Modern Ligation Methods to Access Natural and Modified Proteins

Alice L. Baumann[ab] and Christian P. R. Hackenberger[*ab]

Abstract: Proteins and peptides are gaining increasing interest as tools and targets in fundamental research and drug discovery. Growing research applications have prompted the need for methodologies that produce homogenous peptide and protein material. The development of efficient, chemoselective ligation reactions using unprotected peptide fragments presents a key solution for this challenging task. This review outlines modern ligation methods that enable the synthesis of both native, and also labelled or post-translationally modified peptides and proteins. The ligation methods herein discussed focus on the formation of the backbone amide bond.

Keywords: Chemoselective ligation · Native chemical ligation · Protein modification · Protein semi-synthesis

Alice L. Baumann was born in Walchwil, Switzerland. She studied interdisciplinary sciences at ETH Zurich, where she obtained her MSc in 2015. Following an internship at Novartis AG in Basel, she joined Prof. Christian Hackenberger's group at the Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP) in 2016. Her work is funded by a scholarship from the Deutsche Studienstiftung. In her research, she focuses on the development of a new cysteine-selective bioconjugation reaction.

Christian P. R. Hackenberger completed his graduate studies at the universities of Freiburg and UW Madison (MSc with Samuel Gellman) and his doctoral studies in 2003 at the RWTH Aachen in the group of Carsten Bolm. After a postdoctoral position at MIT with Barbara Imperiali, he started his own group at the Freie Universität Berlin in 2005, funded by the FCI (Liebig scholarship), the DFG (Emmy-Noether program) and the Boehringer Ingelheim Foundation (Plus 3 Award). In 2012, he was appointed Leibniz-Humboldt Professor for Chemical Biology at the Leibniz-Research Institute for Molecular Pharmacology and the Humboldt Universität zu Berlin founded by the Einstein Foundation. His group works on the development of new chemoselective and bioorthogonal reactions and novel approaches to functional protein synthesis and delivery, in particular for the labeling and modification of different antibody formats.

1. Introduction: The Need for Chemical Protein Synthesis

Peptides and proteins are biological macromolecules that are involved in virtually every cellular process. The biological function of a protein is a consequence of its unique folded structure, which in turn is dictated by its specific amino acid sequence. Organisms and chemists have both found ways to further increase the functional repertoire of the 21 proteinogenic amino acids (including selenocysteine) by chemical modifications. Nature uses cellular machineries to install post-translational modification (PTM) on proteins, thereby generating a wide range of chemical alterations of amino acids. These modifications include phosphorylation, acetylation, methylation, or the attachment of carbohydrates or even small proteins such as ubiquitin to the side chains of certain amino acids. PTMs have been shown to profoundly modulate both the structure and function of proteins.[1] On the other hand, chemical synthesis additionally allows for the incorporation of unnatural building blocks into proteins, such as D-amino acids,[2] polymers[3] or other functional tags (*e.g.* drugs,[4] fluorophores, purification tags, *etc.*). Unnatural building blocks help to fulfil the need for functionalized proteins and biopharmaceuticals in science and the pharmaceutical industry.

Critically, in order to study the effect of defined protein modifications, homogenous peptide and protein material is required. While proteins are typically accessed through biological expression, targeted modification remains challenging within cells despite recent advances in the incorporation of unnatural amino acids.[5,6] Particularly with regard to the *site-specific* incorporation of PTMs or other tags in a protein or peptide, chemical synthesis methods as well as chemo-enzymatic approaches have proven indispensable for their ability to produce well-defined (modified) proteins. Unambiguous modifications are crucial for meaningful evaluations of downstream processes.

As starting material for chemical protein synthesis, polypeptides are used, which are nowadays routinely prepared by solid phase peptide synthesis (SPPS).[7,8] For technical reasons, SPPS is usually limited to peptides of 40–50 amino acids in length, necessitating that larger polypeptides be assembled from several smaller fragments. Early work

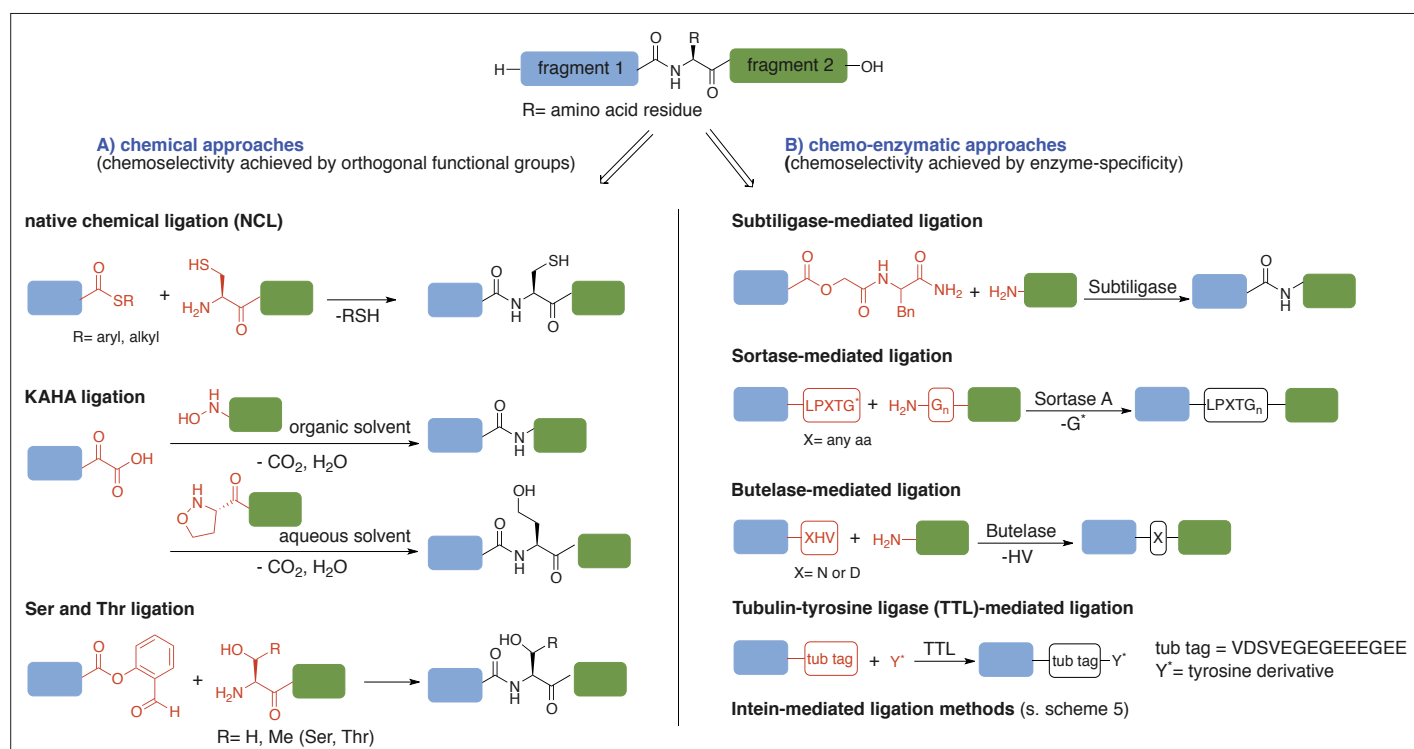*Correspondence:* Prof. C. P. R. Hackenberger[ab]
E-mail: hackenbe@fmp-berlin.de
[a]Leibniz-Forschungsinstitut für Molekulare Pharmakologie im Forschungsverbund Berlin e.V. (FMP)
Campus Berlin-Buch
Robert-Roessle-Str. 10, D-13125 Berlin
[b]Humboldt-Universität zu Berlin
Institut für Chemie
Brook-Taylor-Str. 2, D-12489 Berlin

Scheme 1. Selection of A) chemical and B) chemo-enzymatic peptide and protein conjugation methods using unprotected fragments. KAHA: keto-acid-hydroxylamine.

in this area focused on the condensation of side-chain protected fragments.[9] However, this strategy proved to be limited by unintentional epimerization at the C-terminal amino acid during activation, as well as poor solubility of protected peptide fragments. The development of *chemoselective ligation techniques of unprotected peptides and proteins* has proven to be a key solution to these problems.

This review aims at providing an overview of modern ligation techniques using *unprotected side-chain fragments* and *yielding native peptide bonds*. Methods to chemoselectively modify the side chain functionalities of amino acids are beyond the scope of this discussion and the interested reader is kindly referred to excellent overviews in this area.[6,10] Here, both chemo-enzymatic and purely chemical ligation methods are highlighted, the latter class with a focus on native chemical ligation since it is the most widely used ligation technique to date. Notably, other excellent review articles that cover the subject of peptide ligation exist, including a very recent article from the Payne group on advances in the development of new NCL-inspired ligation technologies.[1,11–15]

2. Overview of Peptide Ligation Techniques Using Unprotected Fragments

Generating native peptide bonds between two unprotected peptides in a chemoselective manner is as challenging as it is useful. Approaches to achieve this goal can be divided in two categories: purely chemical methods and chemo-enzymatic methods (Scheme 1A and B, respectively). The latter rely on the catalytic power of enzymes, usually recognizing a specific amino-acid sequence, to ligate the peptide or protein of interest (POI) to another fragment or to attach a tag to it. In turn, purely chemical methods rely on functional groups that react chemoselectively with each other.

In addition to the aforementioned chemoselectivity, an ideal ligation method should have the following features: It should proceed quickly and under mild conditions, in aqueous buffers and at low molar concentrations, thereby producing no chemical by-products. Furthermore, it should not require catalysts or other chemical additives. From a practical perspective, the starting material should be stable for storage and easily accessible, by biochemical methods, to enable semi-synthetic strategies. The following sections present a selection of the most popular and practical methodologies for both chemical and chemo-enzymatic approaches.

2.1 Native Chemical Ligation

To date, the most robust, practical and widely used method for ligating two unprotected peptide fragments is *native chemical ligation (NCL)*, which was introduced in 1994 by Kent and coworkers.[16,17] Since then, NCL has emerged as an indispensable methodology for chemical protein synthesis.[18]

The reaction mechanism of native chemical ligation involves first a reversible trans-thioesterification reaction *via* a nucleophilic attack of the *N*-terminal Cys-thiolate moiety of the *C*-terminal fragment to the *C*-terminal thioester of the *N*-terminal fragment. In a second step, the resulting activated thioester then undergoes (under physiological conditions) an irreversible, proximity-driven intramolecular S→N acyl shift, yielding a native peptide bond with a cysteine residue at the conjugation junction. The presence of excess free thiol and thus reversibility of the first step guarantees the exquisite regioselectivity of peptides containing other free cysteines.

In order to access *C*-terminal peptide thioesters as starting materials, various methods have been developed (see also section *sequential ligation*), including the synthesis and activation of peptide hydrazide[19] as well as the 3,4-diaminobenzoic-acid (Dbz)[20] and *o*-amino(methyl) aniline (MeDbz) linkers,[21] which are compatible with Fmoc/tBu-SPPS (compare also Scheme 2). Notably, NCL reactions usually proceed with low mM peptide concentrations and in the presence of a chaotropic reagent (*e.g.* 6 M guanidine-HCl).

The NCL method proved to be highly practical because of its compatibility with conditions that are ideal for most peptides and proteins (*i.e.* purely aqueous media and at neutral pH). Moreover, NCL allows for the semi-synthesis of proteins and thus gives rise to homogenous (modified) proteins.
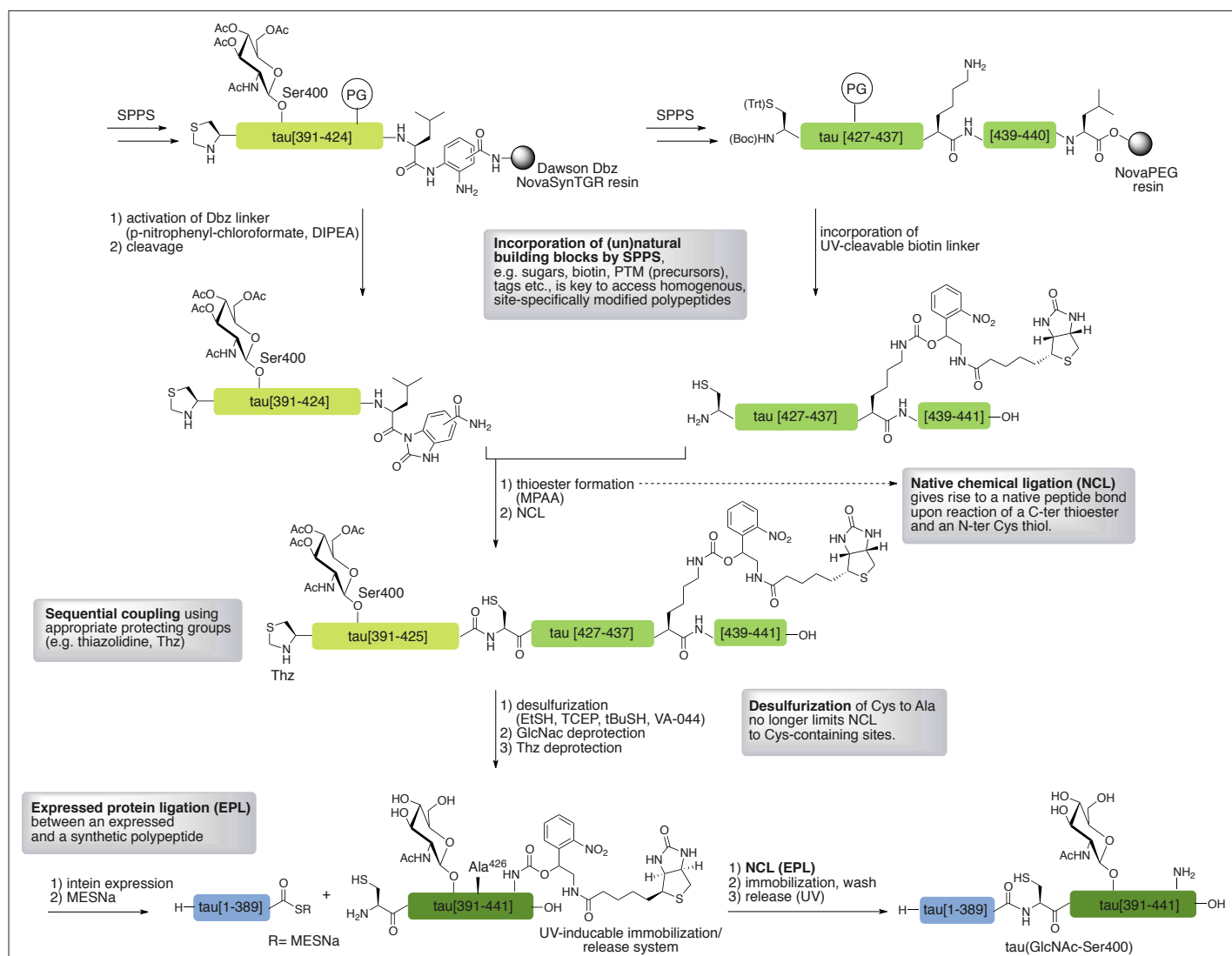
Scheme 2. Chemical synthesis of a site-specifically *O*-GlcNc modified tau protein.[21] Some of the tools in the NCL toolbox, used in this synthesis, are depicted in grey boxes.

The impact of NCL-based peptide ligation becomes apparent when looking at the numerous examples of polypeptides that have been synthesized to date using this method. These include proteins with selectively introduced tags, as well as examples of the total synthesis of native or modified proteins,[11,15] such as the synthesis of a site-selectively *O*-GlcNAc-modified tau protein performed by our laboratory (Scheme 2).[22]

Tau plays an important role in the regulation of mictrotubule stabilization, which is crucial to the integrity of the neuronal network. Hyperphosphorylation of tau is implicated in Alzheimer's disease.[23–25] Another PTM,
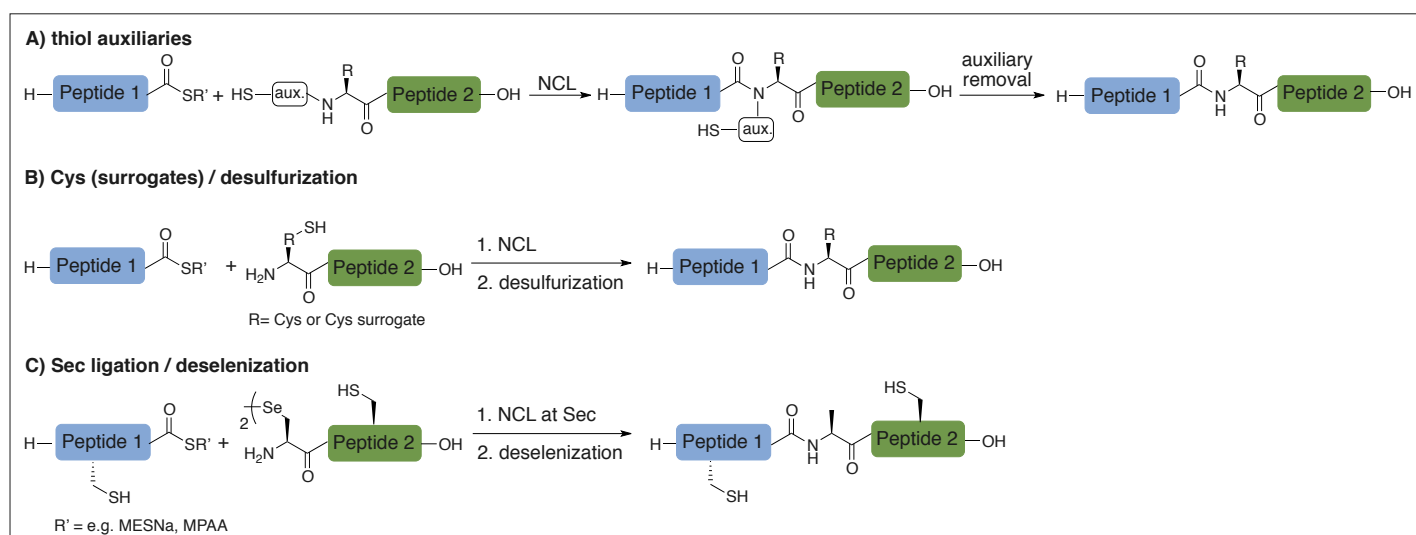
Scheme 3. Ligation methods at Cys-free junctions.

O-linked β-*N*-acetylglucosaminylation (*O*-GlcNAcylation), was found to inhibit tau phosphorylation in a site-specific manner.[26] Our group developed an NCL-based synthetic route, along with a traceless cleavable biotin-photo-linker purification system, which enabled access to site-specifically *O*-GlcNAc-modified tau. Notably, Haj-Yahya and Lashuel recently presented another important contribution on the semisynthesis of post-translationally modified tau.[27] Using our reported tau synthesis, some of the important concepts and elements of the NCL toolbox are highlighted (grey boxes, Scheme 2).

2.2 Strategies beyond NCL: Auxiliaries and Post-ligation Modification

A drawback of the traditional NCL strategy can be that it relies on a native cysteine residue at the conjugation site. In recent years, several strategies have been devised to circumvent the need of a native Cys residue at the conjugation site. These include: use of (*N*-terminal) auxiliaries, Cys surrogates, as well as selenocysteine (Sec), combined with a post-ligation desulfurization or deselenization step (Scheme 3).

Thiol-containing auxiliaries react with the thioester at the *C*-terminus of an *N*-peptide in a mechanism resembling NCL. Upon removal of the auxiliary with suitable reagents, a native peptide or protein sequence is generated (Scheme 3A). Diverse auxiliary-based methods have been developed by, amongst others, the groups of Kent,[28] Wong,[29,30] Brik,[31] Hojo/Nakahara,[32] and most recently Seitz.[33]

In 2001, Yan and Dawson introduced the concept of *post-ligation desulfurization*,[34] which revolutionized the field of NCL. Post-ligation desulfurization utilizes chemical reduction of the cysteine at the conjugation site resulting from NCL to alanine (Scheme 3B). In this way, cysteine-based NCL can be used to generate junctions containing native alanine, a significantly more abundant amino acid (*ca.* 8.9% vs. *ca.* 1.8% abundance of Cys), thereby enlarging the retro-synthetic disconnection possibilities significantly. Large excess of Raney Ni or Pd on Al$_2$O$_3$ were used to reduce the sulfhydryl group, causing side reactions in some cases. In 2007, Danishefsky and co-workers introduced a milder, metal-free desulfurization protocol using TCEP in combination with a water-soluble radical initiator (VA-044) and hydrogen atom source (*e.g.* *t*BuSH).[35,36] Notably, these conditions are chemoselective in the presence of methionine residues and thioesters. Li and co-workers recently reported a yet more elegant protocol, using a combination of borohydride and TCEP, thereby even circumventing the need for a radical-initiator.[37] In the synthesis of semisynthetic tau (Scheme 2), we employed the Danishefsky protocol to desulfurize Cys^{426} to native Ala^{426}, which allowed us to accomplish the ligation at this more accessible site.[22]

Inspired by the Cys→Ala desulfurization methodology, a mad dash began to develop methods for accessing other thiol-derived variants of the canonical amino acids that could analogously be used as *Cys-surrogates* in NCL (Scheme 3B). Following approximately a decade of intensive research, the Cys-surrogate toolbox gained various Fmoc-SPPS-compatible thiol-derivatives of the proteinogenic amino acids and several non-natural amino acids. For the chemical structures of these amino acids and further details, the interested reader is referred to the indicated references.[15,38–42]

Several groups advanced the use of selenocysteine (Sec) in NCL-like transformation with peptide thioesters, giving rise to selenoproteins by *Sec ligation* (Scheme 3C). [43–46] In 2010, the Dawson group discovered that deselenization of a Sec residue can be chemoselectively achieved using TCEP and DTT in the presence of unprotected Cys residues.[47] The power of this methodology has been exemplified for instance by Metanis and co-workers in the synthesis of human enzyme phosphohistidine phosphatase.[48] Another landmark contribution in the field came from the Payne and Metanis groups, who both discovered that Sec can be cleanly converted into Ser with TCEP in the presence of an exogenous oxidant such us oxygen or ozone.[49–51] These discoveries, along with the synthesis of other seleno-derived amino acids,[52] have further broadened the scope of Sec ligation chemistry beyond Ala disconnections.

2.3 Sequential Ligations

Using sequential NCL reactions, *large proteins can also be assembled from more than two fragments* (each of which can be additionally modified). The challenge thereby is to steer the reactivity of the individual fragments such that they react with each other in the desired order.

For the *iterative ligation of peptides in the N→C direction*, two main concepts exist to control the reaction order (Scheme 4): utilizing the varying reactivities of different thioesters (kinetic control) or 'turning on' thioesters through unmasking of their respective precursors at the desired stage of the reaction sequence.

The concept of *kinetically controlled ligation* was first employed by Kent and co-workers in the six-segment assembly of the protein crambin.[53] They utilized the fact that aryl thioesters are more reactive than alkyl thioesters to make the reaction selective (Scheme 4A). This facilitates the incorporation of a bifunctional fragment (*e.g.* Cys on the *N*-terminus and alkyl thioester on the *C*-terminus), which, after the first ligation step, can be activated by adding excess of an aryl thiol additive for a subsequent ligation with a third fragment. A drawback of this otherwise very powerful method is the radical quenching activity of commonly used aryl thiol additives (*e.g.* phenyl thiol or MPAA), which prohibits post-ligation *in situ* radical desulfurization of the ligation product. This limitation has been overcome by the discovery of 2,2,2-trifluoroethanethiol (TFET) as an alternative thiol additive, which increases the rate of the ligation reaction to a similar degree as aryl thiols but does not exhibit their radical quenching properties.[54] Its use has been showcased in a one-pot kinetically-controlled synthesis of sulfated variants of madanin-1 protein.[55]

A conceptually different strategy for iterative ligations in the N→C direction relies on *thioester precursors*, which are activated at a given point in the reaction sequence with suitable reagents to enable ligation to a subsequent fragment (Scheme 4B). Commonly
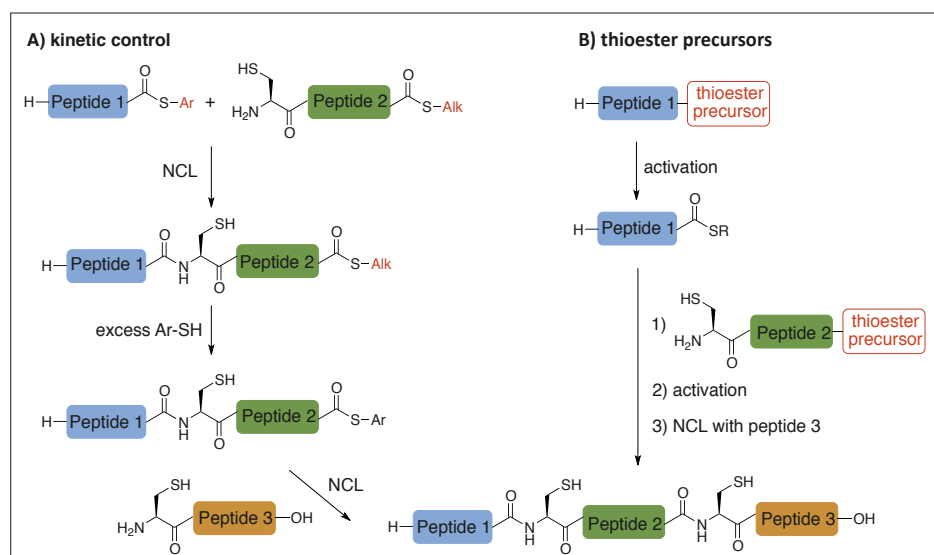
Scheme 4: Strategies for the iterative assembly of multiple peptides. Ar = aryl, Alk = alkyl.

used thioester precursors comprise Dbz-linkers,[20,21] acyl pyrazoles,[56] cryptothioesters,[57–61] peptide acyl hydrazides,[19,62] N-alkyl-Cys,[63,64] o-aminoanilides,[65] C-terminal Cys-activation,[66] as well as thioacids and strained thiolactones.[67,68] In addition, a new method for N→C sequential ligation using thioacid capture ligation and NCL has been reported by Hou et al.[69]

In parallel with the revolutionary advances in the field of NCL discussed above, researchers set out to tackle some of the inherent limitations of the technology. A primary limitation of NCL is the slow ligation rates at sterically demanding amino acids. It turned out that the 21st amino acid selenocysteine (Sec) possesses tailored physical-chemical properties that partially overcome these issues. For example, the lower pKa of Sec (5.2–5.6) compared to Cys (8.2), renders it suitable for ligation at lower pH, which can lead to higher yields due to less thioester hydrolysis. In addition, Sec exists predominantly as selenolate at physiological pH, thus reacting more efficiently than Cys at the same pH. Replacing the thioester moiety with an alkyl selenoester thereby improved ligation rates at sterically hindered C-terminal amino acids.[15,70] For a more substantial overview of recent achievements in the field of selenium-based ligation, as well as examples of synthesized proteins, the reader is referred to the excellent review from the Payne group, published this year.[15]

In addition to methods for N→C protein assembly, researchers have also developed several effective methods for *assembling peptide fragments in the C→N direction*. A popular strategy is to use orthogonal protecting groups for the N-terminal Cys residue to precisely control sequential ligation steps. Thiazolidine (Thz) and acetamidomethyl (Acm) derivatives of cysteine, but also of other thiol amino acids, are the most popular candidates and have been employed in several successful protein syntheses, *e.g.* in the synthesis of tetraubiquitin by Brik and co-workers[71] and in the synthesis of glycosylated interferon-β by Kajihara and co-workers.[72]

2.4 Other Chemoselective Amide Bond-forming Reactions

Despite NCL being the most popular peptide ligation technique, other chemical approaches have been developed that present viable alternatives for the formation of native peptide bonds (Scheme 1A).

In 2006, Bode and co-workers introduced the *ketoacid-hydroxylamine (KAHA) ligation*.[73] This chemoselective ligation method is based on the condensation of a C-terminal α-ketoacid and an N-terminal N-hydroxylamine. The reaction works additive-free and the only side products are CO_2 and H_2O. The α-ketoacid pep-

tide can be prepared on-resin by oxidation of a cyanosulfur-ylide based linker[74] and the N-hydroxylamine moiety can be reliably prepared on resin by a method based on Fukuyama's protocol.[75,76] While the traditional method was limited to organic solvents, the 5-oxaproline KAHA ligation also works in aqueous solvents, yielding a homoserine residue at the ligation site.[77] A concise review by Bode highlights recent advances of the KAHA ligation.[78] Examples of proteins prepared with KAHA ligation include prokaryotic ubiquitin-like protein (Pup),[77] probable cold shock protein A (CspA),[77] ubiquitin-fold modifier-1 (UFM1)[79] and nitrophorin 4.[80]

In 2010, the Chan group described another powerful ligation method: the so-called *Ser or Thr ligation*[81] between an N-terminal Ser/Thr peptide and a C-terminal salicylaldehyde ester peptide. In the first step, an N,O-benzylidene acetal intermediate is generated, followed by an O,N-acyl shift yielding a native peptide bond after a subsequent acidolysis step. The salicylaldehyde ester can be prepared on resin *via* phenolysis of an N-acyl-benzimidazolidinone (Nbz) protected peptide with salicylaldehyde dimethyl acetal.[82] Advantages of the Ser/Thr ligation include its racemization-free mechanism and also its chemoselectivity in the presence of thioester functionalities, making it compatible with NCL. Furthermore, an inherent advantage of this method is the versatility afforded by the high abundance of Ser and Thr residues in natural polypeptides. This ligation strategy has been employed in the chemical synthesis of large proteins such as phosphorylated/methylated nuclear protein HMGA1a[83] or glycosylated interleukin-25 (using a combination of NCL and Ser/Thr ligation),[84] as well as in the macrocyclization of peptides.[85]

Both the KAHA and the Ser/Thr ligation are powerful alternatives to the traditional NCL method, which allow not only for the accession of modified proteins but also the preparation of cyclized peptides. These methods do not rely on a cysteine at the ligation site (such as in traditional NCL), thereby further expanding the scope of possible connection sites. It should be mentioned however, that recent advances in the NCL methodology have also overcome this limitation (see above).

2.5 Enzymatic Amide Bond-forming Methods

In addition to the purely chemical methods discussed above, there are various chemo-enzymatic ligation methods for the ligation of unprotected peptide and protein fragments that make use of the inherent chemoselectivity of enzymes. Prominent examples include the usage of reverse proteolysis (*e.g.* by subtiligase), transpeptidases such as sortase A, as well as other ligation enzymes

(*e.g.* butelase) and split intein-based ligation methods (Scheme 1B).

In the early 1990s, the Wells group introduced the enzyme *subtiligase*, an engineered peptide ligase derived from the bacterial serine protease subtilisin (BPN).[86,87] Two point mutations (BPN S221C P225A) alter the mechanism to favor aminolysis over peptidase activity, thereby facilitating the ligation of a C-terminal ester-containing peptide to an acceptor peptide containing an amine. Subtiligase has been employed in the synthesis of a number of targets, including the ribonuclease A from six segments.[88] This enzyme accepts a wide scope of substrate amino acid sequences on both peptide ends, making it versatile while also limiting its utility for N-terminal bioconjugation. However, Weeks and Wells recently reported a library of subtiligase variants with defined specificities that allow for site-specific modification of protein N-termini.[89] In addition, the same group engineered subtiligase variants that efficiently ligate Cys-free peptides to recombinant protein thioesters fragments.[90] These recent findings, along with the fast reaction times and mild reaction conditions make subtiligase-catalyzed peptide ligation a very powerful tool for peptide ligation and modification.

Sortase A is a transpeptidase, naturally catalyzing the ligation of surface proteins to the cell wall of Gram-positive bacteria.[91] Sortase A recognizes a C-terminal amino acid sequence LPXTG and cleaves it at the threonine residue. The resulting enzyme-substrate acyl intermediate complex can be converted with a second peptide fragment, bearing at least one glycine residue at the N-terminus, to afford the ligation product, which still contains the LPXTG motif.[92] For an overview of applications using sortase A, which despite its principle reversibility became very popular due to the simple access of the starting material, the reader is referred to other excellent reviews.[92–94]

Another peptide ligase, named *butelase 1*, was recently discovered by Nguyen et al. from the Tam group.[95] They isolated this enzyme from *C. ternatea*, a tropical cyclotide-producing plant. Butelase 1 features catalytic efficiencies as high as 1.34×10^6 $M^{-1}s^{-1}$ and k_{cat} of up to 17 s^{-1} and is thus the fastest peptide ligase known to date (> 10k faster than sortase A).[96] Butelase 1 recognizes a C-terminal tripeptide motif Asn/Asp(Asx)-His-Val. It cleaves after Asx and ligates it to the N-terminal residue, thereby displaying broad specificity for the N-terminal amino acids of the substrate to ligate. Butelase 1 mediated ligation reactions are generally completed within minutes, thereby often achieving excellent yield. Moreover, butelase 1 is particularly useful for the efficient macrocyclization of peptides.[97] For intermolecular ligation, an excessive amount of a substrate (>5-fold) is
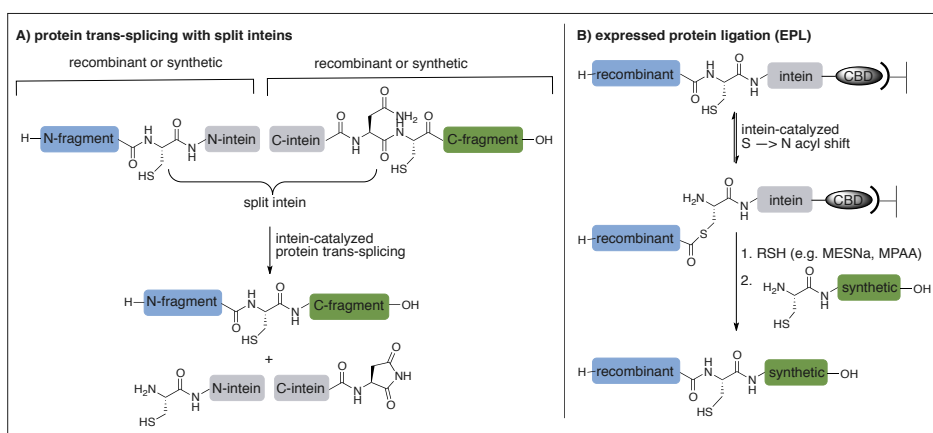
Scheme 5. Strategies for protein semi-synthesis: A) Intein-catalyzed protein trans-splicing, B) mechanism of expressed protein ligation (EPL). CBD = chitin binding domain, MESNa = 2-mercaptoethanesulfonate sodium, MPAA = 4-mercaptophenylacetic acid.

needed to compete with the cleaved dipeptide His-Val. However, this limitation can be overcome by using a thiodepsipeptide as a substrate instead.[98] Additionally, butelase 1 has been employed for the synthesis of protein thioesters.[99]

In our group, another enzymatic method has been developed, employing the enzyme *tubulin tyrosine ligase (TTL)*. In nature, TTL recognizes a 14-mer amino acid *C*-terminal sequence of a protein, termed 'tub tag', and attaches the amino acid tyrosine to it.[100] However, TTL also tolerates tyrosine derivatives, *e.g.* those equipped with click handles, thereby allowing further functionalization of the ligation product.[101,102]

2.6 Protein-splicing and Protein Ligation

Two enzyme-based ligation methods that are particularly useful for the *semi-synthesis of proteins* are expressed protein ligation (EPL) and protein trans-splicing (Scheme 5). Conceptually, both methods rely on an intein-catalyzed protein-splicing step. Inteins (Int) are protein segments that have the ability to excise themselves and join the remaining fragments (the exteins) with a peptide bond in a process called protein-splicing.[103,104] Intein-mediated protein-splicing is spontaneous; hence, all the necessary enzyme activities for these steps are imbedded in the intein segment itself. Inspired by this naturally occurring process, scientists made use of intein-based systems for peptide ligation purposes.

In a process called *trans-splicing of proteins*, so-called 'split inteins' are employed. Recombinant and/or synthetic peptide or protein fragments, each equipped with an intein fragment, associate to form a fully functional intein that bridges the two extein fragments. This intein bridge is removed by protein splicing to afford the protein or peptide of interest (Scheme 5A). Commonly, the SPPS-derived synthetic fragment is fused to the smaller intein fragment, while the recombinant fragment is equipped with the larger intein portion. Since the discovery of split inteins, the feasibility of this approach has been demonstrated in a plethora of examples.[105]

Another milestone around the topic of NCL was achieved when scientists found a way to convert recombinant proteins into protein thioesters. Two decades ago, Cole and co-workers introduced the concept of *expressed protein ligation (EPL)*.[106–109] Similar to the concept explained above, the key to the EPL mechanism is also an intein-catalyzed protein splicing step giving rise to a thioester, which is activated by adding excess thiols (typically MESNa or MPAA). Upon addition of a synthetic (modified) peptide fragment bearing an *N*-terminal cysteine residue, the ligation step yields the semisynthetic protein of interest (Scheme 5B). The intein component is usually fused to an affinity tag, such as the chitin-binding domain (CBD), which facilitates purification.

EPL is useful for the semisynthetic assembly of large modified proteins, but also for native proteins, which are difficult to express in full-length due to folding or solubility problems. In the synthesis of tau (Scheme 2), an EPL step has been employed to conjugate the expressed N-terminal fragment with the synthetic, glycosylated *C*-terminal fragment.[22]

3. Summary and Outlook

Modern chemical ligation methods present a useful solution to the challenge of the total chemical synthesis of proteins. The plethora of chemical and chemo-enzymatic ligation methods that arose over the last decades have greatly enhanced the scope of accessible targets, and importantly also have enabled the incorporation of labels or PTMs.

Among the techniques employed for the coupling of peptide segments, NCL has proven to be particularly practical and robust. Since the seminal report in 1994, NCL has seen revolutionary advances. The use of thiol amino acid surrogates, in combination with the ligation-desulfurization approach and the methodologies that enable sequential ligations of several fragments in an iterative manner, have paved the way for efficient chemical protein synthesis. The power of the NCL toolbox has been demonstrated in the synthesis of many natural and modified targets, including larger proteins consisting of > 120 amino acids.

In our opinion, the limiting factor in the total chemical synthesis of proteins is no longer the ligation itself, but the occasional laborious and time-consuming accession of peptide starting material by SPPS, along with tedious HPLC or other chromatographic purification steps. The field of chemical protein synthesis would therefore greatly benefit from advances that further the automation of these processes in a robust and practical way.

Nevertheless, it is fair to say that the well-equipped toolbox that exists to date for accessing homogenous, natural and modified proteins made it possible to study the effect of molecular changes of a protein in a plethora of biological and pharmaceutical contexts. Without any doubt, those methods have thus truly revolutionized the field of chemical biology.

[1]   P. Siman, A. Brik, *Org. Biomol. Chem.* **2012**, *10*, 5684.
[2]   M. Uppalapati, D. J. Lee, K. Mandal, H. Li, L. P. Miranda, J. Lowitz, J. Kenney, J. J. Adams, D. Ault-Riché, S. B. H. Kent, S. S. Sidhu, *ACS Chem. Biol.* **2016**, *11*, 1058.
[3]   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. Wiedke, 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. H. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884.
[4]   G. Casi, N. Huguenin-Dezot, K. Zuberbühler, J. Scheuermann, D. Neri, *J. Am. Chem. Soc.* **2012**, *134*, 5887.
[5]   L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498.

[6]   K. Lang, J. W. Chin, *Chem. Rev.* **2014**, *114*, 4764.

[7]   R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.

[8]   R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **2016**, *22*, 4.

[9]   T. Kimmerlin, D. Seebach, *J. Pept. Res.* **2005**, *65*, 229.

[10]   E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974.

[11]   J. P. Tam, Q. Yu, Z. Miao, *Biopolymers* **1999**, *51*, 311.

[12]   C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 10030.

[13]   A. F. M. Noisier, F. Albericio, *Amino Acids, Pept. Proteins* **2014**, *39*, 1.

[14]   P. Thapa, R.-Y. Zhang, V. Menon, J.-P. Bingham, *Molecules* **2014**, *19*, 14461.

[15]   S. S. Kulkarni, J. Sayers, B. Premdjee, R. J. Payne, *Nat. Rev. Chem.* **2018**, *2*, 122.

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

[17]   For important earlier studies see: T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, *Justus Liebigs Ann. Chem.* **1953**, *583*, 129.

[18]   S. B. H. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338.

[19]   J.-S. Zheng, S. Tang, Y.-K. Qi, Z.-P. Wang, L. Liu, *Nat. Protoc.* **2013**, *8*, 2483.

[20]   J. B. Blanco-Canosa, P. E. Dawson, *Angew. Chem. Int. Ed.* **2008**, *47*, 6851.

[21]   J. B. Blanco-Canosa, B. Nardone, F. Albericio, P. E. Dawson, *J. Am. Chem. Soc.* **2015**, *137*, 7197.

[22]   S. Schwagerus, O. Reimann, C. Despres, C. Smet-Nocca, C. P. R. Hackenberger, *J. Pept. Sci.* **2016**, *22*, 327.

[23]   E. Koepke, Y.-C. Tung, S. Shaikh, A. C. Alonso, K. Iqbal, I. Grundke-Iqbal, *J. Biol. Chem.* **1993**, *268*, 24374.

[24]   F. Liu, T. Zaidi, K. Iqbal, I. Grundke-Iqbal, R. K. Merkle, C.-X. Gong, *FEBS Lett.* **2002**, *512*, 101.

[25]   L. Martin, X. Latypova, F. Terro, *Neurochem. Int.* **2011**, *58*, 458.

[26]   S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies, D. J. Vocadlo, *Nat. Chem. Biol.* **2008**, *4*, 483.

[27]   M. Haj-Yahya, H. A. Lashuel, *J. Am. Chem. Soc.* **2018**, *140*, 6611.

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

[29]   S. Ficht, R. J. Payne, A. Brik, C.-H. Wong, *Angew. Chem. Int. Ed.* **2007**, *46*, 5975.

[30]   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.

[31]   M. Y. Lutsky, N. Nepomniaschiy, A. Brik, *Chem. Commun.* **2008**, *0*, 1229.

[32]   H. Hojo, C. Ozawa, H. Katayama, A. Ueki, Y. Nakahara, Y. Nakahara, *Angew. Chem. Int. Ed.* **2010**, *49*, 5318.

[33]   S. F. Loibl, Z. Harpaz, O. Seitz, *Angew. Chem. Int. Ed.* **2015**, *54*, 15055.

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

[35]   Q. Wan, S. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248.

[36]   F. W. Hoffmann, R. J. Ess, T. C. Simmons, R. S. Hanzel, *J. Am. Chem. Soc.* **1956**, *78*, 6414.

[37]   K. Jin, T. Li, H. Y. Chow, H. Liu, X. Li, *Angew. Chem. Int. Ed.* **2017**, *56*, 14607.

[38]   L. R. Malins, R. J. Payne, *Curr. Opin. Chem. Biol.* **2014**, *22*, 70.

[39]   L. R. Malins, R. J. Payne, *Aust. J. Chem.* **2015**, *68*, 521.

[40]   B. Premdjee, R. J. Payne, 'Synthesis of Proteins by Native Chemical Ligation–Desulfurization Strategies', in 'Chemical Ligation', Eds. L. D. D'Andrea, A. Romanelli, John Wiley & Sons USA, **2017**, pp. 161.

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

[42]   P. Botti, S. Tchertchian, WO Patent Appl. No. 2006133962, **2006**.

[43]   R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140.

[44]   R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197.

[45]   M. D. Gieselman, L. Xie, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1331.

[46]   R. Quaderer, D. Hilvert, *Chem. Commun.* **2002**, *0*, 2620.

[47]   N. Metanis, E. Keinan, P. E. Dawson, *Angew. Chem. Int. Ed.* **2010**, *49*, 7049.

[48]   P. Sai Reddy, S. Dery, N. Metanis, *Angew. Chem. Int. Ed.* **2016**, *55*, 992.

[49]   L. R. Malins, N. J. Mitchell, S. McGowan, R. J. Payne, *Angew. Chem. Int. Ed.* **2015**, *54*, 12716.

[50]   S. Dery, P. S. Reddy, L. Dery, R. Mousa, R. N. Dardashti, N. Metanis, *Chem. Sci.* **2015**, *6*, 6207.

[51]   N. J. Mitchell, S. S. Kulkarni, L. R. Malins, S. Wang, R. J. Payne, *Chem. Eur. J.* **2017**, *23*, 946.

[52]   S. D. Townsend, Z. Tan, S. Dong, S. Shang, J. A. Brailsford, S. J. Danishefsky, *J. Am. Chem. Soc.* **2012**, *134*, 3912.

[53]   D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2006**, *45*, 3985.

[54]   R. E. Thompson, X. Liu, N. Alonso-Garcia, P. J. B. Pereira, K. A. Jolliffe, R. J. Payne, *J. Am. Chem. Soc.* **2014**, *136*, 8161.

[55]   R. E. Thompson, X. Liu, J. Ripoll-Rozada, N. Alonso-Garcia, B. L. Parker, P. J. B. Pereira, R. J. Payne, *Nat. Chem.* **2017**, *9*, 909.

[56]   D. T. Flood, J. C. J. Hintzen, M. J. Bird, P. A. Cistrone, J. S. Chen, P. E. Dawson, *Angew. Chem. Int. Ed.* **2018**, 10.1002/anie.201805191.

[57]   K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto, A. Otaka, *Chem. Bio. Chem.* **2011**, *12*, 1840.

[58]   J.-X. Wang, G.-M. Fang, Y. He, D.-L. Qu, M. Yu, Z.-Y. Hong, L. Liu, *Angew. Chem. Int. Ed.* **2015**, *54*, 2194.

[59]   N. Naruse, K. Ohkawachi, T. Inokuma, A. Shigenaga, A. Otaka, *Org. Lett.* **2018**, *20*, 2449.

[60]   N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain, O. Melnyk, *Org. Lett.* **2010**, *12*, 5238.

[61]   E. Boll, H. Drobecq, N. Ollivier, A. Blanpain, L. Raibaut, R. Desmet, J. Vicogne, O. Melnyk, *Nat. Protoc.* **2015**, *10*, 269.

[62]   S. Tang, L.-J. Liang, Y.-Y. Si, S. Gao, J.-X. Wang, J. Liang, Z. Mei, J.-S. Zheng, L. Liu, *Angew. Chem. Int. Ed.* **2017**, *56*, 13333.

[63]   H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara, Y. Nakahara, *Tetrahedron Lett.* **2007**, *48*, 25.

[64]   L. A. Elrich, K. S. A. Kumar, M. Haj-Yahya, P. E. Dawson, A. Brik, *Org. Biomol. Chem.* **2010**, *8*, 2392.

[65]   J.-X. Wang, G.-M. Fang, Y. He, D.-L. Qu, M. Yu, Z.-Y. Hong, L. Liu, *Angew. Chem. Int. Ed.* **2015**, *54*, 2194.

[66]   D. Macmillan, A. Adams, B. Premdjee, *Isr. J. Chem.* **2011**, *51*, 885.

[67]   H. Chen, Y. Xiao, N. Yuan, J. Weng, P. Gao, L. Breindel, A. Shekhtman, Q. Zhang, *Chem. Sci.* **2018**, *9*, 1982.

[68]   D. Crich, K. Sana, *J. Org. Chem.* **2009**, *74*, 3389.

[69]   W. Hou, L. Liu, X. Zhang, C. Liu, *R. Soc. Open Sci.* **2018**, *5*, 172455.

[70]   T. Durek, P. F. Alewood, *Angew. Chem. Int. Ed.* **2011**, *50*, 12042.

[71]   K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, *Angew. Chem. Int. Ed.* **2011**, *123*, 6261.

[72]   I. Sakamoto, K. Tezuka, K. Fukae, K. Ishii, K. Taduru, M. Maeda, M. Ouchi, K. Yoshida, Y. Nambu, J. Igarashi, N. Hayashi, T. Tsuji, Y. Kajihara, *J. Am. Chem. Soc.* **2012**, *134*, 5428.

[73]   J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem. Int. Ed.* **2006**, *45*, 1248.

[74]   L. Ju, J. W. Bode, *Org. Biomol. Chem.* **2009**, *7*, 2259.

[75]   S. I. Medina, J. Wu, J. W. Bode, *Org. Biomol. Chem.* **2010**, *8*, 3405.

[76]   H. Tokuyama, T. Kuboyama, A. Amano, T. Yamashita, T. Fukuyama, *Synthesis* **2000**, 1299.

[77]   V. R. Pattabiraman, A. O. Ogunkoya, J. W. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 5114.

[78]   J. W. Bode, *Acc. Chem. Res.* **2017**, *50*, 2104.

[79]   A. O. Ogunkoya, V. R. Pattabiraman, J. W. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 9693.

[80]   C. He, S. S. Kulkarni, F. Thuaud, J. W. Bode, *Angew. Chem. Int. Ed.* **2015**, *54*, 12996.

[81]   X. Li, H. Y. Lam, Y. Zhang, C. K. Chan, *Org. Lett.* **2010**, *12*, 1724.

[82]   Y. Zhang, C. Xu, H. Y. Lam, C. L. Lee, X. Li, *Proc. Natl Acad. Sci. USA* **2013**, *110*, 6657.

[83]   T. Li, H. Liu, X. Li, *Org. Lett.* **2016**, *18*, 5944.

[84]   C. L. Lee, H. Liu, C. T. T. Wong, H. Y. Chow, X. Li, *J. Am. Chem. Soc.* **2016**, *138*, 10477.

[85]   C. T. T. Wong, H. Y. Lam, T. Song, G. Chen, X. Li, *Angew. Chem. Int. Ed.* **2013**, *52*, 10212.

[86]   L. Abrahmsén, J. Tom, J. Burnier, K. A. Butcher, A. Kossiakoff, J. A. Wells, *Biochemistry* **1991**, *30*, 4151.

[87]   T. K. Chang, D. Y. Jackson, J. P. Burnier, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12544.

[88]   D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243.

[89]   A. M. Weeks, J. A. Wells, *Nat. Chem. Bio.* **2018**, *14*, 50.

[90]   S. H. Henager, N. Chu, Z. Chen, D. Bolduc, D. R. Dempsey, Y. Hwang, J. Wells, P. A. Cole, *Nat. Methods* **2016**, *13*, 925.

[91]   O. Schneewind, A. Fowler, K. F. Faull, *Science* **1995**, *268*, 103.

[92]   L. Schmohl, D. Schwarzer, *Curr. Op. Chem. Biol.* **2014**, *22*, 122.

[93]   M. Ritzefeld, *Chem. Eur. J.* **2014**, *20*, 8516.

[94]   M. W.-L. Popp, H. L. Ploegh, *Angew. Chem. Int. Ed.* **2011**, *50*, 5024.

[95]   G. K. T. Nguyen, S. Wang, Y. Qiu, X. Hemu, Y. Lian, J. P. Tam, *Nat. Chem. Biol.* **2014**, *10*, 732.

[96]   G. K. T. Nguyen, Y. Qiu, Y. Cao, X. Hemu, C.-F. Liu, J. P. Tam, *Nat. Protoc.* **2016**, *11*, 1977.

[97]   G. K. T. Nguyen, A. Kam, S. Loo, A. E. Jansson, L. X. Pan, J. P. Tam, *J. Am. Chem. Soc.* **2015**, *137*, 15398.

[98]   G. K. T. Nguyen, Y. Cao, W. Wang, C. F. Liu, J. P. Tam, *Angew. Chem. Int. Ed.* **2015**, *54*, 15694.

[99]   Y. Cao, G. K. T. Nguyen, J. P. Tam, C.-F. Liu, *Chem. Commun.* **2015**, *51*, 17289.

[100] C. Janke, *J. Cell. Biol.* **2014**, *206*, 461.

[101] D. Schumacher, J. Helma, F. A. Mann, G. Pichler, F. Natale, E. Krause, M. C. Cardoso, C. P. R. Hackenberger, H. Leonhardt, *Angew. Chem. Int. Ed.* **2015**, *54*, 13787.

[102] D. Schumacher, O. Lemke, J. Helma, L. Gerszonowicz, V. Waller, T. Stoschek, P. M. Durkin, N. Budisa, H. Leonhardt, B. G. Keller, C. P. R. Hackenberger, *Chemical Science* **2017**, *8*, 3471.

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

[104] N. H. Shah, T. W. Muir, *Chem. Sci.* **2014**, *5*, 446.

[105] C. R. Noren, J. Wang, F. B. Perler, *Angew. Chem. Int. Ed.* **2000**, *39*, 450.

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

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

[108] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.

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