Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments

Lynne E. Canne,[†] Paolo Botti, Reyna J. Simon,[§] Yijun Chen,[‡] Edward A. Dennis,[‡] and Stephen B. H. Kent*

Contribution from Gryphon Sciences, 250 E. Grand Avenue, Suite 90, South San Francisco, California 94080

Received October 15, 1998

Abstract: In this paper we describe "solid phase chemical ligation" (SPCL), the application of the principles of polymer-supported organic synthesis to the construction of large polypeptide chains for the total chemical synthesis of proteins. In this method, each building block used is an *unprotected* peptide segment of 20 or more amino acids. These are consecutively reacted by chemical ligation, the chemoselective reaction of the unprotected peptide segments from aqueous solution, to make the polymer-supported target polypeptide. In a final step, the assembled full-length target polypeptide is released from the aqueous-compatible polymer support. Here we report chemistries for the attachment of the first segment to a polymer support, and for the assembly of the target polypeptide chain starting from the polymer-bound peptide segment. In this solid phase protein synthesis method, large target polypeptide chains can be built efficiently and rapidly by SPCL and, after release from the polymer support, folded to give functional protein molecules. Several examples of the application of SPCL are given: model peptides consisting of 27 and 68 amino acids, and polypeptides corresponding to the proteins C5a (74 amino acids) and MIF (115 amino acids), were each made in good yield and purity from the consecutive solid phase ligation of peptide segments. In addition, we report the total synthesis by SPCL of the enzyme "human group V secretory phospholipase A_2" (GV-PLA_2), which comprises a polypeptide of 118 amino acids containing 6 disulfide bonds. As demonstrated by these examples, SPCL is an important extension of our capabilities for total chemical protein synthesis.

Introduction

Chemical ligation,[1] the chemoselective reaction of unprotected peptides in aqueous solution, has become established as an effective method for the synthesis of large polypeptides that fold efficiently to form active protein molecules.[2] Most chemical protein syntheses by chemical ligation have involved the joining together of just two unprotected peptide segments.[1−4] Sequential chemical ligation in solution of three or more peptide segments is rare,[5] because of the additional chemical manipulations dictated by the need for temporary protection of the "middle" segments, each of which carries both chemoselectively reactive functionalities. The need for a time-consuming and labor-intensive purification step after each ligation is also a significant consideration. In this paper we describe "solid phase chemical ligation" (SPCL), a method that enables the sequential chemoselective reaction of unprotected peptide segments to give a polymer-supported polypeptide product. Our solid phase protein synthesis approach uses the chemical ligation of unprotected peptide segments,[1] and either avoids the necessity for protection of the difunctional middle segments altogether or facilitates the removal of any protecting group used.

General strategies for SPCL are outlined in Schemes 1−3. The key feature of our solid-phase approach is that the target polypeptide can be assembled by the consecutive chemical ligation of unprotected peptide segments in either the N-to-C

* To whom correspondence should be addressed. Phone: (650) 952-7714. Fax: (650) 952-3055. E-mail: skent@gryphonsci.com.

[†] Current address: Exelixis Pharmaceuticals, Suite 1, 280 E. Grand Ave, South San Francisco, CA 94080.

[§] Current address: IntraBiotics, 1255 Terra Bella Ave., Mountain View, CA 94043.

[‡] Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0601.

(1) (a) Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221−225. (b) Liu, C. F.; Tam, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6584−6588. (c) Rose, K. *J. Am. Chem. Soc.* 1994, *116*, 30−34. (d) Muir, T. W.; Williams, M. J.; Ginsberg, M. H.; Kent, S. B. H. *Biochemistry* **1994**, *33*, 7701−7708. (e) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.

(2) Wilken, J.; Kent, S. B. H. *Curr. Opin. Biotechnol.* **1998**, *9*, 412−426.

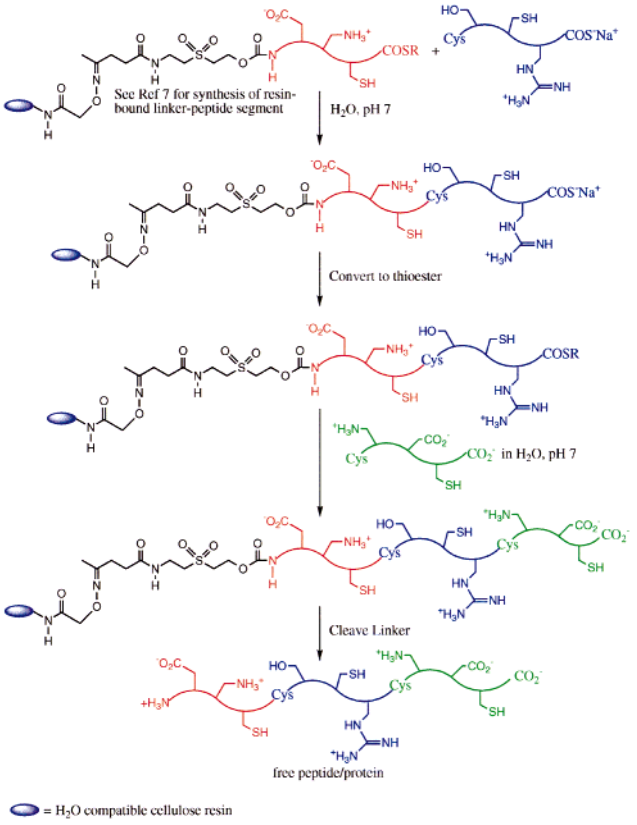
(3) (a) Williams, M. J.; Muir, T. W.; Ginsberg, M. H.; Kent, S. B. H. *J. Am. Chem. Soc.* **1994**, *116*, 10797−10798. (b) Lu, W.; Qasim, M. A.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 8518−8523. (c) Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. *Nature Struct. Biol.* **1996**, *3*, 946−950. (d) Hackeng, T. M.; Mounier, C. M.; Bon, C.; Dawson, P. E.; Griffin, J. H.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7845−7850. (e) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325−4329. (f) Lu, W.; Qasim, M. A.; Laskowski, M., Jr.; Kent, S. B. H. *Biochemistry* **1997**, *36*, 673−679. (g) Lu, W.; Starovasnik, M. A.; Kent, S. B. H. *FEBS Lett.* **1998**, *429*, 31−35. (h) Wilken, J.; Hoover, D.; Thompson, D. A.; Barlow, P. N.; McSparron, H.; Picard, L.; Wlodawer, A.; Lubkowski, J.; Kent, S. B. H. *Chem. Biol.* **1999**, *6*, 43−51. (i) Bolin, K. A.; Anderson, D. J.; Trulson, J. A.; Gantz, I.; Thompson, D. A.; Wilken, J.; Kent, S. B. H.; Milhauser, G. L. *FEBS Lett.*, **1999**, *451*, 125−131.

(4) (a) Baca, M.; Muir, T. W.; Schnölzer, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1995**, *117*, 1881−87. (b) Ueda, H.; Siani, M. A.; Gong, W.; Thompson, D. A.; Brown, G. G.; Wang, J. M. *J. Biol. Chem.* **1997**, *272*, 24966−24960. (c) Campbell, J. J.; Hendrick, J.; Zlotnik, A.; Siani, M. A.; Thompson, D. A.; Butcher, E. C. *Science* **1998**, *279*, 381−384. (d) Dealwis, C.; Fernandez, E. J.; Thompson, D. A.; Simon, R. J.; Siani, M. A.; Lolis, E. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6941−6946.

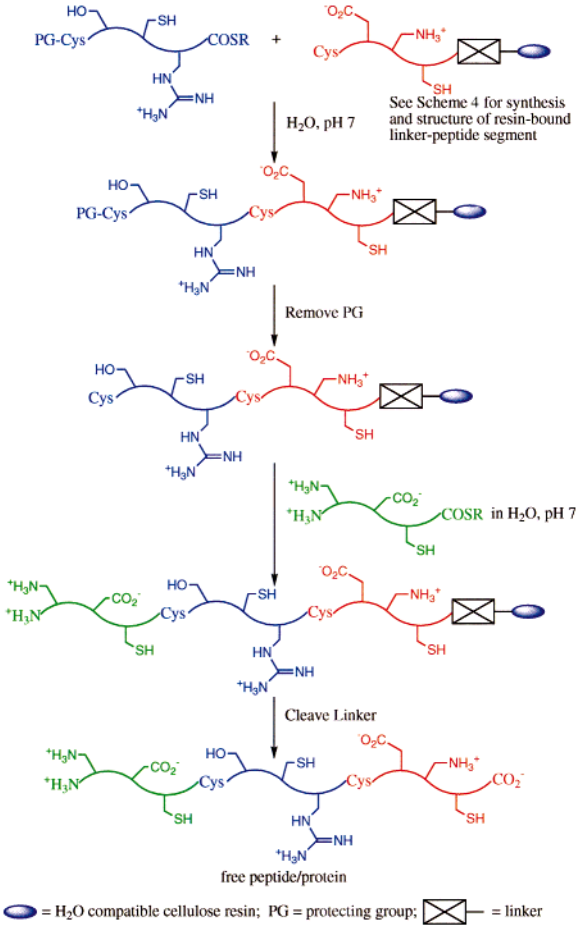
(5) (a) Rau, H. K.; Haehnel, W. *J. Am. Chem. Soc.* **1998**, *120*, 468−476. (b) Figliozzi, G. M.; Siani, M. A.; Canne, L. E.; Robson, B.; Simon, R. J. *Abstracts of Papers*, 24th Symposium of the European Peptide Society, Edinburgh, Scotland; European Peptide Society: Swansea, U.K., 1996; p 202. (c) Love, S. G.; Englebretsen, D. R.; Garnham, B.; Cavanagh, A. C.; Morton, H.; Chernushevich, I. C.; Standing, K. G.; Alewood, P. F. *NATO Series on Molecular Aggregates* 1996.

10.1021/ja9836287 CCC: $18.00
Published on Web 09/09/1999

Scheme 1. Solid Phase Chemical Ligations in the N- to C-Terminal Direction



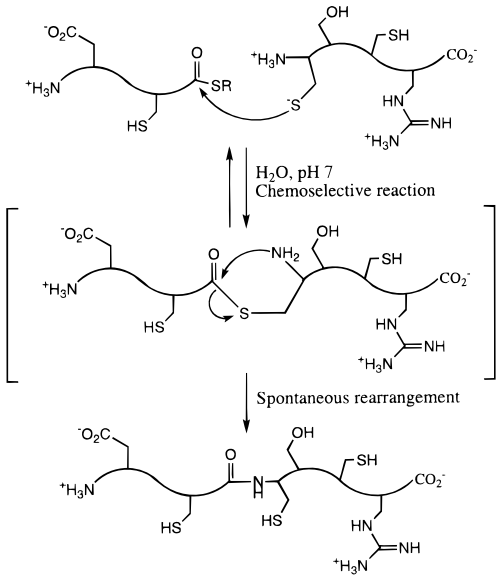
Scheme 2. Solid Phase Chemical Ligations in the C- to N-Terminal Direction



direction (i.e., from the polymer-bound N-terminal segment toward the C-terminal of the target polypeptide chain) or in the C-to-N direction (i.e., from the polymer-bound C-terminal segment toward the N-terminal of the polypeptide chain). In both these approaches, SPCL employs thioester-mediated amide-forming "native" chemical ligation[1e] (Scheme 3) of the unprotected peptide segments in aqueous solution, and takes advantage of the observation that to date racemization has not been observed as a side reaction in coupling reactions performed under the neutral conditions employed in this chemistry.[3b] This absence of racemization is critical to the effective use of a segment condensation strategy in the chemical synthesis of polypeptide chains. The expectation that racemization should be minimal under the native chemical ligation reaction conditions[1e] is now supported by a rapidly growing body of experimental evidence that includes a detailed study of racemization in an example of the use of native chemical ligation in the synthesis of a protein,[3b] the synthesis and near-quantitative correct folding of a number of biologically active proteins,[2] the determination of several high-resolution structures by X-ray crystallography of proteins prepared in good yield by native chemical ligation,[3h,4d] and detailed multidimensional NMR studies of proteins prepared by native chemical ligation.[3g,i]

Another issue in the consecutive native chemical ligation of unprotected peptides in a controlled manner is the potential reactivity of the middle peptide segments, each of which would bear *both* an N-terminal Cys *and* a C-terminal thioester. Such peptides thus would have mutually reactive moieties which would self-condense in an uncontrolled manner to give (primarily cyclic) products of intramolecular reaction.[6] To obviate

Scheme 3. Native Chemical Ligation Mechanism



such undesirable side reactions, we employed either of two strategies. The first was a subtle strategy used for synthesis in the N-to-C direction in which the middle segments are of the form Cys-[peptide]-$^{\alpha}$thiocarboxylate. Effectively, the *absence* of a proton acts as a protecting group for the thioacid, because the ionized thiocarboxylate is unreactive under ligation conditions. In the second strategy, for synthesis in the C-to-N direction, we used a single reversible protecting group on the

(6) (a) Zhang, L.; Tam, J. P. *J. Am. Chem. Soc.* **1997**, *119*, 2363−2370. (b) Camarero, J. A.; Pavel, J.; Muir, T. W. *Angew. Chem., Int. Ed.* **1998**, *37*, 347−349.

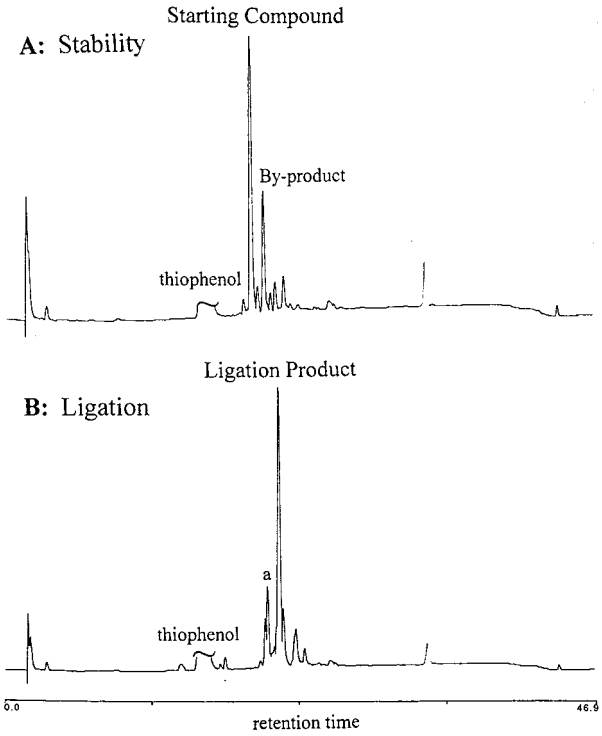


Figure 1. Evaluation of the suitability for use in native chemical ligation of an unprotected peptide segment containing both an unprotected N-terminal Cys residue and a C-terminal $^{\alpha}$thiocarboxylate. The sequence of the peptide was CGFRVREFGDNTA-$^{\alpha}$COSH. (A) Analytical HPLC after the peptide segment had been exposed to ligation conditions (pH 7.0, 1% thiophenol) for 24 h. The major peak is unreacted original peptide-$^{\alpha}$COSH starting material, observed mass 1487.2 ± 0.4 Da, calculated mass 1487.6 Da (average isotopes). The minor peak gave a mass consistent with intramolecular reaction between the N-terminal Cys and the thioacid to form a cyclic product (the observed mass of 1452.7 ± 0.4 Da was 34.5 ± 0.8 Da less than that of the starting material, consistent with loss of H_2S). (B) Analytical HPLC of the Cys-(peptide segment)-$^{\alpha}$COS$^-$ after being exposed overnight to ligation conditions (pH 7.0, 1% thiophenol) in the presence of a peptide-$^{\alpha}$thioester of sequence DSVISLLSGDH-$^{\alpha}$COSR. The major peak represents the desired ligation product between the two peptides (observed mass 2497.7 ± 0.4 Da, calculated mass 2498.7 Da). Peak a is the intramolecular side reaction of the Cys-(peptide segment)-$^{\alpha}$COSH, as seen in the previous HPLC.

N-terminal Cys residue of the otherwise unprotected middle segments and took advantage of polymer-supported chemistry to facilitate its removal. In this paper, chemistries for both strategies are described and examples of the application of the two SPCL strategies to solid phase protein synthesis are given.

Results and Discussion

SPCL in the N-to-C Direction (Scheme 1). In model studies, some results of which are shown in Figure 1, there was only a low level of intramolecular reaction noted when a bifunctional Cys-[peptide]-$^{\alpha}$thiocarboxylate segment underwent prolonged exposure to the reaction conditions employed in native chemical ligation (Figure 1A). This level was further reduced in the presence of a second, thioester-containing peptide (Figure 1B), indicating that the rate of the native chemical ligation reaction competes well with the rate of the intramolecular side reaction(s).

We reasoned that the amount of intramolecular reaction will be of no consequence if a soluble Cys-[peptide]-$^{\alpha}$thiocarboxylate middle segment is added in excess to the resin-bound peptide-thioester, since the ligation reaction will be driven to completion and any unreacted peptide and/or side products remaining in solution can be simply washed away. Thus, for SPCL by consecutive chemical ligation of unprotected peptide segments in the N-to-C direction (Scheme 1) using the Cys-[peptide]-$^{\alpha}$thiocarboxylate chemical tactics, there was no need for temporary protection of *any* of the functionalities on the middle peptide segments.

For this strategy, the N-terminal peptide segment was equipped at the N$^{\alpha}$-moiety with a cleavable linker capable of reacting in a chemoselective fashion with a suitably derivatized water-compatible polymeric support. The cleavable linker for the N$^{\alpha}$-moiety of the N-terminal segment, previously reported as a purification handle for peptides,[7] is a functionalized derivative (BocNHCH$_2$-Msc) of the methylsulfonylethyloxycarbonyl amine protecting group.[8] This linker was ideal for SPCL in that it is easily added to the unprotected amino terminus of peptide-resins, survives the HF treatment used for deprotection and cleavage of peptides from the polystyrene resin used in Boc chemistry stepwise solid phase synthesis of the unprotected segments, is quickly and cleanly cleaved by aqueous base, and was designed to include a protected amine which can be derivatized with any of a variety of functionalities.[7] In the examples to follow, this amine functionality was elaborated with levulinic acid,[9] creating a ketone functionality at the end of the linker. The SPCL polymer support, a cellulose-based amino affinity resin, was derivatized with aminoxyacetic acid,[1c,9] and the modified N-terminal peptide segment was coupled to the resin through the oxime-forming chemoselective reaction[1c,9] between the ketone on the linker-peptide and the aminooxy group on the polymer support (Scheme 1).

The next unprotected peptide segment, bearing an N-terminal Cys and a C-terminal $^{\alpha}$thiocarboxylate, was added under native chemical ligation conditions (pH 7.0, 1% thiophenol).[3e] Once ligation was complete and unreacted reagents had been washed away, the pH was dropped to 4−5 and the $^{\alpha}$thiocarboxylate of the polymer support-bound peptide was converted to a thioester in a chemoselective reaction by the addition of bromoacetic acid in water. At this low pH, other potentially nucleophilic functional groups in the unprotected peptide segments were unreactive. Excess reagent was simply removed by washing the polymer support-bound peptide-$^{\alpha}$thioester, and the pH was returned to 7.0; the polymer support-bound peptide intermediate product was then ready for ligation to the next peptide segment. The cycle of ligation and activation of the resin-bound product was repeated. After the ligation of the final peptide segment, the linker to the polymer support was cleaved by treatment with aqueous base (NaOH, pH 12−14) for 2 min, freeing the full-length peptide, which can then be purified by preparative HPLC.

To illustrate the efficacy of the technique, SPCL in the N- to C-terminal direction was used to synthesize several target polypeptides, each consisting of three peptide segments and requiring two consecutive native ligation steps. Analytical HPLCs of the crude final ligation products released from the polymer support are shown in Figure 2. First, as proof of concept, a model polypeptide target sequence made up of three arbitrary peptide segments was used. The N-terminal segment consisted of 35 residues, the middle segment of 13 residues, and the third segment of 20 residues, to give as a final product after two consecutive native ligations, a large peptide of 68

(7) Canne, L. E.; Winston, R. L.; Kent, S. B. H. *Tetrahedron Lett.* **1997**, *38*, 3361−3364.

(8) *Protecting Groups in Organic Synthesis*, 2nd ed.; Greene, T. W., Wuts, P. G. M., Eds.; John Wiley & Sons: New York, 1991.

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

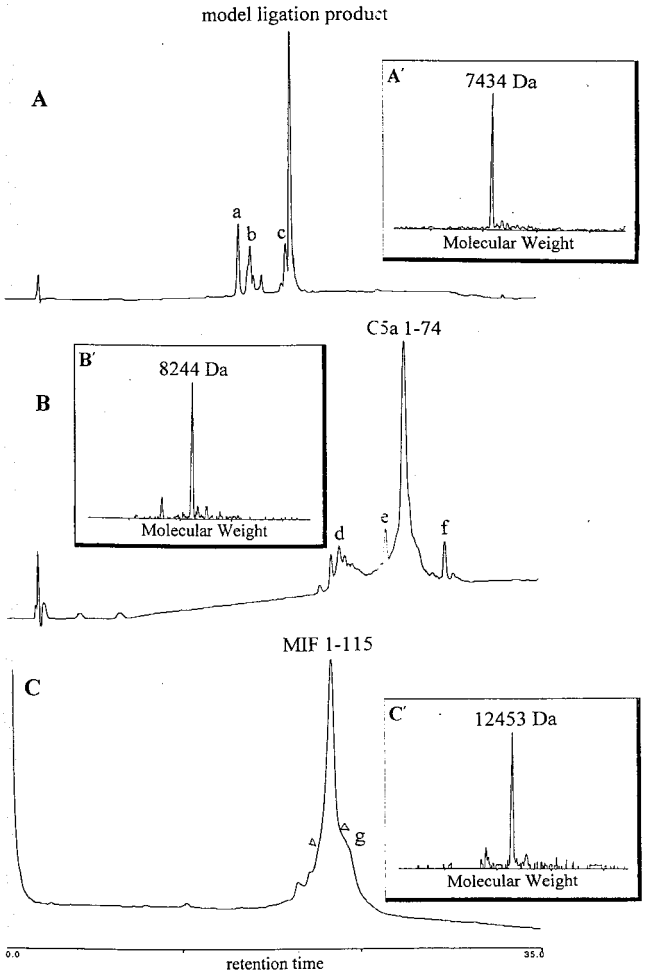


Figure 2. Polypeptide products from three-segment N-to-C strategy solid phase chemical ligation syntheses. Analytical HPLC was used to evaluate the crude products; peaks were identified by ESMS. (A) Total crude products from the N-to-C SPCL synthesis of a 68-residue peptide model system.[10] The major peak has the mass expected for the desired product: observed mass 7434.2 ± 0.5 Da, calculated mass 7434.1 Da (average isotope composition). Peak a is the unreacted first segment, peak b is the unreacted first ligation product, and peak c is the product of ligation between the first segment and the third segment. (B) Total crude products from the N-to-C SPCL synthesis of the 74-residue polypeptide of human anaphylatoxin C5a. The major peak has the mass expected for the desired product: observed mass 8244.7 ± 1.0 Da, calculated mass 8245.7 Da (average isotope composition). The group of peaks under d represent the unreacted first ligation product (C5a 1−46). [The presence of more than one peak is due to minor impurities in the second peptide segment (C5a 21−46) used.] Peak e is the unreacted third peptide segment (C5a 47−74) which did not get completely washed out of the resin, and peak f is the product of ligation between the first and third peptide segments (C5a 1−20 and C5a 47−74). (C) Total crude products from the N-to-C SPCL synthesis of the 115-resiue polypeptide of human MIF. The major peak has the mass expected for the desired product: observed mass 12452.1 ± 2.0 Da, calculated mass 12449.3 Da (average isotope composition). The later eluting shoulder, g, results from ligation of the first and third peptide segments (MIF 1−59 and MIF 81−115).

residues[10] (Figure 2A). The next two examples were mature polypeptide chains corresponding to proteins of biological interest. The 74-residue polypeptide of human C5a,[11] a small protein belonging to the complement system, was synthesized from three segments consisting of residues 1-20, 21-46, and 47-74[12] (Figure 2B). Similarly, the 115-residue polypeptide of the protein macrophage migration inhibitory factor (MIF), a central mediator of inflamatory response,[13] was also made from three segments consisting of residues 1-59, 60-80, and 81-115[14] (Figure 2C). In each of these examples, the crude products contained a high proportion of the target polypeptide and were less complex mixtures than could have been achieved by traditional synthetic methods.

SPCL in the C-to-N Direction (Scheme 2). A generalized strategy for SPCL in the C-to-N direction, by consecutive reaction in aqueous solution, of unprotected peptide segments is outlined in Scheme 2. The strategy employs modification of the *C-terminal* peptide segment to include a cleavable linker functionalized with a group capable of chemoselective reaction with a suitably derivatized water-compatible polymer support. The chemistry for the preparation of a peptide-polymer support linker for this strategy is shown in Scheme 4.

The first (i.e., C-terminal) peptide segment (bearing an N-terminal Cys residue) is coupled to the polymer support, and the next peptide segment (bearing a protected N-terminal Cys and a C-terminal thioester) is added under native chemical ligation conditions (pH 7.0, 1% thiophenol).[3e] Once ligation is complete, unreacted reagents and soluble coproducts are washed away, and the protecting group on the N-terminal Cys is then removed. The growing polymer support-bound peptide is now ready for ligation of the next peptide segment. Consecutive cycles of ligation and deprotection can be performed in this manner to generate the full-length polymer-bound target polypeptide. After the ligation of the final peptide segment is complete, the linker to the polymer support is cleaved, releasing the full-length polypeptide product and any polymer-bound coproducts. After purification, the target polypeptide can be folded to give a fully functional protein molecule.

The strategy for binding the first peptide segment to the solid support must be compatible with the chemical synthesis of the unprotected peptide segments, the conditions of native chemical ligation, and the conditions used for removal of the temporary protecting group on the N-terminal Cys. Oxime formation was chosen as our coupling chemistry.[1c,7,9] Oxime-forming chemoselective reaction proceeds cleanly at low pH (4.6) where amino acid side chain functional groups are unreactive.[1c] Both the aminooxy group and ketone required for the oxime formation can be easily incorporated into either the polymer support used for SPCL or the C-terminal peptide segment. We chose to add the aminooxy group to the water-compatible polymer support; this was simply done by coupling *N*-Boc-aminooxyacetic acid[1c] to the amine of the polymer support and then removing the Boc protection. Incorporation of the ketone into the peptide segment is outlined in Scheme 4.

(10) The target polypeptide sequence was LTEGLHGFHVHEFGDNTAGCTSAGPHFNPLSRKHGCGFRVREFGDNTAVCADPSEEWVQKYVSDLELSA. The sequence of the N-terminal peptide segment was LTEGLHGFHVHEFGDNTAGCTSAGPHFNPLSRKHG-$^{\alpha}$COSH; the sequence of the middle peptide segment was CGFRVREFGDNTAV-$^{\alpha}$COSH; the sequence of the C-terminal segment was CADPSEEWVQKYVSDLELSA.

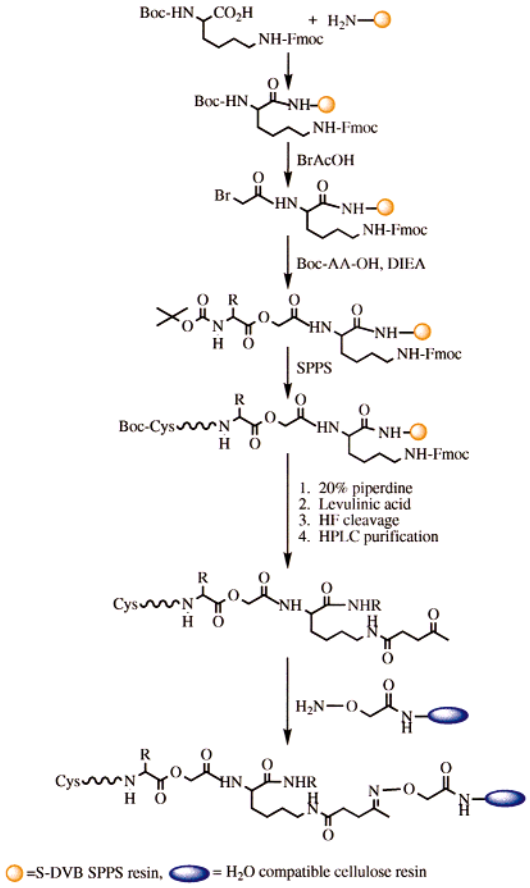
(11) Mollison, K. W.; Mandecki, W.; Zuiderweg, E. R. P.; Fayer, L.; Fey, T. A.; Krause, R. A.; Conway, R. G.; Miller, L.; Edalji, R. P.; Shallcross, M. A.; Lane, B.; Fox, J. L.; Greer, J.; Carter, G. W. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 292−296.

(12) The sequence of C5a(1−20) was TLQKKIEEIAAKYKHSVVKK-$^{\alpha}$COSH. The sequence of C5a(21−46) was CCYDGACVNNDETCEQRAARISLGPK-$^{\alpha}$COSH. The sequence of C5a(47−74) was CIKAFTECCVVASQLRANISHKDMQLGR.

(13) Bernhagen, J.; Mitchell, R. A.; Calandra, T.; Voelter, W.; Cerami, A.; Bucala, R. *Biochemistry* **1994**, *33*, 14144−14155.

(14) The sequence of MIF(1−59) was MPMFIVNTNVPRASVPDGFLSELTQQLAQATGKPPQYIAVHVVPDQLMAFGGSSEPCAL-$^{\alpha}$*COSH*. The sequence of MIF(60−80) was CSLHSIGKIGGAQNRSYSKLL-$^{\alpha}$*COSH*. The sequence of MIF(81−115) was CGLLAERLRISPDRVYINYYDMNAASVGWNNSTFA.

Scheme 4. Preparation of the Polymer-Linker Unprotected C-Terminal Peptide Segment



The temporary protecting group for the N-terminal Cys residue of the middle peptide segments had to be stable to the conditions of Boc chemistry SPPS used for synthesis of the otherwise unprotected peptide-$^{\alpha}$thioester segments, stable to the conditions of native chemical ligation, and yet readily and quantitatively removed when desired. The acetamidomethyl (Acm) [$-CH_2NHCOCH_3$] moiety,[8] used as a protecting group for the side chain thiol of the Cys residues, met all these criteria well. Cys(Acm) is stable to TFA and HF, is stable to nucleophiles and thiols under aqueous ligation conditions; and is readily converted to Cys($-$SH) by treatment with mercuric acetate.[8]

Lastly, the cleavable linker used to modify the first peptide segment for coupling to the water-compatible polymer support had to be stable to Boc chemistry SPPS, stable to the conditions of native chemical ligation and Acm removal, and yet easily and cleanly cleaved after ligations are complete to give the target polypeptide. The linker we designed for this purpose was based on the carboxyamidomethyl (CAM) [$-OCH_2CONH_2$] protecting group for carboxylic acids.[8] The CAM linker meets all the required criteria (stability to TFA and HF, stability to thiols under ligation conditions, stability to cleavage of the Cys(Acm) group) and can be removed in a matter of minutes by treatment with aqueous base.

A 27-residue model peptide was made from three unprotected peptide segments to test the feasibility of the C-to-N SPCL approach. The C-terminal peptide-*O*-CAM-Lys(levulinate)$\cdot$ amide segment was attached to the water-compatible aminooxyacetyl-polymer by oxime-forming chemoselective reaction. The second Cys(Acm)peptide$^{\alpha}$COSR segment was reacted with the polymer support-bound C-terminal peptide by native chemical ligation, and the Acm group was subsequently removed. The final peptide$^{\alpha}$COSR segment was then reacted with the polymer support-bound intermediate product. After each ligation was complete, a sample of the product was cleaved from the polymer support in aqueous base. The analytical data in Figure 3 demonstrated the efficacy of the combination of chemical tactics used and showed efficient synthesis of the desired polypeptide products by the N-to-C SPCL approach.

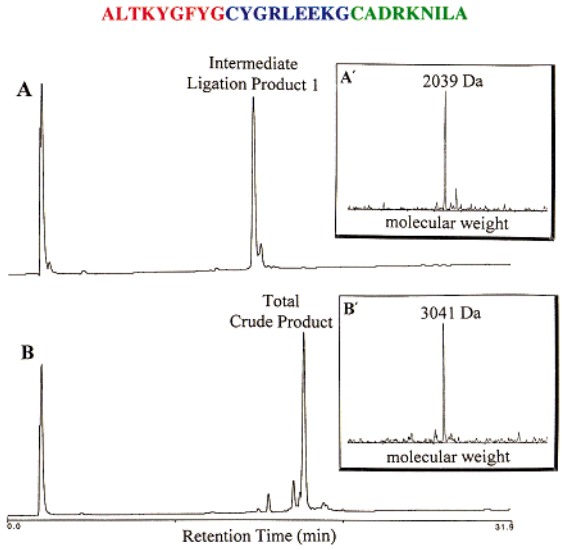


Figure 3. C-to-N strategy SPCL synthesis of a model peptide system. (A) Analytical HPLC of the crude first ligation product after base cleavage from the solid support. The sequence of the target polypeptide and the segments used are shown in the figure. (A′) Hypermass reconstruction of the raw MS data of the main HPLC peak to a single charge state. (B) Analytical HPLC of the crude second and final ligation product after base cleavage from the solid support. (B′) Hypermass reconstruction of the raw MS data of the main HPLC peak to a single charge state.

Solid Phase Protein Synthesis. The human group V secretory phospholipase A2 (GV-PLA2)[15] was chosen to illustrate the total synthesis of a protein by solid phase chemical ligation. GV-PLA_2 is a protein comprising 118 amino acids and containing 6 disulfide bonds; it belongs to a superfamily of enzymes responsible for the hydrolysis of the fatty acid esterified at the *sn-2* position of glycerophospholipids.[16]

The sequence was divided into four peptide segments at suitably located Cys residues for assembly by SPCL using a total of three native chemical ligation reactions. The segments ranged in length from 25 to 33 amino acids and consisted of the sequences GV-PLA_2(1$-$25), GV-PLA_2(26$-$58), GV-PLA_2-(59$-$87), and GV-PLA_2(88$-$118). The procedures used for this larger target were identical to those used for the model system. All peptide solutions ranged in concentration from 11 to 14 mM. Figure 4 shows the analytical data for the intermediate products and for the final ligation product. Electrospray MS of the main peak in the final crude product mixture showed this to be the GV-PLA_2(1$-$118) polypeptide chain [observed mass 13595.3 $\pm$ 1.6 Da, calculated mass 13590.7 Da (average isotopic composition)].

The principal component, corresponding to full-length GV-PLA_2(1$-$118) polypeptide, was purified from the crude product mixture and folded to give a protein with full enzymatic activity (see the Supporting Information).

(15) Balboa, M. A.; Balsinde, J.; Winstead, M. V.; Tischfield, J. A.; Dennis, E. A. *J. Biol. Chem.* **1996**, *271*, 32381$-$32384.
(16) Dennis, E. A. *Trends Biochem. Sci.* **1999**, *22*, 1$-$2; Chen, Y.; Dennis E. A. *Biochim. Biophys. Acta* **1998**, *1394*, 57$-$64.

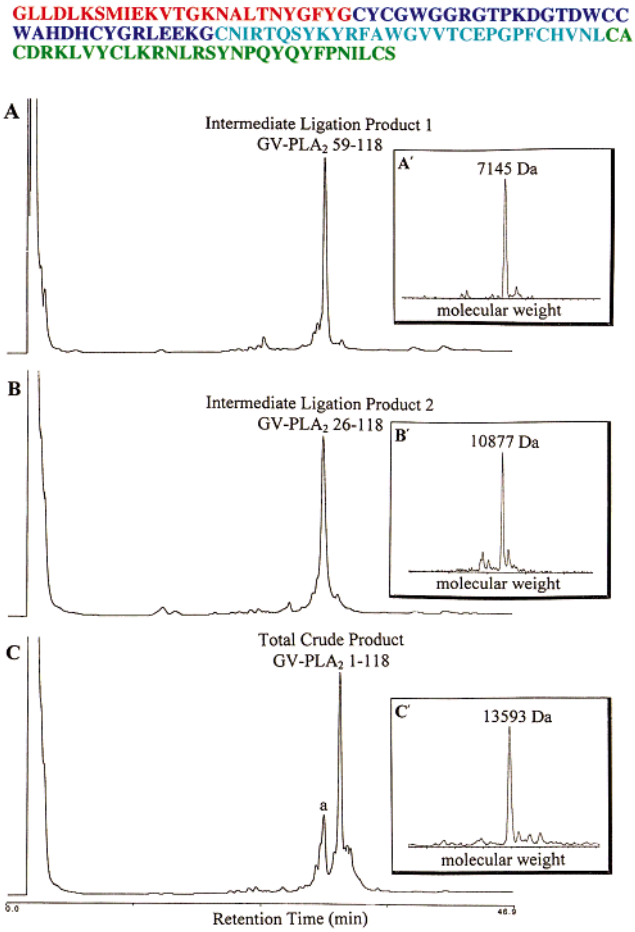


Figure 4. C-to-N strategy SPCL synthesis of GV-PLA$_2$(1−118). The sequence of the target polypeptide and the segments used are shown in the figure. (A) Analytical HPLC of the crude, first ligation product GV-PLA$_2$(59−118) after cleavage from the polymer support. (A′) Hypermass reconstruction of the raw MS data of the main HPLC peak to a single charge state. (B) Analytical HPLC of the crude second ligation product GV-PLA$_2$(26−118) after cleavage from the polymer support. (B′) Hypermass reconstruction of the raw MS data of the main HPLC peak to a single charge state. (C) Analytical HPLC of the crude third and final ligation product GV-PLA$_2$(1−118) after cleavage from the polymer support. Peak a represents a deletion of the second peptide segment GV-PLA$_2$(59−87) and has a mass of 10220 Da. (C′) Electrospray mass spectrometry characterization of the main HPLC peak, hypermass reconstruction to a single charge state of the raw MS data.

Accelerated Solid-Phase Protein Synthesis by SPCL. It was expected that using higher concentrations of the soluble peptide-$^{\alpha}$thioester segments would lead to faster native chemical ligation reactions with the polymer support-bound peptide chain, and thus accelerate the overall syntheses. To test the feasibility of this approach, we repeated the SPCL synthesis of the GV-PLA$_2$ polypeptide chain. Solutions of the incoming peptide segments [GV-PLA$_2$(59−87), -(26−58), -(1−25)] were made as concentrated as possible depending on the solubility of each peptide segment in the pH 7.0 guanidine·HCl buffer used for ligation. The segment GV-PLA$_2$(59−87)$^{\alpha}$COSR was at a concentration of 27 mM, GV-PLA$_2$(26−58)$^{\alpha}$COSR at a concentration of 29 mM, and GV-PLA$_2$(1−25)$^{\alpha}$COSR at a concentration of 50 mM. Each segment was in at least a 10-fold excess over the amount of polymer-bound polypeptide chain used in the reactions. The procedures followed were identical to those outlined above, except each of the three ligation reactions was stopped after only 1 h. The cycle time for the addition of one segment to the polymer-bound peptide (i.e., the ligation reaction, ACM removal, and all washing steps) was approximately 2.5 h, and all three ligations could be completed in an 8 h period.

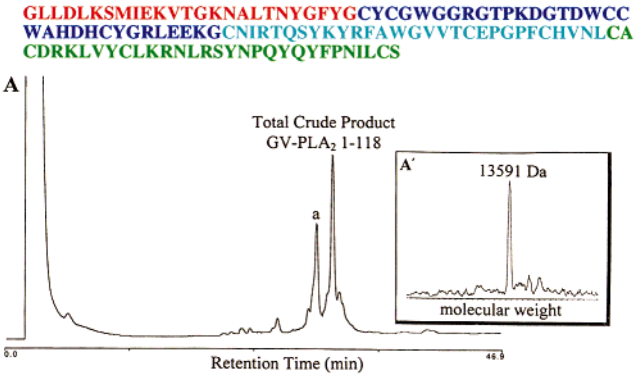


Figure 5. Accelerated C-to-N SPCL synthesis solid-phase chemical ligations to form full-length GV-PLA$_2$. The sequence of the target polypeptide and the segments used are shown in the figure. (A) Analytical HPLC of the crude ligation products from the rapid SPCL assembly of [GV-PLA$_2$(1−118)], after cleavage from the solid support. Peak a represents a deletion of the second peptide segment [(GV-PLA$_2$-(59−87)] and has a mass of 10220 Da. (A′) Hypermass reconstruction of the raw MS data from the main HPLC peak to a single charge state.

The HPLC of the cleaved final ligation products is shown in Figure 5. There are two major peaks. The largest is the desired product, full-length GV-PLA$_2$ [obsd 13592.1 ± 2.3 Da, calcd 13590.7 Da (average isotopic composition)]. The smaller peak is the ligation product that results from the deletion of GV-PLA$_2$(59−87) [i.e., ligation of GV-PLA$_2$(1−25) and (26−58) to GV-PLA$_2$(88−118)], having a mass as observed by ESMS of 10219.4 ± 0.9 Da [calcd 10218.8 Da (average isotopic composition)]. This is an indication that the first ligation was slower than the other two ligations and did not go to completion in the 1 h reaction time allowed. That particular ligation site is Leu-Cys as opposed to the Gly-Cys ligation sites for addition of the other two peptide segments. Steric hinderance associated with the Leu side chain could be the reason for the slower ligation of the second segment. These experiments have demonstrated that the SPCL method shows considerable promise for providing more rapid access to proteins.

Conclusions

The principles of polymer-supported organic synthesis, originally enunciated by Merrifield[17,18] for stepwise solid phase synthesis of peptides, thus are seen to apply equally well to solid phase protein synthesis using SPCL. Specifically, the SPCL method provides for rapid purification of the intermediate polymer support-bound ligation products. Compared with synthesis in solution, this feature of SPCL facilitates the changes in solvent and other conditions necessary for different steps of the synthetic cycle. Similarly, the ease of removal of excess reactant enables faster reaction rates from the use of higher concentrations of each peptide segment. In addition, potential solubility problems for the intermediate product polypeptides will be reduced because their attachment to a swollen, cross-linked polymeric support reduces the tendency to form aggregates.[19]

SPCL thus allows the facile consecutive ligation of several peptide segments and potentially provides for the synthesis of

(17) Merrifield, R.B. *J. Am. Chem. Soc.* **1963**, *85*, 2149−2154.
(18) Merrifield, R.B. *Science* **1986**, *232*, 341−347.
(19) Kent, S. B. H. *Annu. Rev. Biochem.* **1998**, *57*, 957−984.

much larger polypeptide targets.[20] We hope that the results reported here will provide impetus for chemists to build on the SPCL concept to provide ever more facile synthetic access to the world of biologically active proteins. For example, it can be anticipated that scale up will be readily possible by the use of high loadings of peptide on the polymer support. Also, the repetitive nature of the operations in the SPCL method makes it possible to envision automation of the process. The chemistries described in this paper are general; they provide for the assembly of peptides of any sequence, even segments which contain multiple cysteine residues as evidenced by several of the syntheses reported here. Other chemical tactics and other chain assembly strategies can be implemented within the SPCL concept. For example, the peptide-$^{\alpha}$COSR segments could be generated by recombinant means[21] and used in solid-phase protein synthesis by SPCL.

The ability to make proteins directly from gene sequence data is becoming of increasing significance in the emerging "post-genome" era of biomedical research. Chemistry has a pivotal role to play in providing direct access to functional protein molecules that are otherwise known only as predicted amino acid sequences in a computer. The solid phase chemical ligation method described here represents an important extension of the existing armamentarium for total chemical protein synthesis.

Materials and Methods

Peptide Segment Synthesis. Peptides were synthesized in stepwise fashion on a modified ABI 430A peptide synthesizer by established machine-assisted solid phase methods using *in situ* neutralization/HBTU activation protocols for Boc chemistry[22] on Boc-aminoacyl-OCH_2-PAM resins, thioester-generating resins,[23] or -$^{\alpha}$COSH-generating resins[24] as appropriate. The modified N-terminal segment for the N-to-C strategy (Scheme 1) was synthesized as previously described.[7] The modified C-terminal segment for the C-to-N strategy was synthesized as outlined in Scheme 4. After chain assembly was complete, peptides were deprotected and simultaneously cleaved from the resin by treatment with anhydrous HF containing 5% *p*-cresol, lyophilized, and purified by preparative HPLC. Boc-protected amino acids were obtained from Peptides International (Louisville, KY) or NovaBiochem (San Diego, CA). Protection for amino acids was as follows: Arg(Tos), Asp(OChx), Asn(Xan), Cys(pMeBzl) or Cys(ACM), Glu(OChx), His(DNP), Lys-(2ClZ) or Lys(Fmoc), Ser(Bzl), Thr(Bzl), Tyr(2BrZ), Trp(Hoc). DMF and DCM were HPLC grade and used as received. Trifluoroacetic acid was obtained from Halocarbon (River Edge, NJ). Other chemicals were from Fluka (Ronkonkoma, NY) or Aldrich (Milwaukee, WI) and used as received.

Analytical or preparative reversed-phase HPLC was performed on a Rainin HPLC system with 214 nm UV detection, using Vydac C18 analytical or preparative columns. Chromatographic separations were achieved using linear gradients of buffer B in buffer A over 30 or 60 min at l mL/min (analytical) or 15 mL/min (preparative). Buffer A = 0.1% TFA in water; buffer B = 0.1% TFA in acetonitrile. Peptide and protein mass spectrometry was performed on a Sciex API-I electrospray mass spectrometer. Observed masses were derived from the experimental *m/z* values for all observed protonation states of a molecular species, using the program MacSpec (Sciex). Calculated masses were based on the average isotopic composition and were derived using the program MacProMass (Terry Lee and Sunil Vemuri, Duarte, CA).

All synthetic peptide segments were greater than 95% pure by analytical reversed-phase HPLC, and each had the expected mass (±1 part in 10000) compared with the calculated mass (average isotope composition), by electrospray mass spectrometry.

Polymer Support Preparation. For both the N-to-C and C-to-N strategies an amino Spherilose affinity resin was used as polymer support (Isco, Lincoln, NE). The resin was derivatized with Boc-aminooxyacetic acid and was thoroughly washed with DMF. The Boc-aminooxyacetic acid was activated as the succinimide ester[9] and coupled to the amino Spherilose resin for 2−3 h. After coupling was complete as monitored by ninhydrin assay, the resin was washed with DMF, the Boc group was removed by treatment with neat TFA (2 min), and the resin was neutralized with a 10% solution of DIEA in DMF. After further washing with DMF, followed by DCM, the derivatized polymer support was thoroughly dried.

Coupling of the Modified Terminal Peptide Segment to the Polymer Support. The ketone-containing peptide appropriate to the strategy being used (Schemes 1−4) was dissolved in 6 M guanidine·HCl, 0.1 M sodium acetate, and 0.15 M methionine, pH 4.6 (~5 mM peptide) and added to dry aminooxy-functionalized polymer support and allowed to react at room temperature overnight.

Ligation Reactions. The peptide segment to be ligated to the polymer support-bound peptide was dissolved in 6 M guanidine·HCl, 0.1 M sodium phosphate, pH 7.0, 0.15 M methionine, and 1% thiophenol, (10−50 mM peptide) and added to the polymer support, which had been thoroughly washed in the same buffer, and allowed to react at room temperature.

(N-to-C Strategy) Conversion of the Polymer Support-Peptide-$^{\alpha}$Thiocarboxylate to a Thioester. The polymer support-bound peptide-$^{\alpha}$thioacid was thoroughly washed in 6 M guanidine·HCl, O.1 M sodium acetate, and 0.15 M methionine, pH 4.6, and treated with a 50 mM solution of bromoacetic acid in the same buffer for 15 min, followed by thorough washing with the pH 4.6 buffer.

(C-to-N Strategy) Cys(Acm) Deprotection. After ligation of an incoming peptide segment was complete, the Acm group on the N-terminal Cys residue was removed. The resin was washed with 3% aqueous acetic acid followed by treatment for 30 min with a solution of mercury(II) acetate (15 mg/mL) in the acetic acid solution. After thorough washing of the resin with the acetic acid solution followed by 6 M guanidine·HCl and 0.1 M sodium phosphate, pH 7.0, the resin was treated with 20% β-mercaptoethanol in the same pH 7.0 buffer for 30 min.

Cleavage of the Polymer-Msc-Peptide Support (N-to-C Strategy). Polymer support-bound peptide was treated with 6 M guanidine·HCl, 0.1 M sodium acetate, 0.15 M methionine, and 200 mM hydrazine, pH ≈ 14, for 2 min, followed by washing with an equal volume of 6 M guanidine·HCl, 0. 1 M sodium acetate, 0.15 M methionine, and 200 mM hydrazine, pH ≈ 2, and with an equal volume of 6 M guanidine·HCl, 0.1 M sodium acetate, and 0.15 M methionine, pH 4.6. The combined eluants containing the free peptide were subjected to preparative reversed-phase HPLC to yield the purified ligation product.

Cleavage of the Peptide-CAM-Polymer Support (C-to-N Strategy). Polymer support-bound peptide was washed with 8 M urea and 0.1 M sodium phosphate, pH 7.0. The peptide-*O*-CAM-resin was then treated with 0.25 N NaOH in the same urea buffer (resulting pH 12−14) for 2 min, followed by washing with an equal amount of 0.25 N HC1 in the urea buffer (resulting pH 1−2), and then washing with an equal amount the urea buffer, pH 7.0. The combined eluants containing free peptide were purified by semipreparative HPLC and analyzed by electrospray mass spectrometry.

Acknowledgment. We gratefully acknowledge the assistance of Darren Thompson, Jay J. Levy, and Harold P. Amogan in the syntheses of peptide segments, and of our other colleagues at Gryphon Sciences for helpful discussions. The work on GVPLA2 was supported in part by NIH Grant GM20501 and Y.C. was supported by NIH/NCI T32CA 09523.

(20) To date, the largest polypeptide synthesized by the SPCL method described above has involved the chemical ligation of seven unprotected peptide segments to give a product of 28 kDa (Wuyuan Lu, personal communication).

(21) Xu, R.; Ayers, B.; Cowburn, D.; Muir, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 388−393.

(22) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180−193.

(23) (a) Hojo, H.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111−117. (b) Hojo, H.; Kwon, Y.; Kakuta, Y.; Tsuda, S.; Tanaka, I.; Hikichi, K.; Aimoto, *S. Bull. Chem. Soc. Jpn.* **1993**, *66*, 2700−06.

(24) (a) Yamashiro, D.; Li, C. H. *Int. J Pept. Protein Res.* **1988**, *31*, 322−34. (b) Canne, L. E.; Walker, S. M.; Kent, S. B. H. *Tetrahedron Lett.* **1995**, *36*, 1217−1220.

Supporting Information Available: Details of the folding and activity assays for synthetic group V phospholipase A2, including Materials and Methods, Results and Discussion, and a figure and table showing the enzymatic activities obtained (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA9836287