Conformationally Assisted Protein Ligation Using C-Terminal Thioester Peptides

Gangamani S. Beligere and Philip E. Dawson*

Departments of Cell Biology and Chemistry
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
La Jolla, California 92037

Received March 11, 1999

The selective coupling of unprotected peptide fragments is an essential element of many approaches for the total and semisynthesis of proteins.[1] This coupling can be achieved through highly chemoselective reactions of peptides in denaturing aqueous solutions[2] (Figure 1A). A complementary approach has been to use the protein conformation to obtain a regioselective coupling reaction. This method takes advantage of the observation that many proteins remain folded when the polypeptide chain is cleaved by chemical cleavage[3] or limited proteolysis.[4] In this folded conformation, the new peptide N and C termini are placed in close proximity. Semisynthetic proteins have been assembled through reaction of an N-terminal amine with a C-terminal homoserine lactone[1a,3] or by reverse proteolysis using natural[5] or engineered proteases[6] (Figure 1B,C). This approach has been used to make semisynthetic derivatives of both cytochrome *c*[1a,7] and ribonuclease A.[5] Unfortunately, conformationally assisted reactions have been limited due to slow reaction rates.[8]

Recently, peptide thioesters have been utilized for the chemical synthesis of proteins[9,10] and peptides[11] due to a rapid and specific reaction with N-terminal cysteine residues that generates a native peptide bond.[12] One limitation of this "native chemical ligation" approach is the requirement for a naturally occurring or introduced cysteine residue at the site of ligation.[13] Here, we demonstrate that the requirement for a cysteine residue can be eliminated by using a combination of C-α-thioester peptides and conformational assistance.

Chymotrypsin inhibitor 2 (CI2) is a 64 amino acid protein with no disulfide bonds that has been used extensively in studies on protein folding and stability.[14] It has been demonstrated that fragments corresponding to CI2(1−40) and CI2(41−64) can self-associate to form a stable protein with a three-dimensional structure similar to that of the full length protein.[15] Initially, we adapted this system for synthesis by conventional native chemical ligation by altering the sequence slightly, replacing residue 39 with aspartic acid and residue 40 with cysteine.[16] With this sequence alteration, CI2(1−64)T39D/M40C could be synthesized from peptides corresponding to CI2(1−39)T39D-COSR and CI2-(40−64)M40C. The full-length product was synthesized using standard native chemical ligation procedures: ∼1 mM peptide in 6 M Gdn·HCl, 100 mM sodium phosphate, pH 6.3, 2% thiophenol.[17] In these conditions, the peptides are expected to be fully denatured. The reaction proceeded cleanly over 8 h to produce the desired 64 amino acid CI2 polypeptide chain (Table 1-H).

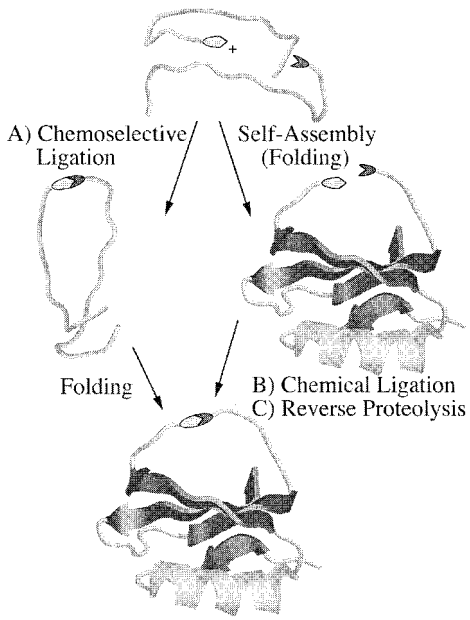

Figure 1. Strategies for the chemical synthesis of proteins.

We hypothesized that this ligation rate would be enhanced if the peptides self-associated to bring the N and C termini into close proximity.[18] To test for conformational assistance in this reaction, the same peptides were reacted in folding conditions of 100 mM sodium phosphate, pH 6.3 and 1 mM of each peptide. The peptides were completely ligated in under 30 min (Table 1-A). This rapid ligation reaction also allowed the efficient synthesis of CI2(1−64)M40C with the sterically hindered amino acid threonine at position 39. In this case, the ligation was complete in 30 min (Table 1-B) compared to ∼48 h for model studies under denaturing conditions.[13] Following purification and folding, the final polypeptide products were fully active as potent inhibitors of subtilisin.

The dramatic enhancement of the ligation reaction posed the question of whether the N-terminal cysteine was necessary for the conformationally assisted reaction. To test this, we replaced the N-terminal cysteine with the native methionine residue at position 40.[19] As shown in Table 1-C, the ligation reaction was complete in just 2 h. In contrast, the same two peptides under

(1) (a) Wallace, C. J. A. *Curr. Opin. Biotech.* **1995**, *6*, 403−410. (b) Wilken, J.; Kent, S. B. H. *Curr. Opin. Biotech.* **1998**, *9*, 412−426.
(2) Muir, T. W.; Dawson, P. E.; Kent, S. B. H. *Methods Enzymol.* **1997**, *289*, 266−298.
(3) Corradin, G.; Harbury, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 3036−3039.
(4) Richards, F. M.; Vithayathil, P. J. *J. Biol. Chem.* **1959**, *234*, 1459−1465.
(5) Homandberg, G. A.; Laskowski, M. *Biochemistry* **1979**, *18*, 586−592.
(6) Suich, D. J.; Ballinger, M. D.; Wells, J. A.; DeGrado, W. F. *Tetrahedron Lett.* **1996**, *37*, 6653−6656.
(7) Wuttke, D. S.; Gray, H. B.; Fisher, S. L.; Imperiali, B. *J. Am. Chem. Soc.* **1993**, *113*, 8455−8456.
(8) One exception is an efficient reassembly of fragments from triose phosphate isomerase, demonstrated by: Vogel, K.; Chmielewski, J. *J. Am. Chem. Soc.* **1994**, *116*, 11163−11164.
(9) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
(10) Dawson, P. E. *Methods Enzymol.* **1997**, *287*, 34−45.
(11) Tam, J. P.; Lu, Y. A.; Chuan-Fa, L.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485−12489.
(12) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. *Liebigs Ann. Chem.* **1953**, *583*, 129−149.
(13) It has recently been demonstrated that ligation reactions can generate X-Cys sites where X is any amino acid except Pro. Additionally Thr, Val, and Ile give slow (>48 h) ligation rates. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. submitted for publication, *Proc. Natl. Acad. Sci. U.S.A.*
(14) Fersht, A. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 3−9.
(15) Gay, G. P.; Fersht, A. R. *Biochemistry* **1994**, *33*, 7958−7963.
(16) The T39D mutation has been shown by Fersht and co-workers to have little effect on the folding and stability of CI2: Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 260−288.
(17) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. *J. Am. Chem. Soc.* **1997**. *119*, 4325−4329. Note: the ligation reactions were performed at a final pH of ∼5 due to TFA from the lyophilized peptide and thiophenol additives.
(18) Similar native ligation rate enhancements have been described with coiled coil sequences: Lee, D. L.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. *Nature* **1996**, *382*, 525−528. In addition, a protein cyclization reaction with a rapid ligation rate has been described. Camarero, J. A.; Pavel, J.; Muir, T. W. *Angew. Chem., Int. Ed.* **1998**, *37*, 347−349.

10.1021/ja9907919 CCC: $18.00
Published on Web 06/17/1999

Table 1. Ligation Reactions under Folding and Denaturing Conditions[a]

	C-terminal thioester Peptide	N-terminal cysteine Peptide	product	t_{50}	t_F
A	(1−39)T39D	(40−64)M40C	(1−64) T39D/ M40C/	2 min	30 min
B	(1−39)	(40−64)M40C	(1−64) M40C	2 min	30 min
C	(1−39) T39D	(40−64)	(1−64) T39D	30 min	2 h
D	(1−39) T39D/ A16G	(40−64) M40C	(1−64) T39D/A16G/ M40C	2 min	30 min
E	(1−39) T39D/A16G	(40−64)	(1−64) T39D/ A16G	30 min	2 h
F	S-pep	S-pro	RNase A	2 hr	10 hr
G	(1−39)T39D	(41−64)	no product		
H	(1−39) T39D	(40−64) M40C	(1−64) T39D/M40C	2 h	8 h
I	(1−39) T39D	(40−64)	no product		
J	S-pep	S-pro	no product		

[a] For entries A−G the ligation was carried out under folding conditions in 100 mM sodium phosphate, pH 6.3, 2% (v/v) thiophenol with 1 mM peptide concentration.[17] [b] Ligation buffer was 6 M guanidine hydrochloride in 100 mM sodium phosphate, pH 6.3, 2% (v/v) thiophenol with 1 mM peptide concentration.[17]

denaturing conditions produced no observable ligation after 24 h (Table 1-I). In addition, the ligation of CI2(1−39)D39T-COSR with the shortened sequence CI2(41−64) failed to ligate under the same folding conditions (Table 1-G).

To be a synthetically useful technique, this conformationally assisted ligation reaction must be compatible with chemically modified or "mutated" protein sequences that destabilize the protein. Detailed thermodynamic and kinetic studies of CI2 by Fersht and co-workers have characterized a modification in the α helix, A16G, that destabilizes both the folded state and the "transition state" of folding by ∼1.2 kcal/mol.[14] As a result, this protein analogue represents a good test for the utility of the conformationally assisted ligation. As shown in Table 1-E, the reaction of CI2(1−39)A16G/T39D-COSR with CI2(40−64) (N-terminal Met) was complete in 2 h, comparable to the ligation with CI2(1−39)T39D-COSR.

Finally, the generality of this approach for protein synthesis was demonstrated through a semisynthesis of ribonuclease A. The use of limited subtilisin digestion to cleave between alanine 20 and serine 21 to produce self-associating fragments corresponding to RNase S (21−124) and the S-peptide(1−20) has been well characterized.[4,5,20] The conformationally assisted ligation reaction consisted of 1 mM RNaseA(21−124) obtained from subtilisin cleavage of bovine RNaseA, 1 mM synthetic RNaseA(1−20)-COSR, in 100 mM sodium phosphate, pH 6.3, 2% thiophenol. The ligation reaction was complete in 10 h (Table 1-F). The resulting protein was found to have catalytic activity similar to that of biologically derived bovine RNaseA.

One limitation to this conformationally assisted ligation approach appears to be the use of Xaa-Yaa ligation sites in which Xaa is a β-branched amino acid and Yaa is not a cysteine.[13] For example, the conformationally assisted ligation of the fully native sequence peptides CI2(1−39)-COSR and CI2(40−64) to form a ^{39}Thr^{40}Met ligation site was extremely slow. However, as discussed earlier, when a C-terminal Thr thioester peptide is ligated to an N-terminal Cys peptide using conformational assistance, the ligation reaction was highly efficient (Table 1B).

We have demonstrated that synthetic C-α-thioester peptides can be ligated to both biologically and chemically derived peptides using the conformational assistance of the protein fold. Recent studies have demonstrated that C-α-thioester peptides can be generated from biological expression using modified protein-splicing fusion proteins[21] and subsequently ligated to N-terminal cysteine-containing peptides. Analogously, semisynthesis of proteins should be possible using conformationally assisted ligation of a biologically expressed N-terminal fragment (α-thioester) and a chemically synthesized C-terminal fragment. This synthetic flexibility should broaden the application of conformationally assisted ligation in semisynthetic systems.

In conclusion, conformational assistance and weakly activated C-α-thioester peptides can be used in combination to achieve a rapid and specific coupling of unprotected peptides. Peptide bond formation does not require an N-terminal cysteine, and ligation between Asp/Met, and Ala/Ser was performed efficiently. This suggests that a variety of ligation junctions should be compatible with this approach. In addition, reactions with N-terminal cysteine residues occurred on a time scale of minutes, allowing difficult couplings such as Thr-Cys to be performed in a straightforward manner. Successful ligation of peptides in two unrelated protein systems, CI2 and RNaseA and a destabilized mutant of CI2 attest to the flexibility of the method.

Many proteins that adopt a folded conformation, following a single chemical or proteolytic cleavage have been characterized.[22] In addition, proteins that remain folded following multiple proteolytic cleavages have been described.[7] The structural diversity of these proteins indicates that self-assembling peptides may be found in many protein systems, especially where the cleavage site is in a loop region of the protein. Consequently, conformationally assisted ligation using C-α-thioester peptides may be a general route for the total and semisynthesis of proteins.

Acknowledgment. We gratefully acknowledge the Skaggs Institute for Chemical Biology for financial support.

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

JA9907919

(19) The thioether side chain of methionine should be unreactive towards the thioester peptide, leaving the terminal amine as the only accessible nucleophile in the folded conformation.

(20) Reverse proteolysis has been used to religate these fragments in 50% DMSO with a yield of ∼50% over 1 week[5]. A small (∼2%) impurity of RNaseS 22−124 is present due to an alternate cleavage site for subtilisin.

(21) (a) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710. (b) Evans, T. C.; Benner, J.; Xu, M.-Q. *Protein Sci.* **1998**, *7*, 2256−2264.

(22) (a) Chaffotte, A. F.; Li, J.-H.; Georgescu, R. E.; Goldberg, M. E.; Tasayco, M. L. *Biochemistry* **1997**, 160400−16048. (b) Honda, S.; Kobayashi, N.; Munekata, E.; Uedaira, H. *Biochemistry* **1999**, *38*, 1203−1213. (c) Taniuchi, H.; Anfinsen, C. B. *J. Biol. Chem.* **1971**, *246*, 2291−2301. (d) Shiba, K.; Schimmel, P. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1880−1884.