C. P. R. Hackenberger and D. Schwarzer

DOI: 10.1002/anie.200801313

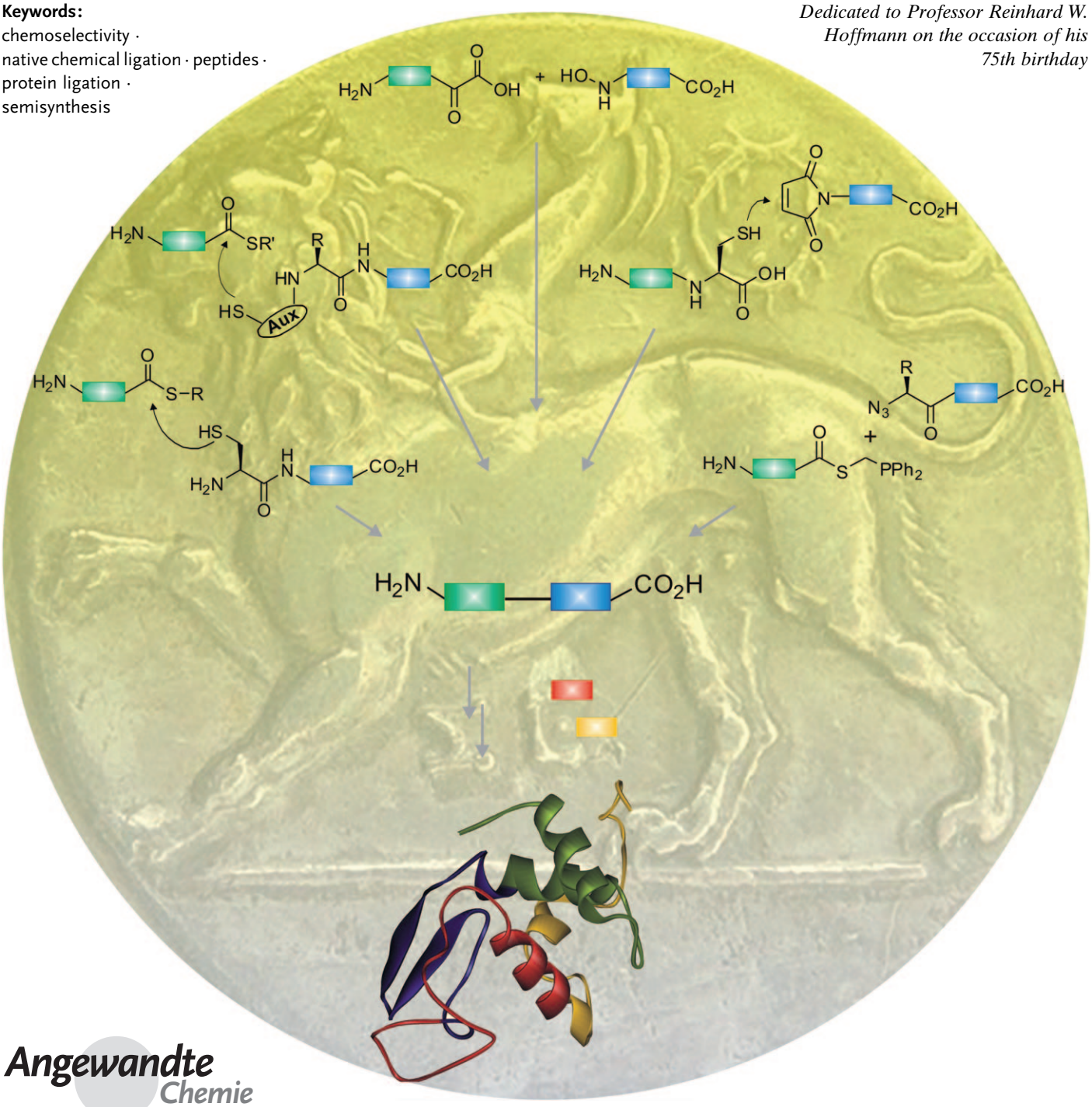
Protein Synthesis

Chemoselective Ligation and Modification Strategies for Peptides and Proteins

Christian P. R. Hackenberger and Dirk Schwarzer**

Keywords:
chemoselectivity · native chemical ligation · peptides · protein ligation · semisynthesis

Dedicated to Professor Reinhard W. Hoffmann on the occasion of his 75th birthday



The investigation of biological processes by chemical methods, commonly referred to as chemical biology, often requires chemical access to biologically relevant macromolecules such as peptides and proteins. Building upon solid-phase peptide synthesis, investigations have focused on the development of chemoselective ligation and modification strategies to link synthetic peptides or other functional units to larger synthetic and biologically relevant macromolecules. This Review summarizes recent developments in the field of chemoselective ligation and modification strategies and illustrates their application, with examples ranging from the total synthesis of proteins to the semisynthesis of naturally modified proteins.

> *"You shall shortly discover that more can be learned by building up than by tearing down, more by joining than by separating, more by reviving the dead than by further killing what is already dead."*[1]
>
> Johann Wolfgang von Goethe, Wilhelm Meister's Wanderjahre

From the Contents

1. Introduction

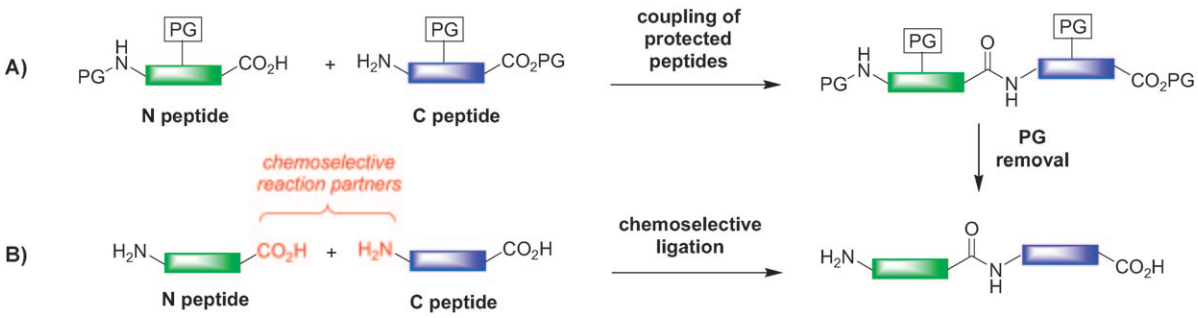
From very early on, chemists identified peptides and proteins as targets for the development of synthetic protocols.[2] Countless research projects have been driven forward by one particular motivation: accessing sufficient peptide or protein material in combination with the potential to site-selectively introduce chemical functionalities to elucidate their structure and function. These routes would thereby complement other routes to obtain protein material such as isolation from natural sources and biochemical expression techniques.

The field of protein synthesis gained further attention when natural protein modifications were identified as the key element for controlling complex cellular processes in living organisms.[3] Researchers thus aimed to control the chemical composition and potential modification of each residue within the polypeptide chain to obtain homogeneous protein material for studying complex biological systems.[3e,f]

Needless to say, the current status of protein synthesis has been preceded by several important investigations—milestones which stimulated and set the basis for further research activities in this field. These important findings started with the first description of the peptide composition by Emil Fischer[2a] followed by early solution-phase condensation strategies.[2b] The development of orthogonal protecting groups later allowed Merrifield to develop the most frequently used method in peptide and protein chemistry, namely, automated solid-phase peptide synthesis (SPPS).[4] Linear peptides of up to 30–40 amino acids can reliably be accessed with SPPS, which led to the coupling of (multiple) peptide fragments as the focus of further investigations.[5] Fragment coupling can be pursued in two ways: in the first route partially protected peptides are coupled to achieve a selective peptide linkage between two termini, followed by the removal of the protecting groups (Scheme 1 A). Although several modes of activation have been reported, including the frequently applied Ag^{I}-mediated coupling of C-terminal N-peptide thioesters with N-terminally unprotected C peptides,[6] this general strategy is limited by the requirement of a final removal of the protecting groups, possible epimerization during the activation at the C terminus of the N-peptide fragment, and occasional aggregation phenomena of longer protected peptide sequences (for an explanation of C and N peptides, see Table 1).[7] Consequently, the second approach has drawn particular attention in the area of peptide and protein chemistry, that is, the *chemoselective ligation of unprotected peptides and proteins* (Scheme 1 B).[8]

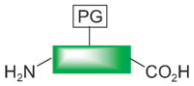
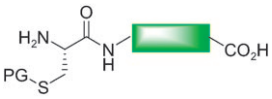
In this Review, we address several important aspects of chemoselective ligation and modification techniques in protein synthesis, which complement excellent previous reviews and highlights on this topic.[9] Various efforts are currently

[*] Dr. C. P. R. Hackenberger
Institute for Chemistry and Biochemistry
Freie Universität Berlin
Takustrasse 3, 14195 Berlin (Germany)
Fax: (+49) 30-838-52551
E-mail: hackenbe@chemie.fu-berlin.de
Dr. D. Schwarzer
Chemical Biology
Leibniz-Institut für Molekulare Pharmakologie (FMP)
Rober-Rössle-Strasse 10, 13125 Berlin (Germany)
Fax: (+49) 30-94793-159
E-mail: schwarzer@fmp-berlin.de



Scheme 1. Fragment coupling strategies: A) Coupling of protected peptides; B) chemoselective ligation of unprotected peptides.

Table 1: Key to the different peptide fragments used in this Review.

Symbol	Species	Remarks
	unprotected peptide	amino acids or functionalized termini important for a scheme are drawn precisely
	protected peptide	PG stands for commonly used protecting groups in Fmoc- or Boc-based SPPS
	partly protected peptide	only the specific protected amino acid is drawn with the protecting group attached
	N peptide	peptides which participate in a ligation reaction; N peptide will end up in the N-terminal region of the ligation product
	C peptide	peptides which participate in a ligation reaction; C peptide will end up in the C-terminal region of the ligation product
	enzyme	
	resin	

address the biological function of proteins by using the synthetic products in subsequent biochemical and biophysical investigations. It is therefore a major goal of this Review to combine an overview of the currently available methods (Sections 2–4) with specific examples in which these methods were directly used for such studies (Sections 6 and 7).

The methods presented in this Review range from purely organic synthetic strategies (Sections 2–4) to molecular biology methods (Sections 2 and 5), thereby stressing the interdisciplinary character of this research area. This Review should give the researcher who is interested in attaching peptides or proteins to their system or intends to derivatize them with desired molecular components an overview of the considerations, requirements, and limitations that have to be taken into account when choosing a particular ligation or conjugation strategy. As some of these methods have only recently been developed, it is important to say that some have not as yet been applied to the ligation or modification of longer unprotected peptide or even protein sequences. Nevertheless, they will be mentioned and discussed here as they demonstrate potential new pathways for the chemoselective ligation and modification of peptides or proteins.[10]

being pursued to further advance chemical ligation methods. These efforts do not only result in inspiring the development of novel methodology to tackle the limitations of current strategies but also result in fascinating studies which precisely

Christian P. R. Hackenberger, born in 1976 in Osnabrück (Germany), studied chemistry at the Albert-Ludwigs Universität Freiburg and at the University of Wisconsin/Madison (MSc in 1999 with Samuel H. Gellman). In 2003 he completed his PhD with Carsten Bolm at the RWTH Aachen. After postdoctoral research with Barbara Imperiali (2003–2005) at the Massachusetts Institute of Technology he moved to the FU Berlin, where he leads an Emmy-Noether group. Since 2008 he has been spokesman of the graduate college "mulitvalency" within SFB 765. His research interests focus on the development of chemoselective ligation strategies and the synthesis of modified peptides and proteins.

Dirk Schwarzer, born in 1972 in Schleswig (Germany), studied chemistry at the Philipps-Universität in Marburg and completed his PhD in 2002 with Mohamed A. Marahiel. After postdoctoral research with Philip A. Cole at the Johns-Hopkins University, in 2006 he moved to the group of Henning D. Mootz at the Universität Dortmund. Since 2007 he has been the head of an Emmy-Noether Group at the Leibniz-Insititut für Molekulare Pharmakologie (FMP) in Berlin. His research interests focus on the investigation of biological questions by using chemical methods, especially the function of posttranslational protein modifications.

2. Chemoselective Peptide Ligation: Capture/Rearrangement, NCL, and EPL

2.1. Early Observations and Experiments

The hallmark of most ligation strategies known today is a capture step that links two peptides in a chemoselective way followed by an intramolecular rearrangement. This "capture/rearrangement" concept is also employed in native chemical ligation (NCL), which represents the most widely used ligation strategy.

Early experiments that led to the capture/rearrangement method were reported in 1953 by Wieland et al., who investigated the chemical properties of amino acid thioesters.[11] Initially, they reported that thiophenol thioesters can undergo intermolecular aminolysis in the presence of amines to give amides. In contrast to these thiophenol thioesters, a glycine thioester of cysteamine **1** could not be synthesized and isolated as such under neutral pH conditions. This observation could be attributed to the additional amino group of cysteamine located in proximity to the thiol moiety. As a result, a rapid intramolecular S→N shift occurred to yield the corresponding amide **2**. This rearrangement occurred under mild acidic conditions and was accelerated at higher pH values. The authors concluded that the observed reaction could be used for peptide synthesis. For this purpose the previously discussed intramolecular rearrangement was combined with an intermolecular thiol–thioester exchange: Specifically, a Val-thiophenol thioester **3** was synthesized and treated with cysteine (**4**; Scheme 2A). The highly reactive aryl thioester rapidly exchanged with the thiol moiety of Cys, which represents the capture step. The resulting Val-S-Cys thioester **5** subsequently rearranged to form a Val-Cys dipeptide (**6**) linked to a native peptide bond.[11]

2.2. Prior Thiol Capture

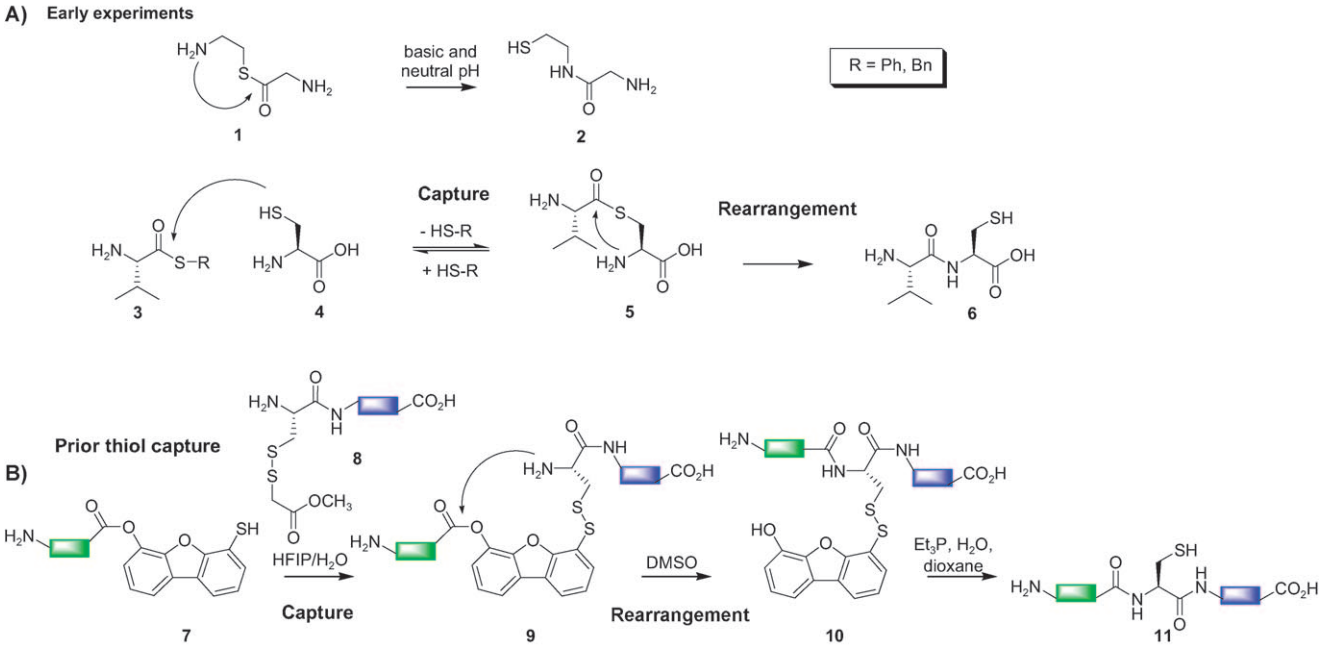
One of the earliest ligation techniques used with unprotected synthetic peptides was the prior thiol capture strategy developed by Kemp et al.[12,13] Initially, the peptide which will end up in the N-terminal region of the ligation product (the N peptide, see Table 1) is attached to an auxiliary template, such as 6-hydroxy-4-mercaptodibenzofuran, through an ester bond in **7** (Scheme 2B).[14] The N peptide can be assembled directly on the template by using optimized solid-phase peptide synthesis protocols. The capture is mediated by a disulfide exchange between a Cys residue located at the N terminus of the C peptide **8** and the free thiol of the template **7** (Scheme 2B). Afterwards an intramolecular O→N transfer takes place in **9** to give a native peptide bond between the N and C peptide in **10**. The template thus fulfils two main functions: on one hand it places the N and C peptide in proximity and on the other hand it serves as a mild activator of the ester-bound C-terminal carboxylic group of the N peptide, thereby increasing its electrophilicity. Finally, reducing agents such as triethylphosphine can be used to release the ligation product **11** from the template.[15]

The prior thiol capture strategy has been optimized[15] and applied in the ligation of various unprotected peptides to yield polypeptides with up to 39 residues.[16] However, another strategy called native chemical ligation (NCL) is routinely used nowadays for peptide ligations.

2.3. Native Chemical Ligation

2.3.1. Method Development and Chemoselectivity

The full potential of native chemical ligation was shown in 1994 by Kent and co-workers for the reaction of unprotected



Scheme 2. A) Aminolysis of thioesters by Wieland and B) Kemp's prior thiol capture strategy. DMSO = dimethylsulfoxide, HFIP = hexafluoroisopropanol.

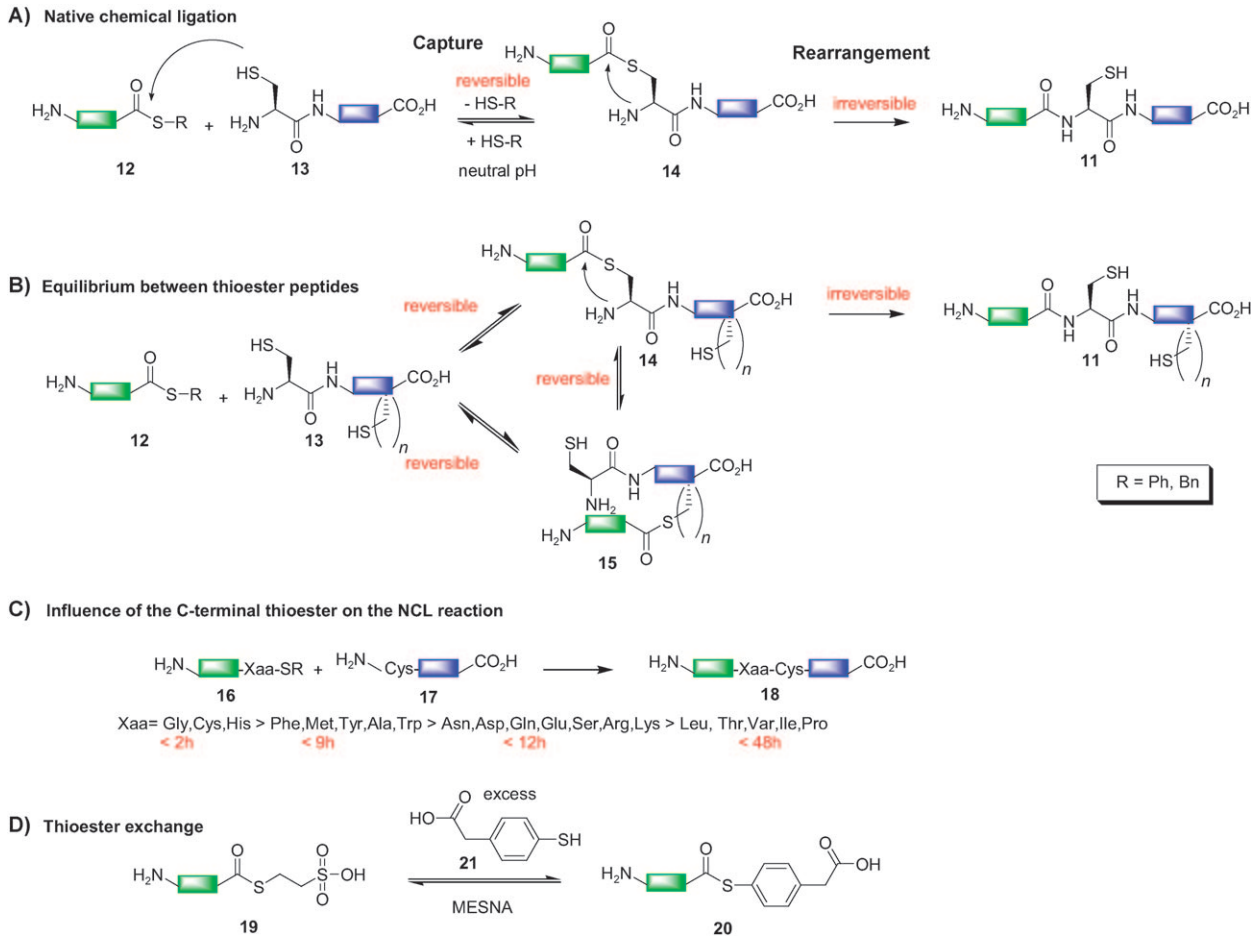
thioesters with N-terminal Cys peptides.[17] This study demonstrated that the reaction is chemoselective and thereby applicable to the ligation of unprotected peptides (Scheme 3 A). NCL is nowadays the most widely used chemoselective ligation technique based on a capture/rearrangement concept. The impact NCL made on chemistry and biochemistry is illustrated by the approximately 900 citations this original article has so far received. The chemoselective capture step is mediated by a reversible thiol–thioester exchange between an electrophilic (aryl) thioester **12** at the C terminus of the N peptide and the nucleophilic thiol of a Cys residue located at the N terminus of the C peptide **13** (Scheme 3 A). In the following rearrangement, the Cys-thioester **14** undergoes a rapid intramolecular S→N transfer, via a favorable five-membered transition state, to form a native peptide bond between the C and the N peptide in polypeptide **11**.[18] Importantly, additional internal Cys residues in the peptide do not interfere with the overall reaction pathway, since the irreversible intramolecular S→N shift can only occur at the unique N-terminal Cys residue. Any internal Cys residues that participate in the formation of thioester **15** rapidly exchange backwards(Scheme 3 B).

In the original study of 1994, NCL was used to synthesize human interleukin 8 (IL-8) from two synthetic peptide fragments.[17] Today, there are several examples of NCL-based peptide ligations for various purposes, including the selective introduction of protein modifications or tags,[19] the total chemical synthesis of posttranslationally modified and unmodified proteins (see Sections 2.4 and 6),[20] and segmental isotopic labeling of proteins for NMR studies.[21,22] In addition, the conjugation of peptides with other macromolecules such as peptide nucleic acid (PNA)[23,24] and DNA[25] can be performed.

2.3.2. *Reaction Conditions for the NCL and Mechanistic Aspects*

A major advantage of NCL is the mild reaction conditions, which allow the ligation to be carried out in buffered aqueous solutions of neutral pH value. This is important for two reasons: On the one hand, strong basic conditions make other residues such as Lys amenable to reactions with thioesters, and furthermore thioesters are not stable under basic conditions. On the other hand, acidic conditions do not favor NCL either, because the reactivity of the Cys thiol and the N-terminal amine of the C peptide is reduced.[9g,18]

Another important factor during the NCL reaction is the nature of the amino acid located at the C terminus of the thioester. Dawson and co-workers have investigated this in



Scheme 3. Native chemical ligation: A) Mechanistic pathway. B) Equilibrium between thioester peptides. C) Influence of residues at the C terminus of the N peptide. D) Thioester exchange with arylthiols. MESNA = sodium 2-mercaptoethanesulfonate.

detail by synthesizing a set of model peptides containing each of the 20 proteinogenic amino acids at the C terminal thioester **16**.[26] These peptides were treated with the C peptide **17** containing an N-terminal Cys residue to form the product **18**, which was monitored over time (Scheme 3C). These experiments indicated that all 20 amino acid thioesters **16** could undergo the NCL reaction, but the side chains affected the ligation rates significantly. The fastest reaction rate was observed for the Gly thioester which reacted quantitatively in less than four hours. In contrast, thioesters containing β-branched amino acids or proline did not result in a quantitative conversion, even after two days.[26] However, sterically hindered thioesters can also be ligated by using more reactive thiols.

Finally, the nature of the thioester has a huge impact on the NCL reaction.[27] Since alkyl thioesters are generally less reactive than aryl thioesters the former are more attractive for synthesis and handling. However, NCL requires a rapid thiol–thioester exchange, which proceeds much more efficiently with aryl thioesters (Scheme 3D). Thus, peptide thioesters are commonly synthesized as their alkyl derivatives **19**, for example, with sodium-2-mercaptoethanesulfonate (MESNA), and converted into the corresponding aryl thioesters **20** in situ by the addition of an excess of aryl thiols, such as the water-soluble (4-carboxymethyl)thiophenol (**21**, MPAA).[28] The thiol–thioester exchange with competing thiols has been investigated in detail by forming thioesters of various reactivity, where the equilibrium depended on the properties of the participating thiol species. The addition of aryl thiols to NCL reactions has an additional positive aspect: the maintenance of the reduced state of the N-terminal Cys of the C peptide, which is critical for the reaction.[29]

In line with the mechanistic aspects discussed above, NCL reactions usually proceed without side reactions and in high yields. Furthermore, NCL reactions can be performed in the presence of denaturing agents or detergents such as urea or sodium dodecylsulfonate (SDS), which can be important when the N or C peptides are not soluble under standard conditions. This advantage was exploited for the synthesis of the potassium channel KcsA.[30] The hydrophobic nature of the N and C peptides necessitated that the reaction had to be carried out in the presence of 1% SDS, which led to a conversion of 80% after only two hours.

Native chemical ligation reactions usually proceed at peptide concentration in the low millimolar range. In this regard, the use of templates which increase the local peptide concentration by binding or association events have proven to be particularly useful: For example, an autocatalytic peptide replicase based on the NCL was been described, in which a 32-residue α-helical peptide based on the leucine zipper domain of the yeast transcription factor GCN4 accelerated its own formation by NCL through association of the two peptide precursors.[31] This concept, also referred to as template-directed ligation,[32] has been extended to NCL with high sequence and diastereoselectivity, with an rate-acceleration of 4100 observed,[33] and the amplification of homochiral ligation products starting from racemic mixtures of the ligated fragments.[34–36]

2.3.3. *Chemical Requirements and Synthesis of the Starting Materials*

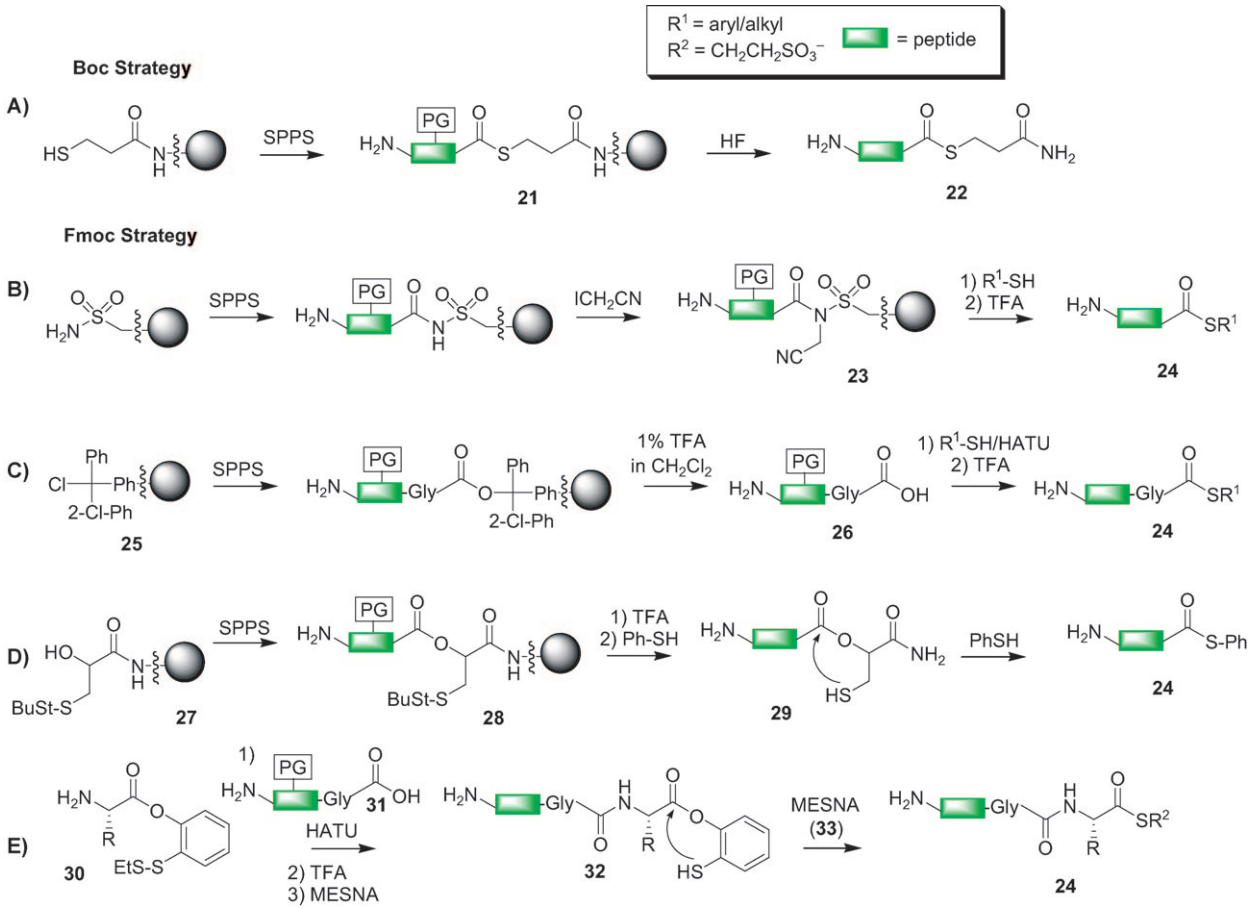
Reasons why NCL has become such a successful tool for peptide ligations are the high stability and accessibility of the starting materials. The readily accessible template, which can be incorporated by standard SPPS methods and facilitates the capture and rearrangement step, is none other than the natural amino acid cysteine which will remain in the ligation product and does not need to be further processed. The second required building block is a C-terminal thioester of the N peptide. Chemical methods for generating peptide thioesters are well established by using Boc as well as Fmoc-SPPS strategies (Boc = *tert*-butoxycarbonyl, Fmoc = 9-fluorenylmethoxycarbonyl).[37] The Boc strategy, however, requires HF to cleave the resin-bound thioester **21** and deliver thioester **22** (Scheme 4A).[38,39] Although special equipment is required for the cleavage reaction (because of the toxic and corrosive nature of HF), this route has been frequently employed in thioester syntheses because of the high overall yields.[37]

Various strategies which take into account the high electrophilicity of thioesters have been developed for the Fmoc-based synthesis of peptide thioesters.[40,41] A prominent example is the use of the alkanesulfonamide "safety-catch" resin,[42] which allows the resulting peptides to be activated with diazomethane,[43] iodoacetonitrile,[44] or by palladium-catalyzed allylation[45] to yield the secondary sulfonamide **23**. The addition of nucleophilic thiols leads to a cleavage of the protected peptides as thioesters (Scheme 4B).[43,46] Following a global deprotection, peptide thioesters **24** can be used for NCL reactions. Alternatively, peptides on HMBA (4-hydroxymethylbenzoic acid) or PAM (4-hydroxymethylphenylacetamidomethyl polystyrene) resins can be converted into thioesters with alkylaluminum thiolates.[47]

Another common approach for the synthesis of thioesters employs a highly acid-sensitive TGT[48] or 2-chlorotrityl resin **25**,[49] which allow cleavage under mild acidic conditions to yield protected peptides **26**. The thioester is introduced by C-terminal activation of the protected peptide in solution after cleavage from the resin and subsequent coupling with a thiol. Global deprotection finally affords an unprotected peptide thioester **24**.[46,50] However, the high risk of racemizing the C-terminal residue by the activation limits this otherwise very useful approach to C-terminal Gly residues or requires careful optimization of the reaction conditions to suppress epimerization of the peptide (Scheme 4C).[51]

An Fmoc-based strategy for the in situ synthesis of thioester **24** by an ester–thioester rearrangement employs the assembly of a carboxy ethyl ester **29**, which bears a disulfide-protected thiol in the β position (Scheme 4D).[52–54] Peptide assembly is performed on the resin-bound S-protected β-mercapto-α-hydroxypropionic acid **27**. After cleavage to yield **28**, thiophenol is added, which removes the disulfide protection and triggers an intramolecular O→S shift to afford thioester **24** (Scheme 4D).[52]

An additional thioester synthesis protocol also utilizing a disulfide-protected thiol in the β position employs phenylester **32**. This reaction is performed in solution and starts with the corresponding amino acid ester **30**, which is coupled with



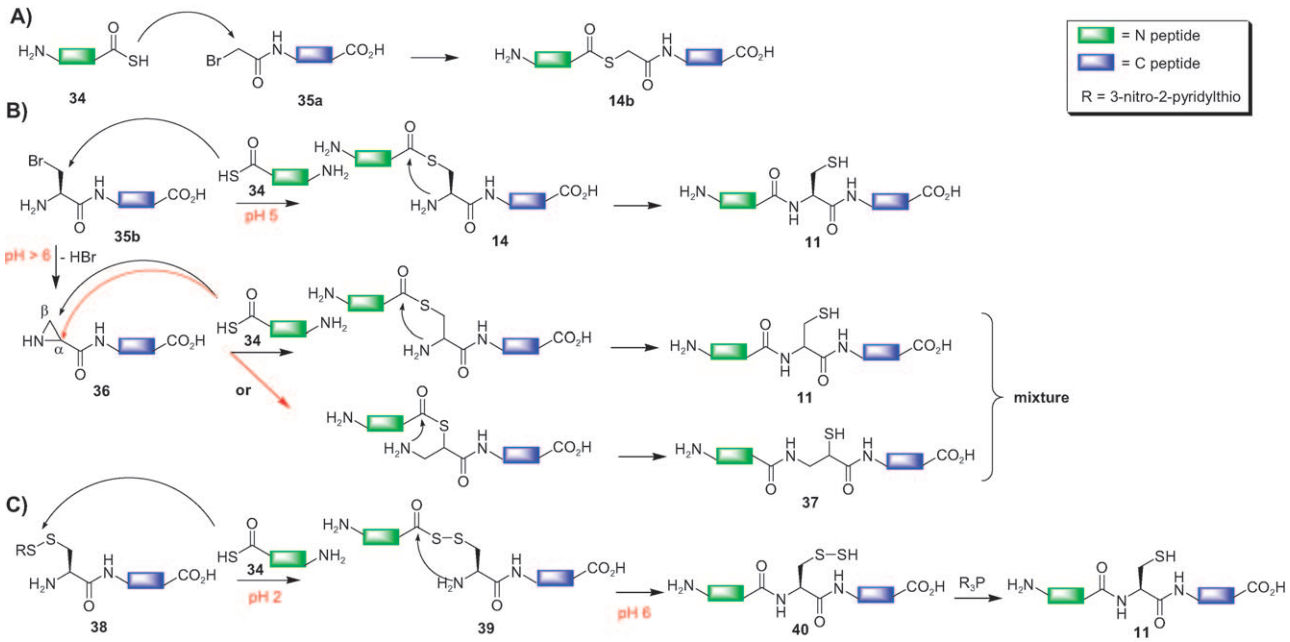
Scheme 4. Strategies for generating peptide thioesters. TFA = trifluoroacetic acid, HATU = 2-(1*H*-7-azabenotriazol-1-yl).

a protected peptide **31**. Following global deprotection, an excess of the thiol is added, which induces the O→S shift and thiol–thioester exchange (Scheme 4E).[53] In both cases the β-disulfide esters can be considered as "thioester precursors", which can be applied in reiterative reactions for the ligation of multiple peptide fragments (for additional strategies see Section 7).[55,56] These compounds were also applied in the synthesis of cyclic peptides,[57] glycopeptides,[58] and peptide ligations at cysteine-free junctions (see Section 3.3.4).

2.3.4. *Alternative Approaches for Peptide Ligations at Cysteine Junctions*

As described in Section 2.3.1, NCL utilizes the rearrangement of a thioester-linked peptide **14** (Scheme 3A), in which an amine nucleophile is positioned in proximity to the thioester moiety. Besides the thiol–thioester exchange reaction, one can envision a second synthetic access to thioesters **14b**, in which an inverted pair of nucleophilic and electrophilic peptides is used, as demonstrated in the reaction of a thioacid N peptide **34** with a C peptide containing a brominated alkyl species at the N terminus (**35a**; Scheme 5A).[59] In this ligation reaction (although not demonstrated for Cys-containing peptides), acidic reaction conditions are not only possible but even favored.[60] Specifically, a C-terminal thiocarboxylic acid **34** performs a nuclophilic attack onto an electrophilic β-bromoalanine residue located at the N terminus of the C peptide **35b**, thereby resulting in the formation of Cys-thioester **14** (Scheme 5B) before the final S→N shift is initiated to furnish **11**. However, in reaction media of pH 6 or higher, the α-bromoalanine residue can undergo 1,3-elimination of HBr, which results in the formation of an aziridine ring in **36**. Subsequent ring opening by the thiocarboxylic acid leads finally to the formation of an undesired side product **37** containing a α-thio-β-alanine residue (Scheme 5B). At pH 5 or lower, the formation of the aziridine ring is largely suppressed due to the protonation of the amino group of the β-bromoalanine.

In an analogous process, capture of the thiol-perthioester derivative can be performed at even lower pH values (Scheme 5C). In this case, the thiocarboxylic acid **34** performs a nucleophilic attack on a disulfide **38** to form a Cys-perthioester **39**. After a S→N shift, the resulting disulfide **40** can be reduced to the thiol **11**.[61] This reaction can be performed at pH < 2, which can be attributed to the activated nature of perthioesters and the strong nucleophilicity of the thiocarboxylic acid.



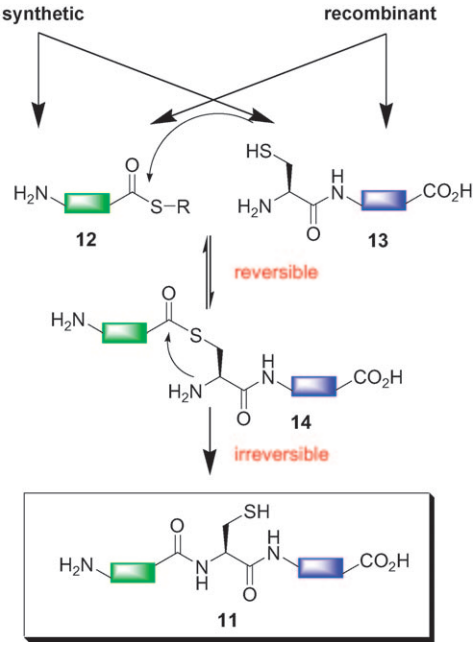
Scheme 5. Chemical ligation with inverted nucleophilic and electrophilic groups. A) Thioester peptide synthesis. B) Thioester ligation resulting in native peptide bonds. C) Cys-perthioester strategy.

2.4. Ligation Methods for Recombinant and Synthetic Peptides

Ligation products from (poly)peptide fragments of synthetic and recombinant sources are generally referred to as "semisynthetic" peptides or proteins, respectively. This strategy combines the advantages of organic synthesis and biochemical techniques, as it is possible to introduce various unnatural functionalities into a biopolymer. In this way, protein semisynthesis bypasses the size limitations of SPPS. The groundwork for modern protein semisynthesis was laid in the early 1970s by Borrás and Offord.[62,63] Their semisynthesis strategy starts with the reaction of the N-terminal amino group of peptides extracted from living organisms with Edman's phenylisothiocyanate (PITC). Next, the ε-amino groups of lysine residues are protected, followed by cleavage of the PITC-derivatized N-terminal amino group. The resulting unique α-amino group could be used for ligation reactions with activated amino acids. The applicability of this approach was demonstrated by incorporating tritium-labeled Phe as the N-terminal amino acid into human proinsulin.[64] However, the requirement to install protection groups on the Lys side chain prior to the ligation reaction limits this approach. Today, chemoselective ligation techniques which do not suffer from these shortcomings are primarily used to link recombinant and synthetic polypeptides.

2.4.1. NCL with Recombinant Peptide Fragments

NCL is especially useful for protein semisynthesis because the capture/rearrangement process is mediated by the natural amino acid cysteine.[9,17,65] Protein semisynthesis is often employed for the selective installation of biophysical probes or modified amino acids into recombinant proteins (Scheme 6).[66,67] To ligate a synthetic thioester **12** with a

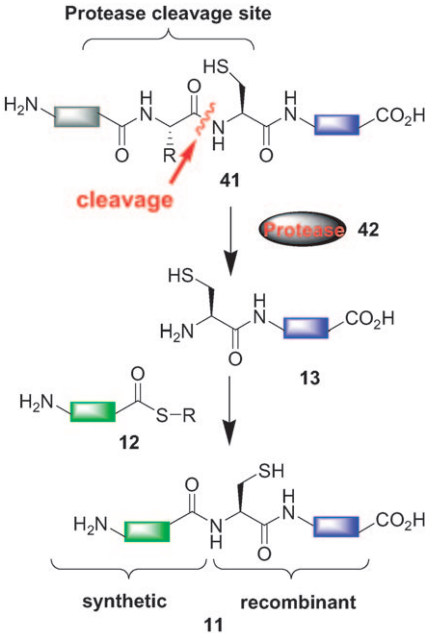


Scheme 6. The concept of protein semisynthesis.

recombinant protein **13** the latter needs to contain an N-terminal Cys residue. The introduction of an N-terminal Cys residue into proteins is commonly achieved by a combination of genetic methods such as site-directed mutagenesis and controlled cleavage with proteases.

In this approach, a Cys residue imbedded in the recognition site of a specific protease **42** (such as factor Xa[65,68] or the TEV protease)[69,70] is introduced at the N terminus of a protein **41** on a genetic level, commonly together with an additional affinity tag.[68–70] After expression and purification,

the recombinant protein is treated with the protease **42**, thereby resulting in the release of the N-terminal Cys protein **13** (Scheme 7). Alternative cleavage modes include the rearrangement of intein fusion proteins[71] and cyanogen bromide (CNBr) mediated cleavage between a Met-Cys motif in a protein.[72] Furthermore, genetic constructs containing a Cys codon following the ATG(Met) start codon have been used to express proteins with an N-terminal Cys residue. In these cases, the endogenous methionine aminopeptidase removes the initial Met residue and liberates the following Cys moiety.[73]

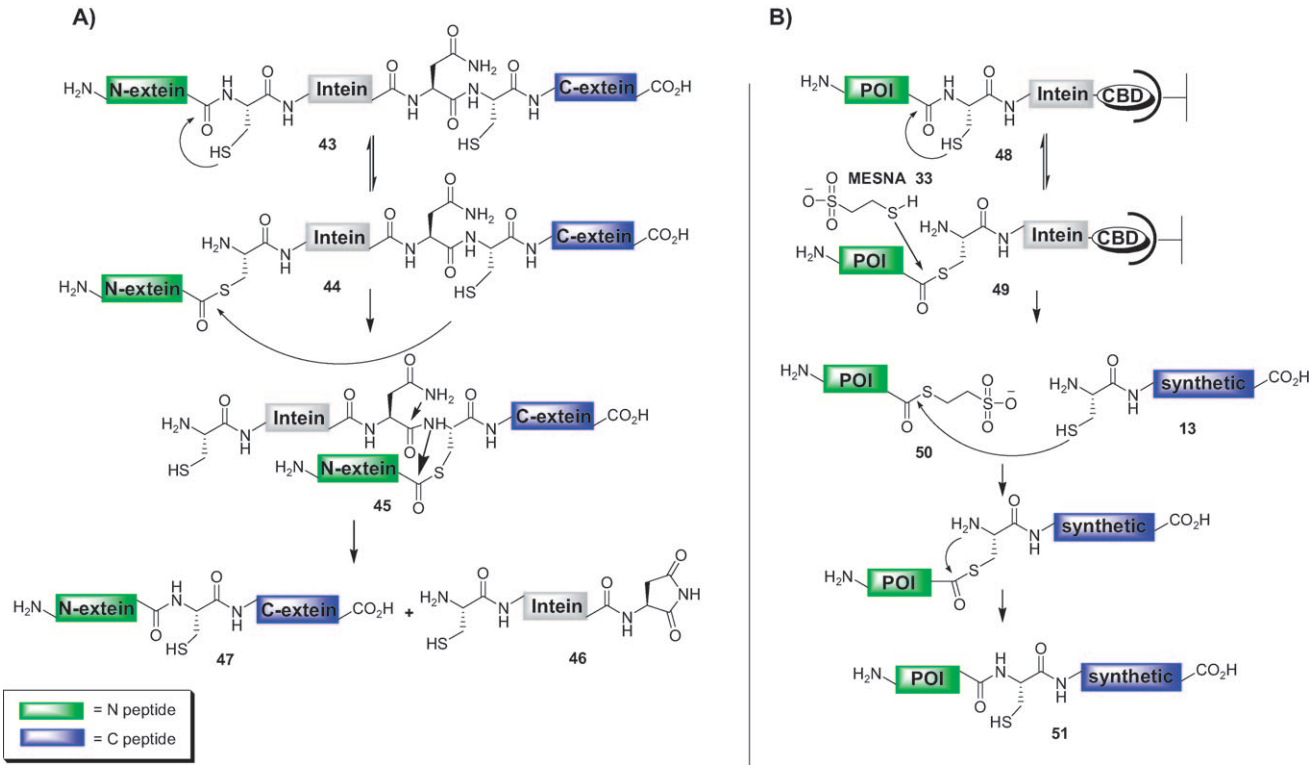


Scheme 7. Generation of recombinant peptides with an N-terminal Cys residue.

2.4.2. *Expressed Protein Ligation*

An obstacle for protein semisynthesis with NCL is the lack of chemical tools to convert recombinant proteins into protein thioesters. Thus, Muir, Sondhi, and Cole introduced a new tool for protein semisynthesis called "expressed protein ligation" (EPL).[74] EPL is based on a naturally occurring enzymatic process called "protein splicing", where an internal protein domain excises itself out of a precursor polypeptide **43** and links the flanking N and C fragments through a native peptide bond.[75–78] In analogy to RNA splicing, the internal protein domain is called an intein and the flanking regions N and C exteins. It is important to note that all necessary enzymatic activities for this process are imbedded in the intein domain itself. The underlying mechanism is illustrated in Scheme 8 A. First, a Cys residue at the N terminus of the intein undergoes an N→S shift to form a Cys-bound thioester **44** between the intein and the N extein.[79] Next, the thioester is transferred to a downstream Cys residue in **45**—in this case the first amino acid of the C extein. An intramolecular rearrangement with an Asn residue at the C terminus of the intein concludes the splicing process, thereby resulting in a succinimide **46** at the C terminus of the intein and a native peptide bond linking the N and C exteins in protein **47**.[80]

To exploit this process for the semisynthesis of a thioester, a protein of interest (POI) is fused with an intein at the genetic level (Scheme 8 B). In the above case, the POI



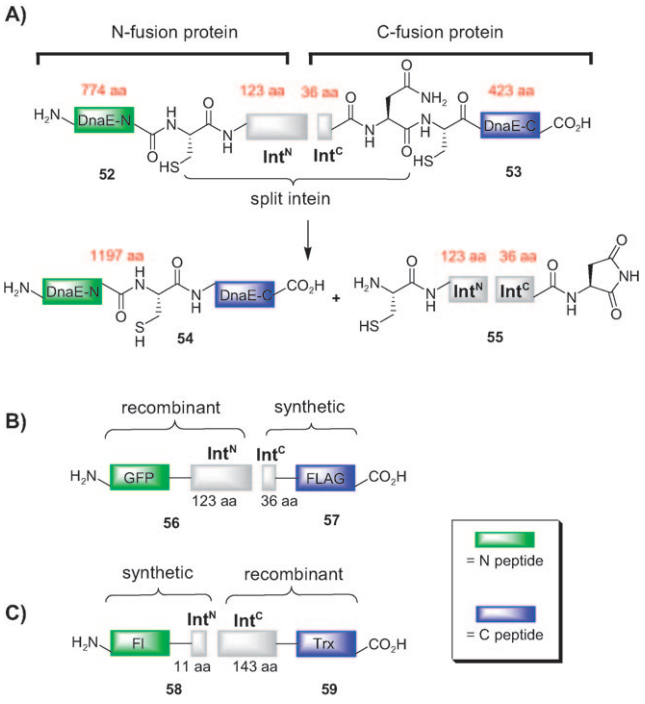
Scheme 8. A) Protein splicing and B) ligation of expressed proteins (EPL). POI = protein of interest; CBD = chitin binding domain.

represents the N extein in the expressed fusion protein **48**. The C extein is replaced by an affinity tag, such as a chitin binding domain (CBD), for purification purpose. The intein catalyzes the reversible N→S shift to **49**. Mutations in the intein prohibit further downstream processes of the splicing process discussed previously. Once the fusion protein is loaded on a resin, the desired protein thioester can be generated by the addition of excess thiols such as sodium 2-mercaptoethanesulfonate (MESNA, **33**), thus leading to the formation of the Cys thioester by a thiol–thioester exchange. The resulting protein thioester **50** can be isolated by simple elution from the resin, to which the intein remains bound. Finally, the thioester can undergo NCL with a synthetic C peptide **51** containing an N-terminal Cys residue. Alternatively, the synthetic peptide can be added directly during the thiol cleavage reaction to form the ligation product **51** in situ (Scheme 8B).

2.4.3. *Trans-Splicing of Proteins*

Another method applicable to ligating synthetic and recombinant peptides is "protein trans-splicing". This reaction is based on the previously discussed action of an intein and can therefore also be considered as a fully enzyme-catalyzed process. This methodology does not require the synthesis of thioesters for the ligation process, instead it employs inteins which are divided into two fragments. These so-called "split inteins" can then reassociate to form a fully functional complex which finally links the N and C exteins by protein splicing (Scheme 9A).

Split inteins were first discovered in the cyanobacterium *Synechocystis sp.* strain PCC6803, in which the catalytic subunit of the DNA polymerase III was expressed as two separate DnaE fragments (Scheme 9A). These fragments represented the N and C exteins of a split intein system, which later rejoins them to yield the full-length DnaE (in this case the polymerase III).[81] Specifically, the 774 amino acids at the N terminus of DnaE are fused with 123 amino acids of the N fraction of the split intein (Int^N) to yield the N-fusion protein **52**. The remaining 36 amino acids of the C-terminal region of the split intein (Int^C) are fused to the missing 423 C-terminal amino acids of DnaE in the C-fusion protein **53**. Upon splicing, DnaE **54** which comprises 1197 amino acids is reconstituted as one covalent polypeptide chain. The non-covalent complex of the Int^N and Int^C fragments **55** is the second component resulting from the splicing event (Scheme 9A).



Scheme 9. Protein trans-splicing. A) Trans-splicing of the split DnaE intein. B) Trans-splicing of a C-terminal semisynthetic protein. C) Trans-splicing of a N-terminal semisynthetic protein.

To utilize split inteins for protein semisynthesis a synthetic Int^C-C-extein peptide with the smaller C-terminal intein fragment can be obtained from solid-phase peptide synthesis (SPPS) and then reacted with a recombinant N-extein-Int^N fusion protein. The feasibility of this approach was demonstrated by the ligation of a synthetic FLAG epitope to a recombinant green fluorescent protein (GFP; Scheme 9B).[82] This method introduced by Giriat and Muir can also be applied inside living cells, and usually proceeds in the mid micromolar range. In this regard mammalian cells expressing the GFP-Int^N construct **56** were incubated with a synthetic Int^C-FLAG peptide **57**. The peptide contained additionally a cell-penetrating peptide, which allowed its shuttling into the living cells.

The natural occurrence of split inteins has led to further genetic and biochemical investigations as to whether regular inteins can be split artificially.[83–86] In this context, a further approach to semisynthetic protein trans-splicing was reported by Mootz and co-workers.[87] The artificially split Ssp DnaB intein was used to ligate a fluorescein (Fl) conjugated peptide **58** to recombinant thioredoxin (Trx) **59**. Importantly, the split site in this intein system is located only 11 amino acids downstream of the N terminus of the intein (Scheme 9C). Thus, this split intein system can be used to ligate synthetic peptides to the N terminus of recombinant proteins, whereas the DnaE system is only suitable for ligations at the C terminus. Furthermore, the size of the synthetic Int^N fragment **58** is only 11 amino acids (compared to 36 in case of the DnaE system), which is very advantageous for the SPPS of the whole synthetic peptide fragment: Since these residues will not end up in the ligation product, their size should be as small as possible to make the Ssp DnaB system an improvement for the semisynthesis of proteins.

3. Peptide and Protein Ligations by Capture/Rearrangement at Cysteine-Free Junctions

As described in Section 2, NCL allows a chemoselective ligation of two unprotected peptide or protein fragments to yield a native amide bond. Although the literature is rich in examples which demonstrate the general utility of NCL, the need for a Cys residue at the ligation site still represents a particular limitation for NCL. Cysteine residues are incorpo-

rated in peptide sequences only to an extent of 1.7 %,[88] which often excludes a retrosynthetic disconnection for peptide ligation by NCL. Additionally, cysteine residues cannot always be regarded as small "non-native" building blocks that replace the naturally occurring residue at the ligation site, as they can undergo disulfide-bond formation which could influence the overall folding pathway. To overcome this limitation, several approaches have been pursued to allow a chemoselective connection of amide bonds at cysteine-free peptide junctions **60**. These methods can be categorized according to their different methodological approaches (Scheme 10). It is important to note that some approaches yield native polypeptides with only naturally occurring amino acids, whereas others furnish amino acid analogues under maintenance of the α-amino acid backbone.

3.1. Ligations with Amino Acids other than Cys

The first approach utilizes amino acids other than Cys at the N terminus of the C-peptide fragment which can participate in the previously discussed capture step (see Section 2) and subsequent nucleophilic attack on an appropriate electrophilic functionality (in most cases a thioester or related moiety).

3.1.1. His Ligation

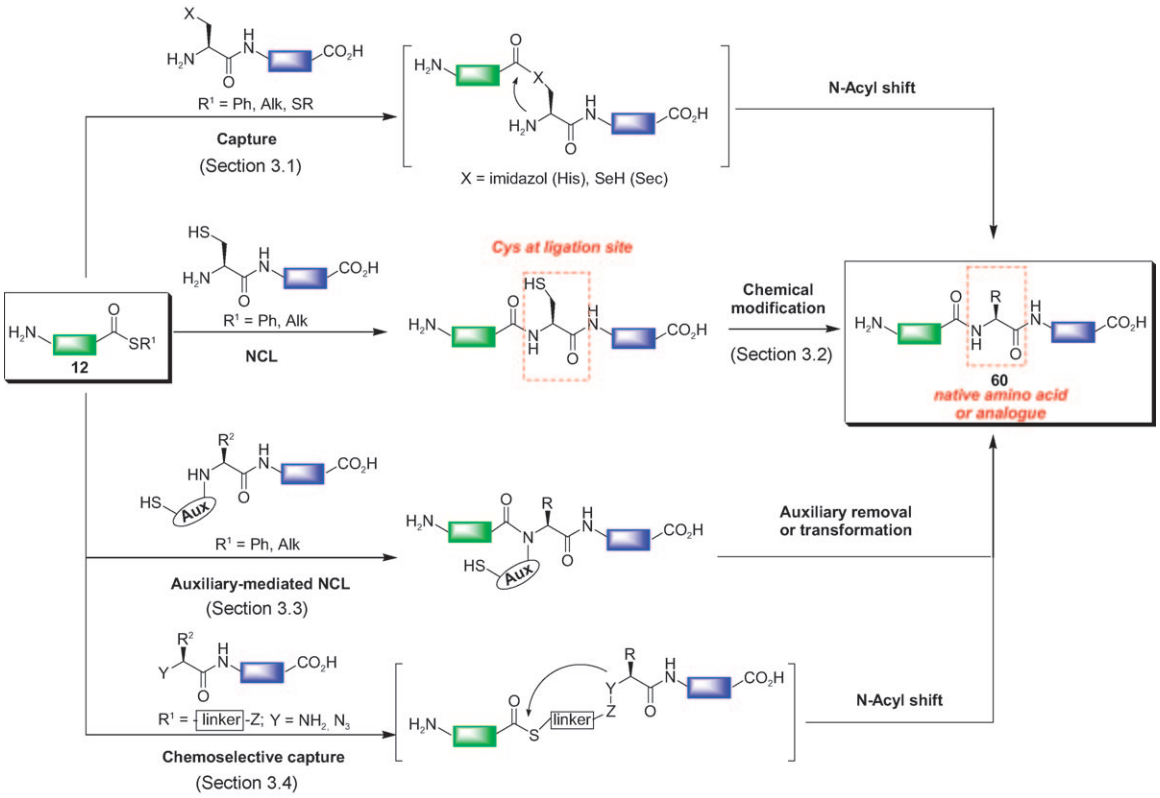
The first example of this approach is the use of His residues at the C terminus of the N peptide **61**. Histidine contains an imidazole moiety as the nucleophile,[89] which is known to accelerate enzymatic acyl transfer reactions under acidic conditions (Scheme 11 A). Although this transformation occurs in rather low yield in ligations with C-terminal peptide thioesters **12**, it becomes applicable under acidic conditions with more active perthioesters **62** (or acyl disulfides), which can be generated by activation of thioacids with Ellman's reagent.[90] This ligation has been demonstrated by Zhang and Tam in the synthesis of His-containing peptides **63** in yields between 60 and 70 %; however, the peptides cannot contain additional nucleophilic side chains.[91]

3.1.2. Sec Ligation

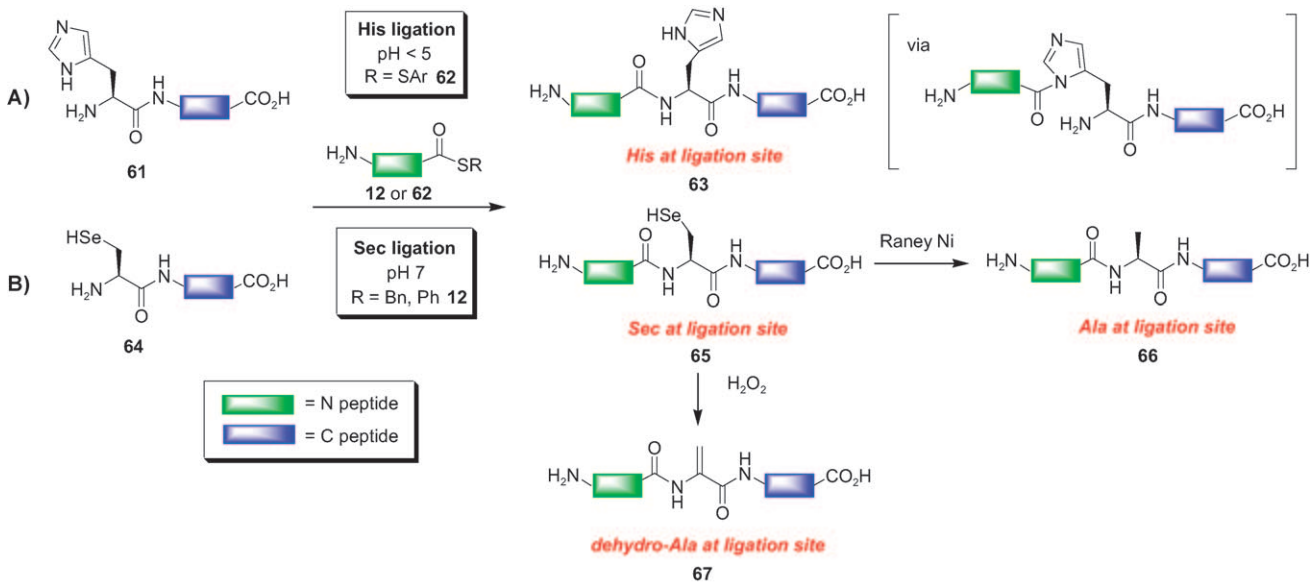
Another strategy involves the use of a selenol nucleophile, such as in **64**, instead of a thiol function for NCL, and leads to a peptide ligation product **65** bearing the rare amino acid selenocysteine (Sec, Scheme 11 B).[92] As the SeH group is more acidic and nucleophilic than a thiol, Sec-mediated ligations by NCL approaches are very fast and occur in high yield. They have been used in the synthesis of various peptide[93] and also protein architectures by EPL.[94] The resulting ligation products have had important applications in NMR spectroscopy[95] and mechanistically relevant investigations of proteins, as for example, in copper-binding studies of the blue copper protein azurin.[96]

3.2. Chemical Modification of NCL Products

The general synthetic strategy described in this section relies on selective chemical transformations of ligation



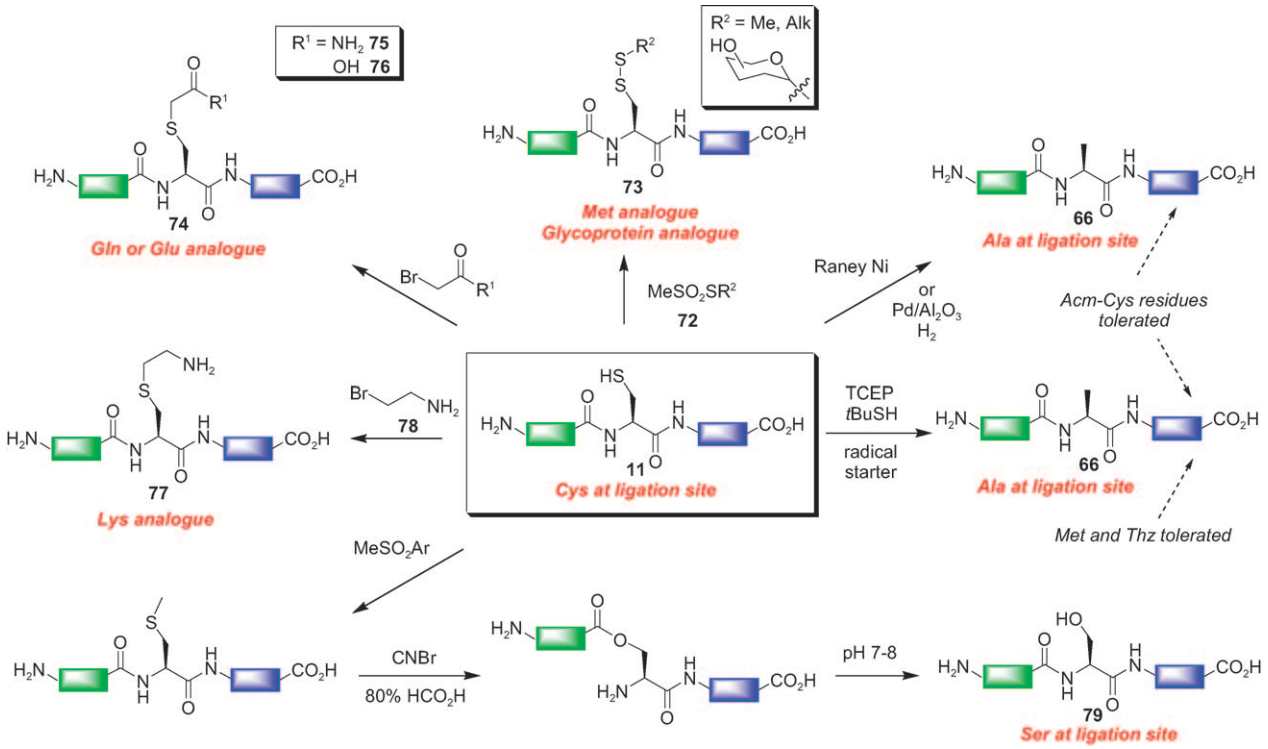
Scheme 10. Synthetic pathways for Cys-free peptides and proteins.



Scheme 11. Ligations at His (A) and Sec residues (B).

products **11** obtained by NCL (Scheme 12).[97] Such transformations are commonly utilized to convert expressed proteins into enzymes with improved enzymatic function[98] or to introduce natural protein modifications site-specifically.[99] The advantage of a chemical modification strategy in combination with NCL arises from the fact that these transformations can address the unique chemical properties of the pivotal amino acid at the native chemical ligation site, for example, a Cys or a Sec residue. Nevertheless, additional thiol functionalities in the ligation product limit the generality of this strategy, as selective transformations are difficult to perform without further protecting-group manipulations.

An additional issue of this approach is based on the fact that chemical transformations have to be performed on the final peptide product. Along with chemoselectivity considerations, harsh reaction conditions could harm the final protein product and thus significantly decrease the overall yield of the ligation or, even worse, affect the overall protein architecture through aggregation phenomena.



Scheme 12. Chemical modification of NCL products: Alkylation and desulfurization strategies. TCEP = Tris-(2-carboxyethyl)phosphine, CNBr = cyanogen bromide.

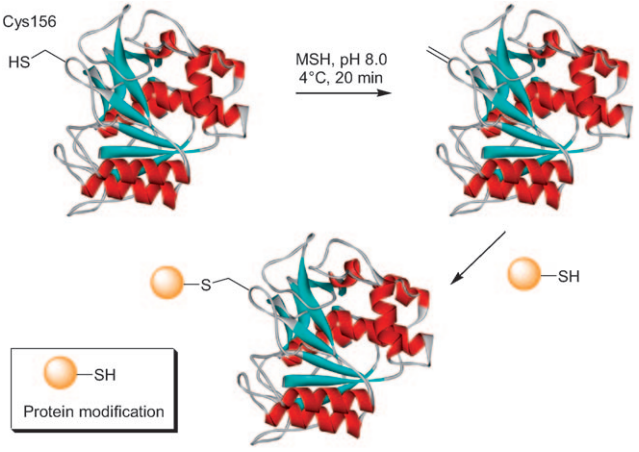
3.2.1. Desulfurization

3.2.1.1. Ala Ligation: Desulfurization of Cys

One of the most frequently used and most straightforward methods to convert thiol-containing ligation products into native polypeptide analogues is to selectively attack the thiol functionality by desulfurization techniques such as by the use of Raney-nickel or nickel boride (Scheme 12). Most importantly, this strategy allows the generation of Ala residues at the ligation site **66**. The method was introduced by Perlstein et al.[100] and has been applied in challenging peptide systems by Yan and Dawson,[101] who also compared palladium- and nickel-mediated processes in regard to the formation of reductive side products. This protocol was applied in the synthesis of various polypeptides[102,103] and further extended by a selective desulfurization of unprotected Cys over Acm-protected Cys residues.[104] This approach allows ultimately the ligation of Cys-containing polypeptides at Ala junctions after a final Acm-deprotection step. However, one of the major obstacles is still the heterogeneous nature of the employed Ni catalyst, which, in certain systems, can result in aggregation,[105] and further side reactions, such as desulfurization of thioethers in Met in thiazolidines (Thz) and the epimerization of secondary alcohols.[106] This disadvantage has been addressed recently by Wan and Danishefsky who reported a metal-free radical-based desulfurization technique to convert Cys residues in the presence of residues such as Met, Acm-protected Cys, and Thz in high yields (Scheme 12).[107]

3.2.1.2. Conversion of Cys into Dehydro-Ala (Dha)

Davis et al. have developed an oxidative elimination strategy employing *O*-mesitylene-sulfonylhydroxylamine (MSH) to convert a surface-exposed Cys residue into a dehydro-Ala residue (Scheme 13).[108] This fast reaction proceeds at 4 °C with full conversion and tolerates the various functional groups in the protein backbone, because methionine residues, in which the thioether moiety could react to give the iminosulfonium salt, can be regenerated by treatment with dithiothreitol (DTT). The resulting dehydro-Ala derivatives allowed a straightforward chemoselective incorporation of various sulfide-containing molecules through a Michael addition (see also Section 4.3.1), thereby allowing conjugation with peptides and carbohydrate derivatives.[109] Additionally, analogues of naturally occurring posttranslationally modified proteins, such as phosphorylated proteins, could be obtained, in which the resulting protein was delivered as a mixture of diastereomers.



Scheme 13. Chemoselective oxidative elimination to Dha proteins according to Davis.

3.2.1.3. Conversion of Sec into Ala or Dehydro-Ala

As mentioned in Section 3.1.2, Sec-containing ligation products can also be further processed. The Raney-nickel-mediated deselenation leads to the isolation of Ala derivatives **66**, whereas an oxidative elimination with H_2O_2 delivers the dehydro-Ala **67** as the final product, which can be further modified as well (see also Section 3.2.1.2, Scheme 11 B).[109–111]

3.2.1.4. Phe Ligation and Val Ligation: Desulfurization of β-Mercapto Bulding Blocks

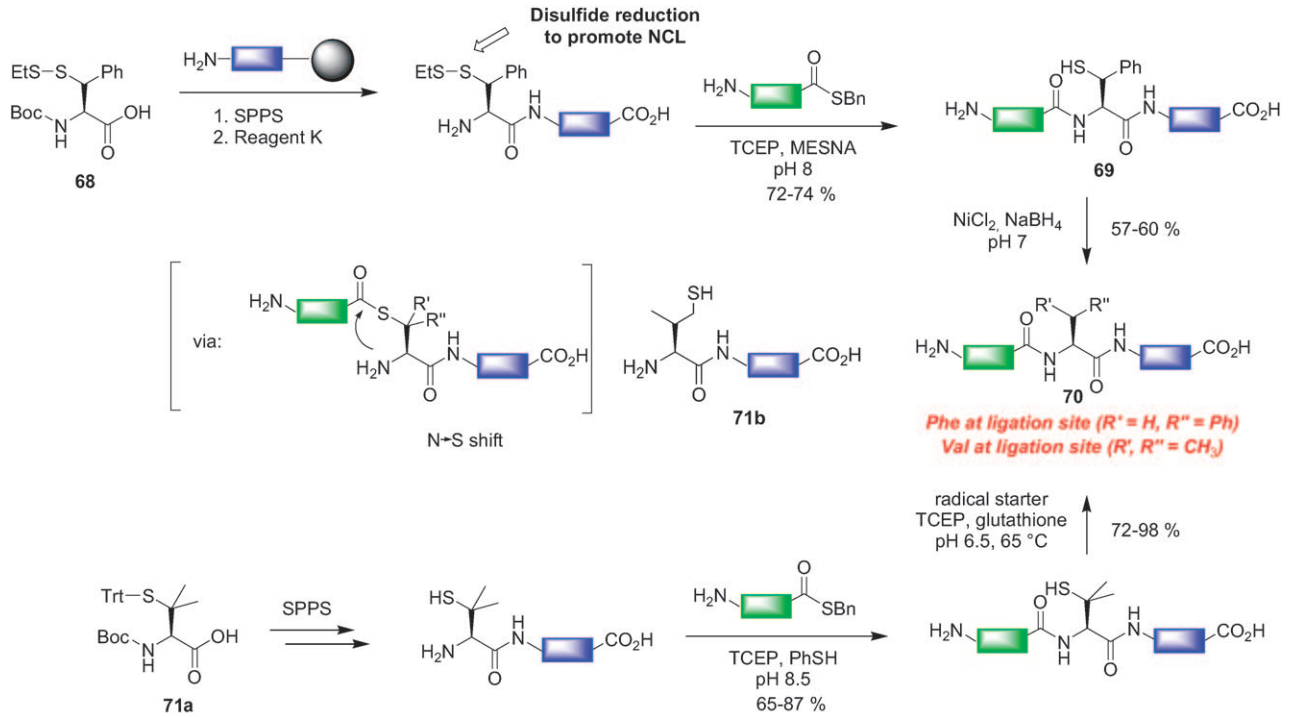
In two recent publications, NCL was extended to β-branched Cys derivatives, which were desulfurized to give natural amino acids after ligation.[112] For example, β-mercaptophenylalanine was used as a template for ligations at Phe junctions (Scheme 14, top).[113] This derivative can be attached to the N terminus of the C peptide as a disulfide building block **68**. After thiol addition, **69** is then generated in situ under NCL conditions. The thiol function at the benzylic position in **70** can be removed with Raney-nickel. This approach has been applied in the ligation of two pentapeptides to give a decapeptide with a Phe residue at the ligation junction in good overall yields.

In a similar approach, a penicillamine building block **71a** has been employed for the ligation at hydrophobic ligation sites, in which the β,β-dimethylcysteine functionality reacted surprisingly fast and in high yields with an N-peptide thioester (Scheme 14, bottom).[114a] The resulting penicillamine containing 11 to 22 amino acids was transformed into the corresponding Val peptides by applying an optimized version of Danisefsky's homogeneous desulfurization method. Most importantly, this two-step procedure has established linkage junctions between various amino acids including the branched hydrophobic residues Leu and Val. In a very recent publication by Danishrfsky et al., another system for the ligation at Val residues was described which employed a (primary) thiol group at the γ position of the N-terminal Val in the C peptide **71b**. This auxiliary delivered the corresponding peptides in an analogous two-step protocol in very good yields even at very hindered Thr-Val junctions.[114b]

3.2.2. Alkylation and Thioalkylation Protocols

3.2.2.1. Conversion of Cys or Sec into Ser as well as Met, Gln, Glu, and Lys Analogues

The nucleophilic properties of thiol and selenol functionalities at the ligation site can also be used in alkylation or thioalkylation reactions. Although the examples mentioned here have mostly been applied to the modification of



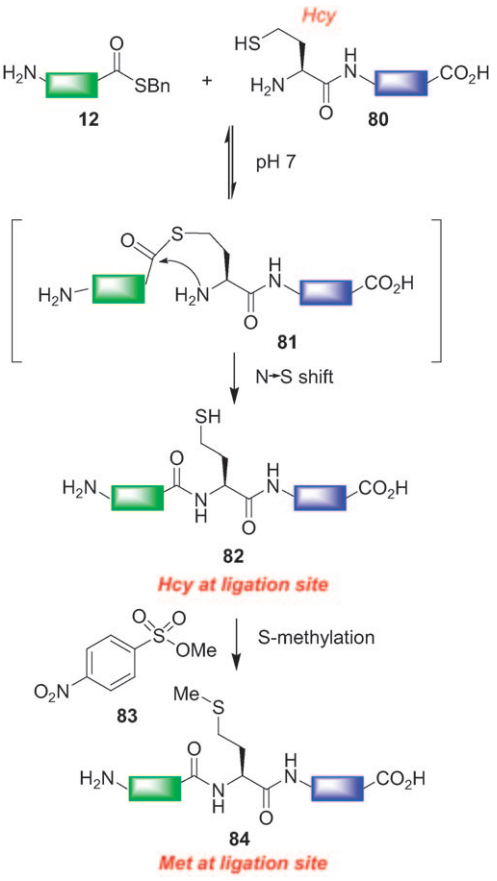
Scheme 14. Native chemical ligation at a Phe (top) or Val junction (bottom) by a two-step ligation and desulfurization protocol.

expressed proteins, they are also applicable to NCL products. For example, Cys residues in proteins can be transformed by thioalkylation with methyl methanethiosulfonate **72** to Met analogues **73**[115] or other important bioconjugates such as disulfide-linked glycoconjugates (Scheme 12),[116] which can be converted into a stable glycosyl thioether linkage by a chemoselective reaction with P^{III} reagents.[117] Additionally, Gln or Glu derivatives **74** can be obtained by reaction with α-iodoacetamide (**75**)[118] or α-iodoacetic acids (**76**, see also Section 7.2),[119] respectively, whereas Lys analogues **77** can be obtained by reaction with aziridines or bromoethylamines (**78**) (Scheme 12).[120] Although these transformations deliver non-native protein analogues of natural occurring amino acids, several studies have demonstrated they have similar biological function as their natural analogues.[119]

In a very recent study the methylation of a Cys residue with methyl-4-nitrobenzenesulfonate was combined with an activation and intramolecular rearrangement with CNBr in formic acid followed by an O→N-acyl shift under slightly basic conditions (pH 7–8).[121,122] This protocol converted Cys residues of NCL ligation products in Ser **79** (Scheme 12) in high yields, and hence has great potential in the synthesis of N-linked glycopeptides as they contain an Asn-Xaa-Ser or Asn-Xaa-Thr consensus sequence (Xaa represents any amino acid except proline).[123]

3.2.2.2. *Met Ligation: Alkylation of Hcy*

Another important ligation strategy has been developed which involves the alkylation of a homocysteine (Hcy) containing polypeptide with no Cys residues.[124,125] Hcy-



Scheme 15. Met ligation: Alkylation of a Hcy-containing NCL product.

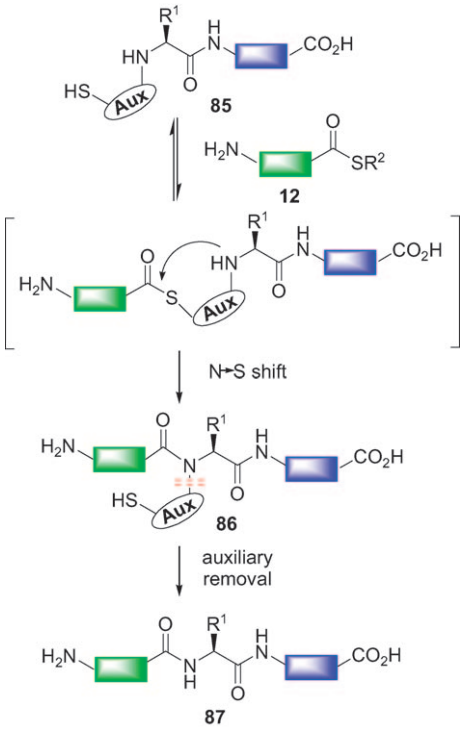
containing polypeptides **82** can be obtained, in analogy to NCL, with a homocysteine at the N terminus (**80**), where the reaction proceeds via a six-membered transition state **81** (Scheme 15). The final alkylation can be performed with methyl *p*-nitrobenzenesulfonate (**83**) under basic conditions,[126] or under neutral or acidic conditions to yield Met-containing peptides **84**.[127] Along those lines, Roelfes and Hilvert have reported the conversion of selenohomocysteine proteins obtained by EPL into selenomethionine analogues to create a selenium analogue of bovine pancreatic polypeptide (bPP).[94b]

3.3. Auxiliary Methods

All the previous methods affect unprotected Cys residues in addition to the Cys residue at the ligation site. In such cases, another approach, that is, the use of an auxiliary strategy, is more applicable.[128] In this method a thiol-containing auxiliary **85** is placed in proximity to the N terminus of a C-peptide fragment, which reacts with thioester **12** in a rearrangement similar to NCL (Scheme 16). The auxiliary has to be removed from the resulting peptide **86** in a final synthetic transformation to furnish the desired native peptide or protein sequence **87**. Thus, the same requirements concerning mild reaction conditions as discussed in the previous section are needed.

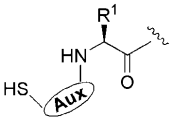
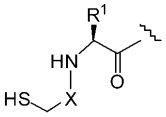
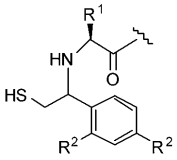
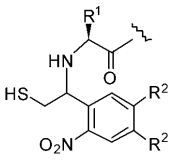
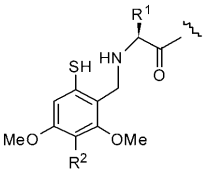
3.3.1. (Oxy-)Ethanethiol Auxiliary

One of the first auxiliary systems employed an ethanethiol (X = CH_2, **88**) or oxyethanethiol (X = OCH_2, **89**) substituent at the N terminus of a C peptide (Table 2, entry 1). The thiol



Scheme 16. Mechanism and scope of the auxiliary-mediated ligation.

Table 2: Auxiliary systems with possible ligation junctions and auxiliary removal conditions.[a]

	Ligation junction	Auxiliary removal conditions
		
 X = CH_2 **88**, X = OCH_2 **89**	Gly-Gly (for **88**) Gly-Gly Phe-Gly Ala-Gly (for **89**)Zn, acetic acid (for **86**)	none (for **85**)
 R^2 = H, OMe **90**	Gly-Gly Ala-Gly His-Gly Lys-Gly	HF (for R^2=H) TFA (for R^2=OMe)
 R^2 = H, OMe **91**	Gly-Gly Ala-Gly	*hν* (310–365 nm)
 R^2 = H **92a** (Dmb) OMe **92b** (Tmb)	Gly-Gly Ala-Gly Lys-Gly	HF (for **92a**) TFA (Hg^{2+}) (for **92b**)

[a] Dmb = 4,6-dimethoxy-2-mercaptobenzylamine, Tmb = 4,5,6-trimethoxy-2-mercaptobenzylamine.

group can participate in the thioester exchange reaction with a second peptide thioester before the secondary amine can undergo the S→N shift. In general, the ethanethiol auxiliary delivered better ligation yields at Gly junctions (up to 90%) compared to the oxyethanethiol analogue. Although the ethanethiol auxiliary cannot be removed from the peptide,[129] the N–O bond of the oxyethanethiol auxiliary can finally be cleaved under reducing conditions with Zn.[130] Both systems have been applied to the synthesis of advanced peptides including peptide cyclization.[129,131]

3.3.2. N^{α}-1-Aryl-2-mercaptoethyl Auxiliary

In the auxiliary system **90** the N terminus of the C peptide is converted into a secondary amine which contains an aryl functionality in the 1-position and a thiol functionality in the 2-position (Table 2, entry 2).[132] This functional group can be introduced into the peptide either by reaction of the parent amine with an α-bromo peptide or by a building block

approach with Boc- or Fmoc-protected N^{α}-(1-aryl-2-mercaptoethyl)amino acid derivatives.[133] The auxiliary system allows sterically nondemanding coupling at Gly and Ala junctions in good overall yield in various peptidic systems (Table 2), including glycopeptides and cyclic peptides, even though the auxiliary has to be removed under highly acidic conditions using HF or TFA/TMSBr.

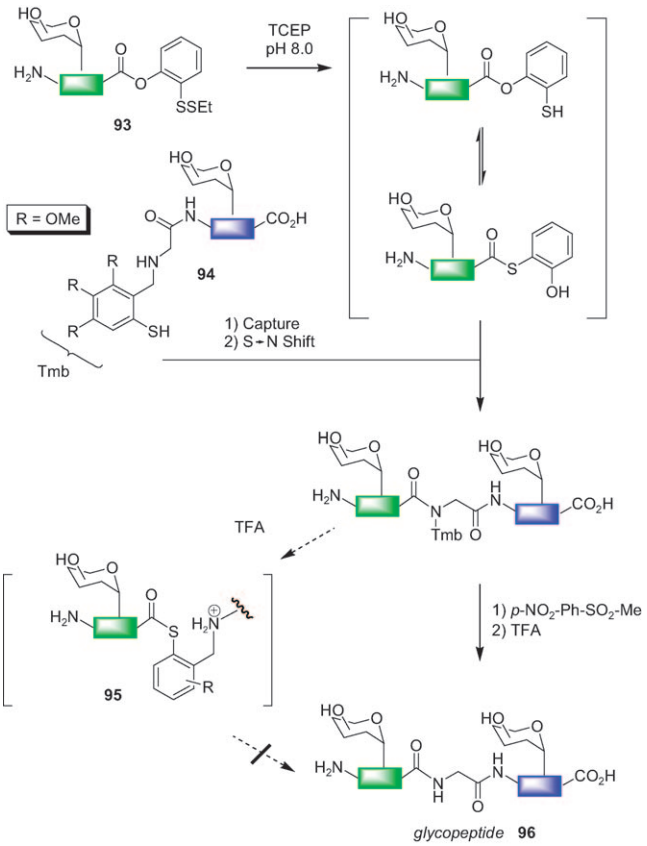
3.3.3. Photocleavable Auxiliaries

Particularly noteworthy is a method that addresses the issue of the removal of the auxiliary under mild conditions introduced by the research groups of Aimoto[134] and Dawson,[135] who developed a photocleavable variant of the previously mentioned 1-aryl-2-mercaptoethyl auxiliary system **90** (Table 2, entry 3). The light-sensitive auxiliary **91** bears an *ortho*-nitrobenzene function, which efficiently ligated two nonapeptides in high yields prior to its photolytic removal. Despite these benign deprotection conditions, the auxiliary system is limited to sterically nondemanding peptide ligations at Gly-Gly or Ala-Gly junctions. Nevertheless, this approach has recently found an important application in the semisynthesis of the protein ubiquitin conjugated to the ε-amino group of a Lys side chain within a histone-derived peptide, and represented the first example of an EPL at a cysteine-free ligation site.[136] In combination with a second EPL, this process was even extended to a chemical ubiquitin modification of proteins. Specifically, a chemically ubiquilated histone H2B was obtained, which stimulated biological responses in mediating intranucleosomal methlyation.[137]

3.3.4. 4,5,6-Trimethoxy-2-mercaptobenzylamine (Tmb) Auxiliary

In 2000, Offer and Dawson introduced a 2-mercaptobenzyl-based auxiliary system **92a** in which an electron-rich arylthiol undergoes thioester exchange with a C-terminal peptide thioester (Table 2, entry 4). A subsequent S→N shift of the secondary amine leads to the formation of a tertiary amide, which is prone towards acidic hydrolysis with TFA.[138] Optimization studies delivered the 4,5,6-trimethoxy-2-mercaptobenzylamine (Tmb) auxiliary **92b** in which the arylthiol had an increased nucleophilicity and the tertiary amide an increased acid lability. This auxiliary was successfully used for the ligation of peptides that contained at least one Gly at the ligation site, with coupling times of 0.5 h (Gly-Gly), 2 h (Lys-Gly), and 5 h (Ala-Gly).[139] This ligation strategy has been used in the synthesis of cytochrome b562, a 106-residue protein.[140]

Recently, Danishefsky et al. have employed the Tmb auxiliary in combination with the C-terminal peptide phenol ester **93** with an *ortho*-disulfide (see Section 2.3.3) in sterically demanding ligations of two glycopeptide fragements **93** and **94**, both containing an N-linked chitobiose residue at a Gly-Gln junction.[141] In this study it was found that the acidic conditions applied for removal of the auxiliary led to an accumulation of the thioester intermediate **95** through protonation of the secondary amine, and thus involving an N→S shift.[18,142] A prior methylation of the arylthiol increased the removal efficiency to furnish glycopeptide **96** (Scheme 17).

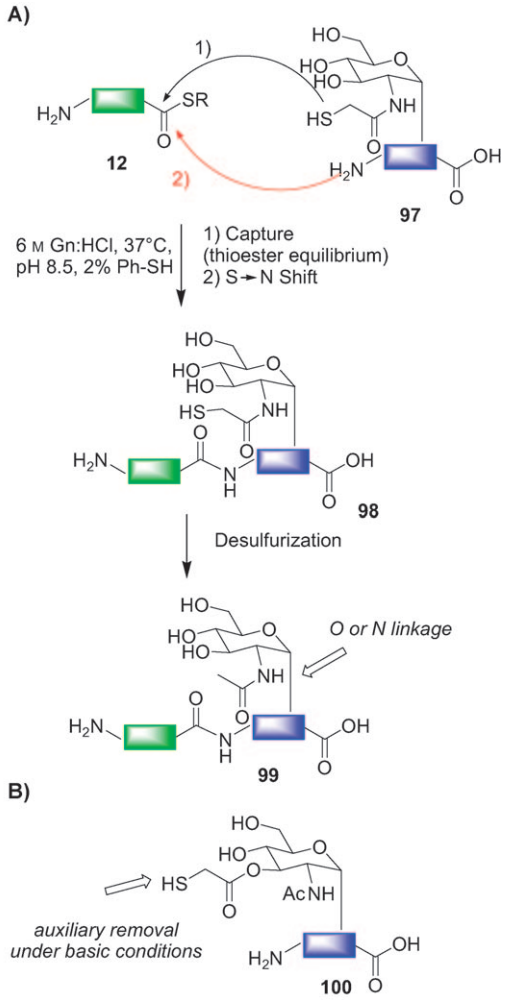


Scheme 17. Cys-free glycopeptide coupling by using a Tmb auxiliary.

3.3.5. Sugar-Assisted Glycopeptide Ligations (SAL)

Recently, Wong et al. introduced a ligation strategy for glycopeptides in which the auxiliary for the capture step is incorporated into the glycan; this method is termed sugar-assisted glycopeptide ligation (SAL, Scheme 18A).[143] Specifically, the thiol function is introduced as a β-thioacetamido substituent at the C2-position in **97** either in a GalNAc moiety (for O-linked glycopeptides)[144] or in a GlcNAc residue (for N-linked glycopeptides).[145] The thiol is proposed to form initially a thioester with a C-terminal peptide thioester **12**. A free amine at the N terminus of the glycopeptide fragment can then perform the S→N shift to generate the native amide bond in **98** under slightly basic conditions. Finally, desulfurization (see Section 3.2.1) yields the native glycan structure **99**. It was demonstrated that glycopeptides carrying up to six additional amino acid residues N-terminal to the glycosylation site were able to undergo SAL, thus allowing a flexible choice of the ligation site to permit bulky amino acid residues in proximity to the glycan.[146] A β-thioacetic ester in the C3-position of a GalNAc residue in auxiliary **100** can also be used; in this case a final basic auxiliary removal with hydrazine is possible (Scheme 18B).[147]

In a very recent study from the same research group, a direct aminolysis approach was described which allowed the coupling between a (glyco-)peptide thioester and another (glyco-)peptide at a cysteine-free junction or sugar auxiliary,



Scheme 18. Sugar-assisted ligation with a A) β-thioacetamide or B) β-thioacetic auxiliary.

respectively.[148] In this ligation strategy, which is based on the early observations by Wieland, a carefully optimized NMP/HEPES buffer at neutral to slightly basic pH is employed, which renders amino groups, including the N terminus of the C peptide, sufficiently nucleophilic to perform the aminolysis with the N-peptide thioester. These conditions simultaneously suppressed thioester hydrolysis efficiently. However, this method cannot be employed as a general chemoselective ligation technique, since additional amino groups in Lys side chains have to be protected.

3.4. Other Methods Relying on Chemoselective Capture Strategies

The methods discussed so far relied on NCL variants in which a thioester equilibrium (or disulfide formation) resembles the capture step before the amide bond is formed by an S→N shift. In contrast, the following methods utilize a different chemoselective reaction (for a detailed discussion on chemoselectivity see Section 4) for the first capture step.

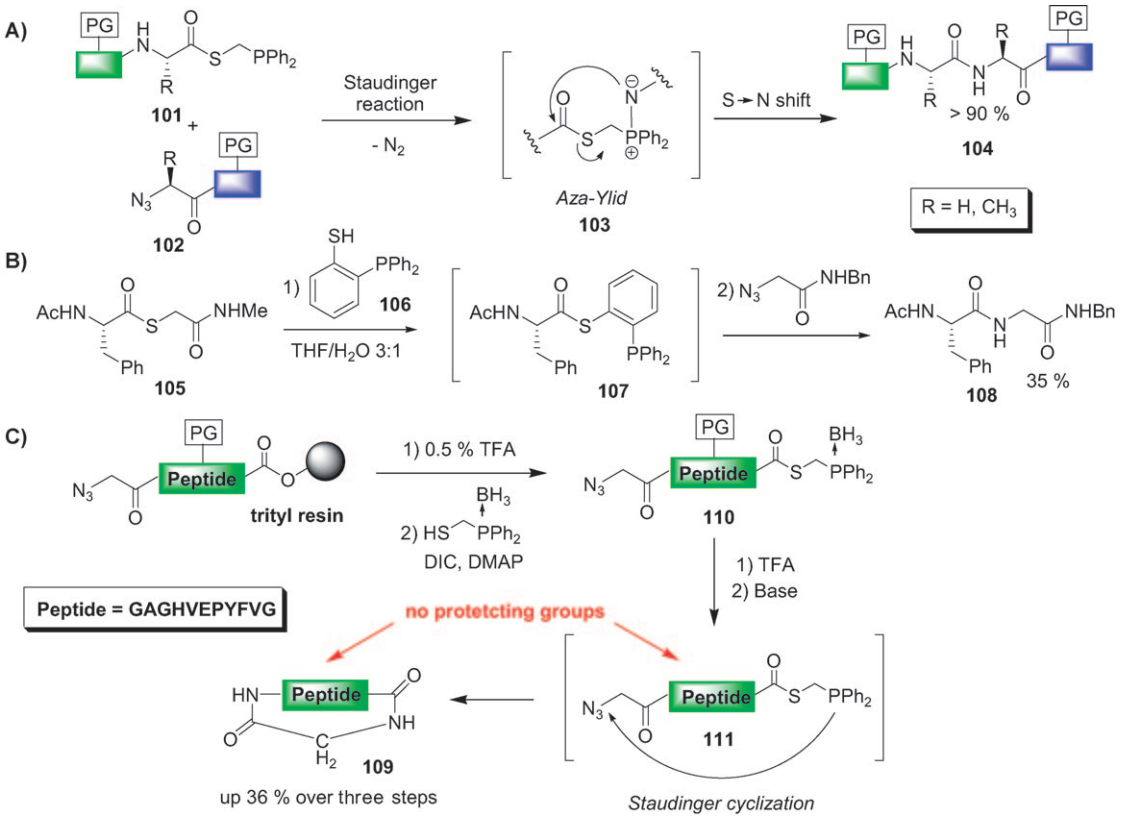
3.4.1. Traceless Staudinger Ligation

This ligation reaction was introduced by the research groups of Raines and Bertozzi in 2000 and represents a traceless variant of the initially reported (nontraceless) Staudinger ligation (see Section 4.2.4).[149,150] Both reactions originate from the well-known Staudinger reaction[151] in which an azide reacts with a phosphine to form an iminophosphorane. The non-traceless Staudinger ligation was identified by Bertozzi as chemoselective and was developed for the chemoselective functionalization of metabolically engineered azidoglycans in cellular environment (see Section 4.2.4).[150]

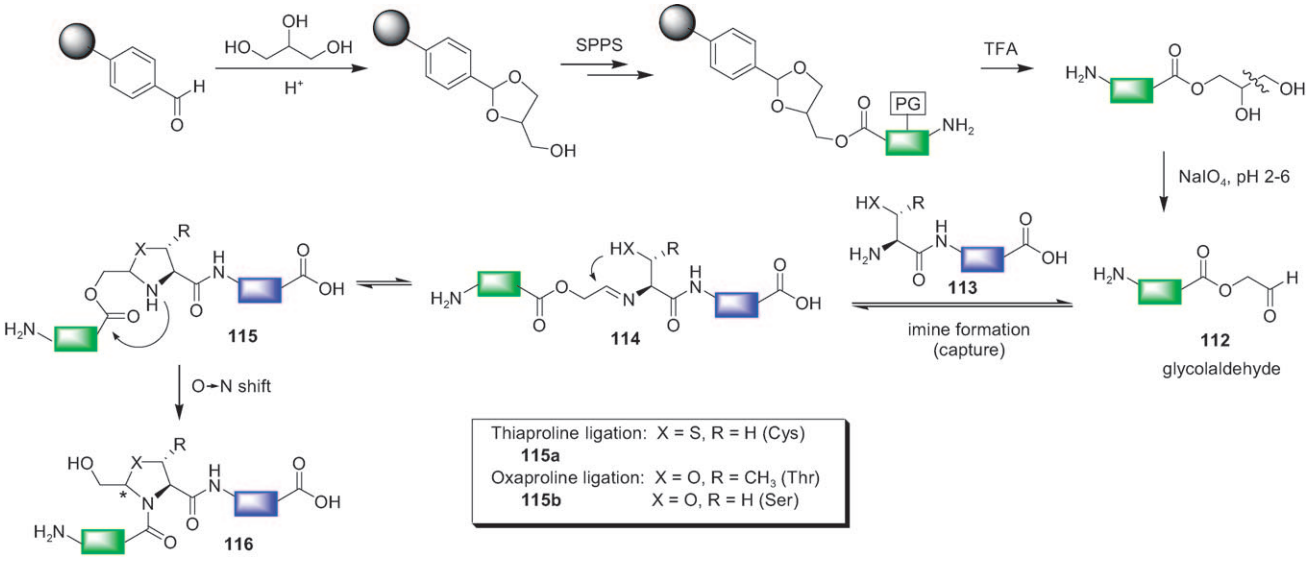
In the traceless Staudinger ligation, the iminophosphorane **103**, formed from an α-azidoamino acid or -peptide **101** and a phosphinothioester **102** (resembling the chemoselective capture step), can undergo an intramolecular attack with an internal thioester to yield a native amide bond in high yields at a Gly-Gly/Ala[149] or even an Ala-Ala junction (**104**) (Scheme 19A).[152] Similarly, an *ortho*-substituted diphenylphosphinoarylthiol **106** can participate in a thioester exchange with **105** to lead to the in situ formation of a phosphinothioester **107**, which can act as a substrate for the Staudinger reaction to give **108** (Scheme 19B).[149a] The traceless Staudinger ligation does not, in contrast to NCL, require a Cys residue at the ligation site. A water-soluble version of the traceless Staudinger ligation has been developed,[153] and detailed mechanistic investigations[154] and applications have been performed. Those studies included site-specific immobilization of peptides on surfaces,[155] ligation of protected (glyco-)peptide fragments,[156] and cyclization of unfunctionalized lactams.[157] Until recently, and in contrast to the nontraceless Staudinger ligation,[158] the chemoselectivity of the traceless Staudinger ligation has rarely been addressed.[156b,c,159] In a very recent publication, the traceless Staudinger ligation was employed in the cyclization of unprotected peptide sequences with 11 amino acids (**109**), which contained all amino acids from the circular protein microcin J25. Although the tolerance of all the naturally occurring amino acids in this protocol has not been proven, it was demonstrated that borane-protected peptide azidophosphinothioesters **110** can be deprotected under acidic conditions, which simultaneously removed the proecting groups at the peptide side chains as well as the borane protecting group to yield phosphonium salts **111**, which cyclized on addition of base (Scheme 19C).[160]

3.4.2. Imine Ligations with Subsequent Pseudoproline Formation

Another ligation process introduced by Liu and Tam furnishes a pseudoproline by reaction of a C-terminal glycolaldehyde peptide **112** with another peptide **113** containing a Cys, Thr, or Ser residue at the N terminus (Scheme 20).[161] This reaction is initiated by the formation of an imine **114** to which the sulfhydryl or hydroxy function of the N-terminal amino acid can add on to form the tautomeric thiazolidine (Thz, **115a**) or oxazolidine (Oxz, **115b**). This peptide derivative can finally undergo an irreversible O→N shift via a bicyclic five-membered intermediate to deliver a



Scheme 19. A) Mechanism of the traceless Staudinger ligation. B) In situ generation of phosphinothioesters for the traceless Staudinger ligation. C) Staudinger cyclization after acidic deprotection of borane-protected phosphinothiols. DIC = diisopropylcarbodiimide.



Scheme 20. Pseudoproline ligation of unprotected N peptides with C-terminal glycoaldehyde C peptides.

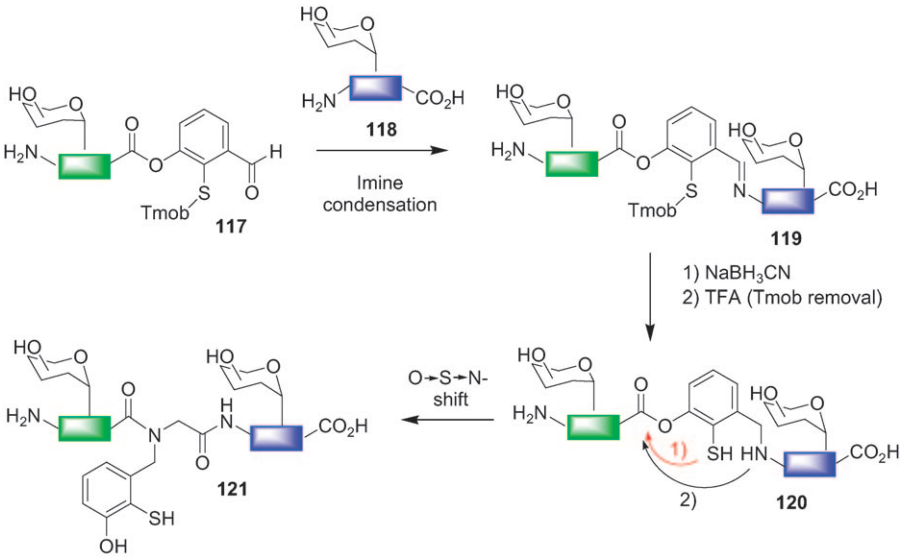
hydroxymethylene-substituted pseudoproline **116** at the ligation junction, with a new stereogenic center at the C2-position.[162] It is important to note that the formation of the thiazolidine can be carried out in aqueous solution, whereas the oxazolidine synthesis requires anhydrous conditions.

This imine ligation reaction was utilized in the assembly of proline-rich analogues of the 59-residue helical antibacterial peptide bactenecin 7.[163] Interestingly, Tam and Miao utilized a three-segment condensation, in which the first junction was constructed by the formation of a thiazolidine of the Cys peptide, a faster step than the formation of the corresponding oxazolidine of the Thr peptide. Simultaneously, an oxidative conversion of a C-terminal glycerol ester of a peptide dimer into the corresponding glycolaldehyde, and subsequent for-

mation of the second pseudoproline was reported as another segment ligation protocol. The synthesized analogues displayed comparable biological activities as the natural product, whereas structural studies revealed a higher population of the polyproline type II helix in the natural polypeptide.

3.4.3. Peptide Ligation by Reductive Amination

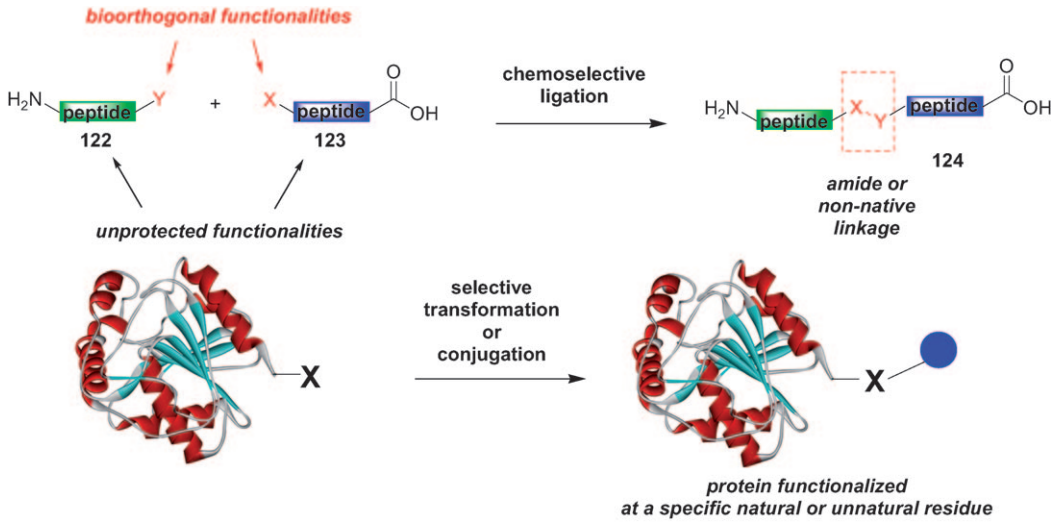
Another Cys-free ligation protocol based on Danishefky's phenolic ester strategy uses a Tmob-protected *ortho*-thiol and O→S→N transfer between various amino acids, one of which should be Gly or Ala (Scheme 21).[164] The strategy relies on the introduction of an alydehyde in the *meta* position to the phenolic ester **117**, which can undergo a reductive amination with the N terminus of an unprotected peptide **118** as the capture step to form **119**. Deprotection of the *ortho*-thiol leads to an O→S shift before the secondary amine can undergo the final S→N transfer (**120**). The auxiliary in **121** can be removed under acidic conditions (in a similar manner as described for the Dawson auxiliary). Besides the deprotection conditions potentially harming the glycosylic linkages in glycopeptides, another drawback of this method is the possible epimerization at the C-terminal amino acid during formation of the phenolic ester (which limits its utility to Gly and Pro residues). The first disadvantage was addressed by a recent protocol in which the phenolic esters were ligated using Ag^{I} salts with a peptide containing a free amine at the N terminus ("auxiliary-free Cys-free" coupling).[165] Although this strategy may also lead to the epimerization of peptide thioesters and, additionally, is not fully chemoselective, as it requires the protection of additional nucleophiles in the peptide side chain, highly complex glycopeptides were ligated in very good overall yields.



Scheme 21. Cys-free glycopeptide coupling by reductive amination. Tmob = 2,4,6-trimethoxybenzyl.

4. Ligation and Modification of Peptides and Proteins with Chemoselective Reactions

Chemoselective ligation strategies allows the selective formation of a covalent bond between highly complex biological molecules without the requirement of protecting-group transformations. Besides the two-step "capture/rearrangement" strategy (see Sections 2 and 3), several organic reactions have recently been identified that allow a "direct" formation of a chemoselective bond in **124** (Scheme 22). In an ideal scenario, these chemoselective reactions proceed rapidly under physiologically benign conditions (neutral pH in aqueous media and at room temperature) to yield a native linking moiety. This bond is constructed from two "bioorthogonal"[166] functional groups in **122** and **123** that do not show any cross-reactivity with other functional groups present in a biological environment.



Scheme 22. Chemoselective reaction for the native or non-native conjugation of polypeptides and selective transformation of natural amino acids.

The development of chemoselective reactions have enabled researchers not only to synthesize homogeneous (un-)natural peptides and proteins by ligation of two (poly-)peptide strands but also to site-selectively functionalize or modify complex biomolecules.[167] These applications have advanced the localization and study of the function of post- and cotranslationally modified proteins even within a cellular environment or a living organism, as for example by site-selective incorporation of biophysical probes, such as fluorescent labels, into a biopolymer (bioorthogonal chemical reporter strategy).[167] In this strategy a unique chemical functionality is introduced by the cellular biosynthetic machinery and afterwards chemoselectively functionalized with an exogenously added biophysical probe, thereby allowing the biomolecules to be tracked in their native environment.[167b]

Within this section we will give an overview of examples for chemoselective reactions in protein synthesis as well as preliminary approaches for the ligation and/or functionalization of polypeptide architectures. These approaches are divided into three different categories: The chemoselective formation of native (Section 4.1) and non-native bonds, which have often been realized in complex protein environments (Section 4.2), as well as the selective transformation of naturally occurring amino acids (Section 4.3).

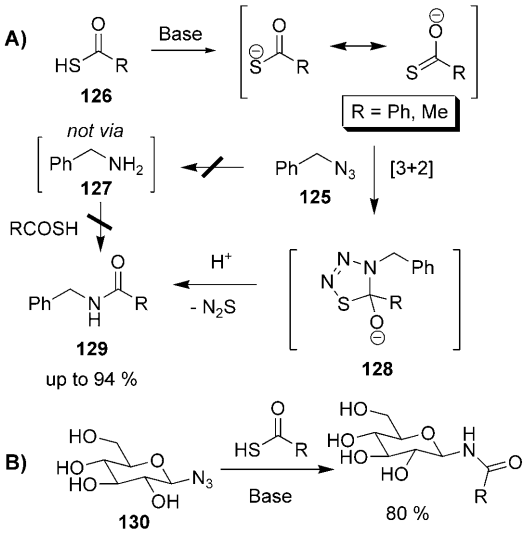
It is important to note that these techniques can be applied to recombinant biological material or isolated natural products; however, special biochemical methods (with the exception of the reactions presented in Section 4.3) have to be used to introduce the required (unnatural) bioorthogonal groups. In other words, an unnatural functional group which will be utilized in chemoselective reactions (as described in several examples in Section 4.2) was already incorporated into the protein before the ligation or modification can occur. In this way, a site-specific modification can be achieved as the unnatural functional group can be addressed selectively from the natural functionalities.

Various methodologies exist for obtaining these modified protein materials decorated with unnatural functional groups by biochemical methods, which are reviewed extensively in the literature. These methods (almost all of which have been used to introduce azide functionalities into biopolymers) include, as mentioned above, the incorporation of unnatural functional groups by using the cellular biosynthetic machinery,[168] for example, in metabolic glycan engineering by unnatural carbohydrate metabolism,[166–168] by unnatural protein translation,[169] by auxotroph expression,[170] by posttranslational[171a,b] or enzymatic[171c,d] modification of protein sequences, or by the semisynthetic strategies discussed in Section 2.3.

4.1. Chemoselective Reactions for Formation of a Native Amide Bond

4.1.1. Thioacid/Azide Amidation

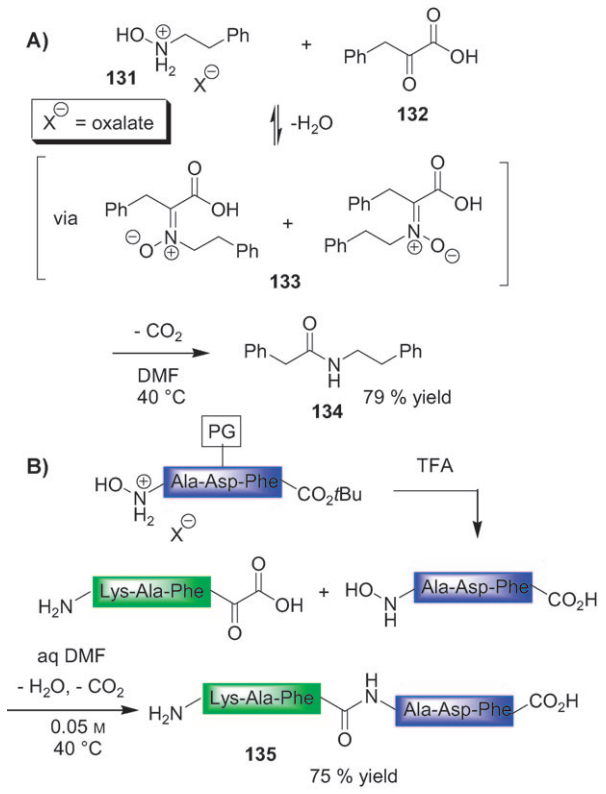
This fascinating reaction involves the formation of an amide bond between azides **125** and thioacids **126**, as described in the 1980s by Hosein Hakimelahi and Just.[172] Interestingly, the formation of **129** does not occur by reduction of the azide to amines **127**, but via the formation of a thiatriazoline derivative **128** through a [3+2] cycloaddition (Scheme 23 A; see also Section 4.2.1).[173] Detailed methodology studies have recently revealed that this reaction can be performed in various (polar and unpolar) organic solvents as well as in water, and many substrates including glycosyl azides **130** are tolerated (Scheme 23 B).[174] Additionally, a $RuCl_3$-promoted variant of the reaction was demonstrated to be high-yielding, even with sterically hindered azides.[175] Although not yet applied to the ligation of longer peptide substrates, this reaction led to the development of a potentially new chemoselective ligation of sulfonamide derivatives.[176]



Scheme 23. A) Mechanism of the thio acid/azide amidation and B) application in the synthesis of β-glycosylamides.

4.1.2. Chemoselective Decarboxylative Amide Ligation

A very promising method published by Bode et al. employs a decarboxylative condensation of *N*-alkylhydroxylamines **131** with α-ketoacids **132** for the chemoselective formation of amide bonds **134** (Scheme 24 A).[177] This reaction does not involve, in contrast to almost all other amide bond forming strategies, an addition–elimination reaction of an activated carboxylic acid derivative, but the decarboxylation and elimination of water from a hemiaminal. The hemiaminal is formed via a nitrone intermediate **133** between the hydroxylamine and the α-ketoacid as demonstrated by mechanistic investigations. Both starting materials are available synthetically,[178] with the hydroxylamine used as an oxalate salt to prevent decomposition. Intensive studies demonstrated that the reaction proceeds at 40 °C in polar solvents in very good yields (70–80 %), and that no epimerization of the α-ketoacids occurs. Most importantly, unprotected peptides **135** containing unprotected Lys and Asp[176] as well as Trp, Tyr, and Arg residues[177] can be ligated, as demonstrated, for example, in the synthesis of a hexapeptide **135** (Scheme 24 B). Possible ligation sites include Phe-Ala, Pro-Ala, Val-Gly, and Ala-Ala. Other hydroxylamine deriv-

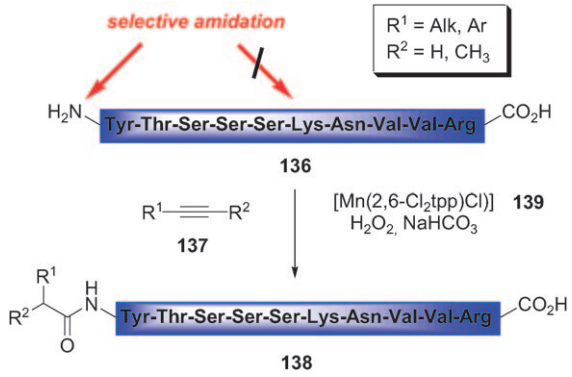


Scheme 24. Chemoselective decarboxylative amide condensation: A) Mechanism and B) application in the synthesis of an unprotected hexapeptide.

atives such as *N*-methoxyalkylamines or isoxazolidine can be used as substrates in this reaction. The latter substrate has been employed in the iterative condensation to give β^3-oligopeptides under aqueous conditions.[179]

4.1.3. *Oxidative Conversion of N-terminal Amines in Peptides into Amides*

The oxidative reaction of an unprotected peptidic N-terminal amine **136** with an alkyne **137** to give amide **138** was recently achieved (Scheme 25).[180] This transformation utilizes a manganese–porphyrin catalyst system **139** with oxone or hydrogen peroxide as the oxidative species. Interestingly, the catalyst can result in the selective transformation of an N-terminal α-amino group of a peptide over other functionalities, such as ε-amino groups of Lys side chains, as demonstrated for various peptide substrates. Moreover, other functional groups do not harm the overall yield of the amide bond formation, although Cys and Met residues are oxidized during the process and have to be reduced again after the transformation with reducing agents. Evidence for the remarkable selectivity has been obtained by structural analysis of a peptide substrate which showed hydrogen bonding of a Lys side chain, thus indicating a conformation effect in addition to electronic effects from the terminal α-amino group.



Scheme 25. Selective oxidative amide synthesis of N-terminal amines in peptides. 2,6-Cl_2tpp = tetrakis(2,6-dichlorphenyl)porphyrin.

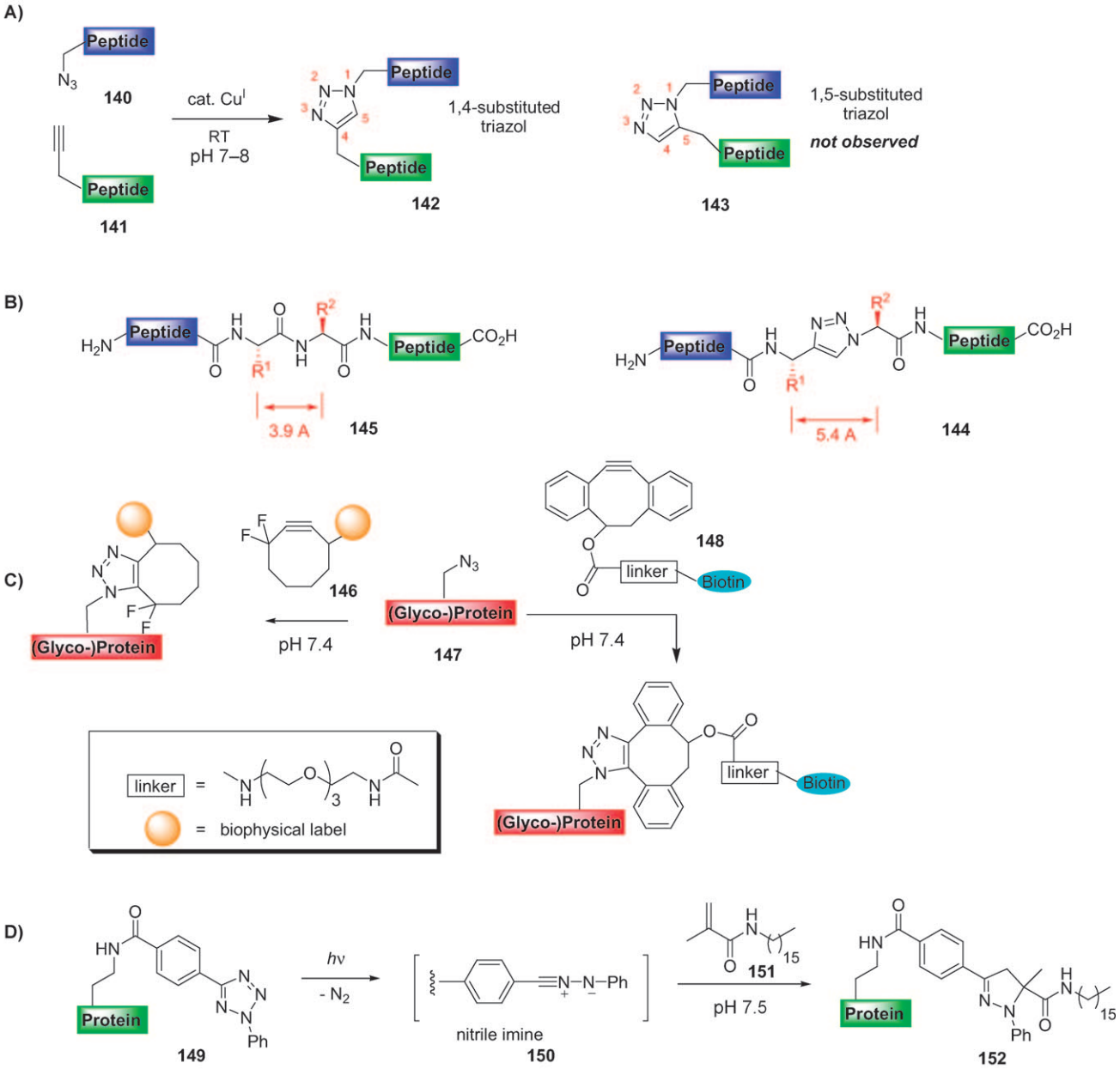
4.2. *Chemoselective Reactions for Formation of Non-Native Bonds*

The reactions described in the following section fall into a particular selectivity pattern: They have in common that two functional ("bioorthogonal") groups react selectively to form a non-native bond in the presence of other functional groups from the protein or peptide environment. Although some of the reactions have already become standard procedures in chemical synthesis and applied in chemical biology, research is still devoted to the identification of new selective, fast, and mild chemoselective reactions which directly deliver biologically active materials in which neither the non-native bond nor the reaction conditions interfere with their biological behavior. Studies are also devoted towards the installation of bioorthogonal functional groups into biopolymers by biochemical techniques. These methods require biologically inert functional groups which can survive biological metabolic pathways.

4.2.1. *1,3-Dipolar Cycloadditions*

4.2.1.1. *Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions: "Click" Coupling*

About 40 years after the initial discovery by Huisgen et al.,[181] the research groups of Meldal[182] and Sharpless[183] developed a Cu^I-catalyzed variant of the [3+2] cycloaddition of azides **140** and alkynes **141**, commonly referred to as the "click" reaction (Scheme 26 A).[184] This cycloaddition yields exclusively the 1,4-substituted triazole isomer, whereas the uncatalyzed version delivers 1,4 and 1,5 isomers (**142** and **143**) in almost equal amounts.[185] Furthermore, the catalyzed reaction is chemoselective to a large number of functional groups, fast, high-yielding, and proceeds under mild conditions in a variety of solvents, including physiological buffer. The reaction has been studied mechanistically,[186] and it has found various applications in materials and polymer science[187] as well as in the specific labeling of bioconjugates,[188] as for example, in monitoring enzyme activities.[189,202b] The copper(I)-catalyzed [3+2] cycloaddition has already been reviewed extensively,[187–193] once even with a major emphasis on the synthesis of peptidomimetics.[194]



Scheme 26. Click coupling: A) Peptide coupling via the Cu^{I}-catalyzed formation of a 1,4-substituted triazole. B) Geometric comparison of a triazole bridge to a native amide bond. C) Strain-promoted cycloaddition of azido proteins. D) Photoinduced 1,3-dipolar cycloaddition between a 2,5-diaryltetrazole-modified protein and an alkene.

The various applications of this reaction in the area of peptide conjugation build upon the three-dimensional orientation of the triazole **144**, which is proposed to be somewhat similar to an amide bond (**145**) in terms of both distance and planarity (Scheme 26B).[194] In addition to Meldal and coworkers, who used click coupling for the generation of peptide libraries in a split-and-mix approach,[195] various other research groups have used this reaction to generate a modification of the heterocyclic peptide backbone[196] in the synthesis of β-turn mimetics,[197] cyclic peptides,[198] *cis*-peptide bond surrogates in proteins,[199] and the construction of peptide–peptide linkages (Scheme 26A).[200] The two reaction partners, azides **140** and alkynes **141**, can be introduced into the peptide substrates by standard SPPS techniques, for example, by a final coupling with an α-azidoamino acid, which can be derived by a copper(II)-catalyzed transfer of a diazo group from the corresponding α-amino acids.[201]

Azides and even alkynes represent functional groups for a straightforward incorporation into biopolymeric materials, and the click reaction can be regarded as chemoselective in biological environments. Thus, this reaction has been used frequently for the bioconjugation of different natural products (including proteins,[202] nonribosomal peptides,[203] and natural products[204]) as well as in the immobilization of

peptides and proteins on surfaces.[205] In the last case, an azide or alkyne is introduced at the C terminus of a protein **172** by EPL to enable attachment to an appropriately functionalized glass surface (see Scheme 29 B in Section 4.2.4).[206,207] Further applications of this reaction include the conjugation of peptides to polymer architectures,[208] the conjugation of radiolabels,[209] and the attachment of carbohydrates[99,210] or lipids[211] to peptide or protein side chains.

4.2.1.2. *Strain-Promoted Cycloaddition*

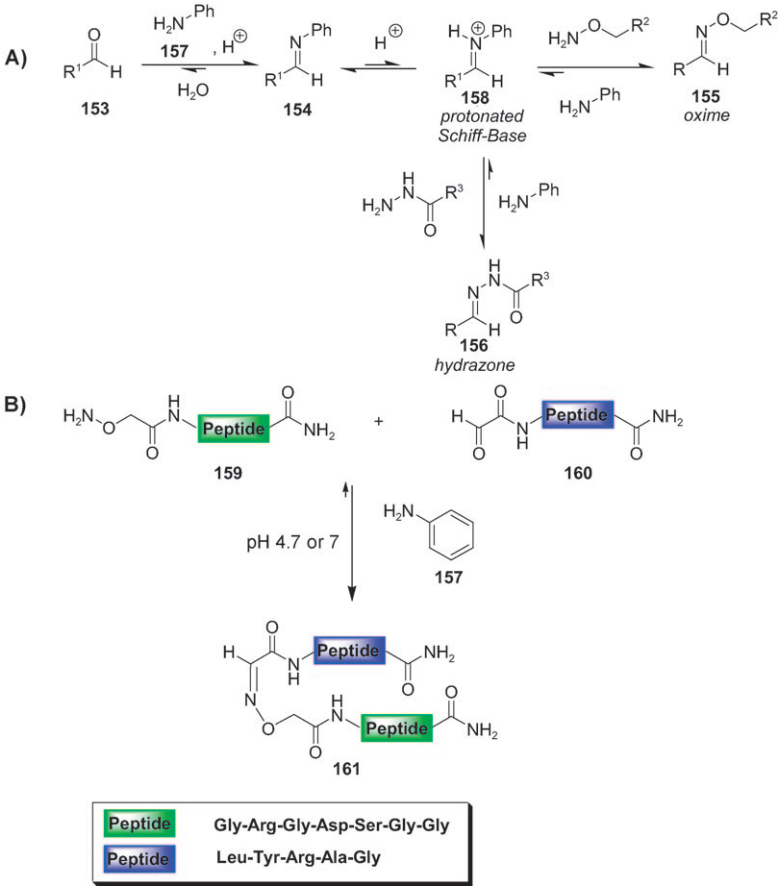
One of the major drawbacks of the copper(I)-catalyzed "click coupling" is the toxicity of copper, which limits its use for in vivo applications.[212] For this purpose, Bertozzi et al. developed a "strain-promoted" cycloaddition reaction in which the ring strain of a cyclic alkyne **146** (cyclooctyne) is used as the driving force for the cycloaddition.[213] The introduction of two fluoride substituents in the α position to the cycloalkyne leads to an acceleration of the reaction and allows a fast bioconjugation of biophysical or functional labels to azido-functionalized (glyco-)proteins **147** for studies in live cells[214] as well as in zebrafish[215] (Scheme 26 C). A dimethoxy-substituted azacyclooctyne was developed to enhance the hydrophilicity and reduce nonspecific binding events.[216] Another recent example employed 4-dibenzocyclooctynols **148** for labeling experiments with azido-containing glycan biopolymers in living cells.[217] In **148**, the two aromatic rings increase the ring strain (and the reactivity for the formation of the triazole), while additionally lowering the tendency of nucleophilic attack to the triple bond by biological nucleophiles.

4.2.1.3. *Photoinduced 1,3-Dipolar Cycloaddition*

In a very recent publication, a photoinducible 1,3-dipolar cyclodaddition was reported (Scheme 26 D).[218] In this reaction, a diaryl-substituted tetrazole derivative in **149**, which can be incorporated into a protein, undergoes a cycloreversion reaction upon irradiation with UV light to form in situ a nitrile imine (**150**). This 1,3-dipole can react with an electron-deficient alkene, such as **151**, to form a pyrazoline cycloadduct **152**. This chemoselective reaction to give **152** is very fast (1 minute) in buffer systems at room temperature, which allowed a chemoselective lipidation strategy by reaction with an appropriate lipid dipolarophile **151**. However, its application in bioorthogonal coupling strategies for similar in vivo studies has still to be demonstrated, as the tetrazole moiety (in contrast to well-established techniques for azides[212a]) has to be incorporated by chemical strategies into the biopolymer. In the original report, the tetrazole was either coupled to a surface-exposed Lys nucleophile of a GFP protein or incorporated by a semisynthetic strategy.[219]

4.2.2. *Oxime and Hydrazone Condensation by Nucleophilic Amine Catalysis*

The imine condensation of ketones and aldehydes **153** with primary amine nucleophiles has drawn significant attention in the area of bioconjugation (Scheme 27; see also Scheme 21).[220] The reaction with primary amines proceeds preferably under acidic conditions with reversible formation of an imine **154**.[221] This imine can be reduced with $NaCNBH_3$ to form a non-hydrolyzable bond. Alternatively, an amine nucleophile with an electron-donating group in the α position can be used to form an oxime **155** or hydrazone **156**. This concept has been employed in a site-specific modification of proteins both in vitro[222] and in vivo,[171,223,224] in which a ketone is installed into a protein by the amber suppressor tRNA method or by metabolic glycan engineering and modified with a hydrazine-containing fluorophore.[225] In two recent protocols, Dawson et al. have now utilized the previously developed concept of nucleophilic amine catalysis during the formation of semicarbazones[226] in oxime and hydrazone ligations to improve the formation of a non-hydrolyzable bond under physiological conditions.[227,228] This concept utilizes aniline (**157**) as a catalyst, which leads to the presence of a significant amount of a protonated Schiff-base **158** even at pH 7. The Schiff base can then rapidly undergo the desired oxime or hydrazone formation (Scheme 27 A).

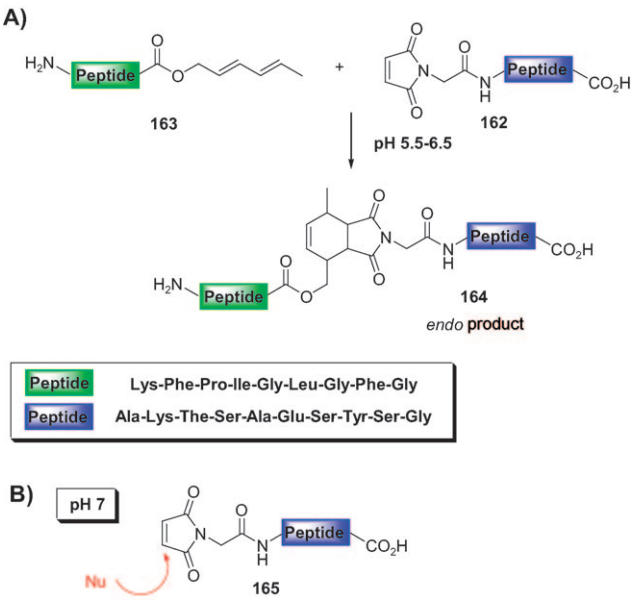


Scheme 27. A) Oxime and hydrazone formation and B) subsequent peptide ligation through nucleophilic amine catalysis.

Both reactions were applied in the ligation of peptides from peptide precursors (as shown Scheme 27B for oxime formation at pH 4.5 or 7 in **161**) and should be applicable for various protein conjugation and ligations strategies in the near future.[227b]

4.2.3. *Peptide Ligation by Diels–Alder Reactions*

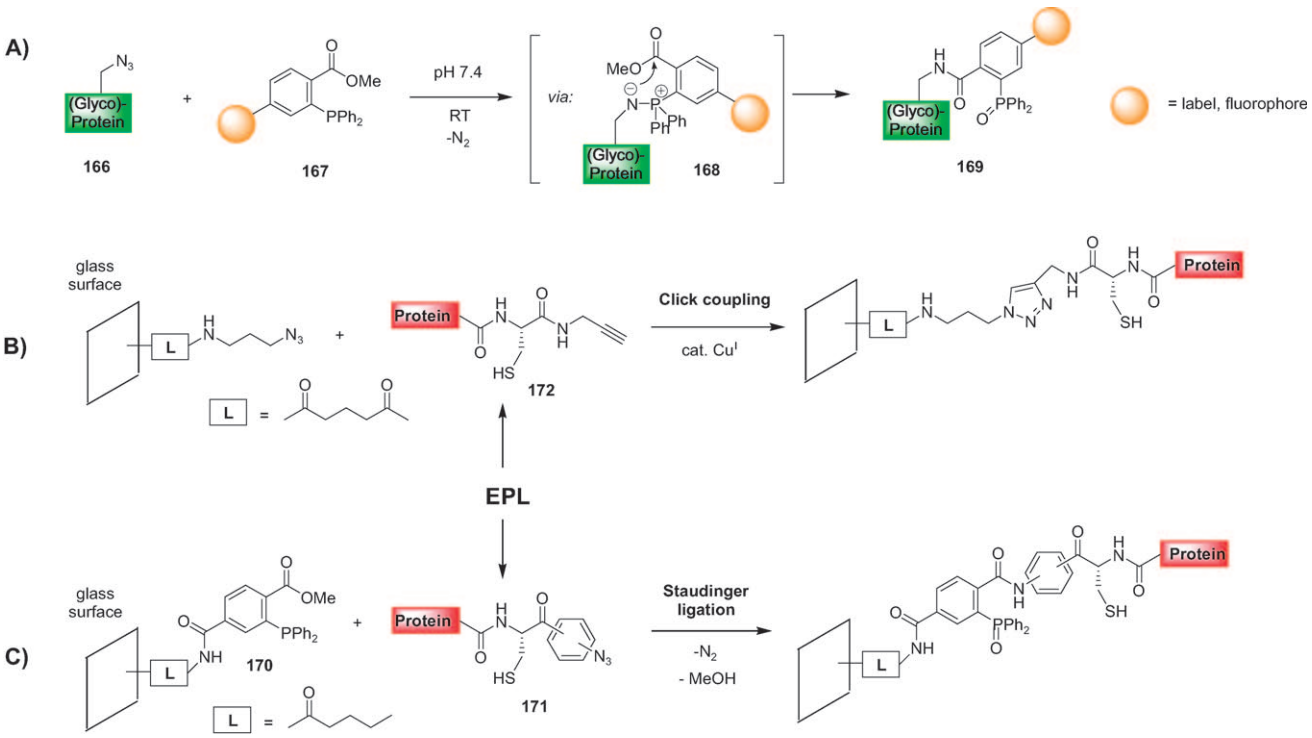
In another application of the maleimido group, Waldmann and co-workers recently identified the [4+2] cycloaddition as a chemoselective reaction for the conjugation of peptides and for the functionalization of proteins (Scheme 28A).[229] In this case, a maleimide-capped peptide **162** reacts as the dienophile with an N peptide containing a C-terminal diene **163** moiety to form an *endo*-Diels–Alder product **164**. The diene unit can be introduced into the peptide by using synthetic methods usually applied to the synthesis of thioesters (see Section 2.3.3). The Diels–Alder reaction has to be performed under slightly acidic conditions to be chemoselective (pH 5.5–6.5), thereby avoiding undesired conjugated addition of amine nucleophiles to **165** at neutral pH (Scheme 28B). These reaction conditions can be advantageous in chemoselective coupling strategies which do not work under the conditions of EPL, as demonstrated for a site-specific modification of the Rab7 protein.



Scheme 28. A) Diels–Alder conjugation of peptides and B) side reaction at neutral pH.

4.2.4. *Staudinger Ligation*

One of the first chemoselective reactions employed for the site-specific functionalization of biopolymers was the Staudinger ligation, which was developed by Bertozzi and coworkers.[150,158a,230] In its non-traceless variant, an iminophosphorane **168** (see Section 3.4.1) is formed between an azido-bearing biopolymer **166** and phosphanylaryl **167** (Scheme 29A). This iminophosphorane **168** can undergo an intramolecular nucleophilic attack with an *ortho*-substituted ester as an internal electrophile to form an amide in the conjugate **169**, which now carries a phosphinoxide scaffold. In its first landmark application, the Staudinger ligation was applied to the selective labeling of sialic acid residues which had been selectively functionalized with azides by metabolic



Scheme 29. A) Mechanism of the Staudinger ligation. B,C) Protein immobilization by click coupling (B) and Staudinger ligation (C).

glycan engineering. This process thus demonstrated the potential of the bioorthogonal reporter strategy.[158a,231] One of the major advantages of this reaction over the click reaction is the fact that no metals are required. The Staudinger ligation has found various very important applications in labeling biomolcules, particularly for reactions in living organisms.[232] However, undesired oxidation of the phosphine and the competing Staudinger reduction sometimes limits its general applicability when quantitative conversions are desired.[233] Additionally, this reaction has been employed for site-selective functionalization of proteins or their modification,[166,168d,234] DNA,[234b,e] and the immobilization of peptides[235] and proteins[236] on appropriately phosphine-functionalized glass surfaces **170** with an azide functionality introduced into the protein target **171** by EPL (Scheme 27C).

4.2.5. *Oxidative Modification of Aniline in Proteins*

Oxidative transformations with $NaIO_4$ have found various applications in protein activation processes, including the oxidative cleavage of N-terminal Ser residues,[237] aldehyde formation from glycans,[238a–c] and cross-linking reactions,[238d] as they only interfere with sulfur-containing amino acids in proteins.[238e] In another chemoselective process with $NaIO_4$, anilines can be coupled to phenylenediamine derivatives in low concentrations at pH 6.5.[239] This reaction was subsequently applied in the modification of proteins **173**, in which the aniline functionality was introduced by unnatural protein expression[240] or semisynthesis, as well as in the functionalization of virus capsids with phenylenediamine-containing peptides **174** to give protein–peptide conjugates **175** (Scheme 30A).[241]

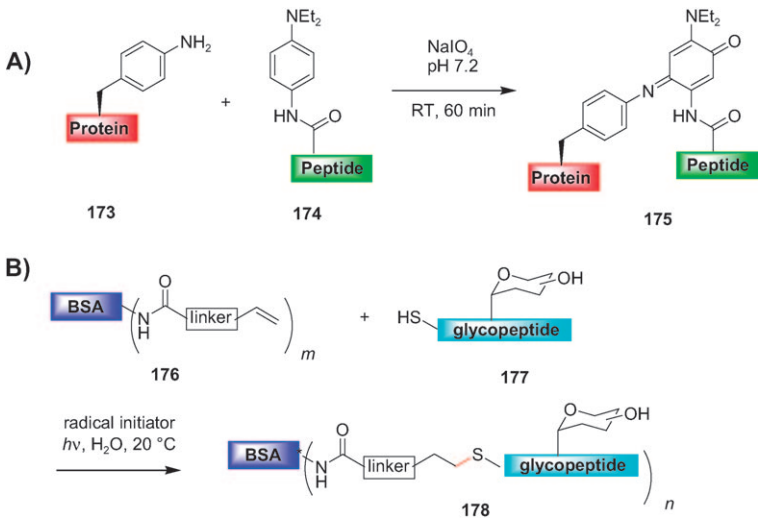
4.2.6. *Radical-Induced Thioether Formation*

A common bioconjugation technique, which has been applied particularly in the conjugation of glycans,[242] is the formation of thioethers from thiols and alkenes under radical conditions.[243] The reactions can be induced photochemically or by the addition of a radical starter such as AIBN. Kunz et al. recently applied this reaction to the functionalization of protein systems, since it tolerates the various functional groups present in peptides and proteins and yielded only disulfides as minor side products.[244] This reaction proceeds under mild conditions (either through a thermally or photochemically induced radical formation) and yields an immune-compatible thioether linkage. Thus it has been applied to the generation of synthetic vaccines **178** by conjugation of glycopeptides **177** to the protein carrier bovine serum albumin (BSA) **176**, which has previously been nonspecifically functionalized with alkene end groups. Although not demonstrated yet, this reaction has the potential to site-selectively modify alkene-functionalized proteins and for the functionalization of cysteine residues in proteins, which nicely leads to the next section.

4.3. *Selective Transformations of Natural Amino Acids*

A particular advantage of the chemical bioconjugation techniques in yielding covalent linkages, described in the following, is the fact that they address naturally occurring amino acids and therefore do not require elaborate biochemical techniques to install unnatural functionalities for chemoselective reactions (see Section 4.2). As already described in Section 3.2, a selective chemical modification of Cys residues can alter the functionality of a protein. In the following examples, chemoselective strategies have been developed to attach peptides to the protein backbone, thereby introducing new functional building blocks into a biopolymer. These examples proceed selectively with only one type of amino acid and therefore differentiate themselves from other labeling approaches which nonselectively functionalize various amino acids, such as the reaction of nucleophilic side chains with succinimide esters. Nevertheless, to obtain a site-selective functionalization of proteins by the methods described here (in contrast to the previous section), one has to rely on spatial control within the overall protein architecture, for example, in a unique solvent-exposed amino acid.

It is important to note that various other bioconjugation techniques exist which allow the site-specific labeling of proteins. These approaches include labeling by specific chelation, such as the binding of biarsenical fluorophores (for example, FlAsH) to tetracysteine motifs[245] or lanthanide binding tags (LBT),[246] enzyme-catalyzed labeling by posttranslational modification (for example, mediated by carrier protein tags),[247] and self-labeling (for example, by the SNAP technology).[248] More information about these methods can be found in the excellent reviews covering this topic.[249]



Scheme 30. A) Oxidative peptide coupling of aniline-containing proteins. B) Glycopeptide conjugation to protein carriers by radical-induced formation of a thioether.

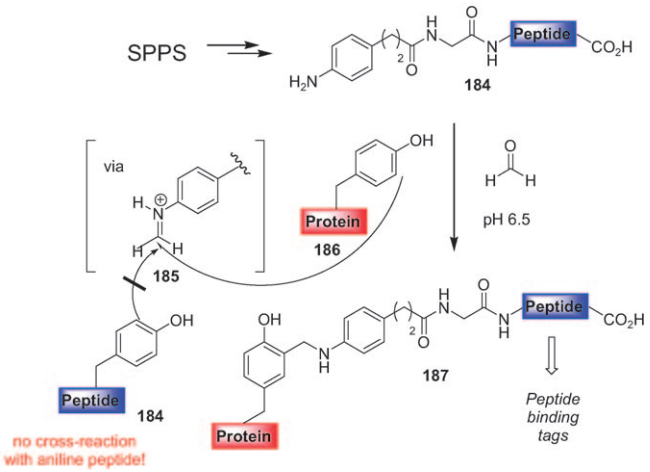
4.3.1. *Conjugated Addition of Sulfur Nucleophiles: MIC Strategy*

The chemoselective Michael addition of sulfur nucleophiles to α,β-unsaturated carbonyl derivatives, such as maleimides, is a well-established conjugation technique that is commonly employed for the attachment of fluorophores to biologically relevant molecules containing sulfur nucleophiles under physiological conditions (pH 7, room temperature). This reaction was employed in a ligation technique in which a maleimidocaproyl (MIC) peptide unit **179** functioned as the electrophilic reactant for a Cys residue in a protein (Scheme 31 A).[250] MIC building blocks **180** ("MIC-OH") can be introduced by SPPS in a straightforward manner even into delicate lipid-modified peptide sequences,[251] or they can be combined with a photoactive group such as benzophenone.[252] Although the site selectivity of this reaction is limited by other Cys residues present in the protein, it has become a widely applied method[253] for selective functionalization of proteins containing only one solvent-exposed Cys. One such example is the synthesis of Ras proteins **181** containing specific protein modifications (Scheme 28 B; for a detailed discussion of this prominent example, see Section 6.3).

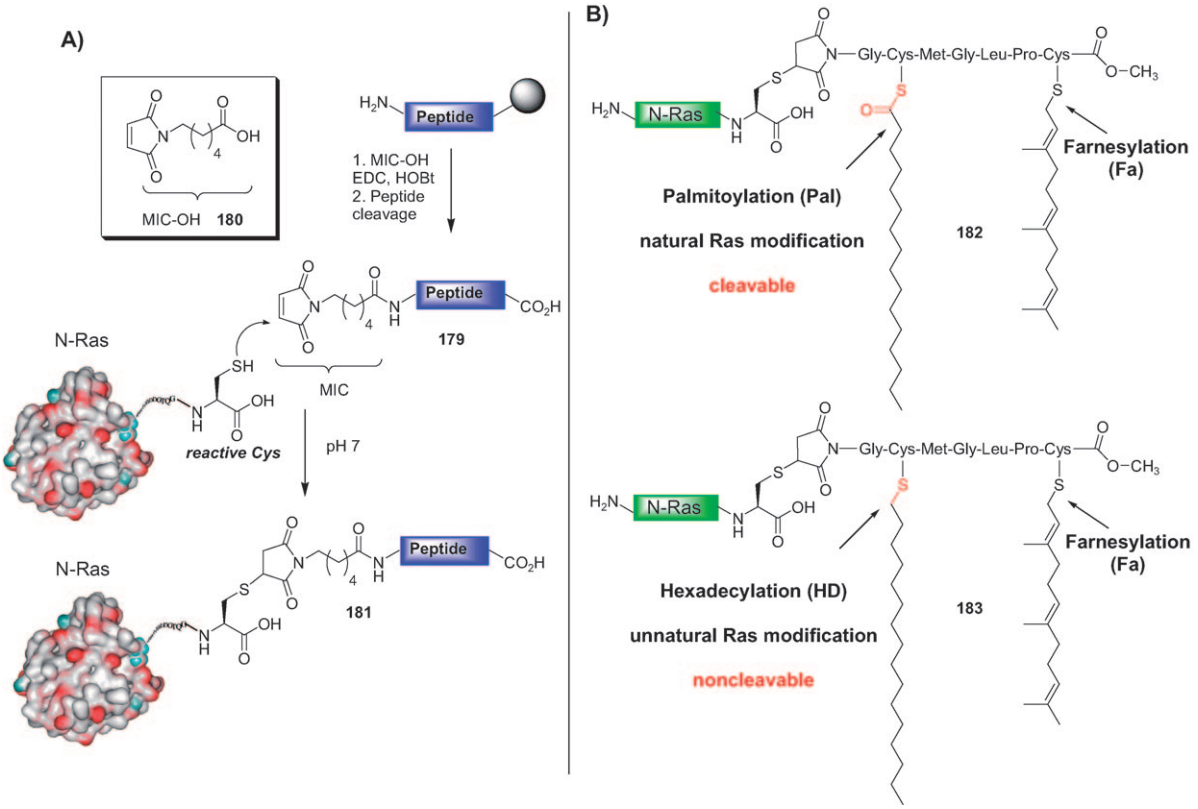
4.3.2. *Mannich Peptide Conjugation to Proteins through Tyr Residues*

Another amino acid that posses unique chemical behavior is Tyr because of its phenolic side chain, which exhibits distinct nucleophilic and electrophilic properties. Moreover, because of the amphiphilicity, Tyr residues can often be found on the surface of proteins and are thus ideal anchors for selective bioconjugation strategies.

A selective conjugation of peptides to proteins has been achieved by a Mannich-type coupling reaction.[254] In this transformation an aniline group displayed at the N terminus of a peptide **184** is reacted with an excess of formaldehyde at pH 6.5 to give an iminium salt **185** (Scheme 32).[255] The latter can undergo electrophilic aromatic substitution at a Tyr



Scheme 32. Mannich-type conjugation of peptides to proteins.



Scheme 31. Conjugated thiol addition: A) MIC-strategy and B) lipidated Ras analogues. EDC = 1-ethyl-3-(3-dimethylaminopropyl)carboiimide, HOBt = *N*-hydroxybenzotriazole, MIC = maleimidocaproyl.

residue in protein **186** and thus lead to the formation of a covalent bond in **187**. This conjugation, which can also functionalize Trp residues,[256,257] has been applied to the attachment of various peptide binding tags, including FLAG epitopes[258] and lanthanide binding tags (LBT),[246] to the protein chymotrypsinogen. Generally, a conversion of up to 84% was achieved at the aniline tag in peptides with several Gly residues, which indicates that a flexible region was crucial for the Tyr attachment to occur. Interestingly, the overall conversion for the Mannich-type conjugation was higher for protein substrates than for peptides or small molecules and led to no cross-reactivity with Tyr residues in the aniline-containing peptides **184**.

Other approaches to selectively functionalize Tyr residues, although not yet demonstrated for peptide attachment, include an alkylation strategy with π-allylpalladium complexes which proceeds at pH 8.5–9 in aqueous buffers.[259] In an application of this strategy, chymotrypsinogen A was transformed into a lipidated (farnesylated) analogue which was embedded into lipid bilayers. Finally, Tyr residues were transformed in a two-step process, first by reaction with diazonium derivatives and second by a Diels–Alder reaction[260] or the formation of an oxime.[261] This concept was demonstrated in a surface modification of a tobacco virus and a bacteriophage.[262]

5. Chemoselective Peptide Ligations Catalyzed by Enzymes

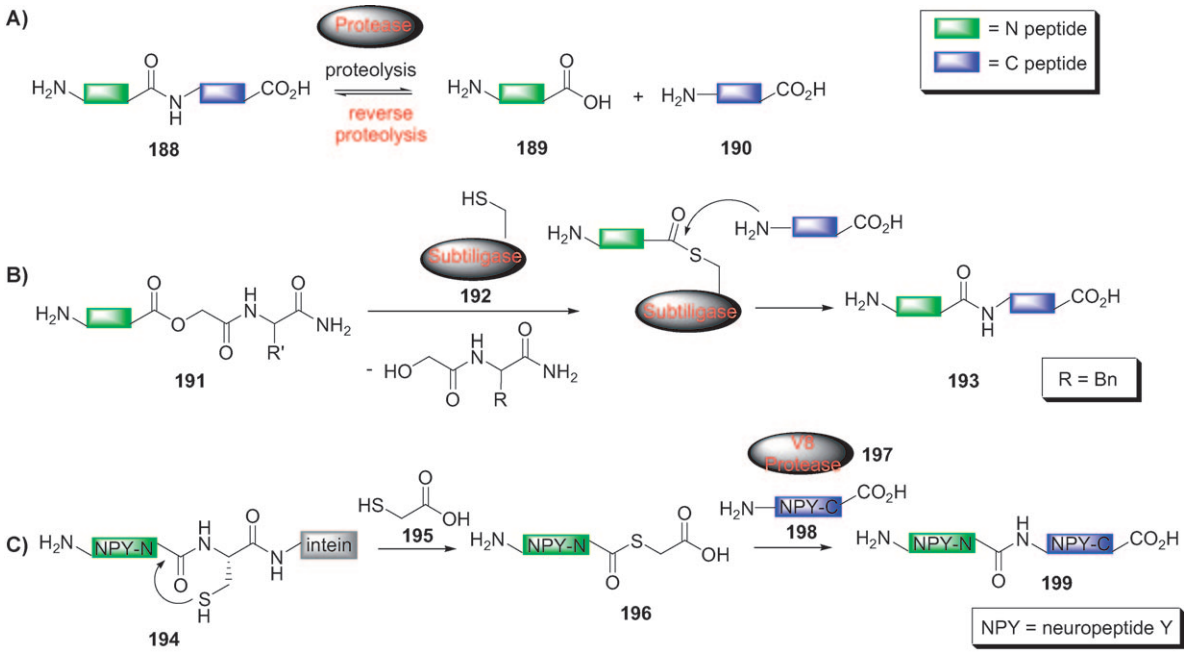
The capture/rearrangement concept has been very successful for achieving chemoselective ligations of peptides. However, alternative strategies do exist which exploit the inherent chemoselectivity of enzymes for ligation purposes. The three most prominent examples include the already discussed usage of split inteins for protein trans-splicing (Section 2.4.3), the reverse proteolysis, and employment of the transpeptidases sortase A.

5.1. Reverse Proteolysis

Whereas enzymes that catalyze the ligation of polypeptides (ligases) are rarely found in nature, proteases which catalyze their cleavage are common. Evolved over millions of years, proteases often cleave peptides at defined positions and with high specificity.[263] For proteases to be used for chemoselective peptide ligations they need to be converted into peptide ligases by reversing their proteolytic activity (see Scheme 30 A). A review of protease usage for peptide ligations is beyond the scope of this Review, and hence we focus only on the basic concepts. More detailed publications reviewing the concepts and applications can been found elsewere.[263–266]

Like all biocatalysts, proteases are capable of catalyzing the reverse reaction, namely a reverse proteolysis. However, under physiological conditions the equilibrium position is shifted towards the side of the cleaved (**189** and **190**) rather than the ligated product (**188**; Scheme 33). There are two basic strategies to manipulate proteases for peptide ligations: the thermodynamically controlled and the kinetically controlled approach.

The thermodynamically controlled approach represents a true reversal of the cleavage reaction by changing the position of equilibrium. Under physiological conditions, reverse proteolysis is prohibited by the energy required for the initial transfer of a proton from the nucleophile of the C peptide onto the carboxylate of the N peptide.[263,265] Organic solvents



Scheme 33. A) General concept of protease-mediated peptide ligations. B) Subtiligase-mediated peptide ligation. C) Expressed enzyme ligation (EEL).

can be added to the reaction media to lower the dielectric constant of the reaction medium and thereby the acidity of the carboxyl group. As a result, proteases in organic solvents or mixtures of water and organic solvents are capable of catalyzing peptide ligation; however, optimization of the reaction conditions and engineering of the proteases is often required for such approaches.[263,267,268]

In the case of the kinetically controlled reverse proteolysis, the equilibrium position remains unchanged. The basic principle of this approach lies in the rapid accumulation of an electrophilic N peptide protease intermediate which can react with the α-amino group of the C peptide.[264] This approach is only feasible with serine or cysteine proteases which possess a covalent enzyme–peptide intermediate in their reaction cycle. The acylation of the protease with the N peptide is achieved by using activated peptides, which usually possess C-terminal esters **191** (or thioesters) as leaving groups (Scheme 33B). Since kinetically controlled reverse proteolysis does not alter the equilibrium, the ligation reaction will have to be stopped when the acyl donor **191** is consumed to avoid cleavage of the ligation product **193**. An important improvement in this regard is represented by the introduction of so-called "substrate mimetics", where the chemical functions mediating the specific recognition of the protease are transferred into the leaving group. As a result, the protease recognition sequence is lost in the ligation product.[269–271]

Protein engineering has been applied to improve the performance of proteases for reverse proteolysis. The most prominent example is the so-called subtiligase, which is a double mutant derived from the bacterial protease subtilisin BPN′ (**192**, Scheme 33B).[272] Mutation of the active site serine residue to a cysteine residue improved the cleavage/ligation ratio, and a second proline to alanine mutation was introduced to reduce the steric crowding at the active site.[272]

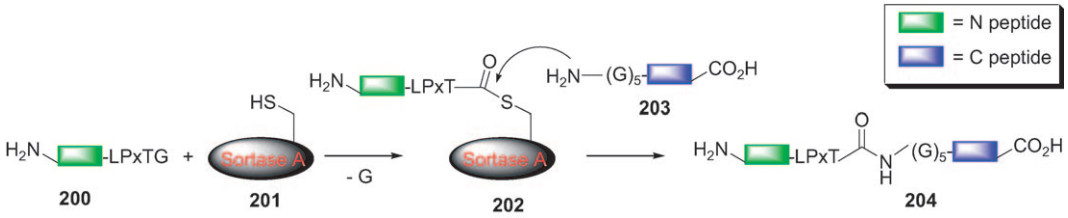
Protease-mediated ligations can be combined with other chemoselective ligation techniques such as EPL. One such procedure termed "expressed enzymatic ligation" (EEL) was applied by Beck-Sickinger and co-workers for the synthesis of the prohormone neuropeptide Y (**199**, 69 amino acids) from two fragments (Scheme 33C).[273] The N fragment **196** (NPY-N) was generated as a C-terminal thioester through EPL by using 2-sulfanylacetic acid (**195**) as the thiol for the thiolysis of the intein fusion protein **194**. The resulting peptide thioester served as a substrate mimetic for the V8 protease **197**, which catalyzed the following ligation with the synthetic C fragment **198** (NPY-C). The advantage of EEL over standard EPL reactions is that proteases are not limited to Cys as the N-terminal amino acid of the C peptide. This feature could be used in this case for the "Cys-free" ligation with the NPY-C peptide carrying the native Ser residue at the N terminus.

5.2. Sortase-Mediated Peptide Ligation

Only a few examples of enzymes are known that naturally catalyze peptide ligations. One prominent example is transpeptidase sortase A (**201**), which has been used successfully for chemoselective peptide ligations. Sortase A is an enzyme that catalyzes the attachment of surface proteins **200** to the cell wall of Gram-positive bacteria such as the pathogenic *Staphylococcus aureus*.[274] The proteins **200** contain a C-terminal sorting signal which includes a highly conserved LPxTG motive. This motive is cleaved by sortase A at the threonine residue, which results in a covalent acyl intermediate **202** attached to the active site cysteine of sortase A.[275] This acylation step is comparable to that catalyzed by serine/cysteine proteases; however the favored deacylating nucleophile of the sortase A is the α-amino group of the peptidoglycan penta-Gly unit (Scheme 34).

To utilize sortase A for chemoselective ligations, the N-terminal peptide needs to include the LPxTG sorting motive at the C terminus, and the C-terminal peptide **203** needs at least one Gly residue as the N-terminal amino acid.[276] Upon addition of sortase, the N peptide is cleaved at the sorting signal and subsequently ligated to the N-terminal Gly moiety of the C peptide. The ability of sortase A to mediate peptide ligations to **204** could be demonstrated for several cases. For example, recombinant green-fluorescent protein (GFP) containing the LPxTG motive at the C terminus was ligated with a folate-conjugated peptide comprising three Gly residues at the N terminus.[276] In another example, sortase A was used to ligate peptide nucleic acid (PNA) to a cell-penetrating peptide which allowed the delivery of the ligation product into living cells.[277] Finally, sortase A was used for the ligation of fluorescently labeled peptides to proteins containing the LPxTG motive exposed on the surface of living cells.[278]

Enzyme-mediated peptide ligation techniques, such as reverse proteolysis or sortase ligation, have gained much attention in recent times.[263,276–278] However, some drawbacks still exist, including the need for activated peptides in kinetically controlled reverse proteolysis and the LPxTG motive which remains in peptides ligated by sortase A. Nevertheless, despite these limitations, enzyme-mediated ligation techniques are a good alternative to conventional ligation techniques following a capture/rearrangement approach (see Section 7.6).



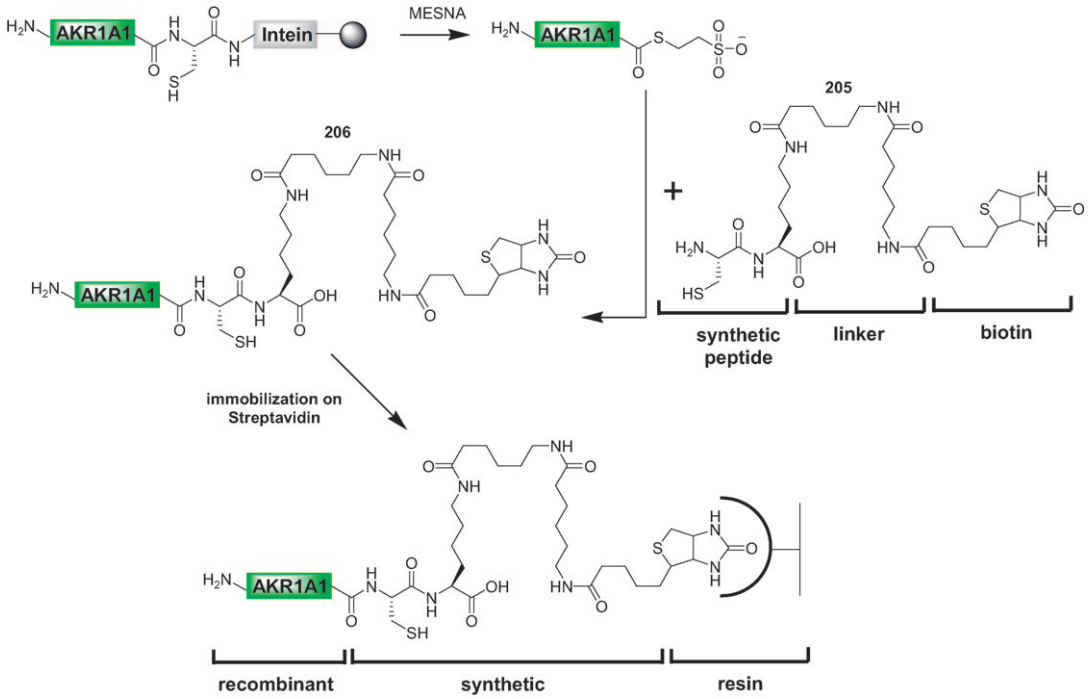
Scheme 34. Sortase-mediated ligation.

6. Protein Semisynthesis: Ligation of Synthetic and Recombinant Peptides

A particularly useful application of chemoselective ligation reactions is the linkage of synthetic peptides and recombinant proteins, which is often referred to as "protein semisynthesis" (see also Section 2.4).[3,9,279–281] This approach combines the advantages of techniques from organic synthesis and recombinant protein expression: on one hand, recombinant peptides are not limited in their size; on the other hand, synthetic peptides are not limited to the 20 proteinogenic amino acids. The installation of a modification into a protein of interest can be achieved by producing the larger fragment of the protein in bacteria and the smaller fragment, including the desired modification, by SPPS. Upon ligation of these two fragments, the full-length modified protein is reconstituted and can be used for biophysical or functional studies. Importantly, protein semisynthesis enables the installation of natural modifications including phosphorylation,[74,282,283] acetylation or methylation,[105] glycosylation,[284,285] and lipidation[286] as well as unnatural modifications or amino acids such as aminophenylalanine,[287] fluorotyrosines,[288] ATP-linked phenylalanine,[289] D-amino acids,[290] or phosphoamino acid mimics.[291] Furthermore, fluorescent probes or affinity tags can be installed in a defined and homogeneous manner, which is an important improvement over nonselective methods which often suffer from a lack of specificity.

6.1. Immobilized Proteins

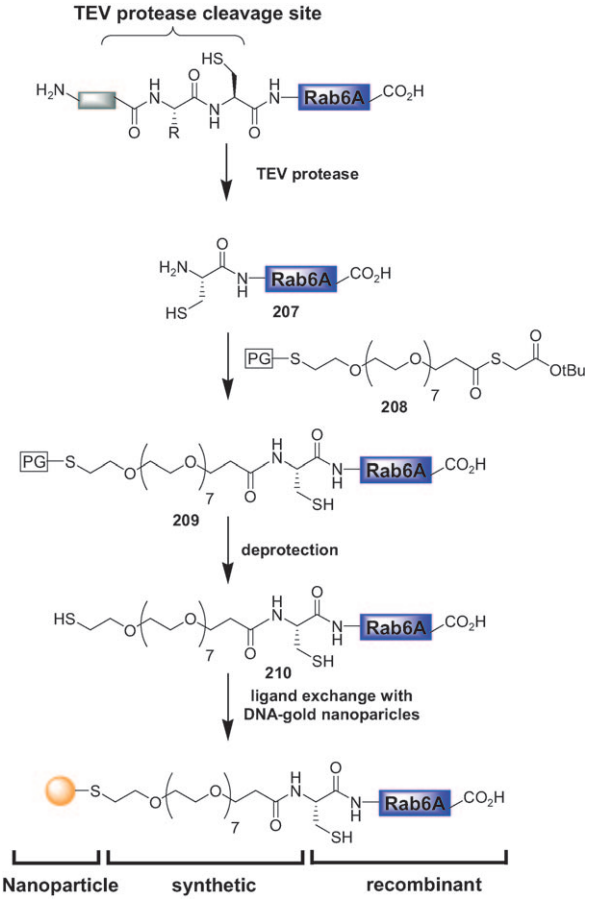
One application is the defined installation of functional groups which can facilitate immobilization on a solid support or nanoparticle (see also Scheme 27).[292,293] For example, proteins immobilized on microarrays are powerful tools for studying protein activity or inhibition on a very small scale and with high throughput. The mode of immobilization is critical for the quality of the resulting microarrays.[294] Common attachment techniques use the cross-linking of surface-exposed functional groups to the solid support. A major drawback of this approach is a random orientation of the attached proteins, which results in inhomogeneous activities. The defined and selective installation of unique functional groups or tags through protein semisynthesis can be used to solve this problem. The feasibility of site-specific protein attachment through semisynthesis has been demonstrated for several proteins, including maltose-binding protein (MBP), glutathione S-transferase (GST), and aldo/keto reductase AKR1A1.[295,296] In these cases, the high affinity of the interaction between biotin and streptavidin was exploited for protein immobilization. The biotin moiety including a linker was introduced in **206** through ligation of a biotin-conjugated peptide **205** to the C terminus of the proteins using the EPL strategy.[295,296] The subsequent immobilization on streptavidin-coated glass slides resulted in a defined orientation of the proteins (homogeneous protein immobilization). The advantage of homogenous protein immobilization was recently demonstrated for AKR1A1. The enzyme was biotinylated either through specific attachment of a biotin moiety using protein semisynthesis (by EPL) or by cross-linking the biotin tag to surface-exposed functional groups (Scheme 35). The kinetic properties of both biotinylated proteins were determined in solution and immobilized on streptavidin-coated slides. Both proteins displayed similar enzymatic activities in solution, comparable with unmodified AKR1A1. However, upon immobilization, the cross-linked



Scheme 35. Protein immobilization on resins illustrated by the example of AKR1A1.

enzyme displayed a tenfold decreased activity, whereas the homogeneous biotinylated enzyme was as active as the unmodified enzyme.[296]

Nanoparticles represent another common target for protein immobilization and have found several biotechnological applications. As in the attachment to glass slides, unspecific protein-immobilization techniques for nanoparticles commonly result in a random orientation. Protein semisynthesis can be used to induce homogeneity in the mode of attachment, as recently demonstrated for the selective attachment of the small GTPase Rab6A on DNA–Gold nanoparticles (Scheme 36).[297] Another beneficial effect



Scheme 36. Conjugation of Rab6A to gold nanoparticles.

can be achieved by introducing a defined linker between the protein and the nanoparticle. To fulfill both needs a polyethylene glycol (PEG) linker **208** was ligated to the N terminus of Rab6A using NCL (**207**). The PEG linker was further deprotected to give a free thiol **209** to allow conjugation to the gold particles. After immobilization of **210**, this highly exposed thiol was found to be the preferential binding site (Scheme 36). After the conjugation had been achieved by a ligand-exchange reaction, biochemical assays confirmed that the attached Rab6A proteins had the expected GTPase activity, in contrast to Rab6A proteins unspecifically conjugated through surface-exposed thiol groups.[297]

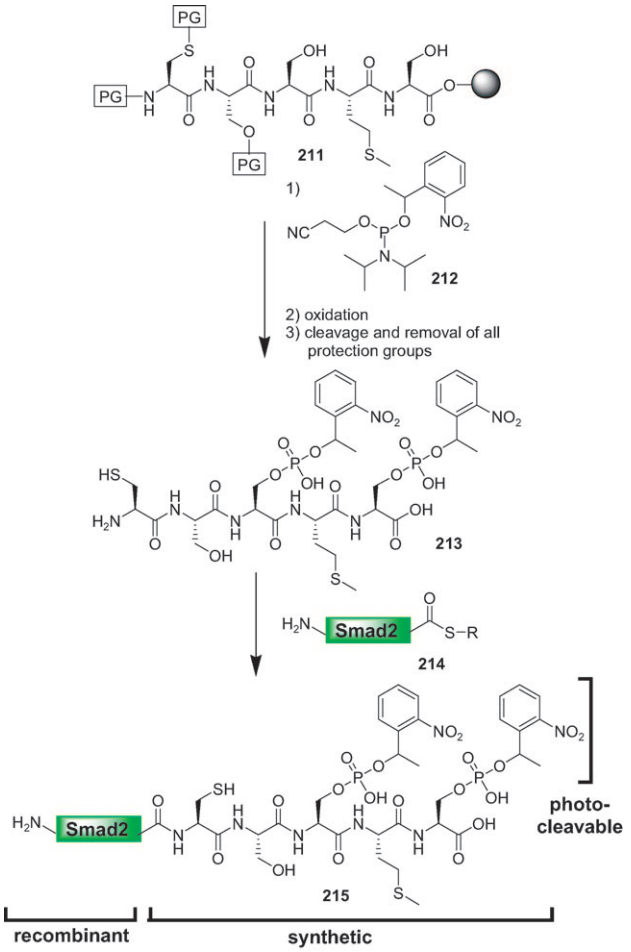
6.2. Protein Phosphorylation

Another important area of application of protein semisynthesis is the introduction of natural modifications in proteins. Several types of post- and cotranslational protein modifications occur naturally in living organisms and they are commonly an important means of regulating enzymatic activities, protein stability, and localization.

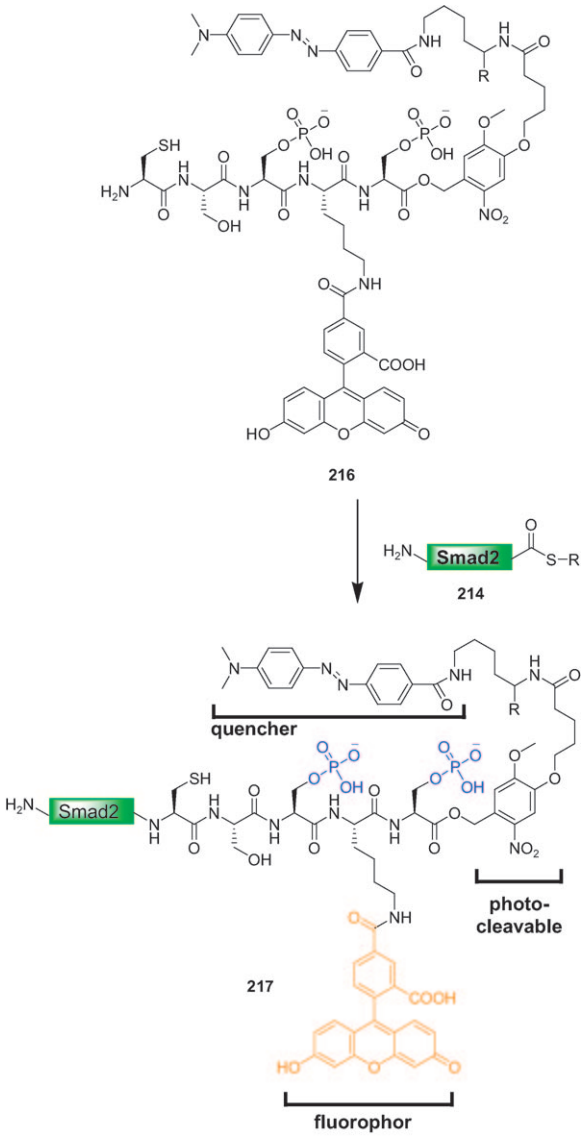
To date the phosphorylation of Ser, Thr, and Tyr residues represent the most prominent modifications found in proteins, thus making the investigation of their physiological function an important scientific task.[298] Since the kinases that introduce the phosphorylations are often unknown or not biochemically accessible, protein semisynthesis has become an important tool for their selective installation.[283,299] In addition, protein semisynthesis can also be applied to install chemical groups that allow the masking and selective unmasking of protein phosphorylations.

A recent example illustrating the potential of the semisynthesis of modified proteins is the generation of caged phosphorylated Smad2 proteins. Smad2 is a signaling protein which is activated in response to the binding of the transforming growth factor β to its cognate receptor.[300] Activation of Smad2 is mediated by phosphorylation of two Ser residues located at the C terminus of the protein. Phosphorylated Smad2 diffuses off its cytosolic binding partner SARA and translocates into the nucleus where it serves as a transcriptional activator.[300] To study this phosphorylation-dependent translocation, the presence or absence of this modification needed to be controlled on a very short timescale. In this regard protein semisynthesis was used to install caged phosphoserine residues at the two phosphorylation sites of Smad2.[301] The caging groups can conceal the presence of the modifications until they are removed by exposure to UV light.[302,303] Smad2 was successfully targeted by two different caging strategies that utilized protein semisynthesis.[301,304] In the first case, two individually caged phosphorylated Ser residues were incorporated by EPL (Scheme 37). The synthesis started with a short synthetic peptide **211**, which corresponds to the five C-terminal amino acids of Smad2. 2-Nitrophenylethyl (NEP) caged phosphate groups were conjugated to the two serine residues by treating the orthogonally deprotected hydroxy groups with the corresponding phosphoramidite **212**. Afterwards, the globally deprotected peptide **213** was ligated to a recombinant Smad2 fragment **214** by EPL. The resulting caged phospho-Smad2 protein **215** was subjected to biochemical studies, which showed that the caged phospho-Smad2 protein mimicked the characteristics of unphosphorylated Smad2 and formed a complex with the SARA protein. The caging group could be removed by exposure to UV light to yield active phosphorylated Smad2.

In a second study, the caging approach for phosphor-Smad2 was extended by generating a C-terminal-phosphorylated Smad2 (**217**) by ligation with **216**, which contained a fluorophore and a photocleavable quencher moiety (Scheme 38).[304] The fluoresence quencher fulfilled two tasks: on the one hand it efficiently suppresses the fluorescence of the fluorophor and on the other hand it conceals the presence of the two phosphate groups on the C-terminal Ser



Scheme 37. Synthesis of Smad2 proteins with photolabile groups.

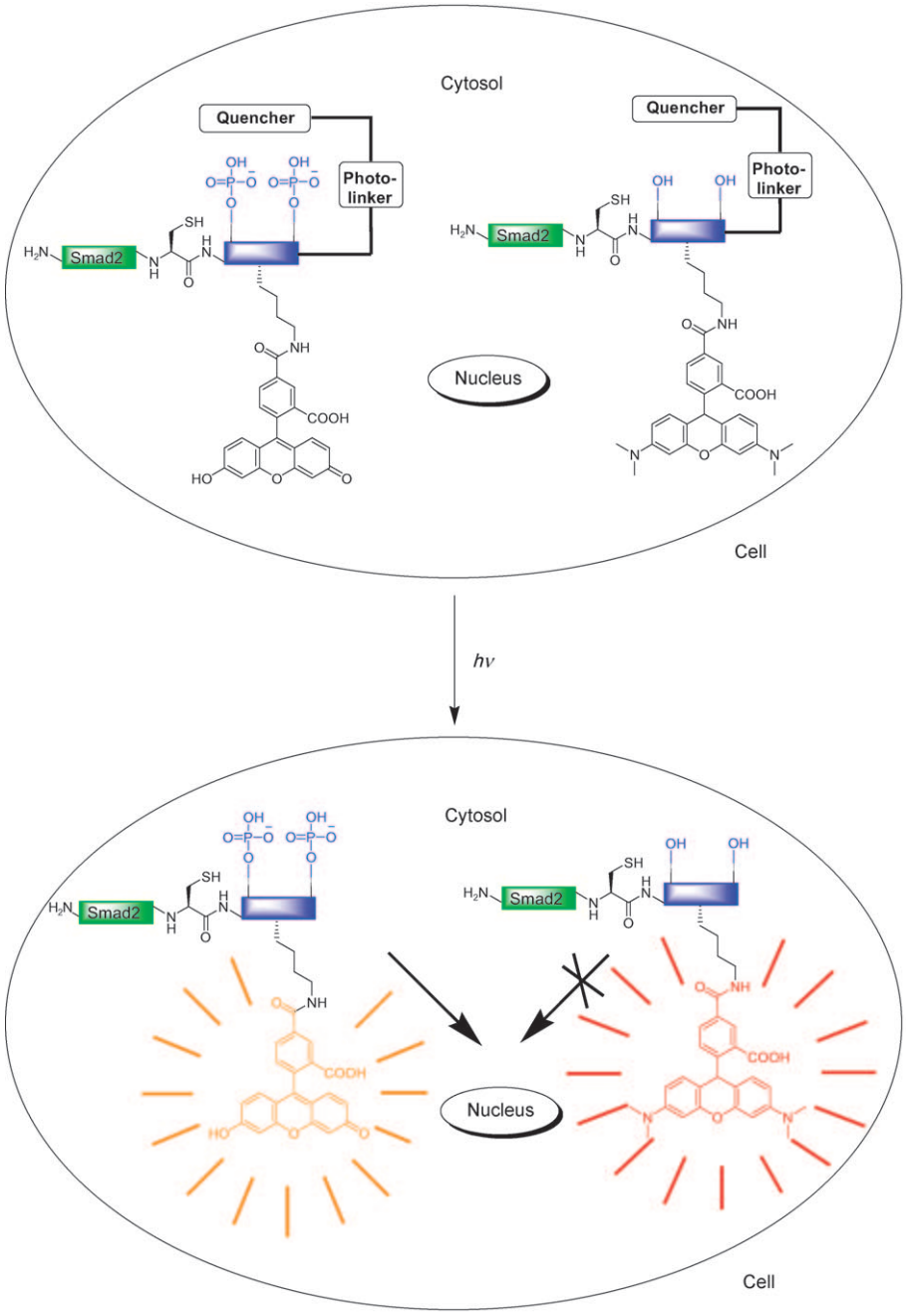


Scheme 38. Synthesis of fluorescent Smad2 proteins.

residues (just like the individually installed NEP caging groups). As a result, the uncaging of Smad2 led to two simultaneous effects: release of the two phosphorserine residues and the emission of a fluorescence signal. This extended caging approach is particular useful for comparing the behavior of modified and unmodified proteins, as demonstrated by comparison studies with a non-phosphorylated Smad2 analogue (Scheme 39). The non-phosphorylated protein contained a fluorophore which emitted at a different wavelength. A mixture of both proteins was introduced into living cells and irradiated, what resulted in the simultaneous uncaging of both proteins, as indicated by fluorescence from the two different fluorophores. As expected, the uncaged phospho-Smad2 protein migrated into the nucleus, whereas the nonphosphorylated protein maintained its cytosolic localization.[304]

The investigation of protein phosphorylation can be hindered by the dynamic character of this modification. The balance between the phosphorylated and unphosphorylated state is maintained by the corresponding kinases and protein phosphatases.[298] Both kinds of proteins are commonly phosphorylated themselves. Investigations of the phosphorylation of phosphatases can be especially difficult, because these enzymes are often capable of auto-dephosphorylation.[305] To overcome this problem, protein semisynthesis has successfully been applied to introduce noncleavable mimics of phosphates at the modification sites (Scheme 40 A).[306,307] One target protein is the low-molecular-weight protein tyrosine phosphatase (LMW-PTP), which is involved in signal transduction processes mediated by the platelet-derived growth factor receptor (PDGFr).[308] LMW-PTP is known to be phosphorylated at two adjacent Tyr residues located in proximity to the active site. Protein semisynthesis with the EPL strategy was used to install phosphonomethylenephenylalanine (**219**, Pmp), a non-hydrolyzable mimetic of phosphotyrosine **218**, at these positions in **220** (Scheme 40 A and B).[309]

The resulting irreversibly phosphorylated proteins were subjected to biochemical and cell-biological assays, which showed that the installed Pmp moieties lowered the phosphatase activity, thus indicating that this modification is a means to down-regulate the enzymatic activity.

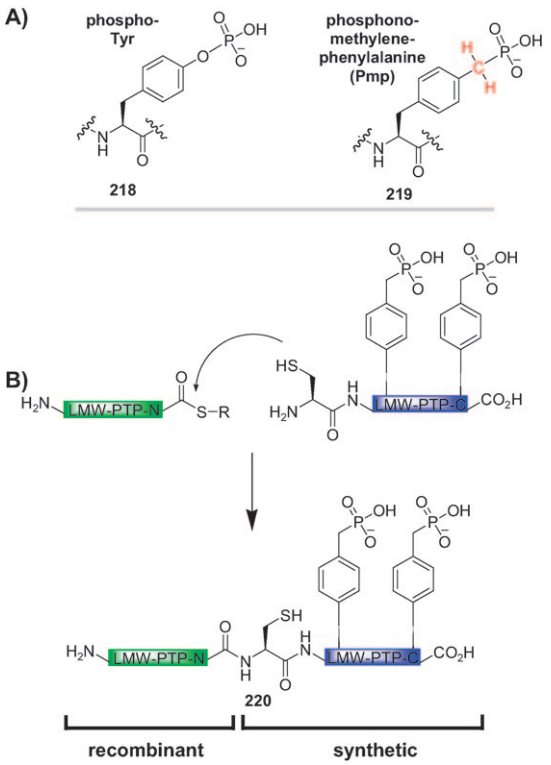


Scheme 39. Functional studies on phosphorylated Smad2 proteins in living cells.

6.3. Protein Lipidation

The attachment of lipid moieties is another pivotal modification of proteins which commonly mediates association with cellular membranes. Prominent examples of lipidated proteins are the small GTPases of the Ras family,[310] which are important regulators of several cellular processes. H-Ras and N-Ras are two isoforms of Ras proteins which are farnesylated and mono- or dipalmitoylated at their C termini (structure **182** in Scheme 31).[310] Whereas the farnesylation (Fa) represents an irreversible modification (thioether bond), the palmitoyl (Pal) groups is linked through a thioester bond, which can later be cleaved (Scheme 31 B). As a result, this modification is commonly lost during purification procedures from natural sources. The MIC ligation strategy (see Section 4.1.3) has been used to overcome this problem and generate homogeneous, lipidated Ras proteins for biochemical and biophysical investigations.[311] A truncated version of the Ras containing a solvent-exposed C-terminal Cys residue was produced in bacteria. Subsequently, MIC-modified peptides resembling the C-terminal tail of Ras with different lipidation patterns were synthesized and conjugated to the protein.[312] The resulting Ras proteins **183** were subjected to biophysical studies and cell biological assays, and it was found that, despite the unnatural linkage, the proteins displayed full biological activity. A further advantage of this approach was the ability to replace the cleavable palmitoyl thioester with a non-cleavable thioether-linked hexadecyl (HD) group (see Scheme 31 B). This irreversibly palmitoylated Ras protein was used to study the cellular trafficking of Ras proteins.[313]

It was long believed that Ras proteins are only active when conjugated to the cell's plasma membrane (PM) in which the attachment process is mediated by the lipid moieties. However, recent studies have demonstrated that a pool of Ras proteins located at the Golgi apparatus might represent an independent signaling platform for cellular processes. Ras proteins access the plasma membrane through the secretory pathway and can return to the Golgi apparatus after removal of the palmitoyl group. The reinstallation of the palmitoyl moiety at the Golgi apparatus redirects the Ras proteins to the plasma membrane, thereby resulting in a de/reacylation cycle that maintains both the PM and the Golgi pool. Detailed investigations that uncovered this Ras cycling employed MIC-generated Ras proteins modified with the thioether-linked HD group instead of the cleavable palmitoyl group (Scheme 31 B). The Ras protein modified with this palmitoyl analogue did not participate in the de/reacylation cycle because of the irreversible nature of this lipidation.[313]



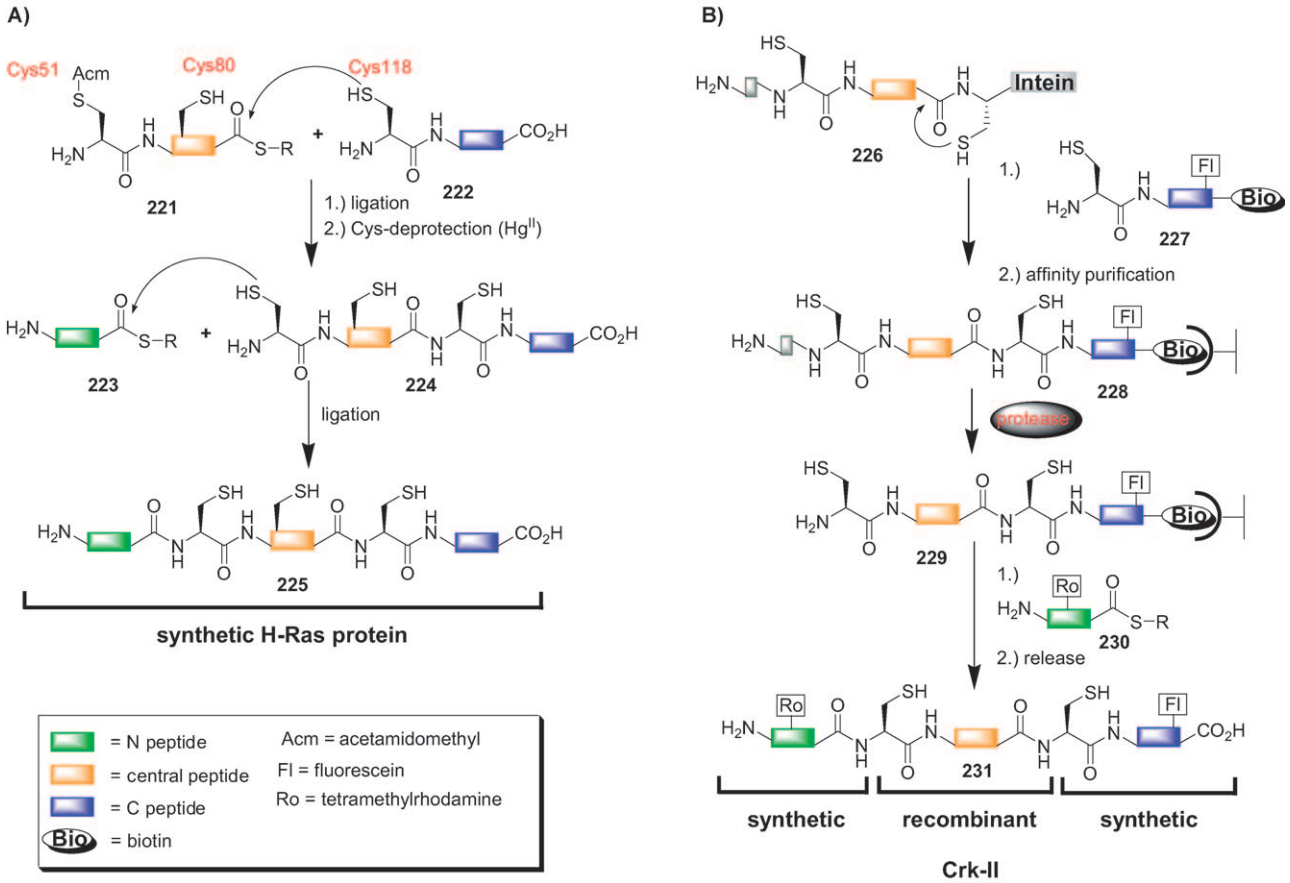
Scheme 40. A) Phosphotyrosine (**218**) and Pmp. B) Semisynthesis of LMW-PTP.

7. Strategies for the Ligation of Multiple Fragments

As discussed previously, chemoselective ligation techniques have expanded the size limitation of SPPS significantly. However, the products formed by the ligation of two segments are still smaller than the majority of natural proteins.[314,315] Although protein semisynthesis can be used to introduce synthetic peptides into proteins, chemical control is only gained over the residues of the small synthetic segments. Therefore, synthetic strategies have been developed to enable the total chemical synthesis of proteins from multiple segments. Additional protection group strategies are used to ensure selective coupling of segments and to avoid side reactions.

7.1. Synthetic H-Ras

A three-segment ligation strategy was applied by Engelhard et al. for the total synthesis of protooncogene H-Ras **225**, which comprises 166 amino acids.[316] This protein contains three natural Cys residues (at positions 51, 80, and 118) which can be used as ligation sites without changing the amino acid composition of the protein. The synthesis strategy used two of these Cys residues, and started with the ligation of the C fragment **221** containing Cys118 at the N terminus and the central fragment **222** comprising a C-terminal thioester (Scheme 41). The N-terminal Cys51 residue of the central



Scheme 41. Strategies for the ligation of multiple fragments: A) synthetic H-Ras, B) solid-phase expressed protein ligation.

fragment represented another natural residue of H-Ras and was protected with the acetamidomethyl (Acm) group. Orthogonal deprotection of the Acm group in **224** was achieved by treatment with Hg^{II} after the ligation.[317] The N-terminal fragment **223** was then attached, which resulted in the fully synthetic H-Ras protein. Subsequent refolding and biochemical analysis revealed that the catalytic properties of the synthetic H-Ras were indistinguishable from biologically produced material.

7.2. Synthetic EPO

A four-segment ligation in combination with various additional modification processes (see Section 3.2) has recently been used for the synthesis of a PEGylated analogue of the glycoprotein hormone erythropoietin (EPO; Scheme 42).[119,318] The natural glycoprotein EPO stimulates the proliferation of erythroid cells. Two of the four peptide segments were modified with branched, negatively charged polymers of a defined length by oxime ligation. These polymers replace glycosylations, to overcome the problem of inhomogeneous glycosylation patterns often encountered in natural glycosylated proteins. The synthetic EPO was assembled from the C to the N fragment by applying the Acm-protection strategy discussed above. As the natural EPO sequence does not incorporate Cys residues at all junction sites, the necessary additional Cys mutations were alkylated with bromoacetic acid (**76**), thereby converting them into an analogue of glutamic acid, which naturally occupies these positions (see Section 3.2.2.1). The biological activity of the synthetic EPO was similar to wildtype EPO; however, the synthetic PEGylated hormone maintained a two- to threefold higher plasma level for several hours when injected into rats.[319]

7.3. Solid-Phase-Expressed Protein Ligation (SPPL)

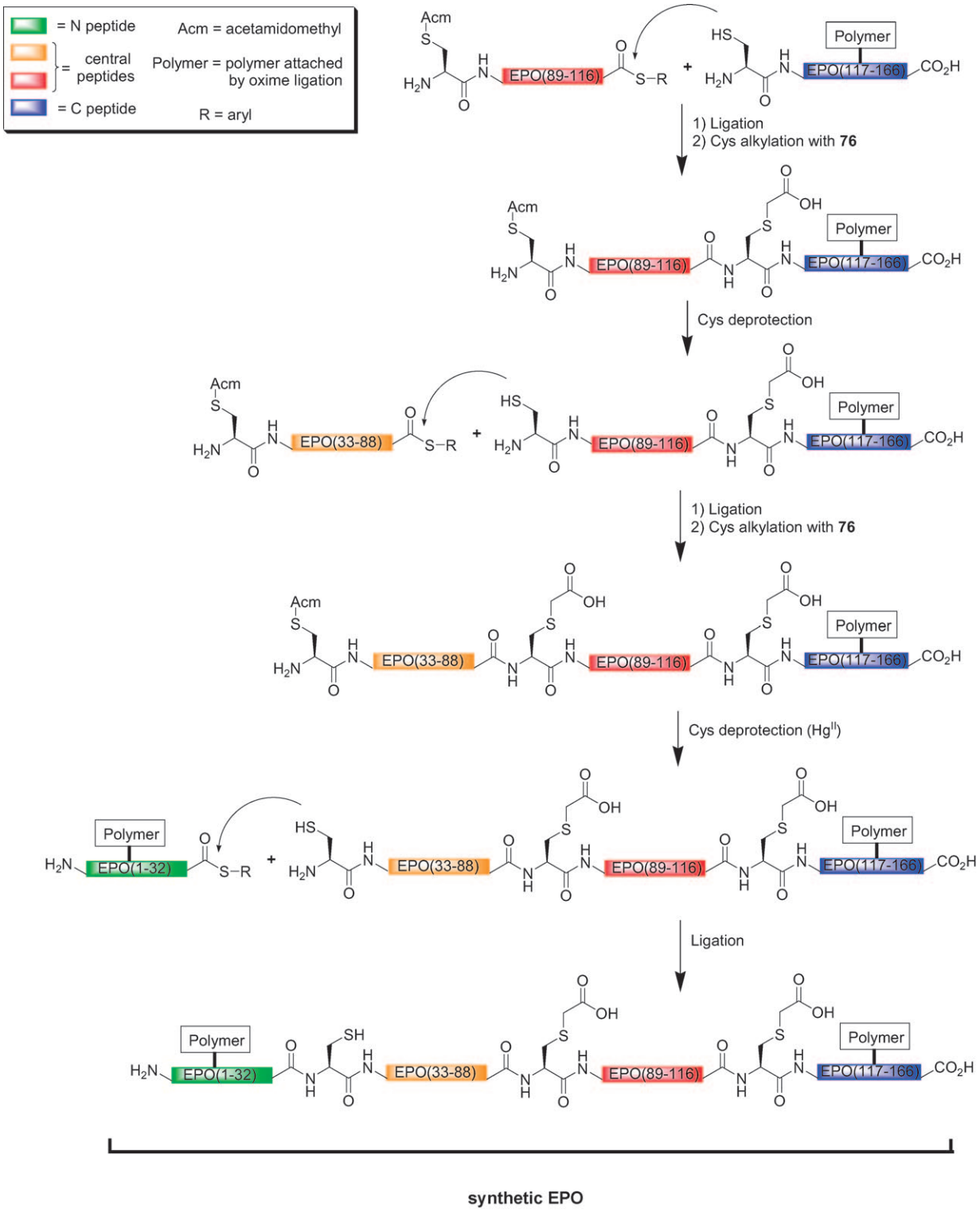
Multiple fragment ligations can also involve recombinant peptides, as demonstrated by Cotton and Muir for the synthesis of a doubly fluorescently labeled Crk-II protein by using NCL and ELP.[320] Crk-II is an adapter protein involved in several signal transduction pathways and composed of one Src-homology domain 2 (SH2) and two Src-homology domains 3 (SH3). The protein is regulated by Tyr phosphorylation, which results in an intramolecular rearrangement caused by the association of the phosphorylated Tyr residue with the SH2 domain. An appropriate FRET pair was ligated to the C terminus and the N terminus of Crk-II to enable this rearrangement to be monitored by fluorescence resonance energy transfer (FRET; Scheme 41 B). The central segment of this three-fragment assembly was produced in bacteria as an intein fusion protein **226**. The C-terminal thioester of the central segment was introduced by EPL and subsequently ligated with the C-terminal peptide **227** containing a fluorescein dye. The necessary protection of the N-terminal Cys residue was achieved by embedding the residue into the recognition site for a protease factor Xa (Scheme 41 B). The ligation product **228** was subsequently treated with the protease to liberate the N-terminal Cys protein **229**, followed by subsequent ligation with an N-terminal tetramethylrhodamine-conjugated peptide thioester **230**. An advantage of this synthetic approach was the assembly on a resin, facilitated by a biotin moiety within the C peptide **227**. Thus, the synthesis strategy resembled the concept of SPPS and was named "solid-phase expressed protein ligation" (SPPL). After release from the resin, biochemical assays of the doubly labeled Crk-II **231** showed the expected FRET change induced by phophorylation, thus confirming its applicability as a reporter for the phosphorylation state of Crk-II.

7.4. Synthetic Ribonuclease A

Besides the most common NCL and EPL approaches, alternative strategies have also been used for the ligation of multiple fragments. The most prominent example is the enzymatic synthesis of fully functional ribonuclease A from six synthetic fragments by the Subtiligase (see also Section 5.1).[321] As in the case of NCL, the N terminus of the peptide fragments needs to be protected to prevent self-ligations. The protection was realized by installing the isonicotinyl (iNOC) group at the α-amine group; this protecting group can be selectively removed under mild reducing conditions. The C terminus of the peptide fragments was esterified with a glycolate-phenylalanine amide moiety to allow efficient alkylation of the Subtiligase. Ribonuclease A was assembled from the C terminal to the N terminal fragment, catalyzed by Subtiligase, followed by cleavage of the iNOC group after each ligation step. The synthetic ribonuclease showed similar activity as recombinant RNase A and the synthetic enzymes generated with other chemoselective ligation techniques. At this point it should be noted that ribonuclease A is a classic target for protein synthesis approaches. It was the first enzyme to be fully synthesized by SPPS in the late 1960s.[322] Synthetic approaches for the installation of native glycosylation patterns have been reported.[41] A total synthesis of ribonuclease A from six fragments by using NCL has also been described.[323]

7.5. Combined Recombinant, Enzymatic, and Chemical Synthesis Strategy (CRECS)

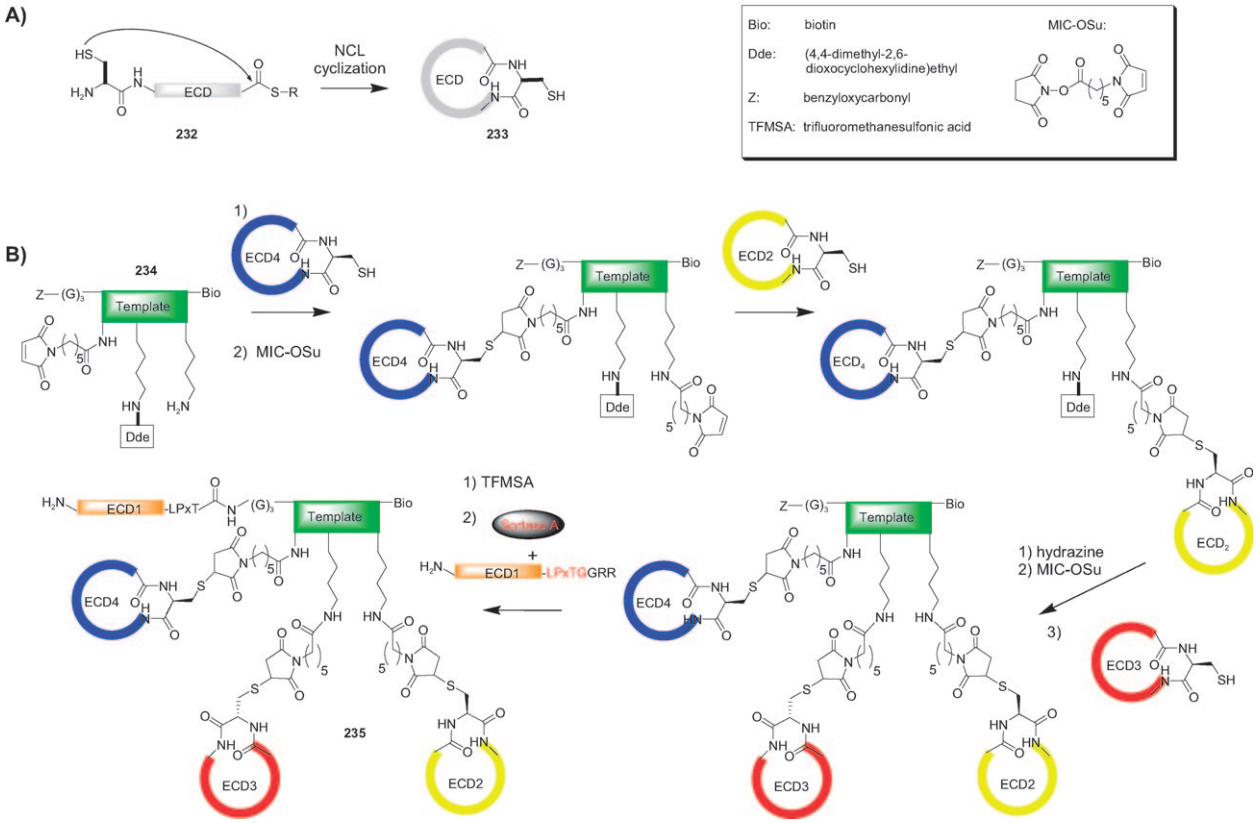
In a very recent publication Beyermann et al. reported the successful application of several ligation techniques discussed in this Review for the synthesis of a protein mimicking the extracellular region of a G-protein-coupled receptor (GPCR, Scheme 43).[324] GPCRs bind their ligands through extracellular binding domains (ECDs) consisting of an N-terminal region (ECD 1) and three loops (ECD 2–4) which link the seven transmembrane helices of the receptors. In the case of the corticotrophin-releasing factor receptor type 1 (CRF1), the natural peptide agonist urocortin 1 binds to the soluble N-terminal ECD1 only.[325] In contrast, another known peptide agonist, sauvagin, requires the whole CRF1 receptor to display high affinity binding, thereby indicating that the three



Scheme 42. Total chemical synthesis of EPO.

loops (ECD 2–4) play an important role in binding this ligand.[326]

To address this question, the binding domains ECD 1–4 were subsequently ligated onto a synthetic peptide template,[327] which served as a scaffold for the spatial arrangement of the extracellular parts of CRF1. The template was constructed from a branched peptide containing a MIC group (see Section 4.3.1), a Dde-protected Lys, an unprotected Lys, and finally a triglycine tail to facilitate ligation reactions catalyzed by sortase A (see Section 5.2). Additional residues of the template were used for spacing the functional groups and to mediate solubility of the construct. Synthetic cyclic



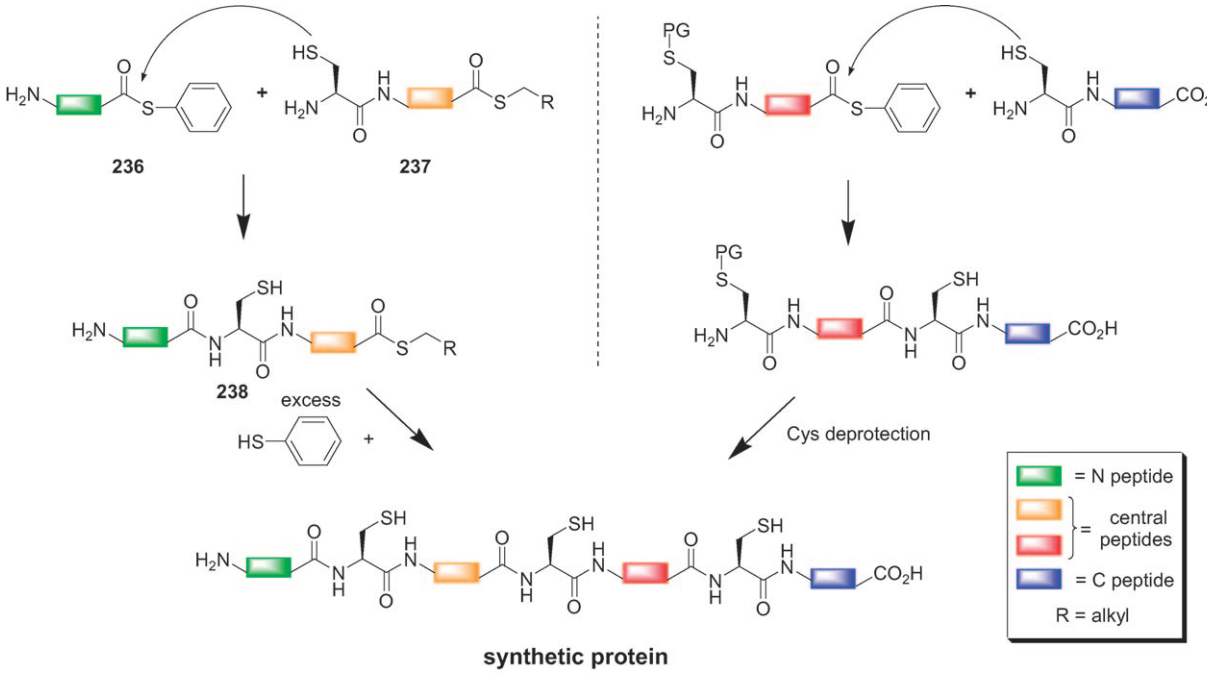
Scheme 43. A) Intramolecular NCL for EDC domain synthesis. B) Synthetic strategy for generating synthetic synthetic CRF1.

peptides **233** derived from the three loop regions of the CRF1 receptor (ECD 2–4) were obtained by intramolecular NCL of linear precursors **232** (Scheme 43 A). First, the unique Cys residue of the cyclic ECD4 was conjugated to the template **234** by the MIC strategy (Section 4.3.1). In an analogous manner, the second and third domains, ECD2 and ECD3, were conjugated after installation of additional MIC groups, first by reaction of MIC-OSu with the unprotected Lys residue and second in the same manner after deprotection of the Dde group with hydrazine. In parallel, the CRF1 N terminus (ECD1) was produced in *E. coli*. The ECD1 domain, which was equipped with a C-terminal sortase A sorting motive, was obtained recombinantly and ligated to the triglycine tail by sortase A (see Section 5.2) to finally yield a synthetic protein mimic **235** of the extracellular fraction of CRF1 (Scheme 43 B). Subsequent binding studies were performed with the synthetic CRF1 receptor and the peptide agonist sauvagin. Only the synthetic CRF1 mimic was able to bind this ligand, unlike ECD1 alone and a control mimic lacking ECD1, therefore demonstrating that the loops contribute to the binding of CRF1 with sauvagin. These reseults show that this approach, which is referred to as the *c*ombined *r*ecombinant, *e*nzymatic, and *c*hemical *s*ynthesis strategy (CRECS strategy), can generate a functional synthetic GPCR that can be used to conduct ligand-binding studies with natural agonists.[324]

7.6. Convergent Strategies for Multiple Fragment Ligations

All the strategies mentioned above require the stepwise assembly of the individual fragments from the C to the N terminus. However, the convergent assembly of multiple fragments in parallel represents a more efficient way for generating whole synthetic proteins. A convergent synthesis approach would require double protection of the central fragment at the N-terminal Cys as well as the C-terminal thioester. In this regard, Danishefsky and co-workers introduced phenolic esters equipped with an adjacent disulfide-protected thiol, which are much less reactive than the corresponding thioesters.[53] The phenolic ester can be converted into a thioester by the addition of an excess of thiol, which reduces the disulfide and thereby triggers an O→S migration (see Schemes 4 E and 15).

An alternative strategy, called the "kinetically controlled ligation", principle was introduced by Kent and co-workers.[328] This approach exploits the different reactivities of aryl and alkyl thioesters for a convergent protein syntheses.[329] Alkyl thioesters are sufficiently unreactive and do not participate in ligation reactions when competing aryl thioesters are present. They even tolerate unprotected N-terminal Cys residues without significant amounts of self-ligation. However, they can be converted into reactive aryl thioesters in a thiol–thioester exchange reaction by adding an excess of aryl thiols. In kinetically controlled ligation procedures, the central peptide fragment **237** contains an N-terminal Cys



Scheme 44. Convergent ligation of multiple fragments by kinetically controlled ligation.

residue and a C-terminal alkyl thioester. This fragment is ligated at the N-terminal Cys residue with aryl thioesters **236**. The C-terminal alkyl thioester of the ligated peptide **238** is subsequently converted into an aryl thioester by thioester exchange to participate in the next ligation reaction (Scheme 44). This methodology allows the ligation of multiple fragments in parallel, thus enabling a fully convergent synthesis of proteins. The feasibility of this approach has been demonstrated for the convergent synthesis of crambin and lysozyme (Scheme 44).[327,330] In the latter case, the lysozyme, which consists of 130 amino acids, was synthesized from four fragments. The folded synthetic lysozyme had full enzymatic activity and the obtained X-ray structure was identical to that of the biologically produced lysozyme.

8. Conclusion

In summary, the repertoire of chemoselective strategies for the synthesis and/or modification of proteins and the impressive functional studies have greatly advanced chemical biology, as these methods made it possible to study the functional consequences of molecular changes within a protein even in a cellular environment. It is, therefore, both the quality and the sheer number of these chemical methods (or in other words, a well-equipped tool box) that represents a true breakthrough in modern synthetic chemistry. Although many of these reactions are very much applicable, chemists and biologists are still devoted to developing further chemoselective methods for the synthesis of modified proteins. In this respect, particular emphasis has been placed on simplifying the accessibility of starting materials and achieving mild reaction conditions that are even transferrable to in vivo applications.

The large number of various ligation techniques reported to date in combination with convergent synthesis schemes provides all the key elements for the total chemical synthesis of natural or fully unnatural proteins with yet to be discovered properties. It is likely that in the near future completely synthetic proteins will be far more than scientific proof of principle constructs, they will become important tools for all sorts of biochemical and biological research.[331]

We acknowledge financial support from the German Science Foundation (DFG) within the Emmy-Noether programm (HA 4468/2-1 and SCHW 1163/3-1) and the SFB 765, the BMBF, as well as the Fonds der Chemischen Industrie (FCI). Furthermore, Dr. Constantin Czekelius, Dr. Karl-Christian Mittermüller, Dr. Remigiusz Serwa, Dr. Verena Böhrsch, and Michaela Mühlberg are thanked for careful proofreading and helpful discussions during the preparation of the manuscript. We thank Prof. Herbert Waldmann for providing a picture of the Ras protein in Scheme 31 A. Nico Rind is thanked for preparation of the frontispiece.

Received: March 18, 2008

[1] J. W. von Goethe, *Goethes Werke*, Weimarer Ausgabe, I. Abtheilung, 25. Band, p. 89.
[2] a) E. Fischer, E. Fourneau, *Ber. Dtsch. Chem. Ges.* **1901**, *34*, 2868–2879; b) for an excellent historical overview on peptide synthesis and ligation strategies, see T. Kimmerlin, D. Seebach, *J. Pept. Res.* **2005**, *65*, 229–260.
[3] a) C. T. Walsh in *Posttranslational Modifications of Proteins*, Roberts, Colorado, **2003**; b) B. G. Davis, *Science* **2004**, *303*, 480–482; c) F. Wold, *Annu. Rev. Biochem.* **1981**, *50*, 783–814; d) C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, Jr., *Angew. Chem.* **2005**, *117*, 7508–7539; *Angew. Chem. Int. Ed.* **2005**, *44*,

7342 – 7372; e) D. P. Gamblin, S. I. van Kasteren, J. M. Calker, B. G. Davis, *Eur. J. Biochem.* **2000**, *275*, 1949 – 1959; f) J.-P. Pellois, T. W. Muir, *Curr. Opin. Chem. Biol.* **2006**, *10*, 487 – 491.

[4] a) R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149 – 2154; b) Nobel lecture: R. B. Merrifield, *Angew. Chem.* **1985**, *97*, 801 – 812; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 799 – 810.

[5] a) M. Bodanszky in *Principles of Peptide Synthesis*, Springer, Berlin, **1984**; b) N. Seewald, H.-D. Jakubke in *Peptides: Chemistry and Biology*, Wiley-VCH, Weinheim, **2002**.

[6] H. Hojo, Y. Matsumoto, Y. Nakahara, E. Ito, Y. Suzuki, M. Suzuki, A. Suzuki, Y. Nakahara, *J. Am. Chem. Soc.* **2005**, *127*, 13720 – 13725.

[7] The term C and N peptide is used in accordance with the terminology intrduced by Tam, see also J. P. Tam, J. Xu, K. D. Eom, *Biopolymers* **2001**, *60*, 194 – 205.

[8] For a recent review on protecting-group-free synthesis in the total synthesis of natural products, see R. W. Hoffmann, *Synthesis* **2006**, 3531 – 3541.

[9] a) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249 – 289; b) J.-P. Pellois, T. W. Muir, *Curr. Opin. Chem. Biol.* **2006**, *10*, 487 – 491; c) T. Durek, C. F. W. Becker, *Biomol. Eng.* **2005**, *22*, 153 – 172; d) B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91 – 118; e) O. Seitz, *Org. Synth. Highl. V* **2003**, *5*, 368 – 383; f) D. Rauh, H. Waldmann, *Angew. Chem.* **2007**, *119*, 840 – 844; *Angew. Chem. Int. Ed.* **2007**, *46*, 826 – 829; g) C. Haase, O. Seitz, *Angew. Chem.* **2008**, *120*, 1575 – 1579; *Angew. Chem. Int. Ed.* **2008**, *47*, 1553 – 1556; h) J. Rademann, *Angew. Chem.* **2004**, *116*, 4654 – 4656; *Angew. Chem. Int. Ed.* **2004**, *43*, 4554 – 4556.

[10] For a recent review concerning new strategies for the formation of amide bonds, see J. W. Bode, *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 765 – 775.

[11] T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, *Justus Liebigs Ann. Chem.* **1953**, *583*, 129 – 149.

[12] a) D. S. Kemp, S. L. Leung, D. J. Kerkman, *Tetrahedron Lett.* **1981**, *22*, 181 – 184; b) D. S. Kemp, D. J. Kerkman, *Tetrahedron Lett.* **1981**, *22*, 185 – 186; c) D. S. Kemp, N. G. Galakatos, *J. Org. Chem.* **1986**, *51*, 1821 – 1829; d) D. M. Coltart, *Tetrahedron* **2000**, *56*, 3449 – 3491.

[13] This method represents a extension of the previously reported "prior amine capture": a) D. S. Kemp, J. A. Grattan, J. Reczekm, *J. Org. Chem.* **1975**, *40*, 3465 – 3466; b) D. S. Kemp, D. J. Kerkman, S.-L. Leung, G. Hanson, *J. Org. Chem.* **1981**, *46*, 490 – 498; c) D. S. Kemp, *Biopolymers* **1981**, *20*, 1793 – 1804; d) for a more recent extension of the amine capture strategy with quinolinium thioester salts: S. Leleu, M. Penhoat, A. Bouet, G. Dupas, C. Papamicaël, F. Marsais, V. Levacher, *J. Am. Chem. Soc.* **2005**, *127*, 15668 – 15669.

[14] D. S. Kemp, N. G. Galakatos, B. Bowen, K. Tam, *J. Org. Chem.* **1986**, *51*, 1829 – 1843.

[15] a) N. Fotouhi, N. G. Galakatos, D. S. Kemp, *J. Org. Chem.* **1989**, *54*, 2803 – 2817; b) D. S. Kemp, R. I. Carry, *Tetrahedron Lett.* **1991**, *32*, 2845 – 2848.

[16] D. S. Kemp, R. I. Lary, *J. Org. Chem.* **1993**, *58*, 2216 – 2222.

[17] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776 – 779.

[18] N→S shifts were reported under acidic conditions for the synthesis of (glyco-)peptide thioesters: F. Nagaike, Y. Onuma, C. Kanazawa, H. Hojo, A. Ueki, Y. Nakahara, Y. Nakahara, *Org. Lett.* **2006**, *8*, 4465 – 4468. For an additional example see Section 3.3.4. These findings indicate that the S→N shift in the NCL leading to the amide bond cannot be generally regarded as fully irreversible. However, the rapid kinetics of the rearrangement result in a fast adjustment of the equilibrium and thereby formation of the thermodynamically favored amide. In the literature the S→N acyl transfer is commonly regarded to as irreversible, which we will continue to do for consistency.

[19] For selected examples see also Section 6.

[20] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960; see also Sections 6 and 7.

[21] J. A. Camarero, A. Shekhtman, E. A. Campbell, M. Chlenov, T. M. Gruber, D. A. Bryant, S. A. Darst, D. Cowburn, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8536 – 8541.

[22] D. Cowburn, T. W. Muir, *Methods Enzymol.* **2001**, *339*, 41 – 54.

[23] C. Dose, O. Seitz, *Org. Lett.* **2005**, *7*, 4365 – 4368.

[24] S. Ficht S, C. Dose, O. Seitz, *ChemBioChem* **2005**, *6*, 2098 – 2103.

[25] M. Lovrinovic, C. M. Niemeyer, *ChemBioChem* **2007**, *8*, 61 – 67.

[26] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068 – 10073.

[27] P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325 – 4329.

[28] For a systematic study, see E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640 – 6646.

[29] For a recent application in kinetically controlled segment ligations, see Section 7. Several NCL-based ligation protocols also include the use of tris(2-carboxyethyl)phosphine hydrochloride (TCEP) as a thiol-free reducing agent.

[30] F. I. Valiyaveetil, R. MacKinnon, T. W. Muir, *J. Am. Chem. Soc.* **2002**, *124*, 9113 – 9120.

[31] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525.

[32] K. Severin, D. H. Lee, J. A. Martinez, M. R. Ghadiri, *Chem. Eur. J.* **1997**, *3*, 1017 – 1024; for a review see R. Issac, Y.-W. Ham, J. Chmielewski, *Curr. Opin. Struct. Biol.* **2001**, *11*, 458 – 463.

[33] K. Severin, D. H. Lee, A. J. Kennan, M. R. Ghadiri, *Nature* **1997**, *389*, 706 – 709.

[34] A. Saghatelian, Y. Yokobayashi, K. Soltani, M. R. Ghadiri, *Nature* **2001**, *409*, 797 – 801.

[35] For recent example in which DNA-templated NCL is employed, see C. Dose, S. Ficht, O. Seitz, *Angew. Chem.* **2006**, *118*, 5495 – 5499; *Angew. Chem. Int. Ed.* **2006**, *45*, 5369 – 5373, and references herein.

[36] For a recent example of a NCL self-replication system employing fluorinated peptides, see C. Jäckel, M. Salwiczek, B. Koksch, *Angew. Chem.* **2006**, *118*, 4305 – 4309; *Angew. Chem. Int. Ed.* **2006**, *45*, 4198 – 4203.

[37] For a recent review, see J. A. Camarero, A. R. Mitchell, *Protein Pept. Lett.* **2005**, *12*, 723 – 728.

[38] L. Zhang, J. P. Tam, *J. Am. Chem. Soc.* **1999**, *121*, 3311 – 3320.

[39] L. E. Canne, S. M. Walker, S. B. Kent, *Tetrahedron Lett.* **1995**, *36*, 1217 – 1220.

[40] For Fmoc deprotection protocols with non-nucleophilic bases to replace commonly employed piperidine, see a) X. Li, T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **1998**, *39*, 8669 – 8672; b) A. B. Clippingdale, C. J. Barrow, J. D. Wade, *J. Pept. Sci.* **2000**, *6*, 225 – 234.

[41] For a recent protocol for the synthesis of glycosylated thioesters based on a double linker resin strategy, see S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chem.* **2005**, *117*, 1677 – 1681; *Angew. Chem. Int. Ed.* **2005**, *44*, 1650 – 1654.

[42] G. W. Kenner, J. R. McDermott, R. C. Sheppard, *Chem. Commun.* **1971**, 636 – 637.

[43] R. Ingenito, E. Bianchi, D. Fattori, A. Pessi, *J. Am. Chem. Soc.* **1999**, *121*, 11369 – 11374.

[44] a) B. J. Backes, A. A. Virgilio, J. A. Ellman, *J. Am. Chem. Soc.* **1996**, *118*, 3055 – 3056; b) B. J. Backes, J. A. Ellman, *J. Org. Chem.* **1999**, *64*, 2322 – 2330.

[45] Y. He, J. P. Wilkins, L. L. Kiessling, *Org. Lett.* **2006**, *8*, 2483 – 2485.

[46] Y. Shin, K. A. Winans, B. J. Backes, S. B. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc.* **1999**, *121*, 11684 – 11698; for a recent self-purifcation strategy of thioesters with a safety-catch linker, see F. Mende, O. Seitz, *Angew. Chem.* **2007**, *119*, 4661 – 4665; *Angew. Chem. Int. Ed.* **2007**, *46*, 4577 – 4580.

[47] a) A. Sewing, D. Hilvert, *Angew. Chem.* **2001**, *113*, 3503 – 3505; *Angew. Chem. Int. Ed.* **2001**, *40*, 3395 – 3396; ; b) D. Swinnen, D. Hilvert, *Org. Lett.* **2000**, *2*, 2439 – 2442.

[48] A. R. Mezo, R. P. Cheng, B. Imperiali, *J. Am. Chem. Soc.* **2001**, *123*, 3885 – 3891.

[49] a) S. Futaki, K. Sogawa, J. Maruyama, T. Asahara, M. Niwa, *Tetrahedron Lett.* **1997**, *38*, 6237 – 6240; b) S. Biancalana, D. Hudson, M. F. Songster, S. A. Thompson, *Lett. Pept. Sci.* **2001**, *7*, 291 – 297; c) R. von Eggelkraut-Gottanka, A. Klose, A. G. Beck-Sickinger, M. Beyermann, *Tetrahedron Lett.* **2003**, *44*, 3551 – 3554.

[50] C. P. R. Hackenberger, *Org. Biomol. Chem.* **2006**, *4*, 2291 – 2295.

[51] a) Y. Kajihara, A. Yoshihara, K. Hiranoa, N. Yamamoto, *Carbohydr. Res.* **2006**, *341*, 1333 – 1340; b) A. C. Nagalingam, S. E. Radford, S. L. Warriner, *Synlett* **2007**, 2517 – 2520.

[52] P. Botti, M. Villain, S. Manganiello, H. Gaertner, *Org. Lett.* **2004**, *6*, 4861 – 4864.

[53] J. D. Warren, J. S. Miller, S. J. Keding, S. J. Danishefsky, *J. Am. Chem. Soc.* **2004**, *126*, 6576 – 6578.

[54] A. P. Tofteng, K. J. Jensen, T. Hoeg-Jensen, *Tetrahedron Lett.* **2007**, *48*, 2105 – 2107.

[55] B. Wu, J. D. Warren, J. Chen, G. Chen, Z. Hua, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 5219 – 5223.

[56] For a conceptually similar strategy in which a Cys-Pro ester is used as an autoactivation unit, see T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2007**, *48*, 1903 – 1905.

[57] J. Chen, J. D. Warren, B. Wu, G. Chen, Q. Wan, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 1969 – 1972.

[58] a) B. Wu, Z. Tan, G. Chen, J. Chen, Z. Hua, Q. Wan, K. Ranganathan, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 8009 – 8011; b) J. Chen, G. Chen, B. Wu, Q. Wan, Z. Tan, Z. Hua, S. J. Danishefsky, *Tetrahedron Lett.* **2006**, *47*, 8013 – 8016.

[59] For a pioneering report of this strategy in the synthesis of a non-native HIV protease with a thioester linkage in the backbone, see M. Schnoelzer, S. B. H. Kent, *Science* **1992**, *256*, 221 – 225.

[60] J. P. Tam, Y. A. Lu, C. F. Liu, J. Shao, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485 – 12489.

[61] C. F. Liu, C. Rau, J. P. Tam, *Tetrahedron Lett.* **1996**, *37*, 933 – 936.

[62] R. E. Offord, *Nature* **1969**, *221*, 37 – 40.

[63] F. Borrás, R. E. Offord, *Nature* **1970**, *227*, 716 – 718.

[64] R. M. Jones, K. Rose, R. E. Offord, *Biochem. J.* **1987**, *247*, 785 – 788.

[65] D. A. Erlanson, M. Chytil, G. L. Verdine, *Chem. Biol.* **1996**, *3*, 981 – 991.

[66] W. Zheng, Z. Zhang, S. Ganguly, J. L. Weller, D. C. Klein, P. A. Cole, *Nat. Struct. Biol.* **2003**, *10*, 1054 – 1057.

[67] A. Romanelli, A. Shekhtman, D. Cowburn, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6397 – 6402.

[68] M. A. Shogren-Knaak, C. L. Peterson, *Methods Enzymol.* **2003**, *375*, 62 – 76.

[69] T. J. Tolbert, D. Franke, C. H. Wong, *Bioorg. Med. Chem.* **2005**, *13*, 909 – 915.

[70] T. J. Tolbert, C.-H. Wong, *Angew. Chem.* **2002**, *114*, 2275 – 2278; *Angew. Chem. Int. Ed.* **2002**, *41*, 2171 – 2174.

[71] a) C. P. R. Hackenberger, M. Chen, B. Imperiali, *Bioorg. Med. Chem.* **2006**, *14*, 5043 – 5048; b) S. Mathys, T. C. Evans, Jr., I. C. Chute, H. Wu, S. Chong, J. Benner, X.-Q. J. Liu, M.-Q. Xu, *Gene* **1999**, *231*, 1 – 13.

[72] D. Macmillan, L. Arham, *J. Am. Chem. Soc.* **2004**, *126*, 9530 – 9531.

[73] I. E. Gentle, D. P. De Souza, M. Baca, *Bioconjugate Chem.* **2004**, *15*, 658 – 663.

[74] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705 – 6710.

[75] T. C. Evans, Jr., J. Benner, M. Q. Xu, *Protein Sci.* **1998**, *7*, 2256 – 2264.

[76] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249 – 289.

[77] R. David, M. P. Richter, A. G. Beck-Sickinger, *Eur. J. Biochem.* **2004**, *271*, 663 – 677.

[78] C. J. Noren, J. Wang, F. B. Perler, *Angew. Chem.* **2000**, *112*, 458 – 476; *Angew. Chem. Int. Ed.* **2000**, *39*, 450 – 466.

[79] L. Saleh, F. B. Perler, *Chem. Rec.* **2006**, *6*, 183 – 193.

[80] Inteins are evolutionary related to the autoprocessing domain of hedgehog proteins which utilizes this process in an intermolecular fassion to link a cholesterol moiety to the C terminus of the hedgehog signaling domain; for a review, see R. K. Mann, P. A. Beachy, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2000**, *1529*, 188 – 202.

[81] H. Wu, Z. Hu, X. Q. Liu, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9226 – 9231.

[82] I. Giriat, T. W. Muir, *J. Am. Chem. Soc.* **2003**, *125*, 7180 – 7181.

[83] W. Sun, J. Yang, X. Q. Liu, *J. Biol. Chem.* **2004**, *279*, 35281 – 35286.

[84] S. Brenzel, T. Kurpiers, H. D. Mootz, *Biochemistry* **2006**, *45*, 1571 – 1578.

[85] M. W. Southworth, E. Adam, D. Panne, R. Byer, R. Kautz, F. B. Perler, *EMBO J.* **1998**, *17*, 918 – 926.

[86] K. V. Mills, B. M. Lew, S. Jiang, H. Paulus, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3543 – 3548.

[87] C. Ludwig, M. Pfeiff, U. Linne, H. D. Mootz, *Angew. Chem.* **2006**, *118*, 5343 – 5347; *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 5218 – 5221.

[88] P. McCaldon, P. Argos, *Proteins Struct. Funct. Genet.* **1988**, *4*, 99 – 122.

[89] T. Wieland, G. Schneider, *Justus Liebigs Ann. Chem.* **1953**, *580*, 159 – 168.

[90] G. L. Ellman, *Arch. Biochem. Biophys.* **1959**, *82*, 70 – 77.

[91] L. Zhang, J. P. Tam, *Tetrahedron Lett.* **1997**, *38*, 3 – 6.

[92] For a review on the incorporation of Sec into proteins, see R. J. Hondal, *Protein Pept. Lett.* **2005**, *12*, 757 – 764.

[93] a) R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197 – 1206; b) M. D. Gieselman, L Xie, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1331 – 1334.

[94] a) M. D. Gieselman, Y. Zhu, H. Zhou, D. Galonic, W. A. van der Donk, *ChemBioChem* **2002**, *3*, 709 – 716; b) G. Roelfes, D. Hilvert, *Angew. Chem.* **2003**, *115*, 2377 – 2379; *Angew. Chem. Int. Ed.* **2003**, *42*, 2275 – 2277; c) R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140 – 5141.

[95] N. P. Luthra, R. C. Costello, J. D. Odom, *J. Biol. Chem.* **1981**, *257*, 1142 – 1144.

[96] a) S. M. Berry, M. D. Gieselman, M. J. Nilges, W. A. van der Donk, Y. Lu, *J. Am. Chem. Soc.* **2002**, *124*, 2084 – 2085; b) M. Ralle, S. M. Berry, M. J. Nilges, M. D. Gieselman, W. A. van der Donk, Y. Lu, N. J. Blackburn, *J. Am. Chem. Soc.* **2004**, *126*, 7244 – 7256.

[97] For reviews concerning the chemical modification of proteins, see a) D. Qi, C.-M. Tann, D. Haring, M. D. Distefano, *Chem. Rev.* **2001**, *101*, 3081 – 3111; b) C.-M. Tann, D. Qi, M. D. Distefano, *Curr. Opin. Chem. Biol.* **2001**, *5*, 696 – 704; c) D. R. W. Hodgson, J. M. Sanderson, *Chem. Soc. Rev.* **2004**, *33*, 422 – 430; d) B. G. Davis, *Curr. Opin. Biotechnol.* **2003**, *14*, 379 – 386; for the first examples of chemical modifications of amino acids in active sites, see: e) K. E. Neet, D. E. Koshland, *Proc. Natl. Acad. Sci.* **1966**, *56*, 1606 – 1611; f) L. Polgar, M. L. Bender, *J. Am. Chem. Soc.* **1966**, *88*, 3153 – 3154; g) P. I. Clark, G. Lowe, *J. Chem. Soc. Chem. Commun.* **1977**, 923 – 924.

[98] For the use of site-specific modifications in the directed molecular evolution of enzyme catalysts, see a) M. T. Reetz, *Tetrahedron* **2002**, *58*, 6595 – 6602; b) M. T. Reetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5716 – 5722.

[99] For an example on a site-specific glycosylation using this approach, see S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105 – 1109.

[100] M. T. Perlstein, M. Z. Atassi, S. H. Cheng, *Biochim. Biophys. Acta Protein Struct.* **1971**, *236*, 174 – 182.

[101] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526 – 533.

[102] a) A. Brik, Y. Y. Yang, S. Ficht, C. H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626 – 5627; b) D. Bang, G. I. Makhatadze, V. Tereshko, A. Kossiakoff, S. B. H. Kent, *Angew. Chem.* **2005**, *117*, 3920 – 3924; *Angew. Chem. Int. Ed.* **2005**, *44*, 3852 – 3856; c) D. Bang, A. V. Gribenko, V. Tereshko, A. A. Kossiakoff, S. B. H. Kent, G. I. Makhatadze, *Nat. Chem. Biol.* **2006**, *2*, 139 – 143; d) M. J. Bayro, J. Mukhopadhyay, G. V. T. Swapna, J. Y. Huang, L. C. Ma, E. Sineva, P. E. Dawson, G. T. Montelione, R. H. Ebright, *J. Am. Chem. Soc.* **2003**, *125*, 12382 – 12383.

[103] For the desulfurization in the glycopeptide-assisted ligation as introduced by Wong et al., see Section 3.3.5.

[104] B. L. Pentelute, S. B. H. Kent, *Org. Lett.* **2007**, *9*, 687 – 690.

[105] S. He, D. Bauman, J. S. Davis, A. Loyola, K. Nishioka, J. L. Gronlund, D. Reinberg, F. Meng, N. Kelleher, D. G. McCafferty, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12033 – 12038.

[106] a) K. Nishide, Y. Shigeta, K. Obata, T. Inoue, M. Node, *Tetrahedron Lett.* **1996**, *37*, 2271 – 2274; b) M. Node, K. Nishide, Y. Shigeta, K. Obata, H. Shiraki, H. Kunishige, *Tetrahedron* **1997**, *53*, 12883 – 12894.

[107] Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408 – 9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248 – 9252.

[108] G. J. L. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 5052 – 5053.

[109] For a recent protocol employing Dha for the synthesis acetylated and methylated Lys analogues, see J. Guo, J. Wang, J. S. Lee, P. G. Schultz, *Angew. Chem.* **2008**, *120*, 6499 – 6501; *Angew. Chem. Int. Ed.* **2008**, *47*, 6465 – 6468; for other chemoselective modifications of Dha, see a) Z. A. Knight, B. Schilling, R. H. Row, D. M. Kenski, B. W. Gibson, K. M. Shokat, *Nat. Biotechnol.* **2003**, *21*, 1047 – 1054; b) F. Rusnak, J. Zhou, G. Hathaway, *J. Biomol. Tech.* **2002**, *13*, 228 – 237.

[110] R. Quaderer, D. Hilvert, *Chem. Commun.* **2002**, *12*, 2620 – 2621.

[111] For an alternative access of Dha-containing proteins, phenylselenocysteins can be introduced by unnatural protein expression and eliminated by H_2O_2: J. Wang, S. M. Schiller, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 6973 – 6975; *Angew. Chem. Int. Ed.* **2007**, *46*, 6849 – 6851.

[112] For a previous discussion on the use of thiol-containing analogues of natural amino acids at ligation junctions, see Ref. [101].

[113] D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064 – 10065.

[114] a) C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912 – 6915; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807 – 6810; b) J. Chen, Q. Wan, Y. Yuan, J. Zhu, S. J. Danishefsky, *Angew. Chem.* **2008**, *120*, 8649 – 8652; *Angew. Chem. Int. Ed.* **2008**, *47*, 8521 – 8524.

[115] a) G. L. Kenyon, T. W. Bruice, *Methods Enzymol.* **1977**, *47*, 407 – 430; b) R. Wynn, F. M. Richards, *Methods Enzymol.* **1995**, *251*, 351 – 356; c) G. DeSantis, P. Berglund, M. R. Stabile, M. Gold, J. B. Jones, *Biochemistry* **1998**, *37*, 5968 – 5973; d) G. DeSantis, X. Shang, J. B. Jones, *Biochemistry* **1999**, *38*, 13391 – 13397.

[116] For a chemoselective functionalization of Cys residues to disulfide-linked glycoconjugates with glycomethane- or phenylthiosulfonates, see a) B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman, J. B. Jones, *Tetrahedron: Asymmetry* **2000**, *11*, 245 – 262; b) D. P. Gamblin, P. Garnier, S. J. Ward, N. J. Oldham, A. J. Fairbanks, B. G. Davis, *Org. Biomol. Chem.* **2003**, *1*, 3642 – 3644; for a strategy employing glycoselenenylsulfides, see c) D. P. Gamblin, P. Garnier, S. van Kasteren, N. J. Oldham, A. J. Fairbanks, B. G. Davis, *Angew. Chem.* **2004**, *116*, 846 – 851; *Angew. Chem. Int. Ed.* **2004**, *43*, 828 – 833; d) G. J. L. Bernardes, D. P. Gamblin, B. G. Davis, *Angew. Chem.* **2006**, *118*, 4111 – 4115; *Angew. Chem. Int. Ed.* **2006**, *45*, 4007 – 4011.

[117] G. J. L. Bernardes, E. L. Grayson, S. Thompson, J. M. Chalker, J. C. Errey, F. E. Oualid, T. D. W. Claridge, B. G. Davis, *Angew. Chem.* **2008**, *120*, 2276 – 2279; *Angew. Chem. Int. Ed.* **2008**, *47*, 2244 – 2247.

[118] D. A. Bochar, L. Tabernero, C. V. Stauffacher, V. W. Rodwell, *Biochemistry* **1999**, *38*, 8879 – 8883.

[119] G. G. Kochendoerfer, S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884 – 887.

[120] a) H. B. Smith, F. C. Hartman, *J. Biol. Chem.* **1988**, *263*, 4921 – 4925; b) C. E. Hopkins, P. B. O'Connor, K. N. Allen, C. E. Costello, D. R. Tolan, *Protein Sci.* **2002**, *11*, 1591 – 1599.

[121] R. Okamoto, Y. Kajihara, *Angew. Chem.* **2008**, *120*, 5482 – 5486; *Angew. Chem. Int. Ed.* **2008**, *47*, 5402 – 5406.

[122] For an ester to amide migration under slightly basic conditions in the synthesis of difficult peptide sequences, see a) I. Coin, R. Dölling, E. Krause, M. Bienert, M. Beyermann, C. Dan Sferdean, L. A. Carpino, *J. Org. Chem.* **2006**, *71*, 6171 – 6177; b) L. A. Carpino, E. Krause, C. Dan Sferdean, M. Schümann, H. Fabian, M. Bienert, M. Beyermann, *Tetrahedron Lett.* **2004**, *45*, 7519 – 7523; c) Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura, Y. Kiso, *Chem. Commun.* **2004**, 124 – 125; d) M. Mutter, A. Chandravarkar, C. Boyat, J. Lopez, S. Dos Santos, B. Mandal, R. Mimna, K. Murat, L. Patiny, L. Saucède, G. Tuchscherer, *Angew. Chem.* **2004**, *116*, 4267 – 4273; *Angew. Chem. Int. Ed.* **2004**, *43*, 4172 – 4178.

[123] Naturally occurring Met residues which would interfere with this strategy can be protected as their sulfoxides and reconverted into the thioethers; see also Ref. [50].

[124] J. P. Tam, Q. Yu, *Biopolymers* **1998**, *46*, 319 – 327; for a desulfurization of Hcy to α-aminoisobutyric acid (Abu) see Ref. [101].

[125] For NCL with an isocysteine residue instead of a Cys that leads to the final incorporation of β-amino acid in the peptide backbone, see C. Dose, O. Seitz, *Org. Biomol. Chem.* **2004**, *2*, 59 – 65.

[126] R. L. Heinrikson, *J. Biol. Chem.* **1971**, *246*, 4090 – 4096.

[127] F. R. Naider, J. M. Becker, *Biopolymers* **1997**, *43*, 3 – 14.

[128] For a recent highlight on cysteine-free ligations, see D. Macmillan, *Angew. Chem.* **2006**, *118*, 7830 – 7834; *Angew. Chem. Int. Ed.* **2006**, *45*, 7668 – 7672.

[129] For an application in the synthesis of disulfide-bridged protein analogues, see S. J. Bark, S. B. H. Kent, *FEBS Lett.* **1999**, *460*, 67 – 76.

[130] L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891 – 5896.

[131] a) W. D. F. Meutermans, S. W. Golding, G. T. Bourne, L. P. Miranda, M. J. Dooley, P. F. Alewood, M. L. Smythe, *J. Am. Chem. Soc.* **1999**, *121*, 9790 – 9796; b) Y. Shao, W. Lu, S. B. H. Kent, *Tetrahedron Lett.* **1998**, *39*, 3911 – 3914.

[132] a) P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831 – 1833; b) C. Marinzi, S. J. Bark, J. Offer, P. E. Dawson, *Bioorg. Med. Chem.* **2001**, *9*, 2323 – 2328.

[133] a) D. L. J. Clive, S. Hisaindee, D. M. Coltart, *J. Org. Chem.* **2003**, *68*, 9247 – 9254; b) S. Tchertchian, O. Hartley, P. Botti, *J. Org. Chem.* **2004**, *69*, 9208 – 9214; c) D. Macmillan, D. W. Anderson, *Org. Lett.* **2004**, *6*, 4659 – 4662.

[134] T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059 – 6061.

[135] C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorg. Med. Chem.* **2004**, *12*, 2749 – 2757.

[136] C. Chatterjee, R. K. McGinty, J.-P. Pellois, T. W. Muir, *Angew. Chem.* **2007**, *119*, 2872 – 2876; *Angew. Chem. Int. Ed.* **2007**, *46*, 2814 – 2818; for a highlight, including the relevance of ubiquilated proteins, see C. P. R. Hackenberger, *ChemBioChem* **2007**, *8*, 1221 – 1223.

[137] R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812 – 816.

[138] J. Offer, P. E. Dawson, *Org. Lett.* **2000**, *2*, 23 – 26.

[139] J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642 – 4646.

[140] D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554 – 6559.

[141] B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. Chem.* **2006**, *118*, 4222 – 4231; *Angew. Chem. Int. Ed.* **2006**, *45*, 4116 – 4125.

[142] This rearrangement was previously observed and utilized, see J. Vizzavona, F. Dick, T. Vorherr, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1963 – 1965; for an application of this reaction in the synthesis of thioesters, see T. Kawakami, M. Sumida, K. Nakamura, T. Vorherr, S. Aimoto, *Tetrahedron Lett.* **2005**, *46*, 8805 – 8807.

[143] A. Brik, C.-H. Wong, *Chem. Eur. J.* **2007**, *13*, 5670 – 5675.

[144] a) A. Brik, Y.-Y. Yang, S. Ficht, C.-H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626 – 5627; b) Y.-Y. Yang, S. Ficht, A. Brik, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 7690 – 7701.

[145] A. Brik, S. Ficht, Y.-Y. Yang, C. S. Bennett, C.-H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 15026 – 15033.

[146] R. J. Payne, S. Ficht, S. Tang, A. Brik, Y.-Y. Yang, D. A. Case, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 13527 – 13536.

[147] S. Ficht, R. J. Payne, A. Brik, C.-H. Wong, *Angew. Chem.* **2007**, *119*, 6079 – 6083; *Angew. Chem. Int. Ed.* **2007**, *46*, 5975 – 5979.

[148] R. J. Payne, S. Ficht, W. A. Greenberg, C.-H. Wong, *Angew. Chem.* **2008**, *120*, 4483 – 4487; *Angew. Chem. Int. Ed.* **2008**, *47*, 4411 – 4415.

[149] a) B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939 – 1941; b) E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141 – 2143; c) B. L. Nilsson, L. L. Kiesling, R. T. Raines, *Org. Lett.* **2001**, *3*, 9 – 12.

[150] For a review, see M. Köhn, R. Breinbauer, *Angew. Chem.* **2004**, *116*, 3168 – 3178; *Angew. Chem. Int. Ed.* **2004**, *43*, 3106 – 3116.

[151] J. Meyer, H. Staudinger, *Helv. Chim. Acta* **1919**, *2*, 635 – 646.

[152] M. B. Soellner, A. Tam, R. T. Raines, *J. Org. Chem.* **2006**, *71*, 9824 – 9830.

[153] A. Tam, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2007**, *129*, 11421 – 11430.

[154] M. B. Soellner, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2006**, *128*, 8820 – 8828.

[155] a) M. B. Soellner, K. A. Dickson, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 11790 – 11791; b) M. Köhn, R. Wacker, C. Peters, H. Schroeder, L. Soulere, R. Breinbauer, C. M. Niemeyer, H. Waldmann, *Angew. Chem.* **2003**, *115*, 6010 – 6014; *Angew. Chem. Int. Ed.* **2003**, *42*, 5830 – 5834; c) A. Watzke, M. Gutierrez-Rodriguez, M. Köhn, R. Wacker, H. Schroeder, R. Breinbauer, J. Kuhlmann, K. Alexandrov, C. M. Niemeyer, R. S. Goody, H. Waldmann, *Bioorg. Med. Chem.* **2006**, *14*, 6288 – 6306; d) A. Watzke, M. Köhn, M. Gutierrez-Rodriguez, R. Wacker, H. Schroeder, R. Breinbauer, J. Kuhlmann, K. Alexandrov, C. M. Niemeyer, R. S. Goody, H. Waldmann, *Angew. Chem.* **2006**, *118*, 1436 – 1440; *Angew. Chem. Int. Ed.* **2006**, *45*, 1408 – 1412.

[156] a) B. L. Nilsson, R. J. Hondal, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 5268 – 5269; b) L. Liu, Z.-Y. Hong, C.-H. Wong, *ChemBioChem* **2006**, *7*, 429 – 432; c) R. Merkx, D. T. S. Rijkers, J. Kemmink, R. M. J. Liskamp, *Tetrahedron Lett.* **2003**, *44*, 4515 – 4518.

[157] a) O. David, W. J. N. Meester, H. Bieräugel, H. Hiemstra, J. H. van Maarseveen, *Angew. Chem.* **2003**, *115*, 4509 – 4511; *Angew. Chem. Int. Ed.* **2003**, *42*, 4373 – 4375; b) G. Masson, T. den Hartog, H. E. Schoemaker, H. Hiemstra, J. H. van Maarseveen, *Synlett* **2006**, 865 – 868.

[158] a) E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007 – 2010; b) N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *Chem. Biol.* **2006**, *1*, 644 – 648.

[159] For the conjugation of unprotected sugar derivatives on protein carriers by the traceless Staudinger ligation, see C. Grandjean, A. Boutonnier, C. Guerreiro, J.-M. Fournier, L. A. Mulard, *J. Org. Chem.* **2005**, *70*, 7123 – 7132.

[160] R. Kleineweischede, C. P. R. Hackenberger, *Angew. Chem.* **2008**, *120*, 6073 – 6077; *Angew. Chem. Int. Ed.* **2008**, *47*, 5984 – 5988.

[161] a) C. F. Liu, J. P. Tam, *J. Am. Chem. Soc.* **1994**, *116*, 4149 – 4153; b) C. F. Liu, J. P. Tam, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6584 – 6588.

[162] J. P. Tam, Z. Miao, *J. Am. Chem. Soc.* **1999**, *121*, 9013 – 9022.

[163] J. P. Tam, Z. Miao, *J. Am. Chem. Soc.* **2000**, *122*, 4253 – 4260.

[164] G. Chen, J. D. Warren, J. Chen, B. Wu, Q. Wan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2006**, *128*, 7560 – 7562.

[165] G. Chen, Q. Wan, Z. Tan, C. Kan, Z. Hua, K. Ranganathan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 7527 – 7531; *Angew. Chem. Int. Ed.* **2007**, *46*, 7383 – 7387.

[166] The term "bioorthogonality" was introduced by C. R. Bertozzi: G. A. Lemieux, C. L. de Graffenried, C. R. Bertozzi, *J. Am. Chem. Soc.* **2003**, *125*, 4708 – 4709; for a subsequent review, see D. D. Dube, C. R. Bertozzi, *Curr. Opin. Chem. Biol.* **2003**, *7*, 616 – 625.

[167] a) J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.* **2005**, *1*, 13 – 21; b) H. C. Hang, C. R. Bertozzi, *Acc. Chem. Res.* **2001**, *34*, 727 – 736.

[168] a) C. P. R. Hackenberger, S. Hinderlich, *ChemBioChem* **2007**, *8*, 1763 – 1765; b) C. T. Campbell, S. G. Sampathkumar, K. J. Yarema, *Mol. BioSyst.* **2007**, *3*, 187 – 194; c) O. T. Keppler, R. Horstkorte, M. Pawlita, C. Schmidt, W. Reutter, *Glycobiology* **2001**, *11*, 11R – 18R; d) for an example of the modification of lipidated biopolymers, see Y. Kho, S. C. Kim, C. Jiang, D. Barma, B. Deb, S. W. Kwon, J. Cheng, J. Jaunbergs, C. Weinbaum, F. Tamanoi, J. Falck, Y Zhao, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12479 – 12484.

[169] For a review, see a) L. Wang, P. G. Schultz, *Angew. Chem.* **2005**, *117*, 34 – 68; *Angew. Chem. Int. Ed.* **2005**, *44*, 34 – 66; b) D. A. Dougherty, *Curr. Opin. Chem. Biol.* **2000**, *4*, 645 – 652; c) J. Xie, P. G. Schultz, *Curr. Opin. Chem. Biol.* **2005**, *9*, 548 – 554.

[170] A. J. Link, M. L. Mock, D. A. Tirell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603 – 609.

[171] a) I. Chen, M. Howath, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, *2*, 99 – 104; b) M. Howarth, A. Y. Ting, *Nat. Protoc.* **2008**, *3*, 534 – 545; c) T. Dierks, B. Schmidt, L. Borissenko, J. Peng, A. Preusser, M. Mariappan, K. von Figura, *Cell* **2003**, *113*, 435 – 444; d) I. S. Carrico, B. L. Carlson, C. R. Bertozzi, *Nat. Chem. Biol.* **2007**, *3*, 321 – 322; see also Refs. [247] and [248].

[172] G. Hosein Hakimelahi, G. Just, *Tetrahedron Lett.* **1980**, *21*, 2119 – 2122.

[173] Electron-poor azides can form thiatriazolines by a stepwise process; for a detailed mechanistic discussion, see R. V. Kolakowski, N. Shangguan, R. R. Sauers, L. J. Williams, *J. Am. Chem. Soc.* **2006**, *128*, 5695 – 5702.

[174] N. Shangguan, S. Katukojvala, R. Greenberg, L. J. Williams, *J. Am. Chem. Soc.* **2003**, *125*, 7754 – 7755; for a similar process involving selenocarboxylates instead of thioacids, see a) X. Wu, L. Hu, *Tetrahedron Lett.* **2005**, *46*, 8401 – 8405; b) X. Wu, L. Hu, *J. Org. Chem.* **2007**, *72*, 765 – 774.

[175] F. Fazio, C.-H. Wong, *Tetrahedron Lett.* **2003**, *44*, 9083 – 9085.

[176] a) R. Merkx, A. J. Brouwer, D. T. S. Rijkers, R. M. J. Liskamp, *Org. Lett.* **2005**, *7*, 1125 – 1128; b) K. N. Barlett, R. V. Kolakowski, S. Katukojvala, L. J. Williams, *Org. Lett.* **2006**, *8*, 823 – 826.

[177] J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem.* **2006**, *118*, 1270 – 1274; *Angew. Chem. Int. Ed.* **2006**, *45*, 1248 – 1252.

[178] For a recent stereoretentive synthesis of C-terminal α-ketoacid peptides by oxidative cleavage of cyanosulfur ylides, see L. Ju, A. R. Lippert, J. W. Bode, *J. Am. Chem. Soc.* **2008**, *130*, 4253 – 4255.

[179] N. Carrillo, E. A. Davalos, J. A. Russak, J. W. Bode, *J. Am. Chem. Soc.* **2006**, *128*, 1452 – 1453.

[180] W.-K. Chan, C.-M. Ho, M.-K. Wong, C.-M. Che, *J. Am. Chem. Soc.* **2006**, *128*, 14796 – 14797.

[181] R. Huisgen, G. Szeimies, L. Moebius, *Chem. Ber.* **1967**, *100*, 2494 – 2507.

[182] C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057 – 3064.

[183] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708 – 2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596 – 2599.

[184] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056 – 2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004 – 2021.

[185] H. C. Kolb, K. B. Sharpless, *Drug Discovery Today* **2003**, *8*, 1128 – 1137.

[186] C. Nolte, P. Mayer, B. F. Straub, *Angew. Chem.* **2007**, *119*, 2147 – 2149; *Angew. Chem. Int. Ed.* **2007**, *46*, 2101 – 2103.

[187] J.-F. Lutz, *Angew. Chem.* **2007**, *119*, 1036 – 1043; *Angew. Chem. Int. Ed.* **2007**, *46*, 1018 – 1025.

[188] P. Wu, V. V. Fokin, *Aldrichimica Acta* **2007**, *40*, 7 – 17.

[189] a) C. M. Salisbury, B. F. Cravatt, *QSAR Comb. Sci.* **2007**, *26*, 1229 – 1238; b) S. A. Sieber, B. F. Cravatt, *Chem. Commun.* **2006**, 2311 – 2319.

[190] R. Breinbauer, M. Köhn, *ChemBioChem* **2003**, *4*, 1147 – 1149.

[191] V. D. Bock, H. Hiemstra, J. H. van Maarseveen, *Eur. J. Org. Chem.* **2006**, 51 – 68.

[192] M. V. Gil, M. J. Arevalo, J. Lopez, *Synthesis* **2007**, 1589 – 1620.

[193] S. Dedola, S. A. Nepogodiev, R. A. Field, *Org. Biomol. Chem.* **2007**, *5*, 1006 – 1017.

[194] Y. L. Angell, K. Burgess, *Chem. Soc. Rev.* **2007**, *36*, 1674 – 1689.

[195] C. W. Tornoe, S. J. Sanderson, J. C. Mottram, G. H. Coombs, M. Meldal, *J. Comb. Chem.* **2004**, *6*, 312 – 324.

[196] W. S. Horne, M. K. Yadav, C. D. Stout, M. R. Ghadiri, *J. Am. Chem. Soc.* **2004**, *126*, 15366 – 15367.

[197] a) K. Oh, Z. Guan, *Chem. Commun.* **2006**, 3069 – 3071; b) Y. Angell, K. Burgess, *J. Org. Chem.* **2005**, *70*, 9595 – 9598.

[198] a) V. D. Bock, R. Perciaccante, T. P. Jansen, H. Hiemstra, J. H. V. Maarseveen, *Org. Lett.* **2006**, *8*, 919 – 922; b) V. D. Bock, D. Speijer, H. Hiemstra, J. H. V. Maarseveen, *Org. Biomol. Chem.* **2007**, *5*, 971 – 975; c) R. A. Turner, A. G. Oliver, R. S. Lokey, *Org. Lett.* **2007**, *9*, 5011 – 5014; d) M. Roice, I. Johannsen, M. Meldal, *QSAR Comb. Sci.* **2004**, *23*, 662 – 673.

[199] A. Tam, U. Arnold, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2007**, *129*, 12670 – 12671.

[200] a) V. Aucagne, D. A. Leigh, *Org. Lett.* **2006**, *8*, 4505 – 4507; b) R. Franke, C. Doll, J. Eichler, *Tetrahedron Lett.* **2005**, *46*, 4479 – 4482.

[201] a) C. J. Cavender, V. J. Shiner, *J. Org. Chem.* **1972**, *37*, 3567 – 3569; b) A. Vasella, C. Witzig, J.-L. Chiara, M. Martin-Lomas, *Helv. Chim. Acta* **1991**, *74*, 2073 – 2077; c) P. B. Alper, S.-C. Hung, C.-H. Wong, *Tetrahedron Lett.* **1996**, *37*, 6029 – 6032; d) P. T. Nyffeler, C.-H. Liang, K. M. Koeller, C.-H. Wong, *J. Am. Chem. Soc.* **2002**, *124*, 10773 – 10778; e) J. T. Lundquist, J. C. Pelletier, *Org. Lett.* **2001**, *3*, 781 – 783; f) S. Yekta, V. Prisyazhnyuk, H.-U. Reissig, *Synlett* **2007**, 2069 – 2072; for a recent one-pot procedure for the diazo-transfer and the azide–alkyne cycloaddition, see H. S. G. Beckmann, V. Wittmann, *Org. Lett.* **2007**, *9*, 1 – 4.

[202] For recent examples, see a) Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192 – 3193; b) A. E. Speers, G. C. Adam, B. F. Cravatt, *J. Am. Chem. Soc.* **2003**, *125*, 4686 – 4687; c) D. C. Dieterich, J. J. Lee, A. J. Link, J. Graumann, D. A. Tirell, E. M. Schuman, *Nat. Protoc.* **2007**, *2*, 532 – 540; d) L. Merkel, H. S. G. Beckmann, V. Wittmann, N. Budisa, *ChemBioChem* **2008**, *9*, 1220 – 1224; e) R. E. Connor, K. Piatkov, A. Varshavsky, D. A. Tirrell, *ChemBioChem* **2008**, *9*, 366 – 369.

[203] H. N. Lin, C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 13998 – 14003.

[204] M. Carmen Galan, E. McCoy, S. E. O'Connor, *Chem. Commun.* **2007**, 3249 – 3251.

[205] For a review, see N. K. Devaraj, J. P. Collman, *QSAR Comb. Sci.* **2007**, *26*, 1253 – 1260.

[206] P.-C. Lin, S.-H Ueng, M.-C. Tseng, J.-L. Ko, K.-T. Huang, S.-C. Yu, A. K. Adak, Y.-J. Chen, C.-C. Lin, *Angew. Chem.* **2006**, *118*, 4392 – 4396; *Angew. Chem. Int. Ed.* **2006**, *45*, 4286 – 4290.

[207] For an immobilization strategy of proteins through unnatural amino acids, see C. Gauchet, G. R. Labadie, C. D. Poulter, *J. Am. Chem. Soc.* **2006**, *128*, 9274 – 9275.

[208] a) A. J. T. Dirks, S. S. van Berkel, N. S. Hatzakis, J. A. Opsteen, F. L. van Delft, J. J. L. M. Cornelissen, A. E. Rowan, J. C. M. van Hest, F. P. J. T. Rutjes, R. J. M. Nolte, *Chem. Commun.* **2005**, 4172 – 4174; b) B. Parrish, R. B. Breitenkamp, T. Emrick, *J. Am. Chem. Soc.* **2005**, *127*, 7404 – 7410; c) P. Wu, A. K. Feldman, A. K. Nugent, C. J. Hawker, A. Scheel, B. Voit, J. Pyun, J. M. J. Frechet, K. B. Sharpless, V. V. Fokin, *Angew. Chem.* **2004**, *116*, 4018 – 4022; *Angew. Chem. Int. Ed.* **2004**, *43*, 3928 – 3932; d) D. T. S. Rijkers, G. W. van Esse, R. Merkx, A. J. Brouwer, H. J. F. Jacobs, R. J. Pieters, R. M. J. Liskamp, *Chem. Commun.* **2005**, 4581 – 4583.

[209] T. L. Mindt, H. Struthers, L. Brans, T. Anguelov, C. Schweinsberg, V. Maes, D. Tourwe, R. Schibli, *J. Am. Chem. Soc.* **2006**, *128*, 15096 – 15097.

[210] a) A. Dondoni, P. P. Giovannini, A. Massi, *Org. Lett.* **2004**, *6*, 2929 – 2932; b) B. H. M. Kuijpers, S. Groothuys, A. R. Keereweer, P. J. L. M. Quaedflieg, R. H. Blaauw, F. L. van Delft, F. P. J. T. Rutjes, *Org. Lett.* **2004**, *6*, 3123 – 3126; c) S. Groothuys, B. H. M. Kuijpers, P. J. L. M. Quaedflieg, H. C. P. F. Roelen, R. W. Wiertz, R. H. Blaauw, F. L. v. Delft, F. P. J. T. Rutjes, *Synthesis* **2006**, 3146 – 3152; d) H. Lin, C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 13998 – 14003; e) Q. Wan, J. Chen, G. Chen, S. J. Danishefsky, *J. Org. Chem.* **2006**, *71*, 8244 – 8249; f) E. Kaltgrad, M. K. O'Reilly, L. Liao, S. Han, J. C. Paulson, M. G. Finn, *J. Am. Chem. Soc.* **2008**, *130*, 4578 – 4579.

[211] H.-J. Musiol, S. Dong, M. Kaiser, R. Bausinger, A. Zumbusch, U. Bertsch, L. Moroder, *ChemBioChem* **2005**, *6*, 625 – 628.

[212] a) J. M. Baskin, C. R. Bertozzi, *QSAR Comb. Sci.* **2007**, *26*, 1211 – 1219; b) J.-F. Lutz, *Angew. Chem.* **2008**, *120*, 2212 – 2214; *Angew. Chem. Int. Ed.* **2008**, *47*, 2182 – 2184.

[213] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046 – 15047; for an earlier publication of the strain-promoted formation of triazols, see G. Wittig, A. Krebs, *Chem. Ber.* **1961**, *94*, 3260 – 3275.

[214] J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16793 – 16797.

[215] S. T. Laughlin, J. M. Baskin, S. L. Amacher, C. R. Bertozzi, *Science* **2008**, *320*, 664 – 667.

[216] E. M. Sletten, C. R. Bertozzi, *Org. Lett.* **2008**, *10*, 3097 – 3099.
[217] X. Ning, J. Guo, M. A. Wolfert, G.-J. Boons, *Angew. Chem.* **2008**, *120*, 2285 – 2287; *Angew. Chem. Int. Ed.* **2008**, *47*, 2253 – 2255.
[218] W. Song, Y. Wang, J. Qu, M. M. Madden, Q. Lin, *Angew. Chem.* **2008**, *120*, 2874 – 2877; *Angew. Chem. Int. Ed.* **2008**, *47*, 2832 – 2835.
[219] In the orginal report it is stated that alkynes behave as efficiently as electron-deficient alkenes as substrates, which could allow an easier application in bioorthogonal chemical reporter strategies.
[220] a) H. F. Gaertner, K. Rose, R. Cotton, D. Timms, R. Camble, R. E. Offord, *Bioconjugate Chem.* **1992**, *3*, 262 – 268; b) K. Rose, *J. Am. Chem. Soc.* **1994**, *116*, 30 – 33; c) L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125 – 1128; d) L. E. Canne, A. R. Ferre-D'Amare, S. K. Burley, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, *117*, 2998 – 3007; e) Y. Singh, O. Renaudet, E. Defrancq, P. Dumy, *Org. Lett.* **2005**, *7*, 1359 – 1362; f) S. E. Cervigni, P. Dumy, M. Mutter, *Angew. Chem.* **1996**, *108*, 1325 – 1328; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1230 – 1232; for a review see: J. A. Borgia, G. B. Fields, *Trends Biotechnol.* **2000**, *18*, 243 – 251.
[221] E. H. Cordes, W. P. Jencks, *J. Am. Chem. Soc.* **1962**, *84*, 832 – 837.
[222] V. W. Cornish, K. M. Hahn, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 8150 – 8151.
[223] Z. Zhang, B. A. C. Smith, L. Wang, A. Brock, C. Cho, P. G. Schultz, *Biochemistry* **2003**, *42*, 6735 – 6746.
[224] L. K. Mahal, K. J. Yarema, C. R. Bertozzi, *Science* **1997**, *276*, 1125 – 1128.
[225] For the expression of *para*-benzoyl-L-Phe-containing proteins, see W. Liu, L. Alfonta, A. V. Mack, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 6185 – 6187; *Angew. Chem. Int. Ed.* **2007**, *46*, 6073 – 6075.
[226] E. H. Cordes, W. P. Jencks, *J. Am. Chem. Soc.* **1962**, *84*, 826 – 831.
[227] a) A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem.* **2006**, *118*, 7743 – 7746; *Angew. Chem. Int. Ed.* **2006**, *45*, 7581 – 7584; b) for a recent application in the ^{18}F-labeling of proteins, see R. R. Flavell, P. Kothari, M. Bar-Dagan, M. Synan, S. Vallabhajosula, J. M. Friedman, T. W. Muir, G. Ceccarini, *J. Am. Chem. Soc.* **2008**, *130*, 9106 – 9112.
[228] A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, *128*, 15602 – 15603.
[229] a) A. D. de Araújo, J. M. Palomo, J. Cramer, M. Köhn, H. Schröder, R. Wacker, C. Niemeyer, K. Alexandrov, H. Waldmann, *Angew. Chem.* **2006**, *118*, 302 – 307; *Angew. Chem. Int. Ed.* **2006**, *45*, 296 – 301; b) A. D. de Araújo, J. M. Palomo, J. Cramer, O. Seitz, K. Alexandrov, H. Waldmann, *Chem. Eur. J.* **2006**, *12*, 6095 – 6109.
[230] For a detailed mechanistic investigation, see F. L. Lin, H. M. Hoyt, H. van Halbeek, R. G. Bergman, C. R. Bertozzi, *J. Am. Chem. Soc.* **2005**, *127*, 2686 – 2695.
[231] S. T. Laughlin, C. R. Bertozzi, *Nat. Protoc.* **2007**, *2*, 2930 – 2944.
[232] J. A. Prescher, D. H. Dube, A. R. Bertozzi, *Nature* **2004**, *430*, 873 – 877.
[233] N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *ACS Chem. Biol.* **2006**, *1*, 644 – 648.
[234] a) K. L. Kiick, E. Saxon, D. A. Tirrell, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 19 – 24; b) R. A. Chandra, E. S. Douglas, R. A. Mathies, C. R. Bertozzi, M. B. Francis, *Angew. Chem.* **2006**, *118*, 910 – 915; *Angew. Chem. Int. Ed.* **2006**, *45*, 896 – 901; c) M. J. Hangauer, C. R. Bertozzi, *Angew. Chem.* **2008**, *120*, 2428 – 2431; *Angew. Chem. Int. Ed.* **2008**, *47*, 2394 – 2397; d) M.-L. Tsao, F. Tian, P. G. Schultz, *ChemBioChem* **2005**, *6*, 2147 – 2149; e) S. C. Hsiao, A. K. Crow, W. A. Lam, C. R. Bertozzi, D. A. Fletcher, M. B. Francis, *Angew. Chem.* **2008**, *120*, 8601 – 8605; *Angew. Chem. Int. Ed.* **2008**, *47*, 8473 – 8477.
[235] M. Köhn, R. Wacker, C. Peters, H. Schroeder, L. Soulere, R. Breinbauer, C. M. Niemeyer, H. Waldmann, *Angew. Chem.* **2003**, *115*, 6010 – 6014; *Angew. Chem. Int. Ed.* **2003**, *42*, 5830 – 5834.
[236] A. Watzke, M. Köhn, M. Gutierrez-Rodriguez, R. Wacker, H. Schroeder, R. Breinbauer, J. Kuhlmann, K. Alexandrov, C. M. Niemeyer, R. S. Goody, H. Waldmann, *Angew. Chem.* **2006**, *118*, 1436 – 1440; *Angew. Chem. Int. Ed.* **2006**, *45*, 1408 – 1412.
[237] For a recent example, see O. Avrutina, H.-U. Schmoldt, D. Gabrijelcic-Geiger, A. Wentzel, H. Frauendorf, C. P. Sommerhoff, U. Diederichsen, H. Kolmar, *ChemBioChem* **2008**, *9*, 33 – 37.
[238] a) L. S. Forsberg, J. H. Pazur, *Carbohydr. Res.* **1979**, *75*, 129 – 140; b) A. M. D'Alessandro, G. D'Andrea, A. Oratore, *Clin. Chim. Acta* **1998**, *274*, 189 – 197; c) D. J. O'shannessy, R. H. Quarles, *J. Appl. Biochem.* **1985**, *7*, 347 – 355; d) B. Liu, L. Burdine, T. Kodadek, *J. Am. Chem. Soc.* **2006**, *128*, 15228 – 15235; e) For a protocol minimzing the oxidation of sulfur by addition of methionine, see J. H. Chen, W. G. Zeng, R. Offord, K. Rose, *Bioconjugate Chem.* **2003**, *14*, 614 – 618.
[239] J. M. Hooker, A. P. Esser-Kahn, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 10256 – 10257.
[240] R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, D. M. Horn, P. G. Schultz, *J. Am. Chem. Soc.* **2003**, *125*, 935 – 939.
[241] Z. M. Carrico, D. W. Romanici, R. A. Mehl, M. B. Francis, *Chem. Commun.* **2008**, 1205 – 1207.
[242] a) C. Kieburg, M. Dubber, T. K. Lindhorst, *Synlett* **1997**, 1447 – 1449; b) D. Ramos, P. Rollin, W. Klaffke, *J. Org. Chem.* **2001**, *66*, 2948 – 2956; c) P. B. van Seeventer, J. A. L. M. van Dorst, J. F. Siemerink, J. P. Kamerling, J. F. G. Vliegenthart, *Carbohydr. Res.* **1997**, *300*, 369 – 373; d) T. Buskas, E. Söderberg, P. Konradsson, B. Fraser-Reid, *J. Org. Chem.* **2000**, *65*, 958 – 963; e) C. D. Heidecke, T. K. Lindhorst, *Chem. Eur. J.* **2007**, *13*, 9056 – 9067; f) J. L. de Paz, C. Noti, P. H. Seeberger, *J. Am. Chem. Soc.* **2006**, *128*, 2766 – 2767.
[243] T. Posner, *Ber. Dtsch. Chem. Ges.* **1905**, *38*, 646 – 657.
[244] S. Wittrock, T. Becker, H. Kunz, *Angew. Chem.* **2007**, *119*, 5319 – 5323; *Angew. Chem. Int. Ed.* **2007**, *46*, 5226 – 5230.
[245] a) B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269 – 272; b) S. A. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis, R. Y. Tsien, *J. Am. Chem. Soc.* **2002**, *124*, 6063 – 6076; c) G. Gaietta, T. J. Deerick, S. R. Adams, J. Bouwer, O. Tour, D. W. Laird, G. E. Sosinsky, R. Y. Tsien, M. H. Ellisman, *Science* **2002**, *296*, 503 – 507; d) A. B. Griffin, S. R. Adams, J. Jones, R. Tsien, *Methods Enzymol.* **2000**, *327*, 565 – 578.
[246] a) K. J. Franz, M. Nitz, B. Imperiali, *ChemBioChem* **2003**, *4*, 265 – 271; b) M. Nitz, K. J. Franz, R. L. Maglathlin, B. Imperiali, *ChemBioChem* **2003**, *4*, 272 – 276; c) B. R. Sculimbrene, B. Imperiali, *J. Am. Chem. Soc.* **2006**, *128*, 7346 – 7352.
[247] a) N. George, H. Pick, H. Vogel, N. Johnsson, K. Johnsson, *J. Am. Chem. Soc.* **2004**, *126*, 8896 – 8897; b) J. Yin, F. Liu, C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 7754 – 7755; see also Refs. [171 a,b].
[248] a) A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86 – 89; b) G. V. Los, A. Darzins, N. Karassina, C. Zimprich, R. Learish, M. G. McDougall, L. P. Encell, R. Friedman-Ohana, M. Wood, G. Vidugiris, K. Zimmerman, P. Otto, D. H. Klaubert, K. V. Wood, *Promega Cell Notes* **2005**, *11*, 2 – 6.
[249] a) H. M. O'Hare, K. Johnsson, A. Gautier, *Curr. Opin. Struct. Biol.* **2007**, *17*, 488 – 494; b) A. Dragulescu-Andrasi, J. Rao, *ChemBioChem* **2007**, *8*, 1099 – 1101; c) E. A. Jares-Erijman, T. M. Jovin in *Laboratory Techniques in Biochemistry and*

Molecular Biology (Ed.: T. W. J. Gadella), Elsevier, **2009**, Chap. 12.

[250] B. Bader, K. Kuhn, D. J. Owen, H. Waldmann, A. Wittinghofer, J. Kuhlmann, *Nature* **2000**, *403*, 223 – 226.

[251] K. Kuhn, D. J. Owen, B. Bader, A. Wittinghofer, J. Kuhlmann, H. Waldmann, *J. Am. Chem. Soc.* **2001**, *123*, 1023 – 1035.

[252] M. Völkert, K. Uwai, A. Tebbe, B. Popkirova, M. Wagner, J. Kuhlmann, H. Waldmann, *J. Am. Chem. Soc.* **2003**, *125*, 12749 – 12758.

[253] G. T. Hermanson in *Bioconjugation Techniques*, Academic Press, San Diego, **1996**.

[254] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942 – 15943.

[255] D. W. Romanini, M. B. Francis, *Bioconjugate Chem.* **2008**, *19*, 153 – 157.

[256] For a selective derivatization of Trp residues with rhodium carbenoids, see J. M. Antos, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 10256 – 10257.

[257] J. M. McFarland, N. S. Joshi, M. B. Francis, *J. Am. Chem. Soc.* **2008**, *130*, 7639 – 7644.

[258] A. Einhauer, A. Jungbauer, *J. Biochem. Biophys. Methods* **2001**, *49*, 455 – 465.

[259] S. D. Tilley, M. B. Francis, *J. Am. Chem. Soc.* **2006**, *128*, 1080 – 1081; for a review on Pd-catalyzed reactions in water, see: a) J. P. Genet, M. Savignac, *J. Organomet. Chem.* **1999**, *576*, 305 – 317; b) B. H. Lipshutz, S. Ghorai, *Aldrichim. Acta* **2008**, *41*, 59 – 72.

[260] J. M. Hooker, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 3718 – 3719.

[261] T. L. Schlick, Z. Ding, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 3718 – 3723.

[262] For other oxidative conversions of Tyr residues, see a) S. Meunier, E. Strable, M. G. Finn, *Chem. Biol.* **2004**, *11*, 319 – 326; b) D. A. Fancy, C. Denison, K. Kim, Y. Xie, T. Holdeman, F. Amini, T. Kodadek, *Chem. Biol.* **2000**, *7*, 697 – 708.

[263] F. Bordusa, *Chem. Rev.* **2002**, *102*, 4817 – 4867.

[264] V. Schellenberger, H.-D. Jakubke, *Angew. Chem.* **1991**, *103*, 1440 – 1452; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1437 – 1449.

[265] H.-D. Jakubke, P. Kuhl, A. Könnecke, *Angew. Chem.* **1985**, *97*, 79 – 87; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 85 – 93.

[266] R. Smits, B. Koksch, *Curr. Top. Med. Chem.* **2006**, *6*, 1483 – 1498.

[267] S. Rival, J. Saulnier, J. Wallach, *J. Biocatal. Biotrans.* **2000**, *17*, 417 – 429.

[268] Y-H- Ye, G.-L. Tian, G.-W. Xing, D.-C. Dai, G. Chen, C.-X. Li, *Tetrahedron* **1998**, *54*, 12585 – 12596.

[269] R. Günther, S. Thust, H.-J. Hofmann, F. Bordusa, *Eur. J. Biochem.* **2000**, *267*, 3496 – 3501.

[270] N. Wehofsky, N. Koglin, S. Thust, F. Bordusa, *J. Am Chem. Soc.* **2003**, *125*, 6126 – 6133.

[271] H. Sekizaki, K. Itoh, E. Toyota, K. Tanizawa, *Chem. Pharm. Bull.* **1996**, *4*, 1585 – 1587.

[272] L. Abrahmsen, J. Tom, J. Burnier, K. A. Butcher, A. Kossiakoff, J. A. Wells. , *Biochemistry* **1991**, *30*, 4151 – 4159.

[273] Z. Machova, R. von Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa, A. G. Beck-Sickinger, *Angew. Chem.* **2003**, *115*, 5065 – 5067; *Angew. Chem. Int. Ed. Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 4916 – 4918.

[274] S. K. Mazmanian, G. Liu, H. Ton-That, O. Schneewind, *Science* **1999**, *285*, 760 – 763.

[275] H. Ton-That, G. Liu, S. K. Mazmanian, K. F. Faull, O. Schneewind, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12424 – 12429.

[276] H. Mao, A. H. Scott, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670 – 2671.

[277] S. Pritz, Y. Wolf, O. Kraetke, J. Klose, M. Bienert, M. Beyermann, *J. Org. Chem.* **2007**, *72*, 3909 – 3912.

[278] M. W. Popp, J. M. Antos, G. M. Grotenbreg, E. Spooner, H. L. Ploegh, *Nat. Chem. Biol.* **2007**, *3*, 707 – 708.

[279] J. P. Pellois, T. W. Muir, *Curr. Opin. Chem. Biol.* **2006**, *10*, 487 – 491.

[280] R. S. Goody, K. Alexandrov, M. Engelhard, *ChemBioChem* **2002**, *3*, 399 – 403.

[281] D. Schwarzer, P. A. Cole, *Curr. Opin. Chem. Biol.* **2005**, *9*, 561 – 569.

[282] M. A. Shogren-Knaak, C. J. Fry, C. L. Peterson, *J. Biol. Chem.* **2003**, *278*, 15744 – 15748.

[283] a) M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson, *Science* **2006**, *311*, 844 – 847; b) E. M. Vogel, B. Imperiali, *Protein Sci.* **2007**, *16*, 550 – 556 and examples in Section 6.2.

[284] a) D. Macmillan, C. R. Bertozzi, *Angew. Chem.* **2004**, *116*, 1379 – 1383; *Angew. Chem. Int. Ed.* **2004**, *43*, 1355 – 1359; b) C. P. R. Hackenberger, C. T. Friel, S. E. Radford, B. Imperiali, *J. Am. Chem. Soc.* **2005**, *127*, 12882 – 12889; c) C. F. W. Becker, X. Liu, D. Olschewski, R. Castelli, R. Seidel, P. H. Seeberger, *Angew. Chem.* **2008**, *120*, 8215 – 8219; *Angew. Chem. Int. Ed.* **2008**, *47*, 8215 – 8219.

[285] For reviews on glycoprotein synthesis, see a) M. J. Grogan, M. R. Pratt, L. A. Marcaurelle, C. R. Bertozzi, *Annu. Rev. Biochem.* **2002**, *71*, 593 – 634; b) B. G. Davis, *Chem. Rev.* **2002**, *102*, 579 – 601; c) L. Liu, C. S. Bennett, C.-H. Wong, *Chem. Commun.* **2006**, 21 – 33; d) C. Haase, O. Seitz, *Top. Curr. Chem.* **2006**, *267*, 1 – 36.

[286] a) A. Rak, O. Pylypenko, T. Durek, A. Watzke, S. Kushnir, L. Brunsveld, H. Waldmann, R. S. Goody, K. Alexandrov, *Science* **2003**, *302*, 646 – 650; b) O. Pylypenko, A. Rak, T. Durek, S. Kushnir, B. E. Dursina, N. H. Thomae, A. T. Constantinescu, L. Brunsveld, A. Watzke, H. Waldmann, R. S. Goody, K. Alexandrov, *EMBO J.* **2006**, *25*, 13 – 23.

[287] M. C. Y. Chang, C. S. Yee, D. G. Nocera, J. Stubbe, *J. Am. Chem. Soc.* **2004**, *126*, 16702 – 16703.

[288] a) M. R. Seyedsayamdost, C. S. Yee, J. Stubbe, *Nat. Protoc.* **2007**, *2*, 1225 – 1235; b) C. S. Yee, M. C. Chang, J. Ge, D. G. Nocera, J. Stubbe, *J. Am. Chem. Soc.* **2003**, *125*, 10506 – 10507.

[289] K. Shen, P. A. Cole, *J. Am. Chem. Soc.* **2003**, *125*, 16172 – 16173.

[290] a) F. I. Valiyaveetil, M. Sekedat, R. Mackinnon, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17045 – 17049; b) F. I. Valiyaveetil, M. Leonetti, T. W. Muir, R. Mackinnon, *Science* **2006**, *314*, 1004 – 1007.

[291] a) Z. Zhang, K. Shen, W. Lu, P. A. Cole, *J. Biol. Chem.* **2003**, *278*, 4668 – 4674; b) W. Zheng, D. Schwarzer, A. Lebeau, J. L. Weller, D. C. Klein, P. A. Cole, *J. Biol. Chem.* **2005**, *280*, 10462 – 10467.

[292] C. M. Niemeyer, *Angew. Chem.* **2001**, *113*, 4254 – 4287; *Angew. Chem. Int. Ed.* **2001**, *40*, 4128 – 4158.

[293] L. Cao, L. van Langen, R. A. Sheldon, *Curr. Opin. Biotechnol.* **2003**, *14*, 387 – 394.

[294] Q. Xu, K. S. Lam, *J. Biomed. Biotechnol.* **2003**, *2003*, 257 – 266.

[295] M. L. Lesaicherre, R. Y. Lue, G. Y. Chen, Q. Zhu, S. Q. Yao, *J. Am. Chem. Soc.* **2002**, *124*, 8768 – 8769.

[296] K. Holland-Nell, A. G. Beck-Sickinger, *ChemBioChem* **2007**, *8*, 1071 – 1076.

[297] C. F. Becker, Y. Marsac, P. Hazarika, J. Moser, R. S. Goody, C. M. Niemeyer, *ChemBioChem* **2007**, *8*, 32 – 36.

[298] T. Hunter, *Cell.* **2000**, *100*, 113 – 127.

[299] For example: J. W. Wu, M. Hu, J. Chai, J. Seoane, M. Huse, C. Li, D. J. Rigotti, S. Kyin, T. W. Muir, R. Fairman, J. Massagué, Y. Shi, *Mol. Cell* **2001**, *8*, 1277 – 1289.

[300] P. M. Siegel, J. Massagué, *Nat. Rev. Cancer* **2003**, *3*, 807 – 821.

[301] M. E. Hahn, T. W. Muir, *Angew. Chem.* **2004**, *116*, 5924 – 5927; *Angew. Chem. Int. Ed.* **2004**, *43*, 5800 – 5803.

[302] D. M. Rothman, M. D. Shults, B. Imperiali, *Trends Cell Biol.* **2005**, *15*, 502 – 510.

[303] G. Mayer, A. Heckel, *Angew. Chem.* **2006**, *118*, 5020 – 5042; *Angew. Chem. Int. Ed.* **2006**, *45*, 4900 – 4921.

[304] M. E. Hahn, J. P. Pellois, M. Vila-Perelló, T. W. Muir, *ChemBioChem* **2007**, *8*, 2100 – 2105.
[305] P. Bouchard, Z. Zhao, D. Banville, F. Dumas, E. H. Fischer, S. H. Shen, *J. Biol. Chem.* **1994**, *269*, 19585 – 19589.
[306] W. Lu, D. Gong, D. Bar-Sagi, P. A. Cole, *Mol. Cell* **2001**, *8*, 759 – 769.
[307] K. Shen, A. C. Hines, D. Schwarzer, K. A. Pickin, P. A. Cole, *Biochim. Biophys. Acta Proteins Proteomics* **2005**, *1754*, 65 – 78.
[308] G. Raugei, G. Ramponi, P. Chiarugi, *Cell. Mol. Life Sci.* **2002**, *59*, 941 – 949.
[309] D. Schwarzer, Z. Zhang, W. Zheng, P. A. Cole, *J. Am. Chem. Soc.* **2006**, *128*, 4192 – 4193.
[310] a) L. Brunsveld, J. Kuhlmann, K. Alexandrov, A. Wittinghofer, R. S. Goody, H. Waldmann, *Angew. Chem.* **2006**, *118*, 6774 – 6798; *Angew. Chem. Int. Ed.* **2006**, *45*, 6622 – 6646; b) D. Olschewski, C. F. W. Becker, *Mol. Biosyst.* **2008**, *4*, 730 – 733.
[311] B. Bader, K. Kuhn, D. J. Owen, H. Waldmann, A. Wittinghofer, J. Kuhlmann, *Nature* **2000**, *403*, 223 – 226.
[312] D. Kadereit, J. Kuhlmann, H. Waldmann, *ChemBioChem* **2000**, *1*, 144 – 169.
[313] O. Rocks, A. Peyker, M. Kahms, P. J. Verveer, C. Koerner, M. Lumbierres, J. Kuhlmann, H. Waldmann, A. Wittinghofer, P. I. Bastiaens, *Science* **2005**, *307*, 1746 – 1752.
[314] T. W. Muir, P. E. Dawson, S. B. Kent, *Methods Enzymol.* **1997**, *289*, 266 – 298.
[315] G. S. Beligere, P. E. Dawson, *Biopolymers* **1999**, *51*, 363 – 369.
[316] C. F. Becker, C. L. Hunter, R. Seidel, S. B. Kent, R. S. Goody, M. Engelhard, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5075 – 5080; for further three-segment ligations, see also a) D. Bang, N. Chopra, S. B. Kent, *J. Am. Chem. Soc.* **2004**, *126*, 1377 – 1383; b) D. Bang, S. B. Kent, *Angew. Chem.* **2004**, *116*, 2588 – 2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534 – 2538.
[317] D. F. Veber, J. D. Milkowski, S. L. Varga, R. G. Denkewalter, R, Hirschmann, *J. Am. Chem. Soc.* **1972**, *94*, 5456 – 5461.
[318] K. Jacobs, C. Shoemaker, R. Rudersdorf, S. D. Neill, R. J. Kaufman, A. Mufson, J. Seehra, S. S. Jones, R. Hewick, E. F. Fritsch, M. Kawakita, T. Shimizu, T. Miyakeet, *Nature* **1985**, *313*, 806 – 810.
[319] It should be noted that the Danishefsky research group has published several glycopeptide ligation strategies which aim towards the synthesis of naturally glycosylated EPO. For examples see Refs. [53, 107, 141, 164, [165].
[320] G. J. Cotton, T. W. Muir, *Chem. Biol.* **2000**, *7*, 253 – 261.
[321] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243 – 247.
[322] a) B. Gutte, R. B. Merrifield, *J. Am. Chem. Soc.* **1969**, *91*, 501 – 502; b) R. G. Denkewalter, D. F. Veber, F. W. Holly, R. Hirschmann, *J. Am. Chem. Soc.* **1969**, *91*, 502 – 503.
[323] D. J. Boerema, V. A. Tereshko, S. B. Kent, *Biopolymers* **2008**, *90*, 278 – 286; for an EPL-based synthesis of modified RNAse A, see: U. Arnold, M. P. Hinderaker, B. L. Nilsson, B. R. Huck, S. H. Gellman, R. T. Raines, *J. Am. Chem. Soc.* **2002**, *124*, 8522 – 8523.
[324] S. Pritz, O. Kraetke, A. Klose, J. Klose, S. Rothemund, K. Fechner, M. Bienert, M. Beyermann, *Angew. Chem.* **2008**, *120*, 3698 – 3701; *Angew. Chem. Int. Ed.* **2008**, *47*, 3642 – 3645.
[325] M. H. Perrin, W. H. Fischer, K. S. Kunitake, A. G. Craig, S. C. Koerber, L. A. Cervini, J. E. Rivier, J. C. Groppe, J. Greenwald, S. Møller Nielsen, W. W, Vale, *J. Biol. Chem.* **2001**, *276*, 31528 – 31534.
[326] J. Klose, K. Fechner, M. Beyermann, E. Krause, N. Wendt, M. Bienert, R. Rudolph, S. Rothemund, *Biochemistry* **2005**, *44*, 1614 – 1623.
[327] The use of topological templates for the selective installation of structural elements to generate synthetic proteins was introduced with the template-assembled synthetic protein (TASP) concept: M. Mutter, S. Vuilleumier, *Angew. Chem.* **1989**, *101*, 551 – 571; *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 535 – 554.
[328] D. Bang, B. L. Pentelute, S. B. Kent, *Angew. Chem.* **2006**, *118*, 4089 – 4092; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985 – 3988.
[329] For a direct on-resin synthesis of peptide phenylthioesters, see D. Bang, B. L. Pentelute, Z. P. Gates, S. B. Kent, *Org. Lett.* **2006**, *8*, 1049 – 1052.
[330] T. Durek, V. Y. Torbeev, S. B. Kent, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4846 – 4851.
[331] During the revision process of this manuscript various other publications appeared which are of particular interest for this Review. These include, for example, a light- and protease-activated O→N-acyl migration in protein splicing (M. Vila-Perelló, Y. Hori, M. Ribó, T. Muir, *Angew. Chem.* **2008**, *120*, 7878 – 7881; *Angew. Chem. Int. Ed.* **2008**, *47*, 7764 – 7767), the application of aqueous cross-metathesis in protein modification (Y. A. Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 9642 – 9643), the semisynthesis of a GPI-anchored prion protein (see Ref. [284c]), the chemoselective labeling of sulfenic acid modified proteins in living cells (K. G. Reddie, Y. H. Seo, W. B. Muse III, S. E. Leonard, K. S. Caroll, *Mol. Biosyst.* **2008**, *4*, 447 – 464), the generation of a FRET-labeled serotonin N-acetyltransferase by a tandem ligation strategy (L. M. Szewczuk, M. K. Tarrant, V. Sample, W. J. Drury III, J. Zhang, P. A. Cole, *Biochemistry* **2008**, *47*, 10407 – 10419), and a chemoselective bioconjugation with tetrazines by an inverse-electron-demand Diels–Alder reaction (M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518 – 13519).

