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 immunoglobulins1-3 and the X-ray crystallographic studies of several antibody molecules and fragments5,7,12,13 led to an advanced understanding of the way in which antibodies meet these opposing requirements.

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. 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 disulfide bridge close to the C-terminus. The two heavy chains are connected via two disulfide linkages in the hinge region.

Fig. 2 Schematic drawing of the strand topology in a V-domain viewed parallel to the strands. (X) 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 disulfide 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 V1-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 Re1 or Au7-9. 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 Mcg10.

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

A considerable loss of accessible surface area is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

**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 VI-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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.

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 is connected with contact formation of the immunoglobulin domains. It amounts to 1760 Å2, 1923 Å2 and 2180 Å2 for 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 Re1 or Au7-9. 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 Mcg10.
Fig. 3 Stereo drawing of a space filling model of human Fe-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 Fe 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.

<table>
<thead>
<tr>
<th>End of VH</th>
<th>D segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>M603: Cys Ala Arg</td>
<td>Asn Tyr Tyr Gly Ser Thr</td>
</tr>
<tr>
<td>New: Cys Ala Arg</td>
<td>Asn Leu Ile Ala Gly Cys Ile</td>
</tr>
<tr>
<td>Kol: Cys Ala Arg</td>
<td>Asp Gly Gly His Gly Phe Cys Ser Ser Ala Ser Cys Phe</td>
</tr>
<tr>
<td>Eu: Cys Ala Gly</td>
<td>Gly Tyr Gly Ile Tyr Ser</td>
</tr>
</tbody>
</table>
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) Cα 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 Eu², reveals a deep curved cleft.
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 Fab12,13,19,20.

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° respectively19,21. In Fab New, with an elbow angle of approximately 137°, there exist a few longitudinal contacts between VL and CL and VH and CH12,22, 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 stabilized 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 binding24. 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 Re177 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 Re1 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 K15. 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 (x) Fab fragment29 binds phosphorycholine. 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 McPC 603, 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 McPC 603 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 M60334, New3, Kol22 and Eu26 are compared in Fig. 5. The sequence alignment and classification in VII, D and J segment18,20 is somewhat arbitrary, especially for the beginning of the J segment as a nucleotide sequence has been determined only for M60335. The additional residues in Kol with the nearly palindromic amino acid sequence -Gly-Phe-Cys-Ser-Ser-Ala-Ser-Gly-Phe-Gly fold into the putative antigen binding site and fill it completely (see Fig. 6a,b). The two cysteines are disulphide bridged and form the start and endpoints of a short antiparallel β-sheet, comprising residues -Cys-Ser-Ala-Ser-Cys.

If in 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 Eu26, 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 78 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-Cys12,13. 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 quadruplicated16 and model build-
The poly-proline segment, a relatively rigid structure, is flanked on both sides by flexible segments. The segment on the C-terminal side is well-defined in the crystal lattice of Kol due to crystal packing interactions, but it lacks internal interactions, that would provide stability in solution. The C terminal segment is disordered and flexible in Kol crystals and in the Fc crystal structure. The rigid hinge segment allows independent movement of the Fab arms and the Fc part. There is direct evidence for flexibility in the crystal lattice of Kol and this is in contrast to the abnormal IgG protein Doh, which lacks a hinge region. The significance of the hinge for Fab-Fc flexibility is obvious.

**Complement binding**

The binding of the Clq component of the C1 complex to antigen-antibody complexes is the first step in the classical pathway of complement activation. The Clq head pieces bind to the CH2 domains of antibodies. Protein A, a constituent of the cell wall of Streptococcus pneumoniae, binds to the Fe part of antibody molecules of certain classes and subclasses, but does not interfere with complement binding. The determination of the crystal structure of the complex between FB (one of the four Fe-binding domains of protein A) and Fc-fragment showed that protein A binds at the CH2-CH3 contact. Fig. 8 shows a space-filling model of the Fβ–Fc complex. The area of CH2 not covered by FB must contain the Clq binding site. In view of the size of the Clq head pieces (mol. wt. 50,000) it appears unlikely that they can bind at the inner sides of CH2, i.e. near the carbohydrate. The most plausible binding site is therefore near the tip of CH2 on the outer side of the domain. It is worth mentioning that this region is disordered in crystals of the Fβ–Fc complex which indicates that this part of the CH2 domain is flexible. Possibly, flexibility is required for antibody Clq interaction.

**Summary and perspectives**

Investigations of the three-dimensional architecture of antibodies have elucidated the folding of the polypeptide chains into domains, and the spatial arrangement of the domains. The structural basis for understanding antibody specificity and antibody flexibility was obtained. Segment flexibility is an important property of antibodies: Flexible segments of the polypeptide chains at the switch and hinge regions allow the Fab fragments to change their shape and their relative orientation. Conformational changes of this kind are necessary to meet the geometric requirements which arise on binding of antibodies to multivalent antigens.

The understanding of the effector functions of antibody molecules is much less complete. One of the central problems is the explanation of the strong enhancement of Clq binding to antigen-antibody complexes as compared to free antibody molecules. Two mechanisms have been considered (for a review see Ref. 39): since Clq is multimeric with at least six antibody binding sites, binding may be enhanced by the formation of antigen-antibody aggregates through crosslinking. Alternatively, antigen binding might induce a conformational change in the Fc-part which enhances affinity for Clq.

There is strong evidence for the importance of aggregation, but a mixed mechanism which involves aggregation and a conformational change cannot be ruled out.

The studies described here were almost exclusively carried out with myeloma or Bence-Jones proteins because these were the only homogeneous immunoglobulins which could be obtained in sufficient quantity. However, in most cases the specificities of such molecules is unknown. Recently, large amounts of homogeneous antibodies elicited against streptococcal or pneumococcal polysaccharides became available from certain rabbit and mouse strains. These sources, and the use of hybrids obtained from myeloma and spleen cells have made it possible to obtain homogeneous antibodies of defined specificity. Structural studies of 'natural' antigen-antibody complexes can be expected to lead to a more complete understanding of antibody function. Crystallographic work on a specific antibody and of its antigen is already in progress.

**Acknowledgements**

We thank Prof. R. Huber for helpful discussions.

**References**


Celltrion, Inc., Exhibit 1082
Fig. 7 Conformation of the hinge region as seen in IgG1 Kcl.

Fig. 8 Space filling model of the FB (protein A) – Fc complex.