Complement regulation at the molecular level: The structure of decay-accelerating factor


*Laboratory of Molecular Biophysics, 1Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England; †Adprotech Ltd., Chesterford Research Park, Saffron Walden, Essex CB10 1XL, England; ‡Faculty of Biomedical and Life Sciences, Division of Virology, University of Glasgow, Church Street, Glasgow G11 5JW, Scotland, and ‌Department of Medical Biochemistry and Immunology, University of Wales College of Medicine, Health Park, Cardiff CF14 4XN, Wales

Communicated by Douglas T. Fearon, University of Cambridge, Cambridge, United Kingdom, November 5, 2003 (received for review September 18, 2003)

The human complement regulator CD55 is a key molecule protecting self-cells from complement-mediated lysis. X-ray diffraction and analytical ultracentrifugation data reveal a rod-like arrangement of four short consensus repeat (SCR) domains in both the crystal and solution. The stalk linking the four SCR domains to the glycosylphosphatidylinositol anchor is extended by the addition of 11 highly charged O-glycans and positions the domains an estimated 177 Å above the membrane. Mutation mapping and hydrophobic potential analysis suggest that the interaction with the convertase, and thus complement regulation, depends on the burial of a hydrophobic patch centered on the linker between SCR domains 2 and 3.

CD55 | complement regulator | pathogen receptor | glycoprotein

The complement system is an important component of the immune response, consisting of more than 30 proteins that function together to provide an initial defense against invasion by pathogens. The classical, alternative, and lectin pathways (CP and AP for classical and alternative pathways, respectively), each activated by different stimuli, converge by means of a series of enzyme-linked cascades to target cells for destruction. Regulation to prevent inappropriate activation against self occurs by means of the proteins encoded in the regulators of complement activation gene clusters, which act at key points in the cascades to prevent pathological consequences. Decay-accelerating factor (CD55) is a member of the regulators of complement activation protein family (1), and its primary function is to inactivate the C3 convertases by dissociating them into their constituent proteins (reviewed in ref. 2). The importance of regulation is highlighted by the large number of human diseases that relate to inappropriate complement activation (3). Complement is also responsible for the primary rejection events in xenotransplantation; therefore, transgenic animals expressing human CD55 potentially provide a route to the generation of readily available organs that will resist rejection. Extended survival times have already been achieved for organs transplanted from human CD55-transgenic pigs to primates (4), and soluble CD55 has been demonstrated to block the Arthus reaction in vivo (5). Incorporation of CD55 into the envelope of baculovirus has proved to be a promising technology for prolonging the lifetime of virions used in genetic therapy (6). CD55 has additional roles, including acting as a binding partner for CD97 (7), a molecule whose expression is up-regulated on leukocytes during inflammatory activation. The significance of the CD97–CD55 interaction is little understood, but it has been implicated in the pathogenesis of multiple sclerosis, because CD97 and CD55, which are absent from normal white matter, are found at high levels in multiple sclerosis lesions (8). CD55 is expressed at a high level on all serum-exposed cells. Perhaps as a consequence of this expression, CD55 has been subverted by many bacterial and viral pathogens which exploit it as a receptor to facilitate cellular infection (reviewed in ref. 9).

CD55 has a C-terminal glycosylphosphatidylinositol (GPI) anchor and consists of a serine/threonine/proline-rich region and four consecutive, membrane-distal, short consensus repeat (SCR) domains characteristic of regulators of complement activation family proteins (reviewed in ref. 2). CD55 is heavily O-glycosylated in the serine/threonine/proline region and contains an N-linked glycan located between SCR domains 1 and 2. To date, our molecular understanding of the biological functions of CD55 has been derived primarily from mutagenesis studies (10, 11) and two structures of CD55 SCR domain pairs (10, 12). These structures suggested that the interactions between the regulator (CD55) and the convertases were likely to involve large areas on the surface of CD55 (the CP convertase contacting much of SCR domains 2 and 3, whereas the AP convertase contacts SCR domains 2, 3, and 4). These interactions were postulated to involve wrapping of the large convertase complexes around the smaller regulator, leading to contacts on both faces of CD55. However, mapping of the key interaction site (as defined by mutagenesis; refs. 10 and 11) for both convertases has been complicated by the fact that the site is centered on residues located at the junction between SCR domains 2 and 3. On the basis of our 1.7-Å x-ray structure of SCR domains 3 and 4 (10), we proposed that many of the residues previously suggested to be the direct contacts of CD55 with the convertases were instead acting indirectly by altering the structure of CD55 at the SCR 2/3 domain interface. Uhrinova et al. (12) reached different conclusions on the basis of their solution structure of SCR domains 2 and 3, in which the key residues were seen to be solvent exposed and not structurally involved in stabilizing the SCR 2/3 domain interface. These inconsistencies and other studies that have shown that dissection of domains from their natural biological context can result in structural perturbation (13) demonstrated that a structure for the four linked SCR domains was required to fully understand the biology of CD55.

Materials and Methods


Abbreviations: SCR, short consensus repeat; CP, classical pathway; AP, alternative pathway; GPI, glycosylphosphatidylinositol; vWF-A, von Willebrand factor type A domain.

Data Deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1OJV, 1OJW, and 1OJY).


§Present address: Astex Technology, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, England.

To whom correspondence should be addressed. E-mail: susan@biop.ox.ac.uk.

© 2004 by The National Academy of Sciences of the USA
Refolded CD55,1234 was further purified by gel filtration to separate folded and unfolded species. Activity of the pure protein was assayed with CP and AP assays (J.W., P.L., D.E., M. Steward, N. Giddings, J. R. Bright, B.P.M., S.M.L., G.P.S., and R.A.G.S., unpublished data).

**Crystallization and Structure Determination.** Crystals of CD55,1234 were grown at 4°C from mother liquor consisting of 0.2 M ammonium sulfate; 20% (wt/vol) monomethoxypolyethylene glycol 5000; 0.1 M sodium acetate, pH 4.6; and 10% (wt/vol) glycerol. The largest crystals were obtained by seeding with crushed crystals. Data were collected from flash-frozen crystals cryoprotected by addition of glycerol to a final concentration of 18%. Data were collected from flash-frozen crystals cooled by liquid nitrogen or by using a freezer. The largest crystals were obtained by seeding with crushed crystals. Data were collected from flash-frozen crystals cryoprotected by addition of glycerol to a final concentration of 18%. Data were collected from flash-frozen crystals cooled by liquid nitrogen or by using a freezer.

Data were processed with either the CCP4 suite programs MOSFLM and SCALA (14) or the HKL suite of programs (15). Molecular replacement (MOLREP; ref. 16) using the coordinates for 50% of the molecule (CD55; ref. 10) found two copies of CD55,1234 in crystal forms A and B. Refinement of these partial models (BUSTER-TNT; refs. 17 and 18) in both crystal forms allowed location of Pt and Au sites in both crystal forms (DMMULTI; ref. 19). Eventually, a complete model was built into maps derived by combining (SIGMAA; ref. 20) experimental and model phases for each crystal form and between the two crystal forms (DMMULTI; ref. 19). The extracted glycans were labeled with the fluorophore 2-amino-3-benzamide by reductive amination (26) and processed through normal-phase HPLC (27). Preliminary assignments for the glycans were confirmed by weak anion exchange HPLC (28), digestions with arrays of exoglycosidases (A), and matrix-assisted laser desorption ionization–time-of-flight MS. Detailed methods, Fig. 4, and Table 3 are published as supporting information on the PNAS web site.

The degree of O-linked glycosylation was estimated from the reduction in mass (as determined by equilibrium analytical ultracentrifugation) of the intact glycoprotein on digestion with neuraminidase combined with the known sequences of the glycans. This method gave an average of 11 O-linked glycans per molecule.

Molecular modeling was performed on a Silicon Graphics Fuel workstation by using the programs INSIGHT2 and DISCOVER (Accelrys, San Diego). N-linked and O-linked glycans were built by using average crystallographic torsion angles for the glycosidic linkages (29, 30). The conformation of the Asn-GlcNac linkage at site 63 was based on average crystallographic values (31). The likely positions of the 11 O-linked glycans were determined by using the program O-LINK (www.cbs.dtu.dk/services/netoglyc), the sites with the highest probability being used. The degree of extension of the tail peptide was assumed to be 2.6 Å per residue, based on the experimentally determined values for the O-glycosylated tail peptide of CD8 (32) and bovine and porcine submaxillary mucins (33, 34).

**Results and Discussion**

**Structure of CD55,1234 in the Crystal and in Solution.** We have demonstrated activity, crystallized (J.W., P.L., D.E.,...
M. Steward, N. Giddings, J. R. Bright, B.P.M., S.M.L., G.P.S., and R.A.G.S., unpublished data), and solved the structure at 2.3 Å (Tables 1 and 2) of all four SCR domains of CD55 (hereafter referred to as CD55_{1234}). The structure of CD55_{1234} is revealed as a linear molecule with the four SCR domains forming an extended rod with overall dimensions of 160 × 50 × 30 Å (Fig. 1a). CD55_{1234} appears to be a relatively rigid molecule as judged by the analysis of three differently packed (Fig. 1 and Table 2) crystal forms, which provide eight independent models of the molecule. Fig. 1b shows two views of an overlay of the eight copies of the molecule, demonstrating that, despite the different crystal packing environments, there is very little variation in the structure of the domain interfaces. Inflexible intermodular junctions have been noted for other multiple SCR domain-containing structures, e.g., the vaccinia complement control protein (35), suggesting that rigidity may be one of the key characteristics of multiple-SCR structures. In particular, the organization of the CD55 SCR 2/3 domain interface is constant across the crystals (Fig. 1b), as demonstrated by the consistent values for the tilt and twist angles (Table 2) that define the orientation of SCR domain 2 with respect to SCR domain 3 for the different copies of the molecule. We can therefore confidently interpret the contribution of residues that influence complement regulation that map to this region. The apparent rigidity seen in this structure (and that observed in our earlier structure of the CD55_{34} fragment; ref. 10) contrasts with the weakly defined module interface in the solution structure of CD55_{34} (12). We have therefore used analytical ultracentrifugation to determine the sedimentation coefficient (s) of CD55_{1234}, which gives an indication of the domain arrangement in solution. Our experimental value for s0, averaged from six independent measurements of s at a range of concentrations, is 2.04 ± 0.13 S. This value compares well with a theoretically determined value of 2.03 S calculated from our coordinates for CD55_{1234} (using the program HYDROPRO; ref. 36). Therefore, the structure reported here from the crystalline environment is a good representation of the architecture present in solution.

**Characterization of the Glycosylation of Erythrocyte-Derived CD55 Allows Construction of a Model for the Complete CD55.** Physiologically, CD55 is heavily glycosylated with both N- and O-linked glycans and a GPI anchor (37). The specific nature of the glycosylation is influenced by the tissue of origin (38). However, because erythrocyte CD55 represents the most abundant and well characterized species, we have targeted it for full chemical characterization.

The structures of the N- and O-linked sugars released from CD55, purified from human red blood cells, were analyzed by a combination of HPLC, exoglycosidase digestions, and MS. The major N-linked oligosaccharides located at Asn-63 in the SCR 1/2 domain interface are galactosylated biantennary complex type sugars, with and without bisecting GlcNAc and terminal sialic acid residues. There are 19 predicted O-linked glycosylation sites (www.cbs.dtu.dk/services/netogly) in the 68-aa, serine/threonine/proline-rich stalk region of CD55. On average, 11 of these sites are occupied per molecule. Seventy-one percent of the O-linked glycans pool consisted of the disialylated core-1 structure Neu5Acα2–3Galβ1–3(Neu5Acα2–6)GalNAc. The stalk region is therefore highly charged. INSIGHTII and DISCOVER (Accelrys) were used to build models of representative glycans onto the crystallographic model for CD55_{1234} by using average crystallographic torsion angles for the glycosidic linkages (29). Earlier analysis of both mucins (39) and CD8 (32) indicate that the addition of GalNAc to multiple serine or threonine residues extends and stiffens the protein backbone. Based on these data, the CD55 serine/threonine/proline-rich region was modeled assuming a 2.6-Å extension per amino acid. This model gives an overall length of 177 Å for the stalk, which is in agreement with earlier data indicating that functional CD55 depends on the SCR domains being separated from the membrane (40). Together with the crystallographic structure of CD55_{1234} and the GPI anchor structure (37), these data provide a full description of the overall architecture of CD55 (Fig. 1c).

Removal of the N-linked glycan is known to have no effect on complement regulation (41), a finding that can now be rationalized, because the N-linked sugar is located on the opposite face to, and directed away from, the major sites of convertase interactions (see discussion below). The N-glycan is predicted to stabilize the structure (42) and, in common with the O-glycans, may serve to protect the protein from proteases.
CD55 Regulates Complement by Means of Hydrophobic Interactions.

The structure for CD55_{234} described herein agrees closely with the previous x-ray structures of CD55_{34} (10) (rms deviation C° positions are 0.5 Å for SCR domains 3 and 4 overlaid on the basis of SCR domain 3 alone). In contrast, CD55_{1234} deviates significantly from the solution structures of CD55_{23} (12), particularly in its relative arrangement of SCR domains 2 and 3, which is not shared by any of the CD55_{23} models deposited in the Protein Data Bank (mean rms deviation C° for SCR domains 2 and 3 overlaid on the basis of SCR domain 3 is 10.4 ± 3.6 Å). Earlier analysis (12) of these solution structures for CD55_{23} suggested that a band of positive charge encircling the 2/3 domain junction was a prominent feature of this molecule and was probably required for functional activity. The major biological functions map to the SCR 2/3 domain interface (11, 12); therefore, significant structural differences between our structure for the intact molecule and these earlier solution structures for the fragment leads to a different interpretation of mutagenesis-derived data.

Fig. 2a shows a representation of the electrostatic surface calculated for the structure of the four SCR domains. Although this surface shows a patch of positive charge encircling the upper portion of the SCR 2/3 domain junction, the dominant features are two negatively charged patches. These patches encompass much of the surface of SCR domain 1 as well as one face of the two membrane-proximal SCR domains (3 and 4). Also prominent is a band of uncharged residues encircling the SCR 2/3 domain junction at the N terminus of SCR domain 3. The dramatic difference between this and the earlier analysis presumably reflects the large difference in both module and specific side-chain orientations between our current and earlier models of CD55.

Fig. 2b and c shows a mapping of residues previously linked to the decay-accelerating activities of CD55. As for the earlier partial structures, mutations affecting CP convertase regulation (Fig. 2b) map to residues in SCR domains 2 and 3, and those effecting AP convertase regulation (Fig. 2c) also involve SCR domain 4. The distribution of these mutations suggests that regulation is effected through an extended interaction between CD55 and the convertases that involves much of the surface of the regulator. In contrast to the earlier structure of CD55_{23}, we find that the critical arginine and lysine residues implicated (12) in convertase recognition are relatively buried in the 2/3 domain interface. We therefore propose that the phenotype conferred by alteration of these residues is the result of structural perturbation, rather than the direct ablation of an active site. In addition, Phe-148 and Leu-171 (previously thought to perform largely structural roles; ref. 12), lie at the center of the uncharged band (also consisting of Phe-169, Ile-172, and Pro-97) encircling the junction of SCR domains 2 and 3.

Fig. 2. Mapping biological function onto the structure of CD55. Two views of CD55 are shown related by a 180° rotation about the vertical axis. The view in Upper is the same view as that shown in Fig. 1a Right. (a) Charge distribution on the surface of CD55. Charges were calculated by using an AMBER force-field in GRASP (52) and colored evenly around 0 from blue (positive) to red (negative) via white at no charge. (b) Mapping of amino acid positions where substitution is known to affect the ability of CD55 to regulate the CP convertases (10, 11). Positions where amino acid substitution is known to reduce activity to 10% or less of wild-type activity are shown in dark blue; those that reduce activity to 50% or less are shown in light blue. (c) Mapping of amino acid positions where substitution is known to affect the ability to regulate the AP convertases (10, 11). Positions where amino acid substitution is known to reduce activity to 10% or less of wild type are shown in dark green; those that reduce activity to 50% or less are shown in light green.

Table 2. Refined model statistics

<table>
<thead>
<tr>
<th></th>
<th>Native1</th>
<th>Native4</th>
<th>Native5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range used (outer shell), Å</td>
<td>16–2.3 (2.37–2.3)</td>
<td>25–2.8 (2.87–2.8)</td>
<td>25–2.6 (2.67–2.6)</td>
</tr>
<tr>
<td>R_{work} (outer shell), %</td>
<td>20.7 (22.7)</td>
<td>25.3 (28.8)</td>
<td>25.2 (27.0)</td>
</tr>
<tr>
<td>R_{free} (outer shell), %</td>
<td>25.2 (26.7)</td>
<td>28.8 (32.9)</td>
<td>29.7 (30.3)</td>
</tr>
<tr>
<td>Rmsd bond lengths, Å/rmsd bond angles, †</td>
<td>0.008/0.8</td>
<td>0.005/0.6</td>
<td>0.005/0.7</td>
</tr>
<tr>
<td>No. of molecules in asymmetric unit</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Waters in model</td>
<td>213</td>
<td>37</td>
<td>143</td>
</tr>
<tr>
<td>Heteroatoms in model glycerol/3O/1Ac</td>
<td>9/2/3</td>
<td>2/2/0</td>
<td>6/3/4</td>
</tr>
<tr>
<td>Tilt/twist SCR domains 1 and 2, **</td>
<td>A, 41/86; B, 41/90</td>
<td>A, 34/174; B, 31/169</td>
<td>A, 31/171; B, 30/171; C, 32/163; D, 34/164</td>
</tr>
<tr>
<td>Tilt/twist SCR domains 3 and 4, **</td>
<td>A, 8/241; B, 6/252</td>
<td>A, 12/248; B, 11/258</td>
<td>A, 10/255; B, 11/258; C, 13/248; D, 14/251</td>
</tr>
<tr>
<td>PDB identifier</td>
<td>1OJY</td>
<td>1OJW</td>
<td>1OJY</td>
</tr>
</tbody>
</table>

All models have no outliers in a Ramachandran plot; chains are denoted A–D as required for each crystal form. Iterative model building and refinement led to the models described and deposited in the Protein Data Bank (PDB) with the identifiers shown. rmsd, rms deviation.

*Reference datasets in heavy-atom phasing.
†Tilt/twist defined by Norman et al. (51). Domain descriptions are as follows: SCR domain 1, Cys-4 to Cys-62; SCR domain 2, Cys-66 to Cys-126; SCR domain 3, Cys-131 to Cys-188; and SCR domain 4, Cys-193 to Cys-251.

CD55 Regulates Complement by Means of Hydrophobic Interactions.
CD55 and the vWF-A domain of factor B generated as described in the text. CD55 is oriented as in Fig. 1 and Phe-169 near the SCR 2 data suggest that this hydrophobic patch (Fig. 3) to Ala did not abolish the ability to regulate the CP convertase. Our sites of direct convertase interaction, because mutation of Phe-169 alone. Leu-171 and Phe-169 have not previously been identified as mutations at the Phe-169 position lower regulation of the AP convertase CD55 to regulate the activity of both convertases, whereas mutation of the vWF-A domain known to abolish factor B–CD55 interaction as distance filters (vWF-A mutation sites must lie within 3 Å of any part of CD55) to select potential complexes. The calculation yielded a set of four solutions with essentially identical domain orientations for the complex. Fig. 3 shows an analysis of the most favorable solution. This proposed complex has some compelling features: the CD55–vWF-A interaction site is formed across the SCR 2/3 domain interface and obscures the hydrophobic patch that we postulate forms a portion of the convertase interaction site. Also, the CD55–vWF-A interaction masks many of the CD55 mutations (10, 11) known to alter AP convertase regulation. This model for the complex leaves factor B peptides, previously implicated in binding of the C3b portion of the convertase (44), exposed and accessible for interaction with the larger C3b; equally exposed are vWF-A residues, where mutation (46) is known to have no effect on the ability of CD55 to regulate the AP convertase.

The mechanisms of decay acceleration by CD55 of the C3

---

**Fig. 3.** Model of the complex between CD55 and the von Willebrand factor type A domain (vWF-A) of factor B. a and d show two views of the complex between CD55 and the vWF-A domain of factor B generated as described in the text. CD55 is oriented as in Fig. 1. Left. b shows a surface representation, and d shows a secondary structure trace for both molecules, which are colored as follows: CD55, the hydrophobic potential was calculated by using the program oesdx and mapped onto the surface in asees (43) colored from green (most favorable for hydrophobic probe interaction) via yellow to white; and vWF-A of factor B, residues implicated in CD55 binding are colored red, and those implicated in C3b binding are colored orange; positions where mutation has no effect on CD55 sensitivity are colored blue. a and c show each molecule rotated 90° away from the view shown in b to allow inspection of the surfaces buried in the interface. CD55 is shown in the same view as that of Fig. 1 Right.

Lukacik et al.
convertases are poorly understood at the molecular level. Binding studies using individual components and AP convertases immobilized on zymosan showed that CD55 had low affinities for the individual components proteins C3b and Bb but much increased affinities for both the C3bB and C3bBb complexes (47). Despite similar affinities for the complexes, CD55 accelerates decay only of the active C3bBb convertase and not its precursor C3bB (48). The C3bB complex may be stabilized by the reported interactions between sites in Ba and C3b (49). These limited data suggest that CD55 must bind sites in both of the components of the convertase. Bound CD55 tends to cause dissociation of the convertase but succeeds only after cleavage of B with release of Ba and consequent weakening of the association between convertase components.

With dissociation of the convertase, affinity for CD55 is lost, releasing the regulator. The solving of the structure of CD55 opens the door to structural studies that will enable testing of these hypothetical models for decay acceleration.

This study provides a complete structural description of an intact and active regulator of complement activation family protein. CD55 is of major therapeutic interest not only in xenotransplantation but also as a potential biopharmaceutical in its own right. We have also described elsewhere (J.W., P.L., D.E., M. Steward, N. Giddings, J. R. Bright, B.P.M., S.M.L., G.P.S., and R.A.G.S., unpublished data) an approach to the functional reconstruction of the GPI anchor by C-terminal modification of complement regulators, including CD55, using a membrane-binding peptide. This greatly increases potency and enables their use in settings where local retention of a soluble exogenously administered complement regulator may be therapeutically important (50).

Solving the structure of CD55, also advances our ability to model its critical interactions with Bb and C2a and thence with convertases. This structure may provide a route to the rational design of synthetic CD55-mimetics, a goal that has been hampered by the lack of structural information on the complex and unstable convertases.

We thank the staff of the European Synchrotron Radiation Facility protein crystallography beams (ID14-1, ID14-2, ID14-3, ID14-4, and ID29), the Medical Research Council beamline (BM14), and the staff of the Synchrotron Radiation Source, Daresbury (stations 9.6, 14.1, 14.2, and 7.2); most of the Laboratory of Molecular Biophysics, Oxford, in particular Ed Lowe, Martin Noble, and Elspeth Garman for assistance with data collection (in-house and at synchrotrons); Russell Wallis for assistance with analytical ultracentrifugation; the staff at Adprotec for assistance in CD55 production; Jim McDonnell for thoughtful reading of this manuscript at an early stage; Gérard Bricogne for the use of a β-version of Buster-TNT; and the Biotechnology and Biological Sciences Research Council for the provision of the matrix-assisted laser desorption ionization–time-of-flight MS instrument in the Glycobiology Institute. This work was funded by a Medical Research Council Case studentship (to P.L.), Biotechnology and Biological Sciences Research Council Grant 43/B16601 (to S.M.L.), Arthritis Research Campaign Grant L0534 (to S.M.L.), and Wellcome Trust Grants GR068590MA (to B.P.M.) and 059011 (to S.M.L. and D.J.E.).