Antibody Structure

DAVID R. DAVIES* AND SUSAN CHACKO

Laboratory of Molecular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, Bethesda, Maryland 20892

Received March 12, 1993

Almost 100 years ago Paul Ehrlich proposed a remarkably modern model for an antibody molecule in which the antibody or "receptor" was branched, permitting multiple sites for binding and for activation of complement. This model also invoked the concept of complementarity in antibody–antigen recognition, consistent with the "lock and key" fit proposed by Fischer for enzymes. In the intervening years a detailed knowledge of the three-dimensional structure and function of this molecule has been obtained (see reviews in refs 1–10b). In this brief Account we shall attempt to summarize rather selectively the present state of the field.

Antibodies have a distorted Y shape (Figure 1) with two arms (the Fab fragments) containing at their tips identical antigen binding sites and with the stem (the Fc fragment) joined to the Fabs by a flexible hinge. Each antibody molecule has two identical heavy (H) and light (L) chains. Each chain has an N-terminal variable domain (VH and VL in the heavy and light chains, respectively) with the remainder consisting of domains (CL, CH1, CH2, CH3, CH4) that are constant for a given class (Figure 2). There are several classes: κ and λ for the light chains; and α, γ, δ, ϵ, and μ for the heavy chains.[11] While the α, γ, and δ chains have three constant domains, the ϵ and μ chains have four. The domain structure is similar for the V and C domains, is often referred to as the immunoglobulin domain structure, and is observed in various modifications in members of an extensive superfamily. The functional unit in antibodies is usually a pair of domains from different chains tightly linked by noncovalent bonds (for example, VH:VL, CH1:CL, CH3:CH3).

The Whole Molecule

Until quite recently, the high-resolution information about the structure of the intact antibody molecule came from studies of its fragments, such as the Fab,[12,13] Fc,[14] and light chain dimers.[15,16] This was because the flexibility of the hinge led to disorder in several of the crystals that were examined (Kol[17] and Zie[18]), so that the Fc could not be observed in the electron density. This disorder did not occur in crystals of two antibodies with hinge deletions (Dob[19] and Mcg[20]), and the structures of both of these were determined at low resolution[21,22] and recently at higher resolution for Mcg.[23] In all of the above crystals the antibody resides on a crystal dyad axis resulting in a symmetrical shape for the molecule. However, the work of many authors using electron microscopy[24] and fluorescence depolarization[25,26] clearly indicated that in solution the hinge provides flexibility so that the two Fabs have considerable rotational flexibility relative to each other and to the Fc. A preliminary study has recently been published[27] of the structure at 3.5-Å resolution of a complete anti-canine lymphoma antibody in which all parts of the molecule are localized in the crystal and visible in the electron density map. In this crystal the molecule is not restricted by crystallographic symmetry and adopts a very asymmetric conformation (Figure 1). In the two Fabs the elbow angles (see below) are

David Davies is the Chief of the Section on Molecular Structure in the Laboratory of Molecular Biology, NIDDK, National Institutes of Health. He was educated at Oxford University and carried out postdoctoral work at Caltech. He has been at the NIH since 1955 and has worked on a variety of nucleic acid and protein structures. His present interests include protein–protein interactions in complex enzymes and antibody–antigen and ligand–receptor interactions.

Susan Chacko is an IRTA Fellow in the Laboratory of Molecular Biology, NIDDK. Her undergraduate education was at the Indian Institute of Technology, Kharagpur, India. She received her Ph.D. at the University of Illinois at Urbana-Champaign and Rice University in the laboratory of Dr. George Phillips. Her research interests include antibody structure, diffuse X-ray scattering, computation, and graphics.

(1) Mariuzza, R. A.; Phillips, S. E. V.; Poljak, R. J. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 139–159.
(2) Marquart, M.; Deisenhofer, J. *J. Immunol. Today* **1982**, *3*, 160–166.
(3) Alzari, P. M.; Lascombe, M.-B.; Poljak, R. J. *Annu. Rev. Immunol.* **1988**, *6*, 555–580.
(4) Chothia, C. *Curr. Opin. Struct. Biol.* **1991**, *1*, 53–59.
(5) Colman, P. M. *Curr. Opin. Struct. Biol.* **1991**, *1*, 232–236.
(6) Colman, P. M. *Adv. Immunol.* **1988**, *43*, 99–132.
(7) Mian, I. S.; Bradwell, A. R.; Olson, A. J. *J. Mol. Biol.* **1991**, *217*, 133–151.
(8) Davies, D. R.; Padlan, E. A.; Sheriff, S. *Annu. Rev. Biochem.* **1990**, *59*, 439–473.
(9) Davies, D. R.; Padlan, E. A. *Curr. Biol.* **1992**, *2*, 254–256.
(10) (a) Sheriff, S. *Immunomethods* **1993**, in press. (b) Sheriff, S. *Immunomethods* **1993**, in press.
(11) Paul, W. E. *Fundamental Immunology*; Raven Press: New York, 1989.
(12) Padlan, E. A.; Segal, D. M.; Spande, T. F.; Davies, D. R.; Rudikoff, S.; Potter, M. *Nature (New Biol.)* **1973**, *245*, 165–167.
(13) Poljak, R. J.; Amzel, L. M.; Avey, H. P.; Chen, B. L.; Phizackerley, R. P.; Saul, F. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3440–3444.
(14) Deisenhofer, J. *Biochemistry* **1981**, *20*, 2361–2370.
(15) Schiffer, M.; Girling, R. L.; Ely, K. R.; Edmundson, A. B. *Biochemistry* **1973**, *12*, 4620–4631.
(16) Epp, O.; Colman, P. M.; Fehlhammer, H.; Bode, W.; Schiffer, M.; Huber, R.; Palm, W. *Biochemistry* **1974**, *14*, 4943–4952.
(17) Marquart, M.; Deisenhofer, J.; Huber, R.; Palm, W. *J. Mol. Biol.* **1980**, *141*, 369–391.
(18) Ely, K. R.; Colman, P. M.; Abola, E.; Hess, A. C.; Peabody, D. R. *Biochemistry* **1978**, *17*, 820–825.
(19) Terry, W. D.; Mathews, B. W.; Davies, D. R. *Nature* **1968**, *220*, 239.
(20) Edmundson, A. B.; Wood, M. K.; Schiffer, M.; Hardman, K. D.; Ainsworth, C. F.; Ely, K. R.; Deutsch, H. F. *J. Biol. Chem.* **1970**, *245*, 2763.
(21) Silverton, E. W.; Navia, M. A.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5140–5144.
(22) Rajan, S. S.; Ely, K. R.; Abola, E. E.; Wood, M. K.; Colman, P. M.; Athay, R. J.; Edmundson, A. B. *Mol. Immunol.* **1983**, *20*, 787–799.
(23) Guddat, L. W.; Herron, J. N.; Edmundson, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4271–4275.
(24) Valentine, R. C.; Green, N. M. *J. Mol. Biol.* **1967**, *246*, 3259–3268.
(25) Cathou, R. E.; Dorrington, K. J. *Biological Macromolecules: Subunits in Biological Systems*; Dekker: New York, 1975; Part C, pp 91–224.
(26) Yguerabide, J.; Epstein, H. F.; Stryer, L. *J. Mol. Biol.* **1970**, *51*, 573–590.
(27) Harris, L. J.; Larson, S. B.; Hasel, K. W.; Day, J.; Greenwood, A.; McPherson, A. *Nature* **1992**, *360*, 369–372.

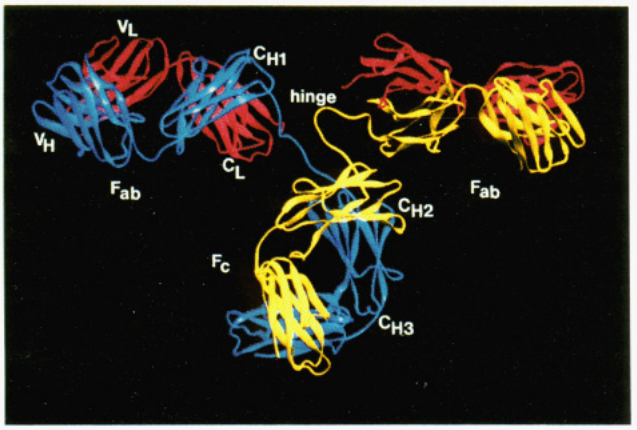

Figure 1. Backbone representation of an intact monoclonal antibody for canine lymphoma. Reproduced with permission from ref 27. Copyright 1992 *Nature*.

not the same, being 159° and 143°. The hinge of this IgG2a antibody adopts a twisted shape and is described as being a tether rather than a hinge. The availability of 2.8-Å data for these crystals will provide an opportunity to examine a complete hinge conformation at this higher resolution. The earlier analysis of Kol,[17] a human IgG1 antibody at 3.0-Å resolution, showed that the hinge contained a segment, Cys-Pro-Pro-Cys, in a polyproline-like conformation.

The Fab

Since 1973, and particularly in the past 4 or 5 years, a large number of Fab structures have been crystallographically determined,[28-72] for which the coordinates of 23 are presently available in the Protein Data Bank.[73,74] As a result, Fabs are rapidly becoming the most studied family of molecules. The solution of new structures has been facilitated by the technique of molecular replacement, powerfully implemented in software packages such as Merlot[75] and XPLOR.[76] This technique takes advantage of the overall similarity of the Fabs and uses the atomic coordinates of previously solved structures to determine a preliminary set of diffraction phases for a new Fab crystal, thus avoiding the need for heavy atom derivatives.

The four domains of the Fab have an overall topological similarity, characterized by two beta-sheets packed closely against each other with a disulfide bridge connecting them (Figure 2). The variable domains of different Fabs have conserved framework structures (FRs) and quite different sequences in the hypervariable loops or complementarity determining regions (CDRs). The six CDRs from the light and heavy chains cluster together in the antigen-binding region to form a continuous hypervariable surface.

(28) Suh, S. W.; Bhat, T. N.; Navia, M. A.; Cohen, G. H.; Rao, D. N.; Rudikoff, S.; Davies, D. R. *Proteins* **1986**, *1*, 74–80.
(29) Saul, F. A.; Amzel, L. M.; Poljak, R. J. *J. Biol. Chem.* **1978**, *253*, 585–597.
(30) Satow, Y.; Cohen, G. H.; Padlan, E. A.; Davies, D. R. *J. Mol. Biol.* **1987**, *190*, 593–604.
(31) Stanfield, R. C.; Fieser, T. M.; Lerner, R. A.; Wilson, I. A. *Science* **1990**, *248*, 712–719.
(32) Herron, J. N.; He, X. M.; Mason, M. L.; Voss, E. W.; Edmundson, A. B. *Proteins: Struct., Funct. Genet.* **1989**, *5*, 271–280.
(33) Herron, J. N.; He, X. M.; Ballard, D. W.; Blier, P. R.; Pace, P. E.; Bothwell, A. L. M.; Voss, E. W.; Edmundson, A. B. *Proteins: Struct., Funct. Genet.* **1991**, *11*, 156–175.
(34) Lascombe, M-B.; Alzari, P. M.; Boulot, G.; Saludjian, P.; Tougard, P.; Berek, C.; Haba, S.; Rosen, E. M.; Nisonoff, A.; Poljak, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 607–611.
(35) Fischmann, T. O.; Bentley, G. A.; Bhat, T. N.; Boulot, G.; Mariuzza, R. A.; Phillips, S. E. V.; Tello, D.; Poljak, R. J. *J. Biol. Chem.* **1991**, *266*, 12915–12920.
(36) Amit, A. G.; Mariuzza, R. A.; Phillips, S. E. V.; Poljak, R. J. *Science* **1986**, *233*, 747–753.
(37) Sheriff, S.; Silverton, E. W.; Padlan, E. A.; Cohen, G. H.; Smith-Gill, S. J.; Finzel, B. C.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8075–8079.
(38) Padlan, E. A.; Silverton, E. W.; Sheriff, S.; Cohen, G. H.; Smith-Gill, S. J.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5938–5942.
(39) Colman, P. M.; Laver, W. G.; Varghese, J. N.; Baker, A. T.; Tulloch, P. A.; Air, G. M.; Webster, R. G. *Nature* **1987**, *326*, 358–363.
(40) Strong, R. K.; Campbell, R.; Rose, D. R.; Petsko, G. A.; Sharon, J.; Margolies, M. N. *Biochemistry* **1991**, *30*, 3739–3748.
(41) Saul, F. A.; Poljak, R. J. *Proteins* **1992**, *14*, 363–371.
(42) Brunger, A. T.; Leahy, D. J.; Hynes, T. R.; Fox, R. O. *J. Mol. Biol.* **1991**, *221*, 239–256.
(43) Rini, J. M.; Schulze-Galmen, U.; Wilson, I. A. *Science* **1992**, *255*, 959–965.
(44) He, X. M.; Ruker, F.; Casale, E.; Carter, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7154–7158.
(45) Tulip, W. R.; Varghese, J. N.; Webster, R. G.; Air, G. M.; Laver, W. G.; Colman, P. M. *Cold Spring Harbor Symp. Quant. Biol.* **1989**, *54*, 257–263.
(46) Tulip, W. R.; Varghese, J. N.; Laver, W. G.; Webster, R. G.; Colman, P. M. *J. Mol. Biol.* **1992**, *227*, 122–148.
(47) Tulip, W. R.; Varghese, J. N.; Webster, R. G.; Laver, W. G.; Colman, P. M. *J. Mol. Biol.* **1992**, *227*, 149–159.
(48) Cygler, M.; Boodhoo, A.; Lee, J. S.; Anderson, W. F. *J. Biol. Chem.* **1987**, *262*, 643–648.
(49) Bizebard, T.; Mauguen, Y.; Skehel, J. J.; Knossow, M. *Acta Crystallogr.* **1991**, *B47*, 549–555.
(50) Brunger, A. T. *Acta Crystallogr.* **1991**, *A47*, 195–204.
(51) Jeffrey, P. D.; Strong, R. K.; Campbell, R. L.; Chang, C.; Sieker, L. C.; Petsko, G. A.; Haber, E.; Margolies, M. N.; Sheriff, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, submitted for publication.
(52) Arevalo, J.; Stura, E. A.; Taussig, M. J.; Wilson, I. A. *J. Mol. Biol.* **1993**, in press.
(53) Lescar, J.; Riottot, M-M.; Souchon, H.; Chitarra, V.; Bentley, G. A.; Navaza, J.; Alzari, P. M.; Poljak, R. J. *Proteins* **1993**, *15*, 209–212.
(54) Rini, J. M.; Stanfield, R. L.; Stura, E. A.; Salinas, P. A.; Profy, A. T.; Wilson, I. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, in press.
(55) Navia, M. A.; Segal, D. M.; Padlan, E. A.; Davies, D. R.; Rao, N.; Rudikoff, S.; Potter, M. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 4071–4074.
(56) Altschuh, D.; Vix, O.; Rees, B.; Thierry, J.-C. *Science* **1992**, *256*, 92–94.
(57) Alzari, P. M.; Spinelli, S.; Mariuzzu, R. A.; Boulot, G.; Poljak, R. J.; Jarvis, J. M.; Milstein, C. *EMBO J.* **1990**, *9*, 3807–3814.
(58) Amzel, L. M.; Poljak, R. J.; Saul, F.; Varga, J. M.; Richards, F. F. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 1427–1430.
(59) Arnold, E.; Jacobo-Molina, A.; Nanni, R. G.; Williams, R. L.; Lu, X.; Ding, J.; Clark, A. D., Jr.; Zhang, A.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H. *Nature* **1992**, *357*, 85–89.
(60) Bentley, G. A.; Boulot, G.; Riottot, M. M.; Poljak, R. J. *Nature* **1990**, *348*, 254–257.
(61) Brady, R. L.; Edwards, D. J.; Hubbard, R. E.; Jiang, J.-S.; Lange, G.; Roberts, S. M.; Todd, R. J.; Adair, J. R.; Emtage, J. S.; King, D. J.; Low, D. C. *J. Mol. Biol.* **1992**, *227*, 253–264.
(62) Colman, P. M.; Webster, R. G. *Biological Organization at High Resolution*; Academic Press: Orlando, FL, 1987; pp 125–133.
(63) Cygler, M.; Rose, D. R.; Bundle, D. R. *Science* **1991**, *253*, 442–445.
(64) Derrick, J. P.; Wigley, D. B. *Nature* **1992**, *359*, 752–754.
(65) Fan, Z.-C.; Shan, L.; Guddat, L. W.; He, X.-M.; Gray, W. R.; Raison, R. L.; Edmundson, A. B. *J. Mol. Biol.* **1992**, *228*, 188–207.
(66) (a) Garcia, K. C.; Verroust, P. J.; Brunger, A. T.; Amzel, L. M. *Science* **1992**, *257*, 502–507. (b) Lascombe, M-B.; Alzari, P. M.; Poljak, R. J.; Nisonoff, A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9429–9433.
(67) Padlan, E. A.; Segal, D. M.; Spande, T. F.; Davies, D. R.; Rudikoff, R.; Potter, M. *Nature (New Biol.)* **1973**, *245*, 165–167.
(68) Poljak, R. J.; Amzel, L. M.; Avey, H. P.; Chen, B. L.; Phizackerley, R. P.; Saul, F. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70*, 3305–3310.
(69) Prasad, L.; Vandonselaar, M.; Lee, J. S.; Delbaere, L. T. J. *J. Biol. Chem.* **1988**, *263*, 2571–2574.
(70) Segal, D. M.; Padlan, E. A.; Cohen, G. H.; Rudikoff, S.; Potter, M.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4298–4302.
(71) Vitali, J.; Young, W. W.; Schatz, V. B.; Sobottka, S. E.; Kretsinger, R. H. *J. Mol. Biol.* **1987**, *198*, 351–355.
(72) Tormo, J.; Stadler, E.; Skern, T.; Auer, H.; Kanzler, O.; Betzel, C.; Blaas, D.; Fita, I. *Protein Sci.* **1992**, *1*, 1154–1161.
(73) Abola, E. E.; Bernstein, F. C.; Bryant, H.; Koetzle, T. F.; Weng, J. *Crystallographic Databases—Information Content, Software Systems, Scientific Applications*; Data Commission of the International Union of Crystallography: Bonn/Cambridge/Chester, 1987; pp 107–132.
(74) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *Eur. J. Biochem.* **1977**, *80*, 319–324.
(75) Fitzgerald, P. M. D. *J. Appl. Crystallogr.* **1988**, *21*, 273–278.
(76) Brunger, A. T. *XPLOR*; Yale University: New Haven, CT.

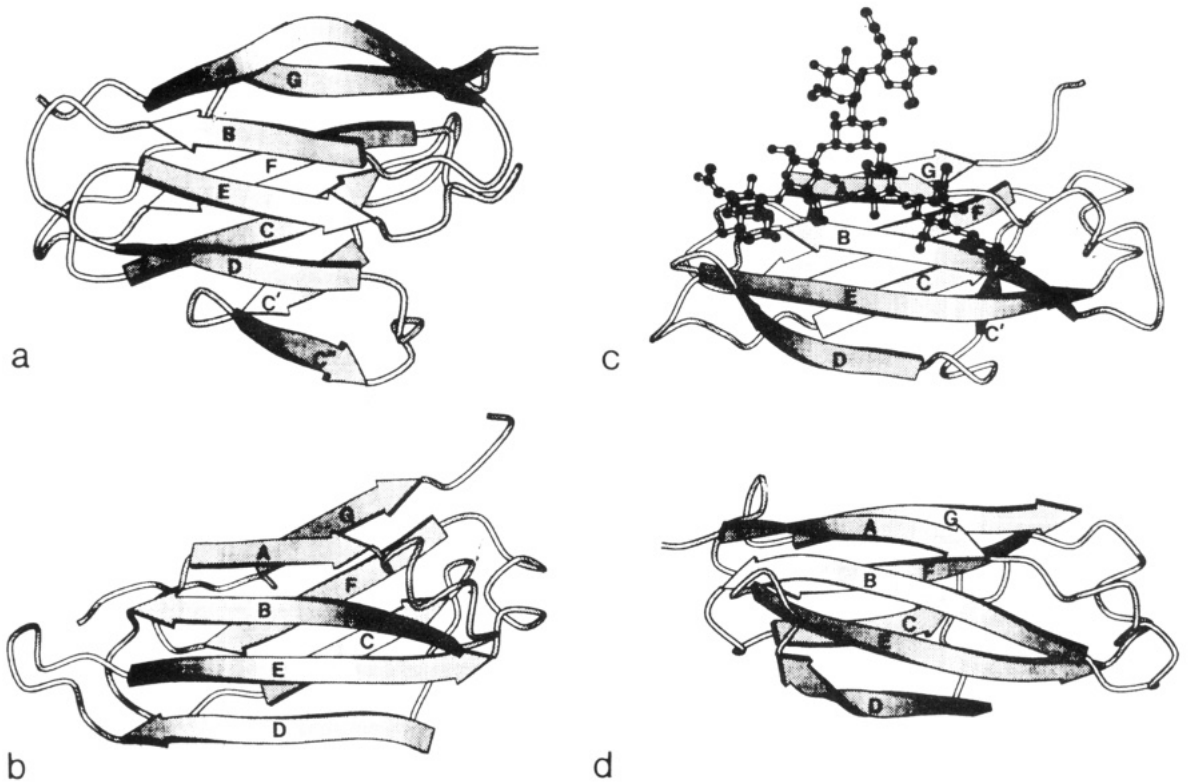

Figure 2. Backbone representation of immunoglobulin domains: (a) variable domain, light chain from J539;[28] (b) constant domain, light chain from J539;[28] (c) CH2 domain with the carbohydrate moiety shown as a ball-and-stick model (from PDB file 1FC2[14,73]); (d) CH3 domain (from PDB file 1FC2[14,73]).

The Antibody Combining Site. An examination of the frequencies of occurrence of amino acids in immunoglobulin domains[77,78] reveals that asparagine, histidine, and tyrosine are more likely to be present in the CDRs than in the framework regions, while cysteine is more likely to be in the FRs. Apolar aliphatic residues (Ala, Val, Ile, Leu) are represented as frequently in the FRs as in other globular proteins but are much less likely to be found in the CDRs. Aromatic residues (Phe, Tyr, Trp), which are usually buried in globular proteins and in the FRs, are more exposed to solvent in the CDRs. These aromatic residues contribute large areas to the antigen-binding surface and may also reduce the loss of conformational entropy upon antigen binding.[77] Examination of the antigen-binding sites of three lysozyme–anti-lysozyme complexes[77] reveals that aromatic residues play a large part in antigen binding. No apolar aliphatic residues were found to interact with the lysozyme.

Chothia and co-workers[79] have examined the known Fab structures and found that five of the six hypervariable loops are limited to a few main-chain conformations, confirming an observation originally based on fewer examples.[80–82] The conformations are largely determined by the interactions of a few residues at specific sites in the CDRs and FRs. These "canonical" conformations can thus be predicted by the size of the CDRs and the occurrence of a specific set of residues that produces a known conformation. Sequence differences in the CDRs alter the antigen-binding surface, and sequence differences in the CDRs and FRs can slightly shift the loop conformations relative to each other. The combined changes account for the wide range of binding specificities found in antibodies. The most variable CDR is H-CDR3, which varies in length from 6 to 14 residues;[83,84] consequently, its conformation cannot be successfully predicted from the existing data base.

The Domain Interactions. The pseudo-twofold axis of rotation between the VH and VL domains varies from 165° to 180°. This angle does not change appreciably between different molecules of the same Fab in the same crystal or in different crystal forms but can change significantly upon binding to the antigen. The VH–VL interface has a large contact surface in the range 1400–1900 Å^2.[85,86] The contacting surface involves both framework and CDR residues, of which the framework residues are highly conserved. An average over 10 Fab structures[86] shows that the four framework regions of the VL contribute 2%, 37%, 7%, and 14% and the three CDRs contribute 7%, 6%, and 25%,

(77) Padlan, E. A. *Proteins* **1990**, *7*, 112–124.
(78) Kabat, E. A.; Wu, T. T.; Bilofsky, H. *J. Biol. Chem.* **1977**, *252*, 669–6616.
(79) Chothia, C.; Lesk, A. M.; Tramontano, A.; Levitt, M.; Smith-Gill, S. J.; Air, G.; Sheriff, S.; Padlan, E. A.; Davies, D.; Tulip, W. R.; Colman, P. M.; Spinelli, S.; Alzari, P. M.; Poljak, R. M. *Nature* **1989**, *342*, 877–883.
(80) Padlan, E. A.; Davies, D. R.; Pecht, I.; Givol, D.; Wright, C. *Cold Spring Harbor Symp. Quant. Biol.* **1977**, *41*, 627–637.
(81) Padlan, E. A. *Q. Rev. Biophys.* **1977**, *10*, 35–65.
(82) de la Paz, P.; Sutton, B. J.; Darsley, M. J.; Rees, A. R. *EMBO J.* **1986**, *5*, 415–425.
(83) Kabat, E. A.; Wu, T. T.; Reid-Miller, M.; Perry, H. M.; Gottesman, K. *Sequences of Proteins of Immunological Interest*; National Institutes of Health: Bethesda, MD, 1987.
(84) Padlan, E. A.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 819–823.
(85) Chothia, C.; Novotny, J.; Bruccoleri, R.; Karplus, M. *J. Mol. Biol.* **1985**, *186*, 651–663.
(86) The VH–VL interface surface area was calculated for Fabs J539, KOL, McPc603, 4-420, R19.9, D1.3, HyHEL-5, HyHEL-10, AN02, and 3D6 using programs MS, MSAV,[86a] and ATMSRF[86b] with a probe radius of 1.7 Å. CDRs and FRs were identified according to the method of Kabat.[83] The percentages do not add up to 100 because of averaging and roundoff errors. (a) Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548–558. (b) Sheriff, S.; Hendrickson, W. A.; Stenkamp, R. E.; Sieker, L. C.; Jensen, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1104–1107.

respectively, to the VL surface area buried in the interface. The four FRs of the VH contribute an average of 0%, 32%, 5%, and 15%, respectively, to the buried interface surface, and the three CDRs contribute an average of 2%, 12%, and 31%, respectively. On each chain, therefore, the FR2 and CDR3 regions contribute at least 60% of the VH–VL interface surface area and are the major determinants of the interface region. It should also be noted that at least 40% of the interface surface area on each chain is formed by hypervariable residues and can therefore be expected to change significantly from one antibody to another.

There are very few or in some cases no solvent inaccessible water molecules in the interface region, indicating a close complementarity of the interacting surfaces. The packing density in the VH–VL interface does not appear to be significantly different from that in the interior of the protein.[87]

The CH1 and CL domain pair also have a pseudo-twofold axis, which varies from 167° to 173°, a smaller variation than for VH–VL. The interface is formed by the ABED strands (see Figure 2) of the two domains.[2] Padlan et al.[88] examined four Fab structures and found that a few hydrophobic residues buried in the interface make most of the interdomain contacts, and these residues are often invariant or highly conserved. The interface region in these four structures also contains a cavity of between 50 and 150 Å^3. λ light chains have a bulky residue (Phe or Tyr) at position 178, which can rotate into this cavity to relieve packing stresses. The cavity can therefore facilitate the packing of either a λ or κ CL with the appropriate CH1.

The "elbow" angle between the two pseudo-twofold axes relating VL to VH and CL to CH ranges from 130° to 180° in different Fabs, revealing the flexibility within the Fab.[8] Some Fabs crystallize in two or more forms with different elbow angles (for example, the Hy-HEL-5 Fab[37]). In other Fab structures the two molecules in the asymmetric unit of a single crystal have different elbow angles.[27,69] The variation in elbow angles results in changes in the contacts between the V and C domains. The three VH and two CH1 residues involved in these contacts are almost invariant and form a "ball-and-socket" joint.[89] Thus, a 3–4-Å movement of these main-chain VH residues relative to the CH1 residues corresponds to a 30–40° change in elbow angle. These conserved residues form a flexible contact surface that allows wide variation in the elbow angle but also serves to prevent formation of a rigid contact between the V and C domains.

Protein G–Fab Complex. The relative invariance of the CH1 domain provides an attractive target for proteins of wide IgG specificity such as protein G, a small cell-surface protein from *Streptococcus*. The structure of a binding domain of protein G has been determined by NMR[90,91] and in complex with an IgG Fab by crystallography.[64] The protein G domain fold consists of a beta-sheet and a helix. This complex is formed in a most interesting way that might be used by other proteins to bind to constant domains. An outer strand of the sheet forms an antiparallel interaction with the last beta-strand (the G strand) in the CH1 domain of the Fab, thus joining together the two sheets of the Fab and the protein G (Figure 3). In addition to the backbone hydrogen bonds formed by linking the two beta-strands together, other specific H-bonds involve side chains of protein G interacting with backbone atoms of the CH1 domain. Also, the hydrophobic interactions involve conserved residues in different IgG isotypes, thus accounting for the broad specificity of binding.

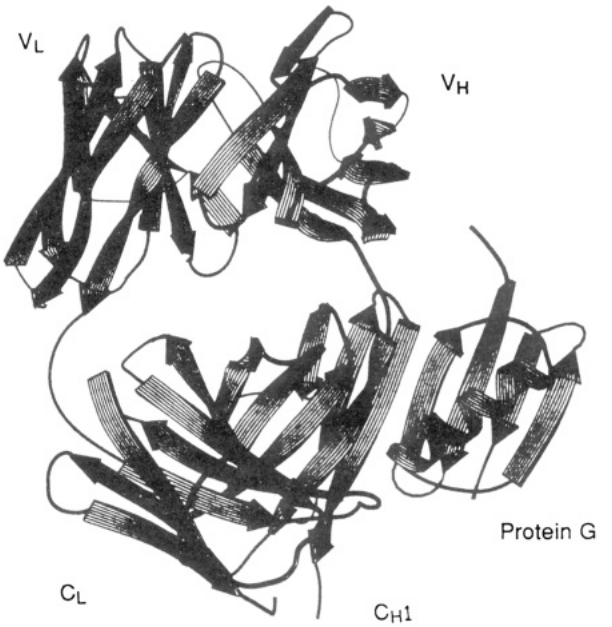

Figure 3. Structure of the Fab–protein G complex. Reproduced with permission from ref 64. Copyright 1992 *Nature*.

The Fc Structure

The Fc is the site of binding for a variety of effector molecules such as C1q and the leukocyte Fc receptor family. These have been reviewed by Burton and Woof[92] and Unkeless et al.[93] and are outside the scope of this Account, although it should be noted that mutant analysis has gone a long way toward defining the binding sites on the Fc for these different proteins.

The structure of human gamma Fc was reported by Deisenhofer[14] at 2.9 Å for the isolated Fc and at 2.8-Å resolution for the Fc complexed with fragment B of protein A. The overall shape of the Fc is that of a rabbit's head, with the two CH2 domains protruding upward from the module formed by the two CH3 domains. The structure of the CH3 domains resembles closely the CH1 (Figure 2b,d) and they are associated in a similar way to the CH1–CL domains of the Fab. The two CH2 domains, although clearly having an immunoglobulin fold, are not directly linked through protein–protein interactions. Instead, there are numerous contacts between the two N-linked carbohydrate chains that are attached to Asn 297 (see Figure 2c) and line the interface between the two domains.

(87) Walls, P. H.; Sternberg, M. J. *J. Mol. Biol.* **1992**, *228*, 227–297.
(88) Padlan, E. A.; Cohen, G. H.; Davies, D. R. *Mol. Immunol.* **1986**, *23*, 951–960.
(89) Lesk, A.; Chothia, C. *Nature* **1988**, *335*, 188–190.
(90) Lian, L-Y.; Yang, J. C.; Derrick, J. P.; Sutcliffe, M. J.; Roberts, G. C.; Murphy, J. P.; Goward, C. R.; Atkinson, T. *Biochemistry* **1991**, *30*, 5335–5340.
(91) Gronenborn, A.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657–661.
(92) Burton, D. R.; Woof, J. M. *Adv. Immunol.* **1992**, *51*, 1–84.
(93) Unkeless, J. C.; Scigliano, E.; Freedman, V. H. *Annu. Rev. Immunol.* **1988**, *6*, 251–281.

There has been a preliminary study of a similar structure for rabbit Fc.[94]

The Fc structures of IgM and IgE differ from the IgG in having three domains per chain. As yet, there has been no crystal structure determination for these Fc, but several modeling studies based on sequence homology with the known structures[95,96] have been made; these have been reviewed by Padlan.[77]

Protein A–Fc Complex. Protein A is a multidomain cell-wall protein from *Staphylococcus aureus* that binds to Fcγ. Fragment B of protein A binds to Fc in the crystal.[14] It consists largely of two alpha-helices which bind in the vicinity of the joint between the CH2 and the CH3 domains. Unlike the protein G interaction with CH1, the protein A interaction involves many side chains on both the Fc and the protein A, thus providing an explanation for the observed lack of binding to human IgG3.[97,98]

Antibody–Antigen Complexes

Protein Antigens. The structures of a number of complexes of monoclonal antibody Fabs with their protein antigens have now been determined. Since the review by Davies et al.,[8] some of these structures have been refined and the data extended to higher resolution. Also for D1.3, the structures of the Fv alone and complexed with lysozyme have been determined at a higher resolution than could be achieved with the Fab.[99,100]

Structures have been reported for three complexes of antibody Fabs with lysozyme[36–38] and two complexes of Fabs with the influenza virus neuraminidase.[39,45] The properties of these complexes have been extensively reviewed[1,5,8,10a,b,101] and will only be summarized here. The complexes are for the most part of high affinity, with K_a of between 10^7 and 10^{10} M^{-1}. The interface region reveals striking shape complementarity between the antibody and antigen, with in most cases few completely buried water molecules, although waters are observed at the periphery of the interface where they are accessible to interaction with bulk solvent. The surface areas removed from solvent accessibility upon complex formation vary from about 1400 Å^2 to about 1800 Å^2. About a dozen hydrogen bonds are found in the interface, together with some salt bridges and a number of van der Waals contacts. Tulip et al.[46] have examined the density of packing in the interface and report that it is significantly lower for the interface of HyHEL-5–lysozyme and the two neuraminidase complexes than in protein interiors. However, a study by Walls and Sternberg,[87] which calculated the ratio between the packing density of each atom and other atoms of the same type, found that the density in the interface region is not significantly different from that in the interior of antibody domains.

Water in the Interface. The role of water molecules in the interface between the antigen and antibody provides an indication of the extent to which the two surfaces are complementary. Presumably, any holes of sufficient magnitude will be occupied by waters. The location of water molecules becomes more certain with the higher resolution of the X-ray data.

Tulip et al.[46] have detected no waters in the neuraminidase–Fab interface that are buried from bulk solvent interactions. Similar reports of few or no buried waters in the interface were made for HyHEL-5 and HyHEL-10 complexed with lysozyme. In the HyHEL-5 complex the two internal waters make hydrogen bond bridges between the two molecules, and in addition there are about seven bridging waters on the periphery of the interface. Fischmann et al.[35] have observed a number of water molecules in the refined 2.5-Å structure of the D1.3 Fab–lysozyme complex that mediate contacts between the antigen and antibody.

Mutant Antigens. Tulip et al.[47] have examined the structures of two neuraminidase mutants complexed with the NC41 antibody. Both mutations (N329D and I368R) were in the antibody–antigen interface region and cause small local changes in the mutation region which decrease the binding affinity by about an order of magnitude. For the N329D mutant, small changes were seen in both the Fab and the neuraminidase. The structure of a mutant of influenza virus hemagglutinin which abolishes binding to a monoclonal antibody has also been reported.[102] Here, too, the replacement of a single residue (Gly) by one with a larger side chain (Asn) caused small local distortions that decreased the binding affinity, in this case to an unmeasurably low level. Similarly, a single-site mutation (K368E) of the N2 neuraminidase resulted in purely local perturbations of the structure but was sufficient to destroy affinity to the S10/1 antibody.[103]

The crystal structure of a mutant lysozyme (R68K) complexed with the Hy-HEL5 Fab has been determined (S.C. and D.R.D., unpublished results). The mutation is in the interface region and results in a 1000-fold decrease in binding affinity. Comparison of the native and mutant complexes shows that there is a local loss of hydrogen bonds associated with the change from arginine to lysine, which can account for the difference in binding affinity. There are no global differences in the other parts of the lysozyme or Fab structure.

An Antibody–Antibody Complex. The structure of a complex between Fab D1.3 and Fab E225 has been reported,[60] where E225 is an anti-idiotope antibody to D1.3, i.e., an antibody to D1.3 that is selected for binding to the variable part of D1.3:

$$\text{E225} \rightarrow \text{D1.3} \rightarrow \text{lysozyme} \quad \text{or} \quad \text{Ab2} \rightarrow \text{Ab1} \rightarrow \text{Ag}$$

Jerne[104] proposed the existence of a network of idiotypes and anti-idiotypes which could be involved in regulation. Anti-idiotypic antibodies could then possess an "internal image" of the antigen with which it would share related epitopes, i.e., the Ag and Ab2 would be structurally related.[105] If Ag binds to a receptor using this epitope, the anti-idiotype could also bind to the

(94) Sutton, B. J.; Phillips, D. C. *Biochem. Soc. Trans.* **1983**, *11*, 130–132.
(95) Pumphrey, R. *Immunol. Today* **1986**, *7*, 174–178.
(96) Padlan, E. A.; Davies, D. R. *Mol. Immunol.* **1986**, *23*, 1063–1075.
(97) Michaelsen, T. E.; Frangione, B.; Franklin, E. C. *J. Immunol.* **1977**, *119*, 558.
(98) Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E. C. *Biochem. Biophys. Res. Commun.* **1976**, *71*, 907.
(99) Boulot, G.; Eisele, J. L.; Bentley, G. A.; Bhat, T. N.; Ward, E. S.; Winter, G.; Poljak, R. A. *J. Mol. Biol.* **1990**, *213*, 617–619.
(100) Bhat, T. N.; Bentley, G. A.; Fischmann, T. O.; Boulot, G.; Poljak, R. J. *Nature* **1990**, *347*, 483–485.
(101) Janin, J.; Chothia, C. *J. Biol. Chem.* **1990**, *265*, 16027–16030.
(102) Knossow, M.; Daniels, R. S.; Douglas, A. R.; Skehel, J. J.; Wiley, D. C. *Nature* **1984**, *311*, 678–680.
(103) Varghese, J. N.; Webster, R. G.; Laver, W. G.; Colman, P. M. *J. Mol. Biol.* **1988**, *200*, 201–203.
(104) Jerne, N. K. *Ann. Immunol. (Paris)* **1974**, *125C*, 373.
(105) Nisonoff, A.; Lamoyi, E. *Clin. Immunol. Immunopathol.* **1981**, *21*, 397–406.

receptor. Examination of the structure of E225 and lysozyme reveals that there is no direct structural resemblance between the antibody site and the epitope on the antigen; i.e., the contact surface on E225 does not mimic the corresponding lysozyme surface. The authors note that since the epitope on lysozyme recognized by D1.3 contains some alpha-helix, this may be hard to mimic in the antibody combining site which consists of the loops from the six CDRs. There are 13 amino acids of D1.3 that make contact with E225. Of these, 10 are CDR residues, 7 of which also contact lysozyme in the D1.3–lysozyme complex. The authors have looked to see whether these residues of D1.3 contact the two proteins (i.e., lysozyme and E225) in the same manner but find that the nature of the contacts (H-bonds, van der Waals contacts, salt bridges) is quite different.

Peptide Antigens. The structures of four complexes between antibodies and their peptide antigens have been reported. They include (a) an antibody to a 19 amino acid peptide from myohemerythrin,[31] where the antibody cross-reacts with myohemerythrin and, more strongly, with apomyohemerythrin; (b) a complex with a nonapeptide from influenza virus hemagglutinin[43] which cross-reacts with the hemagglutinin at low pH; (c) an antibody bound to angiotensin II;[66a] and (d) an antibody complexed to an HIV GP120 peptide.[54,106] For the first two peptides the structures of the native proteins are known, and in both cases the bound peptide conformations differ significantly from their structures in the protein. These results indicate that for these two cases and for the cross-reactivity assays that were used, the antibody recognizes an altered (nonnative) conformation of the protein.

The antibody in (c) was not prepared directly against the angiotensin II peptide. Instead, this antibody (Ab3 in the sequence below) which binds angiotensin II with high affinity ($K_a = 7 \times 10^9$ M^{-1}) is three steps away from the antigen:

$$\text{Ab3} \rightarrow \text{Ab2} \rightarrow \text{Ab1} \rightarrow \text{Ag}$$

The observation that this antibody binds to the original antigen (angiotensin II) strongly suggests that the part of Ab2 that is being recognized by Ab3 is a mimic or internal image of Ag, i.e., of angiotensin II. The conformation observed for the bound angiotensin II is reminiscent of a CDR loop, and the authors report that it resembles in particular the CDR3 loop of REI[16] (the backbone atoms of residues 2–8 of AGII and residues 90–96 of REI superpose with an rms deviation of 0.8 Å). The suggested explanation for the binding of Ab3 to angiotensin II is that the anti-idiotope antibody, Ab2, has a CDR3 that is conformationally very similar to the angiotensin II as bound to its monoclonal antibody, Ab1. The anti-anti-idiotypic antibody, Ab3, binds to this CDR and therefore can also bind to the angiotensin II. It could also be part of the original repertoire of the anti-angiotensin II response. The energetics of peptide binding for this antibody have been investigated by isothermal titration calorimetry.[107] It is observed that binding is favored both enthalpically and entropically, with good agreement between the calculated and experimentally determined thermodynamic quantities.

(106) Stanfield, R. L.; Takimoto-Kamimura, M.; Rini, J. M.; Profy, A. T.; Wilson, I. *J. Mol. Biol.* **1993**, submitted for publication.
(107) Murphy, K. P.; Xie, D.; Garcia, C.; Amzel, L. M.; Freire, E. *Proteins* **1993**, *15*, 113–120.

Carbohydrate Antigens. Carbohydrates are among the most common antigens and have been extensively studied by immunologists. However, until recently there has been no structure of a complex of an Fab with carbohydrate. The structure of J539, an anti-galactan Fab, has been previously reported in the absence of bound ligand,[28] and Cygler et al.[63] have now reported the structure of an Fab, Se155-4, complexed to a tetrasaccharide that is the repeating unit of the polysaccharide *O*-antigen from *Salmonella*. The sequence is

$$\rightarrow 3)\alpha\text{D-Gal}(1\rightarrow 2)[\alpha\text{D-Abe}(1\rightarrow 3)]\alpha\text{D-Man}(1\rightarrow 4)\alpha\text{L-Rha}(1\rightarrow$$

with the 3,6 dideoxy-D-galactose (abequose) forming a branch off the mannose residue. In the complex there is electron density that clearly identifies and positions three sugar residues with no density for the terminal rhamnose. The abequose residue is totally buried in a pocket located between the CDR3 of H and L, where it is anchored by a number of hydrogen bonds, two of which are to a buried water molecule that in turn makes three hydrogen bonds with amino acid residues of the antibody. The other two sugar residues lie on the surface of the combining site. A striking feature of this site, as observed in other antibodies,[77] is the unusual abundance of aromatic residues. The conformation of the trisaccharide is within 14 kJ/mol of the calculated minimum energy conformation.

Haptens and Other Small Antigenic Determinants. The binding of small antigenic determinants such as phosphocholine to McPC603 have been described previously.[8,67,108] Recent antibody complexes with small molecules include BV04-01 with $d(pT)_3$,[33] DB3 with progesterone,[52] Fab 4-4-20 with fluorescein,[32] Fab NQ10/12.5 with 2-phenyloxazolone,[57] Fab AN02 with a DNP spin-label compound,[50,111] and Fab 26-10 with digoxin.[51] These interactions have been summarized by Sheriff[10a,b] in terms of the areas of antibody surface buried in forming the complex, which varies from 156 Å^2 for McPC603 to 389 Å^2 for digoxin. As previously observed for the larger antigens,[8] there does not seem to be any simple correlation between the magnitude of the buried area and the binding constant.

McPC603 binds phosphocholine, and the contacting residues in the combining site have been determined. In 25 mouse phosphocholine antibodies a common VH gene was used and the same residues are invariant or almost invariant. The light chains belong to three different classes, but here again the key CDR3 contacting residues, Tyr-Pro-Leu, are invariant.[109]

The binding site of McPC603 for phosphocholine has been explored by a mutagenesis study.[110] These authors find that almost all of a dozen mutants individually abolish binding, clearly indicating the exquisite design of the site. An exception is the double-mutant D97-(L)L, N101(H)D, which results in a movement of the negative charge while keeping it at the base of the pocket where it can balance the positively charged choline moiety as predicted from the structure.[109]

(108) Davies, D. R.; Metzger, H. *Annu. Rev. Immunol.* **1983**, *1*, 87–117.
(109) Padlan, E. A.; Cohen, G. H.; Davies, D. R. *Ann. Inst. Pasteur/Immunol.* **1985**, *136C*, 271–276.
(110) Glockshuber, R.; Stadlmuller, J.; Pluckthun, A. *Biochemistry* **1991**, *30*, 3049–3054.
(111) Theriault, T. P.; Rule, G. S.; McConnell, H. M. *Protein Structure and Engineering*; Plenum Press: New York, 1990; pp 367–376.

The NQ10/12.5 Fab has been determined with and without bound phenyloxazolone.[57] This antibody belongs to a class of anti-oxazalone antibodies in which one germ line gene for V_κOx1 is used. The identification of the contacting residues permits an evaluation of the contribution of individual somatic point mutations to the maturation of this immune response. Residues L34 and L36, which are sites of mutations associated with increased affinity for 2-phenyloxazolone, are found to interact directly with the hapten.

The structure of the anti-dinitrophenyl-spin-label Fab AN02[50,111] revealed the contacting residues, and for this Fab, too, the germ line sequences are known. Here the majority of the mutated residues are distant from the antigen-binding site and none are in direct contact with the hapten. However, Brunger et al. do not rule out the possibility of long-range effects of some of these mutants and also suggest that two of the changes that occur in the surface of the combining site region may affect the interaction with the carrier protein to which the hapten was coupled.

A complex of an anti-progesterone Fab with and without progesterone[52] shows an interesting change of position of the indole side chain of Trp H100 which can exist in an open and closed conformation. In the absence of progesterone this side chain occupies a cavity created by three aromatic side chains. With progesterone, the Trp H100 side chain is moved to create a hydrophobic pocket for steroid binding.

The structure of cyclosporin complexed to an Fab[56] has been examined. The conformation of the cyclosporin is similar to that observed when it is complexed to cyclophilin and differs from the free cyclosporin conformation.

Conformational Changes on Antigen Binding

Early work on ligand binding to the Mcg L chain dimer led to the suggestion that induced fit might play a role in antibody–antigen interactions. With the increasing number of structures of antibodies free and complexed with antigen, it has become clear that some changes in the antibody structure often accompany binding.[9] There vary from small adjustments of the CDR side chains to quite large movements of VH relative to VL (originally suggested by Colman et al. for the NC41–neuraminidase complex[39]). For the cases where significant changes have been observed, they appear to be of an induced fit nature and result in increased contact between the antibody and antigen. For the D1.3–lysozyme complex,[100] where both Fv and Fab have been crystallized with and without antigen, small differences in the antibody are observed upon binding lysozyme, in particular, a quite small movement of VH relative to VL. There appear to be no significant changes in the CDRs.

The other examples come from small antigenic determinants with varied results. McPC603–phosphocholine shows no noticeable difference between the bound and unbound states, but here the conclusion is not unequivocal because of the presence of a sulfate group (from the crystallization medium) in the binding site at the position otherwise occupied by the phosphate of the phosphocholine. In the anti-ssDNA antibody BV04-01[32] there is a significant movement of the VH relative to VL as well as large changes in three CDRs. These changes open up the combining site when $d(pT)_3$ is bound.

Several peptide-binding Fabs have been examined by Wilson et al. For Fab 17/9 they observe no change in relative positions of VH and VL, but there is a major change in the H-CDR3 without which the peptide would not be able to bind to the antibody.[43] For B13I2 there is a small rigid body movement of H-CDR3.[24] For Fab 50.1 there are changes in H3 and H1 and a strikingly large rotation ($\sim$16°) of VH relative to VL.[106]

The likelihood of VH movement relative to VL may depend on the size of the VH–VL interface,[106] and these authors suggest that this area of contact may be in turn related to the size of the H-CDR3. This loop, as noted earlier, has considerable variation in length, which does show some correlation with the size of the VH–VL interface. Sequence studies by Kabat and Wu[112] show that VH appears to be more important than VL for determining antibody specificity, and of the H-CDRs, H-CDR3 varies the most in antibodies of different specificity. Analysis of the interface between antibody and antigen[8,10] also shows that in all cases but one, VH makes a larger contact with the antigen than VL.

The Use of Protein–Fab Complexes to Promote Crystallization. The protein crystallographer is frequently confronted with the situation in which the protein to be analyzed cannot be crystallized in a suitable form, perhaps because of insolubility or aggregation (for example, membrane proteins or viral capsid proteins). A strategy that might avoid these problems is to prepare a monoclonal antibody against the protein and then to attempt to crystallize the complex, which can sometimes crystallize more readily than the isolated protein. This maneuver has been employed successfully for the HIV reverse transcriptase,[113] where many previous crystallizations had led to poorly diffracting crystals. It should be noted, however, that in this case the same protein was also crystallized in a suitable form by binding an appropriate inhibitor.[114]

Conclusion

Considerable progress has been made since Ehrlich in understanding the molecular basis of antibody specificity and function. Nevertheless, in view of the overwhelming diversity of the immune system, more structural information will be required to fully understand the nature of the antibody–antigen interaction and the underlying principles governing these interactions.

We gratefully thank Eduardo Padlan for many helpful discussions and suggestions during the preparation of this Account.

(112) Kabat, E.; Wu, T. T. *J. Immunol.* **1991**, *147*, 1709–1719.
(113) Arnold, E.; Jacobo-Molina, A.; Nanni, R. G.; Williams, R. L.; Lu, X.; Ding, J.; Clark, A. D., Jr.; Zhang, A.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H. *Nature* **1992**, *357*, 85–89.
(114) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783–1790.