Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of human desArg-C5a

William J. Cook,[a]* Nicholas Galakatos,[b]‡ William C. Boyar,[b] Richard L. Walter[c] and Steven E. Ealick[d]

[a]University of Alabama at Birmingham, Birmingham, AL 35294, USA, [b]Ciba–Geigy Corporation, Summitt, New Jersey, USA, [c]Shamrock Structures LLC, Woodridge, IL 60517, USA, and [d]Cornell University, Ithaca, NY 14853, USA

‡ Current address: Clarus Ventures, Cambridge, MA 02142, USA.

Correspondence e-mail: wjcook@uab.edu

Received 18 September 2009
Accepted 17 November 2009

PDB References: human desArg-C5a, 3hqa; 3hqb.

The anaphylatoxin C5a is derived from the complement component C5 during activation of the complement cascade. It is an important component in the pathogenesis of a number of inflammatory diseases. NMR structures of human and porcine C5a have been reported; these revealed a four-helix bundle stabilized by three disulfide bonds. The crystal structure of human desArg-C5a has now been determined in two crystal forms. Surprisingly, the protein crystallizes as a dimer and each monomer in the dimer has a three-helix core instead of the four-helix bundle noted in the NMR structure determinations. Furthermore, the N-terminal helices of the two monomers occupy different positions relative to the three-helix core and are completely different from the NMR structures. The physiological significance of these structural differences is unknown.

1. Introduction

In the classical activation of the complement protein pathway, the formation of immune complexes triggers a cascade of proteolytic cleavages of complement proteins. The activation of complement component C5 by C5 convertase initiates the assembly of the late complement components, C5b–C9, into the membrane-attack complex. C5a is an anaphylatoxin that is derived from the cleavage of C5. This 74-amino-acid glycoprotein is a potent chemotactic factor for all cells of the myeloid lineage, including neutrophils, eosinophils, basophils and mast cells, causing numerous cellular responses such as chemotaxis, aggregation and adhesion (Ricklin & Lambris, 2007). Most tissue macrophage types, including alveolar macrophages (McCarthy & Henson, 1979), liver Kuppfer cells (Laskin & Pilaro, 1986) and microglia (Yao *et al.*, 1990), also respond to C5a. C5a has been implicated as a causative or aggravating agent in a variety of inflammatory and allergic diseases, such as rheumatoid arthritis (Linton & Morgan, 1999), inflammatory bowel diseases (Woodruff *et al.*, 2003), adult respiratory distress syndrome (Robbins *et al.*, 1987), asthma and allergy (Hawlisch *et al.*, 2004; Gerard & Gerard, 2002; Baelder *et al.*, 2005), ischemia/reperfusion injury (Arumugam *et al.*, 2004) and glomerulonephritis (Welch, 2002; Kondo *et al.*, 2001). Because of its role in various pathologic conditions, C5a is of considerable pharmaceutical interest (Ricklin & Lambris, 2007).

The three-dimensional structures of human and porcine C5a have been determined by NMR methods (Zuiderweg *et al.*, 1989; Zhang *et al.*, 1997; Williamson & Madison, 1990). Although the crystal structure of human C5a has not been

determined, the structure of human complement factor C5 has (Fredslund *et al.*, 2008). These studies revealed a compact structure composed of an antiparallel bundle of four α-helices stabilized by three disulfide linkages. Only one of the three reported NMR structures included all 74 residues (Zhang *et al.*, 1997); in the other two structures the C-terminal residues could not be positioned. Similarly, the C5a portion of the C5 molecule is missing residues 67–71 at the C-terminus. All of these structures are similar to the crystal structure of human C3a, which has been determined at medium resolution (Huber *et al.*, 1980). The C-terminus was well defined in the structure of C3a, but the 14 N-terminal residues were not visible in electron-density maps.

We have now determined the crystal structure of human C5a in two different crystal forms. Each crystal form contained an asymmetrical dimer in the asymmetric unit. However, in contrast to the NMR structures, which showed a tightly packed four-helix bundle, the crystal structure shows a three-helix central core in each monomer connected by short loops located at the surface of the dimer. Furthermore, the N-terminal helices in each half of the dimer occupy completely different positions relative to the central three-helix core.

2. Materials and methods

2.1. Crystallization and data collection

The cloning and expression of recombinant human C5a and desArg-C5a have been described in detail (Toth *et al.*, 1994). The N-terminal threonine residue of human C5a was replaced by methionine to allow the proper initiation of translation in *Escherichia coli*. Purified protein was dissolved in water to a concentration of 20 mg ml^{-1}. Despite extensive screening of crystallization conditions, crystals of C5a were never obtained. However, small crystals of desArg-C5a were grown at 296 K by the hanging-drop method in 2.3–2.5 *M* sodium chloride with 50 m*M* acetate buffer pH 4.8. Two different crystal morphologies were apparent in the same drops. Macro-seeding of each crystal form was required to produce single crystals that were suitable for data collection. This technique produced crystals with dimensions of up to 0.3 mm on an edge. One crystal form was tetragonal ($P4_12_12$) and the other was trigonal ($P3_221$). Both crystal forms contained two monomers in the asymmetric unit.

Intensity data for both crystal forms were collected at room temperature with a Nicolet X-100A area detector at 295 K using Cu $K\alpha$ radiation from a Rigaku RU-300 rotating-anode generator. In order to obtain a complete data set with multiple measurements of all reflections, multiple data sets were collected for the tetragonal crystal form. The crystal-to-detector distance was 12 cm and the detector 2θ value was 15°. Oscillation frames covered 0.25° and were measured for 5 min. A total of 19 450 reflections were processed and merged into 5330 unique reflections (97.5% complete). The $R_{\rm merge}$ value (based on *I*) for the data to 2.58 Å resolution was 0.153. The trigonal crystals were smaller, more difficult to grow and did not diffract nearly as well as the tetragonal crystals. Therefore, only one data set was collected from one crystal. The crystal-to-detector distance was 16 cm and the detector 2θ value was 15°. Oscillation frames covered 0.25° and were measured for 5 min. A total of 8519 reflections were processed and merged into 2707 unique reflections (98% complete). The $R_{\rm merge}$ (based on *I*) for the data to 3.3 Å was 0.098. Indexing and integration of intensity data were carried out using the *XENGEN* processing programs (Howard *et al.*, 1987). Table 1 gives statistics of the data processing.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P4_12_12$	$P3_221$
Crystal data		
Unit-cell parameters (Å)	a = 50.61, c = 117.49	a = 54.72, c = 96.48
$V_{\rm M}$ (Å^3 Da^{-1})	2.26	2.61
Solvent content (%)	46	53
Data collection		
Maximum resolution (Å)	2.58 (2.58–2.76)	3.30 (3.30–3.51)
Redundancy	3.6 (2.8)	3.1 (2.8)
Completeness (%)	97.5 (97.1)	98.1 (93.8)
$R_{\rm merge}$ (%)	15.3 (49.3)	9.83 (33.28)
Overall $I/\sigma(I)$	7.6	1.3
Refinement		
Resolution range (Å)	46.48–2.59 (2.74–2.59)	23.80–3.30 (3.40–3.30)
Reflections	5078	2287
R value	0.213	0.210
Free R value	0.253	0.307
No. of protein atoms	995	959
Coordinate error, maximum-likelihood based (Å)	0.24	0.51
Deviations from ideality		
Bond lengths (Å)	0.009	0.012
Bond angles (°)	1.1	1.5
Dihedral angles (°)	16.0	19.3
Average B factors (Å^2)		
Overall	54.4	63.6
Monomer *A*	52.9	65.3
Monomer *B*	55.9	61.8
MolProbity analysis		
Clash score	9.49 [98th percentile; N = 225, 2.58 ± 0.25 Å]	45.53 [59th percentile; N = 37, >3.00 Å]
MolProbity score	1.89 [98th percentile; N = 6174, 2.58 ± 0.25 Å]	3.46 [62nd percentile; N = 892, 3.3 ± 0.25 Å]
Ramachandran favored (%)	96.83	73.77
Ramachandran outliers (%)	0.00	4.96
Rotomer outliers (%)	1.83	4.81

2.2. Structure determination and refinement

The crystal structure of the tetragonal form was solved using the molecular-replacement program *Phaser* (McCoy *et al.*, 2007). Various search models were tested, including the crystal structure of C3a (Huber *et al.*, 1980), the NMR structure of human C5a (PDB code 1kjs; Zhang *et al.*, 1997) and the C5a portion of the crystal structure of human C5 (PDB code 3cu7; Fredslund *et al.*, 2008). Data from 50 to 2.6 Å resolution were used for each of the two enantiomorphic space groups. No satisfactory solutions were identified using the C3a crystal structure or the C5a NMR structure, even after truncating residues from the N- and C-termini. The initial model used from the C5 structure contained residues Ala681–Ala742

(corresponding to residues 4–65 of C5a). Although this model yielded a potential solution with a log-likelihood score (LLG) of 116, there were also nine clashes between the C^{α} atoms of the two monomers in the asymmetric unit. Visual inspection of the model showed overlapping of the N-terminal helix of one monomer with the C-terminal helix of the other. Therefore, various truncated models were tried. The best log-likelihood score was achieved with a model that only included residues 18–63 (LLG = 167). This solution confirmed the space group to be $P4_12_12$.

Electron-density maps calculated with this partial model were of sufficient quality to allow building of the remaining residues. In particular, the helical density corresponding to most of the residues in the N-terminal helices was apparent (Fig. 1). Calculations were performed using the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). The quality of the maps was significantly improved by using the 'prime-and-switch' phasing technique in *RESOLVE* (Terwilliger, 2000). The graphics program *Coot* was used for model building (Emsley & Cowtan, 2004). Although gaps in the crystal packing clearly indicated where the C-terminal residues were expected to be, residues 67–73 could not be modelled owing to lack of interpretable electron density in this area.

Refinement of the structure was initially performed by simulated annealing using *CNS* (Brünger *et al.*, 1998) with the stereochemical parameter files defined by Engh & Huber (1991). However, the final refinement was performed with *PHENIX* (Adams *et al.*, 2002). No σ cutoff was applied to the data. 5% of the data were randomly selected and removed prior to refinement for analysis of the free *R* factor (Brünger, 1992). The two monomers were restrained by noncrystallographic symmetry where appropriate, although the N-terminal helix and several individual residues could not be included because of local packing differences. Individual *B* factors were included in the refinement. Owing to the relatively low resolution of the structure, no water molecules were included in the model.

For solution of the trigonal crystal structure, a portion of the refined tetragonal structure (residues 18–65) was used as the search model. This partial model was used to avoid bias in case the N-terminal helices of the molecules in the trigonal structure differed from the tetragonal structure. *Phaser* yielded only one solution, with a log-likelihood score of 214; this solution confirmed the space group as $P3_221$. The two monomers in the asymmetric unit had the same relative orientation to each other as in the tetragonal structure. Although the resolution of the data for this crystal form was much lower (3.3 Å), the electron-density map clearly showed that the N-terminal helices of each monomer had the same conformation as in the tetragonal structure. The last seven residues at the C-terminus were not visible in the electron-density maps. The refinement procedure for the trigonal crystal form was similar to the procedure employed for the tetragonal crystal form, except that 10% of the data were used to calculate the free *R* factor. The stereochemical quality of the final models was verified using the program *MolProbity* (Davis *et al.*, 2007). Since the trigonal structure was refined at much lower resolution and appears to be virtually identical to the tetragonal structure, it will not be discussed further. Table 1 contains a summary of the refinement statistics.

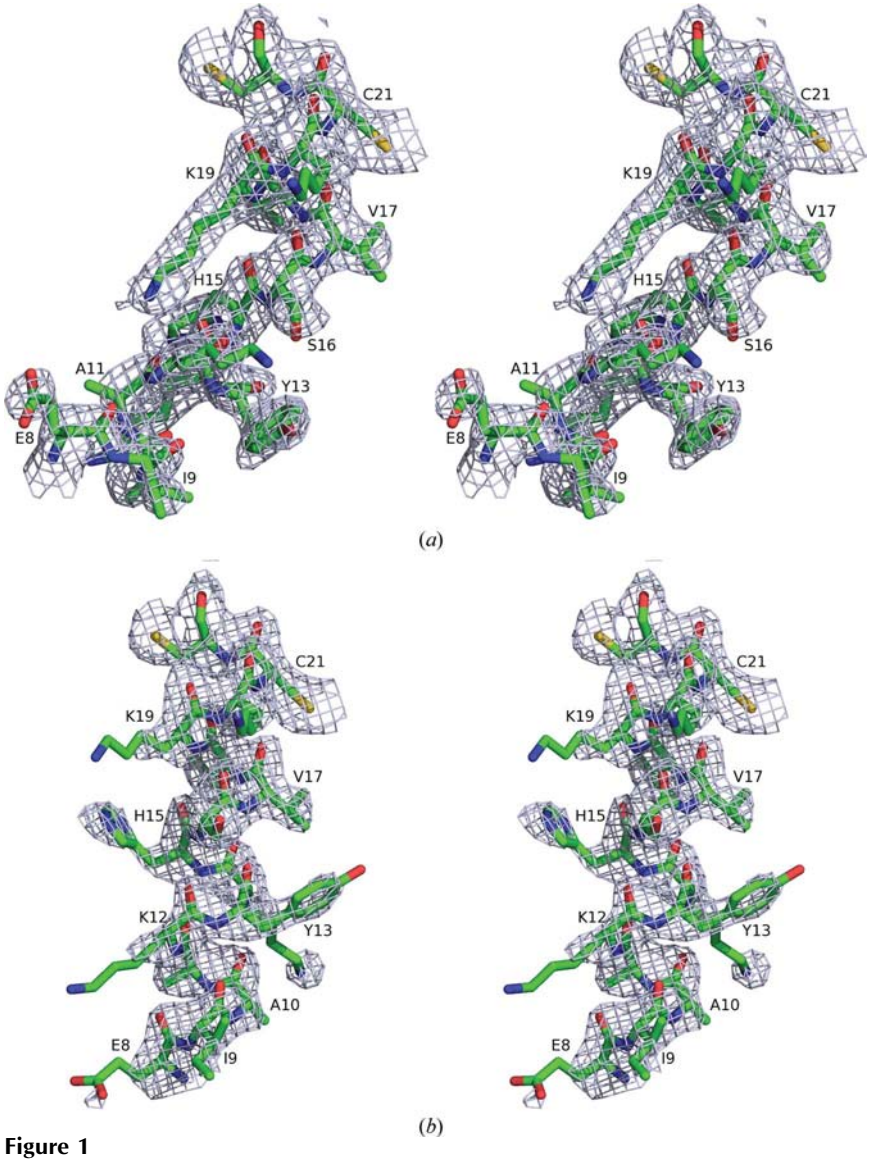

Figure 1
Differences in the N-terminal helices in the two monomers. Residues 8–22 are shown for each monomer. The $2F_o - F_c$ maps are contoured at 1.0σ. (*a*) Monomer *A* (residues 1–17 omitted from the structure-factor calculation). (*b*) Monomer *B* (residues 3–17 omitted from the structure-factor calculation). All figures were created with *PyMOL* (DeLano, 2002).

3. Results

3.1. Structure of desArg-C5a

The protein crystallizes as an asymmetrical dimer with approximate dimensions of 29 × 49 × 51 Å (Fig. 2). The non-crystallographic symmetry is not a simple twofold axis. The polar rotation angle κ was 95.4° based on a superposition of residues 15–66 in each monomer. The core of each monomer is an antiparallel bundle of three helices; these two three-helix bundles form a six-helix bundle in the dimer. This is distinctly different from the NMR structures of C5a and the structure of the C5a portion of intact human C5. The three-helix bundle in each monomer is stabilized by three disulfide bonds (Cys21–Cys47, Cys22–Cys54 and Cys34–Cys55). Residues in these three helices from each monomer form the hydrophobic core of the dimer. The dimer is further stabilized by five intermolecular contacts between the two monomers. Two of these contacts are hydrogen bonds: Lys19 NZ to Ala39 O and Leu41 O. The other three are salt bridges: Lys20 NZ to Glu35 OE2, Glu53 OE1 to Lys49 NZ and Arg62 NH1 to Gln60 OE1.

Although monomer *A* contains four distinct α-helices (residues 3–13, 16–26, 34–40 and 45–63), the N-terminal helix does not form part of the helical bundle. Instead, it projects away from the central core of the structure, giving the dimer a distinctly asymmetrical shape. The two residues between the first two helices (Lys14 and His15) are the two central residues in a type II tight turn. In contrast, monomer *B* only contains three helices (residues 4–26, 34–39 and 45–64) because the short loop between helices I and II in monomer *A* assumes a helical conformation in monomer *B*, thus creating one long helix. As in monomer *A*, the N-terminal portion of this long helix extends away from the core of the structure, creating a very asymmetrical molecule with approximate dimensions of 11 × 18 × 51 Å. In both cases the N-terminal helices are stabilized by contacts with helices from symmetry-related monomers. Other than the 14 N-terminal residues, the overall structures of the two monomers in the asymmetric unit are quite similar. The average root-mean-square value obtained from the superposition of the C^{α} atoms of residues 15–66 is 0.71 Å. If the last four residues at the C-terminus are omitted from the calculation, the value is only 0.55 Å.

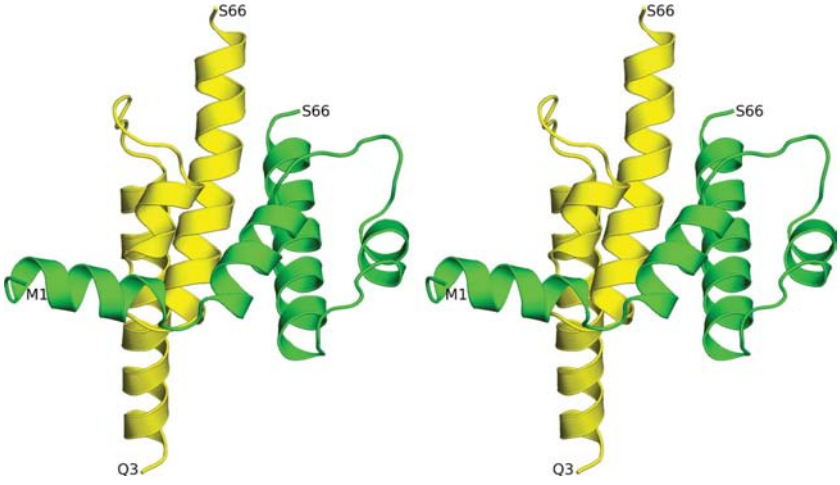

Figure 2
Stereoview of the structure of the C5a dimer. Monomer *A* is shown in green and monomer *B* in yellow. The N- and C-terminal residues of each monomer are labelled.

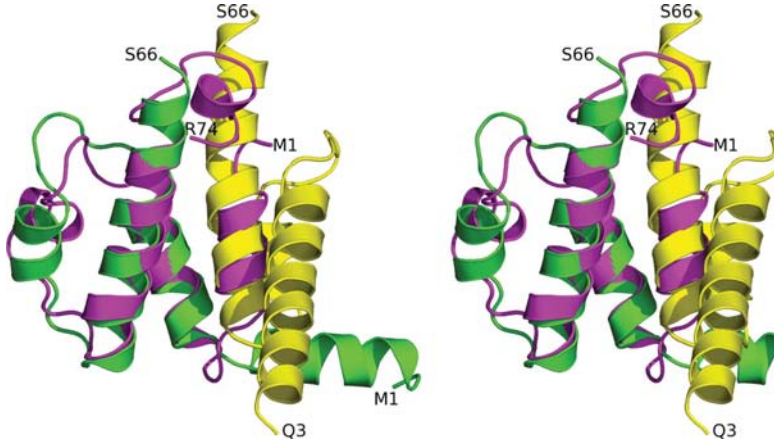

Figure 3
Stereoview of the NMR structure (magenta) superimposed onto monomer *A* (green) of the C5a X-ray structure. The N- and C-terminal residues of the NMR structure and each monomer are labelled. Note that the N-terminal helix of the NMR structure overlaps with the C-terminal helix of monomer *B* (yellow).

3.2. Comparison to C5a NMR structures

Two NMR determinations of the human C5a structure have been reported (Zuiderweg *et al.*, 1989; Zhang *et al.*, 1997), but coordinates have only been deposited for one of them (Zhang *et al.*, 1997). Both structures are composed of four-helix bundles, but there is a major difference between the two NMR structures at the C-terminus. In the structure reported by Zuiderweg *et al.* (1989) only residues 1–63 of the sequence had defined structure; the 11 residues at the C-terminus exhibited the characteristics of random coil. In contrast, the structure reported by Zhang *et al.* (1997) included all 74 residues. In their structure the six C-terminal residues adopt an α-helical conformation, which is connected to the core by a short loop and is situated between the N-terminus and the base of the fourth helix.

The core of the crystal structure of desArg-C5a is similar to the overall NMR structure; in particular, the second and fourth helices are almost exactly superimposable. The major differences occur in the N-terminus, the C-terminus and the loop–helix–loop segment containing residues 27–40 (Figs. 3 and 4). Relative to the other three helices that form the core of the NMR structure, the first 17 N-terminal

residues in each monomer of the dimer occupy completely different positions. In this regard, it is interesting that in the dimer the C-terminal helix of monomer *B* occupies the same relative position as the N-terminal helix in the NMR structure of C5a. The loop–helix–loop region (residues 27–40) varies widely in the NMR structures and none are really similar to the crystal structure. In particular, the side chain of Arg37 assumes a completely different position and helps to stabilize the short loop preceding the helix that begins with Cys34. The helix in this region is longer in both monomers of the dimer than in the NMR structure. Finally, the C-terminal eight residues are not visible in the crystal structure, but the crystal packing rules out an orientation of the C-terminus similar to the NMR structure. It seems likely that the last eight residues in crystallized desArg-C5a adopt an extended random-coil configuration.

It is not clear why C5a assumes a different conformation in the crystal structure compared with the solution structure. C5a was crystallized from a high ionic strength solution (2.3–2.5 *M* NaCl), but this is unlikely to be a reason for the different conformation of the N-terminal helix in the C5a crystal structure. In the C3a structure, which was also crystallized from a high ionic strength solution (3 *M* phosphate), the first 12 residues at the N-terminus are missing, but residues 14–16 of C3a form a turn that closely follows the turn at residues 13–15 in the NMR structure of C5a when the two structures are superimposed. Therefore, the N-terminal helix in C3a probably forms part of the four-helix bundle. Another possibility is stabilization of the dimer by the formation of hydrogen bonds and salt bridges between the monomers. In the solution structure of C5a there are no intramolecular contacts between the N-terminal helix and the rest of the molecule, whereas there are five contacts between the two monomers, as noted above.

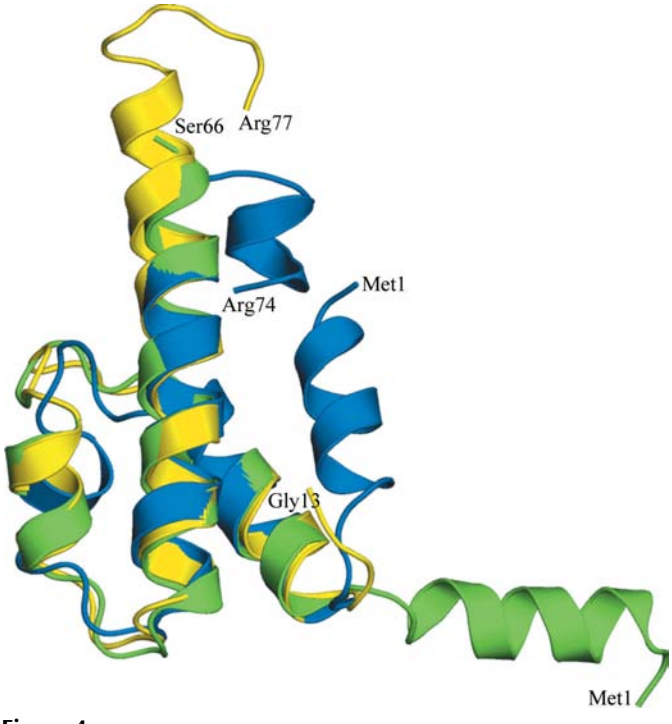

Figure 4
Superposition of the C3a X-ray structure (yellow) and the C5a NMR structure (blue) onto monomer *A* of the C5a X-ray structure (green). The N- and C-terminal residues of each structure are labeled.

3.3. Comparison to the C5a portion of human C5

The crystal structure of human complement factor C5 has been determined (Fredslund *et al.*, 2008). The conformation of desArg-C5a in the crystal structure is similar to the conformation of C5a in the intact C5 molecule, except for the obvious difference at the N-terminus. The possible implications of this change in the orientation of the N-terminal helix are discussed in more detail below. There are also small differences in residues 43–47, which form the end of a loop and the beginning of the fourth helix. In the intact C5 structure the C5a portion assumes a four-helix bundle structure similar to the NMR structure of C5a. The average root-mean-square value obtained from the superposition of C^{α} atoms of residues 16–64 (monomer *A*) is 0.99 Å.

3.4. Comparison to the C3a structure

C3a has 35% amino-acid homology to C5a and its disulfide linkages are located in homologous positions. The crystal structure of human C3a has been reported (Huber *et al.*, 1980), but the published NMR study of human C3a only reported the secondary structure (Nettesheim *et al.*, 1988). Coordinates for the C3a crystal structure were kindly provided to us by J. Deisenhofer. The crystal structure of C3a differs from the crystal structure of desArg-C5a in three important respects (Fig. 4). Two of these differences occur in the N- and C-termini. The first 12 residues at the N-terminus of C3a are not visible in the crystal structure, but the entire C-terminus is well ordered. In fact, the C-terminal portion forms a loop back to the fourth helix in both the NMR structure of C5a and the crystal structure of C3a. The other significant difference occurs in the loop containing residues 28–31. Human C3a has a one-residue insertion at this point relative to human C5a. Other smaller differences include shorter helices corresponding to helix 2 and helix 3 of C5a. The average root-mean-square value obtained from the superposition of the C^{α} atoms of residues 16–29 and 32–66 of C5a (monomer *A*) onto C3a is 0.98 Å.

It is interesting that in the C3a crystal structure two monomers related by the crystallographic dyad form a dimer. The C-terminal helices are arranged antiparallel, with Asn58, Tyr59, Glu62, Leu63, Gln66, Ala68 and Arg69 participating in the contact. The interface between the two monomers in the C3a dimer is thus completely different from the interface of the C5a dimer, where there is no contact between the C-terminal helices.

In contrast to the crystal structure of C3a, the NMR structure of C3a shows a loss of helical structure at residues 67–70 and then no defined structure for the last six residues at the C-terminus, similar to the crystal structure of C5a. Huber *et al.* (1980) suggested that the additional helical structure at the C-terminus in the crystal structure may be explained by

the fact that in the crystal two C3a monomers interact with each other with their C-terminal helices in antiparallel fashion; this association may stabilize the conformation of the last few residues. Unlike the crystal structure of C3a, the NMR structure shows helical structure for residues 8–15, although the seven N-terminal residues have a random conformation (Nettesheim *et al.*, 1988). Even though the tertiary structure was not determined in the NMR study, Nettesheim *et al.* (1988) interpreted the NMR results as indicating that the N-terminal helix does not move independently of the core of the molecule in solution. Thus, it probably forms part of a four-helix bundle like the NMR structure of C5a.

4. Discussion

4.1. Dimer formation and C5a–C5a receptor interactions

The activity of C5a is mediated by a G-protein-coupled receptor (C5aR; Gerard & Gerard, 1991). The receptor has seven transmembrane domains linked by intracellular and extracellular loops, along with an extracellular N-terminus and an intracellular C-terminus. A number of studies have examined the binding of C5a to C5aR (Toth *et al.*, 1994; Mollison *et al.*, 1989; Bubeck *et al.*, 1994; Hagemann *et al.*, 2008). The interaction has been described as a two-site model, in which there is a primary high-affinity contact between basic residues in the core of C5a and acidic residues in the N-terminus of the receptor. The C-terminal tail of C5a then enters a binding pocket formed by hydrophobic residues in the transmembrane domains and charged residues at the base of the extracellular loops to form the second-site interaction.

Specific interactions between C5a and its receptor have been assigned based on data from site-directed mutagenesis of C5a and C5aR. As expected, modifications of the C-terminal residues of C5a, such as His67, Lys68, Leu72 and especially Arg74, lead to a significant decrease in binding with C5aR (Mollison *et al.*, 1989). However, another possibly important residue was noted in the core region (Arg40). The importance of Lys19 and Lys20 was shown in subsequent studies by Bubeck *et al.* (1994), but the possible involvement of Arg40 was not. Toth *et al.* (1994) performed a systematic mutational analysis of C5a and measured the effects on the potency of receptor binding. In addition to Lys19, they suggested a number of other residues that are likely to be important for binding, including His15, Arg37, Leu43, Arg46, Lys49 and Glu53. Most recently, Hagemann *et al.* (2008) used disulfide trapping by random mutagenesis to identify six unique sets of intermolecular interactions for the C5a–C5aR complex. The C5a residues important for binding included His15, Asp24, Cys27, Arg40, Arg46 and Ser66. Although there are studies that suggest that the C5a receptor forms oligomers *in vivo* and *in vitro*, there is no suggestion that the formation of oligomers is induced by binding to a C5a dimer (Klco *et al.*, 2003; Rabiet *et al.*, 2008). On the contrary, it appears that the stimulation and phosphorylation of one C5aR monomer is sufficient to cause dimer formation (Rabiet *et al.*, 2008).

Although the studies do not always agree, it seems likely that at least two regions of C5a (in addition to the C-terminus) are important for binding. One region includes the residues His15, Lys19 and Arg46, whose side chains are relatively close together on the surface of the molecule (Fig. 5). The other region is the surface that contains Asp24, Arg37 and Arg40. Arg37 is conserved in all reported mammalian C5a sequences, but it appears to be unlikely that it has any role in binding to the receptor. Instead, the loop between helices 2 and 3 is partially stabilized by hydrogen bonds between the guanidinium group of Arg37 and the carbonyl O atoms of Asp24 and Cys27 (Fig. 6).

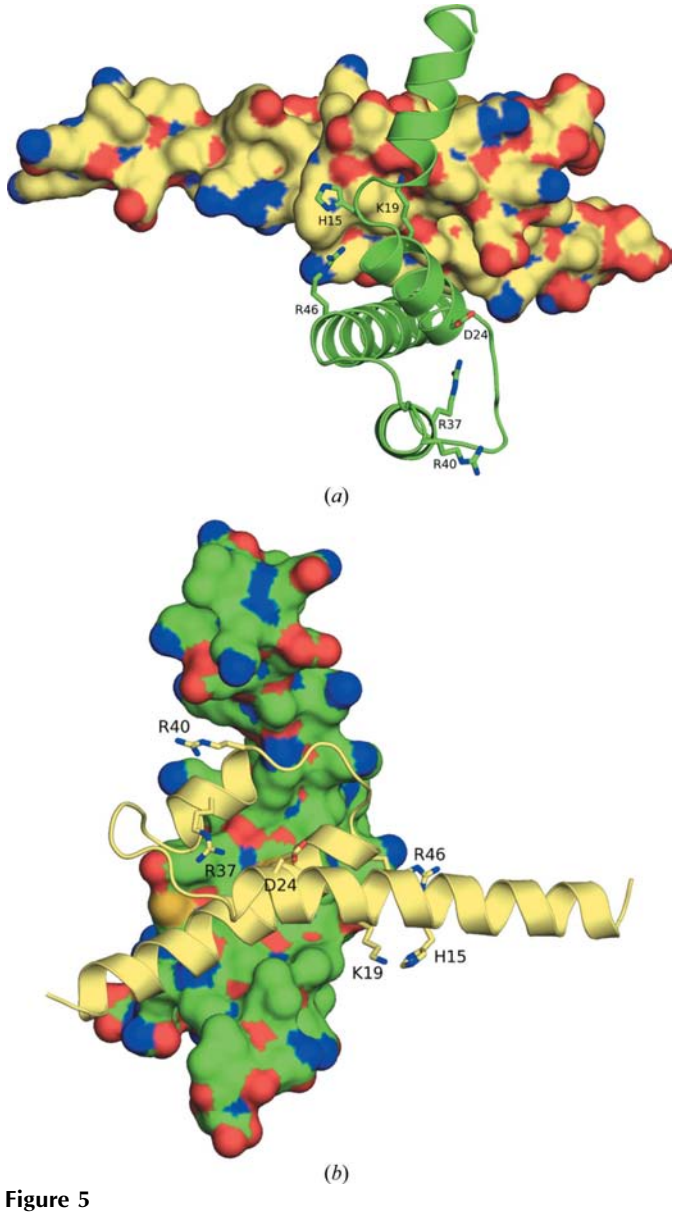

Figure 5
Residues of C5a that may be important for binding to the C5a receptor. The residues proposed to be important for interaction with C5aR (His15, Lys19, Asp24, Arg37, Arg40 and Arg46) are shown as stick models. (*a*) Space-filling model of C5a monomer *B* with a ribbon drawing of monomer *A* (green). (*b*) Space-filling model of C5a monomer *A* with a ribbon drawing of monomer *B* (pale yellow).

In the formation of the dimer, access to some of these key residues is restricted. In monomer *A* the side chain of Lys19 is completely obstructed, while the side chains of His15 and Arg46 are partially obstructed. In monomer *B* the side chain of Arg46 is also partially obstructed, as well as the side chain of Arg40. While it is tempting to speculate that the activity of C5a could be modulated *via* dimer formation, there is no experimental evidence for C5a dimer formation in the literature.

4.2. Implications of the C5a structure for the C5 structure

The structure of C5a demonstrates the extremely flexible nature of the N-terminal helix, which has possible ramifications for the conformation of the C5 molecule. The C5 precursor is processed so that the N-terminal signal polypeptide (residues 1–18) and a short peptide of four basic residues (Arg674-Pro675-Arg676-Arg677) are removed. The two chains, β (residues 19–673) and α (residues 678–1676), are linked by a disulfide bond. C5 convertase activates C5 by cleaving the first 74 residues of the α chain, releasing C5a and generating C5b (β chain + modified α chain). In the structure of C5 (Fredslund *et al.*, 2008), the distance between the last residue of the β chain, Leu673, and the first residue in the anaphylatoxin (C5a) domain, Leu679 (Leu2 in C5a), is approximately 58 Å, which must be spanned by five residues in C5. Obviously, in C5 either the linker region or the N-terminal residues in C5a must have a different conformation. A similar situation was noted in the structure of C3, where the distance between the last residue in the linker region, Gln662, and the first residue in the anaphylatoxin domain, Ser670, is also about 58 Å (Janssen *et al.*, 2005; Fredslund *et al.*, 2006). For the structure of C3, Fredslund *et al.* (2006) suggested two possibilities: (i) the linker region has a different conformation and is positioned closer to the anaphylatoxin domain or (ii) the global conformation of ProC3 is different from that of C3. While those two possibilities cannot be ruled out, the crystal structure of C5a immediately suggests a simpler alternative. When monomer *A* of the C5a dimer is superimposed on the C5a portion of C5, the N-terminal helix points directly at the end of the linker region and extends to within about 18 Å of Glu671 (Fig. 7). This distance would easily accommodate a seven-residue peptide in an extended conformation. In this scenario, no major change in conformation for the linker region or for the global conformation of proC5 would be required. It seems very likely that the N-terminal helix of C3a could undergo a similar movement.

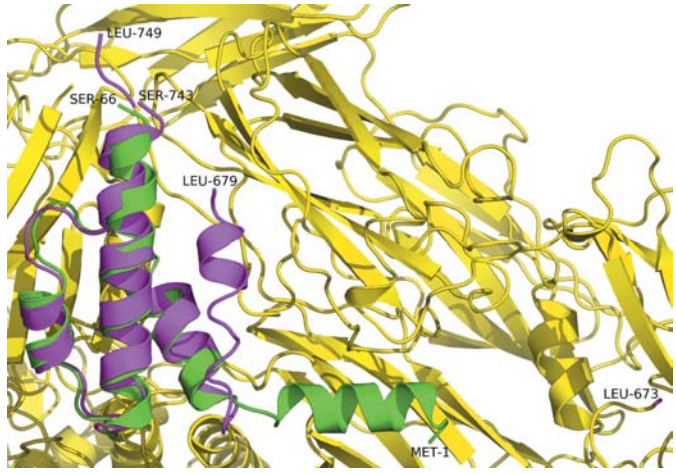

Figure 7
Possible conformation of the anaphylatoxin domain in C5 prior to activation by C5 convertase. Monomer *A* of the C5a dimer (green) is superimposed on the anaphylatoxin domain (magenta) of C5 (yellow). The N- and C-terminal residues of C5a are labelled. Residues 674–678 and 744–748 are missing in the C5 structure.

In summary, the crystal structure of C5a confirms most of the structural features of the NMR-derived structure, but there are several important differences. The most important are the formation of a dimer and the different conformations of the N-terminal helix. The formation of the same dimer in two different space groups with different crystal-packing arrangements argues against an artefact arising from crystal packing. However, the physiological significance of the dimer remains uncertain.

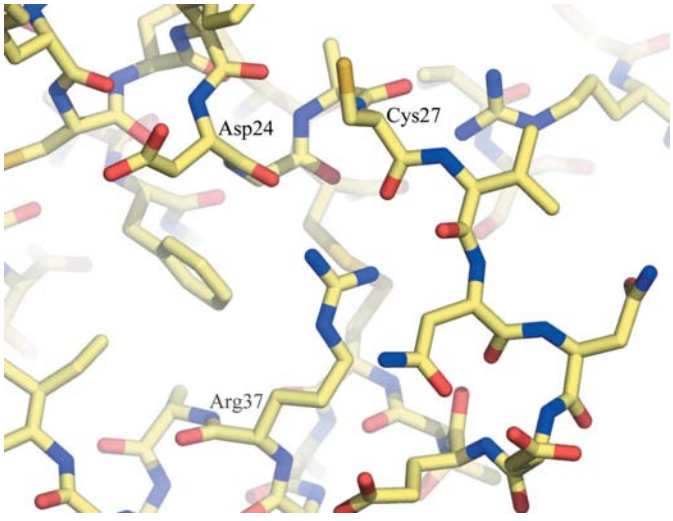

Figure 6
Detailed view of the interactions of Arg37 with Asp24 and Cys27.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Arumugam, T. V., Shiels, I. A., Woodruff, T. M., Granger, D. N. & Taylor, S. M. (2004). *Shock*, **21**, 401–409.
Baelder, R., Fuchs, B., Bautsch, W., Zwirner, J., Köhl, J., Hoymann, H. G., Glaab, T., Erpenbeck, V., Krug, N. & Braun, A. (2005). *J. Immunol.* **174**, 783–789.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Bubeck, P., Grötzinger, J., Winkler, M., Köhl, J., Wollmer, A., Klos, A. & Bautsch, W. (1994). *Eur. J. Biochem.* **219**, 897–904.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B. III, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). *Nucleic Acids Res.* **35**, W375–W383.
DeLano, W. L. (2002). *The PyMOL Molecular Viewer.* http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Fredslund, F., Jenner, L., Husted, L. B., Nyborg, J., Andersen, G. R. & Sottrup-Jensen, L. (2006). *J. Mol. Biol.* **361**, 115–127.
Fredslund, F., Laursen, N. S., Roversi, P., Jenner, L., Oliveira, C. L. P., Pedersen, J. S., Nunn, M. A., Lea, S. M., Discipio, R., Sottrup-Jensen, L. & Andersen, G. R. (2008). *Nature Immunol.* **9**, 753–760.
Gerard, N. P. & Gerard, C. (1991). *Nature (London)*, **349**, 614–617.
Gerard, N. P. & Gerard, C. (2002). *Curr. Opin. Immunol.* **14**, 705–708.
Hagemann, I. S., Miller, D. L., Klco, J. M., Nikiforovich, G. V. & Baranski, T. J. (2008). *J. Biol. Chem.* **283**, 7763–7775.
Hawlisch, H., Wills-Karp, M., Karp, C. L. & Kohl, J. (2004). *Mol. Immunol.* **41**, 123–131.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.
Huber, R., Scholze, H., Paques, E. P. & Deisenhofer, J. (1980). *Hoppe-Seyler's Z. Physiol. Chem.* **361**, 1389–1399.
Janssen, B. J. C., Huizinga, E. G., Raaijmakers, H. C. A., Roos, A., Daha, M. R., Nilsson-Ekdahl, K., Nilsson, B. & Gros, P. (2005). *Nature (London)*, **437**, 505–511.
Klco, J. M., Lassere, T. B. & Baranski, T. J. (2003). *J. Biol. Chem.* **278**, 35345–35353.
Kondo, C., Mizuno, M., Nishikawa, K., Yuzawa, Y., Hotta, N. & Matsuo, S. (2001). *Clin. Exp. Immunol.* **124**, 323–329.
Laskin, D. L. & Pilaro, A. M. (1986). *Toxicol. Appl. Pharmacol.* **86**, 204–215.
Linton, S. M. & Morgan, B. P. (1999). *Mol. Immunol.* **36**, 905–914.
McCarthy, K. & Henson, P. M. (1979). *J. Immunol.* **123**, 2511–2517.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mollison, K. W., Mandecki, W., Zuiderweg, E. R. P., Fayer, L., Fey, T. A., Krause, R. A., Conway, R. G., Miller, L., Edalji, R. P., Shallcross, M. A., Lane, B., Fox, J. L., Greer, J. & Carter, G. W. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 292–296.
Nettesheim, D. G., Edalji, R. P., Mollison, K. W., Greer, J. & Zuiderweg, E. R. P. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 5036–5040.
Rabiet, M.-J., Huet, E. & Boulay, F. (2008). *J. Biol. Chem.* **283**, 31038–31046.
Ricklin, D. & Lambris, J. D. (2007). *Nature Biotechnol.* **25**, 1265–1275.
Robbins, R. A., Russ, W. D., Rasmussen, J. K. & Clayton, M. M. (1987). *Am. Rev. Respir. Dis.* **135**, 651–658.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Toth, M. J., Huwyler, L., Boyar, W. C., Braunwalder, A. F., Yarwood, D., Hadala, J., Haston, W. O., Sills, M. A., Seligmann, B. & Galakatos, N. (1994). *Protein Sci.* **3**, 1159–1168.
Welch, T. R. (2002). *Nature Genet.* **31**, 333–334.
Williamson, M. P. & Madison, V. S. (1990). *Biochemistry*, **29**, 2895–2905.
Woodruff, T. M., Arumugam, T. V., Shiels, I. A., Reid, R. C., Fairlie, D. P. & Taylor, S. M. (2003). *J. Immunol.* **171**, 5514–5520.
Yao, J., Harvath, L., Gilbert, D. L. & Colton, C. A. (1990). *J. Neurosci. Res.* **27**, 36–42.
Zhang, X., Boyar, W., Toth, M. J., Wennogle, L. & Gonnella, N. C. (1997). *Proteins*, **28**, 261–267.
Zuiderweg, E. R., Nettesheim, D. G., Mollison, K. W. & Carter, G. W. (1989). *Biochemistry*, **28**, 172–185.