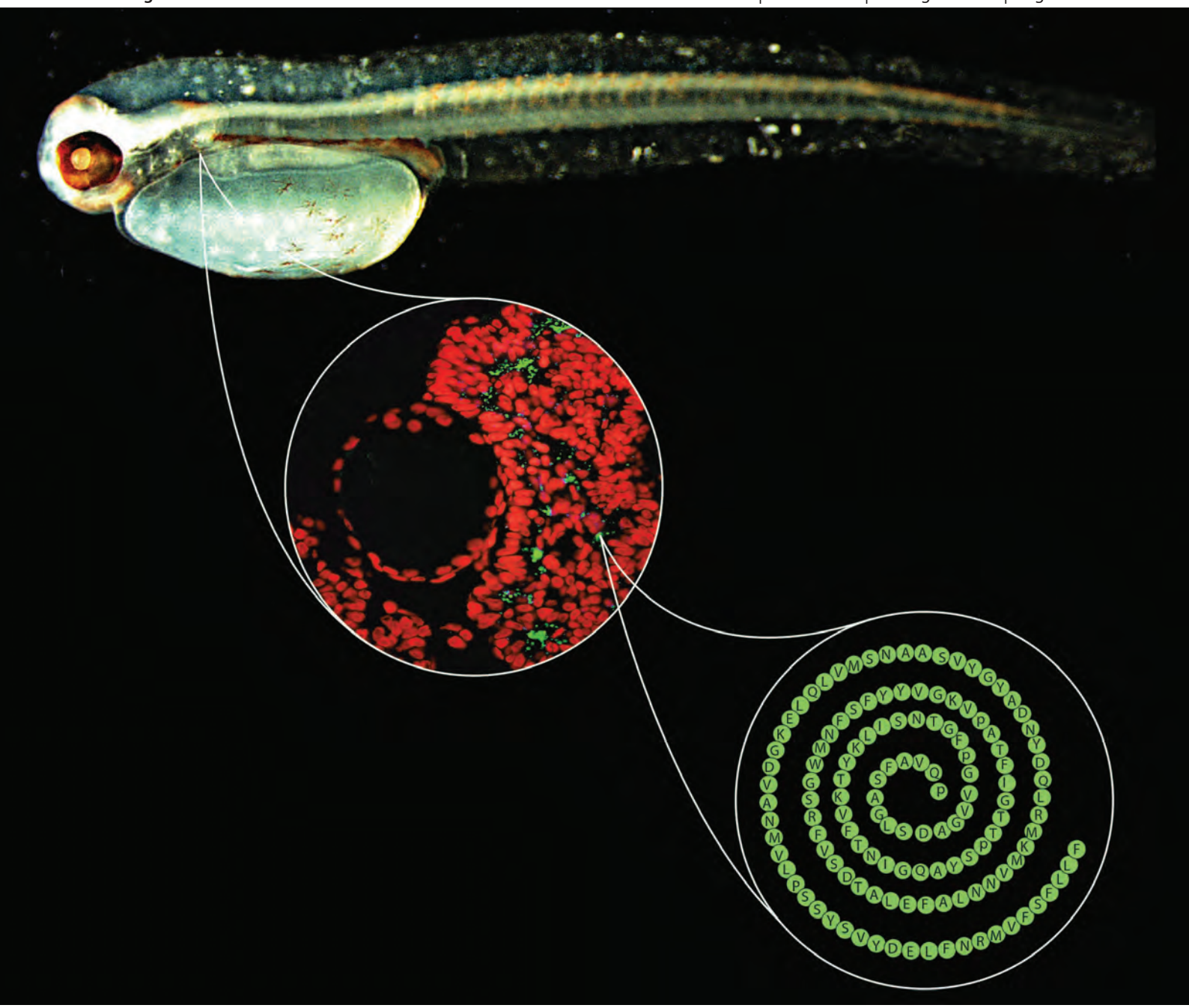
Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 11 | Number 30 | 14 August 2013 | Pages 4901–5048



ISSN 1477-0520

RSC Publishing

PAPER

Margaret A. Brimble *et al.*

Chemical synthesis of a masked analogue of the fish antifreeze potentiating protein (AFPP)

1477-0520(2013)11:30;1-A

Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 4935

Chemical synthesis of a masked analogue of the fish antifreeze potentiating protein (AFPP)†

Sung-Hyun Yang,[a] Joanna M. Wojnar,[a] Paul W. R. Harris,[a,b] Arthur L. DeVries,[c] Clive W. Evans[d] and Margaret A. Brimble*[a,b]

Received 23rd May 2013,
Accepted 12th June 2013

DOI: 10.1039/c3ob41066h

www.rsc.org/obc

A recently identified Antarctic fish protein termed antifreeze potentiating protein (AFPP) is thought to act as an adjunct to the previously characterised antifreeze glycoproteins (AFGPs), the two acting together to inhibit ice crystal growth *in vivo*. Elucidating the functional properties of the new AFPP requires access to large amounts of pure product, but the paucity of natural material necessitates alternative approaches. We therefore embarked on the total chemical synthesis of the AFPP, through a convergent ligation strategy. After many challenges, mostly due to the solubility issues of the peptide fragments, and several revisions of the original synthetic strategy, we have successfully synthesized a masked analogue of AFPP. The key to the successful synthesis was the use of a solubilising tag attached through a hydrolysable linker.

Introduction

Antarctic fishes survive in freezing seawater (*ca.* −1.9 °C) through the production of protective antifreeze proteins.[1] In previous work the DeVries laboratory isolated an Antarctic fish protein[2,3] termed antifreeze potentiating protein (AFPP) which can be considered a novel adjunct to the previously characterized antifreeze glycoproteins (AFGPs) circulating in the blood.[3] Like AFGP, AFPP isolated from Antarctic fish serum binds to ice crystals, inhibiting their growth. Almost all Antarctic fish exposed to freezing seawater that we have analyzed to date, have internalized ice crystals sequestered in macrophages within the spleen.[4,5] Since such crystals will contain adsorbed AFPP and AFGP, these observations raise the possibility that AFPP, in conjunction with AFGP, not only inhibits ice crystal growth in the blood but might also promote removal of potentially lethal circulating ice crystals by facilitating uptake of the ice/antifreeze complexes into spleen macrophages. One possibility is that this is a receptor-driven process (Fig. 1), the experimental evaluation of which will require significant amounts of pure AFPP.

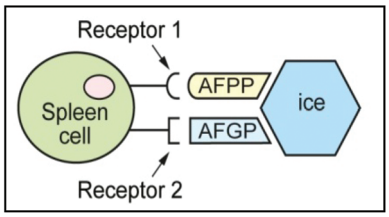


Fig. 1 Possible involvement of AFPP and AFGP in the removal of circulating ice crystals by macrophages in the spleen.

To investigate the role of AFPP in this hypothesis, access to sufficient amounts of pure protein is necessary. Additionally, modified AFPP analogues labelled with specific spectroscopic probes are required to facilitate further studies for functional analysis using biophysical methods (*e.g.* autoradiography; fluorescence microscopy; receptor characterization with photoactive cross-linkers). However, generation of functionally modified AFPP cannot be achieved by recombinant methods, and the paucity of natural material necessitates alternative approaches. Fortunately, the atom-by-atom precision achievable through chemical synthesis enables the site-specific incorporation of labels,[6] allowing us to examine adsorption of AFPP to ice and determine the nature of the postulated molecular interaction with macrophages in the spleen. We therefore embarked upon the total chemical synthesis of the new antifreeze potentiating protein.

[a]*School of Chemical Sciences, University of Auckland, 23 Symonds St., Auckland 1142, New Zealand. E-mail: m.brimble@auckland.ac.nz; Fax: +64 9 3737422; Tel: +64 9 3737599*
[b]*Maurice Wilkins Centre for Molecular Biodiscovery, School of Biological Sciences, The University of Auckland, Auckland 1142, New Zealand*
[c]*Department of Animal Biology, University of Illinois at Urbana–Champaign, 524 Burrill Hall, 407 S Goodwin, Urbana, IL 61801, USA*
[d]*School of Biological Sciences, University of Auckland, 3a Symonds St., Auckland 1142, New Zealand*
†Electronic supplementary information (ESI) available: Experimental procedures and LC-MS chromatograms. See DOI: 10.1039/c3ob41066h

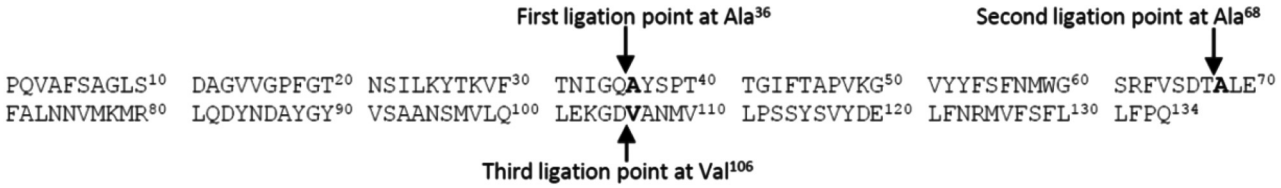


Fig. 2 The derived amino acid sequence of the C1q-AFPP from *Gymnodraco acuticeps* with proposed ligation sites (first ligation strategy) marked with arrows.

Results and discussion

A. First ligation strategy

Available evidence suggests that AFPP is a member of the C1q-like protein family.[7] A C1q-like domain was identified in the Antarctic naked dragonfish *Gymnodraco acuticeps* using molecular cloning techniques and its expression product was found to cross-react with an anti-AFPP.[8] An analogue of the derived C1q-like domain sequence (lacking a glycine at position 129, and with three additional amino acids at the N-terminus likely present in the native product, Fig. 2) was used in this study. This C1q-like analogue (referred to as C1q-AFPP) is 134 amino acids in length, well within what can be achieved[9] by chemical synthesis through native chemical ligation (NCL). NCL is the chemoselective reaction between an N-terminal cysteine and a C-terminal thioester to give a native peptide bond.[10] Unfortunately, the C1q-like domain of AFPP does not contain any natural cysteines, hence a ligation/desulfurization strategy was considered. First introduced by Yan and Dawson in 2001,[11] this method has extended the protein junctions accessible to NCL. In their work, native alanine-containing proteins were accessed through post-ligation desulfurization of cysteine, and more recently a variety of other thiolated building blocks that serve as precursors to natural amino acids have successfully been used in this strategy.[12]

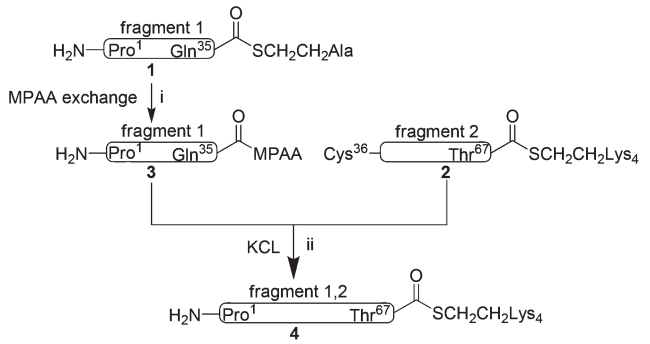
NCL of multiple peptide segments is only suitable for sequential ligations from the C- to the N-terminus. Yet because a convergent synthesis is more efficient,[13] we decided to use kinetically controlled ligation[14] (KCL) as part of a fully convergent ligation strategy. We aimed to split the protein sequence into peptides of approximately equal length. The N-terminal half of the sequence (see Fig. 2) was ideally suited for a KCL, as it had alanines at position 36 and 68. Substituting these for cysteines would allow a kinetically controlled ligation between fragment 1 (Pro^1-Gln^{35} thioaryl ester) and fragment 2 (Cys^{36}-Thr^{67} thioalkyl ester). The relative difference between the rate of ligation between a thioaryl ester and a thioalkyl ester[14] (the basis of the KCL reaction) is further supported by the difference in reactivity between glutamine and threonine thioesters,[15] thereby allowing the desired ligation to occur.

Unfortunately the C-terminal half of the sequence did not have a suitably placed alanine. The alanine at position 107 was most central, yet this ligation position would give a valine thioester for fragment 3. Valine thioesters are notoriously non-reactive, their ligation often requiring multiple days to go to completion.[15] Using one of the other alanines present (Ala93/94) in the sequence as a ligation site would leave a very long (>40 aa) fragment 4, which is approaching the limit for synthesis by solid phase methods. Given the recent reports of ligation at valine (using the valine surrogate dimethylcysteine or penicillamine (Pen))[16,17] we decided on a ligation site at Val^{106} instead.

Synthesis of peptide fragments. We have previously observed that even short AFPP sequences are prone to irreversible aggregation in solution, making their purification and further manipulation extremely difficult. Previous work has shown that addition of charged amino acids such as arginine or lysine to the C- or N-terminus can dramatically reduce aggregation and improve solubility of peptides and proteins.[18] This strategy was successfully applied by Johnson and Kent to the synthesis of hydrophobic peptide thioesters, where the charged amino acids were used as part of the thioester leaving group.[19] We decided on a similar approach for the three thioester-containing fragments, using multiple lysine residues.

Examination of the amino acid composition of C1q-AFPP (Fig. 2) shows that six methionine residues are present, whose methylthioether side chain is easily oxidized under acidic conditions that occur during Boc solid-phase peptide synthesis (SPPS) protocols. This can result in mixtures of reduced and oxidised methionine, which complicates analysis and purification by HPLC methods, and hence we incorporated methionine in its methionine sulfoxide form. Additionally, using Met(O) would increase the overall polarity of the peptides, which may increase their solubility. The sulfoxide would then be reduced[20–22] at the end of the synthesis to give the native methionine.

N-terminal half of C1q-AFPP. Fragment 1 (Pro^1-Gln^{35}-COS-CH_2CH_2-Ala) (**1**) and fragment 2 (Cys^{36}-Thr^{67}-COS-CH_2CH_2-Lys_4) (**2**) were synthesized using *in situ* neutralisation Boc-SPPS under standard reaction conditions (Scheme 1).[23] The synthesis of fragment 1 (**1**) gave a single product in high purity (see Fig. S1†). The peptide was quite insoluble in aqueous acetonitrile mixtures, but was fully soluble in the chaotrope 6 M guanidinium chloride (Gu·HCl). Subsequent exchange of the thioalkyl ester to its thiophenyl ester using the water soluble thiol 4-mercaptophenylacetic acid (MPAA)[24] (1 mM peptide concentration, 100 mM MPAA, 20 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl), pH 6.5) on the crude material proceeded smoothly in 2–3 h, giving Pro^1-Gln^{35}-COS-MPAA (**3**). The synthesis of fragment 2 (**2**) also gave the desired peptide as the major product (see Fig. S2†), and both peptide **2** and peptide **3** were purified under standard conditions.



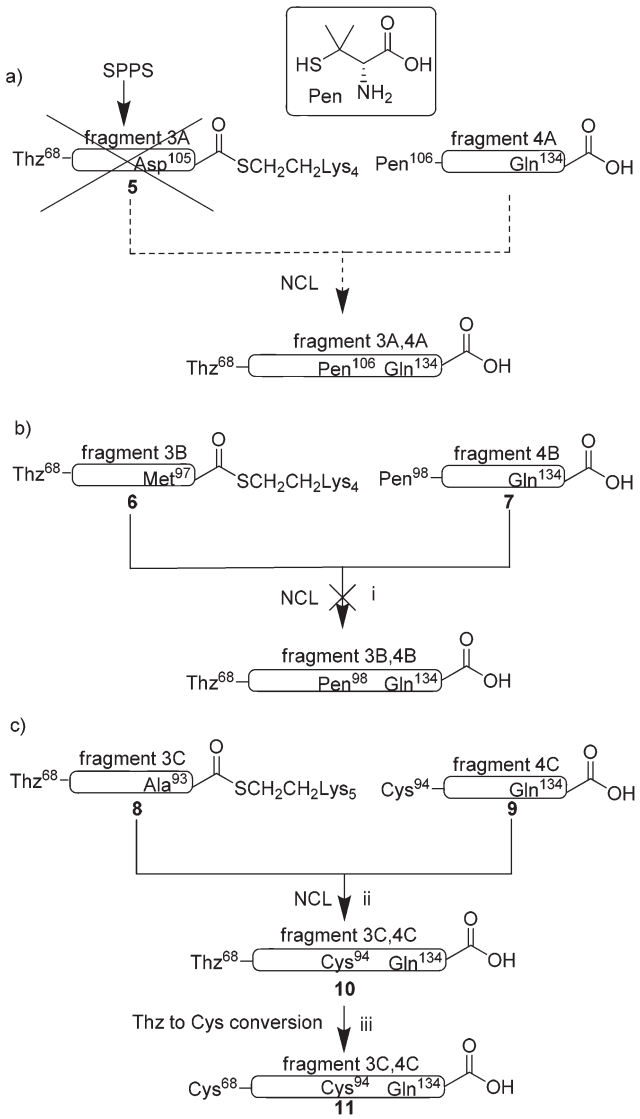
Scheme 1 The synthesis of the N-terminal half of C1q-AFPP, Pro^1-Thr^{67}-COS-CH_2CH_2-Lys_4 (**4**). *Conditions*: (i) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5; (ii) 20 mM TCEP, DMSO–0.2 M Na_2HPO_4 (9 : 1), pH = 6.9.

With the two fragments in hand, their ligation through the KCL reaction was attempted (Scheme 1). Unfortunately, the KCL was hampered by low solubility of both fragments, even with the presence of the Lys-solubilising tag on fragment 2 (**2**). Recent work has shown that ligation reactions can proceed in organic solvents,[25] so the KCL was attempted in DMF. This was once again unsuccessful, as peptide **2** exhibited very low solubility in this system. Furthermore, any addition of Na_2HPO_4 buffer resulted in precipitation of the fragments, and the reaction could not be buffered to optimal pH for KCL. Eventually, dimethyl sulfoxide (DMSO) was found to be the optimal solvent. Dissolution of fragment 1-MPAA (**3**) and fragment 2 (**2**) in DMSO followed by the addition of 20 mM TCEP in Na_2HPO_4 (pH adjusted to 6.9, final ratio 9 : 1) resulted in near quantitative ligation after 70 min, to give the first half of the AFPP (Pro^1-Thr^{67}-COS-CH_2CH_2-Lys_4) (**4**) (see Fig. S3†).

C-terminal half of C1q-AFPP. While fragment 4A (Pen^{106}-Gln^{134}-COOH) could be synthesised successfully, its ligation partner, fragment 3A (Thz^{68}-Asp^{105}-COS-CH_2CH_2-Lys_4) (**5**) (Thz = 1,3-thiazolidine), underwent extensive hydrolysis of the thioester (approx. 30–40%, see Scheme 2a and Fig. S4†) during its synthesis in agreement with previous reports.[26] While side-chain protection of the problematic glutamic and aspartic acid thioesters can avoid this issue, we decided instead to move the ligation point between fragment 3 and 4 to Val^{98} (see Scheme 2b). Fortunately, the new fragment 3B (now Thz^{68}-Met^{97}-COS-CH_2CH_2-Lys_4) (**6**) was synthesized in good yield, and with few byproducts (Fig. S5†).

As fragment 4B (Pen^{98}-Gln^{134}COOH) (**7**) does not contain a thioester, it was conveniently synthesized using Fmoc chemistry aided by a microwave synthesizer. Fmoc-Glu-O*t*Bu was attached *via* its side-chain carboxyl group to a Rink amide linker to give the desired C-terminal glutamine upon acid-mediated cleavage at the end of the synthesis.[27,28]

Penicillamine ligation issues. The penicillamine ligation was attempted under standard NCL conditions (1 mM peptide concentration, 100 mM MPAA, 20 mM TCEP·HCl, pH 6.5). Fragment 3B (**6**) rapidly underwent transthioesterification to the more active MPAA thioester intermediate, which



Scheme 2 Attempts at the synthesis of the C-terminal fragment of C1q-AFPP. (a) Ligation point at Pen^{106} – peptide **5** underwent substantial hydrolysis. (b) Ligation point at Pen^{98} – NCL could not be effected. (c) Ligation point at Cys^{94} allowed the successful synthesis of Thz-peptide **10**, which could be converted to Cys-peptide **11**. *Conditions*: (i) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5; (ii) 100 mM MPAA, 20 mM PBu_3, trifluoroethanol–0.2 M Na_2HPO_4 (1 : 1) pH = 7.0; (iii) 0.2 M $MeONH_2$·HCl, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 4.0.

could be observed by LC-MS. Unfortunately the subsequent step – transthioesterification with the penicillamine followed by intramolecular rearrangement to the desired ligated peptide – proved to be rate-limiting. The very slow ligation reaction was left overnight, during which time a significant portion of fragment 3B (**6**) underwent hydrolysis. While some of the desired ligation product was also observed, the long reaction time in the presence of the reducing agent TCEP caused some of the Met(O) residues in the peptides to be reduced to methionine, further complicating the LC-MS chromatogram. Attempts to optimize the ligation conditions were unsuccessful, and the hydrolysis of fragment 3B (**6**) continued

to be the dominant reaction, leading to very poor yields of the desired product.

Fragment 4C solubility issues. We therefore needed to revise our ligation strategy once more, and hence moved the ligation site to Ala[94]; consequently fragment 4C was now 41 amino acids long (see Scheme 2c). While the synthesis and purification of the shorter fragment 3C (Thz[68]-Ala[93]-COS-CH_2CH_2-Lys$_5$) (**8**) proceeded smoothly (see Fig. S6†), fragment 4C (Cys[94]-Gln[134]COOH) (**9**) was more difficult. As expected, the length of this peptide resulted in significant solubility issues, and standard chromatography conditions were not successful at resolving the peptide from the by-products. After many trials, fragment 4C (**9**) was eventually successfully purified (see Fig. S7†) on a diphenyl reversed-phase column eluted with a 1 : 1 mixture of CH_3CN and ^{i}PrOH containing 0.1% TFA for solvent B, at 55 °C. Preparation of fragment 4C (**9**) was also attempted using the isoacyl dipeptide strategy (using Boc-Ser-(Fmoc-Ser(O^{t}Bu))-OH), under microwave-mediated SPPS. This strategy has been used previously to improve both the synthesis and solubility of otherwise intractable peptides.[29] However, LC-MS analysis showed that a peptide with a single Ser deletion was the major product formed. Hence this strategy was not pursued further.

NCL of fragments 3C and 4C. The first attempted NCL of fragments 3C (**8**) and 4C (**9**) at a conc. 2 mM revealed that both of the fragments exhibit low solubility in standard NCL reaction buffer solution. A slight improvement in solubility was observed when further diluted to 0.5 mM, but still, the peptides were not fully soluble at this concentration, even with sonication and an elevated temperature of 40 °C. Solubility tests on the peptides revealed that fragment 3C (**8**) is sparingly soluble in the NCL buffer while very poor solubility was observed for fragment 4C (**9**).

In order to improve the solubility of both fragments in the ligation reaction, a 1 : 1 mixture of TFE and 0.2 M Na_2HPO_4 was used. While this indeed helped drive the NCL reaction to completion, a significant amount (50%) of the TFE ester of fragment 3C was also observed. This is not surprising, however, as fragment 3C (**8**) was used in excess and presumably unreacted thioester **8** undergoes hydrolysis by trifluoroethanol following completion of the ligation reaction. Furthermore, even upon short-term storage in the freezer, the purified fragment 4C (**9**) unexpectedly underwent dimerization, presumably *via* an intramolecular reaction through the sulfhydryl side chains of cysteine. The dimer could be readily reduced by treatment with TCEP in solution prior to the addition of the NCL partner, fragment 3C thioester (**8**).

Given that the ligation was now being performed in aqueous organic media, we investigated whether neutral organic phosphines would be more suitable and could be used in place of TCEP, which is typically supplied as the hydrochloride salt. Reduction of dimeric fragment 4 (**9**) was tested using three readily available phosphines – tributylphosphine, triphenylphosphine, and tricyclohexylphosphine – in 1 : 1 TFE–0.2 M Na_2HPO_4 or neat TFE. No significant solvent differences were found, and all phosphines performed in a similar manner.

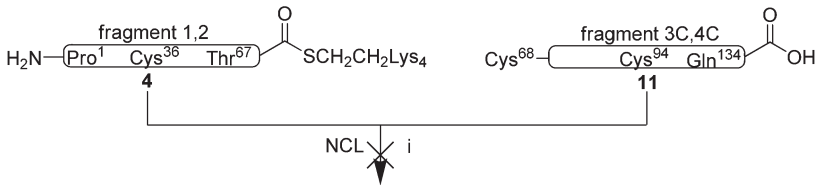
Tributylphosphine (chosen for its convenience, being a liquid) was thus added to the solution of dimeric fragment 4C in a neat TFE solution, and was left standing at room temperature for 30 min. Thz-fragment 3C-thioester (**8**) was then added into the above reaction mixture followed by addition of equivolume of pH adjusted (approx. 7.0) 100 mM MPAA in 0.2 M Na_2HPO_4. The reaction mixture was left overnight at room temperature, with subsequent LC-MS analysis indicating that the ligation went to completion, giving the desired C-terminal half of C1q-AFPP (Thz[68]-Gln[134]COOH) (**10**) (see Fig. S8†). Unreacted excess thioester was converted to its TFE ester under these conditions, but did not affect the yield of the desired product.

Unmasking of Thz to Cys of fragment 3,4. We next attempted a one-pot conversion[30] of the N-terminal thiazolidine group to cysteine of peptide **10**. Unfortunately, decomposition of the ligated Thz[68]-Gln[134]-OH (**10**) was observed overnight when 0.2 M $MeONH_2$·HCl was added directly to the NCL reaction mixture. It was suspected that the presence of tributylphosphine may impede the unmasking of Thz to Cys. Indeed, with the presence of tributylphosphine in the reaction mixture, addition of $MeONH_2$·HCl led to the formation of an unknown side product. By comparison, clean formation of desired Cys-fragment 3,4-OH (**11**) was observed (see Fig. S9†) when the tributylphosphine was completely removed by solid-phase extraction (SPE) prior to the addition of $MeONH_2$·HCl. Interestingly, the extraction step was not required when TCEP was used, allowing the addition of $MeONH_2$·HCl directly to the crude NCL reaction mixture.

Failure of final NCL reaction. The final NCL between fragment 1, 2 (**4**) and Cys-fragment 3–4 (**11**) (Scheme 3) was performed using the same reaction conditions used in the KCL reaction. Unfortunately, no desired product could be detected in the reaction mixture. Due to the ongoing solubility problems of the fragments, a new synthetic strategy was therefore sought.

B. Second, revised ligation strategy using a solubilising tag

The majority of the issues faced during this synthesis could be traced back to the poor solubility of the peptide fragments, especially of the terminal fragment 4 (**9**). Examination of the peptide sequence revealed that each fragment has very little overall charge and therefore was expected to exhibit poor water solubility. The polycationic solubility tags attached to the thioester of fragment 2 and 3 did improve their ease of handling, but unfortunately these are lost during the ligation reaction, exacerbating the solubility problem with each subsequent ligation step. We therefore decided to place a temporary solubilising tag[31] on the C-terminus of fragment 4, attached to a base-labile hydroxymethylbenzoic acid (HMBA) linker. This linker is stable to all the ligation conditions and both Fmoc and Boc SPPS, and being non-thioester based, would not interfere with any of the reactions. It would therefore remain on the peptide throughout the various synthetic and chromatographic steps, to be removed at the very end of the synthesis using aq. buffer at pH 8.[32] Given the failure of the final NCL ligation, we



Scheme 3 Failure of the final NCL reaction (first ligation strategy). *Conditions*: (i) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5.

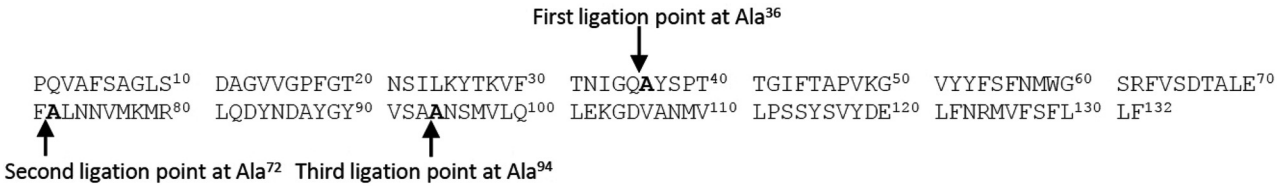
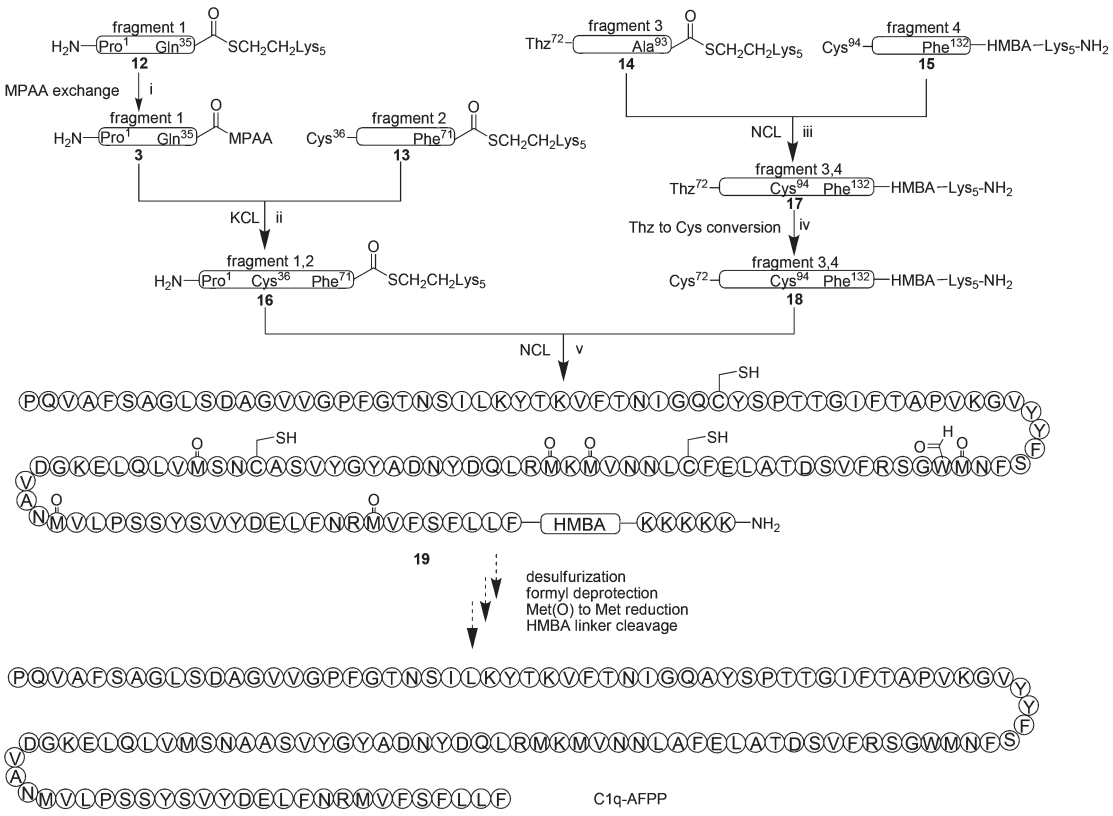


Fig. 3 The second, revised strategy ligation sites marked with arrows.



Scheme 4 Second, revised ligation strategy using a temporary solubilising tag. *Conditions*: (i) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5; (ii) 50 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5; (iii) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 6.5; (iv) 1 M $MeONH_2$·HCl, 50% aq. CH_3CN (+0.1% TFA); (v) 100 mM MPAA, 20 mM TCEP, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 7.0.

also moved the ligation point between fragment 2 and 3 to Phe^{71}/Ala^{72} (Fig. 3), as this would give the more reactive Phe^{71}-thioester for fragment 2. The new ligation point also has consequences for the fragment 1 and 2 KCL reaction. According to Hackeng *et al.*, a phenylalanine thioester is slightly more reactive than a glutamine thioester.[15] However, according to the relative reactivities of aryl *vs.* alkylthioesters,[14] the KCL reaction would still be the favoured one. The new, revised ligation strategy is presented in Scheme 4.

Synthesis of peptide fragments. Fragment 1 (Pro^1-Gln^{35}-COS-CH_2CH_2-Lys_5) (**12**), fragment 2 (Cys^{36}-Phe^{71}-COS-CH_2CH_2-Lys_5) (**13**), and fragment 3 (Thz^{72}-Ala^{93}-COS-CH_2CH_2-Lys_5) (**14**)

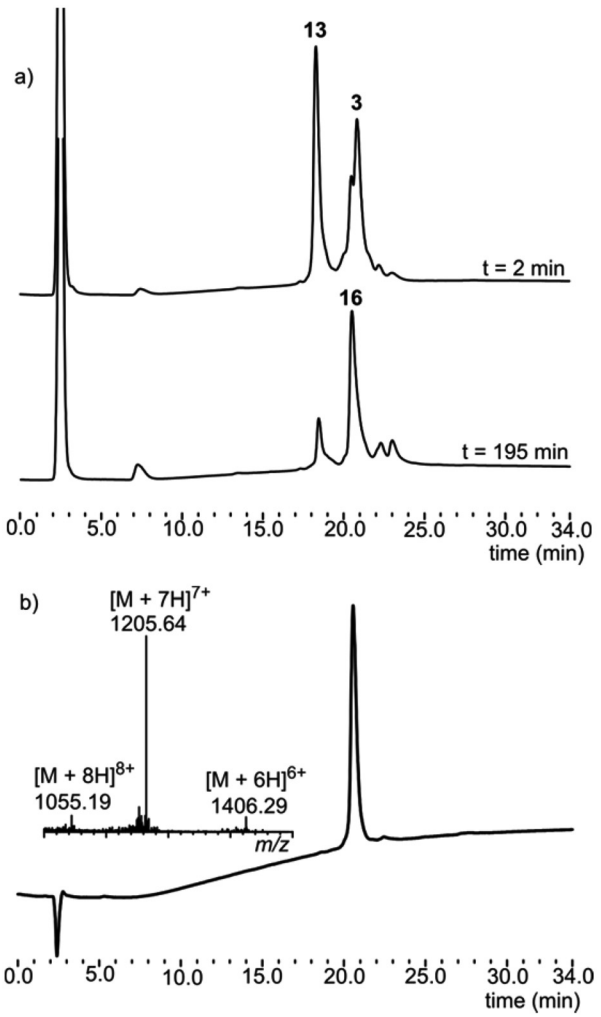


Fig. 4 (a) KCL of fragment 1 MPAA thioester (**3**) and fragment 2 (**13**). (b) The purified N-terminal half of C1q-AFPP (Pro^{1}-Phe^{71}-COS-CH_2CH_2-Lys_5) (**16**). Inset: the ESI-MS of the major peak, expected mass $[M + 7H]^{7+}$ 1205.38, observed mass 1205.64.

were synthesized using *in situ* neutralisation Boc-SPPS under standard reaction conditions. Fragment 4 (Cys^{94}-Phe^{132}-HMBA-Lys_5)‡ (**15**), being a non-thioester, was synthesized using Fmoc SPPS on a Rink amide resin with the aid of a microwave peptide synthesizer. Pleasingly, purification of all peptide fragments could be achieved using standard chromatographic conditions (solvent A = water + 0.1% TFA and solvent B = CH_3CN + 0.1% TFA) (see Fig. S10†).

N-terminal half of C1q-AFPP. The MPAA exchange of fragment 1 (**12**) was carried out as described for the previous strategy, to give the MPAA thioester **3** in 2–3 hours (see Fig. S11†). The KCL of fragment 1 MPAA thioester **3** and the Cys-fragment 2 (**13**), using 50 mM TCEP and a 1 mM final peptide concentration, proceeded smoothly and was complete after approximately 4 hours (see Fig. 4a). The N-terminal half of C1q-AFPP (Pro^{1}-Phe^{71}-COS-CH_2CH_2-Lys_5) (**16**) was obtained after standard chromatographic purification (see Fig. 4b).

C-terminal half of C1q-AFPP. The NCL reaction of fragment 3 (**14**) and 4 (**15**) proceeded smoothly once again under standard conditions (1 mM peptide conc., 100 mM MPAA and 20 mM TCEP), without the need for organic solvent, and was essentially complete after 3 hours (see Fig. 5a), giving the C-terminal half of C1q-AFPP-HMBA-Lys_5 (**17**) after purification (see Fig. 5b).

Unmasking of Thz in fragment 3,4 to Cys. Unfortunately, once again a one-pot Thz to Cys conversion of **17** proved to be problematic. Addition of 0.2 M $MeONH_2$·HCl (pH adjusted to approx. 4) after the completion of the NCL led to the decomposition of the Thz^{72}-Phe^{132}-HMBA-Lys_5-NH_2-fragment (**17**), and the appearance of a new, unknown peak on the LC-MS chromatogram. Use of 2 M $MeONH_2$·HCl, or hydrazine, led to the generation of an unknown by-product or the hydrazide ester, respectively. Various attempts were made to optimize this reaction, including changes to total peptide concentration, $MeONH_2$·HCl concentration, and other aqueous conditions. Our results indicated that a low final concentration of peptide and higher concentration of $MeONH_2$·HCl was required for this reaction to proceed without the formation of by-products.

Eventually, clean conversion was achieved by dissolving the Thz^{72}-Phe^{132}-HMBA-Lys_5-NH_2 (**17**) (after SPE and lyophilisation) in a mixture of 1 : 1 acetonitrile and water (+0.1% TFA) to give a final peptide concentration of about 0.07 mM. $MeONH_2$·HCl was added to give a concentration of 1 M, and the solution was left to react for 2 days, giving the desired Cys^{72}-Phe^{132}-HMBA-Lys_5-NH_2 – fragment (**18**) (see Fig. S12†).

Final NCL. With the second half of the protein finally in hand, the final NCL reaction was attempted. Pro^{1}-Phe^{71}-COS-CH_2CH_2-Lys_5 (**16**) was reacted with Cys^{72}-Phe^{132}-HMBA-Lys_5-NH_2 (**18**) under standard NCL conditions (100 mM MPAA, 20 mM TCEP and a 1 mM final peptide concentration at pH 7.0), to give the masked C1q-AFPP-HMBA-Lys_5-NH_2 (**19**) after 17 hours (see Fig. 6).

Activity testing. The final NCL product (**19**) was tested for antifreeze activity using a Clifton nanoliter osmometer as described elsewhere.[1,33] The defining characteristic of biological antifreeze is the presence of hysteresis activity, measurable as a difference in the freezing and melting points of a small ice crystal in a solution of antifreeze. Soluble, native AFPP has a hysteresis activity of approximately 2 °C at 4 mg mL^{-1} in 50 mM NH_4HCO_3 buffer (pH 7.8), when determined with a small seed crystal, slow cooling and annealing for 1 h at 0.1 °C above the hysteresis freezing point. Despite the presence of a solubilizing tag to aid synthesis, attempts to dissolve synthetic C1q-AFPP at concentrations up to 2 mg mL^{-1} in the same bicarbonate buffer were unsuccessful. This necessitated solubilizing the product first in 1 : 1 acetonitrile–50 mM NH_4HCO_3 with subsequent removal of the acetonitrile by dialysis (2000 MW cut-off) against the bicarbonate buffer and reduction to a final concentration of approximately 4 mg mL^{-1}. Further concentration led to increasing precipitation. Hysteresis determinations were highly variable with a

‡Both the C-terminal Gln and Pro were removed for synthetic convenience in this strategy.

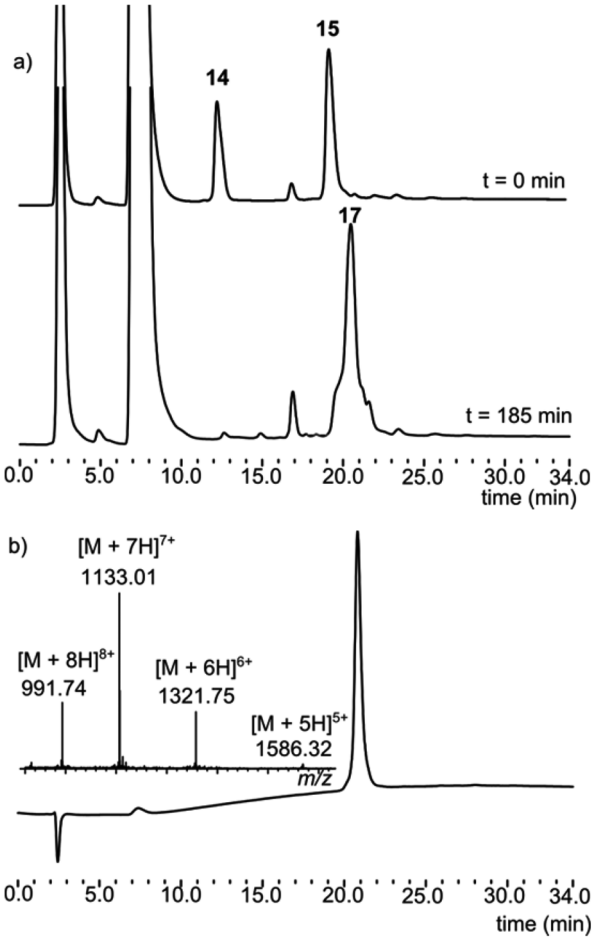


Fig. 5 (a) The NCL of fragment 3 (**14**) and 4 (**15**). (b) The purified C-terminal half of C1q-AFPP-HMBA-Lys_5 (Thz^{71}-Phe^{132}-HMBA-Lys_5-NH_2) (**17**). Inset: the ESI-MS of the major peak, expected mass $[M + 7H]^{7+}$ 1133.18, observed mass 1133.01.

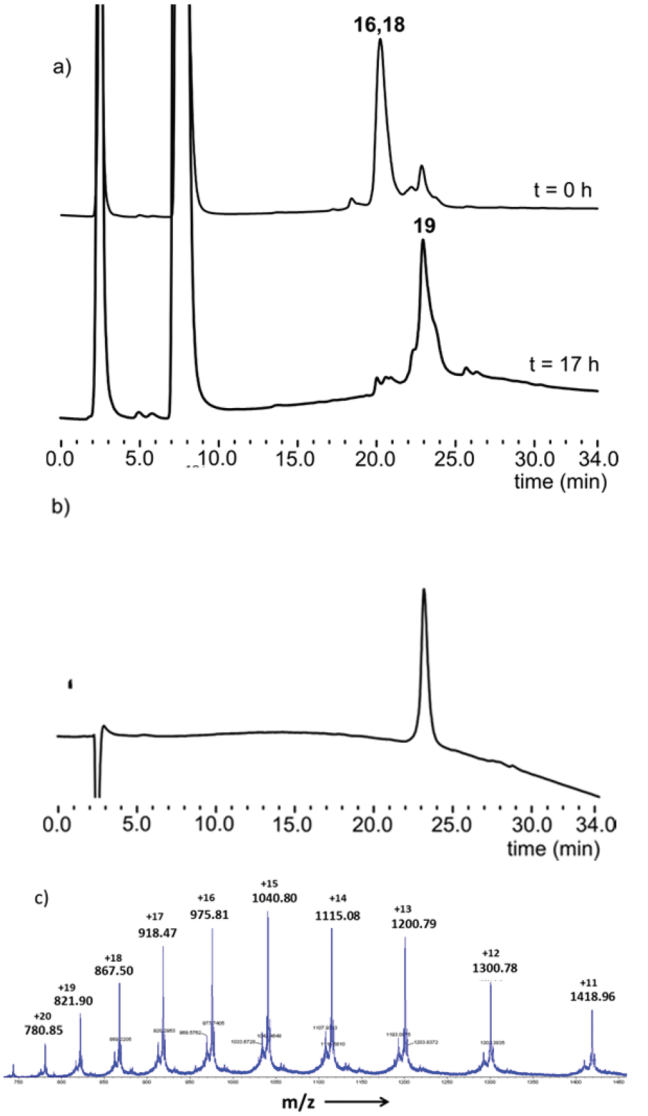


Fig. 6 (a) The final NCL reaction between the N-terminal half **16** and the C-terminal half **18**. (b) The purified C1q-AFPP-HMBA-Lys_5-NH_2 (**19**). (c) ESI-MS of the major peak. The calculated mass is 15 595.71. Deconvolution of the ESI mass spectrum yields an observed mass of 15 597.6 ± 0.12 Da.

maximum of 0.76 °C and a minimum of 0.22 °C. The maximum hysteresis value was observed when the seed ice crystal was very small (<10 μm), the cooling rate was 0.01 °C min^{-1} and the sample was annealed for 1 h at slightly above its hysteresis freezing point. Significantly, the morphology of ice growth at the freezing point was in the form of expanding irregular hexagons (*a*-axis growth), which is the same as the growth direction with native AFPP.

A second distinguishing characteristic of AFPP is that it potentiates the hysteresis activity of AFGP, especially the larger AFGP1–5 isoforms.[2] When synthetic C1q-AFPP was tested in the presence of AFGP1–5, however, no potentiation was observed. Several explanations are possible, one being that the synthetic analogue may not have folded fully into its native conformation, as attested by its poor solubility.

Conclusions

We have demonstrated that a temporary solubilising tag can be successfully used in a convergent native chemical ligation synthesis of a poorly water soluble protein. The solubilising tag proved to be essential to the synthesis of a masked analogue of C1q-AFPP, allowing for improved handling and purification of protein fragments. The 132 amino acid protein with a pentalysine tag was successfully synthesised through a convergent strategy using both kinetically controlled ligation and native chemical ligation to join four separate peptide fragments.

Acknowledgements

This research was supported by the Human Frontier Science Programme (grant RGP003/2009-C).

Notes and references

1 A. L. DeVries, *Science*, 1971, **172**, 1152–1155.
2 Y. Jin, PhD Thesis, University of Illinois, 2003.
3 A. L. DeVries and C.-H. C. Chen, in *The Physiology of Polar Fishes*, eds. A. P. Farrell and J. F. Steffenson, Elsevier, Amsterdam, 2005, pp. 155–201.
4 C. W. Evans, V. Gubala, R. Nooney, D. E. Williams, M. A. Brimble and A. L. DeVries, *Antarct. Sci.*, 2010, **23**, 57–64.
5 C. W. Evans, L. Hellman, M. Middleditch, J. M. Wojnar, M. A. Brimble and A. L. DeVries, *Antarct. Sci.*, 2012, **24**, 259–268.
6 D. J. Lee, K. Mandal, P. W. R. Harris, M. A. Brimble and S. B. H. Kent, *Org. Lett.*, 2009, **11**, 5270–5273.
7 C.-H. C. Cheng, personal communication.
8 C. W. Evans and S. Richter, unpublished work.
9 K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon and A. Brik, *Angew. Chem., Int. Ed.*, 2011, **50**, 6137–6141.
10 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776–779.
11 L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526–533.
12 H. Rohde and O. Seitz, *Pept. Sci.*, 2010, **94**, 551–559.
13 J. B. Hendrickson, *J. Am. Chem. Soc.*, 1977, **99**, 5439–5450.
14 D. Bang, B. L. Pentelute and S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2006, **45**, 3985–3988.
15 T. M. Hackeng, J. H. Griffin and P. E. Dawson, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 10068–10073.
16 C. Haase, H. Rohde and O. Seitz, *Angew. Chem., Int. Ed.*, 2008, **47**, 6807–6810.
17 J. Chen, Q. Wan, Y. Yuan, J. Zhu and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2008, **47**, 8521–8524.
18 A. Kato, K. Maki, T. Ebina, K. Kuwijima, K. Soda and Y. Kuroda, *Biopolymers*, 2006, **85**, 12–18.
19 E. C. B. Johnson and S. B. H. Kent, *Tetrahedron Lett.*, 2007, **48**, 1795–1799.
20 E. Nicolas, M. Vilaseca and E. Giralt, *Tetrahedron*, 1995, **51**, 5701–5710.
21 L. Taboada, E. Nicolas and E. Giralt, *Tetrahedron Lett.*, 2001, **42**, 1891–1893.
22 M. Vilaseca, E. Nicolas, F. Capdevila and E. Giralt, *Tetrahedron*, 1998, **54**, 15273–15286.
23 M. Schnolzer, P. Alewood, A. Jones, D. Alewood and S. B. H. Kent, *Int. J. Pept. Res. Ther.*, 2007, **13**, 31–44.
24 E. C. Johnson and S. B. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640–6646.
25 M. Dittman, J. Sauermann, R. Seidel, W. Zimmerman and M. Engelhard, *J. Pept. Sci.*, 2010, **16**, 558–562.
26 M. Villain, H. Gaertner and P. Botti, *Eur. J. Org. Chem.*, 2003, 3267–3272.
27 G. Breipohl, J. Knolle and W. Stüber, *Int. J. Pept. Protein Res.*, 1990, **35**, 281–283.
28 F. Albericio, R. Van Abel and G. Barany, *Int. J. Pept. Protein Res.*, 1990, **35**, 284–286.
29 Y. Sohma, T. Yoshiya, A. Taniguchi, T. Kimura, Y. Hayashi and Y. Kiso, *Pept. Sci.*, 2007, **88**, 253–262.
30 D. Bang and S. B. H. Kent, *Angew. Chem., Int. Ed.*, 2004, **43**, 2534–2538.
31 M. A. Hossain, A. Belgi, F. Lin, S. Zhang, F. Shabnanpoor, L. Chan, C. Belyea, H.-T. Truong, A. R. Blair, S. Andrikopoulos, G. W. Tregear and J. D. Wade, *Bioconjugate Chem.*, 2009, **20**, 1390–1396.
32 P. W. R. Harris and M. A. Brimble, *Biopolymers*, 2010, **94**, 542–550.
33 A. L. DeVries, in *Methods Enzymol*, ed. L. Packer, Academic Press, New York, London, 1986, pp. 293–303.