chains and two identical light chains with mol. wts of 50,000 and 25,000, respectively. Both types of polypeptide chain are folded into domains: the four domains of the heavy chain are VH, CH1, CH2, and CH3; the light chain consists of the two domains VL and CL. All domains except CH2 are arranged in pairs which are held together by non-covalent forces. Inter-chain disulfide bridges provide further stability.

Among antibody molecules of a given class and species, the V-domains differ considerably in amino acid sequence, whereas the C-domains have identical sequences. The V-domains are composed of about 110 amino acid residues at the N-terminal end of heavy and light chains. The VH-VL pair together forms the antigen binding site; different antibody specificities are the result of different amino acid sequences of the V-domains. The sequence variability in V-domains is most pronounced in a few hypervariable regions. On the other hand the framework residues are well conserved. The constant domains CH2 and CH3 are involved in effector functions such as complement activation and binding to receptors on certain cell types. There is significant homology between the amino acid sequences of all C-domains, and of the framework residues of V-domains.

Proteolytic cleavage at the hinge region yields stable and functional fragments: the antigen-binding fragment Fab, and the Fc fragment (Fc was the first antibody fragment obtained in crystalline form)[6].

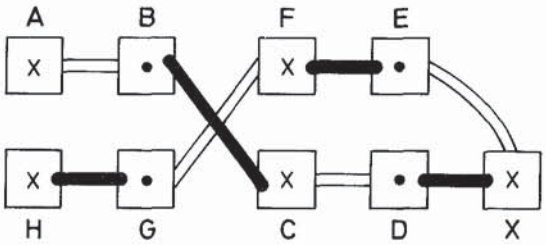

Fig. 2 Schematic drawing of the strand topology in a V-domain viewed parallel to the strands.
(×) and (●) indicate N- and C-terminal ends of the strands pointing towards the observer.

Besides IgG1, several other classes (IgM, IgA, IgD, IgE) and subclasses of immunoglobulins have been identified; the differences between these are located in the constant region of the heavy chain. The two types of light chain (kappa, lambda) can combine with heavy chains of any class.

Domain folding

The general folding pattern in all immunoglobulin domains is very similar. It is shown schematically in Fig. 2 for a V-domain. The folding is characterized by two pleated sheets connected by an internal disulphide bridge linking strands B and C. The two sheets cover a large number of hydrophobic amino acid side chains.

Despite that gross similarity there exist substantial differences when one compares V- and C-domains: C-domains lack strand X, strand D is very short (2–3 amino acids) and connected to strand E. In addition the length of the loop regions in C-domains is different from V-domains, thus changing the overall shape considerably.

VH and VL, on the other hand, show only minor differences when compared with each other (except in the hypervariable regions) as do CL, CH1 and CH3.

CH2 represents yet a third type of domain, differentiated from the other C-domains mainly by the branched carbohydrate chain linked to it. It will be discussed in more detail below.

Domain-domain interaction

Two kinds of domain interactions occur in immunoglobulins: lateral (or trans) interactions and longitudinal (or cis) interactions.

In lateral interactions immunoglobulin domains other than CH2 strongly associate to form modules VL–VH, CL–CH1, CH3–CH3. In V modules VH may be replaced by VL to form light chain V dimers as seen in the Bence-Jones protein fragments Rei or Au[7–9]. In Bence-Jones proteins, which are light chain dimers, one of the light chains simulates the Fab parts of the heavy chain, as described for Meg[10].

V modules associate in a different way than C modules do. In V modules HGCD faces (see Fig. 2) of the domains get into contact, in C modules the ABFE faces are involved.

A considerable loss of accessible surface area[11] is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å², 1923 Å² and 2180 Å² for VL–VH, CL–CH1 modules of IgG Kol[12,13] and the CH3–CH3 module of an human Fc fragment[14,15] respectively. In VL–VH association both framework residues and amino acids from hypervariable segments are involved. A comparison of V-domain amino acid sequences of different animal species shows that the contacting framework residues are highly conserved. Also the constant domain residues participating in lateral contact are either invariant or replaced by homologous residues in different immunoglobulin chains. This low degree of sequence variability for the residues important for lateral contact formation provides an explanation for the fact that different L-chains can associate with different H-chains to give intact immunoglobulins.

In addition to the extensive Van der Waals contacts, there exist a few trans hydrogen bonds, in which mainly polar side chain groups are involved. There are two salt linkages in Kol CL-CH1 contact: Glu 125 light chain – Lys 214 heavy chain, Glu 126 light chain – Lys 148 heavy chain, which have their analgon in CH3 – CH3 pairing: Glu 356 – Lys 439, Glu 357 – Lys 370.

CH2 is an exception, as it forms a single unit without lateral domain interactions (see Fig. 3)*. Instead it interacts with bound carbohydrate, which is attached to Asn 297. The CH2 residues that are involved in carbohydrate contact are, with a few exceptions, structurally in the same positions as the residues that form the CH3–CH3 contact (face ABFE in Fig. 2). This demonstrates that the carbohydrate in CH2 provides a substitute for the C–C contact and presumably helps to stabilize the CH2-domain. The branched carbohydrate forms a few hydrogen bonds with the CH2-domain, but the dominant interactions are hydrophobic in nature. The carbohydrate covers a hydrophobic patch of the protein made up of Phe 241, 243, Val 262, 264, Tyr 296, Thr 260, Arg 301, which would otherwise be exposed to the solvent. The loss of accessible surface area of one CH2 domain is 522 Å², which is only about half as much covered surface area as seen in CH3–CH3 contact (1080 Å²). This observation could explain the apparent 'softness' of those parts of the CH2-domain, as seen in the crystal structure[14,15], which are most remote from the CH3–CH2 interface.

The functional relevance of carbohydrate in antibodies is unclear. It might be involved in intracellular movements of the glycoproteins and in secretion[16–18]. It may well be that the origin of the altered functional properties of carbohydrate-free antibody variants is structural destabilization.

In contrast to the extensive lateral interactions, nonbonded longitudinal interactions along the heavy chain or light chain are much weaker or do not exist at all. However, they are interesting because conformational changes in antibodies affect those interactions.

Fig. 3, which represents the Fc part of an IgG1 molecule shows the CH2–CH3 interaction. With a loss in accessible surface area of 778 Å² this contact has roughly one third of the size of CH3–CH3 contact. The residues that participate in CH2–CH3 contact are highly conserved in all Ig classes, suggesting that this contact is likely to be found in IgG and IgA and as CH3–CH4 contact in IgE and IgM.

* Most readers will need a stereo viewer (commercially available) to see in three dimensions the structures shown in the paired diagrams on pages 162, 163 and 166.

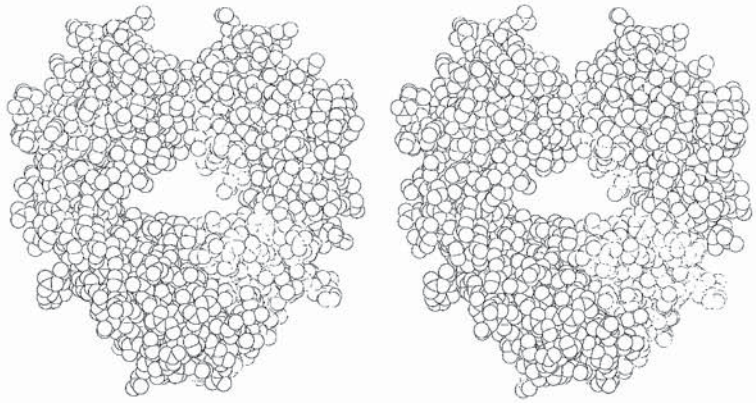

Fig. 3 Stereo drawing of a space filling model of human Fc-fragment.
The molecule is built from two identical polypeptide chains (chain 1, chain 2), and identical carbohydrate groups. Both halves are related by approximate diads.

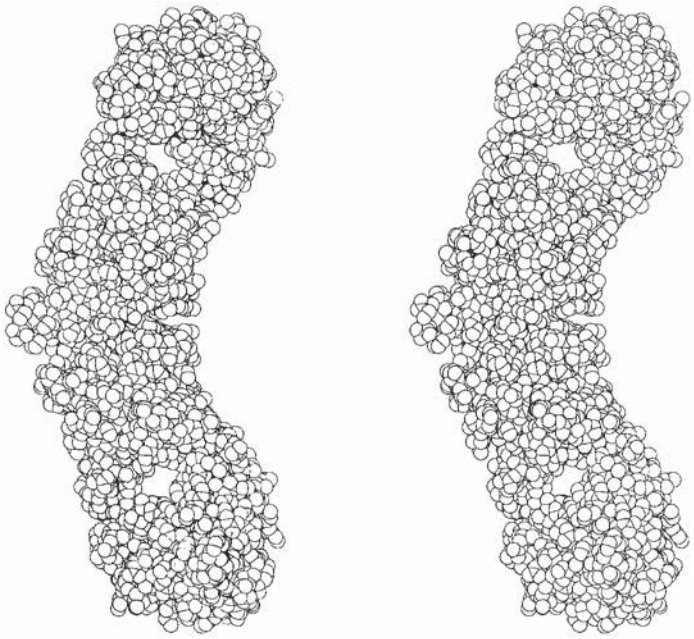

Fig. 4 IgG1 molecule Kol.
The Fab parts and the hinge segment are well ordered in the Kol crystals, the Fc part is disordered and not visible.

PLEASE NOTE

We regret that for technical reasons it has not been possible to reproduce Figs 3, 4, 6, 7 and 8 with the colour coding that allows different parts of the molecules to be distinguished.

The full-colour diagrams, with explanatory legends, can be found in the personal monthly edition of *Immunology Today* dated June 1982.

Fig. 5 Amino acid comparison of residues 98–119 (Eu numbering) of M603, New, Kol and Eu heavy chains. The underlined residues were left out in Fig. 6c.

	End of VH			D segment											
	98														
M603:	Cys	Ala	Arg	Asn	Tyr	Tyr	Gly	Ser	Thr	—	—	—	—	—	—	—
New :	Cys	Ala	Arg	Asn	Leu	Ile	Ala	Gly	Cys	Ile	—	—	—	—	—	—
Kol :	Cys	Ala	Arg	Asp	Gly	Gly	His	Gly	Phe	Cys	Ser	Ser	Ala	Ser	Cys	Phe
Eu :	Cys	Ala	Gly	Gly	Tyr	Gly	Ile	Tyr	Ser	—	—	—	—	—	—	—

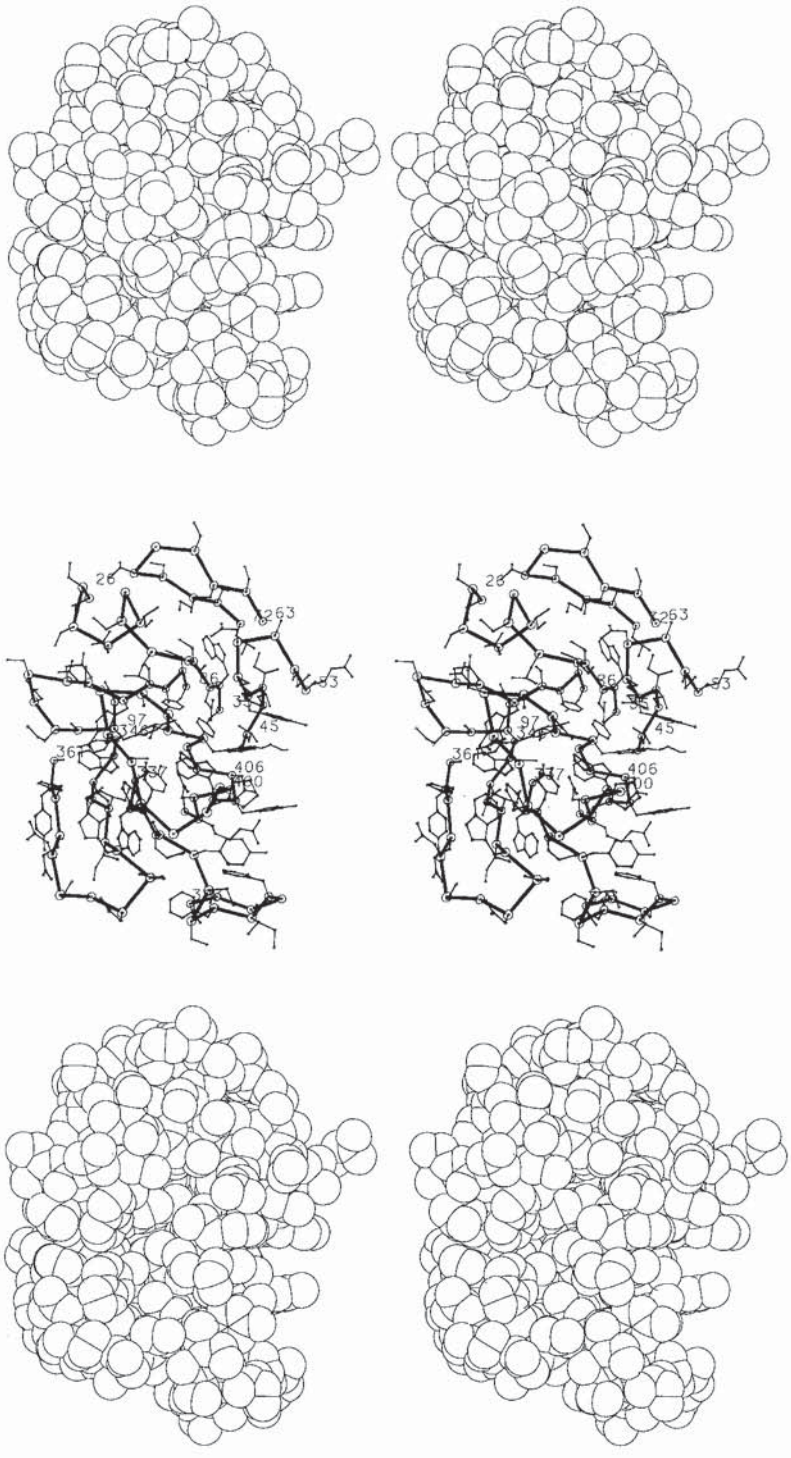

Fig. 6 Antigen binding region of IgG1 Kol.
(a) The extended third hypervariable loop of the heavy chain folds into the putative antigen binding pocket.
(b) C α backbone and sidechains of Kol antigen binding pocket.
(c) Artificial deletion of nine residues in the third hypervariable segment of Kol, which makes it of equal length with IgG1 Eu[28], reveals a deep curved cleft.

		J segment														
							110									119
—		Try	Tyr	Phe	Asp	Val	Try	Gly	Ala	Gly	Thr	Thr	Val	Thr	Val	Ser	Ser
—		—	—	—	Asp	Val	Try	Gly	Gln	Gly	Ser	Leu	Val	Thr	Val	Ser	Ser
Gly	Pro	—	—	—	Asp	Tyr	Try	Gly	Gln	Gly	Thr	Pro	Val	Thr	Val	Ser	Ser
—	—	—	—	—	Pro	Glu	Glu	Tyr	Asn	Gly	Gly	Leu	Val	Thr	Val	Ser	Ser

The CH2–CH3 orientation is found to be somewhat variable and influenced by external forces. In the Fc fragment crystals the two chemically identical chains are in a different environment. As a consequence the CH2–CH3 orientation varies by about 6°. In Fc–Protein A complex crystals this arrangement differs slightly from that of Fc crystals[15].

More drastic changes are observed in VH–CH1 and VL–CL longitudinal contacts, when chemically different Fab fragments are compared. These differences in longitudinal arrangement are most conveniently described by an elbow angle, which is enclosed by the pseudo diads relating VL to VH and CH1 to CL respectively. The elbow angle may vary from more than 170° to 135° when we compare Kol Fab with McPc Fab[12,13,19,20].

In two cases the elbow angles of the same molecule in two different crystal lattices were compared and found to differ by 8° and 17° respectively[19,21]. In Fab New, with an elbow angle of approximately 137°, there exist a few longitudinal contacts between VL and CL and VH and CH1[22,23], whereas there are no non-bonded longitudinal contacts in intact Kol and Fab Kol (see Fig. 4), which are characterized by an open elbow angle. We interpret these observations to mean that in Fab Kol the V–C arrangement is flexible in solution. In the crystal the molecule is stablized by packing interactions; these will be discussed from a different point of view later.

The antigen-binding area

Comparison of amino acid sequences of variable parts has demonstrated the hypervariability of some segments. These were considered to be involved in antigen binding[24]. Indeed, crystal structure analyses of Ig fragment–hapten complexes show that haptens bind in a cleft or depression formed by the hypervariable segments.

The VL dimer of Rei[7,9] may serve as an illustrative example. The symmetrically arranged hypervariable regions form a deep slit-like pocket around the diad relating the two VL monomers. The walls of the slit are lined by tyrosines 49, 91, 96, Asn 34 and Gln 89; the bottom of the pocket is formed by Tyr 36 and Gln 89. A trinitrophenyl group binds to the Rei fragment and fills the binding pocket completely.

Another example of an IgG fragment–hapten complex is Fab New, which is known to bind among other ligands a hydroxy derivative of vitamin K_1[25]. The hypervariable segments of New form a shallow groove with approximate dimensions of 16 × 7 Å and a depth of 6 Å.

McPc 603, a mouse IgA (κ) Fab fragment[20] binds phosphorylcholine. The site of hapten binding is a large wedge shaped cavity, with dimensions 15 × 20 Å and a depth of 12 Å. Only five of the six hypervariable regions contribute to the formation of the cavity: L-chain hypervariable regions one and three, and all three H-chain hypervariable regions. The second hypervariable region of L-chain is screened from the cavity by the first hypervariable loop of L-chain and the third hypervariable loop of H-chain. The deeper cavity in McPc603, as compared to Fab New, is due to longer hypervariable loops. The first hypervariable region of L-chain and the third hypervariable region of H-chain is three residues and the second hypervariable loop of the H-chain is two residues longer in McPc603 than in New.

Phosphorylcholine occupies only a small part of the cavity and interacts via Van der Waals forces, electrostatic interactions, and hydrogen bonds with the protein.

In contrast to the above examples IgG Kol shows no cleft or depression in the antigen-binding region. In IgG Kol the heavy chain has a rather long third hypervariable loop, which contains six residues more than M603 and eight more residues than Fab New. The amino acid sequences of the third hypervariable regions of M603[26], New[23], Kol[27] and Eu[28] are compared in Fig. 5. The sequence alignment and classification in VH, D and J segment[26,29] is somewhat arbitrary, especially for the beginning of the J segment as a nucleotide sequence has been determined only for M603[26]. The additional residues in Kol with the nearly palindromic amino acid sequence -Gly-Phe-Cys-Ser-Ser-Ala-Ser-Cys-Phe-Gly fold into the putative antigen binding site and fill it completely (see Fig. 6a,b). The two cysteins are disulphide bridged and form the start and endpoints of a short antiparallel β-sheet, comprising residues -Cys-Ser-Ser-Ala-Ser-Cys-. If in a model building experiment nine residues are cut from the third hypervariable region of the Kol heavy chain, thus making it of equal length with IgG1 Eu[28], a deep curved cleft appears (Fig. 6c), which easily could accommodate haptens. With respect to the antigen binding area IgG Kol thus looks as if it carried its own hapten in form of an extended third hypervariable loop. Another peculiarity of IgG Kol might be of interest in that context. In the Kol crystal lattice the hypervariable parts of one molecule touch the hinge and spatially adjacent segments of a symmetrically related molecule. This contact consists of three salt linkages (Arg 49 light chain–COOH light chain, Asp 50 light chain–Arg 215 heavy chain, Asp 53 heavy chain–Lys 134 heavy chain), a few hydrogen bonds and extensive Van der Waals interactions. Thus, the lattice contact found in Kol crystals might give an instructive model for antibody–antigen interaction, as antigens are usually macromolecules which cover a much larger part of the antibody than haptens do.

The hinge segment

The hinge segment which covalently links Fab and Fc parts, has a unique primary and spatial structure. Its central region consists of two parallel disulphide-linked poly L-proline helices with an amino acid sequence -Cys-Pro-Pro-Cys-[12,13]. In the IgG1 subclass represented by the Kol molecule the poly-proline double helix is short (Fig. 7). However, in IgG3 the hinge sequence is quadruplicated[30] and model build-

DOCKET ALARM

Explore Litigation Insights

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

Real-Time Litigation Alerts

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

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

Advanced Docket Research

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

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

Analytics At Your Fingertips

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

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

API

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

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

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

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

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