The crystal structure of a complement-1q family protein suggests an evolutionary link to tumor necrosis factor

Lawrence Shapiro* and Philipp E. Scherer†

ACRP30 – adipocyte complement-related protein of 30 kDa or AdipoQ – is an abundant serum protein, secreted exclusively from fat cells, which is implicated in energy homeostasis and obesity [1,2]. ACRP30 is a close homologue of the complement protein C1q, which is involved in the recognition of microbial surfaces [3–5] and antibody–antigen complexes [6,7] in the classical pathway of complement. We have determined the crystal structure of a homotrimeric fragment from ACRP30 at 2.1 Å resolution. The structure reveals an unexpected homology to the tumor necrosis factor (TNF) family. Identical folding topologies, key residue conservations, and similarity of trimer interfaces and intron positions firmly establish an evolutionary link between the TNF and C1q families. We suggest that TNFs – which control many aspects of inflammation, adaptive immunity, apoptosis and energy homeostasis – arose by divergence from a primordial recognition molecule of the innate immune system. The evolutionary connection between C1q-like proteins and TNFs illuminates the shared functions of these two important groups of proteins.

Addresses: *Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine, 1425 Madison Avenue, New York, New York 10029, USA. †Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA.

Correspondence: Lawrence Shapiro
E-mail: shapiro@convex.hhmi.columbia.edu

Received: 29 December 1997
Revised: 16 January 1998
Accepted: 16 January 1998

Published: 2 March 1998

Current Biology 1998, 8: 335–338
http://biomednet.com/elecref/0960982200800335

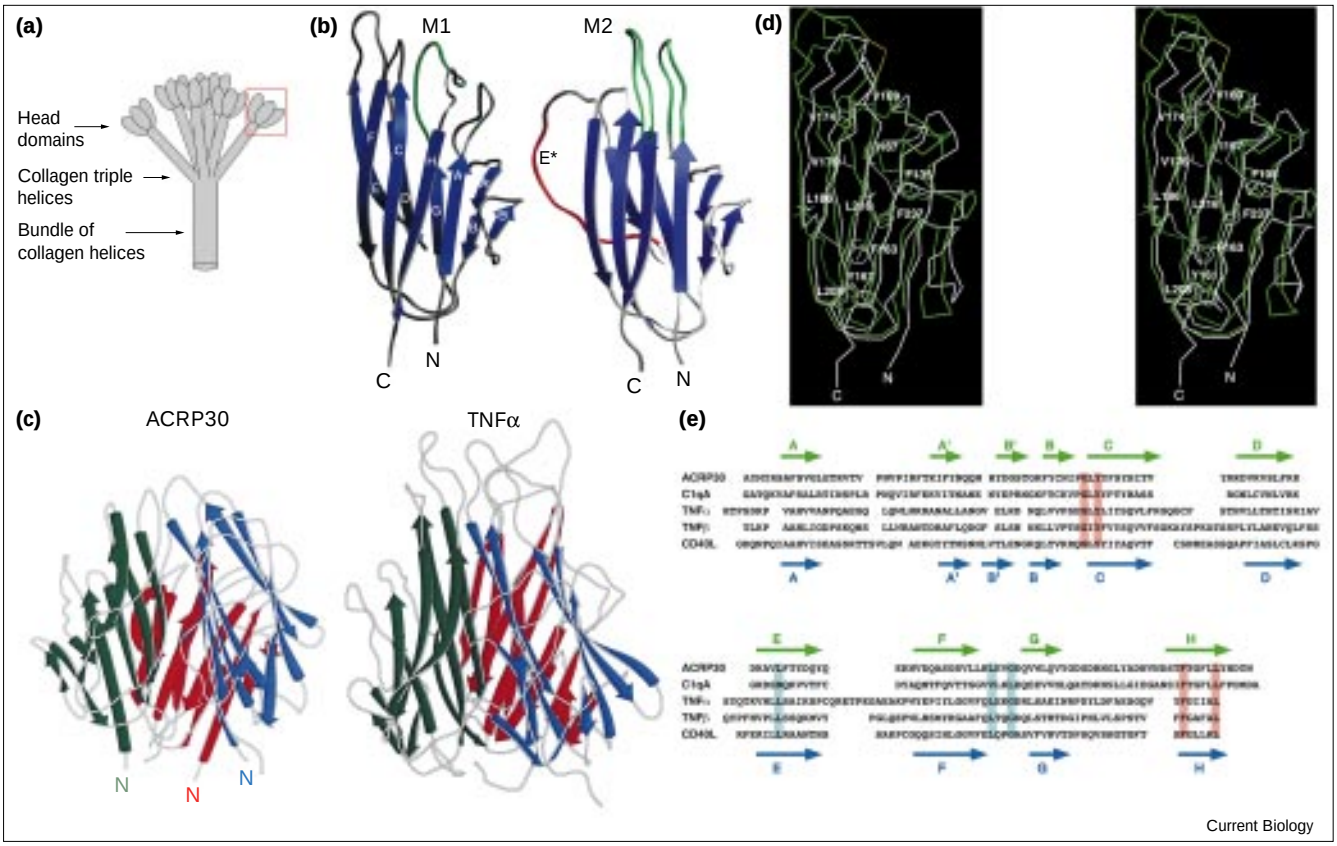
Results and discussion

C1q family proteins are characterized by a distinctive 'globular domain' of about 140 amino acids (the gC1q domain) situated at the carboxyl terminus of a collagen 'stalk'. They often form a characteristic superstructure in which three protomers trimerize to form a collagen triple helix, and these trimers multimerize to form a 'bouquet' (Figure 1a) [6]. Whereas C1q is composed of heterotrimeric subunits, the subunits of ACRP30 are homotrimeric. C1q family members can be expressed as soluble plasma proteins or type II membrane-bound molecules [8]. TNFs are also found in both of these forms [9]. TNFs are trimeric, and although they do not have canonical collagenous regions, they do have stalk regions that, like those of collagens, are often rich in glycine and proline residues. We expressed as a maltose-binding protein fusion a recombinant protein corresponding to the globular domain (residues 111–247) of ACRP30, which has a multimeric structure similar to C1q [1,2,10]. Crystals of this domain were obtained that belong to space group $P6_1$ (a = b = 112.3 Å, c = 71.6 Å), have one trimer per asymmetric unit, and diffract to 2.1 Å Bragg spacings. The structure was determined by multiple isomorphous replacement (MIR), and refined to a final R value of 21.2% (free R value 27.6%) using all data to 2.1 Å resolution (Figure 2, Table 1).

The structure reveals an asymmetrical trimer of β-sandwich protomers, each of which has a ten-strand jelly-roll folding topology identical to TNFs (Figure 1c, [11,12]). We infer from sequence homologies that this fold is common to all C1q family proteins. The trimer (Figure 1c) is bell-shaped, with a wide base; trimer contacts form primarily through a cluster of hydrophobic interactions near the base. The trimer interface near the apex is largely hydrophilic, and contains many ordered water molecules. These overall features are in common with TNF family trimers [11,12]. The collagenous segments, not included in the fragment studied here, would protrude from the amino termini at the center of the trimer. The amino and carboxyl termini are directly adjacent to one another, suggesting that gC1q domains might assemble as either amino-terminal or carboxy-terminal appendages to collagen stalks, in the same way as C-type lectin domains do [13]. C1q proteins with amino-terminal gC1q domains have not, so far, been observed, however. The C-type lectin domains, which also function in innate immunity, have a C1q-like collagen-based superstructure, but have a very different protein fold. This includes a 'neck' region composed of a coiled coil of three α helices that accounts for the major part of the trimer interface [14]; C-type lectin domains lacking this neck region fail to trimerize. In gC1q domains, the globular domain leads directly into the Gly–X–Y (where X and Y are any amino acid) repeats of the collagenous segment, with no intervening neck region.

The folding of collagen triple-helices is thought to nucleate at the carboxyl terminus, where precise alignment of the three polypeptide chains is generally established by a non-collagenous trimerization region; in types VIII and X collagen, these trimerization regions are related to gC1q domains ([15] and references therein). The ACRP30 structure shows that gC1q domains form strong trimers stabilized by a central hydrophobic interface (5324 $Å^2$ buried surface area, excluding water molecules), suggesting a structural

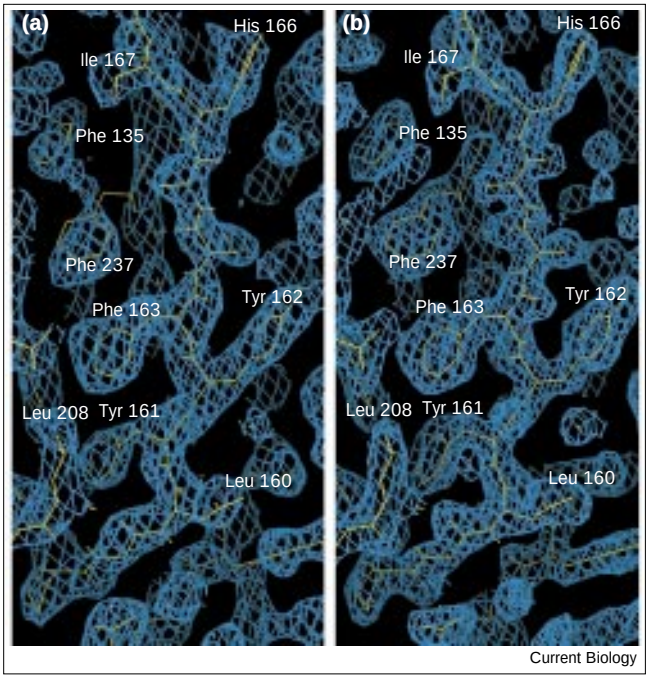
Figure 1



Structure of the C1q family protein ACRP30 and relationship to TNFs. (a) The quaternary structure of C1q proteins. Three protomers form a collagen triple helix capped by a trimeric gC1q region. These trimers oligomerize to form a 'bouquet'. For C1q, the bouquet contains six trimers, resulting in a complex of 18 polypeptide chains, whereas ACRP30 appears to comprise four trimers. A portion corresponding to a single trimeric head domain is boxed in red. (b) Ribbon diagram of an isolated ACRP30 gC1q domain. The β-strand topology (labeled on molecule 1, M1) is identical to that of TNFs, so we have adopted the secondary structure nomenclature used for this family. Disordered loops are drawn in green. The asymmetry of the ACRP30 trimer is illustrated by comparison with M2, which shows the largest structural deviation from M1 (M3, not shown, has a structure similar to M1). The most striking difference is in the E-strand region; this secondary structure element melts to become a loop (E*) in M2 which wraps around the outer edge of its M1 partner as part of their trimer interface. B values for atoms in the E strand are not markedly higher than for other atoms in the structure. The substantial asymmetry of the ACRP30 trimer may relate to asymmetry of C1q which has three related but distinct chains, C1qA, B, and C. (c) Ribbon diagram of the ACRP30 gC1q domain trimer, and comparison with the TNFα trimer. The amino terminus of each ACRP30 protomer is labeled. The collagen stalks, not present in the molecule used in this study, would protrude from these amino termini. The three-fold axis is non-crystallographic, and considerable structural differences are observed between the subunits of the ACRP30 trimer. (d) Stereo view of superposition of Cα traces from ACRP30 (white) and CD40L (green, from PDB entry 1ALY). Forty-eight pairs of corresponding Cα atoms have an r.m.s. displacement of 0.78 Å. This level of structural similarity is on a par with that between different TNFs; for example, TNFα and CD40L superpose such that 65 pairs of corresponding Cα atoms have an r.m.s. displacement of 0.71 Å. Some of the hydrophobic core side chains that are matched in the structure-based sequence alignment are shown. The corresponding side chains take on remarkably similar orientations and thus conserve the overall packing of the cores in the two molecules. (e) Structure-based sequence alignment between TNFs and the gC1q domains from ACRP30 and C1q [29]. Residues that are identical in all five proteins are shaded red, and those that are conserved in four out of the five are shaded light blue. The β-strand regions are indicated for ACRP30 by green arrows above the alignment, and below the alignment in dark blue for CD40L.

basis for their role in collagen folding. These domains can also function in molecular recognition. For example, C1q binds through its head domains to a defined region within the C_H2 domain of antigen-bound immunoglobulin G (IgG) molecules [16]. Chemical modification studies of C1q have implicated two regions of the head domain in immune-complex IgG binding [17]; these are in the C1qB-chain (site 1, localized to residues 114–129) and in the A and C chains (site 2, both around residue 160). Each of these maps to the exterior of the trimer. Site 2 maps to the beginning of the E strand or the DE loop. Site 1 encompasses a face of the molecule that runs from the meandering A–A′ loop through the B′ strand, and is thus attractive as a candidate binding surface. Remarkably, the analogous regions of TNFs are directly involved in interaction with their receptors (see Figure S1 in Supplementary material).

Figure 2



Representative electron density. (a) Experimental MIR/solvent-flattened map computed with structure factors from 8.0 to 2.4 Å Bragg spacings. (b) Electron density map calculated with $2|F_o|-|F_c|$ coefficients and model phases from 8.0 to 2.1 Å resolution. Each map is contoured at 1.0σ.

We superposed the ACRP30 structure on the three known structures of molecules from the TNF family, TNFα, TNFβ and CD40 ligand (CD40L), and used the superpositions (Figure 1d and Figure S2 in Supplementary material) to generate structure-based sequence alignments (Figure 1e). Each of the ten β strands of ACRP30 can be simultaneously superposed with the ten β strands of each TNF; the relative positions and lengths of these β strands are almost identical between ACRP30 and the TNFs. Four residues are conserved throughout both the C1q and TNF families: Tyr161, Gly159, Phe237, and Leu242 (ACRP30 numbering). Each of these residues is important in the packing of the protomer's hydrophobic core. The structures of the hydrophobic cores of gC1q domains and TNFs are similar; side chains in analogous positions often have similar orientations (Figure 1d). TNF and C1q proteins also have similar gene structures: the globular domains of TNFs [18] (with a single exception) and gC1q proteins [19] are each encoded entirely within one exon, and introns in both families are restricted to the respective amino-terminal stalk regions. These data suggest that these two groups of proteins arose by divergence from a common precursor molecule, and establish a C1q/TNF molecular superfamily.

Proteins of the TNF and C1q families share many functional similarities, which can now be examined in the light of their apparent evolutionary connection. Both groups of proteins appear to play major roles in immunity as well as in energy homeostasis. Archetypal TNFs involved in immunity include CD40L, TNFα and FAS ligand (FASL) [9]. CD40L knockout mice are defective in antibody class-switching, and lack all immunoglobulins except IgM in the T-dependent humoral immune response [20]. Curiously, C1q knockout mice also exhibit defects in class-switching (A.J. Cutler, M. Botto, K.A. Davies, D. Gray, M.J. Walport, presented at the IVth International Workshop on C1 and Collectins, 1997) suggesting that these structurally related molecules may play similar roles. TNFs have mono-specific receptors (TNFRs) [21], which are characterized by a distinctive cysteine-repeat motif in their extracellular domains, and apparently arose by co-evolution of the receptor–ligand pairs. All well-characterized TNFRs signal by induced trimerization on binding of trimeric ligands, an uncommon mechanism of signal transduction. C1q has often been implicated in signalling processes, but the search for receptors for C1q family proteins, such as ACRP30, has proved problematic. The conserved trimeric structure of C1q-like proteins suggests that they might also signal through TNFR-like receptors. The highly multimeric superstructure of C1q family proteins could provide a powerful means

Table 1

(a) Diffraction data

Derivative	d_{min} (Å)	Wavelength (Å)	Number of reflections	Completeness (%)	R_{merge} (%)*
Native	2.1	0.9841	28656	93.6	9.1 (19.9)
K_2IrCl_6	2.1	0.9841	54777†	92.7	5.8 (15.7)
HgOAc	2.9	1.5418	2758	95.6	5.8 (13.7)

(b) MIR analysis (15–2.4Å)

		K_2IrCl_6	HgOAc
R_{iso}(I)		0.152	0.316
R_{Cullis}	acentric	0.90	0.67
	centric	0.84	0.63
Phasing power:	acentric	0.5	1.3
	centric	0.4	1.1

3485 reflections; 89.6% complete
MIR $\langle m \rangle$ = 0.346
DM $\langle m \rangle$ = 0.704
18180 reflections; 90.1% complete

(c) Refinement and model

Resolution (Å)	8–2.1
R value	0.212
Free *R* value	0.276
Number reflections	27888
Number total atoms	3132
Protein atoms	2742
Water molecules	390
R.m.s. bond (Å)	0.012
R.m.s. angle (°)	1.899
Average B factor (Å^2)	22.8

*Numbers in parentheses give the R_{merge} in the highest-resolution 0.1 Å shell. †Bijvoet mates were treated as inequivalent for the K_2IrCl_6 data. Heavy atom positions of the derivatives were found by inspection of Patterson maps and subsequent difference Fourier analysis.

for signal transduction through high-valency receptor clustering, even from a single ligand molecule.

Like many complement proteins, TNFα is produced in response to infection and effects multiple responses, including inflammation, cell proliferation and cell death [9]. TNFα also regulates the expression levels of some downstream components of the complement system [22,23] and plays a role in energy homeostasis [24,25] where it is implicated in cachexia, obesity and insulin resistance. It is also a major secretory product of adipocytes [24]. Similar activities have been observed for C1q family proteins: ACRP30 is made exclusively in adipocytes, and its expression is dysregulated in various forms of obesity [2]. ACRP30 secretion is acutely stimulated by insulin [1] and repressed by chronically elevated levels of insulin (P.E.S., unpublished observations). Another C1q-like molecule, the Hib27 protein from Siberian chipmunks, also seems to be involved in energy homeostasis, as its expression is specifically extinguished during hibernation [26]. Functional connection between adipose tissue and immunity has been noted before: adipocytes secrete high levels of complement factors C3 and B, and provide the unique site of synthesis for complement factor D [27].

FASL activates signalling through FAS, which controls apoptosis in the immune system and other cells [28]. FAS signalling provides a selection mechanism in cellular immunity, and also provides a means of killing cells infected by intracellular pathogens through apoptosis. Like other TNFs, FASL may have arisen from a precursor that functioned as a recognition molecule in innate immunity, much like C1q. This suggests that TNFs, which control many aspects of adaptive immunity, may have arisen by divergence from the recognition molecules of the ancient innate immune system.

Supplementary material

Full methodological detail and figures showing a superposition of Cα traces from ACRP30 and TNFα, and a comparison of putative binding surfaces of C1q and TNF family members, are available with the internet version of this paper.

Acknowledgements

We thank C.M. Ogata and W.A. Hendrickson for help with data collection at beam-line X4A, M. Lisanti for help in the earliest stages of this work, and D.H. Fremont, Y. Choi, A.K. Aggarwal and P.D. Kwong for helpful comments. Beamline X4A at the National Syncrotron Light Source at Brookhaven National Lab, a DOE facility, is supported by the Howard Hughes Medical Institute. Coordinates have been deposited in the Brookhaven Protein Data Bank.

References

1. Scherer PE, Williams S, Fogliano M, Baldini G, Lodish HF: A novel serum protein similar to C1q, produced exclusively in adipocytes. *J Biol Chem* 1995, 270:26746-26749.
2. Hu E, Liang P, Spiegelman BM: AdipoQ is a novel adipose-specific gene dysregulated in obesity. *J Biol Chem* 1996, 271:10697-10703.
3. Clas F, Loos M: Antibody-independent binding of the first component of complement (C1) and its subcomponent C1q to the S and R forms of ***Salmonella minnesota***. *Infect Immun* 1981, 31:1138-1144.
4. Clas F, Golecki JR, Loos M: Electron microscopic study showing antibody-independent binding of C1q, a subcomponent of the first component of complement, to serum-sensitive ***salmonellae***. *Infect Immun* 1984, 45:795-797.
5. Alberti S, Marques G, Hernandez-Alles S, Rubires X, Tomas JM, Vivanco F, Benedi VJ: Interaction between complement subcomponent C1q and the ***Klebsiella pneumoniae*** porin OmpK36. *Infect Immun* 1996, 64:4719-4725.
6. Reid KB: Chemistry and molecular genetics of C1q. *Behring Inst Mitt* 1989:8-19.
7. Reid KB, Day AJ: Ig-binding domains of C1q. *Immunol Today* 1990, 11:387-388.
8. Kaul M, Loos M: Collagen-like complement component C1q is a membrane protein of human monocyte-derived macrophages that mediates endocytosis. *J Immunol* 1995, 155: 5795-5802.
9. Tracey KJ, Cerami A: Tumor necrosis factor: a pleiotropic cytokine and therapeutic target. *Annu Rev Med* 1994, 45:491-503.
10. Nakano Y, Tobe T, Choi-Miura NH, Mazda T, Tomita M: Isolation and characterization of GBP28, a novel gelatin-binding protein purified from human plasma. *J Biochem (Tokyo)* 1996, 120:803-812.
11. Jones EY, Stuart DI, Walker NP: Structure of tumour necrosis factor. *Nature* 1989, 338:225-228.
12. Eck MJ, Sprang SR: The structure of tumor necrosis factor-α at 2.6 Å resolution. Implications for receptor binding. *J Biol Chem* 1989, 264:17595-17605.
13. Weis WI, Quesenberry MS, Taylor ME, Bezouska K, Hendrickson WA, Drickamer K: Molecular mechanisms of complex carbohydrate recognition at the cell surface. *Cold Spring Harb Symp Quant Biol* 1992, 57:281-289.
14. Weis WI, Drickamer K: Trimeric structure of a C-type mannose-binding protein. *Structure* 1994, 2:1227-1240.
15. Brass A, Kadler KE, Thomas JT, Grant ME, Boot-Handford RP: The fibrillar collagens, collagen VIII, collagen X and the C1q complement proteins share a similar domain in their C-terminal non- collagenous regions. *FEBS Lett* 1992, 303:126-128.
16. Duncan AR, Winter G: The binding site for C1q on IgG. *Nature* 1988, 332:738-740.
17. Marques G, Anton LC, Barrio E, Sanchez A, Ruiz S, Gavilanes F, Vivanco F: Arginine residues of the globular regions of human C1q involved in the interaction with immunoglobulin G. *J Biol Chem* 1993, 268:10393-10402.
18. Nedwin GE, Naylor SL, Sakaguchi AY, Smith D, Jarrett-Nedwin J, Pennica D, *et al.*: Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization. *Nucleic Acids Res* 1985, 13:6361-6373.
19. Sellar GC, Blake DJ, Reid KB: Characterization and organization of the genes encoding the A-, B- and C-chains of human complement subcomponent C1q. The complete derived amino acid sequence of human C1q. *Biochem J* 1991, 274:481-490.
20. Renshaw BR, Fanslow WCR, Armitage RJ, Campbell KA, Liggitt D, Wright B, *et al.*: Humoral immune responses in CD40 ligand-deficient mice. *J Exp Med* 1994, 180:1889-1900.
21. Naismith JH, Sprang SR: Tumor necrosis factor receptor superfamily. *J Inflamm* 1995, 47:1-7.
22. Kulics J, Circolo A, Strunk RC, Colten HR: Regulation of synthesis of complement protein C4 in human fibroblasts: cell- and gene-specific effects of cytokines and lipopolysaccharide. *Immunology* 1994, 82:509-515.
23. Kawakami Y, Watanabe Y, Yamaguchi M, Sakaguchi H, Kono I, Ueki A: TNF-alpha stimulates the biosynthesis of complement C3 and factor B by human umbilical cord vein endothelial cells. *Cancer Lett* 1997, 116:21-26.
24. Hotamisligil GS, Spiegelman BM: Tumor necrosis factor α: a key component of the obesity-diabetes link. *Diabetes* 1994, 43:1271-1278.
25. Teoman Uysal K, Wiesbrock SM, Marino MW, Hotamisligil GS: Protection from obesity-induced insulin resistance in mice lacking TNF-α function. *Nature* 1997, 389:610-614.
26. Takamatsu N, Ohba K, Kondo J, Kondo N, Shiba T: Hibernation-associated gene regulation of plasma proteins with a collagen-like domain in mammalian hibernators. *Mol Cell Biol* 1993, 13:1516-1521.
27. Flier JS, Cook KS, Usher P, Spiegelman BM: Severely impaired adipsin expression in genetic and acquired obesity. *Science* 1987, 237:405-408.
28. Depraetere V, Golstein P: Fas and other cell death signaling pathways. *Semin Immunol* 1997, 9:93-107.
29. Karpusas M, Hsu YM, Wang JH, Thompson J, Lederman S, Chess L, Thomas D: 2 Å crystal structure of an extracellular fragment of human CD40 ligand *Structure* 1995, 3:1031-1039.

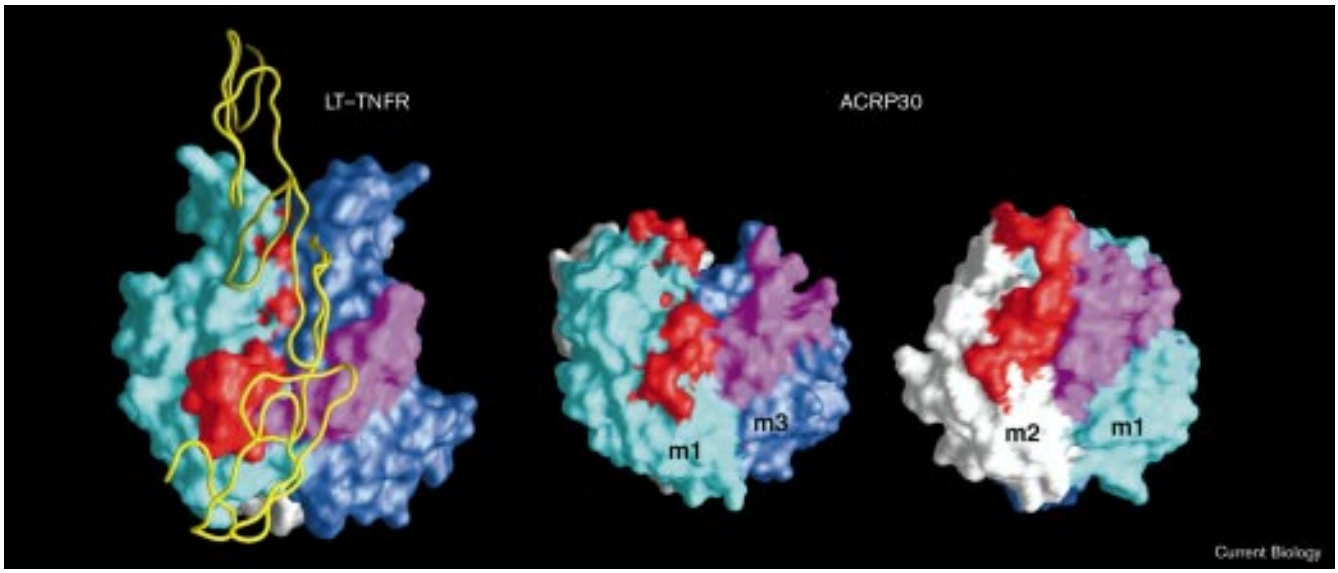
Supplementary material

The crystal structure of a complement-1q family protein suggests an evolutionary link to tumor necrosis factor

Lawrence Shapiro and Philipp E. Scherer

Current Biology 2 March 1998, 8:335–338

Figure S1



Comparison between putative binding surfaces of proteins from the C1q and TNF families. The complex [S4] between lymphotoxin (LT, or TNFβ) and the ectodomain of the 55 kDa TNF receptor (in yellow) is shown on the left. Molecular surfaces of the subunits (m1, m2, and m3) of the TNFβ trimer are colored in blue, cyan, and white (mostly obscured). The receptor binds in a groove formed between two protomers of the trimer. Although three receptors bind symmetrically, for the sake of clarity, only one is shown. The surface formed by the A–A' loop and A' strand of one molecule (m3) is colored magenta, and the surface from the DE loop and E-strand of m1 is colored red. Together, these regions form a major part of the receptor-binding surface of TNFβ. Remarkably, the same molecular regions of C1q have been implicated in IgG binding [17]. Analogous pictures, with the same color scheme, are shown for the C1q-like ACRP30 structure, in the grooves formed between m1 and m3, and m1 and m2. These grooves differ substantially because of the alternate conformation of the E-strand and DE loop adopted by m2 (see Figure 1d), and suggest the likelihood of substantial flexibility in ligand or receptor binding. Some of the loops at the top of m2 and m3 are disordered and therefore not drawn, leading to an appearance in this picture which is artificially shortened.

Materials and methods

Protein expression and crystallization

A fragment corresponding to residues 111–247 of mouse ACRP30 was expressed as a maltose-binding protein fusion in *Escherichia coli* using the pMal-c2 vector (New England Biolabs). Cultures were grown at 30°C, and the fusion protein was isolated by affinity chromatography of the soluble clarified lysate on immobilized amylose. After digestion with factor Xa, the ACRP30 fragment was isolated by ion-exchange chromatography on a mono-Q column, and concentrated to 20 mg/ml. Crystals were grown by the hanging-drop method by mixing equal volumes of protein and a well solution containing 1.8% PEG 4000 and 0.1 M bis-Tris, pH 6.0, at 4°C. These crystals belong to space group $P6_1$ with $a = b = 112.3$ Å, $c = 71.6$ Å, with typical dimensions $60 \times 60 \times 200$ μm. Crystals were cryoprotected by soaking in the well solution supplemented with 15% and then 30% glycerol for 2 min each, and then freezing the crystals, mounted in an ethylon loop, in a cold nitrogen stream at –110°C.

Crystallography

Native and K_2IrCl_6 derivative data were collected on Fuji image plates at beam-line X4A at Brookhaven National Laboratory with the energy tuned to the Selenium K-edge, and the HgOAc derivative was collected on an R-Axis IV detector with a CuKα X-ray source (Table 1). All data was reduced using DENZO and SCALEPACK (HKL Research). Heavy atom positions were found by inspection of Patterson maps and subsequent difference Fourier analysis. The HgOAc and K_2IrCl_6 derivatives each have three sites, with one overlapping between the two. MIR phasing calculations were performed with MLPHARE, and phase extension by solvent flattening and histogram matching (from 4.8 to 2.4 Å over 300 cycles) with DM. MLPHARE and DM are part of the CCP4 suite of programs [S1]. All density calculations and refinement were performed with X-PLOR [S2], molecular superpositions with INSIGHT (Biosym Technologies), and all structure figures were made with the program SETOR [S3].

References

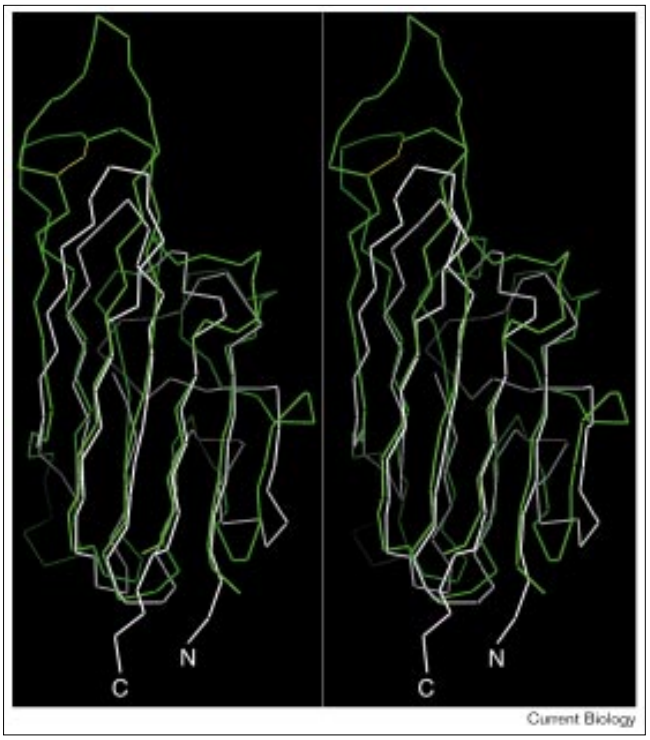
S1. CCP4: The CCP4 program suite: programs for X-ray crystallography. *Acta Crystallogr D* 1994, 50:760-763.

S2. Brünger AT: *X-Plor Version 3.1 Manual*, (Yale University, New Haven, 1993).

S3. Evans SV: SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J Mol Graph* 1993, 11:134-138.

S4. Banner DW, D'Arcy A, Janes W, Gentz R, Schoenfeld HJ, Broger C, *et al.*: Crystal structure of the soluble human 55kd TNF receptor-human TNFβ complex: implications for TNF receptor activation. *Cell* 1993, 73:431-445.

Figure S2



Comparison between ACRP30 and TNFα. Superposition of Cα traces from ACRP30 (white) and TNFα (green, from PDB entry 1TNFA). Fifty pairs of corresponding Cα atoms have an r.m.s. displacement of 0.78 Å. This is representative of the structural similarities between ACRP30 and the members of the TNF family of known structure