THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 22, pp. 15821–15829, May 31, 2013

A Molecular Switch Governs the Interaction between the Human Complement Protease C1s and Its Substrate, Complement C4*

Received for publication, February 25, 2013, and in revised form, April 7, 2013 Published, JBC Papers in Press, April 16, 2013, DOI 10.1074/jbc.M113.464545

Andrew J. Perry[‡1,2], **Lakshmi C. Wijeyewickrema**[‡1], **Pascal G. Wilmann**[‡], **Menachem J. Gunzburg**[‡], **Laura D'Andrea**[‡], **James A. Irving**[‡2,3], **Siew Siew Pang**[‡], **Renee C. Duncan**[‡4], **Jacqueline A. Wilce**[‡], **James C. Whisstock**[‡§5,6], **and Robert N. Pike**[‡5,7]

From the [‡]*Department of Biochemistry and Molecular Biology and the* [§]*Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton, Melbourne, Victoria 3800, Australia*

Background: In the classical pathway of the complement system, activated C1s cleaves C4.
Results: C4 binding parameters and the crystal structure of CCP1-CCP2-SP of C1s zymogen are reported.
Conclusion: C1s must be activated, with repositioning of two loops in its SP domain, before it can bind C4.
Significance: Even when the SP of C1s zymogen is exposed, it cannot bind C4.

The complement system is an ancient innate immune defense pathway that plays a front line role in eliminating microbial pathogens. Recognition of foreign targets by antibodies drives sequential activation of two serine proteases, C1r and C1s, which reside within the complement Component 1 (C1) complex. Active C1s propagates the immune response through its ability to bind and cleave the effector molecule complement Component 4 (C4). Currently, the precise structural and biochemical basis for the control of the interaction between C1s and C4 is unclear. Here, using surface plasmon resonance, we show that the transition of the C1s zymogen to the active form is essential for C1s binding to C4. To understand this, we determined the crystal structure of a zymogen C1s construct (comprising two complement control protein (CCP) domains and the serine protease (SP) domain). These data reveal that two loops (492–499 and 573–580) in the zymogen serine protease domain adopt a conformation that would be predicted to sterically abrogate C4 binding. The transition from zymogen to active C1s repositions both loops such that they would be able to interact with sulfotyrosine residues on C4. The structure also shows the junction of the CCP1 and CCP2 domains of C1s for the first time, yielding valuable information about the exosite for C4 binding located at this position. Together, these data provide a structural explanation for the control of the interaction with C1s and C4 and, furthermore, point to alternative strategies for developing therapeutic approaches for controlling activation of the complement cascade.

The complement system acts as a first line of defense against viral and microbial pathogens (1), but it also plays a central role in inflammatory and autoimmune diseases (2). Initiation of the classical pathway is controlled by the proteases C1r and C1s that form part of the C1 complex. When the C1q component of the complex recognizes ligands, such as antigen-antibody complexes, C1r is autoactivated and subsequently cleaves and activates the zymogen C1s. The primary substrates of active C1s are C2 and C4. These molecules, once cleaved, propagate activation of the complement cascade through to the terminal membrane attack complex (3, 4).

C1s and C1r share common domain architecture (5), containing, from the N terminus, a CUB1-EGF-CUB2 region, followed by two complement control protein modules (CCP1 and CCP2)[8] and a serine protease (SP) domain (6). The CUB1-EGF region in each protease mediates key Ca^{2+}-dependent interactions between C1s, C1r, and C1q (7, 8), whereas the CCP modules play an important role in the interaction between the proteases and their substrates (9, 10). The binding and cleavage of C4 by activated C1s is a key event in propagating and amplifying the activation of the classical pathway of complement: therefore, it is crucial that the binding of C4 by the enzyme is highly regulated to prevent inappropriate activation of the pathway. In addition, understanding the interaction between C1s and C4 is important, not least because blocking the interaction between C1s and C4 may permit therapeutic control of the complement system. In a previous study, we identified a novel exosite for C4 on the C1s SP domain (11). The exosite was centered on loop D

* This work was supported by the National Health and Medical Research Council of Australia and the Australian Research Council.
The atomic coordinates and structure factors (code 4J1Y) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] These authors contributed equally to the work.
[2] National Health and Medical Research Council training fellows.
[3] Present address: Dept. of Medicine, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, CB2 0XY, UK.
[4] Present address: Center for Virology, Burnet Institute, Melbourne, VIC, Australia.
[5] Joint senior and corresponding authors.
[6] Australian Research Council Federation Fellow and an honorary National Health and Medical Research Council Principal Research Fellow. To whom correspondence may be addressed: Dept. of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia. Tel.: 61-3-99029300; Fax: 61-3-99029500; E-mail: James.Whisstock@monash.edu.
[7] To whom correspondence may be addressed: Dept. of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia. Tel.: 61-3-99029300; Fax: 61-3-99029500; E-mail: Rob.Pike@monash.edu.

[8] The abbreviations used are: CCP, complement control protein-like domain; SP, serine protease domain; MASP, mannose-binding lectin-associated serine protease; SPR, surface plasmon resonance; RMSD, root mean square deviation.

(Trp^{570}–Leu^{582}) of C1s, a region that undergoes conformational change in other serine proteases when the zymogen is activated. Most recently, the structure of C4 in complex with MASP-2 (the lectin pathway homologue of C1s) has provided molecular details of this interaction (12), revealing that two key exosites govern MASP-2 interaction with C4; the first is located at the junction between the CCP1 and CCP2 domains, and the second is located on the SP domain.

We postulated that the C1s zymogen-to-active transition may itself represent an important control point that governs interaction between the C1 complex and C4. To test these ideas, we performed surface plasmon resonance (SPR) experiments aimed at understanding the interaction between C1s and C4. We then interpreted these results by determining the crystal structure of the CCP1-CCP2-SP zymogen.

EXPERIMENTAL PROCEDURES

Expression, Refolding, and Purification of the Recombinant Proteins—Mutagenesis of the synthesized cDNA for recombinant C1s CCP12SP (residues Lys^{281}–Asp^{688}) (GenScript, Piscataway, NJ) was carried out as described previously (11) to introduce a cysteine residue at the N terminus of selected proteins. The sequences of all variants were confirmed by double-stranded DNA sequencing. Expression, refolding, and purification of all proteins were carried out as described previously (11). Briefly, after transformation of the vector into *Escherichia coli* strain BL21(DE3)pLysS, the cells were cultured at 37 °C in 2× TY (tryptone/yeast extract) broth with 50 μg/ml ampicillin and 34 μg/ml chloramphenicol to a D_{595} of 0.6, followed by induction with 1 mM isopropyl β-D-thiogalactoside for 4 h. Following induction, the culture was centrifuged (27,000 × *g*, 20 min, 4 °C), the cells were collected in 30 ml of 50 mM Tris-HCl, 20 mM EDTA, pH 7.4, and then frozen at −80 °C. The cells were thawed and sonicated on ice six times for 30 s. After centrifugation at 27,000 × *g* for 20 min, inclusion body pellets were sequentially washed and centrifuged with 10 ml of 50 mM Tris-HCl, 20 mM EDTA, pH 7.4. The washed pellet was resuspended in 10 ml of 8 M urea, 0.1 M Tris-HCl, 100 mM DTT, pH 8.3, at room temperature for 3 h. Refolding was initiated by rapid dilution dropwise into 50 mM Tris-HCl, 3 mM reduced glutathione, 1 mM oxidized glutathione, 5 mM EDTA, and 0.5 M arginine, pH 9.0. The renatured protein solutions were concentrated and dialyzed against 50 mM Tris-HCl, pH 9.0, and renatured proteins were purified on a 5-ml Q-Sepharose column (GE Healthcare). The bound protein was eluted with a linear NaCl gradient from 0 to 400 mM over 35 ml at 1 ml/min. The recombinant proteins were further purified using a Superdex 75 16/60 column (GE Healthcare) in a buffer of 50 mM Tris, 145 mM NaCl, pH 7.4; aliquoted; snap frozen; and maintained at −80 °C. The purity of the protein was confirmed by SDS-PAGE followed by Western blotting and N-terminal sequencing. Typically protein yields were between 2 and 4 mg/liter. Where required, C1s was activated prior to use by incubating overnight at room temperature with C1r as described (11).

Surface Plasmon Resonance Studies—Surface plasmon resonance studies were performed using a BIAcore T100. Activated and zymogen forms of C1s CCP12SP (S632A) and activated C1s CCP12SP (S632A, K575A, R576A, R581A, and K583A) were immobilized on the active flow cell of a BIAcore SA Series S Sensor Chip (GE Life Science), by injecting each protein at a concentration of 0.2 mM. Triplicate samples of 0–3 μM C4 were injected for 60 s at 50 ml/min, with a 5-min dissociation.

The data were analyzed using SCRUBBER2 (BioLogic Software, Campbell, Australia) and SIGMAPLOT version 11.0 (Systat Software, Inc., Chicago, IL) using a two-state model: C1s + C4 ⇆ C1s:C4 ⇆ C1s*C4, where the initial complex (C1s:C4) is converted into a higher affinity complex (C1s*C4) because of conformational change. The rate constants for the first (k_{a1} and k_{d1}) and second (k_{a2} and k_{d2}) step were derived from the analysis using in-built equations for a two-state model in SCRUBBER2. The K_a for the initial binding step (K_1) = k_{a1}/k_{d1}, $K_2 = k_{a2}/k_{d2}$, and the overall association rate constant (K_a) = $K_1(1 + K_2)$. The calculated equilibrium dissociation constant for the entire reaction (K_d) = $1/K_a$. The maximal response units obtained at equilibrium for each concentration of C4 were plotted against the corresponding concentration of C4 and fitted by nonlinear regression using GraphPad Prism version 5.0 to a two-site binding model described by the following equation: $r = R^{Hi}_{max}$ [C4]/(K^{Hi}_d + [C4]) + R^{Lo}_{max} [C4]/(K^{Lo}_d + [C4]), where R represents the response units, R^{Hi}_{max} represents the maximal response units for the high affinity binding site, R^{Lo}_{max} represents the maximal response units for the low affinity binding site, and K^{Hi}_d and K^{Lo}_d represent the equilibrium dissociation constants for the high and low affinity binding sites, respectively.

Crystallization—Crystals of C1s-CCP12-SPz (Q436A,I438A activation loop mutant) were grown using the hanging drop method on siliconized glass coverslips at 20 °C above reservoirs containing 18% (w/v) PEG 3350, 0.2 M potassium nitrate. Drops contained 0.5 μl each of reservoir solution and protein solution at 2.4 mg/ml in 20 mM Tris, 145 mM NaCl, pH 7.4, 0.02% (w/v) NaN_3. The crystals typically formed within 2 weeks and were frozen using liquid nitrogen in synthetic mother liquor containing increasing concentrations of glycerol (5%, 10%, 20% (v/v) final) prior to data collection.

Wild-type C1s-CCP1-CCP2-SPz was crystallized using the hanging drop method at 20 °C above reservoirs containing 17% (w/v) PEG 2000, 0.1 mM succinate/phosphate/glycine buffer mixed at 35:5 ratio (pH 4–10). Drops contained 1 μl each of protein at 7 mg/ml in 20 mM Tris, 145 mM NaCl, pH 7.4, 0.02% (w/v) NaN_3, and reservoir solution. The crystals were cryoprotected using 30% (w/v) glucose in synthetic mother liquor, and the data were collected at cryogenic temperatures on the MX1 beamline at the Australian Synchrotron.

X-ray Data Collection, Structure Determination, and Refinement—Diffraction data were collected using the MX2 Beamline at the Australian Synchrotron at cryogenic temperatures. Two 360° sweeps were acquired for the crystal at 1° oscillation: one containing ice rings caused by ice on the outside of the crystal and one without ice after washing with liquid nitrogen; resolution ranges containing ice rings in the first data set were excluded, as were the final 80 frames collected because of indications of radiation damage. The data were indexed, integrated, scaled, and merged using the Xia2 data reduction pipeline (v0.3.4) (13) using the "−3daii" mode, which internally uses XDS (14–16), Pointless and Aimless (v0.1.16) (17), and the

CCP4 suite (v6.2.0) (18). 5% of the reflections were flagged as a validation set for calculation of R_{free}.

A summary of statistics is provided in Table 1. All diffraction images are deposited in TARDIS (19) and are freely available online. An initial molecular replacement solution was found using the BALBES pipeline web service (20), which internally uses MOLREP (21) with the coordinates for activated C1s (Protein Data Bank code 1ELV, containing the CCP2-SP domains) as a search model and REFMAC5 (22) for postrefinement of the solution. The initial solution contained the CCP2-SP domains (with two monomers in the asymmetric unit); however, BALBES failed to find a molecular replacement solution for the CCP1 domain. Visual inspection of this solution indicated that there was electron density N-terminal to the CCP2 domain. We generated a comparative model of the C1s CCP1 domain using Modeler v9.7 (23) based on the MASP-2 (Protein Data Bank code 1ZJK) CCP1 domain as a template. This model CCP1 domain was used as a second ensemble for molecular replacement using PHASER (24), along with the existing CCP2-SP solution. The best solution (LLG = 3546) contained two conformations of the CCP1 domain (one for each monomer in the asymmetric unit), only one of which was in the correct orientation with its C-terminal end (Pro^{80}) close to the N-terminal end of the CCP2 domain (Val^{81}). The CCP1 domain in this correct orientation was used to construct a complete 2× (CCP1-CCP2-SP) model, which was used as input to the Phenix (25) (v1.7.2 or dev-1048) AutoBuild protocol to generate a starting model (R_{work}/R_{free} = 0.306/0.396). The starting model was subjected to multiple rounds of manual model building using Coot (26) and refinement using Phenix with individual B-factor refinement, torsional non-crystallographic symmetry (NCS) restraints, and translation, libration, and screw rotation displacement refinement. ADP optimization was used in later rounds of refinement. Iterative build omit maps (27) were generated and inspected to reduce model bias. Final coordinates and structure factors have been deposited with the Protein Data Bank under accession code 4J1Y.

Structure Analysis—Fitting for RMSD calculations was performed using the McLachlan algorithm (28) as implemented in the program ProFit (available online). PyMOL 1.2r2 (DeLano Scientific, LLC) was used to produce figures containing molecular graphics. Intermolecular contacts at the CCP interface were determined using the CONTACT program as part of the CCP4 suite (18). PISA (29) was used for the analysis of potential quaternary structures and assemblies and associated surface area of contact measurements.

RESULTS AND DISCUSSION

SPR Experiments Reveal That Zymogen C1s Is Unable to Interact with C4—Our previous attempts at coupling C1s to sensor surfaces for SPR analysis via amine residues failed to reveal significant binding between C1s and C4. Because our work had implicated positively charged residues in C1s in its interaction with C4 (11), we reasoned that the amine coupling approach we initially deployed may interfere with C4 binding by modifying key lysine residues. We therefore developed an SPR-based assay where a form of C1s CCP1-CCP2-SP that was biotinylated at the N terminus was immobilized to a streptavidin-coated chip.

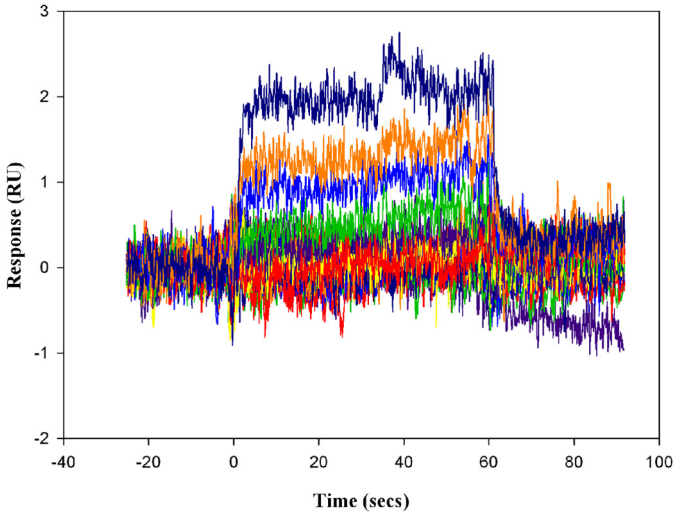

FIGURE 1. **SPR curves for 0–3 μM C4 flowed over biotinylated zymogen form CCP1-CCP2-SP C1s attached to streptavidin immobilized to the chip.** The obtained responses were fitted to a two-step binding equation within the program SCRUBBER.

Using this approach, we were able to investigate the ability of zymogen and activated forms of a S632A mutant (the active site Ser residue had to be mutated to Ala to prevent cleavage of C4) to bind C4. The zymogen form was unable to bind C4 effectively, with very low binding responses only observed at high concentrations of C4 (Fig. 1), in complete contrast to the activated form (Fig. 2). The data obtained with the activated form allowed us to ascertain the kinetic mechanism and rate constants underlying the binding of C4 by C1s. Because previous data (12) strongly suggest that two binding sites for C4 exist on C1s, we plotted the SPR response units as a function of C4 concentration to yield a curve that could be fitted using a two-site binding model (Fig. 2*B*). This analysis gave two K_d values: one for a low affinity site of 540 nM and one for a high affinity site of 43 nM. The low affinity equilibrium dissociation constant would most likely correlate with an initial binding step in a two-step/state mechanism, and the high affinity one would most likely correlate with a second, locking step in the interaction. Indeed, the SPR data obtained for the binding of C4 by activated CCP1-CCP2-SP C1s could be fitted to a two-state binding model: C1s + C4 $\leftrightarrows$ C1s:C4 $\leftrightarrows$ C1s*C4, where the initial complex (C1s:C4) is converted into a higher affinity complex (C1s*C4) because of conformational change. The following rate constants were obtained: $k_{a1} = 7.9 \times 10^5$ ($\pm 5.7 \times 10^3$) $M^{-1}\cdot s^{-1}$; $k_{d1} = 0.48$ ($\pm$ 0.003) s^{-1}; $k_{a2} = 2.5 \times 10^{-4}$ ($\pm 1 \times 10^{-6}$) s^{-1}; and $k_{d2} = 6.5 \times 10^{-5}$ ($\pm 2 \times 10^{-5}$) s^{-1}. The individual rate constants for each of the two steps in the binding mechanism yielded an overall association rate constant (K_a) of 7.8×10^6 $M^{-1}\cdot s^{-1}$, and the K_d for the overall reaction was estimated to be 128 nM. The K_d for the initial binding step was estimated to be 607 nM, in reasonable concordance with the low affinity binding constant obtained from two-site equilibrium binding analysis (540 nM). The estimated K_d for the overall reaction is between the values for the low and high affinity sites estimated from the equilibrium analysis, as might be expected. Thus, the constants obtained

by fitting the equilibrium data to the two-site binding model correlate with calculated constants from the data obtained from the SPR kinetic analyses, indicating that the use of the two-state binding model to analyze the SPR data was justified.

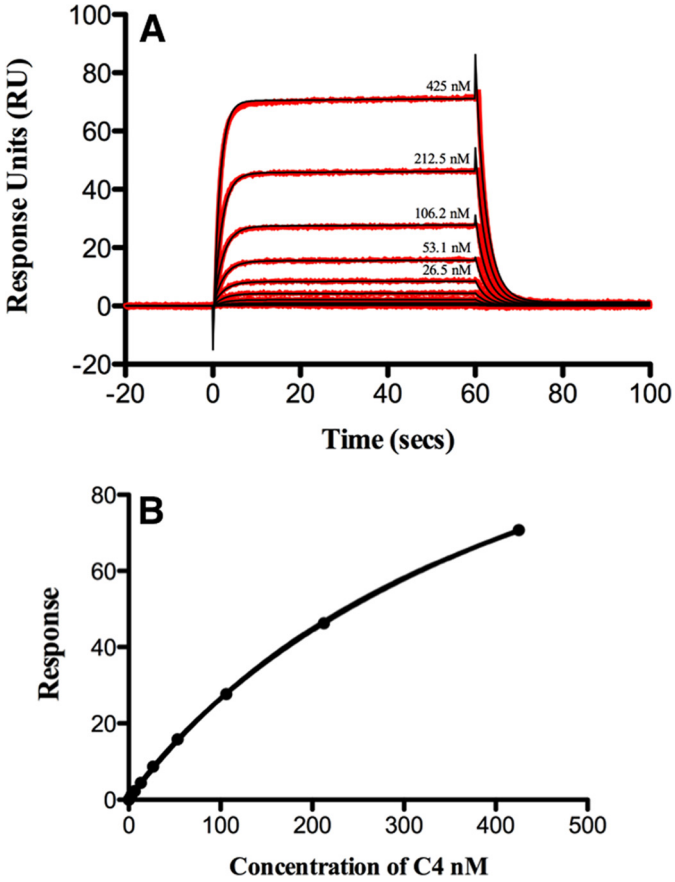

FIGURE 2. *A*, SPR curves for 0–425 nM C4 flowed over biotinylated activated CCP1-CCP2-SP C1s attached to streptavidin immobilized to the chip. The obtained responses (*red lines*) were fitted to a two-state binding equation in Biacore T100 Evaluation Software (*fitted curves* shown by *black lines* superimposed on the obtained responses). The experiments were conducted in triplicate and showed excellent overlap. *B*, the response units obtained at equilibrium for each concentration of C4 were plotted against the C4 concentration and fitted using a two-site binding model on GraphPad Prism (regression coefficient for the fit = 0.99).

The data given here are consistent with an initial low affinity interaction with a K_d value of ~500 nM. The concentration of C4 in plasma varies between individuals, but lies in the range from 0.4 to 1.7 μM. The initial binding interaction would thus be expected to occur with an affinity in the range of the physiological concentration of C4 and would represent the encounter complex between the protease and the substrate. The locking step yielding a high affinity interaction would then be predicted to bring the cleavage/activation loop in C4 into the active site of C1s.

The X-ray Structure of Zymogen C1s CCP1-CCP2-SP and Structural Comparison with the Active Form—Our kinetic data suggested that the transition from the zymogen to the active form is an important determinant that governs the interaction between C1s and C4. The structure of active C1s (CCP2-SP) is known (30). Accordingly, to understand the conformational changes that occur upon enzyme activation, we determined the 2.7 Å structure of the C1s CCP1-CCP2-SP zymogen. The construct contained two mutations (Q436A and I438A) in the activation loop (the region targeted for cleavage by C1r) (Fig. 3 and Table 1), designed to stabilize the mutant against cleavage by C1r. Two molecules were present in the asymmetric unit. Throughout this work, we refer to the (more complete) molecule A. We also determined a low resolution (3.5 Å) structure of wild-type zymogen CCP1-CCP2-SP; this structure was essentially identical to the higher resolution mutant structure described here (Fig. 4).

Several regions in C1s zymogen structure were not visible in electron density (Glu^{297}, Gly^{396}, Glu^{433}–Ile^{438}, and Lys^{599}–Ala^{607}). Most notably, they include the activation loop region (residues Glu^{433}–Ile^{438}). This region was also disordered in the lower resolution wild-type structure. Together, these data suggest that the activation loop is exposed and flexible in the C1s zymogen and that once appropriately positioned within the C1 complex, this region would be readily available for cleavage.

We compared the structure of the SP of the C1s zymogen determined here with the activated SP in the previously determined structure of the CCP2-SP form (30) (Fig. 5). As with other members of the chymotrypsin family (31–33) (Table 2),

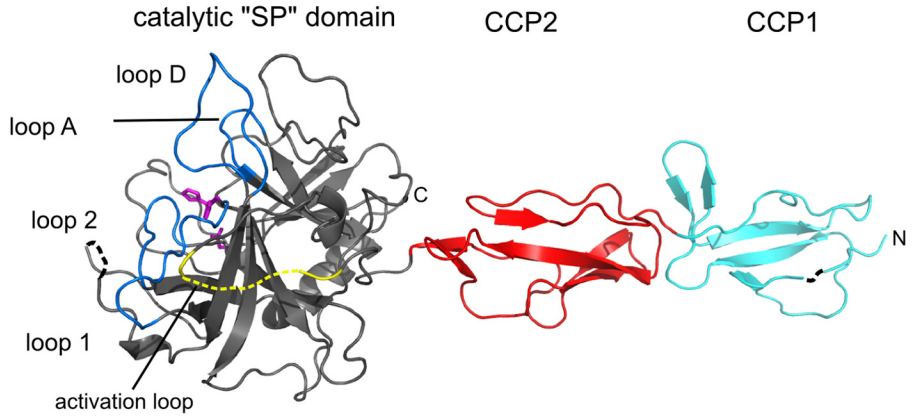

FIGURE 3. **The catalytic domain of C1s zymogen (*gray*) shows a chymotrypsin-like fold composed of two six-stranded β-barrels and a short α-helix at the extreme C-terminal end, whereas the CCP domains (*red* and *cyan*) are a "sushi domain" fold.** One of the two C1s-CCP12-SP zymogen molecules in the asymmetric unit is shown, represented as a cartoon ribbon. The SP, CCP1, and CCP2 domains and the N and C termini of each molecule are labeled. The catalytic triad (Ser$^{632[195]}$, His$^{475[57]}$, and Asp$^{529[102]}$) is shown as *magenta sticks*, and loops 1, 2, A, and D are colored *blue*. The activation loop region is colored *yellow*. Regions not visible in electron density are shown by *dashed lines*.

TABLE 1

Data collection and refinement statistics

The values in parentheses are for the highest resolution shell (2.66–2.73 Å). The loops with no observed electron density and side chains with partial or no observed electron density beyond Cβ atoms are listed (C1s-CCP1-CCP2-SPz, Q436A,I438A activation loop mutant).

Data collection	
Space group	C222
Cell dimensions	
a, *b*, *c* (Å)	149.9, 158.9, 78.7
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	63.8 (2.66)
I/σ	20.6 (2.5)
R_{merge} (%)	9.3 (112.5)
$R_{pim(I)}$ (%)	2.8 (39.2)
Completeness (%)	99.4 (97.8)
Redundancy	12.2 (9.3)
Total observations	331,946 (18,225)
Refinement	
Resolution (Å)	63.8 (2.66)
Unique reflections	27,106 (2614)
R_{work}/R_{free} (%)	20.1/25.8
No. atoms	5661
Protein	5620
Ligand/ion	0
Water	41
B-factors	
Protein	103.4
Water	70.10
RMSDs	
Bond lengths (Å)	0.011
Bond angles (°)	1.52
Ramachandran favored (%)	93.7
Ramachandran outliers (%)	0.7
Molprobity score[a]	2.40
Molprobity clashscore[a]	11.11
Number of protein residues	741
Missing loops	
Molecule A: Glu^{297}, Gly^{396}, Glu^{433}--Ile^{438}, and Lys^{599}--Ala^{607}	
Molecule B: Gly^{396}, Phe^{432}--Gly^{441}, Thr^{494}--Ala^{498}, Lys^{575}--Arg^{576}, Lys^{599}--Tyr^{610}, and Gly^{24}	
Missing side chains	
Molecule A: Lys^{296}, Glu^{297}, Asp^{298}, Lys^{310}, Lys^{369}, Lys^{420}, Lys^{515}, Glu^{518}, Glu^{597}, Glu^{608}, Lys^{623}, Lys^{677}, and Gln^{680}	
Molecule B: Gln^{297}, Lys^{351}, Lys^{369}, Asn^{395}, Lys^{420}, Lys^{447}, Thr^{494}, Lys^{499}, Lys^{501}, Glu^{518}, Val^{519}, Lys^{536}, Lys^{540}, Leu^{559}, Met^{560}, Arg^{578}, Arg^{581}, Lys^{583}, Arg^{586}, Lys^{596}, Glu^{597}, Lys^{623}, Lys^{629}, Lys^{644}, and Lys^{677}	

[a] Molprobity score and clashscore are in the 92nd and 97th percentile, respectively, compared with structures at this resolution.

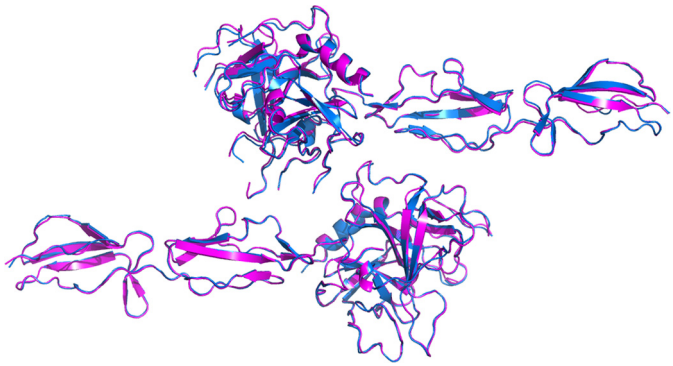

FIGURE 4. **The structure of wild-type C1s (*magenta*) determined at 3. 5 Å resolution superimposed with the 2.7 Å activation loop mutant structure (*blue*).** RMSD over backbone CA atoms is 0.445 Å.

the catalytic domain of C1s undergoes significant structural changes upon activation. Most notably, large conformational changes and rearrangements are observed in loops 1 (Gly^{621}–Asp^{631}), 2 (Trp^{655}–Thr^{661}), and D (Trp^{570}–Leu^{582}) (Fig. 5*A*). These loops form the "activation domain" of chymotrypsin family serine proteases (34) and surround the acyl (nonprime) side of the substrate binding site. Of these changes, the move-

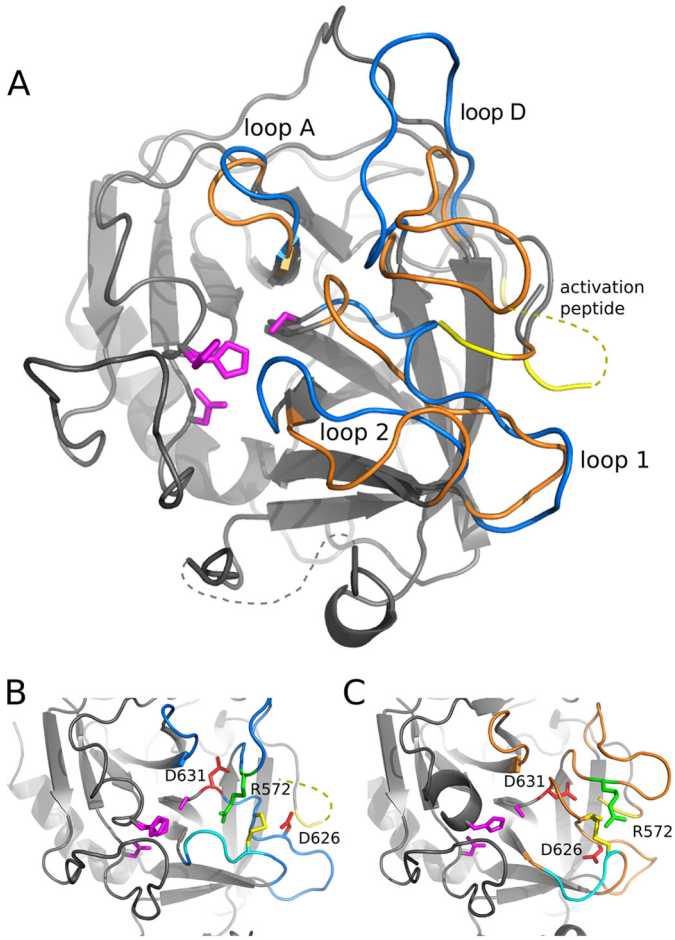

FIGURE 5. *A*, superimposition of the catalytic domain of zymogen and loops showing displacement from active (1ELV) C1s, represented as cartoon ribbons. Regions of the backbone that show significant displacements between the active (*orange*) and zymogen (*blue*) form are labeled with the nomenclature of Perona and Craik (36). Regions not visible in electron density are shown by *dashed lines* (part of the activation peptide (*yellow*) and loop 3). Loop 1 spans residues 621–631, loop 2 spans residues 655–661, and loop D spans residues 570–582. *B* and *C*, atomic details of key side chains in the zymogen (*B*) and active (*C*) protease are shown: Asp^{626} (*red*), which is found at the base of the S1 pocket in the active form, and Asp^{631} (*red*), which makes ionic interactions with the cleaved activation peptide N terminus. Arg^{572}, which in the zymogen bridges loops 1 and 2 via backbone hydrogen bonds and occludes the S1 pocket, is shown in *green*. The disulfide bond (Cys^{628}–Cys^{659}) between loops 1 and 2 is shown with *yellow sticks*. The side chains of the catalytic triad are represented as *magenta sticks* (Ser$^{632[195]}$, His$^{475[57]}$, and Asp$^{529[102]}$).

ments in the D-loop are particularly striking, with residue shifts of up to 11 Å apparent.

In the C1s zymogen, the oxyanion hole is unformed as a consequence of the distortion of the peptide backbone in the loop 1 region (Ser^{627}–Asp^{631}). The amide of Gly$^{630[193]}$ (throughout the manuscript, residue numbers in brackets following the native C1s numbering refer to the canonical chymotrypsinogen numbering) on loop 1 is flipped ~180° and is ~7.7 Å away from its position in the active enzyme. The catalytic triad of zymogen C1s (Ser$^{632[195]}$, His$^{475[57]}$, and Asp$^{529[102]}$) is essentially positioned in the active conformation. His$^{475[57]}$ is best modeled as two alternative conformations: one mimicking the active conformation and a second where the side chain is rotated away from S632[195] such that it makes no side chain hydrogen bonds.

TABLE 2
Named loops of C1s and related serine proteases
RMSD values are also shown between superimposed catalytic domains of C1s active and zymogen forms and other related serine proteases. The values shown are the RMSD/number Cα atoms.

Loop/domain name	C1s (zymogen/1ELV)[a]	C1r (1GPZ/1MD8)	MASP-2 (1ZJK/1Q3X)	Chymotrypsinogen (2CGA/1YPH)
SP domain	2.4/235 [0.76/219] Cys^{425}–Ser^{683}	1.93/229 [0.87/215] Cys^{434}–Glu^{686} (451–703)	1.97/246 (0.84/227) Cys^{434}–Phe^{686}	1.57/238 (0.63/228) Cys^{1}–Asn^{245}
Activation peptide	429–441 (cleaved at $Arg^{437} \downarrow Ile^{438}$)			5–18 (cleaved at $Arg^{15} \downarrow Ile^{16}$)
A	456–460 (shorter than in chymotrypsinogen)	465–470 (482–488)	463–468	34–41
B	478–484			59–66
E	493–502			74–82
C	514–525 (longer in C1s, poorly defined)			95–98
D	570–582	579–590 (596–607)	573–585	145–151
3	593–612 (longer in C1s, 598–609 not observed)			170–175
1	621–631	626–636 (643–653)	622–631	185–187
2	655–661	659–666 (676–683)	654–665	217–223

[a] RMSD values are based on all equivalent Cα atoms. RMSD of equivalent Cα atoms in core regions aligned by ProFit (iterative alignment) are in square brackets. The Protein Data Bank codes are: MASP-2 zymogen, 1ZJK (37), and active form, 1Q3X (38); C1r zymogen, 1GPZ (39), and active form, 1MD8 (40); chymotrypsinogen, 2CGA (41); and chymotrypsin, 1YPH. Where Protein Data Bank crystal structure numbering does not match the native numbering (i.e., C1r structures), full-length native sequence numbering is indicated in brackets.

The primary (S1) substrate binding site of the C1s zymogen is unformed. The acidic side chain of Asp^{626} on loop 1 (equivalent to the specificity determining Asp^{189} in trypsin), which sits at the base of the S1 pocket in the active form, is exposed to the solvent in a position away from the S1 pocket in the zymogen (Fig. 5, *B* and *C*). The top of the S1 pocket is also completely occluded by residues from loop 2 (Gly^{656}–Cys^{659}) and Arg^{572} from loop D. Most notably, Arg^{572} from the loop D bridges loop 1 and 2 via hydrogen bonds to the backbone carbonyl oxygens of Gly^{630} and Gly^{656} (Fig. 5*B*).

Comparison of the C1s Zymogen with the MASP-2/C4 Complex Suggests a Conformational Switch in loop D Governs C4 Binding—In mammals, a second proteolytic system, the mannose-binding lectin assembly, feeds into the complement system. Like C1, the mannose-binding lectin complex includes proteases that are suggested to be analogous to C1r (MASP-1) and C1s (MASP-2; which shares similar domain structure and 36% sequence identity in the protease domain to C1s). Recently, the structure of MASP-2 in complex with C4 was determined (12). By analogy, these data provide important insight into the interaction between C1s and C4.

Superimposition of the active form of C1s with MASP-2 reveals that the conformation of the key loops (and in particular loop D) that contact C4 are essentially in the same conformation (Fig. 6, *A* and *B*, and 7). Furthermore, both MASP-2 and C1s contain a cluster of basic residues that contact the C4 sulfotyrosine sequence. In particular, in the active form of C1s, we note that Arg^{576} and Arg^{581} are positioned such that they would be able to interact with the sulfotyrosine and aspartyl residues on C4. Together, these data suggest that activated C1s could bind to C4 in a similar orientation.

In contrast, similar superimposition experiments with the SP domain of the C1s zymogen reveal that the position of loop D (and 576–579 in particular) would be anticipated to sterically interfere with C4 binding (Fig. 6*C*). Interestingly, in the C1s zymogen, changes to loop D are such that Arg^{576} is ~10 Å [$C\beta$-$C\beta$] from its position in the active enzyme, oriented away from the sulfotyrosine residues.

To test these ideas, we used a mutant that combined the S632A mutation with mutations to alanine of all four basic residues in loop D that are predicted to act as an SP domain exosite for C4 (Lys^{575}, Arg^{576}, Arg^{581}, and Lys^{583}). We subjected this mutant (biotinylated at an N-terminal Cys residue) to further SPR analysis (Fig. 8) in comparison with the S632A mutant described previously (Fig. 2). Our data revealed that the activated form of the mutant was as impaired in its ability to bind C4 as the zymogen form of the enzyme. Together, these data support the idea that this region does indeed form a binding exosite for C4 (11).

The CCP1 Domain of C1s Likely Contains an Additional Exosite for C4 Binding—The three-dimensional structure of the C1s CCP1 domain has not been previously reported, because the published structure of active C1s does not contain this domain (30). Previously, the CCP1 and CCP2 domains of C1s and MASP-2 have been implicated in C4 binding (10, 35). In the MASP-2/C4 complex (12), the "C345C" domain of C4 interacts with the CCP1 and CCP2 domains of MASP-2, particularly via long range electrostatic interactions with Glu^{333} and

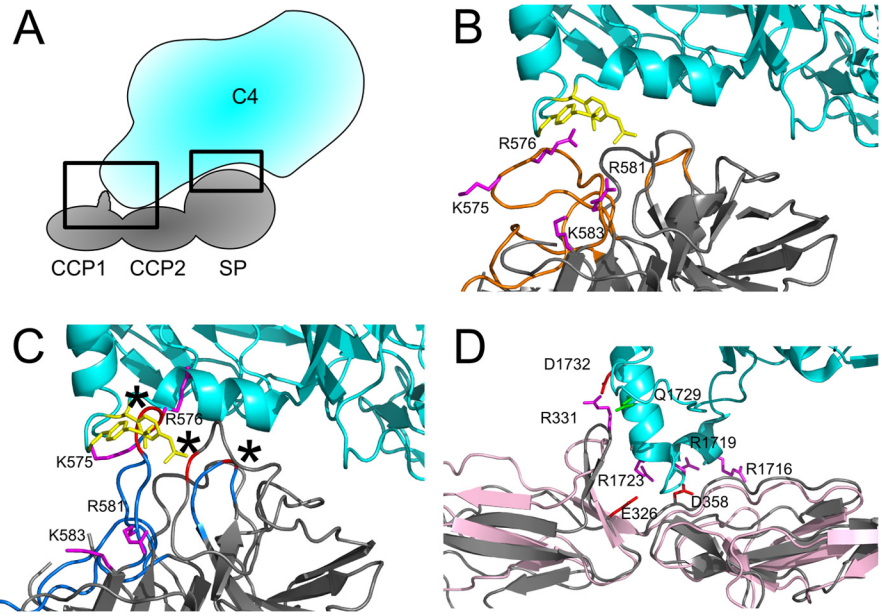

FIGURE 6. *A*, schematic representation of the MASP-2/C4 complex and the proposed C1s/C4 complex. *Boxed regions* in *A* are shown in *B–D*. The coordinates of active (*orange loops*) and zymogen (*blue loops*) C1s were superimposed on the MASP-2 SP domain from the MASP-2/C4 complex structure (Protein Data Bank code 4FXG), maintaining the relative orientation of C4 (*cyan*). *B* and *C*, the potential C1s/C4 exosite interface between active (*B*) and zymogen (*C*) was compared with that of MASP-2/C4. Active C1s has essentially the same backbone conformation as activated MASP-2, and by analogy with MASP-2, the positively charged residues Lys^{575}, Arg^{576}, Arg^{581}, and Lys^{583} (*magenta*) would be anticipated to form a binding site for the C4 sulfotyrosine tail. Active C1s does not make any large steric clashes with C4 in this arrangement. In contrast, the sulfotyrosine exosite interface is deformed in the C1s zymogen, in particular because of the conformational change in loop D. The C1s zymogen makes steric clashes with C4 in this orientation (Arg^{576}–Arg^{578} in loop D, Asn^{457} in loop A, and Gln^{493}, backbone is colored *red* and indicated with an *asterisk* at clash regions where proposed interfacial heavy atoms are closer than 2 Å). The C4 activation loop (residues 750–760), which makes steric clashes with C1s zymogen loops 1 and 2 because of the unformed S1 pocket, is omitted for clarity. The two sulfotyrosine residues on the C-terminal end of C4 are represented as *yellow sticks*. *D*, comparison of the CCP domain exosite between MASP-2/C4 and a potential C1s/C4 complex. C1s was superimposed with MASP-2/C4 based on aligning the CCP2 domain only to compare the potential exosite interfaces. In this orientation, C1s is predicted to make conserved long range electrostatic interactions with C4 via C1s-Glu^{326} (MASP-2 Glu^{333}) and C1s-Asp^{358} (MASP-2 Asp^{355}) to Arg^{1716}, Arg^{1719}, and Arg^{1724} on C4. Based on our model, we further predict that C1s-Arg^{331} may make an additional interaction with C4-Asn^{1729} and ionic interactions with C4-Asp^{1732}.

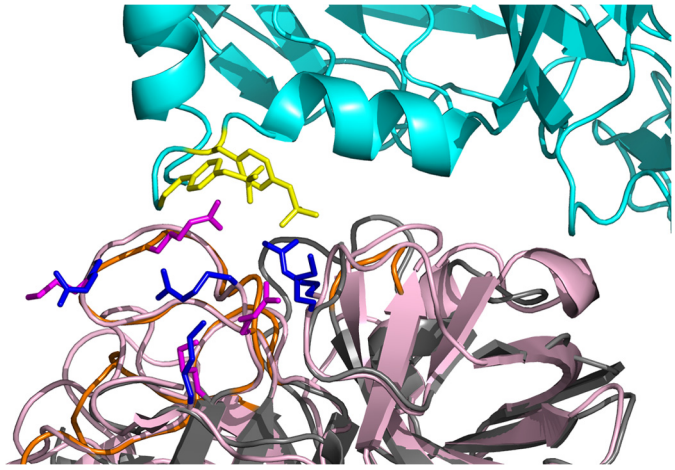

FIGURE 7. **Activated C1s (*gray* with *orange loops*) superimposed with activated MASP-2 (*pink*) from the MASP-2/C4 complex.** Positively charged residues that form the proposed exosite for the C4 sulfotyrosine tail (*yellow*) are colored *magenta* (C1s) and *blue* (MASP-2).

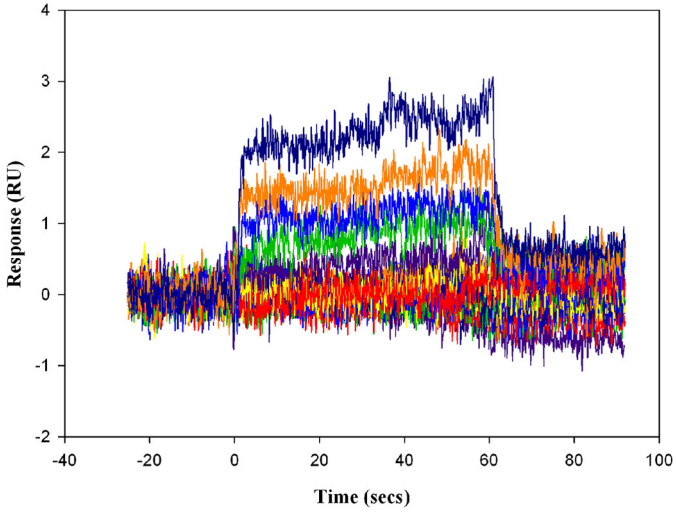

FIGURE 8. **SPR curves for 0–3 μM C4 flowed over biotinylated activated form CCP1-CCP2-SP C1s with mutations to all four positively charged residues in the exosite on the SP domain, attached to streptavidin immobilized to the chip.** The obtained responses were fitted to a two-step binding equation within the program SCRUBBER.

Asp^{365}; these positions are conserved in C1s (Glu^{326} and Asp^{358}, respectively) (Fig. 6*D*). The CCP1 domain of C1s contains an extended β-hairpin loop (327–334), which is structurally equivalent to the 333–340 loop of MASP-2. We therefore suggest that, as in the MASP-2 complex, the 327–334 loop may comprise an exosite for interaction with C4 (Fig. 6*D*). In addition, we note that Arg^{331} (not conserved in MASP-2) on the tip of the C1s CCP1 β-hairpin loop is placed such that it would be predicted to make a salt bridge with Asp^{1732} of C4.

Concluding Statements—Our own work (11) and that of others has suggested the importance of exosites on the CCP (10) and SP domains (12) for the interaction between C1s/MASP-2 and C4. Here, we have shown by SPR that the C1s zymogen is

unable to form significant interactions with C4. In contrast, active C1s is able to form a tight complex with C4. These latter data obtained using SPR could be fitted to a two-state binding model, yielding an overall K_d value of 32 nM.

In support of these ideas, we demonstrate through structural studies that the conformations of key loops maintained in the C1s zymogen form would predictably occlude interaction with C4. Mutagenesis studies performed here and previously (11) further reveal that positively charged residues Lys^{575}, Arg^{576}, Arg^{581}, and Lys^{583} in loop D are essential for binding C4. Together our data suggest that the C1s SP domain contains an exosite under conformational control that is able to bind C4 upon the transition from the zymogen to the active form.

Acknowledgments—We thank Prof. Anna Blom (Lund University, Malmo, Sweden) for initial assistance with the SPR work and useful discussions about the data obtained in this study. We also thank Prof. James Huntington (Cambridge Institute for Medical Research, Cambridge, UK) for insightful discussions surrounding the structure of the enzyme. We thank the Australian synchrotron (MX-2 Beamline), the Monash Protein Crystallography platform, and the Monash Biomedical Proteomics facility for valuable technical assistance and support.

REFERENCES

1. Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J. D. (2010) Complement. A key system for immune surveillance and homeostasis. *Nat. Immunol.* **11,** 785–797
2. de Cordoba, S. R., Tortajada, A., Harris, C. L., and Morgan, B. P. (2012) Complement dysregulation and disease. From genes and proteins to diagnostics and drugs. *Immunobiology* **217,** 1034–1046
3. Rosado, C. J., Buckle, A. M., Law, R. H., Butcher, R. E., Kan, W. T., Bird, C. H., Ung, K., Browne, K. A., Baran, K., Bashtannyk-Puhalovich, T. A., Faux, N. G., Wong, W., Porter, C. J., Pike, R. N., Ellisdon, A. M., Pearce, M. C., Bottomley, S. P., Emsley, J., Smith, A. I., Rossjohn, J., Hartland, E. L., Voskoboinik, I., Trapani, J. A., Bird, P. I., Dunstone, M. A., and Whisstock, J. C. (2007) A common fold mediates vertebrate defense and bacterial attack. *Science* **317,** 1548–1551
4. Hadders, M. A., Beringer, D. X., and Gros, P. (2007) Structure of C8α-MACPF reveals mechanism of membrane attack in complement immune defense. *Science* **317,** 1552–1554
5. Forneris, F., Wu, J., and Gros, P. (2012) The modular serine proteases of the complement cascade. *Curr. Opin. Struct. Biol.* **22,** 333–341
6. Duncan, R. C., Wijeyewickrema, L. C., and Pike, R. N. (2008) The initiating proteases of the complement system. Controlling the cleavage. *Biochimie* **90,** 387–395
7. Villiers, C. L., Arlaud, G. J., and Colomb, M. G. (1985) Domain structure and associated functions of subcomponents C1r and C1s of the first component of human complement. *Proc. Natl. Acad. Sci. U.S.A.* **82,** 4477–4481
8. Busby, T. F., and Ingham, K. C. (1990) NH_2-terminal calcium-binding domain of human complement C1s− mediates the interaction of C1r− with C1q. *Biochemistry* **29,** 4613–4618
9. Kardos, J., Gál, P., Szilágyi, L., Thielens, N. M., Szilágyi, K., Lõrincz, Z., Kulcsár, P., Gráf, L., Arlaud, G. J., and Závodszky, P. (2001) The role of the individual domains in the structure and function of the catalytic region of a modular serine protease, C1r. *J. Immunol.* **167,** 5202–5208
10. Rossi, V., Bally, I., Thielens, N. M., Esser, A. F., and Arlaud, G. J. (1998) Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* **273,** 1232–1239
11. Duncan, R. C., Mohlin, F., Taleski, D., Coetzer, T. H., Huntington, J. A., Payne, R. J., Blom, A. M., Pike, R. N., and Wijeyewickrema, L. C. (2012) Identification of a catalytic exosite for complement component C4 on the serine protease domain of C1s. *J. Immunol.* **189,** 2365–2373
12. Kidmose, R. T., Laursen, N. S., Dobó, J., Kjaer, T. R., Sirotkina, S., Yatime, L., Sottrup-Jensen, L., Thiel, S., Gál, P., and Andersen, G. R. (2012) Structural basis for activation of the complement system by component C4 cleavage. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 15425–15430
13. Winter, G. (2009) Xia2. An expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43,** 186–190
14. Kabsch, W. (1988) Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr.* **21,** 67–72
15. Kabsch, W. (1988) Evaluation of single-crystal x-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.* **21,** 916–924
16. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26,** 795–800
17. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 72–82
18. Collaborative Computational Project, Number 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763
19. Androulakis, S., Schmidberger, J., Bate, M. A., DeGori, R., Beitz, A., Keong, C., Cameron, B., McGowan, S., Porter, C. J., Harrison, A., Hunter, J., Martin, J. L., Kobe, B., Dobson, R. C., Parker, M. W., Whisstock, J. C., Gray, J., Treloar, A., Groenewegen, D., Dickson, N., and Buckle, A. M. (2008) Federated repositories of x-ray diffraction images. *Acta Crystallogr. D Biol. Crystallogr.* **D64,** 810–814
20. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008) BALBES. A molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 125–132
21. Vagin, A., and Tepllyakov, A (1997) MOLREP. An automated program for molecular replacement. *J. Appl. Crystallogr.* **30,** 1022–1025
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
23. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815
24. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
25. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX. A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221
26. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
27. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Adams, P. D., Read, R. J., Zwart, P. H., and Hung, L. W. (2008) Iterative-build OMIT maps. Map improvement by iterative model building and refinement without model bias. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 515–524
28. McLachlan, A. D. (1982) Rapid comparison of protein structures. *Acta Crystallogr. A* **38,** 871–873
29. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797
30. Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G. J., and Fontecilla-Camps, J. C. (2000) Crystal structure of the catalytic domain of human complement C1s. A serine protease with a handle. *EMBO J.* **19,** 1755–1765
31. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T., and Xuong, N. H. (1970) Chymotrypsinogen. 2.5-angstrom crystal structure, comparison with α-chymotrypsin, and implications for zymogen activation. *Biochemistry* **9,** 1997–2009
32. Madison, E. L., Kobe, A., Gething, M. J., Sambrook, J. F., and Goldsmith, E. J. (1993) Converting tissue plasminogen activator to a zymogen. A regulatory triad of Asp-His-Ser. *Science* **262,** 419–421
33. Vijayalakshmi, J., Padmanabhan, K. P., Mann, K. G., and Tulinsky, A. (1994) The isomorphous structures of prethrombin2, hirugen-, and

PPACK-thrombin. Changes accompanying activation and exosite binding to thrombin. *Protein Sci.* **3,** 2254–2271
34. Huber, R., and Bode, W. (1978) Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* **11,** 114–122
35. Duncan, R. C., Bergström, F., Coetzer, T. H., Blom, A. M., Wijeyewickrema, L. C., and Pike, R. N. (2012) Multiple domains of MASP-2, an initiating complement protease, are required for interaction with its substrate C4. *Mol. Immunol.* **49,** 593–600
36. Perona, J. J., and Craik, C. S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **272,** 29987–29990
37. Gál, P., Harmat, V., Kocsis, A., Bián, T., Barna, L., Ambrus, G., Végh, B., Balczer, J., Sim, R. B., Náray-Szabó, G., and Závodszky, P. (2005) A true autoactivating enzyme. Structural insight into mannose-binding lectin-associated serine protease-2 activations. *J. Biol. Chem.* **280,** 33435–33444
38. Harmat, V., Gál, P., Kardos, J., Szilágyi, K., Ambrus, G., Végh, B., Náray-Szabó, G., and Závodszky, P. (2004) The structure of MBL-associated serine protease-2 reveals that identical substrate specificities of C1s and MASP-2 are realized through different sets of enzyme-substrate interactions. *J. Mol. Biol.* **342,** 1533–1546
39. Budayova-Spano, M., Lacroix, M., Thielens, N. M., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2002) The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J.* **21,** 231–239
40. Budayova-Spano, M., Grabarse, W., Thielens, N. M., Hillen, H., Lacroix, M., Schmidt, M., Fontecilla-Camps, J. C., Arlaud, G. J., and Gaboriaud, C. (2002) Monomeric structures of the zymogen and active catalytic domain of complement protease C1r. Further insights into the c1 activation mechanism. *Structure* **10,** 1509–1519
41. Wang, D., Bode, W., and Huber, R. (1985) Bovine chymotrypsinogen A x-ray crystal structure analysis and refinement of a new crystal form at 1.8 A resolution. *J. Mol. Biol.* **185,** 595–624