Protein Assembly by Orthogonal Chemical Ligation Methods

Bradley L. Nilsson,[†] Robert J. Hondal,[‡,§] Matthew B. Soellner,[†] and Ronald T. Raines[*,†,‡]

Departments of Chemistry and Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

Received December 14, 2002; E-mail: raines@biochem.wisc.edu

Chemical synthesis harbors the potential to provide ready access to natural proteins as well as to create nonnatural ones. Indeed, numerous proteins have already been assembled from synthetic peptides.[1,2] "Native chemical ligation"—the coupling of a peptide[3] (or protein[4]) containing a C-terminal thioester with another peptide containing an N-terminal cysteine residue—has been especially efficacious.[1] Emerging strategies for protein assembly avoid the need for a cysteine residue at the ligation junction. The Staudinger ligation is one such strategy.[5]

In our version of the Staudinger ligation, a peptide containing a C-terminal phosphinothioester reacts with a peptide containing an N-terminal azide to give an amide with no residual atoms, as in Scheme 1.[5b,c] The initial intermediate is an iminophosphorane, which rearranges to an amidophosphonium salt. Hydrolysis yields the amide and a phosphine oxide. Previously, we showed that this reaction can be used to form dipeptides in high isolated yield (>90%) and with the retention of α-carbon stereochemistry.[5b,c] Herein, we describe the first use of the Staudinger ligation to couple peptides on a solid support. We use the fragment thus produced to assemble a protein via native chemical ligation.[3,4] The synthesis of a protein by this route expands the versatility of chemical approaches to protein production.

Scheme 1
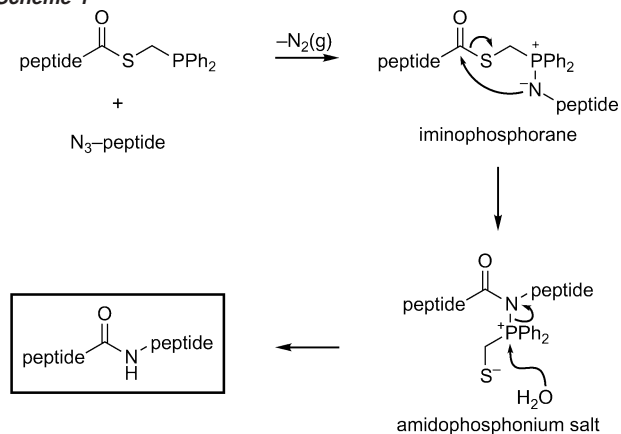
As a model protein, we chose ribonuclease A (RNase A;[6] 124 amino acid residues), which was the first protein to succumb to total synthesis.[7] Herein, RNase A was assembled from three fragments, which comprised residues 1−109, 110−111, and 112−124. The route is shown in Scheme 2.

RNase A(110−111) was synthesized as a C-terminal phosphino-thioester using a sulfonamide-linker ("safety-catch") resin.[8] The fully loaded resin was activated with iodoacetonitrile.[8b] Treatment with an excess of diphenylphosphinomethanethiol[5b,c] in DMF for

† Department of Chemistry.
‡ Department of Biochemistry.
§ Current address: Department of Biochemistry, College of Medicine, University of Vermont, Burlington, VT 05405.

18 h liberated FmocCys(Trt)Glu(OtBu)SCH₂PPh₂, which was isolated in 64% yield.

RNase A(112−124) was synthesized as an N-terminal azide. The n − 1 peptide was synthesized by using standard Fmoc-protection and HATU activation on a hydroxyethylpolystyrene (PEGA) resin, which has diverse solvent compatibility. α-Azido glycine[5a,b] (residue 112) was used to cap the n − 1 peptide by its activation with PyBOP, HOBt, and DIPEA in DMF. The identity of N₃CH₂C(O)-Asn(Trt)ProTyr(tBu)ValProValHis(Trt)Phe Asp(OtBu)AlaSer(tBu)-Val was confirmed by cleaving a small amount of the α-azido peptide from the resin with TFA and analyzing by MALDI mass spectrometry.

RNase A(110−111) and RNase A(112−124) were coupled by Staudinger ligation directly on the PEGA resin. Four equivalents of RNase A(110−111) in DMF/H₂O (10:1) was incubated with the resin over 12 h. After cleavage from the resin, side-chain depro-tection, and HPLC purification, RNase A(110−124) was isolated in 61% yield.

RNase A(110−124) was synthesized again by this route, now incorporating an NMR probe of protein structure. Specifically, [^{13}C′,^{13}Cα,^{15}N]proline[9] was inserted at position 114. The Asn113−Pro114 peptide bond resides in the cis (E) conformation in the properly folded protein, but in a mixture of cis and trans conforma-tions in peptide fragments.[10]

RNase A(1−109) was produced by biosynthesis as a C-terminal thioester with N-methylmercaptoacetamide, as described previous-ly.[4c,11] Both unlabeled and labeled RNase A(110−124) contained an N-terminal cysteine residue. Ligation of RNase A(1−109) and RNase A(110−124) in aqueous buffer, folding, and purification yielded intact RNase A. Its molecular mass was verified by MALDI mass spectrometry.

The ability to incorporate labeled amino acids at specific sites is a distinct advantage of producing proteins by chemical synthe-sis.[1,2] The route in Scheme 2 was used to incorporate [^{13}C′,^{13}Cα,^{15}N]-Pro114 into RNase A, and a 1D HSQC NMR experiment was used to probe for proper folding of the resulting protein. In Fmoc-[^{13}C′,^{13}Cα,^{15}N]ProOH, the carbamyl C−N bond was a mixture of cis and trans isomers (Figure 1). In contrast, the Asn113−[^{13}C′,^{13}Cα,^{15}N]Pro114 peptide bond in labeled RNase A was a single species, consistent with this C−N bond being only in the cis conformation. Moreover, the chemical shift of the α-proton of [^{13}C′,^{13}Cα,^{15}N]Pro114 in the synthetic RNase A was identical to that of unlabeled Pro114 in natural RNase A.[12]

Enzymatic activity provides an extremely sensitive measure of protein structure.[13] The enzymatic activity of the RNase A synthesized as in Scheme 2 ($k_{cat}/K_M = 0.94 \times 10^7$ M^{-1} s^{-1}) was nearly equal to that of the wild-type enzyme produced by recombinant DNA technology ($k_{cat}/K_M = 1.1 \times 10^7$ M^{-1} s^{-1}).[11,14]

Thus, the solid-phase assembly of peptides with the Staudinger ligation has not only been realized, but used to assemble a functional enzyme. This method for amide bond formation is orthogonal and complementary to other ligation methods.[5] The enzyme created

Scheme 2
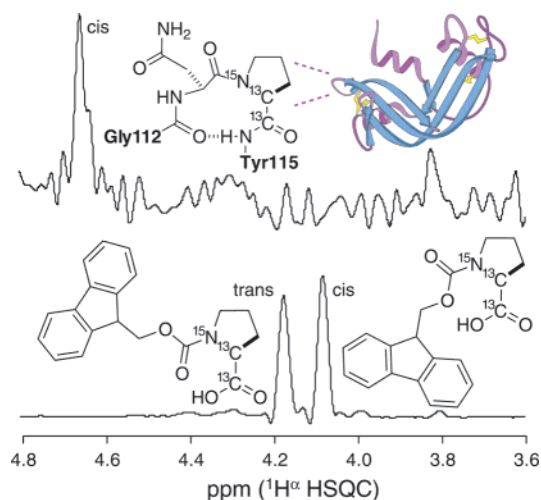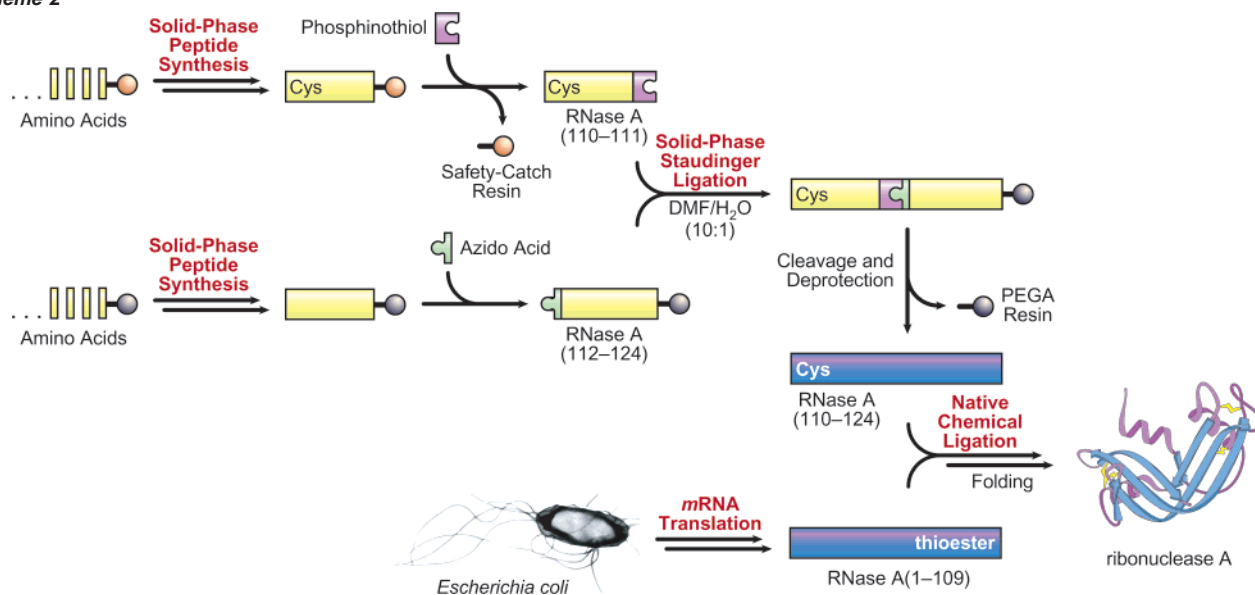
Figure 1. Observation of the α-proton of [^{13}C′,^{13}C$^{\alpha}$,^{15}N]proline by 1D ^{15}N-CT-HSQC NMR spectroscopy. Top: [^{13}C′,^{13}C$^{\alpha}$,^{15}N]Pro114 RNase A in D$_2$O (δ 4.66 ppm). Bottom: Fmoc[^{13}C′,^{13}C$^{\alpha}$,^{15}N]ProOH in D$_2$O/CD$_3$-OD (1:1), assigned by 2D HMBC NMR spectroscopy.

herein is remarkable in that its peptide bonds were synthesized in four distinct processes: *m*RNA translation by a ribosome, solid-phase peptide synthesis, native chemical ligation, and solid-phase Staudinger ligation (Scheme 2). We anticipate that the two solid-phase processes alone could be the basis for an automated means to assemble proteins, making enzymes and other proteins more accessible targets for synthetic chemistry.

Supporting Information Available: Procedures and additional data for syntheses and analysis reported herein (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.
(2) (a) Tam, J. P.; Yu, Q.; Miao, Z. *Biopolymers* **1999**, *51*, 311−332. (b) Borgia, J. A.; Fields, G. B. *Trends Biotechnol.* **2000**, *15*, 243−251. (c) Miranda, L. P.; Alewood, P. F. *Biopolymers* **2000**, *55*, 217−226. (d) Tam, J. P.; Xu, J.; Eom, K. D. *Biopolymers* **2001**, *60*, 194−205.
(3) (a) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. *Liebigs Ann. Chem.* **1953**, *583*, 129−149. (b) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
(4) (a) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710. (b) Severinov, K.; Muir, T. W. *J. Biol. Chem.* **1998**, *273*, 16205−16209. (c) Evans, T. C., Jr.; Benner, J.; Xu, M. Q. *Protein Sci.* **1998**, *7*, 2256−2264.
(5) For use of the Staudinger ligation to form peptide bonds, see: (a) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939−1941. (b) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2001**, *3*, 9−12. (c) Soellner, M. B.; Nilsson, B. L.; Raines, R. T. *J. Org. Chem.* **2002**, *67*, 4993−4996. For other important precedents, see: (d) Staudinger, H.; Meyer, J. *Helv. Chim. Acta* **1919**, *2*, 635−646. (e) Bosch, I.; Romea, P.; Urpí, F.; Vilarrasa, J. *Tetrahedron Lett.* **1993**, *34*, 4671−4674. (f) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007−2010. (g) Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. *Org. Lett.* **2000**, *2*, 2141−2143.
(6) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045−1066.
(7) (a) Gutte, B.; Merrifield, R. B. *J. Am. Chem. Soc.* **1969**, *91*, 501−502. (b) Hirschmann, R.; Nutt, R. F.; Veber, D. F.; Vitali, R. A.; Varga, S. L.; Jacob, T. A.; Holly, F. W.; Denkewalter, R. G. *J. Am. Chem. Soc.* **1969**, *91*, 506−507. (c) Merrifield, R. B. *Science* **1984**, *232*, 341−347.
(8) (a) Kenner, G. W.; McDermott, J. R.; Sheppard, R. C. *J. Chem. Soc., Chem. Commun.* **1971**, 636−637. (b) Backes, B. J.; Ellman, J. A. *J. Org. Chem.* **1999**, *64*, 2322−2330.
(9) [^{13}C′,^{13}C$^{\alpha}$,^{15}N]Proline was a gift from the National Stable Isotope Resource at Los Alamos (SIR).
(10) An, S. S. A.; Lester, C. C.; Peng, J.-L.; Li, Y.-J.; Rothwarf, D. M.; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc.* **1999**, *121*, 11558−11566.
(11) (a) Hondal, R. J.; Nilsson, B. L.; Raines R. T. *J. Am. Chem. Soc.* **2001**, *123*, 5140−5141. Also see: (b) Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. *J. Am. Chem. Soc.* **2002**, *124*, 8522−8523.
(12) Rico, M.; Bruix, M.; Santoro, J.; Gonzalez, C.; Neira, J. L.; Nieto, J. L.; Herranz, J. *Eur. J. Biochem.* **1989**, *183*, 623−638.
(13) Knowles, J. R. *Science* **1987**, *236*, 1252−1258.
(14) Assays for ribonucleolytic activity used 6-carboxyfluorescein-dArU(dA)$_2$−6-carboxytetramethylrhodamine as the substrate (Kelemen, B. R.; Klink, T. A.; Behlke, M. A.; Eubanks, S. R.; Leland, P. A.; Raines, R. T. *Nucleic Acids Res.* **1999**, *27*, 3696−3701).

JA029752E