# Efficient chemical synthesis of homogenous human complement protein C3a†

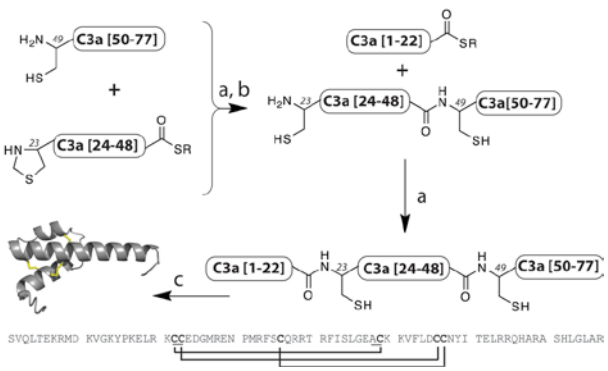
**Artin Ghassemian[a,b,‡], Ching-I Anderson Wang[a,‡], Mei-Kwan Yau[a], Robert C. Reid[a], Richard J. Lewis[a], David P. Fairlie[a], Paul F. Alewood[a], and Thomas Durek[a,*]**

***Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX***
**DOI: 10.1039/b000000x**

**We report the total chemical synthesis of human C3a by one-pot native chemical ligation of three unprotected peptide segments followed by efficient *in-vitro* folding, which yielded the target molecule in high yield and excellent purity. The synthetic material was fully active and facilitated determination of the C3a crystal structure at 2.1Å resolution.**

The anaphylatoxins C3a and C5a are key mediators of the complement system, which represents the first line of immunological defense for the recognition and elimination of microbes and pathogens.[1] They selectively bind to their respective G protein-coupled receptors (C3aR and C5aR) triggering a variety of pro-inflammatory processes and have recently been linked to a number of infectious, inflammatory, neurodegenerative and autoimmune diseases.[2] Currently, C3a is commercially produced in relatively low yields by biological means (recombinant expression or purification from plasma)[3, 4] resulting in a market value of approximately US$ 5000 per mg of protein. In addition, these approaches often require protein purification tags and additional steps for their removal and suffer from the inherent lability of C3a in biological fluids. C3a is rapidly inactivated within seconds by carboxypeptidase–mediated cleavage of a single C-terminal arginine residue ($Arg^{77}$).[5, 6] The resulting protein, C3a-desArg (also termed acylation-stimulating protein, ASP), does not bind C3aR, lacks any pro-inflammatory activity but instead has been shown to stimulate triglyceride synthesis and glucose uptake in adipose tissue.[7-9] To circumvent the problems associated with the isolation of C3a from biological sources we sought a total chemical synthesis approach enabling preparation of homogenous full length C3a.

Human C3a is a 77 residue protein containing three intramolecular disulfide bonds between C22-C49, C23-C56, and C36-C57.[10, 11] Because polypeptides of this size are difficult to obtain in high purity by standard stepwise solid phase peptide synthesis (SPPS),[12] we envisioned a fragment ligation approach by employing Kent's native chemical ligation (NCL).[13] The 77 amino acid polypeptide chain is retro-synthetically split into three peptide segments of about similar length and the three polypeptides are then joined consecutively in the C- to N-terminal direction by NCL as shown in Scheme 1. Temporary protection of $Cys^{23}$ was realized by substitution with thiazoldine-4-carboxylic acid (Thz).[14] The individual fragments C3a[1-22]-α-thioester, C3a[23-48]-α-thioester and C3a[49-77] were prepared by highly optimized *in situ* neutralization *tert*-butyloxycarbonyl (Boc) SPPS in high yield and purity.[15] Both ligations and the intermediate conversion of $Thz^{23}$ to $Cys^{23}$ were carried out in one-pot fashion.[14] The first NCL was carried out in 6M guanidine hydrochloride, 200 mM $Na_2HPO_4$, 50 mM Tris(2-carboxyethyl)phosphine•HCl (TCEP) and 50 mM 4-mercaptophenylacetic acid (MPAA), pH 7.1. The reaction was complete after 6h after which methoxyamine HCl was added and the pH adjusted to 4.0. After completion of the deprotection reaction, the mixture was readjusted to pH 7.1 and the second ligation initiated by adding C3a[1-22]-α-thioester (complete after 6 h). This one-pot approach afforded fully reduced C3a[1-77] without purification of intermediate products in short time (~24 h), good yield (41%) and high purity (SI Figure 1).




**Scheme 1**: Chemical synthesis of human C3a. a: native chemical ligation; b: Thz → Cys conversion; c: folding and disulfide formation. The amino acid sequence is given in the lower section with cysteines in bold, ligation sites underlined and connectivity of the three disulfides indicated.

The final folding of a cysteine-rich polypeptide chain into its 3D structure can often be problematic because they tend to yield multiple disulfide isomers that are difficult to separate and result in lower overall yield. This often necessitates optimization of folding conditions on a case-by-case basis. We found that folding and correct formation of the three disulfide bonds of C3a in 50 mM $Na_2HPO_4$, pH 7.5 supplemented with a redox system of 8 mM reduced glutathione, 1 mM oxidized glutathione was quantitative after 2 h at room temperature (Figure 1A) and native C3a was recovered in excellent yield after preparative RP-HPLC

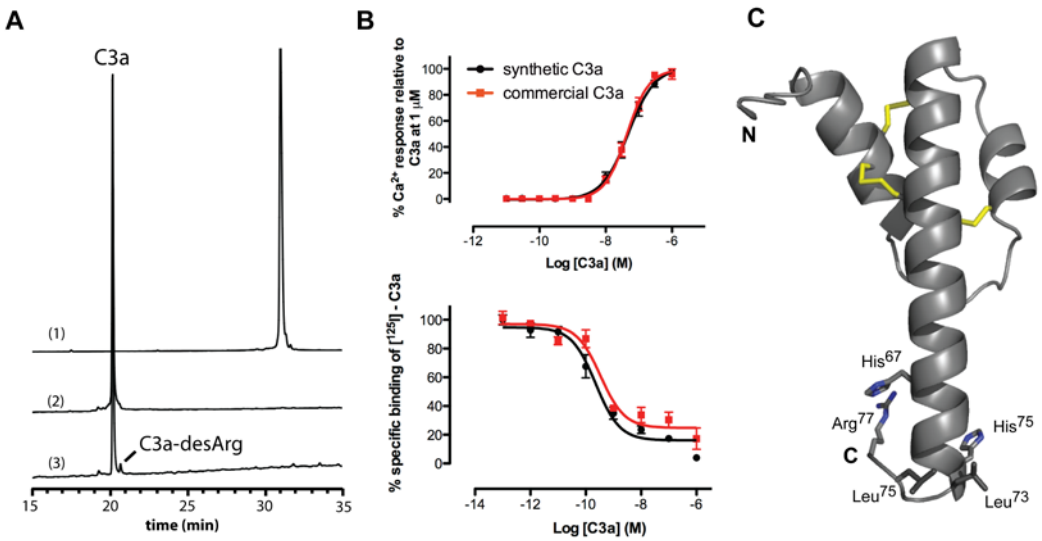



**Figure 1**: (A) RP-HPLC analysis of fully reduced C3a (1), the crude folding mixture after 2h (2) and commercially available material (3). (B) Concentration response curves for synthetic (●) and commercially available (■) C3a as determined by a fluorescence $Ca^{2+}$ release assays (upper panel) and by competitive radioligand binding assay using [$^{125}$I]-C3a binding to HMDM (lower pane). (C) Ribbon representation of the structure of synthetic C3a at 2.1 Å. Pharmacophore residues are shown in stick representation.

purification (61%). MS analysis suggested formation of the three disulfide bonds [observed MW: 9088.71±0.2 Da; calculated MW: 9088.65 Da (average isotope composition)]. Further HPLC analysis indicated that synthetic C3a was homogenous and devoid of truncated side-products, unlike commercially available material from biological sources (Figure 1A).

To confirm that C3a was folded in the correct biologically active conformation, intracellular $Ca^{2+}$ release and [$^{125}$I]-C3a radioligand binding assays were performed on human monocyte-derived macrophages (HMDM) expressing C3aR (Figure 1B). $EC_{50}$s of 48 ± 10 nM (synthetic) and 43 ± 7 nM (commercial) determined in the $Ca^{2+}$ release assay, as well as $IC_{50}$s of 0.2 ± 0.1 nM (synthetic) and 0.3 ± 0.1 nM (commercial) in the competitive radioligand binding assay are in excellent agreement with values reported previously[16] and indicate that the synthetic material is essentially indistinguishable from plasma-derived C3a. Furthermore, studies conducted with the C3aR-selective antagonist SB290157[17] as well as receptor desensitization experiments demonstrate that synthetic C3a is indeed selective for C3aR (SI Figure 4, 5).

To confirm the covalent structure of the synthetic material, we determined the 3D structure of synthetic full-length C3a by X-ray crystallography. Rhomboid-shaped crystals were obtained by the hanging-drop vapour diffusion method using 49–51% (v/v) 2-methyl-2,4-pentanediol (MPD) as the precipitant. Crystals appeared within 6 days after setting up the drops. The structure of C3a was determined by molecular replacement using the structure of C5a-desArg[18] as a search model and refined to a resolution of 2.1 Å, with a final *R* factor of 20.8% and an $R_{free}$ of 26.7%. Peptide backbone and amino acid side chains are generally well defined, except for residues 1-9, which are completely missing. Overall, the C3a structure is highly similar to the 3.2 Å structure of C3a-desArg described by Huber and Deisenhofer (Figure 1C).[11]

## Conclusions

In conclusion**,** we have demonstrated that total chemical synthesis is an efficient alternative for obtaining access to multi-milligram quantities of homogenous C3a. This approach gives complete control over the structure and will enable generation of unique chemical analogs of C3a (such as protease-resistant or fluorescently labeled variants) that will prove useful in future complement research. Based on our experience,[19] the one-pot NCL approach developed by Bang and Kent[14] is the most promising strategy for chemically accessing small and medium sized proteins up to 100 amino acids in length.

## Notes and references

[a] *Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Brisbane, Queensland, Australia 4072. Tel: +61-7-3346-2985 Fax: +61-7-3346-2090* thomas.durek@gmail.com

[b] *present Address: McMaster University, Hamilton, ON, Canada, L8S 4L8*

*Acknowledgments: This work was supported by a National Health and Medical Research Council (NHMRC) Program grant (569927 to RJL, PFA) and by The University of Queensland.*

† Electronic Supplementary Information (ESI) available: Full details on the synthesis, pharmacological assays and structure determination. Atomic coordinates and structure factors for C3a have been deposited in the PDB (ID: 4I6O).See DOI: 10.1039/b000000x/

‡ these authors contributed equally

1. D. Ricklin, G. Hajishengallis, K. Yang and J. D. Lambris, *Nat. Immunol.*, 2010, **11**, 785-797.
2. H. Qu, D. Ricklin and J. D. Lambris, *Mol. Immunol.*, 2009, **47**, 185-195.

3. E. P. Paques, H. Scholze and R. Huber, *Hoppe. Seylers Z. Physiol. Chem.*, 1980, **361**, 977-980.
4. G. Bajic, L. Yatime, A. Klos and G. R. Andersen, *Protein Sci.*, 2012.
5. V. A. Bokisch and Mullereb.Hj, *J. Clin. Invest.*, 1970, **49**, 2427-&.
6. S. L. Mueller-Ortiz, D. C. Wang, J. E. Morales, L. Li, J. Y. Chang and R. A. Wetsel, *J. Immunol.*, 2009, **182**, 6533-6539.
7. H. C. Wilken, O. Gotze, T. Werfel and J. Zwirner, *Immunol. Lett.*, 1999, **67**, 141-145.
8. J. Elsner, M. Oppermann, W. Czech, G. Dobos, E. Schopf, J. Norgauer and A. Kapp, *Eur. J. Immunol.*, 1994, **24**, 518-522.
9. P. J. Daffern, P. H. Pfeifer, J. A. Ember and T. E. Hugli, *J. Exp. Med.*, 1995, **181**, 2119-2127.
10. T. E. Hugli, *J. Biol. Chem.*, 1975, **250**, 8293-8301.
11. R. Huber, H. Scholze, E. P. Paques and J. Deisenhofer, *Hoppe. Seylers Z. Physiol. Chem.*, 1980, **361**, 1389-1399.
12. S. B. Kent, *Annu. Rev. Biochem.*, 1988, **57**, 957-989.
13. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776-779.
14. D. Bang and S. B. Kent, *Angew. Chem. Int. Ed.*, 2004, **43**, 2534-2538.
15. M. Schnölzer, P. Alewood, A. Jones, D. Alewood and S. B. Kent, *Int. J. Pept. Protein Res.*, 1992, **40**, 180-193.
16. C. C. G. Scully, J. S. Blakeney, R. Singh, H. N. Hoang, G. Abbenante, R. C. Reid and D. P. Fairlie, *J. Med. Chem.*, 2010, **53**, 4938-4948.
17. R. S. Ames, D. Lee, J. J. Foley, A. J. Jurewicz, M. A. Tornetta, W. Bautsch, B. Settmacher, A. Klos, K. F. Erhard, R. D. Cousins, A. C. Sulpizio, J. P. Hieble, G. McCafferty, K. W. Ward, J. L. Adams, W. E. Bondinell, D. C. Underwood, R. R. Osborn, A. M. Badger and H. M. Sarau, *J. Immunol.*, 2001, **166**, 6341-6348.
18. W. J. Cook, N. Galakatos, W. C. Boyar, R. L. Walter and S. E. Ealick, *Acta Crystallogr. Sect. D. Biol. Crystallogr.*, 2010, **66**, 190-197.
19. R. A. Morales, N. L. Daly, I. Vetter, M. Mobli, I. A. Napier, D. J. Craik, R. J. Lewis, M. J. Christie, G. F. King, P. F. Alewood and T. Durek, *ChemBioChem*, 2010, **11**, 1882-1888.