

ingly, *Wnt3a* is neither expressed in the AER nor implicated in its formation during mouse limb development (15, 30). Several other *Wnt* genes have also been shown to be expressed differentially between mouse and chick embryos, both in the developing central nervous system and in limb buds, including some in the murine AER (15, 32). Different *Wnt* genes could substitute for one another as long as they activate the same intracellular signaling pathway mediated by β -catenin/LEF1. It is therefore probable that another species of *Wnt* that is expressed in the mouse AER plays the same role as *Wnt3a* in the chick.

Both WNT3a and WNT7a proteins act, at least in part, on the mesoderm, where they activate distinct targets; WNT3a induces *Lef1* whereas WNT7a induces *Lmx1*, which implies that receptors for both factors must be present on the surface of mesenchymal cells. In spite of previous data suggesting that all members of the highly transforming class of *Wnt* genes act through β -catenin, our results indicate that the induction of *Lmx1* expression by WNT7a signaling is not mediated by β -catenin and LEF1. Precedent exists for more divergent *Wnt* genes, such as *Wnt5a*, to act through distinct signaling cascades (33). Transcriptional activation of downstream genes by distinct WNT pathways allows for their different inductive roles in the same tissue during development.

REFERENCES AND NOTES

- R. L. Johnson and C. L. Tabin, *Cell* **90**, 979 (1997).
- M. Kengaku, V. Twombly, C. Tabin, *Cold Spring Harbor Symp. Quant. Biol.* **L12**, 421 (1998); Y. Yang and A. McMahon, personal communication.
- A. Vogel, C. Rodriguez, J. C. Izpisua-Belmonte, *Development* **122**, 1737 (1996).
- P. H. Crossley, G. Minowada, C. A. MacArthur, G. R. Martin, *Cell* **84**, 127 (1996).
- See supplementary figures at www.sciencemag.org/feature/data/976400.shl.
- A mouse *Wnt3a* cDNA encoding the entire open reading frame and a deleted mutant of β -catenin containing the internal *Armadillo* repeats that acts as a stable constitutively activated variant (23) were individually cloned into the shuttle vector SLAX-13 and then subcloned into retroviral vector RCAS(BP)A (34). Retrovirus was produced using a line 0 chick embryo fibroblast and harvested as described (35). Embryos at stage 9 to 11 were injected in the fore- or hindlimb primordia. This protocol results in widespread infection of the limb ectoderm by the time the embryos are harvested at stage 22 to 24 for whole-mount in situ hybridization (7, 8, 17). A significant number of samples (>10) were analyzed in each case. To generate embryos primarily infected in the limb ectoderm, injection was targeted onto the external surface of the ectoderm underneath the vitelline membrane.
- E. Laufer *et al.*, *Nature* **386**, 366 (1997); C. Rodriguez-Esteban *et al.*, *ibid.*, p. 360.
- M. Kengaku *et al.*, unpublished data.
- H. Ohuchi *et al.*, *Development* **124**, 2235 (1997).
- Y. Yokouchi, K. Ohsugi, H. Sasaki, A. Kuroiwa, *ibid.* **113**, 431 (1991); M. A. Ros *et al.*, *ibid.* **116**, 811 (1992).
- S. Sawai, K. Kato, Y. Wakamatsu, H. Kondoh, *Mol. Cell. Biol.* **10**, 2017 (1990).
- M. A. Ros, M. Sefton, M. A. Nieto, *Development* **124**, 1821 (1997).
- G. T. Wong, B. J. Gavin, A. P. McMahon, *Mol. Cell. Biol.* **14**, 6278 (1994).
- S. J. Du, S. M. Purcell, J. L. Christian, L. L. McGrew, R. T. Moon, *ibid.* **15**, 2625 (1995); R. T. Moon, J. D. Brown, M. Torres, *Trends Genet.* **13**, 157 (1997).
- B. A. Parr, M. J. Shea, G. Vassileva, A. P. McMahon, *Development* **119**, 247 (1993).
- B. J. Gavin, J. A. McMahon, A. P. McMahon, *Genes Dev.* **4**, 2319 (1990); C. N. Dealy, A. Roth, D. Ferrari, A. M. Brown, R. A. Koshier, *Mech. Dev.* **43**, 175 (1993).
- R. D. Riddle *et al.*, *Cell* **83**, 631 (1995).
- A. Vogel, C. Rodriguez, W. Warnken, J. C. Izpisua Belmonte, *Nature* **378**, 716 (1995).
- P. A. Parr and A. P. McMahon, *ibid.* **374**, 350 (1995).
- C. Yost *et al.*, *Genes Dev.* **10**, 1443 (1996); B. Rubinfeld *et al.*, *Science* **275**, 1790 (1997).
- J. R. Miller and R. T. Moon, *Genes Dev.* **10**, 2527 (1996).
- L. L. McGrew, C. J. Lai, R. T. Moon, *Dev. Biol.* **172**, 337 (1995).
- N. Funayama, F. Fagotto, P. McCreary, B. M. Gumbiner, *J. Cell Biol.* **128**, 959 (1995).
- P. J. Morin *et al.*, *Science* **275**, 1787 (1997).
- J. Behrens *et al.*, *Nature* **382**, 638 (1996); O. Huber *et al.*, *Mech. Dev.* **59**, 3 (1996).
- E. Brunner, O. Peter, L. Schweizer, K. Basler, *Nature* **385**, 829 (1997); J. Riese *et al.*, *Cell* **88**, 777 (1997).
- M. Molenaar *et al.*, *Cell* **86**, 391 (1996).
- M. van de Wetering *et al.*, *ibid.* **88**, 789 (1997).
- A. E. Munsterberg, J. Kitajewski, D. A. Bumcrot, A. P. McMahon, A. B. Lassar, *Genes Dev.* **9**, 2911 (1995); M. Maroto *et al.*, *Cell* **89**, 139 (1997); T. L. Greco *et al.*, *Genes Dev.* **10**, 313 (1996).
- S. Takada *et al.*, *Genes Dev.* **8**, 174 (1994).
- S. Noramly, J. Pisenati, U. Abbott, B. Morgan, *Dev. Biol.* **179**, 339 (1996); J. Gastrop, R. Hoevenagel, J. R. Young, H. C. Clevers, *Eur. J. Immunol.* **22**, 1327 (1992).
- M. Hollyday, J. A. McMahon, A. P. McMahon, *Mech. Dev.* **52**, 9 (1995).
- D. C. Slusarski, J. Yang-Snyder, W. B. Busa, R. T. Moon, *Dev. Biol.* **182**, 114 (1997); D. C. Slusarski, V. G. Corces, R. T. Moon, *Nature* **390**, 410 (1997).
- S. H. Hughes, J. J. Greenhouse, C. J. Petropoulos, P. Suttrave, *J. Virol.* **61**, 3004 (1987).
- B. A. Morgan and D. M. Fekete, *Methods Cell Biol.* **51**, 185 (1996).
- M. Levin, R. L. Johnson, C. D. Stern, M. Kuehn, C. Tabin, *Cell* **82**, 803 (1995); R. D. Riddle, R. L. Johnson, E. Laufer, C. Tabin, *ibid.* **75**, 1401 (1993).
- We thank A. McMahon and Y. Yang for sharing data before publication and providing the chick *Wnt3a* probe and the mouse *Wnt3a* cDNA, and B. Eshelman for technical expertise. This work was supported by a grant from the American Cancer Society to C.J.T., grants from NIH and the G. Harold and Leyla Mathers Charitable Foundation to J.C.I.B., a fellowship from the Japan Society for the Promotion of Science to M.K., and a fellowship from the Fulbright/Spanish Ministry of Education to J.C.

26 November 1997; accepted 21 April 1998

X-ray Crystal Structure of C3d: A C3 Fragment and Ligand for Complement Receptor 2

Bhushan Nagar,* Russell G. Jones,* Russell J. Diefenbach,†
David E. Isenman, James M. Rini‡

Activation and covalent attachment of complement component C3 to pathogens is the key step in complement-mediated host defense. Additionally, the antigen-bound C3d fragment interacts with complement receptor 2 (CR2; also known as CD21) on B cells and thereby contributes to the initiation of an acquired humoral response. The x-ray crystal structure of human C3d solved at 2.0 angstroms resolution reveals an α - α barrel with the residues responsible for thioester formation and covalent attachment at one end and an acidic pocket at the other. The structure supports a model whereby the transition of native C3 to its functionally active state involves the disruption of a complementary domain interface and provides insight into the basis for the interaction between C3d and CR2.

Serum complement protein C3 is a central component of host defense because its proteolytic activation is the point of convergence of the classical, alternative, and lectin pathways of complement activation. C3 cleavage products mediate many of the effector functions of humoral immunity, including inflammation, opsonization, and

cytolysis. Proteolytic cleavage of C3 into C3a and C3b exposes an internal thioester bond that through transacylation mediates covalent attachment of C3b to the surface of foreign pathogens (1). Although surface-bound C3b is itself a ligand for complement receptor 1 (CR1; also known as CD35), it can subsequently be degraded into the successively smaller fragments iC3b and C3dg, tagging the pathogen for recognition by other receptors, including the B cell complement receptor CR2 (CD21) (1). The interaction between B cell CR2 and antigen-bound iC3b or C3dg is an essential component of a normal antibody response (2), making an important link between the innate and adaptive arms of the immune system (3). The C3d fragment (a CR2-

Department of Biochemistry and Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, M5S 1A8, Canada.

*These authors contributed equally to this work.

†Present address: Centre for Virus Research, Westmead Hospital, University of Sydney, Westmead, NSW 2145; Australia.

‡To whom correspondence should be addressed at the Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada, M5S 1A8. E-mail: james.rini@utoronto.ca

binding, 35-kD, protease-resistant fragment of C3dg), when conjugated to a soluble antigen, is a potent molecular adjuvant (4).

A Cys¹⁰¹⁰ → Ala¹⁰¹⁰ mutant form of human C3d (C3 residues 996 through 1303; prepro-C3 numbering used throughout) was expressed in *Escherichia coli* in both its native and selenomethionyl-labeled forms. The alanine mutation was introduced to ensure correct disulfide bonding between Cys¹¹⁰¹ and Cys¹¹⁵⁸ (Cys¹⁰¹⁰ contributes the thiol moiety of the thioester, which would only be expected to form in native C3). A rosette inhibition assay confirmed that the recombinant C3d fragment was functionally active and bound CR2 to an extent comparable to that of serum-derived C3dg (5). Crystals were grown and the structure was determined by multi-wavelength anomalous diffraction (MAD) phasing from data collected at the selenium absorption edge (6). The structure was refined to an R_{cryst} of 19.3% and an R_{free} of 23.3% for the data in the 5.0 to 1.8 Å resolution range (Table 1). The final model contains residues 1 through 294 (vector

residues methionine and leucine plus C3 residues 996 through 1287) of the 310–amino acid construct, as well as 110 solvent molecules.

C3d is an α - α barrel with overall dimensions of approximately 30 Å × 30 Å × 50 Å (Fig. 1, A and B). This α -barrel structure, although uncommon, has been seen in glucoamylase, endoglucanase, and the β subunit of protein farnesyltransferase (7). C3d shows no significant sequence similarity with these proteins. The topology of the molecule is such that consecutive helices alternate from outside to inside, resulting in a core of six parallel helices (α 1, α 3, α 5, α 8, α 10, and α 12), surrounded by a second set of six parallel helices (T1, α 2, α 4, α 7, α 9, and α 11) running antiparallel to those of the core (Fig. 1A). The most NH₂-terminal helix (T1) is one of five (T1 through T5) short segments of 3₁₀ helix found in the structure. Although the core of the protein is formed mainly by apolar side chains, there are several polar residues pointing into the core, hydrogen bonded to buried water molecules. The two opposing ends of

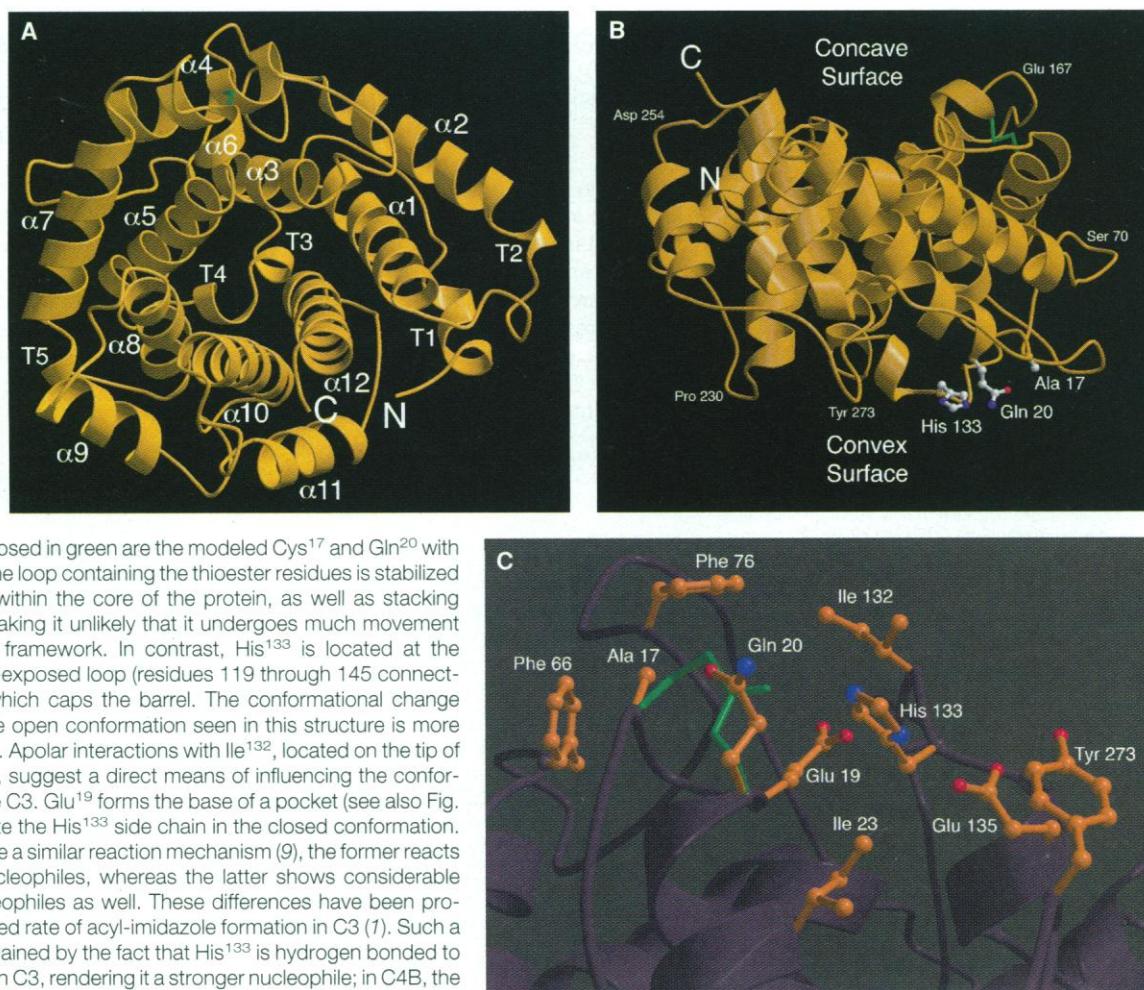
the barrel are structurally distinct (Fig. 1B): One is a convex surface presenting the amino acid side chains responsible for covalent attachment to antigen, whereas the other is a more concave surface containing an extended acidic pocket.

C3 attaches to pathogens covalently (8), predominately through ester linkages. To ensure covalent attachment to only nonself elements, an activated short-lived form of C3 is generated in close proximity to the cell surface of pathogenic microorganisms (1). Mechanistically, activation is triggered by proteolytic cleavage of C3, which initiates a series of steps (9) that closely parallel those elucidated for the B isotype of human C4 (10). Key among these steps is nucleophilic attack on the buried thioester linkage (formed by the side chains of Cys¹⁰¹⁰ and Gln¹⁰¹³) by His¹¹²⁶. The acyl-imidazole-activated Gln¹⁰¹³ side chain is then exposed to solvent, where it can be attacked by cell surface nucleophiles (for example, hydroxyl groups from glycolipids or glycoproteins of pathogens), leading to covalent attachment of C3b to antigen. In

Fig. 1. C3d structure. (A)

Ribbon representation of C3d. The single intra-chain disulfide bond is shown in green here and in (B). (B) Ribbon representation viewed into the side of the α - α barrel. Shown are the residues responsible for covalent attachment to antigen: Ala¹⁷ (Cys¹⁷ in wild-type C3d), Gln²⁰, and His¹³³ on the convex end of the molecule. Additional residues have been labeled to serve as reference points. Red and blue label oxygen and nitrogen atoms, respectively, here and in (C). (C)

The thioester region. Superimposed in green are the modeled Cys¹⁷ and Gln²⁰ with an intact thioester bond. The loop containing the thioester residues is stabilized by helix α 2, well packed within the core of the protein, as well as stacking interactions with Phe⁶⁶, making it unlikely that it undergoes much movement relative to the α - α barrel framework. In contrast, His¹³³ is located at the midpoint of a long surface-exposed loop (residues 119 through 145 connecting helices α 4 and α 5), which caps the barrel. The conformational change required to bring about the open conformation seen in this structure is more likely to occur in this region. Apolar interactions with Ile¹³², located on the tip of the His¹³³-containing loop, suggest a direct means of influencing the conformation of this loop in native C3. Glu¹⁹ forms the base of a pocket (see also Fig. 2B) that may accommodate the His¹³³ side chain in the closed conformation. Although C3 and C4B share a similar reaction mechanism (9), the former reacts primarily with hydroxyl nucleophiles, whereas the latter shows considerable reactivity with amino nucleophiles as well. These differences have been proposed to reflect an increased rate of acyl-imidazole formation in C3 (7). Such a rate increase could be explained by the fact that His¹³³ is hydrogen bonded to negatively charged Glu¹³⁵ in C3, rendering it a stronger nucleophile; in C4B, the residue equivalent to Glu¹³⁵ is a serine, which is probably unable to perform a similar function. [The figure was prepared with MOLSCRIPT (33) and RASTER3D (34).]



a competing reaction, the acyl-imidazole intermediate is rapidly hydrolyzed, thereby limiting the lifetime of the transacylation-competent state. The residues involved in forming the thioester bond [Cys¹⁷ (1010 in C3; here mutated to Ala¹⁷) and Gln²⁰ (1013 in C3)] and the acyl-imidazole intermediate [Gln²⁰ and His¹³³ (1126 in C3)] are surface-exposed in C3d and located on the convex surface of the molecule (Fig. 1, B and C). The Ala¹⁷ residue is situated on

the β turn preceding α 1, and Gln²⁰ is the NH₂-terminal residue of α 1, positioning them in close proximity. Although the side chain of Gln²⁰ is somewhat disordered and Cys¹⁷ has been mutated to alanine in this structure, modeling shows that the cysteine sulfur and the carbonyl carbon of the glutamine side chain can be brought to within 2.0 Å of each other (close to a typical S–C bond length) with little movement of the peptide backbone (Fig. 1C). It would ap-

pear, then, that the relative disposition of these residues is probably close to that found in native C3. In contrast, His¹³³ is misoriented and too far away from the modeled thioester linkage (4.1 Å) to form the acyl-imidazole intermediate with Gln²⁰ (Fig. 1C). Small changes in the relative orientation of the reactive groups in the active sites of enzymes are crucial in governing the rate of catalysis (11). In this case, a conformational change in the main

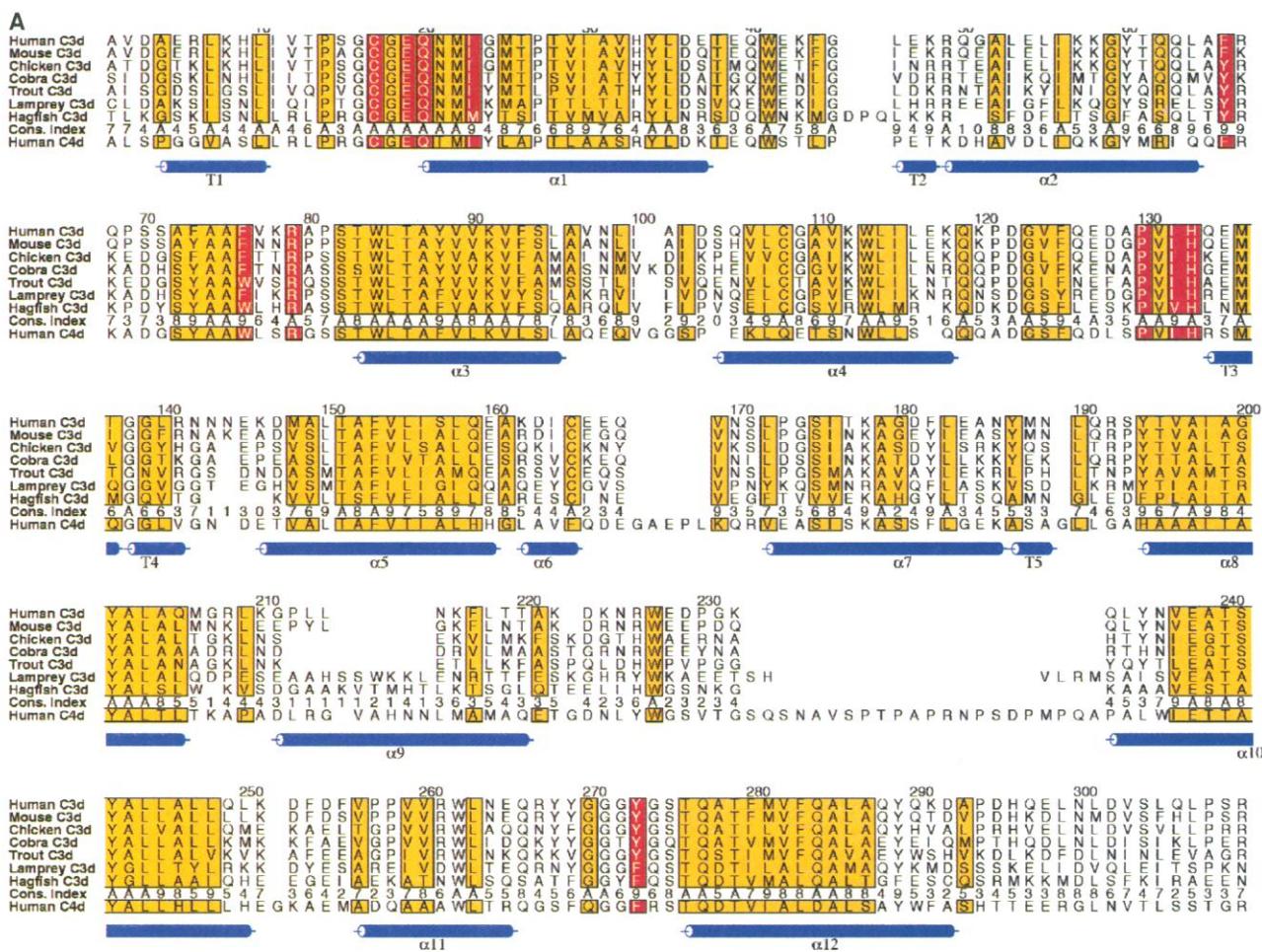
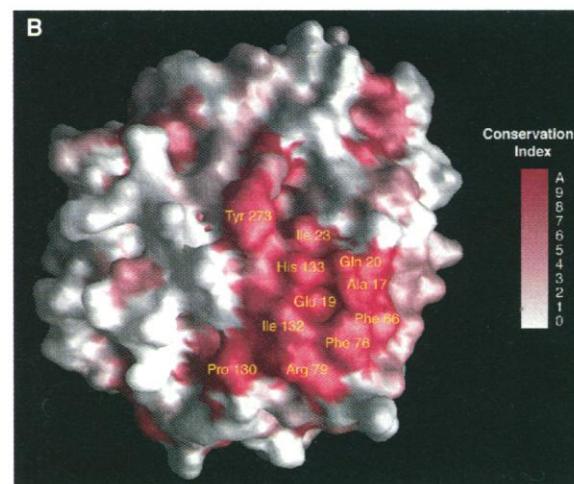


Fig. 2. Sequence conservation of C3d. **(A)** Multiple sequence alignment of selected species of C3d and human C4d (B isotype) (21). Residues shaded in yellow are at least 90% buried in the C3d structure, and those shaded in red are residues composing the contiguous surface patch labeled in **(B)**. Numbers correspond to the degree of conservation in C3d sequences only: 0 (not conserved) to A (highly conserved), as determined by the program AMAS (32). In human C4d, approximately 75% of the core residues, as well as the putative domain interface residues, are highly conserved [a conservation index (cons. index) of 7 or higher when included in the AMAS calculation], which suggests that it will adopt a similar fold and possess the analogous domain interface. The helical segments in human C3d are indicated by blue cylinders. [The figure was prepared with ALSCRIPT (35).] **(B)** Mapping of residue conservation as determined in **(A)** onto the surface of C3d; white (not conserved) to progressively darker red (highly conserved). [The figure was prepared with GRASP (36).] The conserved patch includes most of the surface apolar residues shown in Fig. 1C.



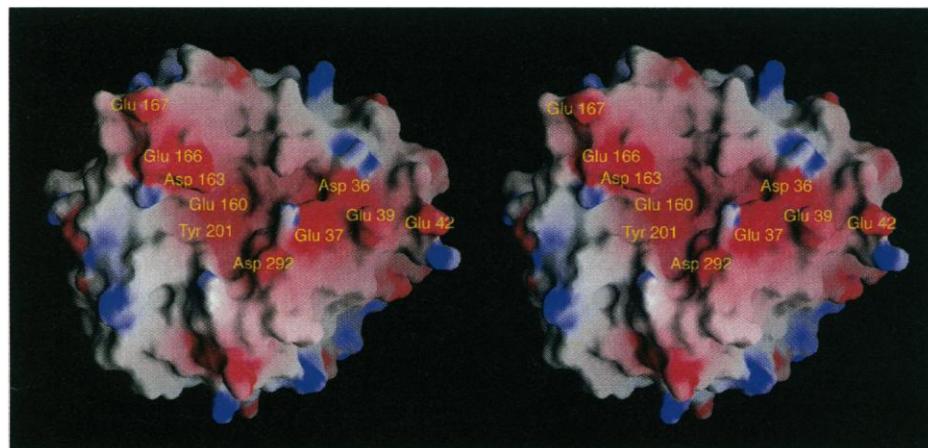


Fig. 3. Stereo view of an electrostatic surface rendition of C3d, showing the acidic pocket on the concave end of the molecule. Acidic and basic residues are colored red and blue, respectively. Labeled are the surface-exposed residues that form the pocket. The contour level is at ± 10 kT. [The figure was prepared with GRASP (36).]

Table 1. Structure determination and refinement. The structure was determined by the MAD phasing method. Three data sets were collected at the Se absorption edge corresponding to the peak, edge, and a remote wavelength. Atomic positions for the 8 Se atoms were obtained by means of Patterson methods, with the aid of the program HEAVY (24), and difference Fourier maps. The three data sets were then treated as a multiple isomorphous replacement problem, using the remote wavelength as “native” and the peak and edge wavelengths as “derivatives.” The selenium positions were refined and phases were calculated with the program PHASES (25). The resulting experimental map was of very high quality, with an overall figure of merit of 0.8 to 2.0 Å (0.9 after solvent flattening) and was traced with the program O (26). The initial R_{cryst} and R_{free} values (7% of the reflections set aside for the test set) of the model before refinement were 37.1 and 37.4%, respectively, in the 8.0 to 1.8 Å resolution range against the native data set. The model was then refined with the program X-PLOR (27). rsm, root mean square deviation. After the model was subjected to simulated annealing, successive rounds of manual rebuilding, and positional/B-factor refinement, as well as the application of an overall anisotropic B factor to the data, the R_{cryst} and R_{free} values were reduced to 19.3 and 23.3%, respectively (the 5.0 to 1.8 Å range). The final model contained 294 residues, 104 water molecules, and six glycerol molecules. Electron density for the six NH_2 -terminal residues (1 through 6), residues 42 through 46, residue 143, and residues 166 through 171 was weak. No electron density was present for residues 295 through 310. Relative residue accessibility was calculated with the program NACCESS (28), secondary structure was assigned with the use of the program PROMOTIF (29), and structural similarity was searched for with the program DALI (30). The multiple sequence alignment was performed with ClustalW (27), and conservation indices were calculated with AMAS (32).

| | Native | MAD (SeMet) | | |
|---|---------|---------------------------------|--------|--------|
| | | Remote | Edge | Peak |
| Diffraction data | | | | |
| Wavelength (Å) | 0.993 | 0.9678 | 0.9794 | 0.9792 |
| Resolution (Å) | 1.8 | 2.0 | 2.0 | 2.0 |
| Temperature (°C) | -160 | -160 | -160 | -160 |
| Measured reflections (<i>n</i>) | 103308‡ | 128694 | 125648 | 130972 |
| Unique reflections (<i>n</i>) | 29818‡ | 41245 | 40622 | 40791 |
| Completeness (%) | 80.4§ | 91.2 | 89.3 | 89.8 |
| R_{sym}^* | 0.046 | 0.052 | 0.054 | 0.055 |
| Sites (<i>n</i>) | — | 8 | 8 | 8 |
| Phasing power† | | | | |
| Dispersive | — | — | 2.5 | 1.8 |
| Anomalous | — | 3.4 | 4.8 | 5.0 |
| Figure of merit (before solvent flattening) | | 0.801 | | |
| Refinement statistics | | | | |
| Resolution (Å) | 5.0–1.8 | rmsd bond length (Å) | | 0.006 |
| Reflections (<i>n</i>) ($F > 2\sigma$) | 24149 | rmsd bond angle (°) | | 1.203 |
| R_{cryst} | 0.193 | rmsd B values (Å ²) | | 2.01 |
| R_{free} | 0.233 | | | |

* $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections. †Phasing power, root mean square (rms) $F_{\text{H}} / \text{rms } \epsilon$, where ϵ is lack of closure and F_{H} is the calculated heavy atom structure factor. ‡Bijvoet mates were averaged for the native data set. §91% complete in the 2.07 to 2.02 Å shell; 59% complete in the 2.02 to 1.79 Å shell.

chain would be required to bring the thioester linkage and histidine side chain into position for nucleophilic attack and covalent bond formation (see Fig. 1C). We see no evidence of disorder or alternate conformations in the polypeptide segments containing these residues, which suggests that in native C3 this region must be held in a strained conformation capable of forming the acyl-imidazole intermediate, termed here the “closed” conformation. This structure showing the large separation between the histidine and glutamine residues (see Fig. 1C) would then represent the relaxed or “open” conformation.

A number of surface-exposed apolar residues, including Ile²³, Phe⁶⁶, Phe⁷⁶, Ile¹³², and Tyr²⁷³ (Fig. 1C), flank the thioester region. Sequence comparisons of the C3d region from diverse species show that these residues are also highly conserved (Fig. 2A). In fact, when the degree of similarity for each position in the sequence (Fig. 2) is mapped onto the structure, we find a large, contiguous, surface-exposed patch, including the apolar residues, whose degree of conservation is as high as that of residues in the core (Fig. 2B). We interpret the conservation of this patch to be a reflection of an important structural role, namely, that of a domain interface in native C3. We further suggest that the domain interface not only serves to protect the thioester from solvent hydrolysis but, in addition, provides some, if not all, of the strain energy required to hold this region in the closed conformation. At the same time, interactions with His¹³³ must prevent nucleophilic attack on the thioester in native C3. Upon C3 cleavage and subsequent disruption of this interface, the strain energy would be transferred to the covalent bond formed between His¹³³ and Gln²⁰, enhancing the reactivity of the acyl-imidazole intermediate while holding the structure in the closed conformation. After nucleophilic attack and relief of the strain, conversion to the open conformation would occur, precluding the possibility of reforming the acyl-imidazole linkage. It has been shown (10) that the Cys¹⁷ thiolate anion released after acyl-imidazole formation serves to base catalyze the subsequent nucleophilic attack, an observation supported by the relative dispositions of Ala¹⁷, Gln²⁰, and His¹³³ (Fig. 1C).

The presence of a domain interface burying the thioester region and providing the strain energy to hold native C3 in the closed conformation is consistent with a number of chemical inactivation experiments. Even in the absence of proteolytic activation, small nucleophiles such as methylamine and ammonia are known to cleave the thioester linkage, resulting in a slow (~10-hour) conformational change to

the so-called C3b-like state (12). The associated increase in volume and polarity of the cysteine and modified glutamine side chains (spontaneous hydrolysis by water would lead to glutamic acid) would be expected to disrupt the steric and chemical complementary of the domain interface, thermodynamically destabilizing the molecule. Furthermore, mildly denaturing agents are known to promote covalent attachment of C3 to hydroxyl-bearing nucleophiles or the competing thioester hydrolysis reaction (13), again presumably by disrupting the domain interface. Mutations of Cys¹⁷, Glu¹⁹, Gln²⁰, and His¹³³ show, however, that variations in the interface, including those leading to the loss of the thioester linkage (Cys¹⁷ → Ala¹⁷, Glu¹⁹ → Gln¹⁹, and Gln²⁰ → Asn²⁰), do not necessarily preclude a C3-like conformation (14). These results are consistent with the ability of proteins to accommodate changes in core packing or protein-protein interfaces while still retaining function (15). In any case, steric clashes associated with increased bulk, as with methylamine addition, would be expected to be the most destabilizing.

Efforts have been made to identify C3d residues involved in binding the CR2 receptor, and a peptide segment has been proposed (residues 228 through 239 in our structure) (16). However, no consensus has yet been reached, because extensive mutagenesis of this segment (in part located on the convex surface) did not substantially affect binding (17). The CR2 receptor itself consists of 15 to 16 short consensus repeats of 60 to 70 amino acids, only the two NH₂-terminal repeats of which are involved in C3d binding (18). Because C3d interacts with pathogen surfaces through a covalent link at Gln²⁰, it is unlikely that the thioester-containing end of the α-α barrel would be involved in CR2 interactions. The structure shows that the opposite end of the barrel (the concave surface in Fig. 1B), which would be completely accessible for receptor interactions even when bound to cell surfaces, presents a potential site of interaction—that of an extended pocket formed primarily by acidic residues (Fig. 3). A number of the residues found in the pocket are highly conserved among diverse C3d sequences (Asp³⁶, Glu¹⁶⁰, and Tyr²⁰¹). This pocket corresponds in location to the substrate binding sites in the α-α barrels of glucoamylase, endoglucanase, and farnesyltransferase. With respect to the receptor, inhibition studies have identified CR2-derived peptides, containing basic residues, that are important in C3d binding (19).

Homology modeling [based on the structure of a two-domain fragment of human factor H (20)] shows that these basic residues, among others, form an extensive positively charged surface on the two NH₂-terminal repeats of human CR2. These observations make the negatively charged pocket, on the concave end of the barrel, an attractive candidate for the site of CR2 interaction.

REFERENCES AND NOTES

1. S. K. Law and A. W. Dodds, *Protein Sci.* **6**, 263 (1997).
2. M. B. Fischer *et al.*, *J. Immunol.* **157**, 549 (1996); J. M. Ahearn *et al.*, *Immunity* **4**, 251 (1996).
3. D. T. Fearon and R. H. Carter, *Annu. Rev. Immunol.* **13**, 127 (1995).
4. P. W. Dempsey, M. E. Allison, S. Akkaraju, C. C. Goodnow, D. T. Fearon, *Science* **271**, 348 (1996).
5. Expression and purification were done as follows. A human C3d cDNA corresponding to residues 996 through 1303 containing the Cys¹⁰¹⁰ → A¹⁰¹⁰ mutation (14) was inserted into the bacterial expression plasmid pET 15b (Novogen), in which the NH₂-terminal oligo-His encoding sequence has been deleted. The encoded C3d fragment contained an additional eight amino acids (MLDAERLK) (21) at the NH₂-terminus that would not be present in enzymatically produced C3d, ML being vector derived and DAERLK corresponding to residues at the COOH-terminus of C3g. Protein expression in transformed *E. coli* strain BL21 (DE3) grown in Luria Bertoni broth with ampicillin was induced with 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 28°C for 12 hours. Selenomethionine (SeMet) C3d was expressed in the same cells grown in ampicillin-containing M9 minimal medium supplemented with 0.4% glucose, MgSO₄ (100 mg/ml), CaCl₂ (15 mg/ml), and thiamine (1.25 mg/ml). After IPTG induction, the medium was further supplemented with the amino acids K, L, I, V, T, and F (100 mg/ml each) and seleno-L-methionine (50 mg/ml) (22). Both native C3d and the SeMet C3d were purified from the soluble fraction of the bacterial lysate by DEAE-Sepharose, then by Mono Q HR10/10 fast protein liquid chromatography (FPLC) (Pharmacia), both at pH 7.1, and finally by Mono S HR5/5 FPLC (Pharmacia) at pH 6. Amino acid composition analysis of SeMet C3d showed that the replacement of the eight Met positions by SeMet was near quantitative. Recombinant C3d Cys¹⁰¹⁰ → Ala¹⁰¹⁰ was assessed for binding to human Raji B cell-associated CR2 with a rosette inhibition assay (17). Serum-derived C3dg was used as a positive control. As monomeric ligands, both proteins inhibited 50% of rosette formation at 0.2 μM.
6. Crystallization and data collection were done as follows. Crystals were grown by the vapor diffusion method from protein drops [10 mg/ml in 10 mM Pipes (pH 6.0) and 100 mM NaCl] equilibrated against well solution (1 ml) containing 12% polyethylene glycol 20K and 100 mM MES buffer (pH 6.5). Rodlike crystals grew within 1 day in space group P2₁2₁2₁ (a = 60.8 Å, b = 64.4 Å, c = 87.1 Å), with one molecule in the asymmetric unit (50% solvent content). The SeMet-substituted crystals were grown under the same conditions in the presence of 10 mM dithiothreitol (DTT). Although DTT was essential for good crystal growth, both biochemical criteria and the structure showed that the single intrachain disulfide of C3d remained intact. All data sets were collected at the Cornell High-Energy Synchrotron Source on beamline F2 with the use of the Princeton 1K charge-coupled device detector. The data were collected from crystals flash-frozen
7. A. Aleshin, A. Golubev, L. M. Firsov, R. B. Honzatko, *J. Biol. Chem.* **267**, 19291 (1992); P. M. Alzari, H. Souchon, R. Dominguez, *Structure* **4**, 265 (1996); H. W. Park, S. R. Boduluri, J. F. Moomaw, P. J. Casey, L. S. Beese, *Science* **275**, 1800 (1997).
8. S. K. Law and R. P. Levine, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2701 (1977); B. F. Tack, R. A. Harrison, J. Janatova, M. L. Thomas, J. W. Prahl, *ibid.* **77**, 5764 (1980).
9. M. Gadjeva *et al.*, *J. Immunol.*, in press.
10. A. W. Dodds, X. D. Ren, A. C. Willis, S. K. Law, *Nature* **379**, 177 (1996).
11. A. D. Mesecar, B. L. Stoddard, D. E. Koshland Jr., *Science* **277**, 202 (1997).
12. D. E. Isenman, D. I. Kells, N. R. Cooper, H. J. Müller-Eberhard, M. K. Pangburn, *Biochemistry* **20**, 4458 (1981).
13. S. K. Law, *Biochem. J.* **211**, 381 (1983).
14. L. Isaac and D. E. Isenman, *J. Biol. Chem.* **267**, 10062 (1992); L. Isaac *et al.*, *Biochem. J.* **329**, 705 (1998).
15. A. E. Eriksson *et al.*, *Science* **255**, 178 (1992); F. M. Richards and W. A. Lim, *Quart. Rev. Biophys.* **26**, 423 (1993); J. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1 (1996).
16. J. D. Lambiris, V. S. Ganu, S. Hirani, H. J. Müller-Eberhard, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4235 (1985).
17. R. J. Diefenbach and D. E. Isenman, *J. Immunol.* **154**, 2303 (1995).
18. T. Hebell, J. M. Ahearn, D. T. Fearon, *Science* **254**, 102 (1991).
19. H. Molina *et al.*, *J. Immunol.* **154**, 5426 (1995).
20. P. N. Barlow *et al.*, *J. Mol. Biol.* **232**, 268 (1993).
21. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
22. G. D. Van Duyne, R. F. Standaert, P. A. Karplus, S. L. Schreiber, J. Clardy, *J. Mol. Biol.* **229**, 105 (1993).
23. Z. Otwinowski, in *Proceedings of the CCP4 Study Weekend*, L. Sawyer, N. Isaacs, S. Bailey, Eds. (Science and Engineering Research Council Daresbury Laboratory, Warrington, UK, 1993), pp. 56–62.
24. T. C. Terwilliger, S.-H. Kim, D. Eisenberg, *Acta Crystallogr.* **A43**, 1 (1987).
25. W. Furey and S. Swaminathan, *Methods Enzymol.* **277** (part B), 590 (1997).
26. T. A. Jones, J. Y. Zou, S. W. Cowan, Kjelgaard, *Acta Crystallogr.* **A47**, 110 (1991).
27. A. Brünger, *X-PLOR, v3.1 A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
28. S. J. Hubbard and J. M. Thornton, NACCESS computer program, Department of Chemistry and Molecular Biology, University College London, 1993.
29. E. G. Hutchinson and J. M. Thornton, *Protein Sci.* **5**, 212 (1996).
30. L. Holm and C. Sander, *J. Mol. Biol.* **233**, 123 (1993).
31. J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res.* **22**, 4673 (1994).
32. C. D. Livingstone and G. J. Barton, *Comput. Appl. Biosci.* **9**, 745 (1993).
33. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
34. E. A. Merritt and M. E. P. Murphy, *Acta Crystallogr.* **D50**, 869 (1994).
35. G. J. Barton, *Protein Eng.* **6**, 37 (1993).
36. A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **11**, 281 (1991).
37. This work was supported by Medical Research Council of Canada grants to J.M.R. (MT-12458) and D.E.I. (MT-7081) and by an Ontario Graduate Scholarship to B.N. Coordinates have been deposited in the Brookhaven Protein Data Bank.

8 January 1998; accepted 20 March 1998