

# 3-Mercaptopropionic acid-mediated synthesis of peptide and protein thioesters†

Jaskiranjit Kang, Jonathan P. Richardson and Derek Macmillan\*

Received (in Cambridge, UK) 10th September 2008, Accepted 10th October 2008

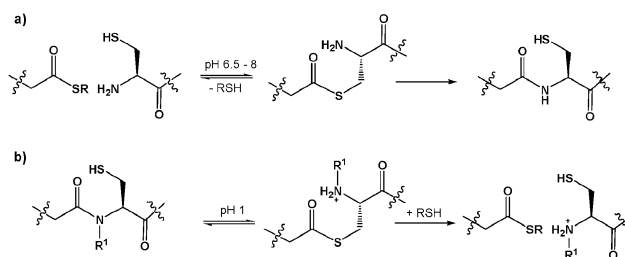
First published as an Advance Article on the web 5th November 2008

DOI: 10.1039/b815888f

**Peptides and proteins fragment sequence-specifically in the presence of 3-mercaptopropionic acid to afford thioesters which can be used in native chemical ligation reactions.**

Native chemical ligation (NCL) is an extremely useful method for the production of synthetic and semi-synthetic proteins.<sup>1</sup> Since it is widely accepted that approximately fifty amino acid residues represents the limit for efficient automated solid phase peptide synthesis (SPPS) the use of recombinant methods for the production of the required thioester and cysteine-containing components has been widely exploited.<sup>2</sup> However, while the number of available synthetic methods for the production of the thioester component has risen dramatically,<sup>3</sup> micro-organism-derived thioesters are most frequently generated by the commercially available intein-fusion expression system.<sup>4</sup> Recently, several variations on a central theme of thioester production *via* an *N* to *S* acyl shift, that are mechanistically more similar to the intein system, have been reported (Scheme 1).<sup>5</sup>

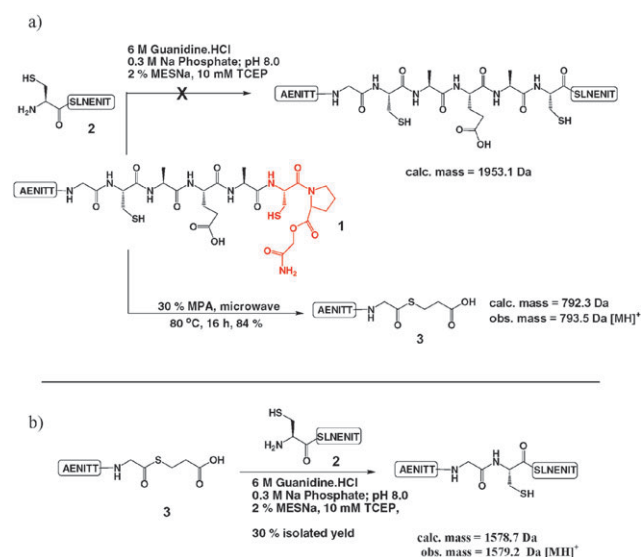
Following the work of Kawakami and Aimoto<sup>5d</sup> on the use of cysteinylprolyl esters (CPEs) for the transient production of thioesters which can participate in NCL reactions, we had attempted to utilise CPEs in our studies geared towards the semi-synthesis of erythropoietin (EPO) and had managed to apply them with limited success (Scheme 2(a)).<sup>6</sup> Consequently, we attempted to “rescue” CPE **1** employing 3-mercaptopropionic acid (MPA) to produce thioesters *via* a *N* → *S* acyl shift in aqueous MPA as reported by Hojo and co-workers.<sup>5b</sup> In doing so we observed formation of an undesired glycine thioester **3** at the expense of our desired alanine thioester. Interestingly the “isolated” cysteine residue appeared to facilitate conversion to a glycine-MPA thioester in the absence of any additional stimulus such as a CPE, a mercaptomethyl proline residue,<sup>5b</sup> or cysteine *N*-alkylation<sup>5e</sup> and led us to predict that some selectivity may be achievable in peptide and protein fragmentation leading to thioesters in a single step, using MPA as the sole thioesterification reagent.<sup>7</sup> The desired alanine thioester was observed in trace quantities only and although it was possible that any Ala thioester could be further processed to the Gly-thioester, it was not yet clear if this was the case. To validate that the glycine thioester **3** had



**Scheme 1** (a) NCL, and (b) acid mediated *N* → *S* thioester formation compared. R<sup>1</sup> is usually an acyl-transfer facilitating group such as alkyl, benzyl, or δ-mercaptopmethyl prolyl.<sup>5</sup>

indeed formed it was isolated by semi-preparative HPLC and successfully ligated to model peptide **2** (Scheme 2(b)).<sup>8</sup>

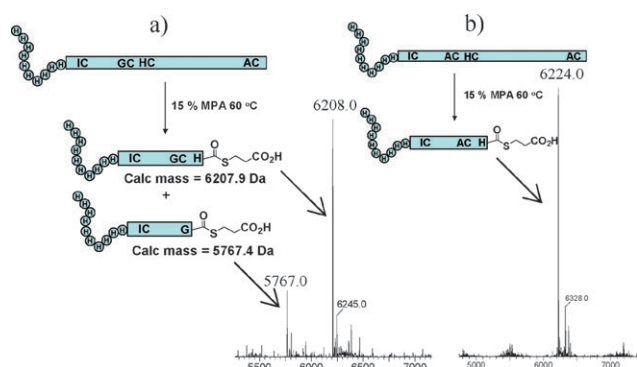
To determine whether or not MPA could effect this transformation more widely we investigated the MPA mediated fragmentation of a recombinant 21 kDa protein sample, His<sub>10</sub>-WT human erythropoietin (EPO). This 166 residue protein contains four cysteine residues in the arrangement I<sup>7</sup>C, G<sup>29</sup>C, H<sup>33</sup>C and A<sup>161</sup>C and was treated with 20% MPA over time periods of 1–48 h and at temperatures ranging from 40 to 80 °C. The protein appeared to fragment in 20% v/v MPA at temperatures as low as 40 °C (Fig. 1(a)) and surprisingly there appeared to be an accumulation of the His<sub>10</sub>-EPO(1-28)-Gly-MPA thioester and His<sub>10</sub>-EPO(1-32)His-MPA thioester



**Scheme 2** (a) Attempted CPE-mediated ligation between peptide **1** (complete sequence: AENITTCGAEAC-CPE) and peptide **2** (sequence: CSLNENIT) and subsequent MPA treatment.<sup>6</sup> (b) Successful ligation reaction employing the Gly-thioester **3**.<sup>8</sup>

Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ. E-mail: d.macmillan@ucl.ac.uk; Tel: 020-7679 4684

† Electronic supplementary information (ESI) available: Experimental procedures for the production of **1**, all model peptides and proteins and their reactions with MPA including HPLC, LC-MS and NMR characterisation of selected products. See DOI: 10.1039/b815888f



**Fig. 1** (a) MPA-mediated fragmentation of His<sub>10</sub>-WT human EPO with 20% MPA for 16 h at 60 °C gives thioesters corresponding to cleavage between GC and HC junctions but not between IC or AC. (b) A G28A mutation appears sufficient to suppress mixed thioester formation.

peptides, with the His-thioester peptide appearing the most abundant. The peptides corresponding to cleavage at the I<sup>7</sup>C junction or between A<sup>161</sup>C were not observed. While it proved difficult to separate this pair of thioesters, the mutation G28A was sufficient to suppress thioesterification at this site (Fig. 1(b)) and the His<sub>10</sub>EPO-1-32(G28A)-MPA thioester was isolated in 28% yield.

The presumed peptide thioesters were isolated by reverse-phase HPLC and also successfully ligated to model peptide **2** in the presence of 6 M guanidine-HCl, 300 mM Na phosphate buffer; pH 7.4, 25 mM 4-mercaptophenylacetic acid (MPAA),<sup>9</sup> 20 mM TCEP to afford the expected NCL products.<sup>8</sup> This experiment further suggested that MPA facilitated protein fragmentation at cysteine residues and that the relative degree of fragmentation may vary depending on the nature of the amino acid adjacent to the cysteine residue. Combining these results led us to postulate an approximate reactivity ranking of H ≥ G > A > I, a trend similar to that observed in the reactivity of thioesters in regular NCL reactions, the reverse process.<sup>10</sup> Consequently we predicted that useful thioesterification may be limited to NCL's favoured three residues (Gly, His, Cys).

Seven peptides (Table 1) were prepared to probe this hypothesis in more detail. Each peptide was related to the recombinant EPO sample that had undergone fragmentation

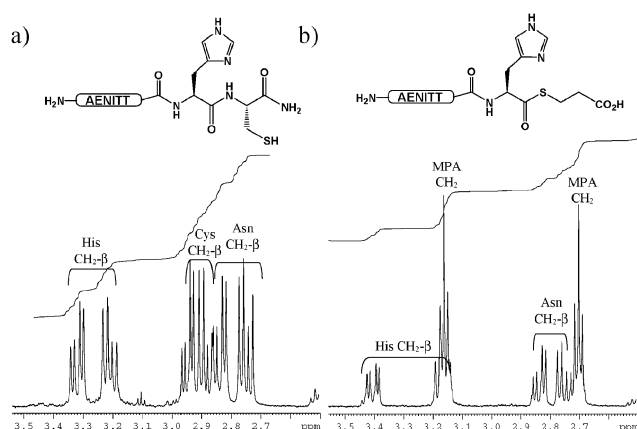
and contained two possible positions at which fragmentation could take place, though in each case the amino acid adjacent to the cysteine residue had been altered such that the outcome of the cleavage reaction could be predicted. In keeping with the observed reactivity profile of peptide thioesters in native chemical ligation reactions we allowed Gly-Cys, Cys-Cys, Ile-Cys and His-Cys "pairs" to compete with each other. Gly, Cys, and His thioesters are known to participate in native chemical ligation reactions with particularly fast kinetics whereas Ile is among the slower thioesters to react.<sup>10</sup> Consequently, we expected that peptides which contained exclusively GC junctions or a GC junction combined with CC or HC junctions to yield mixtures of thioesters whereas peptides that contained IC or protected GC(S-Acm) to yield singly-cleaved compounds and this was indeed found to be the case. Thioester formation at HC junctions appeared to be the most rapid and most extensive followed by cleavage at GC and CC junctions. Interestingly the peptide containing the GC and HC pair (entry 1) appeared to fragment in a very similar manner to the recombinant protein fragment providing a mixture of His and Gly thioesters.

Having demonstrated that the thioesterification reaction appeared to be selective, we conducted another small screen of test peptides to probe for particularly favourable or unfavourable thioesterification sites.<sup>8</sup> Ten short peptides of sequence H-AENITTXC-NH<sub>2</sub> (where X = A, D, E, F, G, K, P, S, V, W) were cleaved from the solid support. LC-MS of the crude products indicated that the peptides were highly pure though in some cases (with the exception of FC, PC, KC, WC and VC containing peptides) two resolved peaks with identical masses were observed which were tentatively assigned as the *N*- and the *S*- peptide.<sup>5f</sup> Each crude sample was isolated in near quantitative yield so each peptide was used directly in MPA-mediated thioesterification reactions. The peptides were exposed to forcing conditions of 20% MPA at 80 °C for 24 h, anticipating that Gly-thioester formation should be near complete and Val-thioester formation negligible. Analysis of the reaction mixtures after 20 h revealed that, as expected, Gly-thioester formation had progressed to near completion whereas Pro and Val showed negligible, though detectable, thioester formation. Notably the Asp-Cys containing peptide had been completely consumed yet did not give rise to a thioester product, the major species corresponding to a product which had lost both the Asp and Cys residues. In contrast

**Table 1** Peptides designed to test for selectivity in MPA-mediated fragmentation

Entry	Peptide sequence <sup>a</sup>	X <sup>8</sup> C cleavage <sup>b</sup>		X <sup>12</sup> C cleavage <sup>b</sup>		X <sup>12</sup> C : X <sup>8</sup> C <sup>c</sup>	Isolated yield <sup>d</sup> (%)
		Calc. <i>m/z</i>	Obs. <i>m/z</i>	Calc. <i>m/z</i>	Obs. <i>m/z</i>		
1	H-AENITTGCAEHC-NH <sub>2</sub>	793.3	793.5	1233.5	1233.6	≈ 1 : 1	28
2	H-AENITTGC(Acm)AEHC-NH <sub>2</sub>	793.3	n/o	1304.5	1304.6	> 9 : 1	39
3	H-AENITTICAEHC-NH <sub>2</sub>	849.4	n/o	1289.5	1289.7	> 9 : 1	60
4	H-AENITTICAEGC-NH <sub>2</sub>	849.4	n/o	1209.5	1209.7	> 9 : 1	n.d.
5	H-AENITTGCAEGC-NH <sub>2</sub>	793.3	793.5	1153.4	1153.6	≈ 1 : 1	26
6	H-AENITTGCAECC-NH <sub>2</sub>	793.3	793.5	1199.4	1199.6	≈ 1 : 1	28
7	H-AENITTGCAEIC-NH <sub>2</sub>	793.3	793.5	1209.5	n/o	1 : 9	n.d.

<sup>a</sup> Peptides prepared on Rink amide MBHA resin. <sup>b</sup> Each 50 μL reaction contained approx. 8 mM peptide and was analysed by LC-MS after 36 h at 50 °C (unoptimised) in 20% MPA. The calculated mass corresponds to the MPA thioester. <sup>c</sup> Determined by LC-MS. <sup>d</sup> Unoptimised yield of the full length (11 mer) thioester. n.d. = not determined and applies to peptides where isolation was hindered by overlapping MPA derived peaks.



**Fig. 2**  $^1\text{H}$  NMR analysis of test peptide H-AENITTHC-NH<sub>2</sub> (a) before and (b) after MPA-mediated thioesterification.

the Glu-Cys containing peptide (like the Ala-Cys, Ser-Cys, and Phe-Cys containing peptides) gave rise to an approximately 50 : 50 mix of starting material and product. The Ser-Cys containing peptide (both the starting material and the product) appeared to have further condensed with MPA under the reaction conditions, presumably on the primary hydroxyl group of serine, to afford an MPA ester. This modification of serine, along with Ser-MPA thioester formation, was negligible when the reaction was repeated at 60 °C for 16 h with the vast majority of the starting material remaining unchanged. Besides Val and Pro, Lys and Trp appeared to have performed least efficiently in thioester formation at 80 °C. From these results we concluded that, of the nine additional fragmentation junctions tested, only Gly-Cys junctions and His-Cys are likely to give rise to thioesters under more desirable conditions (temperatures ranging from 50 to 60 °C). The model experiments with test peptides also provided access to multimilligram quantities of products allowing  $^1\text{H}$  and  $^{13}\text{C}$  NMR analysis of the proposed thioester products. The  $^1\text{H}$  NMR spectrum of H-AENITTHC-NH<sub>2</sub> and H-AENITTGC-NH<sub>2</sub> before and after exposure to MPA clearly indicate the loss of cysteine, the incorporation of MPA (Fig. 2) along with the associated downfield shift in the resonance of the  $\alpha$ -hydrogens of the C-terminal residue.<sup>8</sup> Thioester formation was further supported by the appearance of a signal at 205 ppm, typical for the carbonyl carbon of a thioester, in the  $^{13}\text{C}$  NMR spectrum of H-AENITTHC-NH<sub>2</sub> after treatment with MPA.  $^{13}\text{C}$  labelling studies are currently underway to further confirm thioester formation.

In summary, through our investigations employing CPE terminated peptides, we observed an interesting one-step thioesterification reaction in which peptide and protein samples containing GC, CC and HC junctions preferentially fragment at these positions in the presence of MPA.<sup>11</sup> While fragmentation apparently occurs most readily at histidine the extent of racemisation needs detailed investigation though a short thioester, Ac-Ala-His-SCH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H derived from peptide Ac-Ala-(L)-His-Cys-NH<sub>2</sub> upon exposure to 20% MPA for 48 h at 55 °C surprisingly exhibits undetectable racemisation when compared (by HPLC) with Ac-Ala-(D)-His-Cys-NH<sub>2</sub> exposed to identical conditions.<sup>8</sup> Optimisation

of the reaction conditions and exploration of alternative experimental protocols, e.g. use of microwave irradiation, in combination with site-directed mutagenesis, may ultimately yield a direct, small molecule-mediated, method for the production of bacterially expressed protein thioesters to complement the intein-fusion system.

The authors acknowledge financial support from The Wellcome Trust and the EPSRC.

## Notes and references

- (a) P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**(5186), 776; (b) P. E. Dawson and S. B. H. Kent, *Annu. Rev. Biochem.*, 2000, **69**, 923; (c) D. Macmillan, *Angew. Chem., Int. Ed.*, 2006, **45**(46), 7668; (d) C. Haase and O. Seitz, *Angew. Chem., Int. Ed.*, 2008, **47**(9), 1553.
- (a) V. Muralidharan and T. W. Muir, *Nat. Methods*, 2006, **3**(6), 429; (b) T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. USA*, 1998, **95**(12), 6705; (c) C. J. Noren, J. M. Wang and F. B. Perler, *Angew. Chem., Int. Ed.*, 2000, **39**(3), 451; (d) R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder and T. W. Muir, *Nature*, 2008, **453**(7196), 812.
- (a) Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman and C. R. Bertozzi, *J. Am. Chem. Soc.*, 1999, **121**(50), 11684; (b) R. Ingenito, E. Bianchi, D. Fattori and A. Pessi, *J. Am. Chem. Soc.*, 1999, **121**(49), 11369; (c) D. Swinnen and D. Hilvert, *Org. Lett.*, 2000, **2**(16), 2439; (d) H. Hojo, E. Haginoya, Y. Matsumoto, Y. Nakahara, K. Nabeshima, B. P. Toole and Y. Watanabe, *Tetrahedron Lett.*, 2003, **44**(14), 2961; (e) J. A. Camarero, B. J. Hackel, J. J. D. Yoreo and A. R. Mitchell, *J. Org. Chem.*, 2004, **69**(12), 4145; (f) O. S. Franziska Mende, *Angew. Chem., Int. Ed.*, 2007, **46**(24), 4577; (g) S. Manabe, T. Sugioka and Y. Ito, *Tetrahedron Lett.*, 2007, **48**(5), 849; (h) T. J. Hogenauer, Q. Wang, A. K. Sanki, A. J. Gammon, C. H. L. Chu, C. M. Kaneshiro, Y. Kajihara and K. Michael, *Org. Biomol. Chem.*, 2007, **5**, 759; (i) S. Ficht, R. J. Payne, R. T. Guy and C.-H. Wong, *Chem. Eur. J.*, 2008, **14**(12), 3620; (j) J. B. Blanco-Canosa and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2008, **47**, 6851.
- www.neb.com.
- (a) N. Ollivier, J.-B. Behr, Q. El-Mahdi, A. Blanpain and O. Melnyk, *Org. Lett.*, 2005, **7**(13), 2647; (b) F. Nagaike, Y. Onuma, C. Kanazawa, H. Hojo, A. Ueki, Y. Nakahara and Y. Nakahara, *Org. Lett.*, 2006, **8**(20), 4465; (c) H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara and Y. Nakahara, *Tetrahedron Lett.*, 2007, **48**(1), 25; (d) T. Kawakami and S. Aimoto, *Tetrahedron Lett.*, 2007, **48**(11), 1903; (e) H. Hojo, Y. Murasawa, H. Katayama, T. Ohira, Y. Nakahara and Y. Nakahara, *Org. Biomol. Chem.*, 2008, **6**, 1808; (f) K. Nakamura, M. Sumida, T. Kawakami, T. Vorherr and S. Aimoto, *Bull. Chem. Soc. Jpn.*, 2006, **79**, 1773; (g) B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2006, **45**(25), 4116.
- Although the reaction was not successful, when the CPE peptide was later re-prepared with the cysteine residue adjacent to glycine protected with an S-Acm group the desired CPE ligation product was observed and the Ala-MPA thioester could be isolated (after treatment of the CPE peptide with MPA).
- N→S acyl transfer has been observed in a synthetic peptide containing a Gly-Cys junction, in ref. 5f, upon exposure to 71% TFA in CDCl<sub>3</sub> as part of a two-step thioesterification procedure. The glycine thioester so produced was not isolated or used in ligation reactions due to difficulties encountered as a consequence of the two-step procedure adopted, (the S-peptide spontaneously reverted to the N-peptide upon attempted isolation). Neither the potential selectivity of the thioesterification, nor application to recombinant samples was subsequently explored.
- See ESI† for experimental details.
- E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**(20), 6640.
- T. M. Hackeng, J. H. Griffin and P. E. Dawson, *Proc. Natl. Acad. Sci. USA*, 1999, **96**(18), 10068.
- L. Zhang and J. P. Tam, *Tetrahedron Lett.*, 1997, **38**, 3.