1 The Building Blocks of the Complement System

Dimitrios Morikis and John D. Lambris

CONTENTS

I. OVERVIEW

The complement was discovered more than a century ago (1895) and its proteins were purified to homogeneity thereafter. However, the first three-dimensional structure of a complement component appeared in the literature in 1980. This was the crystal structure of a fragment of human C3a and desArg-C3a by Huber et al.[1] determined by x-ray diffraction. Subsequently, the solution structure of human C5a derived from nuclear magnetic resonance (NMR) spectroscopy was published by Zuiderweg et al. in 1989.[2] In 1990, two three-dimensional structures of complement proteins were deposited in the Protein Data Bank (PDB, then the Brookhaven PDB, since renamed Research Collaboratory for Structural Bioinformatics; see *www.rcsb.org/pdb/*).[3] These were the solution three-dimensional structures of porcine desArg-C5a,[4] and the 16th module of human factor H.[5] We do not include published structural studies dealing with only secondary structure determination by NMR that were popular during that time, which coincides with the early steps of multidimensional biomolecular NMR development. The number of deposited complement protein structures was 25 at the end of the 1990s, and increased to 56 at year-end 2003. In addition, the structures of 2 complement inhibitors, 4 viral complement regulators, and at least 12 complement-like proteins were deposited in the PDB at the end of 2003. Five of the complement or complement-related structures are low-resolution

models; three were determined with the use of x-ray and neutron solution scattering and two by cryoelectron microscopy. Not all of the deposited structures are unique or complete. Some are complexes, multimers, fragments, domains or modules, mutants, structures with various bound ligands or inhibitors, structures with various sample or crystallization conditions (e.g., pH), structures of proteins crystallized in various space groups, proteins from various species, or NMR structures deposited twice as an ensemble of low-energy structures and an averaged minimized structure. Several articles in the literature describe structure–function relations, based on the deposited structures.

Now, for the first time, we have a critical and representative number of complement protein structures that provide insights into the structural biology of the complement system. The goal of this book is to accomplish the task of collective presentation in a single source for the available structures of complement components, regulators, receptors, related proteins, and selected inhibitors, and to discuss structure–function correlations. We hope that this publication will provide the basis for additional theoretical work for researchers who are working (both experimentally and theoretically) in the areas of protein dynamics, protein–protein interactions, signaling, and design of inhibitors with potential to become therapeutic agents. We believe that this publication will provide momentum and excitement for further work in structure determination of complement proteins and protein complexes. Several years of structural work are forecast, owing to the complexity and multiplicity of complement interactions. Future studies will aid in obtaining a better picture of the structure that underlies complement function.

II. STRUCTURAL VIEW OF THE COMPLEMENT ACTIVATION PATHWAYS

Table 1.1 presents a summary of atomic resolution complement structures deposited at the PDB, together with their PDB codes, species, physical methods for structure determination, and original references. All cascades of complement activation, the alternative, classic, lectin, and common pathways, are structurally represented in the PDB.

Figure 1.1 shows a representation of the protein–protein interactions involved in complement activation. Initiation of complement activation occurs by the binding of the C1 complex (C1qsrrs) to antigen–antibody complexes (classical pathway), the binding of mannose-binding lectin (MBL) to carbohydrates on the surfaces of pathogens (lectin pathway), and the spontaneous activation–inactivation reaction (tick-over activation) of C3, generation of C3b, and binding of C3b to pathogen surfaces (alternative pathway). Four enzymes called convertases are essential for complement activation. These are C3 convertases C3bBb and C4b2a, and C5 convertases C3bBb3b and C4b2a3b. Convertases are composed of fragments of C2, C3, C4, and factor B. C2 and factor B are complement proteases. The role of convertases is to cleave C3 and C5, generating anaphylatoxins C3a and C5a, and components C3b and C5b that carry on the complement activation pathways. Functionally, the end results of complement activation are (a) inflammatory response, which directs

TABLE 1.1
Structures of Complement Proteins Deposited with the Protein Data Bank

Protein[a,b]	PDB Code	Species	Method	Reference
		Nonprotease Complement Components		
C1q	1PK6	*Homo sapiens*	X-ray	6
C1q	1C28	*Mus musculus*	X-ray	7
C3d	1C3D	*Homo sapiens*	X-ray	8
C3d (in complex with CR2)	1GHQ	*Homo sapiens*	X-ray	9
C3dg	1QQF	*Rattus norvegicus*	X-ray	10
C3dg	1QSJ	*Rattus norvegicus*	X-ray	10
C4Adg	1HZF	*Homo sapiens*	X-ray	11
C5a	1KJS	*Homo sapiens*	NMR	12
C5a-desArg	1C5A	Porcine	NMR	4
C8γ	1IW2 (pH 7), 1LF7 (pH 4)	*Homo sapiens*	X-ray	13
		Complement Proteases		
C1r zymogen, SP-CCP modules	1MD7	*Homo sapiens*	X-ray	14
C1r-C1r zymogen, SP-CCP-CCP modules	1GPZ	*Homo sapiens*	X-ray	15
C1r active catalytic domain, SP-CCP modules	1MD8	*Homo sapiens*	X-ray	14
C1r, EGF-like module	1APQ	*Homo sapiens*	NMR	16
C1s, Catalytic domain, SP-CCP	1ELV	*Homo sapiens*	X-ray	17
C1s, CUB1-EGF domain	1NZI	*Homo sapiens*	X-ray	18
Factor B, SP domain	1DLE	*Homo sapiens*	X-ray	19
Factor B, VWF module	1Q0P	*Homo sapiens*	X-ray	20
Factor D	1DSU, 1HFD, 1BIO, 1DFP, 1DIC, 1DST, 1FDP	*Homo sapiens*	X-ray	21–26
MASP-2, CUB1-EGF-CUB2	1NT0	*Rattus norvegicus*	X-ray	27
		Complement Receptors and Regulators		
CR1 (CD35)	1GKG, 1GKN, 1PPQ	*Homo sapiens*	NMR	28,29
CR2 (CD21)	1LY2	*Homo sapiens*	X-ray	30
CR2 (CD21) (in complex with C3d)	1GHQ	*Homo sapiens*	X-ray	9

(continued)

TABLE 1.1 (CONTINUED)
Structures of Complement Proteins Deposited with the Protein Data Bank

Protein[a,b]	PDB Code	Species	Method	Reference
Factor H	1HCC, 1HFH, 1HFI	*Homo sapiens*	NMR	5,31
Factor H	1HAQ	*Homo sapiens*	X-ray and neutron solution scattering and homology modeling	32
DAF (CD55)	1H03, 1H04, 1H2P, 1H2Q, 1OJV, 1OJW, 1OJY, 1OK1, 1OK2, 1OK3, 1OK9, 1UOT	*Homo sapiens*	X-ray	33,34
DAF (CD55)	1NWV	*Homo sapiens*	NMR	35
DAF (CD55) (in complex with an echovirus)	1M11, 1UPN	*Homo sapiens*	Cryoelectron microscopy	36,37
MCP (CD46)	1CKL	*Homo sapiens*	X-ray	38
CD59	1CDQ, 1CDR, 1CDS, 1ERG, 1ERH	*Homo sapiens*	NMR	39,40

MASP-2, MBL-associated serine protease; MBL, mannose-binding lectin; CR1, complement receptor 1; CR2, complement receptor 2; DAF, decay-accelerating factor; MCP, membrane cofactor protein; SP, serine protease; VWF, von Willebrand factor; CCP, complement control protein; EGF, epidermal growth factor; CUB, complement-urchin-bone.

[a] We do not discriminate among fragments or components, homomultimers, presence of glycans, ions, or other molecules, and amino acid insertions or deletions, but we indicate heterocomplexes and the C1r-C1r dimer. A complete description of the structures can be found in the Protein Data Bank.[3]
[b] Additional entries of complement-related proteins and viral and semi-synthetic or synthetic complement inhibitors, reviewed in this volume, can be found on the companion CD.

Source: Protein Data Bank (www.rcsb.org/pdb/).

immune system cells and molecules at the points of infection; (b) opsonization, the process of coating foreign pathogens for recognition and elimination by phagocytes; (c) direct lysis of pathogen cells by disrupting the hydrophobicity of the lipid bilayers of pathogen cell membranes; and (d) aid to the adaptive immune system by participating in both T- and B-cell responses (acting as a link between innate and adaptive immunity). In addition, complement prevents the formation of large immune complexes and helps eliminate them from circulation, through the processes of opsonization and phagocytosis. End proteins of complement action are C3a, C4a, C5a, C3b, C4b, C3d, C4d, and $C5b678(9)_n$ (MAC) (Figure 1.1).

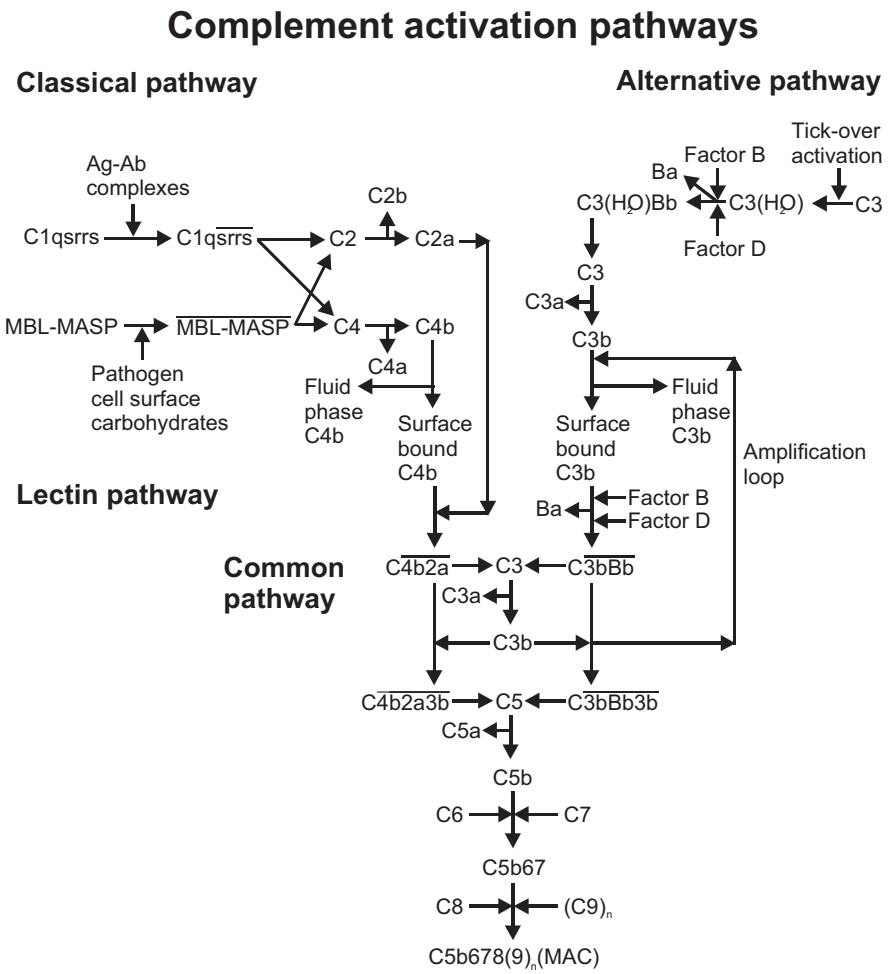

FIGURE 1.1 Cascades of complement activation pathways. All pathways (classical, lectin, alternative, and common) are shown. The overline bars denote activated complexes.

Figure 1.2A shows the natural regulators of complement activation (RCAs) and the complement components they act on. Complement regulation is important to direct complement action against non-self agents and in essence to discriminate non-self from self. When complement regulation breaks down, autoimmune conditions appear. Figure 1.2B shows schematically the stepwise processes of the regulation of the complement system. Much of regulation of complement activation occurs on, but is not limited to, the C3 convertase enzymes. Figure 1.3 shows the cell membrane–bound complement receptors and the complement components that they interact with.

Figure 1.4 shows a graphical representation of the structures involved in complement activation, albeit at coarse detail in terms of interactions, for those proteins with available structures (Table 1.1). Figure 1.4 shows that several complement proteins are modular and some are homo- or heteromultimers. Modules without currently available structures are represented with simple geometric shapes in

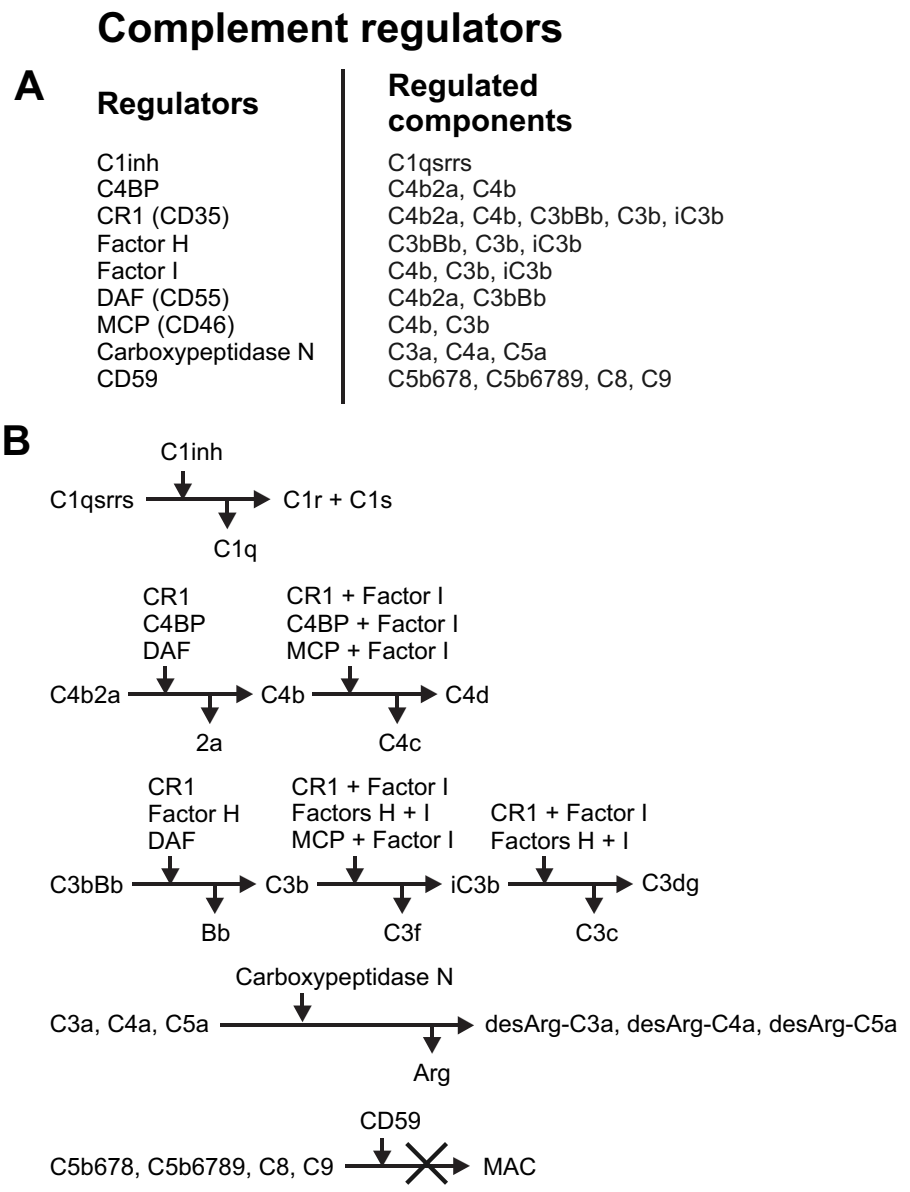

FIGURE 1.2 (A) Summary of regulators of complement activation (RCA) and their corresponding regulated complement proteins or protein complexes. (B) Cascades of complement regulation.

Figure 1.4. The same applies in the case of homomultimers, where only structures of a single monomer are shown in Figure 1.4. Heteromultimers are not overlaid in Figure 1.4, for simplicity. The choice of the interactions shown in Figure 1.4 was

Complement receptors

Cell-bound receptors	Interacting components
CR1 (CD35)	C3b, iC3b, C4b
CR2 (CD21)	iC3b, C3dg, C3d
CR3 (CD11b/18)	iC3b
CR4 (CD11c/18)	iC3b
C3aR	C3a
C5aR (CD88)	C5a
C1qR	C1q, MBL

FIGURE 1.3 Summary of cell membrane–bound complement receptors and their interacting complement proteins.

based on the availability of structures when we were preparing this manuscript. Figure 1.4 does not discriminate fragments from whole proteins, nor does it depict sequence alterations (e.g., mutations, additions, deletions), or the presence of solvent atoms, ions, glycans, or other small molecules. Here we use the terms 'domain' and 'module' indiscriminately when this is appropriate, according to Campbell's[42] definition for domains, modules, and repeats: "Domain is a compact structural unit in a protein; the amino acid sequence need not be contiguous. Modules are a subset of domains; they are contiguous in sequence, and are repeatedly used as building blocks in functionally diverse proteins; they have identifiable amino acid patterns that can be described by a consensus sequence. A repeat is a sequence unit that does not occur as a single copy; several repeats are needed to form a superstructure."* The following groups of complement components (and their respective fragments or modules) are structural and functional homologues: (C1r, C1s), (C2, factor B), (C3, C4), (C3bBb, C4b2a), (C3bBb3b, C4b2a3b), and (C3a, C4a, C5a) (Figures 1.1 and 1.4).

Figure 1.5 shows a graphical representation of available representative structures of RCAs and complement receptors, which interact with complement proteins. In the case of proteins with several repeats of complement control protein (CCP) modules, only the longest sequential fragment is shown in Figure 1.5. The CCP modules are also known as short consensus repeat (SCR) modules.

A comprehensive review of CCP modules is given by Soares and Barlow in Chapter 2. The discussion in Chapter 2 goes beyond structure and structure–function correlations. Soares and Barlow provide a bioinformatics analysis of sequences and structures, computational structural homology modeling, and evolutionary analysis of CCP modules.

The classical pathway is represented by the structure of components C1q and C4Ad and proteases C1r and C1s (Figure 1.4, Table 1.1). C1q, C1r, and C1s, and the modeling of their complexes are reviewed by Arlaud, et al. in Chapter 3. C4Ad is reviewed by Isenman and van den Elsen in Chapter 5.

The alternative pathway is represented by the structures of components C3a and C3d and proteases factor D and factor B (Figure 1.4; Table 1.1). factors B and D are reviewed by Xu et al. in Chapter 4. C3d is reviewed by Isenman and van den Elsen in Chapter 5 and the complex of C3d with complement receptor 2 (CR2 or

* Reproduced with permission from I.D. Campbell (2003). *Biochem. Soc. Trans.* 31, 1107–1116.

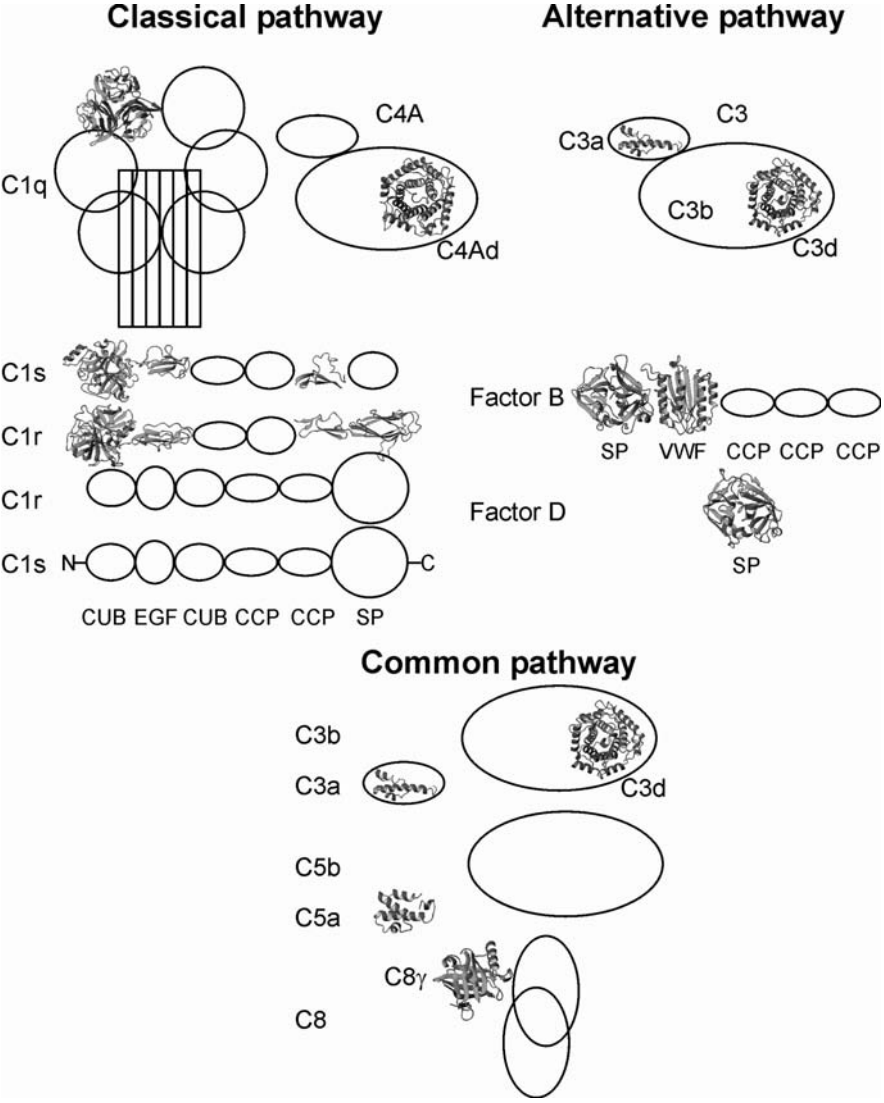

FIGURE 1.4 A structural view of the complement activation pathways. This figure was made with currently available structures of complement proteins, protein fragments, or protein complexes. References and codes of PDB files are given in Table 1.1. Molecular graphics were prepared with the program MOLMOL[41] in all figures.

CD21) is reviewed by Hannan et al. in Chapter 6. C3a is reviewed by Morikis et al. in Chapter 7.

The lectin pathway is not reviewed because at the time we prepared this manuscript structures had not been deposited at the PDB. Later, the structure of a fragment of MBL-associated serine protease (MASP-2) was determined,[27] which is shown in Figure 1.6 (Table 1.1). MASP-2 is a structural and functional homologue of C1r, C1s (Figures 1.4 and 1.6). Structural and functional homologues also include MBL and C1q.

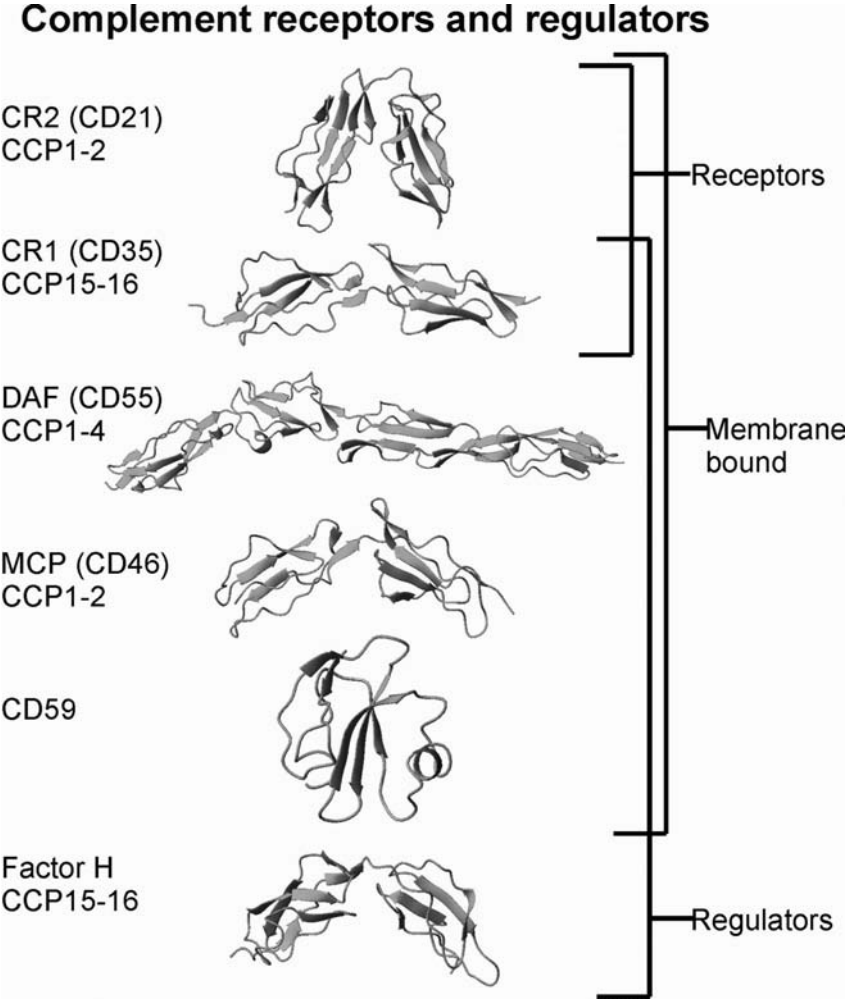

FIGURE 1.5 A structural view of complement receptors (CRs) and regulators of complement activation (RCAs). This figure was made with currently available structures of CR and RCA fragments. References and codes of PDB files are provided in Table 1.1.

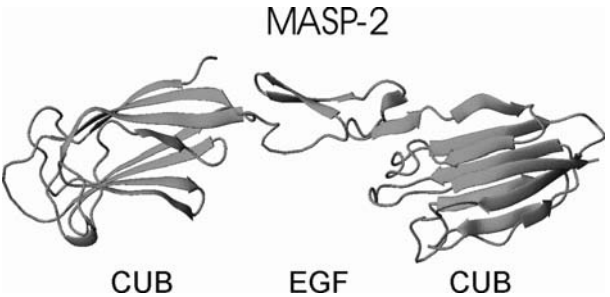

FIGURE 1.6 The structure of the CUB1-EGF-CUB2 modules of rat mannose-binding protein associated serine protease-2 (MASP-2;[27] PDB code 1NT0).

The common pathway is represented here by the structures of C3a, C5a, and C8γ (Figure 1.4, Table 1.1). C3a and C5a are reviewed by Morikis et al. in Chapter 7. C8γ is reviewed by Lebioda and Sodetz in Chapter 10.

The RCAs are represented by the structure of factor H, complement receptor 1 (CR1 or CD35), decay-accelerating factor (DAF or CD55), membrane cofactor protein (MCP or CD46), and CD59 (protectin). Membrane proteins interacting with complement proteins are represented by CR1 (CD35), CR2 (CD21), DAF (CD55), MCP (CD46), and CD59. Complement receptors are represented by CR1 (CD35), CR2 (CD21), and complement-like repeats CR3, CR7, CR8, LB5, and Tva. The CCP modules are the most common building blocks of the reviewed RCAs, membrane proteins, and receptors (Figure 1.5). The structures of factor H, MCP (CD46), and viral complement regulator vaccinia control protein (VCP) are reviewed by Soares and Barlow in Chapter 2. The structure of CR1 (CD35) is reviewed by Krych-Goldberg et al. in Chapter 8. The structure of CR2 (CD21) free and in complex with C3d is reviewed by Hannan et al. in Chapter 6. The structure of DAF (CD55) is reviewed by Lukacik et al. in Chapter 9. The structure of the extracellular region of CD59 is reviewed by Morgan and Tomlinson in Chapter 11. The structures of complement-like repeats CR3, CR7, CR8, LB5, and Tva are reviewed by Dolmer and Gettins in Chapter 12. This chapter also discusses the effect of Ca^{2+} binding to complement-like repeats using fluorescence emission spectroscopy and isothermal titration calorimetry.

Low-resolution structures of several complement proteins, RCAs, antibodies, and complexes are reviewed by Perkins and Furtado in Chapter 13. This chapter demonstrates the utility of synchrotron x-ray and neutron solution scattering, analytical ultracentrification, and constrained molecular modeling for coarse structure determination.

Finally, the structures of two classes of peptide complement inhibitors, which are promising to be developed into therapeutics, are presented together with the paths that led to their discovery, design, applications, and optimization of activity. These are compstatin, a C3 inhibitor, and its analogues and the cyclic antagonists of human C5a receptor (C5aR or CD88). Compstatin and its analogues are reviewed by Morikis and Lambris in Chapter 14. This chapter also discusses the conformational interconversion of compstatin using molecular dynamics simulations. The cyclic antagonists of human C5aR (CD88) are reviewed by Taylor and Fairlie in Chapter 15. The selection of the compstatin analogues and C5aR antagonists for inclusion in this volume was based on the fact that they are peptides and the process of their activity optimization demonstrates the power of structural biology in rational drug design. A massive literature exists for complement inhibitors, recently reviewed in Lambris and Holers.[43]

III. STRUCTURAL CLASSIFICATION OF COMPLEMENT PROTEINS

Branden and Tooze[44] and Lesk[45] demonstrate early efforts for the classification of protein structure into families of structural motifs. However, complement structures are absent from these treatises. With the increase of deposited structures at the PDB, several resources for classification schemes and structural analyses have been

developed and are available on the Web. The PDB (*www.rcsb.org/pdb/*),[3] Swiss PDB (*www.expasy.org/*),[46] and the National Center for Biotechnology Information (NCBI) (*www3.ncbi.nlm.nih.gov/*) provide links to a large array of excellent resources for protein structure analysis, classification, and visualization.

We have chosen to present the CATH domain classification for the complement system. A tool for understanding structure (*www.biochem.ucl.ac.uk/bsm/*),[47] CATH classifies domains of high-resolution crystal structures (less than 3 Å) or NMR structures of proteins. CATH presents four main hierarchical levels: class (C-level), architecture (A-level), topology (T-level), and homologous superfamily (H-level).[47] Further levels of classification also exist. Class provides secondary structure information for the domain and can be mainly alpha (meaning α-helical), mainly beta (meaning β-sheet), alpha-beta (meaning mixed α-helical and β-sheet), or low secondary structure. Architecture provides information on the spatial arrangement of elements of secondary structure that form the shape of the domain, without taking into account connectivity of elements of secondary structure. Topology or fold family provides information on the spatial arrangement and connectivity of elements of secondary structure that form the shape of the domain. Homologous superfamily cluster structures according to sequence similarity and structural homology. These domains are thought to have a common ancestor.[47]

Table 1.2 shows the CATH classification of complement structural motifs. According to CATH, there are six defined building blocks (and one irregular) in terms of architecture, and nine building blocks in terms of topology for the 16 complement proteins (with 56 entries in the PDB) reviewed here (Table 1.1). These proteins belong to 10 structurally homologous superfamilies (Table 1.2). In certain instances in Table 1.2, we used the alternative protein classification method SCOP (Structural Classification of Proteins).[48] With the exception of MASP-2, the structures in Tables 1.1 and 1.2 are reviewed in subsequent chapters.

IV. BEYOND STRUCTURE: EXAMPLES OF STRUCTURE-BASED STUDIES FOR COMPLEMENT AND COMPLEMENT-RELATED PROTEINS OR INHIBITORS

The determination of three-dimensional structures of complement proteins has been the basis for the elucidation of protein complex formation and the study of conformational dynamics by NMR and molecular dynamics simulations. Computational modeling, based on three-dimensional structures, has also been used for electrostatic calculations aimed at predicting the ionization properties of complement proteins and the driving forces for protein–protein recognition and association. These studies include structural modeling of C1 from its components (Chapter 3),[49] NMR studies that address backbone mobility for CCPs of CR1 (CD35) and MCP (CD46)[29] and VCP,[50] molecular dynamics studies for compstatin (Chapter 14),[51] electrostatic calculations for C3d-CR2 association,[52] docking studies for DAF (CD55) with von Willebrand factor type A domain of factor B (Chapter 9),[34] and docking studies of a C5aR antagonist peptide to C5aR (CD88).[53]

TABLE 1.2
CATH Classification of Complement Structures[a]

Protein	PDB Code	Class	Architecture	Topology	Homologous Superfamily
		Nonprotease Complement Components			
C1q	1C28, 1PK6	Mainly beta	Sandwich	Jelly rolls	Lymphokine
C3d	1C3D, 1GHQ	Mainly alpha	Alpha/alpha barrel	Glycosyltransferase	Isomerase
C4Adg	1HZF				
C5a, C5a-desArg	1KJS, 1C5A	Mainly alpha	Up-down bundle	Influenza virus matrix protein; chain A, domain 1	Complement factor
C8γ	1IW2, 1LF7	Mainly beta	Barrel	Serratia metallo proteinase inhibitor, subunit I	Retinol transport
		Complement proteases			
C1r, SP module	1GPZ, 1MD7, 1MD8	Mainly beta	Barrel	Thrombin, subunit H	Trypsin-like serine proteases
C1s, SP module	1ELV				
Factor B, SP module	1DLE				
Factor D	1DSU, 1HFD, 1BIO, 1DFP, 1DIC, 1DST, 1FDP				
C1r, CCP module	1GPZ, 1MD7, 1MD8	Mainly beta	Ribbon	Complement module; domain 1	Complement module, domain 1
C1s, CCP module	1ELV				
C1r, EGF-like module	1APQ	Mainly beta	Ribbon	Laminin	Laminin
C1s, EGF-like module	1NZI				
MASP-2, EGF-like module	1NT0				
C1s, CUB-like module[a]	1NZI	Mainly beta	Sandwich	Jelly rolls	Spermadhesin
MASP-2, CUB-like module[a]	1NT0				
Factor B, VWF module[a]	1Q0P	Alpha-beta	3-Layer (aba) sandwich	Rossmann fold	Cell adhesion

(continued)

TABLE 1.2 (CONTINUED)
CATH Classification of Complement Structures[a]

Protein	PDB Code	Class	Architecture	Topology	Homologous Superfamily
Complement Receptors and Regulators					
CR1 (CD35)	1GKG, 1GKN, 1PPQ	Mainly beta	Ribbon	Complement module; domain 1	Complement module; domain 1
CR2 (CD21)	1LY2, 1GHQ				
Factor H	1HCC, 1HFH, 1HFI				
DAF (CD55)	1H03, 1H04, 1OJV, 1NWV				
MCP (CD46)[a]	1CKL				
CD59	1CDQ, 1CDR, 1CDS, 1ERG, 1ERH	Mainly beta	Ribbon	CD59	CD59

[a] When CATH[47] classification was not available, we used the SCOP[48] classification or the CATH classification of homologous structures.

Source: Protein Data Bank (www.rcsb.org/pdb/).

V. PERSPECTIVE

Significant progress has been made in the determination of the three-dimensional structure of complement components, regulators, receptors, and inhibitors. However, the structural biology of the complement system is far from complete. Table 1.3 shows a summary of selected complement proteins that will aid the reader in identifying samples for structural studies. Since protein size and complexity are issues in structural studies, the molecular mass and number of chains are also given in Table 1.3.

The power and utility of structure determination to develop structure–function or structure–activity correlations have been demonstrated in all chapters of this volume. Knowledge of structure brings us closer to understanding function, when compared to sequence alone. This is because structure allows us to discern the physicochemical principles that underlie biological processes. However, proteins are not static, as seen in pictures generated by time-averaged or ensemble-averaged three-dimensional representations of their structures. Indeed, proteins are dynamic molecules — they shake, twist, and swing. Proteins experience a variety of local, segmental, and global motions at a variety of time scales. Knowledge of structure allows us to understand the protein dynamics. In combination, knowledge of protein structure and dynamics allows us to understand protein–protein association, protein–ligand binding, and other interactions. Finally, knowledge of protein structure, dynamics, and interactions brings us closer to understanding the physicochemical

TABLE 1.3
Size and Complexity of Proteins in Complement Research[a]

Protein	Approximate Molecular Mass (kDa)	Number of Chains
C1q	410	18
C1r	83	1
C1s	85	1
C2	102	1
C3	190	2
C4	205	3
C5	196	2
C6	125	1
C7	120	1
C8	150	3
C9	166	1
Factor B	100	1
Factor D	24	1
Properdin	224	4
MBL	540	18
MASP-1	94	1
MASP-2	76	1
C1inh	105	1
C4BP	550	7
Factor H	150	1
Factor I	100	2
CR1 (CD35)	160–250	1
CR2 (CD21)	140	1
CR3 (CD11b/18)	265	2
CR4 (CD11c/18)	245	2
DAF (CD55)	70	1
MCP (CD46)	45–70	1
C1qR	65	1
C3aR	54	1
C5aR (CD88)	45	1
CD59	20	1
Carboxypeptidase N	280	4

[a] Most entries derive from *The Merck Manual of Diagnosis and Therapy*, section 12, chapter 146, Merck & Co., 1995–2004 (also available at: *www.merck.com/mrkshared/mmanual/tables/146tb2.jsp* and *www.merck.com/mrkshared/mmanual/tables/146tb3.jsp*).

basis of function. The physicochemical origins of function provide quantitation and understanding beyond simple (yet important) biological observation, and they form the starting point for the prediction of protein properties, and eventually of protein function, and for the design of pharmaceutical regulators in cases of defective function. In the absence of experimental data or when incomplete experimental data

are available for dynamics and interactions, biomolecular simulation can be a significant tool toward the prediction of the properties of proteins and protein complexes. Biomolecular simulation at the atomic level is possible only when three-dimensional structures are available as starting points. The structures provide the atomic three-dimensional coordinates that form the spatial arrangement of the biomolecular building blocks and the spatial arrangement of the physicochemical properties of the building blocks.

VI. EPILOGUE

In conclusion, we intended to assemble a volume containing the current state-of-the-art of the structural biology of the complement system with structure–function correlations. Judging from the number of unsolved complement structures and the rate of new structures deposited at the PDB, we predict that additional structures will be determined in the near future to identify the missing links in multidomain proteins. We also expect the determination of more protein complex structures that mediate protein interactions which are responsible for recognition, binding, signaling, and function. The structural biology of the complement system provides an unexplored fertile ground for the study of dynamics and thermodynamics, both experimentally and theoretically. On the theoretical end, we expect that computational simulation of complement protein dynamics will provide insights on protein motions at time scales inaccessible by experiment. We also expect that calculations of the energetics of complex formation and stability will provide insights on the thermodynamics involved in these processes.

ACKNOWLEDGMENTS

This work was supported by grants AI-30040 and GM-56698 from the National Institutes of Health. We thank Dr. Philippe Valadon for providing the molecular structure visualization program RasTop, and Dr. Helen Berman for giving us permission to use the PDB coordinate files.

SUPPLEMENTARY MATERIAL ON CD

The companion CD contains the following: (a) a complete set of figures, including color figures, when available, and corresponding figure captions, (b) all PDB coordinate files[3] for the structures discussed in the book, (c) additional PDB files and references for complement-related proteins and viral, semi-synthetic, or synthetic complement regulators or inhibitors, reviewed in this volume, and (d) the biomolecular structure visualization program RasTop (P. Valadon, RasTop, version 2.0.3, Philippe Valadon, La Jolla, CA, 2000–2003; *www.geneinfinity.org/ rastop/*),[54] which reads and displays PDB coordinate files. RasTop is based on RasMol (R. Sayle. RasMol, version 2.6, Biomolecular Structures Group, Glaxo Wellcome Research & Development, Stevenage, United Kingdom, 1992–1999; *www.openrasmol.org/*).[55]

REFERENCES

1. R Huber, H Scholze, EP Paques, J Deisenhofer. *Hoppe-Seylers Z Physiol Chemie*, 361:1389–1399, 1980.
2. ERP Zuiderweg, DG Nettesheim, KW Mollison, GW Carter. *Biochemistry*, 28:172–185, 1989.
3. HM Berman, J Westbrook, Z Feng, G Gilliland, TN Bhat, H Weissig, IN Shindyalov, PE Bourne. *Nucleic Acids Res*., 28:235–242, 2000.
4. MP Williamson, VS Madison. *Biochemistry*, 29:2895–2905, 1990.
5. DG Norman, PN Barlow, M Baron, AJ Day, RB Sim, ID Campbell. *J. Mol. Biol*., 219:717–725, 1991.
6. C Gaboriaud, J Juanhuix, A Gruez, M Lacroix, C Darnault, D Pignol, D Verger, JC Fontecilla-Camps, GJ Arlaud. *J. Biol. Chem*., 278:46974–46982, 2003.
7. L Shapiro, PE Scherer. *Curr. Biol*., 8:335–338, 1998.
8. B Nagar, RG Jones, RJ Diefenbach, DE Isenman, JM Rini. *Science*, 280:1277–1281, 1998.
9. G Szakonyi, JM Guthridge, DW Li, K Young, VM Holers, XJS Chen. *Science*, 292:1725–1728, 2001.
10. G Zanotti, A Bassetto, R Battistutta, C Folli, P Arcidiaco, M Stoppini, R Berni. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol*., 1478:232–238, 2000.
11. JMH van den Elsen, A Martin, V Wong, L Clemenza, DR Rose, DE Isenman. *J. Mol. Biol*., 322:1103–1115, 2002.
12. XL Zhang, W Boyar, MJ Toth, L Wennogle, NC Gonnella. *Proteins Struct. Function Genet*., 28:261–267, 1997.
13. E Ortlund, CL Parker, SF Schreck, S Ginell, W Minor, JM Sodetz, L Lebioda. *Biochemistry*, 41:7030–7037, 2002.
14. M Budayova-Spano, W Grabarse, NM Thielens, H Hillen, M Lacroix, M Schmidt, JC Fontecilla-Camps, GJ Arlaud, C Gaboriaud. *Structure*, 10:1509–1519, 2002.
15. M Budayova-Spano, M Lacroix, NM Thielens, GJ Arlaud, JC Fontecilla-Camps, C Gaboriaud. *EMBO J*., 21:231–239, 2002.
16. B Bersch, JF Hernandez, D Marion, GJ Arlaud. *Biochemistry*, 37:1204–1214, 1998.
17. C Gaboriaud, V Rossi, I Bally, GJ Arlaud, JC Fontecilla-Camps. *EMBO J*., 19:1755–1765, 2000.
18. LA Gregory, NM Thielens, GJ Arlaud, JC Fontecilla-Camps, C Gaboriaud. *J. Biol. Chem*., 278:32157–32164, 2003.
19. H Jing, YY Xu, M Carson, D Moore, KJ Macon, JE Volanakis, SVL Narayana. *EMBO J*., 19:164–173, 2000.
20. AA Bhattacharya, JML Lupher, DE Staunton, RC Liddington. *Structure*, 12:371–378, 2004.
21. SVL Narayana, M Carson, O Elkabbani, JM Kilpatrick, D Moore, X Chen, CE Bugg, JE Volanakis, LJ Delucas. *J. Mol. Biol*., 235:695–708, 1994.
22. H Jing, YS Babu, D Moore, JM Kilpatrick, XY Liu, JE Volanakis, SVL Narayana. *J. Mol. Biol*., 282:1061–1081, 1998.
23. LB Cole, NM Chu, JM Kilpatrick, JE Volanakis, SVL Narayana, YS Babu. *Acta Crystallogr. D Biol. Crystallogr.*, 53:143–150, 1997.
24. LB Cole, JM Kilpatrick, NM Chu, YS Babu. *Acta Crystallogr. D Biol. Crystallogr.*, 54:711–717, 1998.
25. S Kim, SVL Narayana, JE Volanakis. *J. Biol. Chem*., 270:24399–24405, 1995.
26. H Jing, KJ Macon, D Moore, LJ DeLucas, JE Volanakis, SVL Narayana. *EMBO J*., 18:804–814, 1999.

27. H Feinberg, JCM Uitdehaag, JM Davies, R Wallis, K Drickamer, WI Weis. *EMBO J.*, 22:2348–2359, 2003.
28. BO Smith, RL Mallin, M Krych-Goldberg, XF Wang, RE Hauhart, K Bromek, D Uhrin, JP Atkinson, PN Barlow. *Cell*, 108:769–780, 2002.
29. JM O'Leary, K Bromek, GM Black, S Uhrinova, C Schmitz, XF Wang, M Krych, JP Atkinson, D Uhrin, PN Barlow. *Protein Sci.*, 13:1238–1250, 2004.
30. AE Prota, DR Sage, T Stehle, JD Fingeroth. *Proc. Natl. Acad. Sci. U.S.A.*, 99:10641–10646, 2002.
31. PN Barlow, A Steinkasserer, DG Norman, B Kieffer, AP Wiles, RB Sim, ID Campbell. *J. Mol. Biol.*, 232:268–284, 1993.
32. M Aslam, SJ Perkins. *J. Mol. Biol.*, 309:1117–1138, 2001.
33. P Williams, Y Chaudhry, IG Goodfellow, J Billington, R Powell, OB Spiller, DJ Evans, S Lea. *J. Biol. Chem.*, 278:10691–10696, 2003.
34. P Lukacik, P Roversi, J White, D Esser, GP Smith, J Billington, PA Williams, PM Rudd, MR Wormald, DJ Harvey, MDM Crispin, CM Radcliffe, RA Dwek, DJ Evans, BP Morgan, RAG Smith, SM Lea. *Proc. Natl. Acad. Sci. U.S.A.*, 101:1279–1284, 2004.
35. S Uhrinova, F Lin, G Ball, K Bromek, D Uhrin, ME Medof, PN Barlow. *Proc. Natl. Acad. Sci. U.S.A.*, 100:4718–4723, 2003.
36. YN He, F Lin, PR Chipman, CM Bator, TS Baker, M Shoham, RJ Kuhn, ME Medof, MG Rossmann. *Proc. Natl. Acad. Sci. U.S.A.*, 99:10325–10329, 2002.
37. D Bhella, IG Goodfellow, P Roversi, D Pettigrew, Y Chaudhry, DJ Evans, SM Lea. *J. Biol. Chem.*, 279:8325–8332, 2004.
38. JM Casasnovas, M Larvie, T Stehle. *EMBO J.*, 18:2911–2922, 1999.
39. CM Fletcher, RA Harrison, PJ Lachmann, D Neuhaus. *Structure*, 2:185–199, 1994.
40. B Kieffer, PC Driscoll, ID Campbell, AC Willis, PA Vandermerwe, SJ Davis. *Biochemistry*, 33:4471–4482, 1994.
41. R Koradi, M Billeter, K Wuthrich. *J. Mol. Graph.*, 14:51–55, 1996.
42. ID Campbell. *Biochem. Soc. Trans.*, 31:1107–1114, 2003.
43. JD Lambris, VM Holers, Eds. *Therapeutic Interventions in the Complement System*. Humana Press, Totowa, NJ, 2000.
44. C Branden, J Tooze. *Introduction to Protein Structure*, 2nd ed., Garland Publishing, New York, 1999.
45. AM Lesk. *Introduction to Protein Architecture*. Oxford University Press, Oxford, 2001.
46. E Gasteiger, A Gattiker, C Hoogland, I Ivanyi, RD Appel, A Bairoch. *Nucleic Acids Res.*, 31:3784–3788, 2003.
47. CA Orengo, AD Michie, S Jones, DT Jones, MB Swindells, JM Thornton. *Structure*, 5:1093–1108, 1997.
48. AG Murzin, SE Brenner, T Hubbard, C Chothia. *J. Mol. Biol.*, 247:536–540, 1995.
49. GJ Arlaud, C Gaboriaud, NM Thielens, M Budayova-Spano, V Rossi, JC Fontecilla-Camps. *Mol. Immunol.*, 39:383–394, 2002.
50. CE Henderson, K Bromek, NP Mullin, BO Smith, D Uhrin, PN Barlow. *J. Mol. Biol.*, 307:323–339, 2001.
51. B Mallik, JD Lambris, D Morikis. *Proteins Struct. Function Genet.*, 53:130–141, 2003.
52. D Morikis, JD Lambris. *J. Immunol.*, 172:7537–7547, 2004.
53. BO Gerber, EC Meng, V Dotsch, TJ Baranski, HR Bourne. *J. Biol. Chem.*, 276:3394–3400, 2001.

54. P Valadon. RasTop version 2.0.3, La Jolla, CA: Philippe Valadon, 2000–2003.
55. R Sayle. RasMol 2.6, Stevenage, UK: Biomolecular Structures Group, Glaxo Wellcome Research & Development, 1992–1999.