JOC*Article*

A Cassette Ligation Strategy with Thioether Replacement of Three Gly-Gly Peptide Bonds: Total Chemical Synthesis of the 101 Residue Protein Early Pregnancy Factor [$\psi(CH_2S)^{28-29,56-57,76-77}$]

Darren R. Englebretsen,*,† Bronwyn Garnham, and Paul F. Alewood*

Centre for Drug Design and Development, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia

p.alewood@mail.imb.uq.edu.au

Received September 15, 2001

The 101 residue protein "early pregnancy factor" (EPF), also known as human chaperonin 10, was synthesized from four functionalized, but unprotected, peptide segments by a sequential thioether ligation strategy. The approach exploits the differential reactivity of a peptide-$NHCH_2CH_2SH$ thiolate with XCH_2CO-peptides, where X = Cl or I/Br. Initial model studies with short functionalized (but unprotected) peptides showed a significantly faster reaction of a peptide-$NHCH_2CH_2SH$ thiolate with a $BrCH_2CO$-peptide than with a $ClCH_2CO$-peptide, where thiolate displacement of the halide leads to chemoselective formation of a thioether surrogate for the Gly-Gly peptide bond. This rate difference was used as the basis of a novel sequential ligation approach to the synthesis of large polypeptide chains. Thus, ligation of a model bifunctional N^{α}-chloroacetyl, C-terminal thiolated peptide with a second N^{α}-bromoacetyl peptide demonstrated chemoselective bromide displacement by the thiol group. Further investigations showed that the relatively unreactive N^{α}-chloroacetyl peptides could be "activated" by halide exchange using saturated KI solutions to yield the highly reactive N^{α}-iodoacetyl peptides. These findings were used to formulate a sequential thioether ligation strategy for the synthesis of EPF, a 101 amino acid protein containing three Gly-Gly sites approximately equidistantly spaced within the peptide chain. Four peptide segments or "cassettes" comprising the EPF protein sequence (BrAc-[EPF 78−101] **12**, ClAc-[EPF 58−75]-[$NHCH_2CH_2$-SH] **13**, ClAc-[EPF 30−55]-[$NHCH_2CH_2SH$] **14**, and Ac-[EPF 1−27]-[$NHCH_2CH_2SH$] **15**) of EPF were synthesized in high yield and purity using Boc SPPS chemistry. In the stepwise sequential ligation strategy, reaction of peptides **12** and **13** was followed by conversion of the N-terminal chloroacetyl functional group to an iodoacetyl, thus activating the product peptide for further ligation with peptide **14**. The process of ligation followed by iodoacetyl activation was repeated to yield an analogue of EPF (EPF $\psi(CH_2S)^{28-29,56-57,76-77}$) **19** in 19% overall yield.

Introduction

In past years solid-phase,[1−5] solution-phase,[6−9] and convergent solid-phase[10,11] chemical peptide synthesis techniques have been used to synthesize reliably highly pure, fully biologically active proteins.[12,13] Despite these successes, chemical protein synthesis methods remain limited in the length of protein that may be obtained in high yield and purity. In stepwise solid-phase syntheses of longer peptides, the yields have invariably been low due to the difficulties inherent in separating the target protein from a large number of resin-bound peptide byproducts formed during chain assembly and from side products generated in the final cleavage step(s). Convergent synthesis strategies require coupling of fully protected peptide segments. While these strategies are often successful, they require high levels of expertise, are slow, and can be complicated by unforseen difficulties in purification and coupling of the poorly soluble fully protected peptide segments.[10]

In addition to established chemical protein synthesis techniques, the recently described chemical ligation technologies represent an important advance in chemical protein synthesis.[14−19] In its simplest form, a chemical

† Present address: Chemistry Department, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.

(1) Gutte, B.; Merrifield, R. B. *J. Biol. Chem.* **1971**, *246*, 1922−1941.
(2) Clark-Lewis, I.; Aebersold, R.; Ziltener, H.; Schrader, J. W.; Hood, L. E.; Kent, S. B. H. *Science* **1986**, *231*, 134−139.
(3) Kent, S. B. H. *Annu. Rev. Biochem.* **1988**, *57*, 957−989.
(4) Schneider, J.; Kent, S. B. H. *Cell* **1988**, *54*, 363−368.
(5) Kent, S. B. H.; Alewood, D.; Alewood, P.; Baca, M.; Jones, A.; Schnolzer, M. In *Innovations and Perspectives in Solid-Phase Peptide Synthesis*; Epton, R., Ed.; Intercept Ltd.: Andover, 1992; pp 1−22.
(6) Yakima, H.; Fujii, N. *J. Am. Chem. Soc.* **1981**, *103*, 5867−5871.
(7) Sakakibara, S. *Bioploymers (Pept. Sci.)* **1995**, *37*, 17−28.
(8) Nishiuchi, Y.; Inui, T.; Nishio, H.; Bódi, J.; Kimura, T.; Tsuji, F. I.; Sakakibara, S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13549−13554.
(9) Quibell, M.; Packman, L. C.; Johnson, T. *J. Am. Chem. Soc.* **1995**, *117*, 11656−11668.
(10) Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Tetrahedron* **1993**, *49*, 11065−11133.
(11) Benz, H. *Synthesis* **1994**, 337−358.
(12) Wilken, J.; Kent, S. B. H. *Curr. Opin. Biotechnol.* **1998**, *9*, 412−426.
(13) Borgia, J. A.; Fields, G. B. *Tibtech* **2000**, *18*, 243−251.
(14) Schnolzer, M.; Kent. S. B. H. *Science* **1992**, *256*, 221−225.
(15) Baca, M.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11638−11642.

10.1021/jo016121e CCC: $22.00
Published on Web 02/23/2002

ligation strategy requires the assembly of two peptide segments which have chemoselective, mutually reactive functional groups at the appropriate N- and/or C-termini. The specific reaction of these N- and C-terminal functionalities results in covalent coupling of the two peptide segments, giving a product[14] which has the "ligation site" peptide bond replaced by an amide bond isostere or surrogate. There are a number of benefits of methods for protein synthesis by chemoselective ligation. The unprotected peptide segments, prepared by efficient SPPS techniques, may be purified quickly by standard RP-HPLC techniques. The high specificity of the ligation reaction between the N- and C-terminal functional groups means that side chain protection is unnecessary, simplifying synthesis and purification of the peptide segments. An added advantage of using unprotected peptides is that high concentrations, in denaturing aqueous solutions, of both reactants can be employed, thus enhancing the rate of the bimolecular ligation reaction.[15] The highly specific chemoselective ligation reaction usually goes to completion and results in few or no unwanted side products.[15] Finally, chemical ligation is readily undertaken in aqueous or aqueous−organic solutions, which allows convenient monitoring by RP-HPLC and LC-MS analysis.

Protein synthesis by chemical ligation was first exemplified by the synthesis of a fully active HIV-1 protease analogue[15] by thioester-forming ligation. A feature of the thioester amide bond surrogate is that it may be used as a probe to study the role of backbone amide bonds in the activity of the protein. For example, enzymatic characterization of a "chemically mutated" HIV-1 protease analogue demonstrated the crucial importance of the hydrogen bonding contribution from the amide NH of Ile50 in promoting the activity of HIV-1 protease, a result which was only obtainable by studying a protein synthesized using ligation chemistry.[16] We have also described the synthesis of HIV-1 protease analogues by both thioether and thioester ligation strategies.[17] In this work significant differences in the activity of the various analogues were found and were related to graduated disruption of the substrate-binding ability of the ligated analogues compared to the native protein.[17] The choice of the thioether surrogate has been pursued in our work[17] because of the superior chemical stability of the thioether bond compared to that of the comparatively base-labile thioester.[15]

Other ligation chemistries have been described. A particularly elegant example is native chemical ligation, where two peptide fragments, one bearing a C-terminal glycine thioester and the other an N-terminal cysteine residue, react to form a protein containing a native amide bond at the segment junction.[17,18] Native ligation has been extended recently to X-Cys sites, where X can be any of the 20 natural amino acids, although Val, Ile, and Pro show slow ligation rates.[20] A more general version of native chemical ligation which extends to sites other than X-Cys has been reported.[21] Native chemical ligation is also used in the solid-phase chemical ligation method.[19]

Native chemical ligation has been recently extended beyond chemical synthesis by the development of expressed protein ligation, in which recombinant proteins bearing C-terminal thioester functional groups are used in subsequent native ligation syntheses.[22,23] Other methods of chemical ligation have been used to synthesize active HIV-1 protease analogues containing a pseudoproline linkage,[24,25] and peptide dendrimers containing thiazolidine, oxime, and hydrazone linkages.[26]

Chemical protein synthesis via ligation strategies is limited by the current chemistry employed in stepwise solid-phase peptide synthesis. Peptide segments up to ∼60 amino acids may be synthesized in high yield and purity,[3] thus imposing a limit of about 120 amino acids on the length of proteins resulting from ligation of two segments. We are interested in further exploiting and developing solution-phase *sequential ligation strategies* to enable the assembly of longer proteins (e.g., 100−200 amino acids). In this report we describe the development of a sequential ligation method for protein synthesis based on chemoselective thioether formation. The fundamental principles were developed through studies of small peptides, and the application of this strategy is demonstrated by the synthesis of an analogue of the 101 amino acid protein "early pregnancy factor" EPF [$\psi(CH_2S)^{28-29,56-57,76-77}$], also known as human chaperonin 10.[27,28]

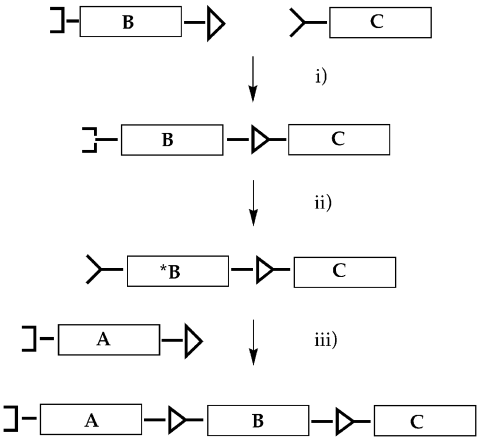
Results and Discussion

Concept and Strategy. In the sequential peptide ligation strategy a target protein sequence is divided into a series of readily synthesized peptide segments or cassettes where each suitably functionalized peptide has the potential to be "inserted" regioselectively into the synthetic scheme (Scheme 1). Each segment (except those at the N- and C-termini of the target protein) has at its C-terminus a defined chemoselective functional group, while at its N-terminus there is a masked functional group which can be readily activated.[56]

The first step of protein synthesis is chemoselective reaction of the functional groups at the C- and N-termini of segments **B** and **C**, respectively, to give product **BC**, with the N-terminus of segment **B** incorporating a masked functional group. The masked N-terminal func-

(16) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
(17) (a) Englebretsen, D. R.; Bergman, D.A.; Garnham, B. G.; Brinkworth, R. I.; Fairlie, D. P.; Alewood, P. F. In *Proceedings of the 14th American Peptide Symposium*; Kaumaya, P., Ed.; Mayflower: Kingswinford, 1995; pp 589−590. (b) Wilce, J.A.; Love, S. G.; Richardson, J. S.; Alewood, P. F.; Craik, D. J. *J. Biol. Chem.* **2001**, *276*, 25997−26003.
(18) Muir, T. W.; Dawson, P. E.; Kent, S. B. H. In *Methods in Enzymology*; Fields, G. B., Eds.; Academic Press: New York, 1997; pp 266−298.
(19) Canne, L. E.; Botti, P.; Simon, R. J.; Chen, Y.; Dennis, E. A.; Kent, S. B. H. *J. Am. Chem. Soc.* **1999**, *121*, 8720−8727.
(20) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068−10073.
(21) Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891−5896.
(22) Severinov, K.; Muir, T. W. *J. Biol. Chem.* **1998**, *273*, 16205−16209.
(23) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710.
(24) Liu, C.-F.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 7−312.
(25) Tam, J. P.; Yu, Q.; Mioa, Z. *Bioploymers (Pept. Sci.)* **1999**, *51*, 311−332.
(26) Shao, J.; Tam, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3893−3899.
(27) Cavanagh, A. C.; Morton, H. *Eur. J. Biochem.* **1994**, *222*, 551−560.
(28) Monzini, N.; Legname, G.; Marcucci, F.; Gromo, G.; Modena, D. *Biochim. Biophys. Acta* **1994**, *1218*, 478−480.

SCHEME 1. General Principle of Protein Synthesis by Sequential Ligation[a]



[a] (i) Chemoselective reaction of the functional groups at the C- and N-termini of cassettes B and C, respectively, with cassette B also incorporating a masked functional group at its N-terminus. (ii) The masked N-terminal functional group of the resulting ligated product, BC, is transformed to the activated intermediate *BC.(iii) Ligation of the N, C terminal bifunctional cassette peptide A with activated peptide *BC yielding the peptide ABC which contains two ligation sites and a masked functional group at its N-terminus. Further repetition of the process results in elongation of the protein chain.

tional group of **BC** is then transformed to the activated intermediate ***BC**. Subsequent ligation of the bifunctional cassette **A** with activated peptide ***BC** yields the peptide **ABC** which contains two ligation sites and a masked functional group at its N-terminus. Further repetition of the process results in elongation of the protein chain. In this strategy the major requirement for successful sequential ligation is the incorporation of an unreactive but readily transformed masked N-terminal functional group. In this report we describe how this may be readily achieved via thioether ligation chemistry.

Thioether Ligation. Thioether linkages are generated by simple nucleophilic displacement, by a thiol-bearing peptide, of the halogen from the N^{α}-haloacetyl group of a second peptide. Peptide and protein syntheses which incorporate thioether ligation strategies were rendered practical by the discovery that N^{α}-chloro- and bromo-acetylated peptides survive treatment by liquid HF[29−31] used during the cleavage of peptides synthesized via Boc SPPS.[57] N^{α}-Haloacetylated peptides may also be synthesized via Fmoc SPPS.[32] N^{α}-Haloacetylated peptides have been coupled to cysteine-bearing peptides in the synthesis of cyclic peptides and peptide oligomers,[29,30] peptide−protein immunogens,[33] a protein construct,[34] a peptide−porphyrin template complex,[35] a template-assembled synthetic protein,[36] and a designed transmembrane protein.[32] However, the thioether linkages generated in these products were not strictly peptide backbone amide bond surrogates.

Thioether amide bond surrogates were first introduced into short peptide analogues using either preformed pseudodipeptide subunits or by direct formation of a thiomethylene bond using on-resin procedures.[37,38] Previously, we have used solution ligation procedures to synthesize HIV-1 protease and transthyretin analogues containing Gly-ψ(CH$_2$S)-Gly thioether amide bond surrogates.[17a,b] Earlier studies of short protected peptide analogues have demonstrated that the methodology can in principle be extended to ligation sites other than Gly-Gly, shown for example by the synthesis of the hindered Leu-Nle thioether amide surrogate.[38] In this study we describe in detail the conditions for the sequential ligation of C-terminal 2-mercaptoethylamido functionalized peptides and N^{α}-haloacetyl peptides to generate multiple thioether linkages which contain a thioether surrogate of a Gly-Gly peptide bond.

Synthesis of [*N*-(2-Mercaptoethylamido)]-Functionalized Peptides. Peptide segments were assembled on a linker derived from 2-aminoethanethiol in order to introduce a C-terminal *N*-(2-mercaptoethyl)amide moiety.[39] C-Terminal *N*-(2-mercaptoethylamide)-functionalized peptides have been synthesized using linkers generated by on-resin procedures.[40,41] However, in our approach we preferred to introduce a fully characterized linker to a well-behaved aminomethyl-polystyrene resin.[5] A clear parallel exists with Boc aminoacyl-OCH$_2$Pam-polystyrene resins, which incorporate a purified, characterized chemical "handle" or linker onto a well-characterized resin[5,42] and are the resins of choice over in-situ derivatized "Merrifield resins" which retain undesirable resin-bound functional groups. Accordingly, we synthesized the linker 4-[N^{α}-Boc-2-aminoethylmercapto]-methylphenoxyacetic acid (Boc-AMPA **1**) in two steps.[39] The Boc-AMPA linker was subsequently attached to an aminomethyl-polystyrene resin and used to assemble C-terminally 2-mercaptoethylamido functionalized peptides using Boc chemistry with in situ neutralization.[43] Selected peptides were functionalized at their N-terminus using the anhydrides of either chloroacetic and bromoacetic acid. The peptides were liberated from the resin and side chain protecting groups simultaneously removed using standard treatment with HF in the presence of *p*-cresol scavenger. After purification by preparative RP-HPLC, all purified peptides[58] showed single peaks by RP-HPLC analysis, while ES-MS analysis showed the expected molecular ion masses.

(29) Lindner, L.; Robey, F. A. *Int. J. Pept. Protein Res.***1987**, *30*, 794−800.
(30) Robey, F. A.; Fields, R. L. *Anal. Biochem*. **1989**, *177*, 373−377.
(31) Ivanov, B. B.; Robey, F. A. *Pept. Res.* **1996**, *9*, 305−307.
(32) Englebretsen, D. R.; Choma, C. T.; Robillard, G. T. *Tetrahedron Lett*. **1998**, *39*, 4929−4932.
(33) Kolodny, N.; Robey, F. *Anal. Biochem*. **1990**, *187*, 136−140.
(34) Muir, T. W.; Williams, M. J.; Ginsberg, M. H.; Kent, S. B. H. *Biochemistry* **1994**, *33*, 7701−7708.
(35) Choma, C. T.; Kaestle, K.; Akerfeldt, K. S.; Kim, R. K.; Groves, J. T.; DeGrado, W. F. *Tetrahedron Lett*. **1994**, *35*, 6191−6194.
(36) Nefzi, A.; Sun, X.; Mutter, M. *Tetrahedron Lett*. **1995**, *36*, 229−230.
(37) Spatola, A. F.; Darlak, K. *Tetrahedron* **1988**, *44*, 821−833.
(38) Paladino, J.; Thurieau, C.; Morris, A. D.; Kucharczyk, N.; Rouissi, N.; Regoli, D.; Fauchere, J.-L. *Int. J. Pept. Protein Res*. **1993**, *42*, 284−293.
(39) Englebretsen, D. R.; Garnham, B. G.; Bergman, D. A.; Alewood, P. F. *Tetrahedron Lett*. **1995**, *36*, 8871−8874.
(40) Rietman, B. H.; Smulders, R. H. P. H.; Eggen, I. F.; Van Vliet, A.; Van De Werken, G.; Tesser, G. I. *Int. J. Pept. Protein Res*. **1994**, *44*, 199−206.
(41) Winger, T. M.; Ludovice, P. J.; Chaikof, E. L. *Bioconjugate Chem*. **1995**, *6*, 323−326.
(42) Mitchell, A. R.; Kent, S. B. H.; Engelhard, M.; Merrifield, R. B. *J. Org. Chem*. **1978**, *43*, 2845−2852.
(43) Schnölzer, M.; Alewood, P. F.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res*. **1992**, *40*, 180−193.

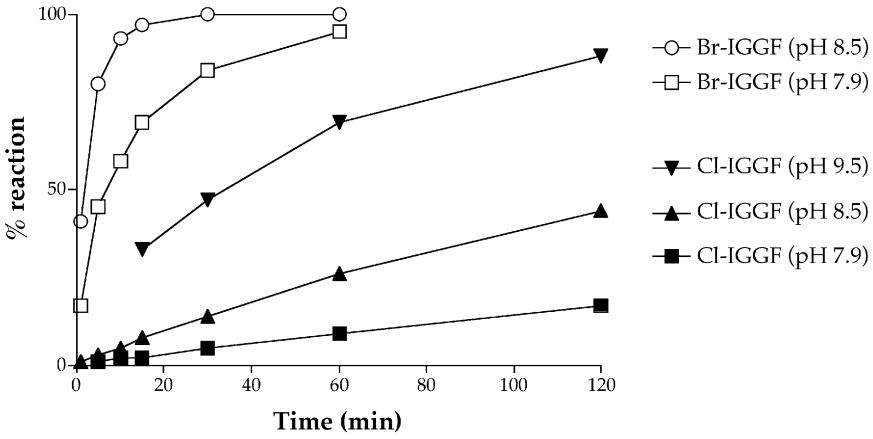


FIGURE 1. Graph of the pH dependence of the reaction of ClAc-IGGF and BrAc-IGGF with LPGKWKPKMI-[NH$(CH_2)_2$-SH]. In both cases the dependence of the rate on pH corresponds to pH dependent increase of ionization of the peptide thiol (pK_a 8.2).

Variables Affecting the Rate of the Thioether Ligation Reaction. Surprisingly, there are few reports of detailed physicochemical studies on the reaction of thiols with haloacetyl compounds in aqueous environments.[44] We thus carried out a short study to determine the effect of pH and the halide leaving group on the rate of the thioether formation by ligating the model peptides LPGKWKPKMI-[NH$(CH_2)_2$-SH] **2** and X-Ac-IGGF (X = Cl, **3**; X = Br, **4**) to give product LPGKWKPKMIG-ψ-(CH_2S)-GIGGF **5**. The peptide sequence chosen for this study, LPGKWKPKMIGGIGGF, HIV-1 PR [38−53], with the GG ligation site underlined, was selected as a model as both the starting peptide segments (**2**, **3**, and **4**) and ligated product **5** were all well resolved and thus conveniently monitored by RP-HPLC analysis. In these experiments a peptide concentration of 2 mM was selected as this represented a typical concentration used for ligating long peptide segments.[14,15,17]

Ligation of LPGKWKPKMI-[NH$(CH_2)_2$-SH] **2** with ClAc-IGGF **3** or BrAc-IGGF **4** at several pH values (Figure 1) shows that, for a given halogen leaving group, the rate of formation of ligation product **5** increased with increasing pH. This result was expected on the basis that the concentration of thiolate anion, the active nucleophilic species, would increase with pH. Figure 1 also shows that at identical pH values (7.9 or 8.5) the initial reaction rate of BrAc-IGGF **4** was approximately an order of magnitude greater than that of ClAc-IGGF **3**.

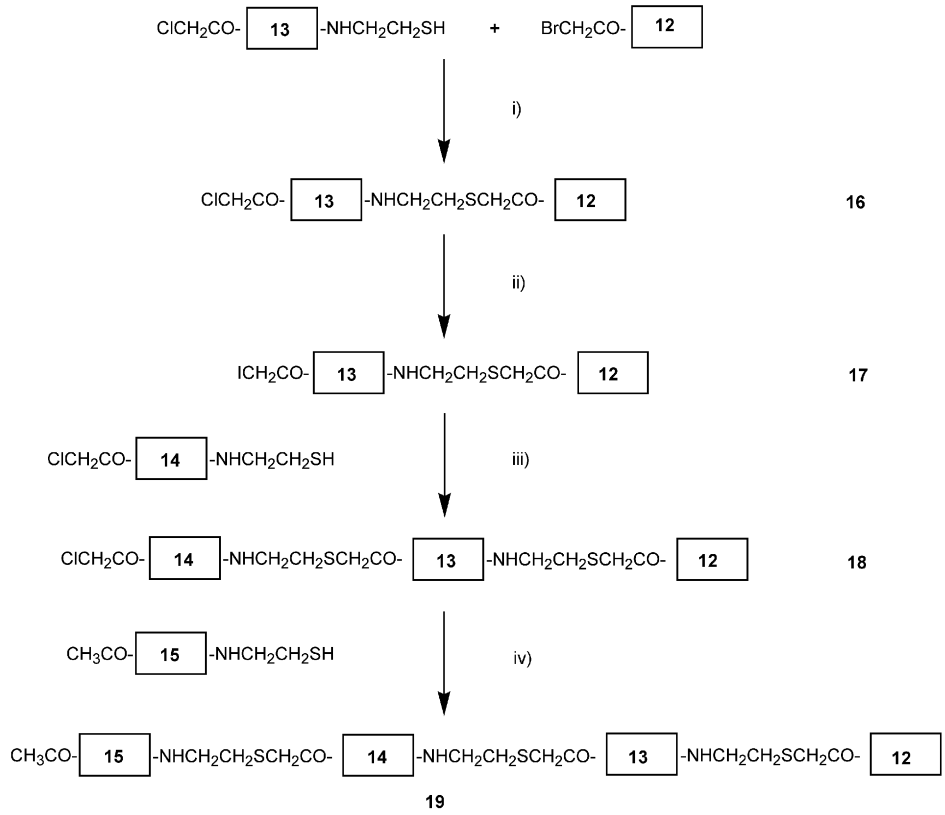
The differences in reaction rates of α-chloro- and α-bromo-acetylated peptides were further explored in a competition experiment at pH 8.5. To distinguish between the relative rate of displacement of chloride versus bromide, ClAc-IGGS **6** (used instead of ClAc-IGGF **3** to improve starting material and product resolution by RP-HPLC analysis) and BrAc-IGGF **4** were reacted with thiolated peptide **2**. The product ratio of 97:3 (by RP-HPLC integration) for LPGKWKPKMIG-ψ(CH_2S)-G-IGGF **5** versus LPGKWKPKMIG-ψ(CH_2S)-GIGGS **7** (derived from bromide and chloride displacement, respectively) showed the high chemoselectivity of thiolated peptide **2** for the N^{α}-bromoacetyl over the N^{α}-chloroacetyl peptide. Unexpectedly, a similar competition study between BrAc-IGGF **4** and IAc-IGGS **8** for **2** at pH values of 8.5, 7.0, and 6.0 showed no significant difference in reactivity toward the thiolated peptide, with approximately equal ratios of **5** and **7** (from bromide and iodide displacement, respectively) being observed (data not shown). Importantly, at pH 7.0 the ligation reaction for both haloacetyl peptides **4** and **8** was complete after 2 h.

Chemoselective Control of Thioether Ligation. To investigate whether chemoselective control was achievable during the ligation of an N^{α}-bromoacetyl peptide with another peptide bearing both an N^{α}-chloroacetyl group and a C-terminal 2-mercaptoethylamido group, ClAc-LPGKWKPKMI-[NH$(CH_2)_2$-SH] **9** was reacted with BrAc-IGGF **4** at pH 8.5. Selective displacement of bromine by thiolate was observed to give ClAc-LPGKWKPKMI-G-ψ(CH_2S)-G-IGGF **10** as the sole product after 1 h. Other possible cyclic or oligomeric products resulting from thiol displacement of chloride were not observed by RP-HPLC and ES-MS analyses. We envisaged that the N^{α}-chloroacetyl group of the ligated product **10** may act as a "masked" functional group under the right conditions. While in principle N^{α}-chloroacetylated peptides could react further with another C-terminal thiol-peptide under more forcing conditions (e.g., higher pH and/or longer reaction time) to give a peptide with two thioether ligation sites, in practice the reaction of chloroacetylated peptides with thiol functions was quite slow, even at pH 9.5 (Figure 1). Accordingly, we investigated altering the reactivity of N^{α}-chloroacetyl peptides by substituting the chlorine atom with either a bromine or iodine atom, and thus increasing the reactivity of the peptide for subsequent ligation with another thiolated peptide.

Activation of ClAc-IGGF 7 by Halide Exchange. The conversion of ClAc-IGGF **3** to IAc-IGGF, **11**, was readily achieved within 90 min using saturated KI solution (~7.5 M) over the pH range 2 to 7. Conversion could also be achieved using saturated KI in 6 M urea. Denaturants, acting as chaotropes, such as urea are often essential if high mM concentrations of peptide segments are to be achieved for chemoselective ligations. ES-MS analysis of RP-HPLC fractions from these experiments also showed that no hydrolysis of the iodoacetyl to a hydroxyacetyl group had occurred. The rate of conversion of ClAc-IGGF **3** to IAc-IGGF **11** was much greater than conversion of **3** to BrAc-IGGF **4** (<20% after 5 h) using either 6 M Gu·HBr or saturated KBr solutions, and

(44) Lindley, H. *Biochem. J.* **1962**, *82*, 418−425.

SCHEME 2. Synthesis of EPF by Sequential Thioether Ligation[a]



[a] (i) Ligation of ClAc-EPF(58−75)-[NH(CH$_2$)$_2$SH] **13** and BrAc-EPF(78−101) **12** to give ClAc-EPF(58−75)-G-ψ(CH$_2$S)-G-(78−101) **16**. (ii) Conversion of **16** to IAc-EPF(58−75)-G-ψ(CH$_2$S)-G-(78−101) **17**. (iii) Ligation of ClAc-EPF(30−55)-[NH(CH$_2$)$_2$SH] **14** and IAc-EPF(58−75)-G-ψ(CH$_2$S)-G (78−101) **17** to give ClAc-EPF(30−55)-G-ψ(CH$_2$S)-G-(58−75)-G-ψ(CH$_2$S)-G-(78−101) **18**. (iv) Simultaneous exchange and ligation of Ac-EPF(1−27)-[NH(CH$_2$)$_2$SH] **15** with ClAc-EPF(30−55)-G-ψ(CH$_2$S)-G-(58−75)-G-ψ(CH$_2$S)-G-(78−101) **18** yielding ligated EPF: Ac-EPF(1−27)-G-ψ(CH$_2$S)-G-(30−55)-G-ψ(CH$_2$S)-G-(58−75)-G-ψ(CH$_2$S)-G-(78−101) **19**.

therefore chloride to bromide exchange was not investigated further. It should be noted, however, that in 6 M Gu·HCl the rate of the reverse reaction (conversion of a bromoacetyl peptide to chloroacetyl) was significant (data not shown). Thus in further experiments only urea was used as denaturant.

Synthesis of Early Pregnancy Factor (EPF) [ψ(CH$_2$S)$^{28-29,56-57,76-77}$]. The preliminary reactivity studies underpinned the principles of successful sequential thioether ligation. Thus a bromoacetyl group was highly reactive toward thiolate at pH 8.5, while an accompanying chloroacetyl functionality was relatively unreactive. However, the chloroacetyl group could be converted quantitatively within 90 min to the reactive iodoacetyl group, which could then react selectively with a thiol in the presence of another chloroacetyl group. On the basis of these simple principles, we undertook to establish sequential peptide ligation through the synthesis of an analogue of the 101 residue "early pregnancy factor" (EPF)[27,28] that contained three thioether ligation sites.

EPF, an extracellular cytokine,[27] is identical in sequence to the intracellular "folding" protein human chaperonin 10 and provided a convenient synthetic target as it possesses three approximately equidistantly distributed Gly-Gly sites (Figure 2), suitable for replacement by the thioether amide bond surrogate. Inspection of the crystal structures of two chaperonin 10 homologues[45,46] indicated that the selected Gly-Gly residues lay in exposed loops and that replacement of the peptide bonds

Ac-AGQAFRKFLPLFDRVLVERSAAETVTKGGIML
PEKSQGKVLQATVVAVGSGSKGKGGEIQPVSVKV
GDKVLLPEYGGTKVVLDDKDYFLFRDGDILGKYVD

FIGURE 2. Amino acid sequence of EPF. In the sequential ligation synthesis of this protein from four peptide cassettes, amide bonds between the underlined Gly-Gly dipeptides were replaced by a thioether surrogate.

by thioether surrogates would not interfere with critical intra- or intermolecular hydrogen bonds required to maintain structure. In addition, activity assays using the rosette inhibition test were available for EPF,[47] so it would be possible to assess the impact of the thioether amide bond surrogates on activity of the synthetic EPF in direct comparison to the native protein.

The route chosen to synthesize EPF [ψ-(CH$_2$S)$^{28-29,56-57,76-77}$] is shown in Scheme 2. Four specifically functionalized peptide segments [BrAc-EPF-(78−101) **12**; ClAc-EPF(58−75)-[NHCH$_2$CH$_2$SH] **13**; ClAc-EPF(30−55)-[NHCH$_2$CH$_2$SH] **14**; and Ac-EPF(1−27)-[NHCH$_2$CH$_2$SH] **15** were prepared and purified in good yield. The assembly of EPF [ψ(CH$_2$-S)$^{28-29,56-57,76-77}$] **19** was then performed by sequential segment ligation

(45) Hunt, J. F.; Weaver, A. J.; Landry, S. J.; Gierasch, L.; Deisenhofer, J. *Nature* **1996**, *379*, 37−45.
(46) Xu, Z.; Horwich, A. L.; Sigler, P. B. *Nature* **1997**, *388*, 741−750.
(47) Morton, H.; Hegh, V.; Clunie, G. J. A. *Nature* **1974**, *249*, 459−460.

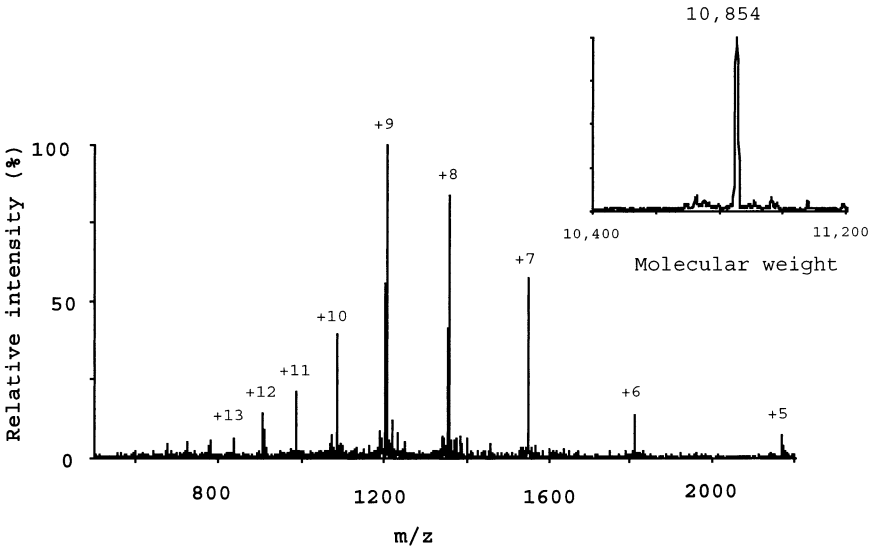


FIGURE 3. ES-MS mass spectrum of EPF [$\psi(CH_2S)^{28-29,56-57,76-77}$]. Inset: mathematical reconstruction of the mass spectrum showing appearance of ligated EPF as a singly charged species.

reactions. For each step we employed a small excess of the thiol-containing peptide segment and followed the ligation by LC-MS analysis. All reactions shown in Scheme 2 were complete within 4 h. To avoid minor losses of material we combined the final halo-exchange and ligation steps when reacting **15** and ClAc-EPF(30−55)-G-ψ(CH_2S)-G-(58−75)-G-ψ(CH_2S)-G-(78−101) **18**. The ligated EPF **19** was purified to yield 11.7 mg of protein with a cumulative yield of 19% based on BrAc-EPF -(78−101). RP-HPLC analysis showed a single peak[48] and ES-MS analysis (Figure 3) confirmed the high purity of the protein with good agreement between the observed molecular weight of the ligated protein (10853 ± 2 Da) and the calculated value (10852.8 Da, average isotope mass). Preliminary assays for EPF activity using a rosette inhibition test indicated that the ligated protein possessed full biological activity.[47,48] In addition, ES-MS analysis under native ionization conditions showed that the ligated protein appeared to associate into heptamers in a fashion similar to native chaperonin 10.[48]

Early sequential-type ligation methods focused on laborious semisynthetic strategies using specially tailored recombinant proteins.[49,50] Later, so-called "convergent chemical ligation" strategies use a combination of ligation chemistries for protein synthesis; examples include a "tethered dimer" HIV-1 protease[51] and a heterodimeric cMyc-Max protein construct containing two N-termini.[52] Most recently sequential native ligation methods have been used in a model synthesis of a 95 amino acid protein module,[19] to insert a short fluorophore-bearing synthetic peptide between two recombinant protein domains[53] and to synthesize a three zinc finger protein.[54]

We have shown by the synthesis of fully active[48] EPF [$\psi(CH_2S)^{28-29,56-57,76-77}$] the general utility of a novel sequential thioether ligation strategy for chemical protein synthesis. An interesting feature of this sequential ligation strategy is that each component peptide segment could, in principle, act as an interchangeable cassette. This approach would be particularly useful for chemical syntheses of variant proteins where each of the analogue-containing cassette peptides would be inserted during the synthesis of the full length protein, thereby avoiding the need for multiple full length syntheses.

Conclusion

We have described a novel thioether cassette strategy for the chemical ligation of unprotected peptide segments. The demonstration of "tunable" *N*-haloacetyl-peptide chemistry opens up many possibilities for future chemoselective ligation approaches. The successful synthesis of the fully active "early pregnancy factor" analogue, EPF [$\psi(CH_2S)^{28-29,56-57,76-77}$] with three thioether peptide bond replacements should make this approach a useful addition to the armamentarium of chemical protein synthesis methods. Further, the method should be compatible with other ligation chemistries (e.g., native chemical ligation) and other strategies (e.g., solid-phase chemical ligation).

Experimental Section

Materials and Methods. Diisopropylcarbodiimide, 2-aminoethanethiol, and α-bromoacetic acid were supplied by Aldrich (Milwaukee). 4-Hydroxymethylphenoxyacetic acid was supplied by Sigma (St. Louis, MO). α-Iodoacetic acid was purchased from BDH Chemicals Ltd. (Poole, UK). α-Chloroacetic acid was purchased from Ajax Chemicals (Sydney, Australia). Boc-L-amino acid-O-CH_2-Pam resins (Pam: phenylacetamidomethyl) were purchased from ABI Inc. (Foster City, CA). Aminomethyl polystyrene resin [100−200 mesh, 1% DVB cross-linked, amine substitution 0.83 mmol/g] and Boc-L-

(48) Love, S.; Englebretsen, D. R.; Garnham, B. G.; Cavanagh, A. C.; H Morton, H.; Alewood, P. F. In. *The NATO Series on Molecular Aggregates*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; pp 171−181

(49) Gaertner, H. F.; Offord, R. E.; Cotton, R.; Timms, D.; Camble, R.; Rose, K. *J. Biol. Chem.* **1994**, *269*, 7224−7230.

(50) Gaertner, H. F.; Offord, R. E.; Cotton, R.; Timms, D.; Camble, R.; Rose, K. *Bioconjugate Chem.* **1994**, *5*, 333−338.

(51) Baca, M.; Muir, T. W.; Schnolzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 1881−1887.

(52) Canne, L. E.; Ferré-D'Amaré, Y.; Burley, S. K.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 2998−3007.

(53) Cotton, G. J.; Ayers, B.; Xu, R.; Muir, T. W. *J. Am. Chem. Soc.* **1999**, *121*, 1100−1101.

(54) Beligere, G. S.; Dawson, P. E. *Biopolymers (Pept. Sci.)* **1999**, *51*, 363−369.

amino acids were from the Peptide Institute (Osaka, Japan). Dichloromethane, diisopropylethylamine, *N*,*N*-dimethylformamide, di-*tert*-butyl pyrocarbonate, and trifluoroacetic acid were obtained from Auspep (Melbourne, Australia). RP-HPLC grade acetonitrile was purchased from Millipore-Waters (Sydney, Australia). HBTU (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) was purchased from Richelieu Biotechnologies (Quebec, Canada). Deionized water was used throughout and was prepared by a Milli-Q water purification system. Screw-cap glass peptide synthesis reaction vessels (20 mL) with sintered glass filter frit were obtained from Embell Scientific Glassware (Queensland, Australia). An all-Kel-F apparatus was used for HF cleavage. Argon, helium, and nitrogen (all ultrapure grade) were from BOC gases (Queensland, Australia). ^{1}H NMR spectra were recorded on a 300 MHz instrument in $CDCl_3$ and chemical shifts are reported in parts per million (ppm) downfield from $(CH_3)_4Si$. Thin-layer chromatography was performed using silica gel plates. Unless otherwise noted in the text, analytical RP-HPLC was performed using Vydac C4 or C18 columns (5 μm, 0.46 cm × 15 cm) with linear gradients of 5−85% B over 40 min at a flow rate of 1 mL/min (A = 0.1% aqueous TFA; B = 90% CH_3CN, 10% H_2O, 0.09% TFA) and monitoring at 214 nm.

Mass spectra were acquired on a PE Sciex API III triple quadrupole mass spectrometer equipped with an ion spray atmospheric pressure ionization source. Samples (10 μL) were injected into a moving solvent (30 μL/min; 50/50 CH_3CN/0.05% TFA) coupled directly to the ionization source via a fused silica capillary interface (50 μm i.d. × 50 cm length). Sample droplets were ionized at a positive potential of 5 kV and entered the analyzer through an interface plate and subsequently through an orifice (100−120 μm diameter) at a potential of 80 V. Full scan mass spectra were acquired over the mass range of 400 to 2000 daltons with a scan step size of 0.1 Da. Molecular masses were derived from the observed *m*/*z* values using the program MacSpec 3.3 (PE-Sciex Toronto, Canada). Calculated average masses were determined using the MacBiospec program (PE-Sciex Toronto, Canada).

Synthesis of 4-[*N*-Boc-2-aminoethylmercapto]methylphenoxyacetic Acid (AMPA linker) 1. 4-Hydroxymethylphenoxyacetic acid (1.82 g, 10 mmol) and 2-aminoethanethiol hydrochloride (1.14 g, 10 mmol) were dissolved in 15 mL of TFA, and stirred at room temperature. Thin-layer chromatography (*n*-butanol:acetic acid:water, 4:1:1) showed complete reaction after 70 min. The mixture was diluted with 150 mL of water and lyophilized. The residue was dissolved in 100 mL of 0.1% aqueous TFA, filtered to remove insoluble material, and lyophilized. The white fluffy powder (4-[2-aminoethylmercapto]methylphenoxyacetic acid·TFA salt, 3.07 g, 86%) was dissolved in 22 mL of 0.1 M NaOH. Dioxane (25 mL) was added and the solution was cooled to 0 °C. Di-*tert*-butyl pyrocarbonate (2.25 g, 10.3 mmol, 1.2 equiv) was dissolved in 15 mL of dioxane and added to the stirred solution. After 35 min the solution was warmed to room temperature. After 55 min thin-layer chromatography of the cloudy solution indicated incomplete reaction. Ten milliliters of a 0.1 M NaOH solution was then added, whereupon the solution cleared, and after 75 min reaction was complete. The solution was diluted with 250 mL of water and extracted twice with 100 mL of ethyl acetate to remove unreacted di-*tert*-butyl pyrocarbonate The aqueous layer was carefully acidified to pH 1 with 10 mL of 6 M HCl and extracted twice with 100 mL of ethyl acetate. The combined organic layers were washed twice with 25 mL of water and twice with a saturated NaCl solution. The organic phase was dried with $MgSO_4$, and the ethyl acetate was removed in-vacuo leaving an oil, which was dissolved in ethyl acetate (30 mL) and then triturated with petroleum ether (100 mL) and cooled to 0 °C. After standing for 30 min the resultant precipitate was collected by filtration and washed with 40 mL of ethyl acetate:petroleum ether (1:3). The solid was dried in-vacuo to yield 1.97 g of 4-[*N*-Boc-2-aminoethylmercapto]-methylphenoxyacetic acid (AMPA linker **1**, 5.8 mmol, 58% overall yield), mp 83−85 °C (uncorr.) Anal. Calcd for $C_{16}H_{23}NO_5S$: C, 56.30; H, 6.74; N, 4.11. Found: C, 56.37; H, 6.81; N, 4.03. ^{1}H NMR ($CDCl_3$): δ 7.25 (d, 2H, *J* = 8.7 Hz), 6.87 (d, 2H, *J* = 8.6 Hz), 4.65 (s, 2H), 3.67 (s, 2H), 3.26 (m, br, 2H), 2.52 (t, 2H), 1.45 (s, 9H).

Functionalization of Aminomethyl Polystyrene with 4-Methyl-[*N*-Boc-2-aminoethylmercapto]methylphenoxyacetic Acid. Aminomethyl polystyrene·HCl resin (1.1 mmol amine) was swollen in DMF:DIEA 3:1 for 10 min. The resin was collected by filtration and washed with DMF. The AMPA linker **1** (1.5 mmol, 1.4 equiv) and HBTU (1.5 mmol) were dissolved in 4 mL of DMF and DIEA (460 μL, 2.6 mmol) was added. The activated linker solution was then added to the resin. After mixing for 25 min, the ninhydrin assay indicated complete substitution of resin-bound amine by the linker.

Peptide Synthesis. Peptides were synthesized manually on a 0.50 mmol scale using HBTU activation of Boc-amino acids with in situ neutralization chemistry as previously described.[41] The syntheses were performed on Boc-Phe-OCH$_2$-Pam, Boc-Ser(OBzl)-OCH$_2$-Pam, and appropriately functionalized aminomethyl polystyrene resins. The following amino acid side chain protection was used: Boc-Lys(ClZ)-OH, Boc-Ser(Bzl)-OH, Boc-Trp(CHO)-OH, Boc-Thr(Bzl)-OH, Boc-Arg-(Tos)-OH, Boc-Asp(OChx)-OH and Boc-Glu(OChx)-OH, Boc-Asn(Xan)-OH and Boc-Gln(Xan)-OH, Boc-Tyr(2BrZ)-OH. Each residue was coupled for 10 min and coupling efficiencies were determined by the quantitative ninhydrin reaction.[55] Prior to HF cleavage, the N-terminal Boc protecting group was removed (100% TFA) followed by formyl group removal where required (1.5 mL of ethanolamine in 25 mL of DMF/5% water (2 × 30 min). The peptide-resins were washed with DMF and then DCM and dried on a sinter with vacuum under a gentle stream of nitrogen. Selected peptides were haloacetylated at their N-terminus, as follows. The amino-peptide-resin was washed with DMF, neutralized for 2 min with 20% DIEA/DMF, washed with DMF, and haloacetylated for 10 min using the desired haloacetic acid anhydride (5−10 equiv) dissolved in DMF. After a final DMF wash the N^{α}-haloacetyl peptide resin was washed with DCM and dried as above. All dried peptide-resins were cleaved using 10 mL of *p*-cresol:HF 1:9

(55) Sarin, V.; Kent, S. B. H.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem*. **1981**, *117*, 147−157.

(56) In principle, the ligation chemistries selected could involve differential regioselective chemistry for each cassette.[19,52] Further, more general versions of this strategy can be readily envisaged. These could include N- to C-terminal protein synthesis where each cassette is incorporated at the C-terminus of the growing peptide chain and supported solid-phase cassette ligation where the C- or N-terminal cassette is attached to a support with subsequent addition of each cassette in a fashion similar to stepwise SPPS.

(57) Boc SPPS refers to use of *tert*-butoxylcarbonyl and benzyl temporary and permanent protecting groups, respectively. Abbreviations use are the following: CH_3CN, acetonitrile; DCM, dichloromethane; DIEA, diisopropylethylamine; DMF, *N*,*N*-dimethylformamide; Da, Dalton; ES-MS, electrospray mass spectrometry; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HF, anhydrous hydrogen fluoride; LC-MS, liquid chromatography mass spectrometry; NMR, nuclear magnetic resonance; Pam, phenylacetamidomethyl polystyrene resin, RP-RP-HPLC, reversed-phase high performance liquid chromatography; SPPS, solid-phase peptide synthesis; TFA, trifluoroacetic acid. Standard IUPAC single and triple letter codes for amino acids are used throughout.

(58) **Terminology.** In this report we refer to the N-terminally functionalized α-chloro-, bromo-, and iodo-acetylated peptides by the respective prefixes ClAc, BrAc, and IAc. Thus N^{α}-chloroacetyl-Ile-Gly-Gly-Phe is denoted as ClAc-IGGF. The C-terminal 2-mercapto ethylamido functional group is enclosed in brackets; thus Leu-Pro-Gly-Lys-Trp-Lys-Pro-Lys-Met-Ile-NH(CH$_2$)$_2$-SH is represented as LPGK-WKPKMI-[NH(CH$_2$)$_2$-SH]. The thioether surrogate for Gly-Gly dipeptides, -[NH-CH$_2$-CH$_2$-S-CH$_2$-CO]-, is represented using IUPAC notation where the amide bond (CONH) which is replaced by the thioether surrogate is denoted -ψ(CH$_2$-S)-; thus the product of the ligation of LPGKWKPKMI-[NH(CH$_2$)$_2$-SH] and BrAc-IGGF is shown as LPGKWKPKMIG-ψ(CH$_2$-S)-GIGGF.

(59) Note that the pK_a of the thiol group in both cysteine and cystamine is approximately 8.2.

(−5 °C, 90 min). HF was removed in-vacuo, and the peptides were precipitated with 10 mL of ether, collected by filtration on a sintered glass funnel under a flow of nitrogen, and washed twice with 25 mL of ether. Following dissolution in 50% acetic acid, peptide solutions were diluted with water and lyophilized.

Peptide Isolation and Characterization. Crude peptides (up to 70 mg) were purified by preparative RP-HPLC using either a Vydac C-4 (10 μm, 2.2 cm × 25 cm) or Waters Delta-pak C-18 (10 μm, 1.9 cm × 30) column with a linear gradient from 0 to 60% B over 60 min at 8 mL/min. Fractions were collected based on monitoring of the column effluent at 230 nm. Those fractions shown by ES-MS analysis to contain solely the target peptide were combined and lyophilized. Purified peptides were homogeneous by HPLC analysis and showed the expected molecular ion masses by ES-MS: ClAc-IGGS **6** M_r 409.1, calcd 409.8; ClAc-IGGF **3** M_r 469.3, calcd 470.0; BrAc-IGGF **4** M_r 514.0, calcd 514.4; IAc-IGGS **8** M_r 500.8, calcd 501.3; LPGKWKPKMI-[NH(CH_2)$_2$-SH] **2** M_r 1256.9, calcd 1257.6; ClAc-LPGKWKPKMI-[NH(CH_2)$_2$-SH] **9** M_r 1332.4, calcd 1334.1; BrAc-TKVVLDDKDYFLFRDGDILGKYVD {BrAc-EPF(78−101)} **12** M_r 2956.4, calcd For $C_{133}H_{202}N_{30}O_{41}$-Br, 2957.2; ClAc-EIQPVSVKVGDKVLLPEY-[NH(CH_2)$_2$-SH] {ClAc-EPF(58−75)-[NH(CH_2)$_2$-SH]} **13** M_r 2149.5, calcd for $C_{97}H_{160}N_{22}O_{28}SCl$, 2149.9; ClAc-IMLPEKSQGKVLQATVVAV-GSGSKGK-[NH(CH_2)$_2$-SH] {ClAc-EPF(30−55)-[NH(CH_2)$_2$-SH]} **14** M_r 2749.4, calcd for $C_{118}H_{209}N_{33}O_{35}S_2Cl$, 2749.6; Ac-AGQAFRKFLPLFDRVLVERSAAETVTK-[NH(CH_2)$_2$-SH], {Ac-EPF(1−27)-[NH(CH_2)$_2$-SH]} **15** M_r 3151.9, calcd for $C_{143}H_{233}N_{40}O_{38}S$, 3152.6 (all calcd weights are average isotope masses).

Reaction of LPGKWKPKMI-[NH(CH_2)$_2$SH] 2 and X-Ac-IGGF. BrAc-IGGF **4** (0.72 μmol) was dissolved in 350 μL of 6 M Gu·HCl, 100 mM Tris pH 8.5. The solution was added to 0.93 mg of **2** (0.74 μmol). Samples (20 μL) were withdrawn at intervals, quenched with 80 μL of 20% TFA in water, and analyzed by RP-HPLC. The experiment was repeated at pH 7.9. This experiment was also done using ClAc-IGGF **3** instead of **4** at pH values of 7.9, 8.5, and 9.5.

Competition between (ClAc-IGGS 6 and BrAc-IGGF 4) or (BrAc-IGGF 4 and IAc-IGGS 8) for LPGKWKPKMI-[NH(CH_2)$_2$SH] 2. 6 (1.95 μmol), **4** (1.84 μmol), and **2** (1.62 μmol) were dissolved in 900 μL of 6 M Gu·HBr, 100 mM Tris pH 8.5 to give an initial peptide concentration of ∼2 μM. At intervals 200 μL samples were quenched in a mixture of 200 μL of 50% RP-HPLC solvent B and 200 μL of 50% TFA. The quenched samples were analyzed by RP-HPLC (Vydac C18, 0.46 × 15 cm, 0−40% B over 40 min at 1 mL/min). Peaks from RP-HPLC analyses were trapped and identified by ion spray mass spectrometry. In similar experiments the competition between **4** and **8** for **2** was studied at pH values of 8.5, 7.0, and 6.0.

Reaction of ClAc-LPGKWKPKMI-[NH(CH_2)$_2$SH] 9 with BrAc-IGGF 4. 4 (1.46 μmol) and **10** (1.39 μmol) were dissolved in 700 μL of 6 M GuHBr, 100 mM Tris pH 8.5. Samples were quenched after 1 and 2 h for RP-HPLC and ES-MS analysis as described above.

Conversion of ClAc-IGGF 3 to IAc-IGGF 11. 3 (0.72 μmol) was dissolved in 340 μL of freshly prepared saturated KI, 100 mM phosphate pH 7.0 (peptide concentration approximately 2 mM). During the course of the exchange reaction samples were taken for RP-HPLC and ES-MS analysis. The experiment was repeated using 0.1% TFA solution saturated with KI (pH 2). The exchange was also studied at a peptide concentration of 22 mM in saturated KI, 100 mM phosphate, pH 7.0.

Synthesis of EPF 19 by Sequential Ligation: Ligation of ClAc-EPF(58−75)-[NH(CH_2)$_2$SH] 13 and BrAc-EPF (78−101) 12. 13 (13.4 μmol) and **12** (11.0 μmol) were dissolved in 5 mL of 8 M urea, 100 mM Tris·HCl, pH 8.5. A sample taken after 40 min for RP-HPLC analysis showed the ligation to be almost complete. After 2 h the reaction was quenched with 500 μL of acetic acid and purified by RP-HPLC. Fractions were analyzed by ES-MS, and those containing the target peptide were pooled and lyophilized. The ligated peptide ClAc-EPF-(58−75)-G-ψ(CH_2S)-G-(78−101) **16** (M_r 5025.0, calcd 5025.3, 5.0 μmol) was dissolved in 5 mL of 8 M urea saturated with NaI (pH 3.5) to effect halide exchange. After 30 min RP-HPLC and ES-MS analysis showed complete exchange, and the IAc-EPF(58−75)-G-ψ(CH_2S)-G-(78−101) **17** was isolated by RP-HPLC to give 4.9 μmol of peptide (M_r 5117.1, calcd For $C_{230}H_{360}N_{52}O_{69}SI$ 5116.7).

Ligation of ClAc-EPF(30−55)-[NH(CH_2)$_2$SH] 14 and IAc-EPF(58−75)-G−ψ(CH_2S)-G(78−101) 17. 17 (4.9 μmol) and **14** (8.5 μmol) were dissolved in 4 mL of 8 M urea, pH 8.5. The reaction was quenched with 100 μL of acetic acid after 3 h. RP-HPLC purification gave 4.8 μmol of ClAc-EPF(30−55)-G-ψ(CH_2S)-G-(58−75)-G-ψ(CH_2S)-G-(78−101) **18** (M_r 7736.5, calcd for $C_{348}H_{567}N_{85}O_{104}S_3Cl$, 7737.5).

Simultaneous Exchange and Ligation of Ac-EPF(1−27)-[NH(CH_2)$_2$SH] 15 with ClAc-EPF(30−55)-G-ψ(CH_2S)-G-(58−75)-G-ψ(CH_2S)-G-(78−101) 18. 15 (3.4 μmol) and **18** (2.3 μmol) were dissolved in 3 mL of 8 M urea, 100 mM Tris·HCl, 1 M KI, pH 9.5. RP-HPLC analysis after 30 min indicated little reaction had occurred. KI, 2.0 g, was therefore added 60 min after the start of the reaction, and after a further 2 h RP-HPLC analysis indicated that the reaction was almost complete. After a total of 4 h the mixture was diluted with 10 mL of 8 M urea, pH 8.5, and then quenched with 100 μL of acetic acid. RP-HPLC purification yielded 1.0 μmol of ligated EPF: Ac-EPF(1−27)-G-ψ(CH_2S)-G-(30−55)-G-ψ(CH_2S)-G-(58−75)-G-ψ(CH_2S)-G-(78−101) **19**. The ligated protein was characterized by RP-HPLC and ES-MS analysis (M_r 10853 ± 2, calcd for $C_{491}H_{798}N_{125}O_{142}S_4$ 10852.3).

Acknowledgment. This work was supported in part by the Australian Research Grants commission. We thank Drs. John Wade and Les Miranda for helpful comments in preparing the manuscript and Dianne Alewood for preparing some of the schemes and figures. We also thank Dr. Joel D. A. Tyndal and Dr. Nathan P. Cowieson, Institute for Molecular Bioscience, for their assistance with the cover art.

Supporting Information Available: Experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

JO016121E