Genetic Expression of an Amyloid Peptide Fragment and Analysis of Formylated Products

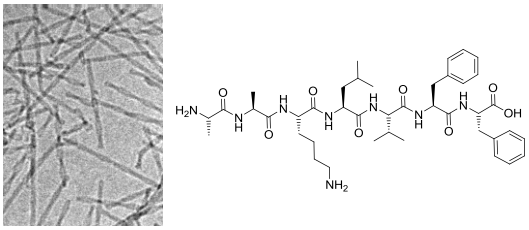
G. Cheng, C. Krasel,[†] H. G. Zhou,[†] D. Chappell, I. W. Hamley*

School of Chemistry, Food Science and Pharmacy, University of Reading, Whiteknights, Reading RG6 6AD, UK
[†] Currently at Institut für Pharmakologie & Toxikologie, Karl-von-Frisch-Strasse 1 35041 Marburg, Germany

* I.W.Hamley@reading.ac.uk. *Also at Diamond Light Source, Harwell Science and Innovation Campus, Chilton, Didcot OX11 0DE, UK*

Received Date (will be automatically inserted after manuscript is accepted)

ABSTRACT



The model amyloid peptide AAKLVFF was expressed as a His-tagged fusion protein with the immunoglobulin-binding domain B1 of streptococcal protein G (GB1), a small (56 residues), stable, single-domain protein. We show that expression of this model amyloid peptide is possible, and is not hindered by aggregation. Formylation side reactions during the CNBr cleavage are investigated via synthesis of selectively formylated peptides.

Expression of proteins using recombinant DNA methods is now ubiquitous, however expression of amyloid-forming proteins is potentially complicated by aggregation of the peptide during or after expression. Indeed, short amyloid forming segments can drive a non-fibrillizing protein into the amyloid state. This was shown by Teng and Eisenberg who demonstrated that when 6-8 residue fragments from amyloid-forming human proteins tau, α-synuclein, PrP prion and amyloid β (Aβ) were inserted into a region of the enzyme RNase A amyloid formation occured. [1] Baxa *et al.* showed that fusion of the Ure2p prion protein with various proteins led to amyloid formation.[2] The functions of the proteins (barnase, carbonic anhydrase, glutathione S-transferase and green

fluorescent protein) were not substantially influenced showing that they retained their native structures. Amyloid fibrils have been used as scaffolds to present proteins such as cytochrome c, which retains its function on fibrils formed from a fusion construct with the amyloid forming SH3 domain.[3] It is therefore of great interest to determine methods by which amyloid-forming short peptides can be expressed without driving amyloid assembly of the host protein in the fusion construct.

The Hamley group have recently investigated the self-assembly of peptide NH_2-AAKLVFF-COOH in detail. [4-7] This peptide is based on sequence Aβ(16-20), KLVFF, from the amyloid β peptide, extended at the N terminus with two hydrophobic alanine residues. In water at sufficiently high concentration, this peptide forms twisted fibrils (as observed by TEM and cryo-TEM),[4, 6, 7] whereas in methanol it forms nanotubes.[5-7] Both these structures have also been reported for the related peptide Aβ(16-22), KLVFFAE, by varying pH in aqueous solution.[8, 9] The formation of β-sheet structures in both solvents has been confirmed by FTIR, x-ray diffraction experiments in solution and circular dichroism (CD) spectroscopy on dried films.[4-7] In dilute solution, the CD spectra reveal the absence of β-sheet ordering but instead show features of a disordered conformation with a possible contribution from aromatic stacking interactions resulting from the phenylalanine residues. Solution NMR experiments were used to examine the solubility of the peptide in the two solvents, and to determine the critical aggregation concentration.[7]

In the present paper, we report on the expression of AAKLVFF using a recombinant protein. This work was motivated by the desire to explore the synthesis of the peptide on a larger scale using appropriate hosts (eg. bacteria) and this work provides proof-of-concept of this. It is noted that the cyanogen bromide cleavage step leads to formylated peptide, and the nature of the formylated product was investigated in detail via synthesis of AAKLVFF formylated either just at the N terminus or additionally at the K residue. This indicates that CNBr produces a peptide with backbone formylation at the N terminus. Formylation at the ε-amino group in lysine was achieved by blocking the peptide N terminus in azido-AAKLVFF. Formic acid /acetic anhydride did not give a formylated product, however, successful formylation was achieved with *p*-nitrophenol in a borate buffer (pH=10) / acetonitrile (1 : 1). Formylation of peptides and proteins has been investigated previously,[10-12] however we are not aware of prior reports on this in the preparation of recombinant amyloid-type peptides.

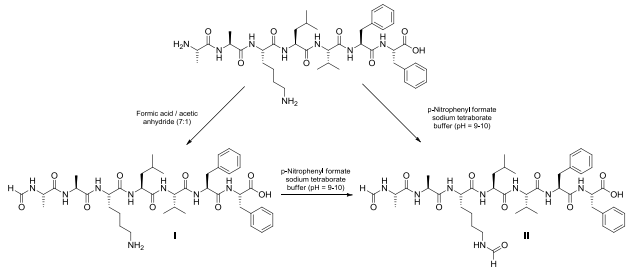
We used the GB1 fusion protein domain to express the model amyloid forming peptide AAKLVFF. Theoretically**,** cleaving the GB1-AAKLVFF fusion protein sequence by CNBr leads to sequences shown in Table 1, based on the UniProt Knowledgebase (Swiss-Prot and TrEMBL).[13] The corresponding peptide masses ($[M+H]^+$) from the sequence are also shown in Table 1.

mass	peptide sequence
6024.4586	HIAAGACCTTTACAGTTACT GAACATATGGCGGCGAAACT GGTGTTCTTTTAAGGATCCK TFTVTEHM
5229.8631	GGTACCATGGGCAGCAGCCA TCATCATCATCATCACACTT ACAAATTAATCCTTAATGGT M
795.4763	AAKLVFF

Table 1 Theoretical analysis of peptide masses ($[M+H]^+$) cleaved from GB1-AAKLVFF

The ESI-MS spectrum of the product obtained from RP-HPLC was compared with that of synthetic AAKLVFF, revealing the $[M + H]^+$ peak 795.4767 (genetic) / 795.4765 (synthetic), (calc. 795.4771) and the doubly protonated peak $[M+2H]^+$ at 398.2419 (genetic) / 398.2418 (synthetic) (calc. 398.2426) in both spectra (SI Fig.1). In addition, fragmention analysis confirmed the sequence AAKLVFF. However, an ion at m/z 851.6071 was also detected in the mass spectrum of the cleaved sequence AAKLVFF corresponding to an additional mass of 56 Da which is consistent with formylation of both primary amines by the formic acid.

In order to investigate the formation of formylated byproducts in more detail, formylation of a sample of AAKLVFF in 98% formic acid was conducted in the presence of acetic anhydride[10] in addition to a control experiment which consisted of the formylation of synthetic AAKLVFF in 70% formic acid with an excess of cyanogen bromide. The experiments were both monitored by RP-HPLC and after 6 hours the formylation using 98% formic acid displayed a new peak at 10.03 min (Peptide **I**, Scheme 1) whereas, in the control experiment, a new peak was observed at 10.59 min after 7 days (Peptide **II**) as shown in SI Fig 2.

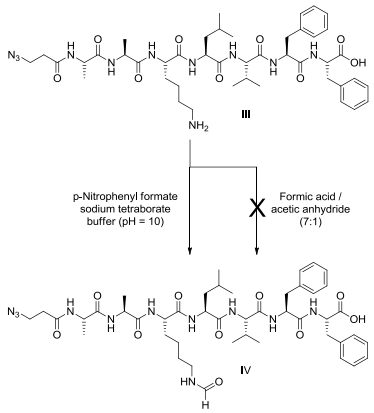


Scheme 1. Formylation of peptide AAKLVFF to give peptides **I** and **II**.

Analysis of the fragmentation peaks in the mass spectrum of peptide **I** revealed the $[M+H]^+$ peak 823.4708

(calc. 823.4720) indicating that it was the mono-formylated product, formylated at the backbone N terminus (Scheme 1).

In order to investigate further the single selective formylation reaction, the peptide azido-AAKLVFF **III** blocked at the N-terminus with 3-azidopropanoic acid which contained only the primary amine on the lysine side-chain was synthesized (Scheme 2). In this case, formylation was not observed in 98% formic acid /acetic anhydride, however, successful formylation to produce peptide **IV** was achieved based on a method reported by Dempsey,[10] using *p*-nitrophenol in a borate buffer (pH=10) / acetonitrile mixture (1 : 1) (Scheme 2). As a consequence, formylation of peptide AAKLVFF was attempted using these same conditions and the isolated product displayed an identical retention time (t_R) with peptide **II**. The ESI-MS spectrum of the product showed the $[M + H]^+$ ion at 851.4673, thereby confirming bis-formylation, both on the backbone and on the lysine side-chain (calc. 851.4669).



Scheme 2. Formylation of peptide azido-AAKLVFF (**III**) to produce peptide **IV**.

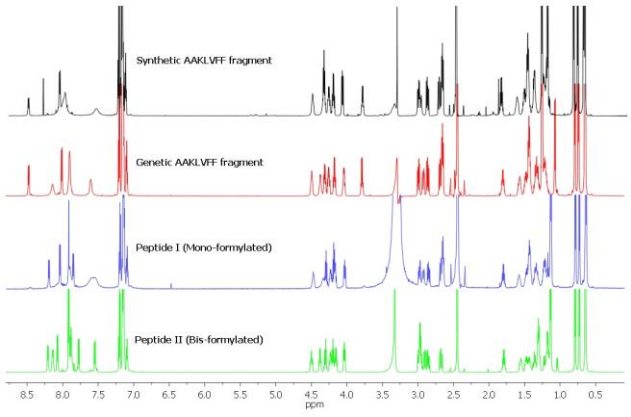


Figure 1. Comparison of ^{1}H-NMR (DMSO-d_6) of genetic and synthetic AAKLVFF fragments, Peptide **I** (formylation on backbone) and Peptide **II** (formylation on backbone and lysine side chain).

The genetic AAKLVFF was isolated from its mono-formylated (Peptide **I**) and bis-formylated (Peptide **II**) analogues following repeated semi-preparative HPLC (SI Fig. 2b). All of the ^{1}H-NMR data have been analysed (Figure 1), and are reported in the experimental section (Supporting Information).

The self-assembly of peptides **I**, **II** and **III** was studied using FTIR spectroscopy, circular dichroism (CD) and cryogenic-transmission electron microscopy (cryo-TEM). We have previously reported in detail on the self-assembly of synthetic AAKLVFF in water, using these techniques. Fig. 2 shows representative cryo-TEM images of these peptides, showing self-assembly into fibrils. Due to insolubility in water, peptide **II** was dissolved in a 1:1 water/acetonitrile mixture. Peptide **I** exhibits short rigid fibrils, which at high magnification can be seen to be twisted. In contrast, peptide **II** shows longer, less straight fibrils which at higher magnification do not, in general, show twisting. The difference in morphology of these samples highlights the influence of packing and electrostatic interactions on the intermolecular stacking of the β-sheets. Peptide **III** formed a very dense network of extended fibrils as shown in Fig.2c. It was also necessary to dilute this sample (compared to conditions used for **I** and **II**) in order to image the fibrillar network by cryo-TEM. The presence of the azido group appears to enhance fibrillisation of AAKLVFF, possibly due to enhanced hydrophobic interactions.

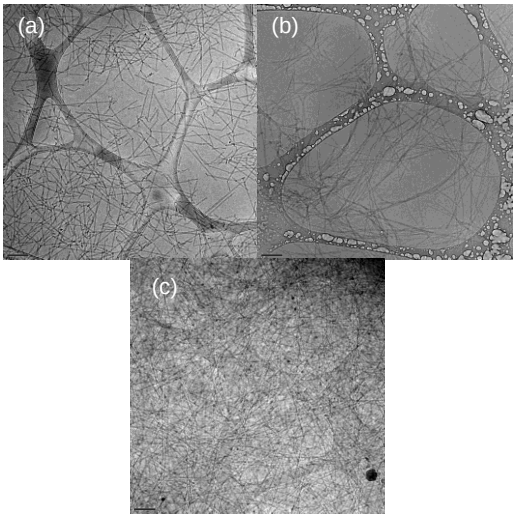


Figure 2. Cryo-TEM images (a) mono-formylated peptide **I** (2 wt% in H_2O), (b) bis-formylated peptide **II** (2 wt% in 1:1 H_2O:CH_3CN) , (c) azido-AAKLVFF, peptide **III** (0.012 wt% in H2O). The scale bars represent 200 nm.

These fibrils are rich in β-sheet as shown by the representative FTIR spectrum for peptide III in Fig.3a (and our previous extensive characterization of

AAKLVFF[4, 6, 7]). FTIR spectroscopy in the amide I region provides information on secondary structure. Fig.3a shows a peak at 1625 cm^{-1} associated with β-sheet and a smaller peak at 1672 cm^{-1} often ascribed to the presence of TFA counterions.[14-16] The CD spectrum (Fig.3b) measured under the same conditions shows a maximum at around 190 nm and abroad minimum around 210 nm, also consistent with β-sheet structure.

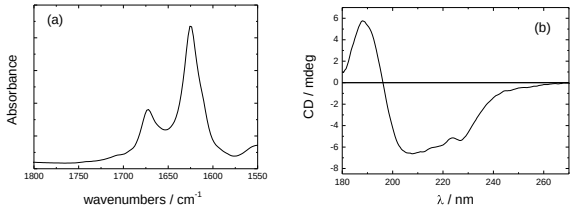


Figure 3. (a) FTIR and (b) CD data for an 0.98 wt% solution of peptide **III** in D_2O.

In summary, we have shown that genetic engineering methods can be used to produce a model amyloid peptide. CNBr cleavage leads to a product that is formylated at the N terminus and on the lysine ε-amino group. Self-assembly of mono- and bis- formylated AAKLVFF leads to fibrils with distinct structures. A model system with an azido-functionalized N terminus (blocked for formylation studies) also self-assembles, with a much greater density of fibrils than for the other peptides and with well resolved β-sheet features in the FTIR and CD spectra. The azido-functionalized peptide is also under investigation as a building block to construct three dimensional chemical structures, e.g. star-type macromolecules or peptide-functionalized hyperbranched molecules, by "click" chemistry.

Acknowledgment This work was supported by EPSRC grant EP/F048114/1 to IWH that supported GC and EPSRC Platform grant EP/G026203/1 that supported DC. We thank Dr Valeria Castelletto and Ms Claire Moulton for assistance with FTIR and CD spectroscopy. We are grateful to Steve Furzeland and Derek Atkins (Unilever, Colworth, UK) for the cryo-TEM experiments. We thank the University of Reading for the award of a RETF studentship for Mr H.G.Zhou.

Supporting Information Available Experimental methods, Scheme showing DNA sequence of fusion construct used to synthesize GB1-AAKLVFF, electro-spray mass spectra, RP-HPLC chromatograms.

1. Teng, P. K.; Eisenberg, D. *Protein Engineering, Design and Selection* **2009,** 22, (8), 531-536.
2. Baxa, U.; Speransky, V.; Steven, A. C.; Wickner, R. B. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** 99, (8), 5253-5260.
3. Baldwin, A. J.; Bader, R.; Christodoulou, J.; MacPhee, C. E.; Dobson, C. M.; Barker, P. D. *J. Am. Chem. Soc.* **2006,** 128, (7), 2162-2163.
4. Castelletto, V.; Hamley , I. W.; Harris, P. J. F. *Biophys. Chem.* **2008,** 138, 29-35.
5. Krysmann, M. J.; Castelletto, V.; McKendrick, J. M. E.; Hamley , I. W.; Stain, C.; Harris, P. J. F.; King, S. M. *Langmuir* **2008,** 24, 8158-8162.
6. Castelletto, V.; Hamley , I. W.; Harris, P. J. F.; Olsson, U.; Spencer, N. *J. Phys. Chem. B* **2009,** 113, 9978-9987.
7. Hamley, I. W.; Nutt, D. R.; Brown, G. D.; Miravet, J. F.; Escuder, B.; Rodríguez-Llansola, F. *J. Phys. Chem. B* **2010,** 114, 940-951.
8. Mehta, A. K.; Lu, K.; Childers, W. S.; Liang, S.; Dong, J.; Snyder, J. P.; Pingali, S. V.; Thiyagarajan, P.; Lynn, D. G. *J. Am. Chem. Soc.* **2008,** 130, 9829-9835.
9. Liang, Y.; Pingali, S. V.; Jogalekar, A. S.; Snyder, J. P.; Thiyagarajan, P.; Lynn, D. G. *Biochemistry* **2008,** 47, (36), 10018-10026.
10. Dempsey, C. E. *J. Chem. Soc., Perkin Trans. 1* **1982,** 11, 2625-2629.
11. Goodlett, D. R.; Armstrong, F. B.; Creech, R. J.; van Breemen, R. B. *Anal. Biochem.* **1990,** 186, (1), 116-120.
12. Duewel, H. S.; Honek, J. F. *J. Protein Chem.* **1998,** 17, (4), 337-350.
13. http://expasy.org/tools/peptide-mass.html, In 2010.
14. Pelton, J. T.; McLean, L. R. *Anal. Biochem.* **2000,** 277, 167-176.
15. Gaussier, H.; Morency, H.; Lavoie, M. C.; Subirade, M. *Appl. Environ. Microbiol.* **2002,** 68, (10), 4803-4808.
16. Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *Biochemistry* **2004,** 43, (22), 6893-6898.
17. Huth, J. R.; Bewley, C. A.; Jackson, B. M.; Hinnebusch, A. G.; Clore, G. M.; Gronenborn, A. M. *Protein Science* **1997,** 6, (11), 2359-2364.
18. Bradford, M. M. *Anal. Biochem.* **1976,** 72, (1-2), 248-254.
19. Choikhet, K.; Glatz, B.; Rozing, G. *LC GC Europe* **2003,** 16.
20. Williams, S. *Journal of Chromatography A* **2004,** 1052, (1-2), 1-11.