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. 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. However, while the number of available synthetic methods for the production of the thioester component has risen dramatically, micro-organism-derived thioesters are most frequently generated by the commercially available intein-fusion expression system. Recently, several variations on a central theme of thioester production via an N → S acyl shift, that are mechanistically more similar to the intein system, have been reported (Scheme 1).

Following the work of Kawakami and Aimoto 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). This 166 residue protein contains four cysteine residues in the arrangement I7C, G29C, H33C and A161C 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 His10-EPO(1-28)-Gly-MPA thioester and His10-EPO(1-32)His-MPA thioester

Scheme 1 (a) NCL, and (b) acid mediated N → S thioester formation compared. R1 is usually an acyl-transfer facilitating group such as alkyl, benzyl, or δ-mercaptopropyl.5

indeed formed it was isolated by semi-preparative HPLC and successfully ligated to model peptide 2 (Scheme 2(b)).

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, His10-WT human erythropoietin (EPO). This 166 residue protein contains four cysteine residues in the arrangement I7C, G29C, H33C and A161C 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 His10-EPO(1-28)-Gly-MPA thioester and His10-EPO(1-32)His-MPA thioester

Scheme 2 (a) Attempted CPE-mediated ligation between peptide 1 (complete sequence: AENITTGCAEA-CPE) and peptide 2 (sequence: CSLNENIT) and subsequent MPA treatment.5 (b) Successful ligation reaction employing the Gly-thioester 3.
peptides, with the His-thioester peptide appearing the most abundant. The peptides corresponding to cleavage at the I1C junction or between A161C 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 His10EPO-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), 20 mM TCEP to afford the expected NCL products. 8 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. 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. 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. Ten short peptides of sequence H-AENITTXC-NH2 (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. 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>
<thead>
<tr>
<th>Entry</th>
<th>Peptide sequence&lt;sup&gt;b&lt;/sup&gt;</th>
<th>X&lt;sup&gt;13&lt;/sup&gt;C cleavage&lt;sup&gt;b&lt;/sup&gt;</th>
<th>X&lt;sup&gt;12&lt;/sup&gt;C cleavage&lt;sup&gt;b&lt;/sup&gt;</th>
<th>X&lt;sup&gt;13&lt;/sup&gt;C : X&lt;sup&gt;12&lt;/sup&gt;C&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Isolated yield&lt;sup&gt;d&lt;/sup&gt; (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H-AENITTTGCAHCH-NH2</td>
<td>793.3</td>
<td>793.5</td>
<td>1233.5</td>
<td>1233.6</td>
</tr>
<tr>
<td>2</td>
<td>H-AENITTTGCAACHC-NH2</td>
<td>793.3</td>
<td>n/o</td>
<td>1304.5</td>
<td>1304.6</td>
</tr>
<tr>
<td>3</td>
<td>H-AENITTTCAECNH2</td>
<td>849.4</td>
<td>n/o</td>
<td>1289.5</td>
<td>1289.7</td>
</tr>
<tr>
<td>4</td>
<td>H-AENITTTCAECGC-NH2</td>
<td>849.4</td>
<td>n/o</td>
<td>1209.5</td>
<td>1209.7</td>
</tr>
<tr>
<td>5</td>
<td>H-AENITTTCAEFGC-NH2</td>
<td>793.3</td>
<td>793.5</td>
<td>1153.4</td>
<td>1153.6</td>
</tr>
<tr>
<td>6</td>
<td>H-AENITTTCAEGGC-NH2</td>
<td>793.3</td>
<td>793.5</td>
<td>1199.4</td>
<td>1199.6</td>
</tr>
<tr>
<td>7</td>
<td>H-AENITTTCAEGCC-NH2</td>
<td>793.3</td>
<td>793.5</td>
<td>1209.5</td>
<td>n/o</td>
</tr>
</tbody>
</table>

<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 thioester peptide. <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.
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


6 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-acetyl group the desired CPE ligation product was observed and the Ala-MPA thioester could be isolated (after treatment of the CPE peptide with MPA).

7 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 CDCl3 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.

8 See ESi† for experimental details.

