

CHEMICAL SYNTHESIS OF PROTEINS

Bradley L. Nilsson,¹ Matthew B. Soellner,¹
and Ronald T. Raines^{1,2}

*Departments of ¹Chemistry and ²Biochemistry, University of Wisconsin—Madison,
Madison, Wisconsin 53706; email: raines@biochem.wisc.edu*

Key Words peptide synthesis, protein chemistry, proteomics, structural genomics

■ **Abstract** Proteins have become accessible targets for chemical synthesis. The basic strategy is to use native chemical ligation, Staudinger ligation, or other orthogonal chemical reactions to couple synthetic peptides. The ligation reactions are compatible with a variety of solvents and proceed in solution or on a solid support. Chemical synthesis enables a level of control on protein composition that greatly exceeds that attainable with ribosome-mediated biosynthesis. Accordingly, the chemical synthesis of proteins is providing previously unattainable insight into the structure and function of proteins.

CONTENTS

INTRODUCTION	91
PEPTIDE SYNTHESIS	93
PROTEIN SYNTHESIS BY PEPTIDE LIGATION	95
PEPTIDE LIGATION WITH SULFUR	96
Prior Thiol Capture	96
Native Chemical Ligation	97
Expressed Protein Ligation	98
Chemical Modification of Ligated Peptides	100
Acyl-Initiated Capture	101
PEPTIDE LIGATION WITH SELENIUM	102
GENERAL STRATEGIES FOR PEPTIDE LIGATION	103
Conformationally Assisted Ligation	104
Removable Auxiliaries	104
Staudinger Ligation	107
PROSPECTUS	111

INTRODUCTION

The human genome contains approximately 30,000 genes (83, 144). Scientists from a broad range of disciplines are now working to reveal the structure and function of the proteins encoded by these genes. Their findings could lead to the solution of a multitude of problems in biology and medicine. In addition to

structure-function analyses of extant proteins, protein chemists are working to create new proteins with desirable properties, either by *de novo* design or by altering natural frameworks.

The study of natural proteins and the creation of nonnatural ones require the ability to access and manipulate proteins. The isolation of proteins from their natural source is often tedious, idiosyncratic, and impractical. In contrast, the production of proteins with recombinant DNA (rDNA) technology, either in a heterologous host or *in vitro*, can provide access to large quantities of protein and allow for the exchange of 1 of 20 common amino acid residues for another. However, aggregation often limits the yield of properly folded proteins produced with rDNA. Moreover, the restrictions of the genetic code severely limit the possible modifications.

The chemical synthesis and semisynthesis of proteins harbor the potential to overcome many of the disadvantages of current protein production methods (19, 29, 78). In particular, chemical synthesis using established solid-phase techniques are rapid to effect, easily automated, and facilitate purification. Accordingly, the application of existing and emerging synthetic methods could facilitate research in all aspects of protein science.

Chemical synthesis enables the facile incorporation of nonnatural functionality into proteins. The genetic code limits the components of natural proteins to approximately 20 α -amino acids. Methods that overcome this limitation but still rely on the ribosome are similarly limited to a subset of α -amino acids and α -hydroxy acids (65, 91). In marked contrast, the nonnatural functionality made available by chemical synthesis is limited only by the constraints of the periodic table and the imagination of protein chemists.

The desire to synthesize proteins is not new. On December 12, 1902, Emil Fischer (47) delivered his Nobel Prize lecture in Stockholm, Sweden, saying in part:

Of the chemical aids in the living organism the ferments—mostly referred to nowadays as enzymes—are so pre-eminent that they may justifiably be claimed to be involved in most of the chemical transformations in the living cell. The examination of the synthetic glucosides has shown that the action of the enzymes depends to a large extent on the geometrical structure of the molecule to be attacked, that the two must match like lock and key. Consequently, with their aid, the organism is capable of performing highly specific chemical transformations which can never be accomplished with the customary agents. To equal Nature here, the same means have to be applied, and I therefore foresee the day when physiological chemistry will not only make extensive use of the natural enzymes as agents, but when it will also prepare synthetic ferments for its purposes.

A century later, Fischer's vision is becoming reality. Enzymes and other proteins not only are accessible targets for synthetic chemistry, but also are poised to become dominant targets of the twenty-first century. Herein, we discuss current efforts toward preparing proteins synthetically, focusing on the development of powerful

new methodologies for splicing peptide fragments in a convergent strategy for the total chemical synthesis of proteins.

PEPTIDE SYNTHESIS

The chemical synthesis of proteins is now possible because of the prodigious advances in peptide synthesis that have occurred over the last century. Fischer's 1901 synthesis of glycyl glycine is the first reported synthesis of a dipeptide and is also the first instance of the term "peptide" used to refer to a polymer of amino acids (48). His 1907 synthesis of an octadecapeptide consisting of 15 glycine and 3 leucine residues was a remarkable achievement, despite his inability to control its amino acid sequence (46).

An important advance in peptide synthesis was Bergmann & Zervas' 1932 introduction of reversible protection for the α -amino group (14). With the emergence of protecting group strategies, it became possible to synthesize small-peptide hormones. For example, in 1953 du Vigneaud and coworkers (43) reported a solution-phase synthesis of the octapeptide hormone oxytocin. Even though 50 years had passed since Fischer's first synthesis of a peptide, these types of syntheses were still only accomplished with considerable effort (95).

The advent of solid-phase methods heralded a revolution for peptide synthesis (94). In 1963, Merrifield (96) described the first solid-phase synthesis of a peptide, a tetrapeptide. He attached an amino acid to an insoluble support via its carboxyl group and then coupled the next amino acid, which had a protected amino group and an activated carboxyl group. The amino-protecting group was removed, and the next amino acid was coupled in a similar manner. Within a few years, Merrifield (97) reported the development of an instrument for the automated synthesis of peptides. In short order, Gutte & Merrifield (59, 60) used this new strategy to achieve the first synthesis of an enzyme, ribonuclease A (RNase A), albeit in low overall yield. Concurrently, a team led by Hirschmann (66) reported the chemical synthesis of RNase S (which consists of residues 21–124 of RNase A) by solution-phase segment condensation reactions.

Automated solid-phase peptide synthesis is commonplace today. In the most common strategy (Figure 1), an amino acid with both α -amino group and side chain protection is immobilized to a resin. The α -amino-protecting group is typically an acid-sensitive *tert*-butoxycarbonyl (Boc) group or a base-sensitive 9-fluorenylmethoxycarbonyl (Fmoc) group (146). These α -amino-protecting groups can be removed quickly and completely, and a protected amino acid with an activated carboxyl group can then be coupled to the unprotected resin-bound amine. The coupling reactions are forced to completion by using an excess of activated soluble amino acid. The cycle of deprotection and coupling is repeated to complete the sequence. With side chain deprotection and cleavage, the resin yields the desired peptide (61, 134).

The efficiency of solid-phase peptide synthesis continues to improve. New solid supports have increased the length of accessible peptides (90). New linkers

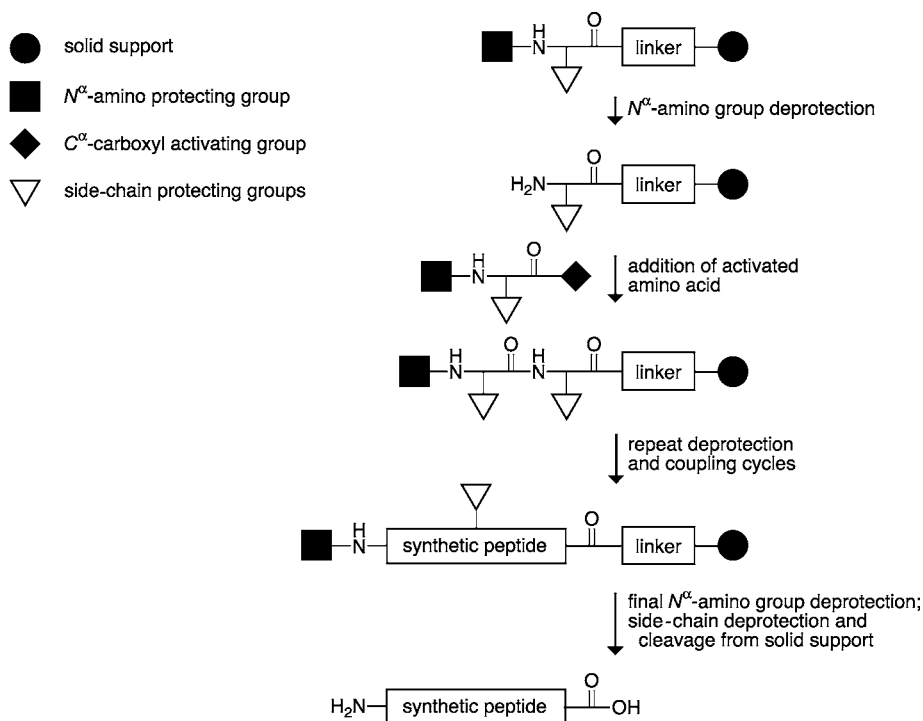


Figure 1 Solid-phase peptide synthesis.

between the support and the peptide have diversified the conditions that can be used to liberate a synthetic peptide (132). New side chain protection strategies have minimized deleterious side reactions. Finally, new carboxyl-activating groups have increased the speed and efficiency of amino acid couplings while reducing the risk of epimerization (3, 99).

Solid-phase peptide synthesis alone has enabled the total chemical synthesis of some proteins. Since the pioneering work of Merrifield, proteins that contain as many as 166 amino acid residues have been synthesized in this fashion (95). These syntheses have, in some cases, been critical to the structure-function analyses of the target proteins. Notably, the chemical synthesis of HIV-1 protease enabled the structural characterization of protease inhibitor complexes (98, 135, 148). In addition, an enantiomer of this protease was synthesized with D-amino acids to demonstrate its chiral specificity for a peptide substrate derived from D-amino acids (41). Another notable example is a recent synthesis of a modified B1 domain of streptococcal G protein (112). This synthesis incorporated a completely nonnatural dibenzofuran-based β -turn mimic as a conformational probe. These studies highlight the true power and potential of the total chemical synthesis of proteins in enhancing our understanding of protein structure and function.

Despite the significant gains made in advancing the technology of solid-phase peptide synthesis, there remain limitations. Modern peptide synthesis is typically limited to peptides consisting of no more than 40 residues (2, 22). Peptides and proteins of greater length can be prepared, although not routinely. Hence, most proteins cannot be synthesized by the stepwise assembly of amino acid monomers.

The convergent assembly of protected or partially protected peptide segments, both in solution and on solid phase, is one approach used to access proteins that contain more than 40 residues (1, 2, 4, 84, 122). The 238-residue precursor of the green fluorescent protein (GFP) was synthesized with this segment-condensation approach (110). Twenty-six peptide fragments corresponding to parts of the GFP were synthesized, assembled in solution, and deprotected. The resulting protein exhibited a fluorescence spectrum indistinguishable from that of the biosynthetic protein upon standing in solution. In addition to producing a large, complex protein in its native form by chemical synthesis, this study also demonstrated unequivocally that the formation of the GFP fluorophore is not dependent on any external cofactors.

PROTEIN SYNTHESIS BY PEPTIDE LIGATION

The formation of an amide bond between protected peptide fragments by the attack of an amino group on an activated carboxyl group can be problematic, especially with large peptides (32). The problem is that both the intrinsic rate constant (k) for this reaction and the reactant concentrations are often too low for a substantial reaction velocity (v , where $v = k[\text{peptide 1}][\text{peptide 2}]$). In other words, the reactivity between an amino group and an activated carboxyl group is not inherently high enough for acyl transfer to overcome the relatively low solubility of peptides.

In nature, peptide bond formation occurs by sequential transfer of the C-terminal acyl group of a nascent peptide chain to the α -amino group of an aminoacyl tRNA. The C-terminal acyl group is activated only as an ester with a tRNA, and the transfer is aided, in large part, simply by the reactants being held in close proximity by the ribosome (129). Proteases have been engineered to perform in a similar manner (18, 30, 106) and used for the convergent synthesis of proteins as large as RNase A (70).

New chemical strategies have emerged that take advantage of enforced proximity to couple peptides. These peptide ligation methods provide a practical and powerful means to assemble synthetic peptides into proteins. The peptides can be protected or unprotected, and coupling can occur in an aqueous or organic solvent, in solution, or on a solid support.

Peptide ligation strategies utilize three steps (Figure 2). An initial capture step links the peptides by a chemical reaction that is more rapid than intermolecular acyl transfer to an amine and that uses functional groups with reactivity that is orthogonal to that in proteins. The pivotal acyl transfer step is thus rendered

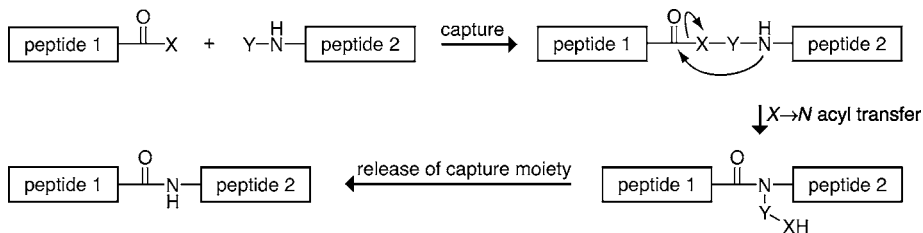


Figure 2 A general strategy for peptide ligation.

intramolecular and hence can occur with maximal efficiency (92, 116). In the last step, the capture moiety is released, either in a discrete chemical process or spontaneously.

Below, we discuss current strategies for peptide ligation, along with their relative advantages and disadvantages. We also highlight some notable applications of these strategies. Our focus is on ligation strategies that yield native amide bonds. We do not discuss approaches that yield nonnatural bonds between peptide fragments (44, 140, 138).

PEPTIDE LIGATION WITH SULFUR

Prior Thiol Capture

Modern methods of peptide ligation trace their origin to Kemp's prior thiol capture strategy (Figure 3). In this approach, 4-hydroxy-6-mercaptodibenzofuran is used as the association element to bring the peptide coupling partners together (49, 72, 74–76). This strategy uses a highly efficient thiol-disulfide exchange reaction in a capture step prior to the acyl transfer reaction. Acyl transfer occurs with a half-life ranging from 0.1 to 50 h, depending on the bulk of the side chain of the C-terminal acyloxy-derived residue (77), and without racemization of the coupled amino acids (88). Upon completion of the acyl transfer reaction, phosphine reduction of the mixed disulfide yields the native peptide.

The utility of this approach was demonstrated by the ligation of a variety of peptide fragments. In one study, the C-terminal 29-residue fragment of BPTI (bovine pancreatic trypsin inhibitor) was synthesized from four segments, each possessing an N-terminal cysteine residue and a C-terminal dibenzofuran (49). The synthesis commenced with the two fragments at the extreme C terminus and proceeded sequentially toward the N terminus. The N-terminal cysteine residues were protected orthogonally until they were needed for coupling, when they were deprotected and derivatized as mixed disulfides. In a separate study, a 39-residue peptide and a 25-residue peptide were synthesized in high yield, utilizing side chain protection only at cysteine residues (73). The ligation of fully protected peptides was 50-fold slower.

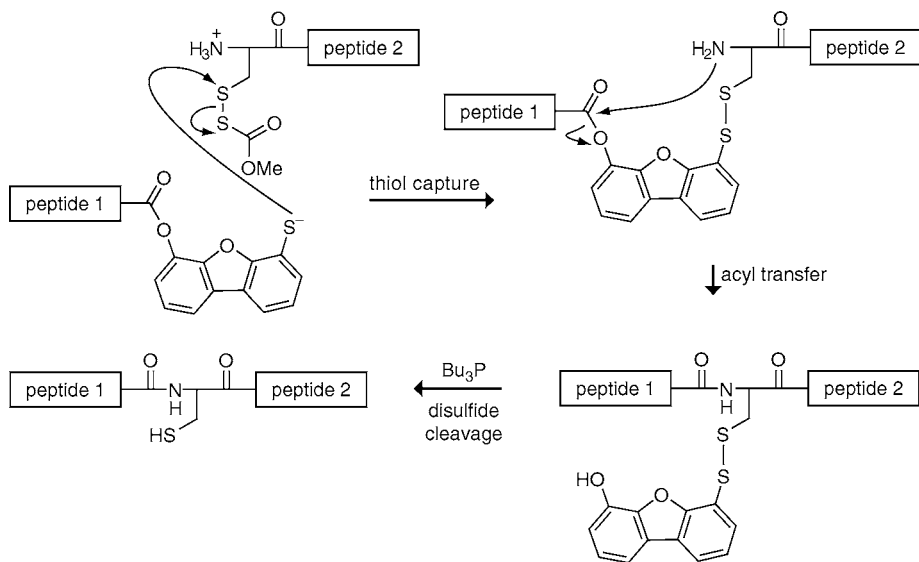


Figure 3 Prior thiol capture.

Kemp's prior thiol capture strategy is a seminal contribution to the development of peptide ligation concepts and methodology. It represents the first demonstration of the chemoselective ligation of unprotected peptide fragments. In addition, it represents the first systematic application of the use of proximity effects to evoke acyl transfer via an intramolecular reaction.

Native Chemical Ligation

Currently, the most common ligation method is native chemical ligation (Figure 4). In 1953, Wieland et al. (147) discovered the chemical foundation for this reaction, when the reaction of ValSPh and CysOH in aqueous buffer was shown to yield

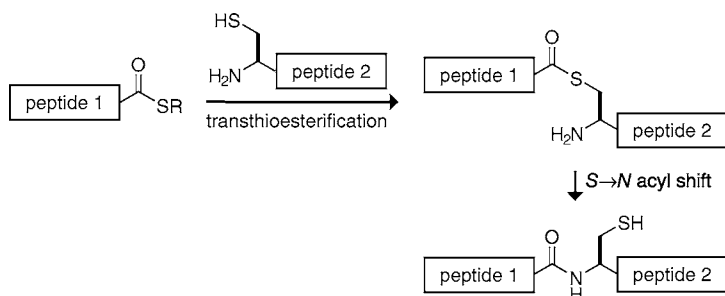


Figure 4 Native chemical ligation.

the dipeptide ValCysOH. The reaction proceeded through the intermediacy of a thioester containing the sulfur of the cysteine residue. In the 1990s, Kent developed this reaction into a practical method to ligate large peptide fragments (38, 39, 78). In native chemical ligation, the thiolate of an N-terminal cysteine residue of one peptide attacks the C-terminal thioester of a second peptide to effect transthioesterification. An amide linkage forms after rapid $S \rightarrow N$ acyl transfer.

Native chemical ligation is a powerful approach for the total chemical synthesis of proteins. Ligations are effected in aqueous solution with unprotected peptides (104). A chaotrope such as guanidinium chloride or urea is added to discourage peptide aggregation and structure. Transthioesterification with embedded cysteine residues does not lead to amide bond formation and hence is reversible. The ligation rate can be modulated by the addition of exogenous thiols (37). Ligation can also occur between cysteine and any of the other common amino acids, although ligation to hindered amino acids such as valine, isoleucine, and proline occurs relatively slowly (63). Although normally carried out in solution, native chemical ligation has been used to couple unprotected peptide fragments on a solid support (23, 25, 27).

The application of native chemical ligation in the synthesis of proteins is burgeoning. Some recent examples of target proteins include the human matrix Gla protein (84 residues) (64), the anticoagulant microprotein S (116 residues) (62), and the human neutrophil pro α -defensin-1 (75 residues) (149). Proteins with posttranslational modifications that are inaccessible with rDNA techniques have been synthesized successfully with native chemical ligation. These include the glycoproteins diptericin (58 residues) (128) and lymphotactin (93 residues) (86). The prion protein has also been synthesized with appropriate mimics of its glycosylphosphatidylinositol anchor (8).

A powerful aspect of the chemical synthesis of proteins is the ability to incorporate nonnatural residues as probes of structure and function or for therapeutic purposes. Native chemical ligation has yielded many important examples of modified synthetic proteins. The protooncogene H-Ras (166 residues) and its binding partner, the Ras-binding domain (81 residues), have been synthesized with the site-specific incorporation of fluorescent tags to probe the interactions between these two proteins (10, 11). In a particularly noteworthy example, a synthetic polyethylene glycol-modified erythropoiesis protein (166 residues) has been synthesized by native chemical ligation (79). The appended polymer chains serve to prolong its half-life in vivo and hence its potent biological activity.

A one-pot total synthesis of a crambin has been performed in high yield by the sequential coupling of three synthetic polypeptide chains (9). This synthesis of crambin (46 residues) did not require purification of the intermediates and afforded high yields of intact protein. Additionally, folding of the synthetic protein was performed in the reaction mixture prior to purification.

Expressed Protein Ligation

An intein is a natural splicing element that is analogous to the intron of nucleic acids (111, 117). An intein mediates its own excision from a peptide

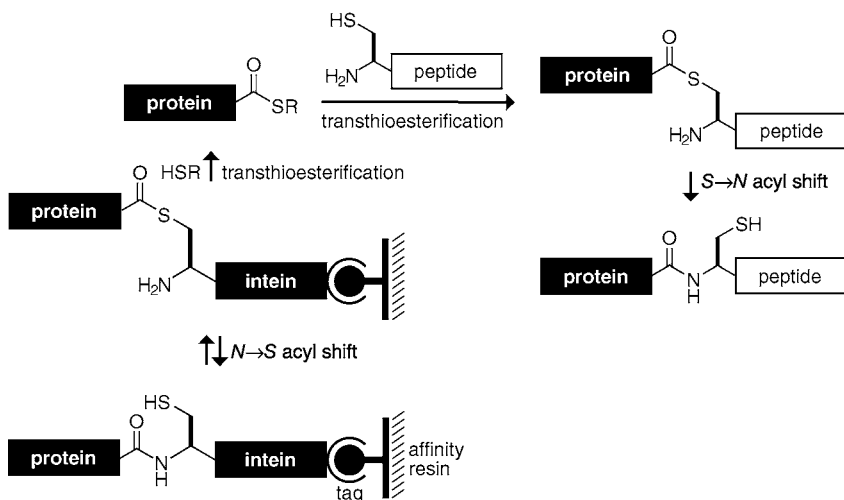


Figure 5 Expressed protein ligation.

sequence through a series of acyl transfer reactions that ultimately results in the splicing of its flanking peptides. In 1998, Evans et al. and Muir and colleagues (45, 105, 126) appropriated the chemistry of intein-mediated, posttranscriptional protein splicing for protein semisynthesis. In their expressed protein ligation (Figure 5), the fusion of a target protein to a modified intein is used to install a thioester at the C terminus of the target protein. The fusion protein can be bound to an affinity resin, where the incipient thioester can be trapped with a small-molecule thiol, thus eluting the target protein as a simple C-terminal thioester and leaving the remainder of the fusion protein bound to the affinity resin.

The semisynthesis of proteins using expressed protein ligation represents a powerful union of chemistry and biology. In effect, the use of rDNA to produce protein fragments increases the size of proteins that are accessible to native chemical ligation (compare Figure 5 with Figure 4). The bulk of a target protein can be accessed by biosynthesis and then ligated to a peptide prepared by chemical synthesis.

Expressed protein ligation has achieved widespread use in protein structure-function analyses (36, 57, 103). In particular, the semisynthesis of proteins that are phosphorylated (31, 115), glycosylated (58, 141, 145), or labeled with isotopes (17, 35, 121) or fluorophores (34, 118, 124) on specific residues is leading to notable insights. In addition, expressed protein ligation has been used to produce a functional single-chain antibody with a C-terminal thioester, which could be the basis of a new class of diagnostic and therapeutic agents (136).

Expressed protein ligation has facilitated protein “prosthesis.” In two examples, reverse turns in RNase A were replaced with synthetic reverse-turn mimics, dinipectic acid (7) and dimethylproline (6) (Figure 6). The incorporation of these

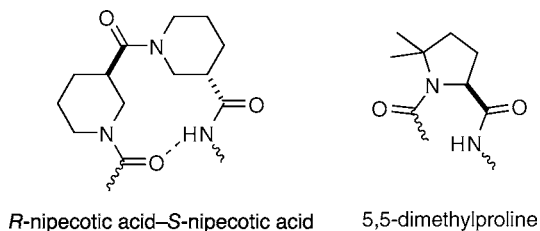


Figure 6 Prosthetic modules installed into RNase A by expressed protein ligation (6, 7).

reverse-turn mimics increased the thermal stability of the protein, and the variant containing dimethylproline folded more quickly.

Intein-mediated protein splicing has been induced to occur in living cells, enabling new biological applications. One strategy involves splitting an intein into two fragments (54, 101, 102). When these fragments associate noncovalently, the activity of the intein is regained and the ensuing splicing event can produce a functional protein. Fragment association can be modulated by a cell-permeable small molecule. The binding of a small molecule has also been used to activate an intact intein, which can then activate a target protein (24).

Chemical Modification of Ligated Peptides

The utility of the native chemical ligation approach has been expanded by chemical modification. In these strategies, an additional step allows ligation with a thiol-containing residue, but yields a protein that does not contain a thiol. To date, this approach has been used in only a few specific instances.

One such example is native chemical ligation at cysteine followed by desulfurization (150). This strategy employs standard native chemical ligation using a cysteine residue. After the ligation step, chemical desulfurization of the cysteine residue gives alanine at the ligation junction (Figure 7). Desulfurization has been effected by hydrogen gas and a palladium catalyst. Three proteins have been synthesized with this strategy: the streptococcal protein G B1 domain (56 residues), a variant of barnase (110 residues), and the cyclic antibiotic microcin J25 (21 residues). All thiols in a protein target are reduced by known desulfurization conditions, but disulfide bonds are unaffected, making this strategy applicable to proteins that contain cystine but not cysteine residues.

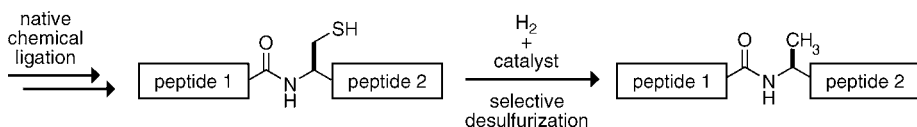


Figure 7 Ligation at an incipient alanine residue.

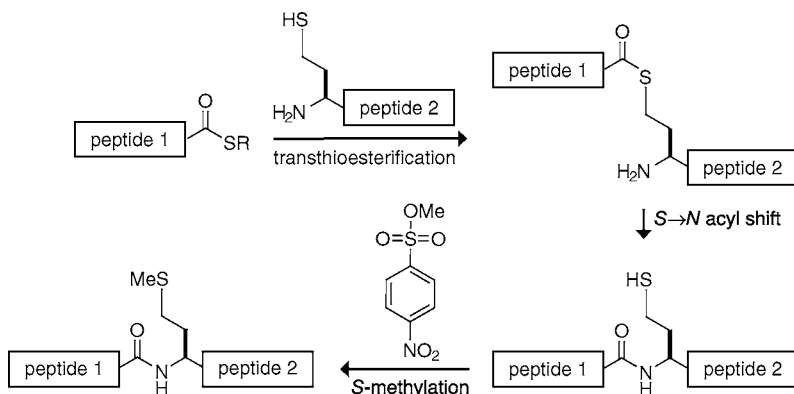


Figure 8 Ligation at an incipient methionine residue.

A second example yields methionine at the splice site by ligation with homocysteine followed by its chemical methylation (139). This ligation proceeds initially in a fashion nearly identical to native chemical ligation (compare Figure 8 with Figure 4). Homocysteine effects $S \rightarrow N$ acyl transfer efficiently, despite having an additional methylene group in its side chain. After acyl transfer, sulfur methylation is effected with methyl *p*-nitrobenzenesulfonate. An excess of the methylating reagent is used, but the methylation of amino groups is avoided by limiting reaction times. Again, this method is limited to targets that lack a cysteine residue.

Acyl-Initiated Capture

Another original approach to peptide ligation is the in situ generation of a cysteine thioester by acyl-initiated rather than side chain-initiated capture (137). In this method, Tam et al. (137) reacted a peptide having a C-terminal thioacid with a second peptide having β -bromoalanine as its N-terminal residue to generate

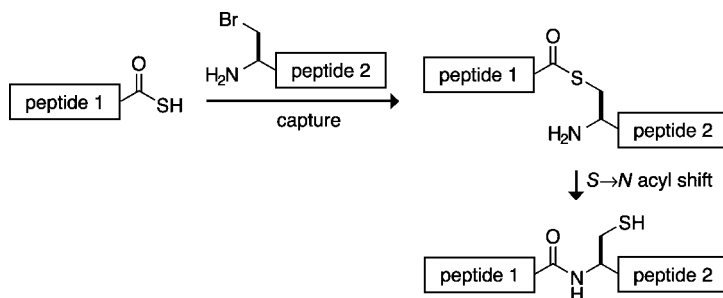


Figure 9 Acyl-initiated capture.

a thioester analogous to that formed during native chemical ligation (compare Figure 9 with Figure 4). This intermediate then undergoes $S \rightarrow N$ acyl transfer, resulting in a native amide bond with a cysteine residue at the ligation junction. A disadvantage of this method is that bromoalanine is prone to aziridine formation and other undesirable side reactions.

PEPTIDE LIGATION WITH SELENIUM

Selenocysteine is a natural amino acid residue with a low natural abundance (81, 82). During the biosynthesis of selenoproteins, selenocysteine (Sec or U) is incorporated by the ribosomal translation of mRNA and has its own tRNA^{Sec} and codon, UGA (which is also the *opal* stop codon). Decoding a UGA codon for selenocysteine requires a unique structure in the 3'-untranslated region of the mRNA called a selenocysteine insertion sequence element. The production of eukaryotic selenocysteine-containing proteins in prokaryotes is problematic because eukaryotic and prokaryotic cells use a different selenocysteine insertion sequence element (5).

The feasibility of using selenocysteine in native chemical ligation has been demonstrated with model systems (53, 67, 119). pH rate profiles have demonstrated that a ligation with a selenocysteine residue occurs much more readily than does one with cysteine (67), as expected from the lower pK_a of selenols and the somewhat higher reactivity of selenates. For example, at pH 5.0 the reaction with selenocysteine is 10³-fold faster than that with cysteine.

Selenocysteine has been used to mediate expressed protein ligation (67, 68). rDNA technology was used to prepare a fragment corresponding to residues 1–109 of RNase A with a C-terminal thioester. Standard solid-phase methods were used to synthesize a peptide corresponding to residues 110–124, but with selenocysteine rather than cysteine as residue 110. The thioester fragment and the peptide fragment were then ligated, and the product was folded and purified. The integrity of the desired C110U variant was verified by mass spectrometry and its wild-type enzymatic activity. The data indicate that C110U RNase A is not only an intact protein, but also a correctly folded enzyme with a selenosulfide (Se–S) bond between Sec-110 and Cys-56.

A semisynthesis of a selenocysteine-containing variant of azurin has also been achieved with expressed protein ligation (15). Cys-112 was replaced with selenocysteine, and fragment 112–128 was ligated to fragment 1–111. The variant azurin protein displayed unique spectral properties that were used to reveal the nature of the enzyme-copper interaction in the native protein.

The isomorphous replacement of sulfur with selenium can stabilize a protein. Specifically, proteins containing selenosulfide bonds should have greater conformational stability in a reducing environment than proteins with disulfide bonds, as the reduction potential of a selenosulfide bond is much less than that of the corresponding disulfide bond (16). This use of selenocysteine to stabilize an enzyme represents another form of protein prosthesis.

Selenomethionine has been incorporated into bovine pancreatic polypeptide with a strategy similar to that used for methionine (120) (compare with Figure 8). Here, native chemical ligation of a peptide fragment having an N-terminal selenohomocysteine residue was used to effect native chemical ligation. Subsequent *Se*-methylation yielded a selenomethionine residue. This method could have general applicability, as the methylation of selenium should be chemoselective at acidic pH.

GENERAL STRATEGIES FOR PEPTIDE LIGATION

A limitation of native chemical ligation is its intrinsic reliance on having a cysteine residue at the ligation juncture. Cysteine is the second least common amino acid, comprising only 1.7% of all residues in proteins (Figure 10). Hence, most proteins cannot be prepared in their native form by any method that allows peptides to be coupled only at cysteine residues. However, nonnative cysteine residues have been added to enable protein synthesis by native chemical ligation (79) or semisynthesis by expressed protein ligation (69, 142).

The addition of nonnative cysteine residues can incur risk. Of the functional groups in the 20 common amino acids, the thiolate of cysteine is by far the most reactive toward disulfide bonds, $O_2(g)$, and other electrophiles (12, 125). In addition, cysteine can suffer base-catalyzed β -elimination to form dehydroalanine, which can undergo further deleterious reactions (50). The side chains of nonnative

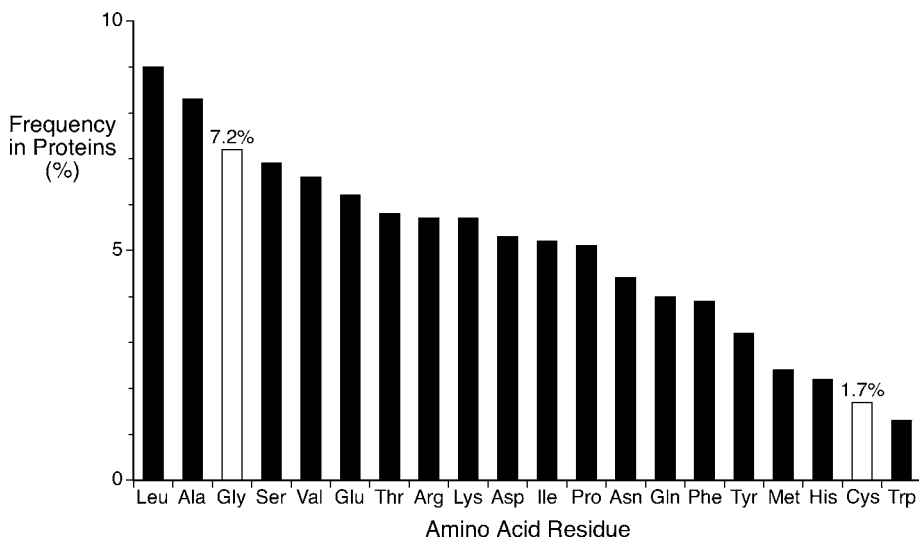


Figure 10 Frequency of occurrence of amino acid residues in the primary structures of 1021 unrelated proteins (89). The most common residues at ligation junctions in synthetic proteins are cysteine (as in native chemical ligation) and glycine (see below).

cysteine residues have been alkylated [e.g., with bromoacetate to form 4-thiahomoglutamate (79)] to minimize these risks.

The removal of the cysteine limitation by applying a more general ligation technology (Figure 2) would extend greatly the utility of protein synthesis (33). An ideal technique would enable complete and rapid ligation between any two amino acid residues without detectable epimerization. Below, we describe three such strategies.

Conformationally Assisted Ligation

A peptide with a C-terminal thioester can be ligated with a second peptide that does not contain a cysteine residue at its N terminus if there is a strong conformational preference for the association of the two fragments (13). For example, chymotrypsin inhibitor 2 has been divided into two peptides that correspond to residues 1–40 and 41–64. These fragments self-associate to form a native-like complex. When fragment 1–40 was synthesized as a C-terminal thioester, complete ligation to fragment 41–64 (residue 41 is methionine) was complete in less than 2 h. This route was used to prepare chymotrypsin inhibitor 2 labeled site-specifically with donor and acceptor dyes for studies of protein folding and stability (42).

The S-peptide (residues 1–20) and S-protein (residues 21–124) fragments of RNase A associate strongly in solution. The incubation of S-peptide having a C-terminal thioester with S-protein induces a conformationally assisted ligation. The ligation was complete in 10 h to yield intact RNase A (13).

Conformationally assisted ligation has limitations. The method can be used only for the ligation of peptides that have an inherent propensity for self-association. In addition, ligations to form an Xaa-Yaa peptide bond are slow when Xaa is a β -branched residue (13).

Removable Auxiliaries

An emerging strategy in protein synthesis has been the use of auxiliaries that act as cysteine surrogates to mediate the chemical ligation of peptide fragments (Figure 11). These auxiliaries are appended synthetically to the N terminus of a

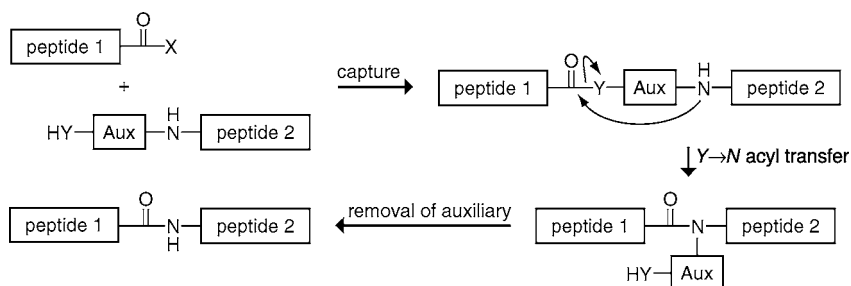


Figure 11 Auxiliary-mediated peptide ligation.

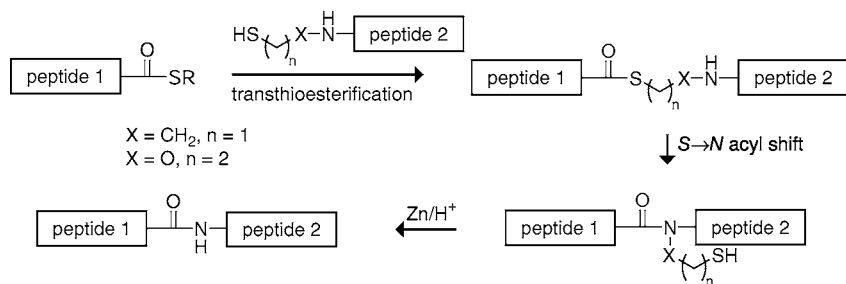


Figure 12 Native chemical ligation with N^α -ethanethiol and N^α -oxyethanethiol auxiliaries.

peptide. A second peptide is activated at its C terminus in a manner that allows capture of the auxiliary. After the capture step, a $Y \rightarrow N$ acyl transfer reaction produces an amide. The final step is removal of the extraneous atoms of the auxiliary, which often involves an additional synthetic step.

The first auxiliaries used for peptide ligation at noncysteine residues were N^α -ethanethiol and N^α -oxyethanethiol (26) (Figure 12). Peptides containing these auxiliaries at their N terminus were synthesized and coupled to peptides with a C-terminal thioester under conditions similar to those used in native chemical ligation. The most efficient couplings occurred when both the N- and C-terminal residues were glycine. Yields were lower if one residue was phenylalanine and the other was glycine. No ligation was detected between two nonglycyl residues. After ligation, the N^α -oxyethanethiol (but not the N^α -ethanethiol) auxiliary was removed efficiently by reduction with zinc under acidic conditions. The N^α -oxyethanethiol auxiliary has also been used in the cyclization of synthetic peptides (127).

N^α -2-Mercaptobenzylamine is the basis for a second class of auxiliaries for peptide ligation (Figure 13). Peptides incorporating these auxiliaries at the N terminus react with thioesters in a manner similar to native chemical ligation (compare with Figure 4). Acyl transfer occurs via a six-membered ring. The auxiliaries, with appropriate phenyl-ring substitution, can be removed under acidic conditions after acyl transfer.

Offer & Dawson have reported three N^α -2-mercaptobenzylamine auxiliaries. The simple mercaptobenzylamine auxiliary ($R' = R'' = R''' = \text{H}$ in Figure 13) was the first of this class to be used for peptide ligation (114). This auxiliary enabled acyl transfer if a glycine residue flanked the ligation junction and a β -branched residue did not. The simple mercaptobenzylamine auxiliary could not be removed by acid treatment of the rearranged peptides. The dimethoxyl auxiliary ($R' = R'' = \text{OMe}$ and $R''' = \text{H}$ in Figure 13) was reported subsequently, although its use was demonstrated in only a single peptide ligation in which both residues flanking the ligation junction were glycine (71). Treatment with a trifluoromethanesulfonic acid/thioanisole/trifluoroacetic acid cocktail removed the dimethoxyl auxiliary completely from the acyl transfer product to give a native peptide bond.

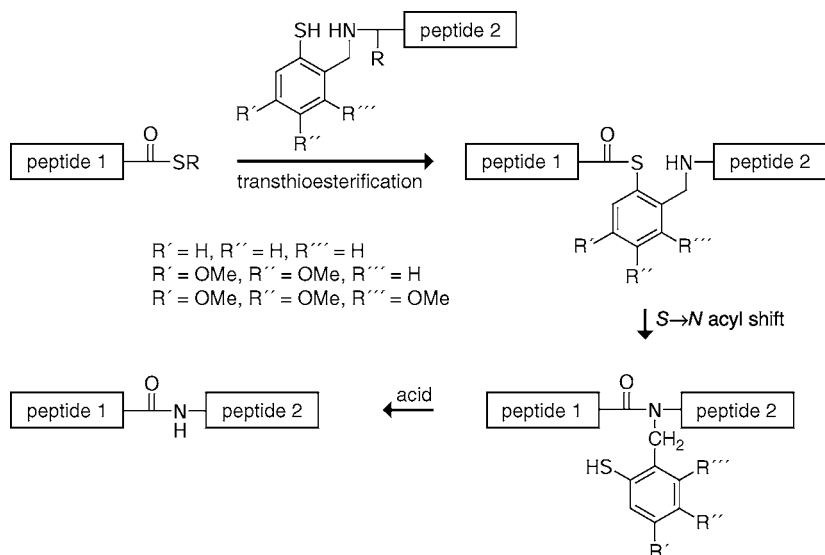


Figure 13 N^{α} -2-Mercaptobenzylamine-assisted chemical ligations.

The trimethoxyl auxiliary ($R' = R'' = R''' = \text{OMe}$ in Figure 13) is the most recently reported ligation agent of this class (113). The half-lives for acyl transfer were 0.2, 2.0, and 5.0 h for Gly-Gly, Lys-Gly, and Gly-Ala, respectively. No acyl transfer occurred between alanine residues. The auxiliary could be removed after acyl transfer using either TFA (trifluoroacetic acid) or HF (hydrofluoric acid). The 63-residue SH3 (Src homology 3) domain of α -spectrin was prepared by chemical ligation between Lys-27 and Gly-28 using the trimethoxyl auxiliary.

An N^{α} -(1-phenyl-2-mercaptoethyl) group represents a third class of auxiliary-assisted chemical ligation (Figure 14). These agents are used in a manner similar to that of the N^{α} -2-mercaptobenzylamine auxiliaries (compare with Figure 13). The N^{α} -(1-phenyl-2-mercaptoethyl) auxiliaries incorporate a more nucleophilic alkyl thiol in their design to enhance the efficiency of transthioesterification in the capture step, and mediate acyl transfer via a more favorable five-membered ring. Again, substitution on the phenyl ring affects the lability of the auxiliary after acyl transfer.

Initial studies of the N^{α} -(1-phenyl-2-mercaptoethyl) class of auxiliaries were undertaken using compounds with an unsubstituted ring ($R' = R'' = H$ in Figure 14) (87). Transthioesterification occurred efficiently regardless of sequence; thus acyl transfer limits the ligation rate. Half-lives for acyl transfer for Gly-Gly and Ala-Gly coupling were 30–40 and 120–140 min, respectively; acyl transfer could not be used to induce Gly-Ala and Ala-Ala coupling. No removal of the N^{α} -(1-phenyl-2-mercaptoethyl) auxiliary from rearranged products has been reported to date.

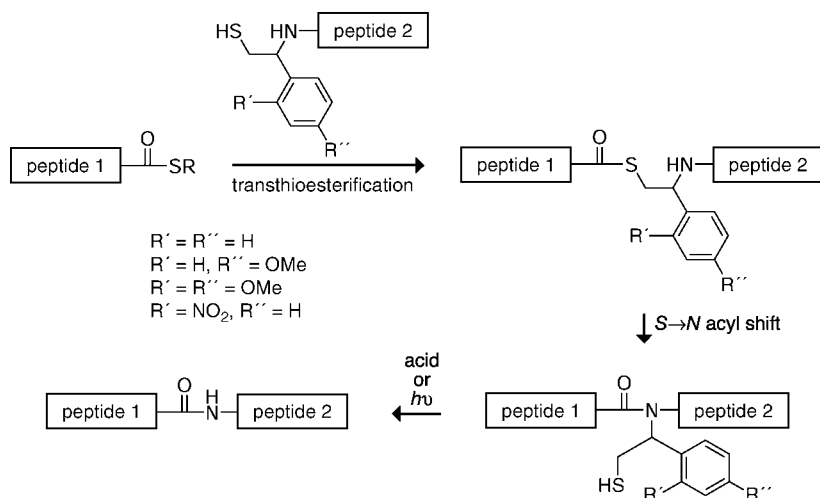


Figure 14 N^α -(1-Phenyl-2-mercaptoethyl)-assisted chemical ligation.

Two ring-substituted forms of the same auxiliary were reported concurrently. Both the N^α -(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary ($\text{R}' = \text{H}$ and $\text{R}'' = \text{OMe}$ in Figure 14) and the N^α -(1-(2,4-dimethoxyphenyl)-2-mercaptoethyl) auxiliary ($\text{R}' = \text{OMe}$ and $\text{R}'' = \text{OMe}$) ligate efficiently at Xaa-Gly sites where Xaa was Gly, His, Ala, or Lys. Acyl transfer yields were 76% to 98%, with the typical rearrangement completed in 16 h. Ligations using the dimethoxyl auxiliary for Xaa-Gly (where Xaa \neq Gly) required longer reaction times (40 h). The auxiliaries could be removed after acyl transfer by using HF (methoxyl) or TFA (dimethoxyl). These auxiliaries have been applied to the total chemical synthesis of cytochrome b562 (106 residues) (85) and for the cyclization of 11-residue cyclic peptides (28).

The N^α -(1-(2-nitrophenyl)-2-mercaptoethyl) auxiliary ($\text{R}' = \text{NO}_2$ and $\text{R}'' = \text{H}$ in Figure 14) is photolabile. Capture and acyl transfer was demonstrated by the ligation of two 10-residue peptides to produce a Gly-Gly peptide bond. Acyl transfer was completed after 24 h. The auxiliary was removed efficiently by photochemical irradiation at 365 nm for 6 h in an aqueous buffer, pH 7.4. Although the use of the nitro auxiliary is still subject to the glycine limitation, it could be removed under far milder conditions than could its analogs.

Staudinger Ligation

The Staudinger reaction has inspired a simple alternative to native chemical ligation with the promise of removing the cysteine limitation. In the Staudinger reaction, a phosphine is used to reduce an azide to an amine: $\text{PR}_3 + \text{N}_3\text{R}' + \text{H}_2\text{O} \rightarrow \text{O} = \text{PR}_3 + \text{H}_2\text{NR}' + \text{N}_2$ (55, 56, 133). This reaction occurs via a stable intermediate, an

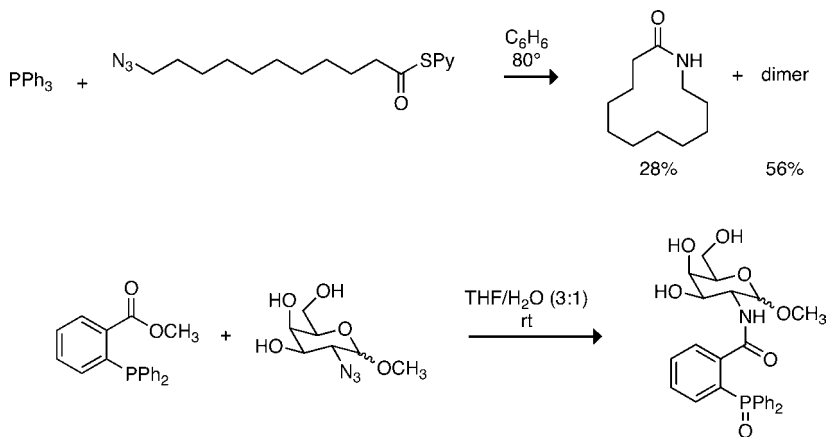


Figure 15 Precedents for the Staudinger ligation of peptides. *Top*: Vilarrasa and coworkers (20); *bottom*: Bertozzi and coworkers (123).

iminophosphorane ($\text{R}_3\text{P}^+ - \text{NR}'$, which is also known, less precisely, as an “aza-ylide”). Vilarrasa and others have shown that the nitrogen of the iminophosphorane can be acylated, both in intermolecular (i.e., three-component) and intramolecular (two-component; Figure 15) ligations (20, 21, 51, 52, 100, 143). Hydrolysis of the resulting amidophosphonium salt gives an amide and a phosphine oxide. By rearranging the functional groups in the two components, Saxon & Bertozzi (123) have shown that the phosphine itself can serve as the acyl donor (Figure 15). Their two-component ligation enables the formation of an amide bond between two molecules, but leaves a phosphine oxide in the amide product.

A “traceless” Staudinger ligation has been developed for the chemical synthesis of proteins (108, 109). In the capture step, a peptide that has a C-terminal phosphinothioester reacts with a peptide that has an N-terminal azide to form an iminophosphorane (Figure 16). Attack of the iminophosphorane nitrogen on the thioester leads to an amidophosphonium salt that is acyclic, which is key for traceless ligation. Hydrolysis of the amidophosphonium salt produces an amide bond between the two peptides and releases a phosphine oxide. No extraneous atoms remain in the amide product.

In essence, the Staudinger ligation of peptides using a phosphinothiol couples the energetics of native chemical ligation with that of the Staudinger reaction. This coupling is apparent in a comparison of functional group interconversions during the two ligation reactions (Figure 17). The dissimilar groups are the reactants and products of the Staudinger reaction, which is exergonic (55, 56). Accordingly, the thermodynamic driving force of the Staudinger ligation exceeds that of native chemical ligation. The Staudinger ligation is being used increasingly in the chemical synthesis of proteins, as well as in other areas of synthetic chemistry and chemical biology (80).

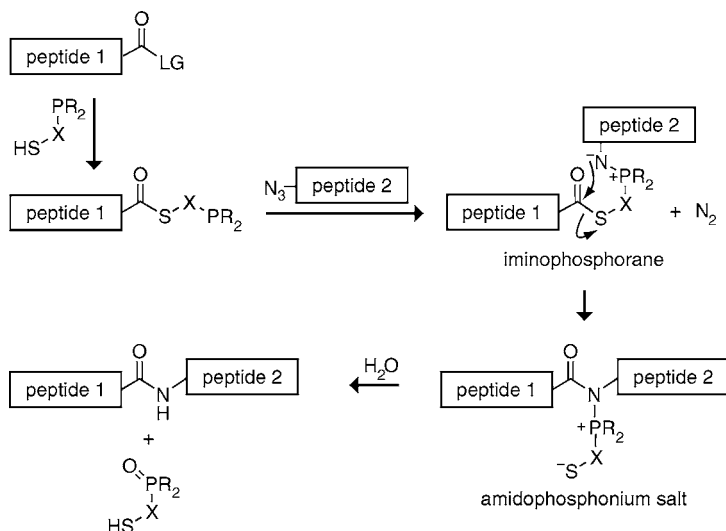


Figure 16 Staudinger ligation of peptides mediated by a phosphinothiol.

Several phosphinothiol compounds have been assessed for the ability to mediate the Staudinger ligation. Compound **1** was the first phosphinothiol tested in this capacity (108) (Figure 18). Its coupling to *N*^α-acetyl amino acids and incubation of the resulting thioester with N₃GlyNH₂Bn [where N₃Gly refers to a N₃CH₂C(O) unit] resulted in dipeptide formation, albeit in low yields (≤35%). Still, phosphinothiol **1** mediated the first traceless Staudinger ligation (108).

Phosphinothiol **2** was a second-generation phosphinothiol (109, 107, 130, 131). Staudinger ligation with phosphinothiol **1** occurs through an intermediate with a six-membered ring. Reducing the size of this ring would bring the nucleophilic

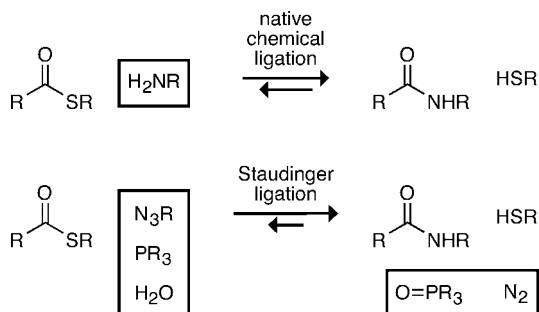


Figure 17 Comparison of functional group interconversions during native chemical ligation and phosphinothiol-mediated Staudinger ligation. Dissimilar groups are boxed.

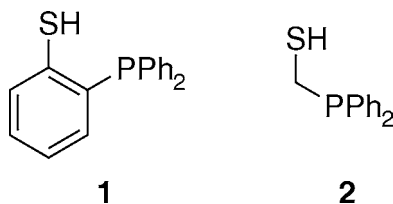


Figure 18 Phosphinothiols used in the Staudinger ligation of peptides.

iminophosphorane nitrogen more proximal to the electrophilic thioester carbon and thus could improve the yields for the Staudinger ligation products. To access a transition state with a smaller ring, the *o*-phenyl group of phosphinothiol **1** was replaced with a single methylene group. Thioesters of phosphinothiol **2** with AcOH, AcGlyOH, and AcPheOH were prepared either by transthioesterification or by coupling with dicyclohexylcarbodiimide. These phosphinothioesters formed dipeptides with N₃GlyNHBn in >90% yields (109). Additionally, AcGlySCH₂PPh₂ was coupled with the D- and L-enantiomers of N₃Ser(Bn)NHBn, N₃PheNHBn, and N₃Asp(OBn)NHBn (131). These ligations, which likewise proceeded in >90% yield, confirmed that the Staudinger ligation retains the stereochemistry at the α -carbon of the azido acid. As with known auxiliary-assisted ligations (see above), the Staudinger ligation with phosphinothiols **1** and **2** is less efficient when a glycine residue is not at the ligation junction (B.L. Nilsson, M.B. Soellner & R.T. Raines, unpublished results).

The Staudinger ligation with phosphinothiol **2** has been exploited in the assembly of RNase A (107). After folding, this semisynthetic RNase A had a molecular mass and enzymatic activity that were indistinguishable from that of biosynthetic RNase A. Moreover, the semisynthesis incorporated a ¹³C-enriched proline residue at position 114 for nuclear magnetic resonance spectroscopic analysis of protein conformation. This semisynthesis is unique in invoking four distinct chemical processes to create the peptide bonds of a protein: mRNA translation by the ribosome, solid-phase peptide synthesis, expressed protein ligation, and the Staudinger ligation (Table 1). Access to this arsenal greatly empowers modern protein chemists.

TABLE 1 Methods to form peptide bonds during the semisynthesis of RNase A (107)

Bond	Ligation method
1–109	mRNA translation by ribosome (sequential)
109–110	Expressed protein ligation (convergent)
110–111	Solid-phase peptide synthesis (sequential)
111–112	Staudinger ligation (convergent)
112–124	Solid-phase peptide synthesis (sequential)

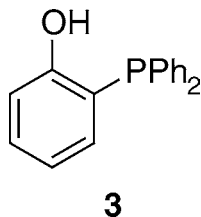


Figure 19 Phosphinophenol used in the Staudinger ligation of peptides at nonglycyl residues.

Phosphinophenol **3** has been used to promote the Staudinger ligation of two nonglycyl amino acids (**93**) (Figure 19). Under specialized reaction conditions, Gly-Phe and Ala-Phe couplings occurred in 32% to 36% yields. The reaction conditions, however, are not readily adapted to unprotected peptides in aqueous solution, as amino groups (e.g., N-terminal and lysyl side chain) are acylated by esters of phosphinophenol **3** (**93**). Moreover, thioester fragments generated by expressed protein ligation or other methods cannot be converted readily to esters with phosphinophenol **3**, which highlights an intrinsic advantage of phosphinothiols in mediating the Staudinger ligation (Figure 16). Still, the success with phosphinophenol **3** provides an important precedent in the development of the Staudinger ligation as a sequence-independent method for peptide ligation.

PROSPECTUS

The development and refinement of peptide ligation methods for the convergent synthesis of proteins has entered an explosive phase of growth. Native chemical ligation and expressed protein ligation have already enabled the synthesis and semisynthesis of more than 100 proteins (38, 103). The junctions between peptides in these proteins, however, must be at the N terminus of a cysteine residue. Removable auxiliaries and the Staudinger ligation provide the means to ligate peptides between any two residues in a protein sequence. The ligation of peptides on a solid support is likely to facilitate protein folding [by minimizing aggregation (40)] and enable protein production to be automated from start to finish. Accordingly, synthetic chemistry is poised to make ever-increasing contributions to protein science for the foreseeable future.

ACKNOWLEDGMENTS

We are grateful to T.W. Muir and L.L. Kiessling for contributive discussions. Work on protein chemistry in the laboratory of R.T.R. has been supported by a fellowship from the John Simon Guggenheim Memorial Foundation and a grant from the National Institutes of Health (GM44783).

The Annual Review of Biophysics and Biomolecular Structure is online at
<http://biophys.annualreviews.org>

LITERATURE CITED

1. Aimoto S. 1999. Polypeptide synthesis by the thioester method. *Biopolymers* 51:247–65
2. Albericio F. 2004. Developments in peptide and amide synthesis. *Curr. Opin. Chem. Biol.* 8:211–21
3. Albericio F, Carpino LA. 1997. Coupling reagents and activation. *Methods Enzymol.* 289:104–26
4. Albericio F, Lloyd-Williams P, Giralt E. 1997. Convergent solid-phase peptide synthesis. *Methods Enzymol.* 289:313–36
5. Arner ES, Sarioglu H, Lottspeich F, Holmgren A, Bock A. 1999. High-level expression in *Escherichia coli* of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the selA, selB and selC genes. *J. Mol. Biol.* 292:1003–16
6. Arnold U, Hinderaker MP, Köditz J, Golbik R, Ulbrich-Hoffmann R, Raines RT. 2003. Protein prosthesis: A nonnatural residue accelerates folding and increases stability. *J. Am. Chem. Soc.* 125:7500–1
7. Arnold U, Hinderaker MP, Nilsson BL, Huck BR, Gellman SH, Raines RT. 2002. Protein prosthesis: a semisynthetic enzyme with a β -peptide reverse turn. *J. Am. Chem. Soc.* 124:368–69
8. Ball HL, King DS, Cohen FE, Prusiner SB, Baldwin MA. 2001. Engineering the prion protein using chemical synthesis. *J. Pept. Res.* 58:357–74
9. Bang D, Kent S. 2004. A one-pot total synthesis of crambin. *Angew. Chem. Int. Ed. Engl.* 43:2534–38
10. Becker CFW, Hunter CL, Seidel RP, Kent SBH, Goody RS, Engelhard M. 2001. A sensitive fluorescence monitor for the detection of activated Ras: total chemical synthesis of site-specifically labeled Ras binding domain of c-Raf1 immobilized on a surface. *Chem. Biol.* 8:243–52
11. Becker CFW, Hunter CL, Seidel R, Kent SBH, Goody RS, Engelhard M. 2003. Total chemical synthesis of a functional interacting protein pair: the protooncogene H-Ras and the Ras-binding domain of its effector c-Raf1. *Proc. Natl. Acad. Sci. USA* 100:5075–80
12. Bednar RA. 1990. Reactivity and pH dependence of thiol conjugation to *N*-ethylmaleimide: detection of a conformational change in chalcone isomerase. *Biochemistry* 29:3684–90
13. Beligere GS, Dawson PE. 1999. Conformationally assisted protein ligation using C-terminal thioester peptides. *J. Am. Chem. Soc.* 121:6332–33
14. Bergmann M, Zervas L. 1932. Über ein allgemeines Verfahren der Peptid-Synthese. *Ber. Deutsch. Chem. Ges.* 65:1192–201
15. Berry SM, Gieselman MD, Nilges MJ, van der Donk WA, Lu Y. 2002. An engineered azurin variant containing a selenocysteine copper ligand. *J. Am. Chem. Soc.* 124:2084–85
16. Besse D, Siedler F, Diercks T, Kessler H, Moroder L. 1997. The redox potential of selenocysteine in unconstrained cyclic peptides. *Angew. Chem. Int. Ed. Engl.* 36:883–85
17. Blaschke UK, Cotton GJ, Muir TW. 2000. Synthesis of multi-domain proteins using expressed protein ligation: strategies for segmental isotopic labeling of internal regions. *Tetrahedron* 56:9461–70
18. Bordusa F. 2002. Proteases in organic synthesis. *Chem. Rev.* 102:4817–68
19. Borgia JA, Fields GB. 2000. Chemical synthesis of proteins. *Trends Biotechnol.* 18:243–51

20. Bosch I, Romea P, Urpí F, Vilarrasa J. 1993. Alternative procedures for the macrolactamisation of ω -azido acids. *Tetrahedron Lett.* 34:4671–74
21. Bosch I, Urpí F, Vilarrasa J. 1995. Epimerization-free peptide formation from carboxylic-acid anhydrides and azido derivatives. *J. Chem. Soc. Chem. Commun.* pp. 91–92
22. Bray BL. 2003. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat. Rev. Drug Discov.* 2:587–93
23. Brik A, Keinan E, Dawson PE. 2000. Protein synthesis by solid-phase chemical ligation using a safety catch linker. *J. Org. Chem.* 65:3829–35
24. Buskirk AR, Ong Y-C, Gartner ZJ, Liu DR. 2004. Directed evolution of ligand dependence: small-molecule-activated protein splicing. *Proc. Natl. Acad. Sci. USA* 101:10505–10
25. Camarero JA, Cotton GJ, Adeva A, Muir TW. 1998. Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.* 51:303–16
26. Canne LE, Bark SJ, Kent SBH. 1996. Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* 118:5891–96
27. Canne LE, Botti P, Simon RJ, Chen Y, Dennis EA, Kent SBH. 1999. Chemical protein synthesis by solid phase ligation of unprotected peptide segments. *J. Am. Chem. Soc.* 121:8720–27
28. Cardona VMF, Hartley O, Botti P. 2003. Synthesis of cyclic peptides from unprotected precursors using removable N^{α} -(1-(4-methoxyphenyl)-2-mercaptoethyl) auxiliary. *J. Pept. Res.* 61:152–57
29. Casi G, Hilvert D. 2003. Convergent protein synthesis. *Curr. Opin. Struct. Biol.* 13: 589–94
30. Chang TK, Jackson DY, Burnier JP, Wells JA. 1994. Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA* 91:12544–48
31. Cole PA, Courtney AD, Shen K, Zhang Z, Qiao Y, et al. 2003. Chemical approaches to reversible protein phosphorylation. *Acc. Chem. Res.* 36:444–52
32. Coltart DM. 2000. Peptide segment coupling by prior ligation and proximity-induced intramolecular acyl transfer. *Tetrahedron* 56:3449–91
33. Cotton GJ, Muir TW. 1999. Peptide ligation and its application to protein engineering. *Chem. Biol.* 6:R247–56
34. Cotton GJ, Muir TW. 2000. Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed ligation. *Chem. Biol.* 7:253–61
35. Cowburn D, Muir TW. 2001. Segmental isotopic labeling using expressed protein ligation. *Methods Enzymol.* 339:41–54
36. David R, Richter MP, Beck-Sickinger AG. 2004. Expressed protein ligation. Method and applications. *Eur. J. Biochem.* 271:663–77
37. Dawson PE, Churchill MJ, Ghadiri MR, Kent SBH. 1997. Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* 119:4325–29
38. Dawson PE, Kent SBH. 2000. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* 69:923–60
39. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. 1994. Synthesis of proteins by native chemical ligation. *Science* 266: 776–79
40. De Bernardez Clark E, Schwarz E, Rudolph R. 1999. Inhibition of aggregation side reactions during in vitro protein folding. *Methods Enzymol.* 309:217–36
41. del Milton RC, Milton SCF, Kent SBH. 1992. Total chemical synthesis of a D-enzyme: The enantiomers of HIV-1 protease show demonstration of reciprocal chiral substrate specificity. *Science* 256:1445–48
42. Deniz AA, Laurence TA, Beligere GS, Dahan M, Martin AB, et al. 2000. Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of

- chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. USA* 97:5179–84
43. Du Vigneaud V, Ressler C, Swan JM, Roberts CW, Katsoyannis PG, Gordon S. 1953. The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* 75:4879–80
44. Eom KD, Miao Z, Yang Z-L, Tam JP. 2003. Tandem ligation of multipartite peptides with cell-permeable activity. *J. Am. Chem. Soc.* 125:73–82
45. Evans TC, Benner J, Xu MQ. 1998. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* 7:2256–64
46. Fischer E. 1907. Synthetical chemistry in its relation to biology (Faraday Lecture). *J. Chem. Soc. Chem. Commun.* 91:1749–65
47. Fischer E. 1966. Syntheses in the purine and sugar group. In *Nobel Lectures, Chemistry 1901–1921*, pp. 21–35. Amsterdam: Elsevier
48. Fischer E, Fourneau E. 1901. Ueber einige Derivate des Glykocolls. *Ber. Deutsch. Chem. Ges.* 34:2868–77
49. Fotouhi N, Galakatos NG, Kemp DS. 1989. Peptide synthesis by prior thiol capture. 6. Rates of the disulfide bond forming capture reaction and demonstration of the overall strategy by synthesis of the C-terminal 29-peptide sequence of BPTI. *J. Org. Chem.* 54:2803–17
50. Friedman M. 1999. Chemistry, biochemistry, nutrition, and microbiology of lysinoalanine, lanthionine, and histidinoalanine in food and other proteins. *J. Agric. Food Chem.* 47:1295–319
51. Garcia J, Urpí F, Vilarrasa J. 1984. New synthetic tricks—triphenylphosphine-mediated amide formation from carboxylic-acids and azides. *Tetrahedron Lett.* 25:4841–44
52. Garcia J, Vilarrasa J, Bordas X, Banaszek A. 1986. New synthetic tricks—one-pot preparation of N-substituted phthalimides from azides and phthalic-anhydride. *Tetrahedron Lett.* 27:639–40
53. Gieselmann MD, Xie L, van der Donk WA. 2001. Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.* 3:1331–34
54. Giriat I, Muir TW. 2003. Protein semi-synthesis in living cells. *J. Am. Chem. Soc.* 125:7180–81
55. Gololobov YG, Kasukhin LF. 1992. Recent advances in the Staudinger reaction. *Tetrahedron* 48:1353–406
56. Gololobov YG, Zhmurova IN, Kasukhin LF. 1981. Sixty years of Staudinger reaction. *Tetrahedron* 37:437–72
57. Goody RS, Alexandrov K, Engelhard M. 2002. Combining chemical and biological techniques to produce modified proteins. *ChemBioChem.* 3:399–403
58. Grogan MJ, Pratt MR, Marcaurelle LA, Bertozzi CR. 2002. Homogeneous glycopeptides and glycoproteins for biological investigation. *Annu. Rev. Biochem.* 71:593–634
59. Gutte B, Merrifield RB. 1969. The total synthesis of an enzyme with ribonuclease A activity. *J. Am. Chem. Soc.* 91:501–2
60. Gutte B, Merrifield RB. 1971. The synthesis of ribonuclease A. *J. Biol. Chem.* 246:1922–41
61. Guy CA, Fields GB. 1997. Trifluoroacetic acid cleavage and deprotection of resin-bound peptides following synthesis by Fmoc chemistry. *Methods Enzymol.* 289:67–83
62. Hackeng TM, Fernandez JA, Dawson PE, Kent SB, Griffin JH. 2000. Chemical synthesis and spontaneous folding of a multidomain protein: anticoagulant microprotein S. *Proc. Natl. Acad. Sci. USA* 97:14074–78
63. Hackeng TM, Griffin JH, Dawson PE. 1999. Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* 96:10068–73
64. Hackeng TM, Rosing J, Spronk HMH, Vermeer C. 2001. Total chemical synthesis of human matrix Gla protein. *Protein Sci.* 10:864–70

65. Hendrickson TL, de Crecy-Lagard V, Schimmel P. 2004. Incorporation of nonnatural amino acids into proteins. *Annu. Rev. Biochem.* 73:147–76
66. Hirschmann R, Nutt RF, Veber DF, Vitali RA, Varga SL, et al. 1969. Studies on the total synthesis of an enzyme. V. The preparation of enzymatically active material. *J. Am. Chem. Soc.* 91:507–8
67. Hondal RJ, Nilsson BL, Raines RT. 2001. Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* 123:5140–41
68. Hondal RJ, Raines RT. 2002. Semisynthesis of proteins containing selenocysteine. *Methods Enzymol.* 347:70–83
69. Huse M, Muir TW, Xu L, Chen YG, Kuriyan J, Massague J. 2001. The TGF β receptor activation process: an inhibitor-to substrate-binding switch. *Mol. Cell* 8:671–82
70. Jackson DY, Burnier J, Quan C, Stanley M, Tom J, Wells JA. 1994. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* 266:243–47
71. Kawakami T, Akaji K, Aimoto S. 2001. Peptide bond formation mediated by 4,5-dimethoxy-2-mercaptobenzylamine after periodate oxidation of the N-terminal serine residue. *Org. Lett.* 3:1403–5
72. Kemp DS, Carey RI. 1989. Boc-L-Dmt-OH as a fully N,S-blocked cysteine derivative for peptide synthesis by prior thiol capture. Facile conversion of N-terminal Boc-L-Dmt-peptides to H-Cys(Scm)-peptides. *J. Org. Chem.* 54:3640–46
73. Kemp DS, Carey RI. 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.* 58:2216–22
74. Kemp DS, Fotouhi N. 1987. Peptide synthesis by prior thiol capture. V. The scope and control of disulfide interchange during the acyl transfer step. *Tetrahedron Lett.* 28:4637–40
75. Kemp DS, Galakatos NG. 1986. Peptide-synthesis by prior thiol capture. 1. A convenient synthesis of 4-hydroxy-6-mercaptodibenzofuran and novel solid-phase synthesis of peptide-derived 4-(acyloxy)-6-mercaptodibenzofurans. *J. Org. Chem.* 51:1821–29
76. Kemp DS, Galakatos NG, Bowen B, Tan K. 1986. Peptide synthesis by prior thiol capture. 2. Design of templates for intramolecular O,N-acyl transfer. 4,6-Disubstituted dibenzofurans as optimal spacing elements. *J. Org. Chem.* 51:1829–39
77. Kemp DS, Galakatos NG, Dranginis S, Ashton C, Fotouhi N, Curran TP. 1986. Peptide synthesis by prior thiol capture. 4. Amide bond formation: the effect of a side-chain substituent on the rates of intramolecular O,N-acyl transfer. *J. Org. Chem.* 51:3320–24
78. Kent S. 2003. Total chemical synthesis of enzymes. *J. Pept. Sci.* 9:574–93
79. Kochendoerfer GG, Chen S-Y, Mao F, Cressman S, Traviglia S, et al. 2003. Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* 299:884–87
80. Kohn M, Breinbauer R. 2004. The Staudinger ligation—a gift to chemical biology. *Angew. Chem. Int. Ed. Engl.* 43:2–12
81. Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehtab O, et al. 2003. Characterization of mammalian selenoproteomes. *Science* 300:1439–43
82. Kryukov GV, Gladyshev VN. 2004. The prokaryotic selenoproteome. *EMBO Rep.* 5:538–43
83. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921
84. Lloyd-Williams P, Albericio F, Giralt E. 1992. Convergent solid-phase peptide synthesis. *Tetrahedron* 49:11065–133

85. Low DW, Hill MG, Carrasco MR, Kent SBH, Botti P. 2001. Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci. USA* 98:6554–59
86. Marcaurelle LA, Mizoue LS, Wilken J, Oldham L, Kent SBH, et al. 2001. Chemical synthesis of lymphotactin: a glycosylated chemokine with a C-terminal mucin-like domain. *Chem. Eur. J.* 7:1129–32
87. Marinz C, Bark SJ, Offer J, Dawson PE. 2001. A new scaffold for amide ligation. *Bioorg. Med. Chem.* 9:2323–28
88. McBride BJ, Kemp DS. 1987. Peptide synthesis by prior thiol capture. III. Assessment of levels of racemization during two typical thiol capture coupling reactions. *Tetrahedron Lett.* 28:3435–38
89. McCaldon P, Argos P. 1988. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* 4:99–122
90. Meldal M. 1997. Properties of solid supports. *Methods Enzymol.* 289:83–104
91. Mendel D, Cornish VW, Schultz PG. 1995. Site-directed mutagenesis with an expanded genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 24:435–62
92. Menger FM. 1985. On the source of intramolecular and enzymatic reactivity. *Acc. Chem. Res.* 18:128–34
93. Merckx R, Rijkers DTS, Demink J, Liskamp RMJ. 2003. Chemoselective coupling of peptide fragments using the Staudinger ligation. *Tetrahedron Lett.* 44:4515–18
94. Merrifield B. 1986. Solid phase synthesis. *Science* 232:341–47
95. Merrifield B. 1996. The chemical synthesis of proteins. *Protein Sci.* 5:1947–51
96. Merrifield RB. 1963. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* 85:2149–54
97. Merrifield RB, Stewart JM, Jernberg N. 1966. Instrument for automated synthesis of peptides. *Anal. Chem.* 38:1905–14
98. Miller M, Schneider J, Sathyanarayana BK, Toth MV, Marshall GR, et al. 1989. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* 246:1149–52
99. Miranda LP, Alewood PF. 1999. Accelerated chemical synthesis of peptides and small proteins. *Proc. Natl. Acad. Sci. USA* 96:1181–86
100. Mizuno M, Muramoto I, Kobayashi K, Yaginuma H, Inazu T. 1999. A simple method for the synthesis of *N*^β-glycosylated-asparagine and -glutamine derivatives. *Synth. Stuttg.* 162–65
101. Mootz HD, Blum ES, Tyszkiewicz AB, Muir TW. 2003. Conditional protein splicing: a new tool to control protein structure and function in vitro and in vivo. *J. Am. Chem. Soc.* 125:10561–69
102. Mootz HD, Muir TW. 2002. Protein splicing triggered by a small molecule. *J. Am. Chem. Soc.* 124:9044–45
103. Muir TW. 2003. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* 72:249–89
104. Muir TW, Dawson PE, Kent SBH. 1997. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* 289:266–98
105. Muir TW, Sondhi D, Cole PA. 1998. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. USA* 95:6705–10
106. Nakatsuka T, Sasaki T, Kaiser ET. 1987. Peptide segment coupling catalyzed by the semisynthetic enzyme thiostubtilisin. *J. Am. Chem. Soc.* 109:308–10
107. Nilsson BL, Hondal RJ, Soellner MB, Raines RT. 2003. Protein assembly by orthogonal chemical ligation methods. *J. Am. Chem. Soc.* 125:5268–69
108. Nilsson BL, Kiessling LL, Raines RT. 2000. Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.* 2:1939–41
109. Nilsson BL, Kiessling LL, Raines RT. 2001. High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide. *Org. Lett.* 3:9–12

110. Nishiuchi Y, Inui T, Nishio H, Bódi J, Kimura T, et al. 1998. Chemical synthesis of the precursor molecule of the *Aequorea* green fluorescent protein, subsequent folding, and development of fluorescence. *Proc. Natl. Acad. Sci. USA* 95: 13549–54
111. Noren CJ, Wang JM, Perler FB. 2000. Dissecting the chemistry of protein splicing and its applications. *Angew. Chem. Int. Ed. Engl.* 39:450–66
112. Odaert B, Jean F, Boutillon C, Buisine E, Melnyk O, et al. 1999. Synthesis, folding, and structure of the β -turn mimic modified B1 domain of streptococcal protein G. *Protein Sci.* 8:2773–83
113. Offer J, Boddy CNC, Dawson PE. 2002. Extending synthetic access to proteins with a removable acyl transfer auxiliary. *J. Am. Chem. Soc.* 124:4642–46
114. Offer J, Dawson PE. 2000. N^{α} -2-Mercaptobenzylamine-assisted chemical ligation. *Org. Lett.* 2:23–26
115. Ottesen JJ, Huse M, Sekedat MD, Muir TW. 2004. Semisynthesis of phospho-variants of Smad2 reveals a substrate preference of the activated T β RI kinase. *Biochemistry* 43:5698–706
116. Page MI, Jencks WP. 1971. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. USA* 68:1678–83
117. Paulus H. 2000. Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* 69:447–96
118. Pellois JP, Hahn ME, Muir TW. 2004. Simultaneous triggering of protein activity and fluorescence. *J. Am. Chem. Soc.* 126:7170–71
119. Quaderer R, Sewing A, Hivert D. 2001. Selenocysteine-mediated native chemical ligation. *Helv. Chim. Acta* 84:1197–206
120. Roelfes G, Hilvert D. 2003. Incorporation of selenomethionine into proteins through selenohomocysteine-mediated ligation. *Angew. Chem. Int. Ed. Engl.* 42: 2275–77
121. Romanelli A, Shekhtman A, Cowburn D, Muir TW. 2004. Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc. Natl. Acad. Sci. USA* 101:6397–402
122. Sato T, Aimoto S. 2003. Use of thiosulfonate for the protection of thiol groups in peptide ligation by the thioester method. *Tetrahedron Lett.* 44:8085–87
123. Saxon E, Bertozzi CR. 2000. Cell surface engineering by a modified Staudinger reaction. *Science* 287:2007–10
124. Scheibner KA, Zhang Z, Cole PA. 2003. Merging fluorescence resonance energy transfer and expressed protein ligation to analyze protein-protein interactions. *Anal. Biochem.* 317:226–32
125. Schneider CH, de Weck AL. 1968. Studies on the direct neutral penicilloylation of functional groups occurring on proteins. *Biochim. Biophys. Acta* 168:27–35
126. Severinov K, Muir TW. 1998. Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem.* 273:16205–9
127. Shao Y, Lu W, Kent SBH. 1998. A novel method to synthesize cyclic peptides. *Tetrahedron Lett.* 39:3911–14
128. Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. 1999. Fmoc-based synthesis of peptide- α thioesters: application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.* 121:11684–89
129. Sievers A, Beringer M, Rodnina MV, Wolfenden R. 2004. The ribosome as an entropy trap. *Proc. Natl. Acad. Sci. USA* 101:7897–901
130. Soellner MB, Dickson KA, Nilsson BL, Raines RT. 2003. Site-specific protein immobilization by Staudinger ligation. *J. Am. Chem. Soc.* 125:11790–91
131. Soellner MB, Nilsson BL, Raines RT. 2002. Staudinger ligation of α -azido acids

- retains stereochemistry. *J. Org. Chem.* 67: 4993–96
132. Songster MF, Barany G. 1997. Handles for solid-phase peptide synthesis. *Methods Enzymol.* 289:126–74
133. Staudinger H, Meyer J. 1919. New organic compounds of phosphorus. III. Phosphinimethylene derivatives and phosphinimines. *Helv. Chim. Acta* 2:635–46
134. Stewart JM. 1997. Cleavage methods following Boc-based solid-phase peptide synthesis. *Methods Enzymol.* 289:29–44
135. Swain AL, Miller MM, Green J, Rich DH, Schneider J, et al. 1990. X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proc. Natl. Acad. Sci. USA* 87:8805–9
136. Sydor JR, Mariano M, Sideris S, Nock S. 2002. Establishment of intein-mediated protein ligation under denaturing conditions: C-terminal labeling of a single-chain antibody for biochip screening. *Bioconjug. Chem.* 13:707–12
137. Tam JP, Lu YA, Liu CF, Shao J. 1995. Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. USA* 92:12485–89
138. Tam JP, Xu JX, Eom KD. 2001. Methods and strategies of peptide ligation. *Biopolymers* 60:194–205
139. Tam JP, Yu Q. 1998. Methionine ligation strategy in the biomimetic synthesis of parathyroid hormone. *Biopolymers* 46:319–27
140. Tam JP, Yu QT, Miao ZW. 1999. Orthogonal ligation strategies for peptide and protein. *Biopolymers* 51:311–32
141. Tolbert TJ, Wong CH. 2004. Conjugation of glycopeptide thioesters to expressed protein fragments: semisynthesis of glycosylated interleukin-2. *Methods Mol. Biol.* 283:255–66
142. Valiyaveetil FI, Sekedat M, Muir TW, MacKinnon R. 2004. Semisynthesis of a functional K⁺ channel. *Angew. Chem. Int. Ed. Engl.* 43:2504–7
143. Velasco MD, Molina P, Fresneda PM, Sanz MA. 2000. Isolation, reactivity and intramolecular trapping of phosphazide intermediates in the Staudinger reaction of tertiary phosphines with azides. *Tetrahedron* 56:4079–84
144. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. 2001. The sequence of the human genome. *Science* 291:1304–51
145. Warren JD, Miller JS, Keding SJ, Danishefsky SJ. 2004. Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc.* 126: 6576–78
146. Wellings DA, Atherton E. 1997. Standard Fmoc protocols. *Methods Enzymol.* 289:44–67
147. Wieland T, Bokelmann E, Bauer L, Lang HU, Lau H, Schafer W. 1953. Polypeptide syntheses. VIII. Formation of sulfur containing peptides by the intramolecular migration of aminoacyl groups. *Liebigs Ann. Chem.* 583:129–49
148. Wlodawer A, Miller M, Jaskólski M, Sathyanarayana BK, Baldwin E, et al. 1989. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* 245:616–21
149. Wu Z, Prahl A, Powell R, Erickson B, Lubkowski J, Lu W. 2003. From pro defensins to defensins: synthesis and characterization of human neutrophil pro α -defensin-1 and its mature domain. *J. Pept. Res.* 62:53–62
150. Yan LZ, Dawson PE. 2001. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* 123:526–33