Chemical synthesis of proteins

Powerful new enzymatic and chemical methods for coupling unprotected peptide fragments are making the assembly of large synthetic proteins possible. By allowing the use of novel backbones and the incorporation of multiple unnatural amino acids at specific sites, these methods promise large expansion of the repertoire of protein molecules accessible to research.

Chemistry & Biology December 1994, **1**:201–203

Molecular biology has revolutionized the study of protein structure and function. Not only has the microbial production of enzymes and other proteins in useful amounts become routine in recent years, but systematic alteration of protein sequence by site-directed mutagenesis is now a standard tool for dissecting chemical mechanism. The recent development of new methods for ligating peptides, however, promises to make chemical synthesis of large proteins an attractive alternative to biosynthesis, particularly for the construction of novel molecules containing non-natural amino acids or other structural modifications.

Attempts to synthesize proteins chemically have traditionally followed one of two basic strategies: stepwise assembly from their constituent amino acids on a solid support, or convergent coupling of peptide segments in solution. Although solid-phase peptide synthesis has been optimized to the extent that proteins over 100 amino acids in length can be prepared in favorable cases [1], such syntheses still represent heroic undertakings and the number of side products that accumulate over the many coupling steps that are necessary often makes purification of the final product laborious. Convergent coupling has the considerable advantage that synthesis and purification of peptide segments up to 30 residues long is relatively straightforward; since a variety of unusual amino acids can be introduced into each segment, this approach is also highly adaptable to the preparation of multiple analogs of naturally-occurring proteins. But broad implementation of the fragment condensation strategy has been limited by the poor solubility of protected peptide segments and the tendency of α-carboxy-activated peptides to racemize.

Many of these difficulties can be circumvented by enzymatic coupling procedures. Proteolytic enzymes, particularly serine proteases, have been used extensively in peptide synthesis because of their selectivity, ability to function under mild reaction conditions, and minimal requirements for protecting groups [2]. Although it might seem counterintuitive to attempt to synthesize a protein using a protease, problems due to proteolytic degradation of the final peptide product by the catalyst can be minimized by manipulating the reaction conditions or through genetic engineering. For example, replacement of the active site serine in subtilisin with cysteine converts the protease into an acyl transferase: the modified enzyme is readily acylated by esters and subsequently deacylated by amines, but has greatly diminished amidase activity compared with wild-type subtilisin [3]. However, substitution of sulfur for oxygen engenders steric crowding in the active site, limiting the activity of the enzyme. Wells and colleagues [4] have relieved this crowding by mutating a second residue in the active site. The resulting double mutant, dubbed 'subtiligase,' efficiently couples unprotected peptides in water.

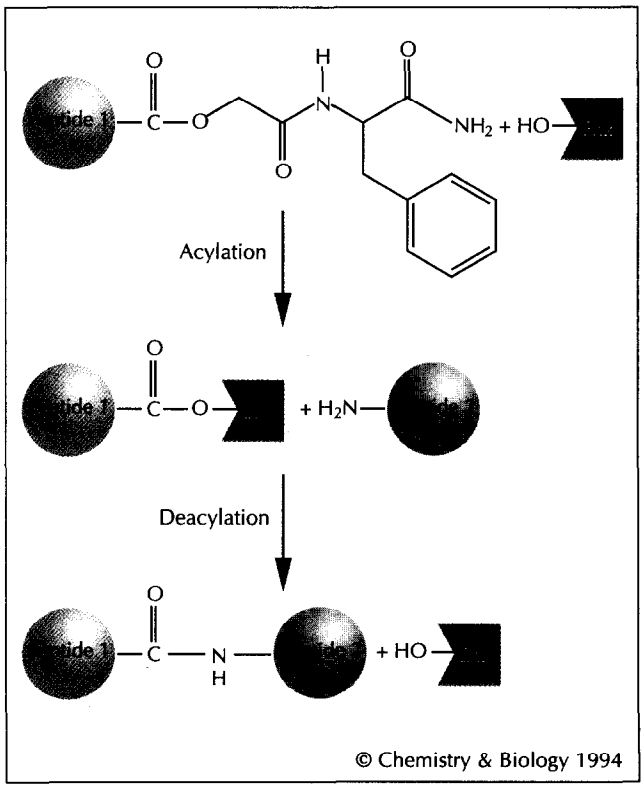

Fig. 1. General scheme for ligating peptide fragments using subtiligase. In the first step, the enzyme attacks peptide 1, which is esterified with glycolate-phenylalanyl amide, to produce a thiol-acyl enzyme intermediate. In the second step, the α-amine of peptide 2 attacks the thiol-acyl bond on subtiligase and generates the free enzyme plus the ligated peptide product.

The acyl donor in a typical subtiligase-catalyzed ligation (Fig. 1) is a peptide whose carboxyl terminus is esterified with a glycolate-phenylalanine amide group. This ester efficiently acylates subtiligase because of the enzyme's preference for glycine and phenylalanine in the two positions directly following the 'cleavage' site, P1' and P2', respectively. Deacylation of the enzyme by a fully deprotected peptide (or protein) corresponding to the desired acceptor fragment then yields the ligated product. Nearly quantitative yields have been obtained for reactions between model esters and dipeptide nucleophiles [4]. Such an approach has also been successfully applied to the post-translational modification of natural and recombinant proteins. For instance, biotinylated and heavy-atom derivatives of N-methionyl human growth hormone were prepared by ligating peptides that contain biotin or mercury to the hormone's N-terminus [5]. Wong *et al.* [6] have used analogous enzymatic methods to synthesize a variety of N- and O-linked glycopeptides.

Blockwise assembly of large proteins by repeated couplings of long peptides is also possible. The enzyme RNase A (126 amino acids) was synthesized in good yield and high purity by subtiligase-catalyzed ligation of six peptide fragments ranging in length from 12 to 30 residues [7]. RNase A mutants containing the unnatural amino acid 4-fluorohistidine in place of the histidines at positions 12 and 119 were also prepared to examine whether the basic nature of these residues is important in catalysis [7]. Given that the incorporation of multiple unnatural amino acids into a protein backbone is not currently possible using recombinant technologies, this application of subtiligase illustrates the potential utility of chemical approaches to protein total synthesis. By expanding the set of possible changes that can be made, our ability to probe structure–function relationships in proteins with the precision of physical organic chemistry and to design new materials with novel properties will increase dramatically. Site-specific incorporation of NMR probes, structural constraints, and catalytic cofactors into proteins are only some of the exciting possibilities that spring to mind.

The utility of subtiligase-catalyzed peptide coupling is a consequence, in part, of its high reaction rates and relatively broad specificity. The glycolate esters used as acyl donors are stable and easily prepared by solid-phase methods, so that side-reactions, especially racemization, which are often problematic in the chemical coupling of activated peptides are suppressed. The fact that the enzyme tolerates a wide range of P1, P1', P2 and P2' amino acids also provides great latitude in the choice of ligation site. Site-directed mutagenesis has been used to broaden this specificity further [4,6] and to increase the stability of the enzyme [5,6]. The design of subtiligase variants capable of coupling peptides in 4 M guanidine hydrochloride may be particularly important [5], as highly-structured peptides and proteins are otherwise poor substrates for the ligase.

New non-enzymatic methods for coupling peptide fragments, including highly structured peptides, have the potential to complement and extend enzymatic ligation procedures considerably. Several research groups [8–10] have reported powerful chemical ligation strategies that can be carried out under fully denaturing conditions with high efficiency. In these reactions, unprotected peptide fragments are specifically brought together in such a way as to juxtapose the acyl donor and amine acceptor. The high effective molarity of the reactive groups then ensures facile intramolecular formation of the desired peptide bond. This approach, which bypasses the need for highly activated acyl donors, is illustrated by the recent work of Kent and coworkers [10].

As shown in Fig. 2, native chemical ligation can be achieved by chemoselectively reacting a peptide bearing a thioester at its carboxyl terminus with a second peptide capped at its amino end by a cysteine. The resulting thioester ligation product undergoes a rapid, geometrically favorable rearrangement involving acyl transfer from the thiol to the amine, yielding a final product that contains a native amide bond at the ligation site. Both peptide segments are conveniently prepared by conventional techniques, and coupling is rapid and selective in aqueous buffer, occurring without significant side reactions. In fact, solubilizing agents such as

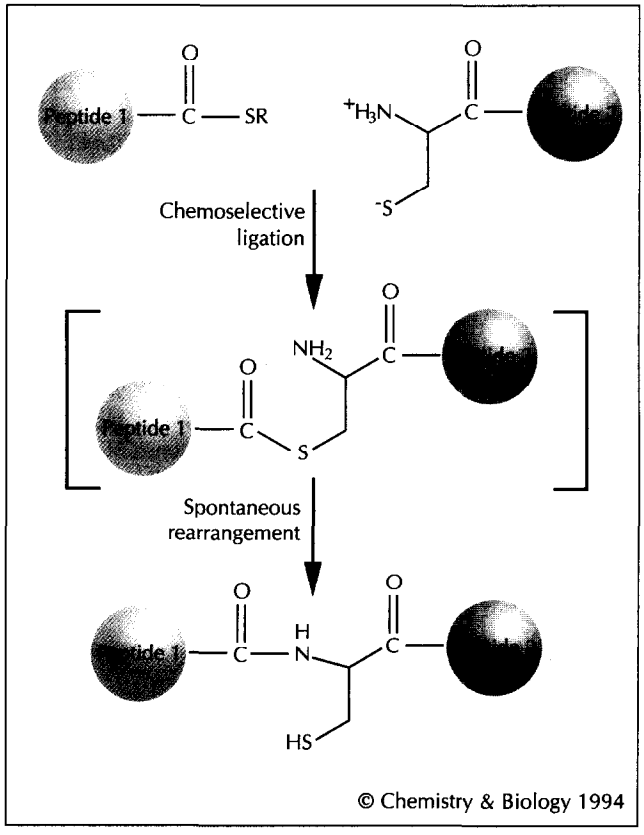

Fig. 2. Two-step approach for chemically coupling two peptides. Peptide 1, derivatized as a thiol ester, reacts with the thiol side chain of the N-terminal cysteine of peptide 2 to give a thioester ligation product. The intermediate rapidly rearranges to give a peptide bond at the ligation site.

urea and guanidine hydrochloride appear to enhance the chemical efficiency of reaction by concentrating and denaturing the peptide segments. The synthesis of interleukin-8, a 72-residue cytokine that contains multiple disulfide bonds, demonstrates the utility of this approach for the preparation of moderately-sized polypeptides [10]; the synthesis of larger proteins and protein analogs is limited only by the availability of larger, appropriately modified peptide segments.

Although native chemical ligation strategies require a cysteine at the coupling site, this is not likely to be a significant limitation in the preparation of proteins with novel architectures. The thiol side chain is a useful catalytic group and/or metal ligand in many proteins. It is also readily converted through selective alkylation to a variety of anionic, cationic and neutral functionalities. Indeed, the high yields and comparatively easy implementation of the chemical approach are likely to render it an effective alternative to enzymatic coupling methods in many instances, and a successful marriage of the two techniques may dramatically increase the size range of polypeptides directly accessible by total chemical synthesis.

Of course, chemists wishing to synthesize proteins need not be bound by nature's choices. In addition to expanding the repertoire of amino acid side chains that can be incorporated into proteins, chemical synthesis makes backbone engineering possible: peptide segments and protein domains can be joined by linkages other than the amide bond. This point is illustrated by variants of HIV-1 protease containing novel thiol ester linkages [11,12], which were constructed by combining a peptide fragment possessing a C-terminal thioester with a second segment bearing a bromoacetyl group at its amino end. Alkylation of the thiol occurs rapidly and in high yield. When the thioester replaces the natural peptide bond between Gly51 and Gly52, the alkylated product remains fully active [11], but when it is used to replace the Gly49–Ile50 peptide bond, catalytic activity is reduced [12]. This observation provided direct evidence for the notion that specific backbone hydrogen bonds from the enzyme active site flaps to the substrate are important for the catalytic function of the protease. Backbone engineering has also been exploited for the rapid assembly of a four-helix bundle on a templating peptide [13] and for semisynthesis of analogs of human granulocyte colony stimulating factor containing a single backbone acyl hydrazone bond [14].

In sum, versatile new methods for coupling peptides provide the basis for a general and modular approach to the chemical synthesis of large proteins. Accordingly, the field of protein chemistry is likely to experience something of a renaissance. For decades to come, wide implementation of these technologies will provide exciting research opportunities for chemists and biologists interested in exploring protein structure and function.

References

1. Muir, T.W. & Kent, S.B.H. (1993). The chemical synthesis of proteins. *Curr. Opin. Biotechnol.* **4**, 420–427.
2. Kullman, W. (1987). *Enzymatic Peptide Synthesis.* CRC Press, Boca Raton.
3. Nakatsuka, T., Sasaki, T. & Kaiser, E.T. (1987). Peptide segment coupling catalyzed by the semisynthetic enzyme thiolsubtilisin. *J. Am. Chem. Soc.* **109**, 3808–3810.
4. Abrahmsén, L., Tom, J., Burnier, J., Butcher, K.A., Kossiakoff, A. & Wells, J.A. (1991). Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**, 4151–4159.
5. Chang, T.K., Jackson, D.Y., Burnier, J.P. & Wells, J.A. (1994). Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA*, in press
6. Wong, C.-H., Schuster, M., Wang, P. & Sears, P. (1993). Enzymatic synthesis of N- and O-linked glycopeptides. *J. Am. Chem. Soc.* **115**, 5893–5901.
7. Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J. & Wells, J.A. (1994). A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **266**, 243–247.
8. Kemp, D.S. & Carey, R.I. (1993). Synthesis of a 39-peptide and a 25-peptide by thiol capture ligations: observation of a 40-fold rate acceleration of the intramolecular O,N-acyl-transfer reaction between peptide fragments bearing only cysteine protective groups. *J. Org. Chem.* **58**, 2216–2222.
9. Liu, C.-F. & Tam, J.P. (1994). Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. *J. Am. Chem. Soc.* **116**, 4149–4153.
10. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B.H. (1994). Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779.
11. Schnölzer, M. & Kent, S.B.H. (1992). Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineering HIV protease. *Science* **256**, 221–225.
12. Baca, M. & Kent, S.B.H. (1993). Catalytic contribution of flap-substrate hydrogen bonds in 'HIV-1 protease' explored by chemical synthesis. *Proc. Natl. Acad. Sci. USA* **90**, 11638–11642.
13. Dawson, P.E. & Kent, S.B.H. (1993). Convenient total synthesis of a 4-helix TASP molecule by chemoselective ligation. *J. Am. Chem. Soc.* **115**, 7263–7266.
14. Gaertner, H.F., Rose, K., Cotton, R., Timms, D., Camble, R. & Offord, R.E. (1992). Construction of protein analogues by site-specific condensation of unprotected fragments. *Bioconjugate Chem.* **3**, 262–268.

Donald Hilvert, Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, USA.