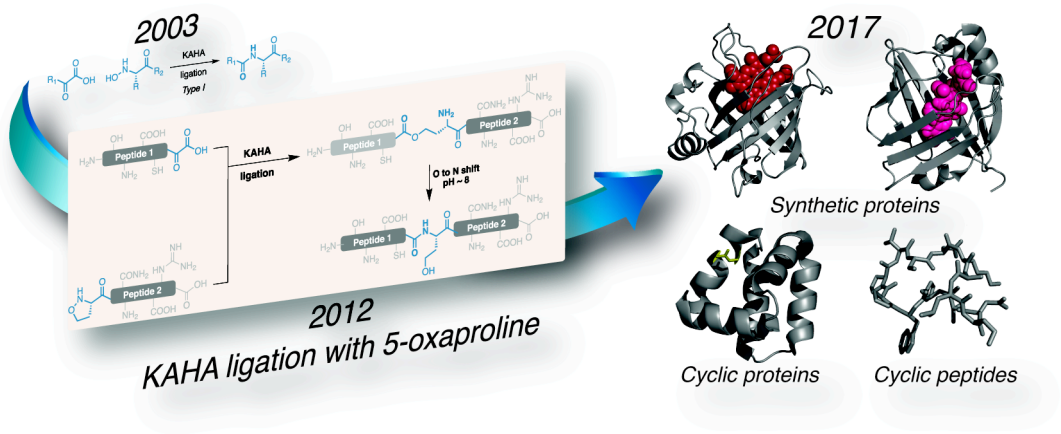
Chemical Protein Synthesis with the α-Ketoacid-Hydroxylamine Ligation

Originally published in:
Accounts of Chemical Research 50(9), https://doi.org/10.1021/acs.accounts.7b00277

Chemical Protein Synthesis with the α-Ketoacid-Hydroxylamine (KAHA) Ligation

Jeffrey W. Bode*
Laboratorium für Organische Chemie,
Department of Chemistry and Applied Biosciences,
ETH-Zürich, 8093 Zürich, Switzerland
bode@org.chem.ethz.ch

TOC Graphic



Conspectus

The coupling of an α-ketoacid and a hydroxylamine (KAHA ligation) affords amide bonds under aqueous, acidic conditions without the need for protecting groups or coupling agents. Translating this finding into a general approach to chemical protein synthesis required the identification of methods to incorporate the key functional groups into unprotected peptide segments – ideally using well established Fmoc-solid phase peptide synthesis protocols. A decade of effort has now lead to robust, convenient methods for preparing peptide bearing free or masked C-terminal α-ketoacids and N-terminal hydroxylamines. The facile synthesis of the segments and the aqueous, acidic conditions of the KAHA ligation make it ideal for the construction of small proteins (up to 200 residues), including SUMO and related modifier proteins, betatrophin and other protein hormones, nitrophorin 4, S100A4, and the cyclic protein AS-48.

Key to the successful development of this protein synthesis platform was the identification and gram-scale synthesis of (S)-5-oxaproline. This hydroxylamine monomer is completely stable to standard methods and practices of solid phase peptide synthesis while still performing very well in

the KAHA ligation. This reaction partner – in contrast to all other examined – affords *esters*, rather than amides as the primary ligation product. The resulting *depsi*-peptides often offer superior solubility and handling and have been key in the chemical synthesis of hydrophobic and ampiphilic proteins. Upon facile O-to–N acyl shift, peptides bearing a non-canonical homoserine residue at the ligation site are formed. With proper choice the ligation site, the incorporation of this unnatural amino acid does not appear to affect the structure or biological activity of the protein targets.

The development of the chemical methods for preparing and masking peptide α-ketoacids and hydroxyalmines, the preparation of several protein targets by convergent ligation strategies, and the synthesis of new hydroxylamine monomers affording either natural or unnatural residues at the ligation site are discussed. By operating under acidic conditions and with distinct preference for the ligation site, these efforts establish KAHA ligation as a complementary method to the venerable native chemical ligation (NCL) for chemical protein synthesis.

This account documents both the state of the KAHA ligation and the challenges in identifying, inventing, and optimizing new reactions and building blocks needed to interface KAHA ligation with Fmoc-solid phase peptide chemistry. With these challenges largely addressed, peptide segments ready for ligation are formed directly upon resin cleavage, facilitating rapid assembly of four to five segments into proteins. This work sets the stage for applications of the KAHA ligation to chemical biology and protein therapeutics.

1. Introduction

At the beginning of 2011, we had a problem: our flagship research project, the α-ketoacid–hydroxylamine (KAHA) ligation (had stalled Figure 1a).[1] The amide-forming reaction itself worked perfectly fine, but our goal of developing it into a method for chemical protein synthesis seemed forever out of reach. We had devised a robust method for making most peptide segments with C-terminal α-ketoacids,[2,3] but we struggled with both the synthesis and stability of the N-terminal peptide hydroxylamines.[4] After exhausting many different hydroxylamine derivatives, several different protecting group strategies, and increasingly desperate purification protocols, it was time to face the fact that many peptide hydroxylamines may not be sufficiently stable for use in protein chemical synthesis.

To complicate matters, we had just completed our move from the University of Pennsylvania to ETH Zürich. We had new labs, a motivated group of talented students and postdocs, the chance to invest in new equipment and research directions. The KAHA team, once the largest part of our research group, had dwindled to two group members– Ayo Ogunkoya and Dr. Vijaya Pattabiraman – making a last stand at peptide chemistry. We had retreated from protein synthesis entirely, and Vijay and Ayo were back to making dipeptides. A third member, Ivano Pusterla, was wrapping up mechanistic studies on the KAHA ligation that, while fascinating, suggested that the unique mechanism of the KAHA ligation might make it ill suited for protein synthesis (Figure 1b).[5] He was looking for a new project – probably on one of the other more promising topics in our group.

Ayo and Vijay were examining various types of peptide hydroxylamines, most of which had problems. Type I – *O*-unprotected hydroxylamines – displayed very good reactivity and stability in KAHA ligations with simple substrates. But when installed onto an α-amino acid, these hydroxylamines were prone to oxidation to the corresponding oxime. Furthermore, Ivano's mechanistic studies on the ligation with Type I hydroxylamines revealed the formation of a nitrone intermediate was essential, a process likely to be slowed in the presence of water. Type II hydroxylamines – those bearing an *O*-acyl group on the hydroxylamine oxygen – proceeded very well in water (Figure 1c) but these more activated substrates were prone to elimination when constructed from α-amino acids.[6] Their cousins, *O*-alkyl hydroxylamines, were completely stable but reluctant to undergo KAHA ligation.

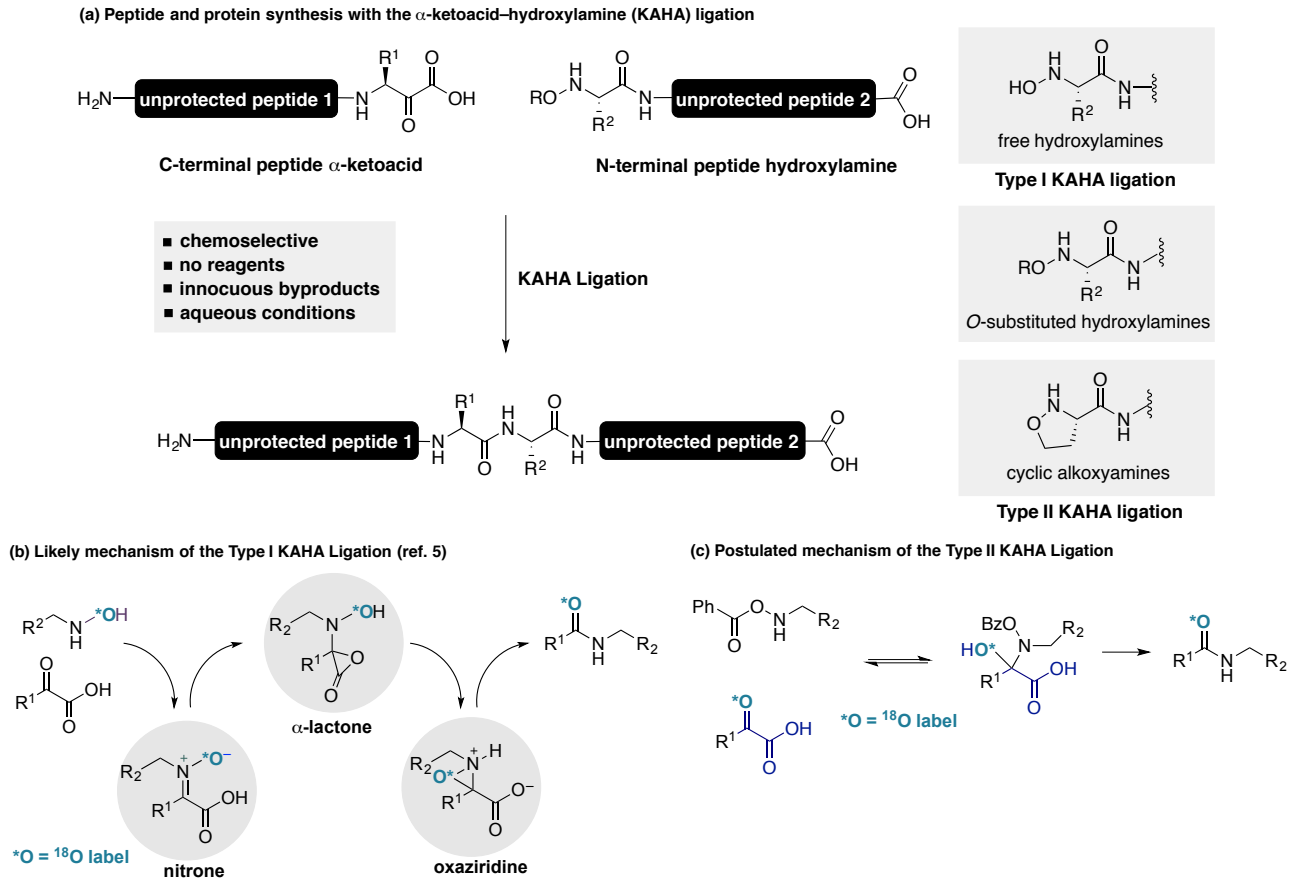


Figure 1. (a) Chemoselective coupling of unprotected peptide segments with the α-ketoacid–hydroxylamine (KAHA) ligation. The KAHA ligation works with either substituted or unsubstituted hydroxylamines by distinct mechanisms. Cyclic hydroxylamines, particularly (*S*)-5-oxaproline, have proven to be the best combination of stability, reactivity, and synthetic access. (b) Likely mechanism of the Type 1 KAHA ligation. (c) Anticipated (but still unproven) mechanism of the Type II KAHA ligation with O-acyl hydroxylamines.

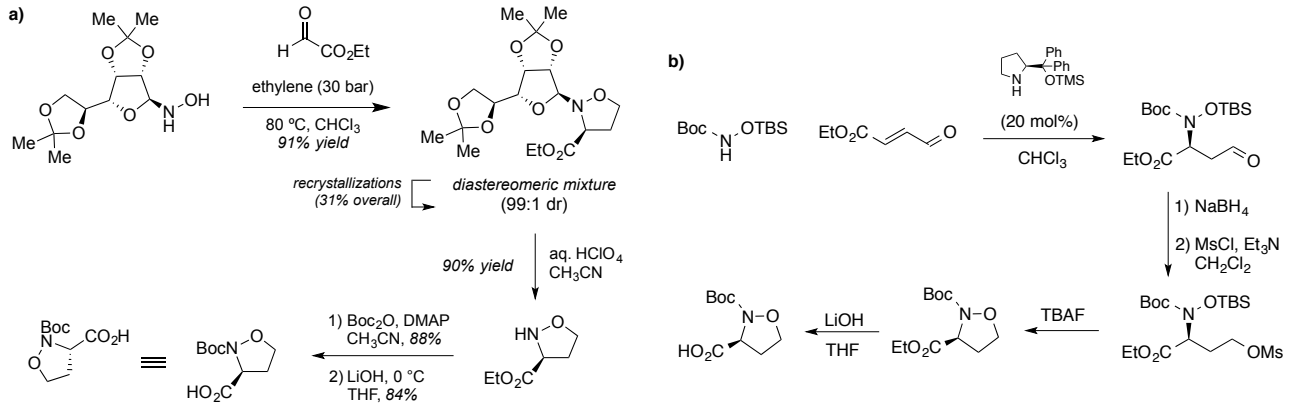
2. 5-Oxaproline

From our work on β-peptide synthesis with the KAHA ligation, we were aware of another class of hydroxylamines – isoxazolidines.[7] These remarkably stable and easily prepared compounds underwent KAHA ligation under aqueous conditions. Their main limitation – at least in the context of protein synthesis – was the need for higher reaction concentrations (typically 200 millimolar!). Although we had proposed such cyclic hydroxylamines for peptide synthesis years before, no one had tried them – probably because the PI kept saying that even if they worked in principle, they would not operate at the lower concentrations needed for protein synthesis.

Without my knowledge, Ayo began to examine the cyclic hydroxylamine 5-oxaproline for KAHA ligation. Its synthesis had been reported in 1981 by my ETH colleague Andrea Vasella,[8] using the sugar derived chiral auxiliary approach that we employed for preparing the β-peptide monomers. The synthesis required a high-pressure cycloaddition with ethylene, gave a mixture of diastereoisomers, and could only be run on limited scale with the equipment available to us

(Scheme 1a). Despite this, Ayo scrapped out enough of the cycloaddition product to couple it to a small peptide and perform the first successful ligation with (*S*)-5-oxaproline. The (*S*)-Boc-5-oxaproline was easily installed onto the N-terminus of a solid supported peptide and the cleaved, unprotected peptide completely stable to purification – a nice contrast to our experience with almost every other type of hydroxylamine. To the PI's surprise, the ligations proceeded well at lower concentrations (10–20 millimolar) than expected.

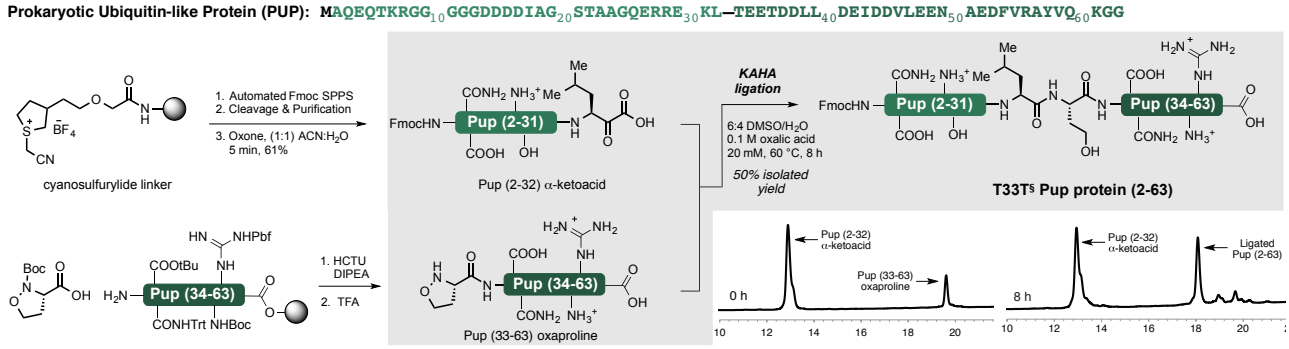
The lengthy, poorly selective, and unscalable route to (*S*)-5-oxaproline initially limited the amount we could produce, graduate student Claudia Murar and Thibault Harmand eventually devised a new, scalable, organo-mediated approach that could be easily executed on a decagram scale (Scheme 1b). This route has been used to prepare kilogram quantities of Boc-5-(*S*)-oxaproline.



Scheme 1. The original and revised route to Boc-(*S*)-5-oxaproline.

3. Protein Syntheses with 5-Oxaproline

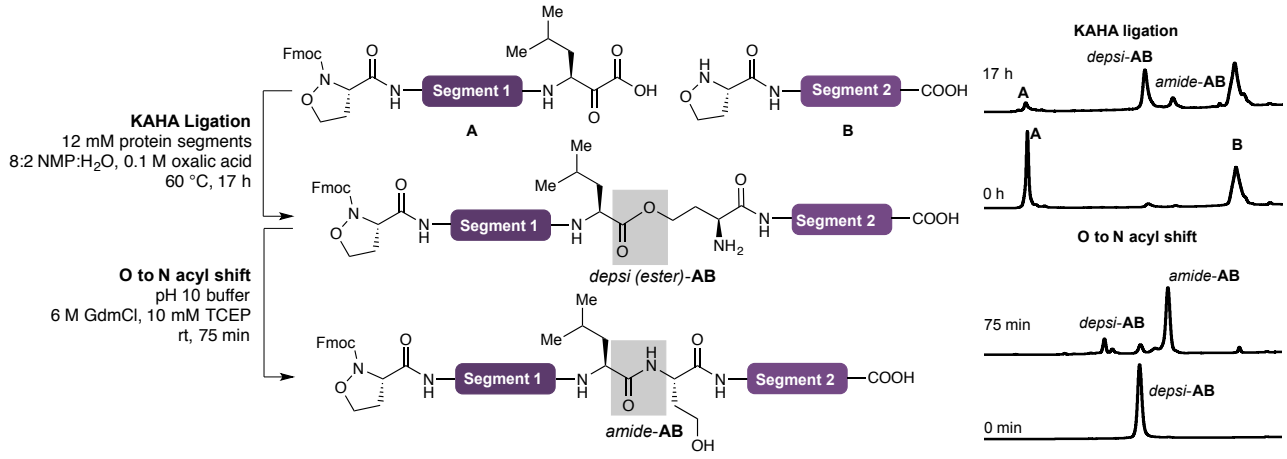
The product of this ligation – the non-canonical amino acid homoserine – was not our first choice. But after years of work to establish a route to the peptide α-ketoacids we were eager to establish that KAHA ligation could really be used to make longer peptides and small proteins. As a first target, we chose the synthesis of Pup – the modifier protein from *Prokaryotes* that fulfills the role of ubiquitin in higher organism.[9] This 63 residue peptide could be split into two segments of about 30 residues each, which smoothly ligated to give the desired product (Scheme 2).[10] Vijay followed it up with the synthesis of CspA, another small protein. Both proteins, when placed in folding buffer, adopted the expected conformation as judged by circular dichorism and were further characterized by mass spectrometry. The use of 5-oxaproline for the first protein syntheses with KAHA ligation made a nice communication and got our protein synthesis project back on track.



Scheme 2. Synthesis of Pup, a small modifier protein from *Mycobacterium*, by KAHA ligation of two unprotected segments.

4. Ester-formation in KAHA Ligation with 5-Oxaproline

Reinvigorated by this success, we quickly started several new protein synthesis projects. In addition to learning that solid phase synthesis of long (>40 residue) peptide segments on hundred milligram scales was anything but easy, we soon made a shocking discovery about the KAHA ligation with 5-Opr. During the course of investigations, particular on cyclic peptides and one of our more challenging protein targets, we occasionally noticed small peaks in the HPLC traces with the identical mass of our ligation products. A thorough and systematic evaluation by graduate students Florian Rohrbacher and Thomas Wucherpfenning soon revealed a surprise – the primary ligation products were esters, not amides (Scheme 3)![11]



Scheme 3. Esters – rather than the expected amides – are the primary products of KAHA ligation with (*S*)-5-oxaproline. In analogy to Kiso's isopeptides, these esters readily rearrange to the amide products under basic conditions. The data shown is from a sequence in the chemical synthesis of irisin.

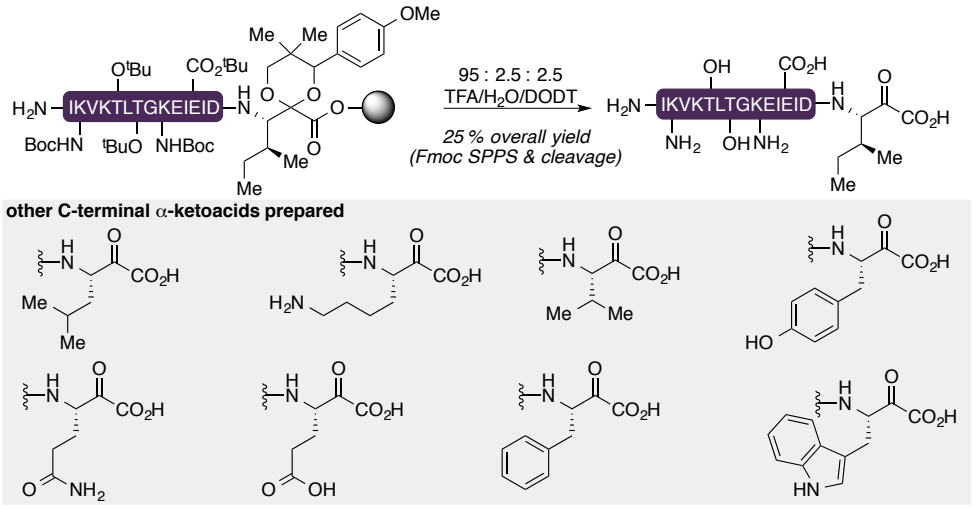
We had completely missed the formation of esters in our first reports on protein synthesis. As very well known from the beautiful work of Kiso on iso-acyl peptides,[12] these esters readily rearrange to amides under slightly basic conditions. As soon as we exposed our ligation products to folding buffer, they underwent O-to-N shift and the proteins adopted the correct secondary and tertiary

structures. This finding also led to new insights on the mechanism of the KAHA ligation. We knew already that the mechanisms of Type I and Type II KAHA ligations were different; the ester formations suggested that KAHA ligations with 5-Opr proceeded via the formation of a nitrilium intermediate.[13] This is trapped by the homoserine oxygen to give an iminoether, which under acidic conditions is known to open to the ester.[14] The trace amount of amide formed during the KAHA ligations arises from direct interception of the nitrilium by water.

5. Traceless Synthesis of Peptide α-Ketoacids

At this point in our studies, we still faced one major synthetic challenge: the preparation of peptide α-ketoacids bearing Cys and Met residues. Although our cyanosulfurylide chemistry works remarkably well for peptides that lack these residues, it was still a two-step protocol that complicated the use of KAHA ligation in applications such as the synthesis of macrocyclic peptides. We had long sought an acid-labile α-ketoacid protecting group stable to the basic conditions of Fmoc-peptide synthesis and removable without epimerization. In principle, we knew how to do this. In practice, we could not synthesize the necessary monomers – the chemistry simply did not exist.

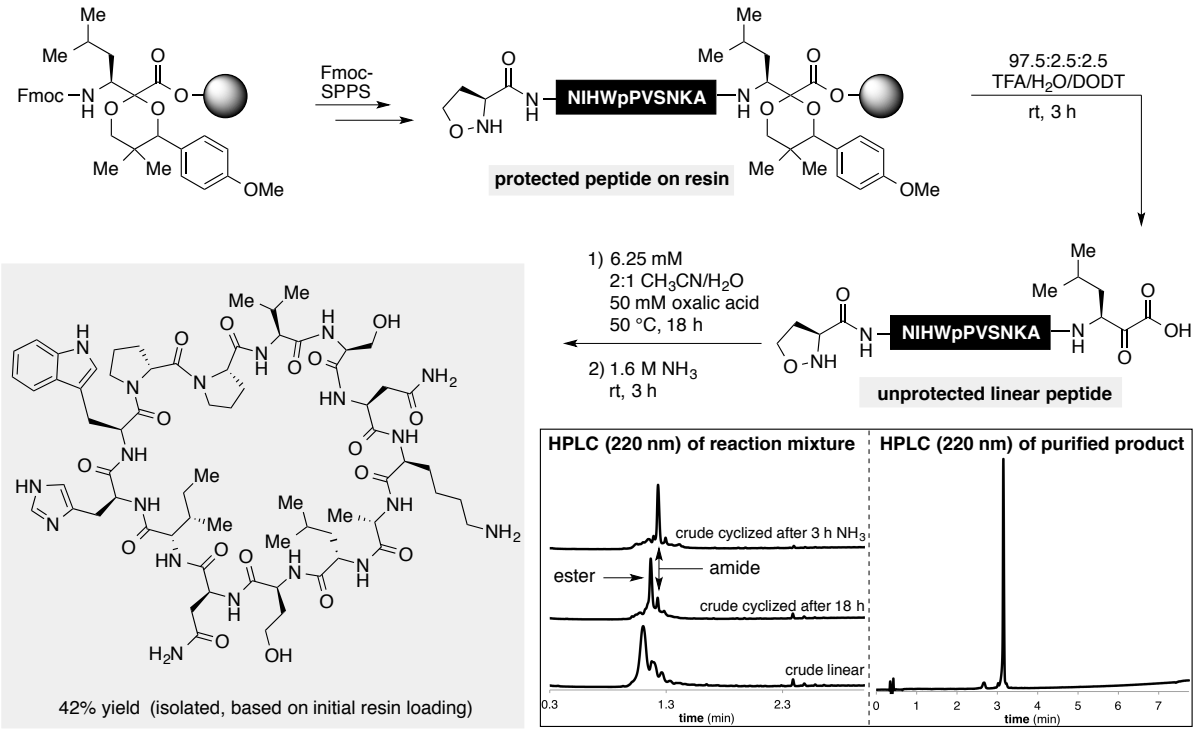
Efforts begun with Javier Ruiz-Rodriguez, advanced by MSc student Felix Limberg, finished by Thomas Wucherpfennig, and refined by Frédéric Thuaud, André Zwicky, and Florian Rohrbacher finally led to a nearly perfect solution (Scheme 4).[15] Felix developed a new protecting group for peptide α-ketoacids that could be removed – without epimerization – under conditions used for cleavage of peptides from standard resins. This worked brilliantly – but actually getting the protecting group onto the amino acid proved remarkably difficult. Our first approach – Mukaiyama acetalization – was low yielding and applicable only to Leu and Phe derivatives. Despite this, it was a critical advance in our research program. Diligent studies and optimization by Fred and Florian led to a new transacetalization protocol that not only improved yields and scalability but also extended this approach to many more α-ketoacids.[16] These are now routinely prepared on 20+ gram scale both in our lab and by industrial partners.



Scheme 4. Preparation of C-terminal peptide α-ketoacids with acid-labile, protected monomers.

6. Direct Synthesis of Macrocyclic Peptides

Peptide macrocycles are of enormous interest as platforms for modern pharmaceutical development. They are synthesized effectively by an established, but laborious protocol – Fmoc-SPPS on Cl-Trityl resin, cleavage to give the partially protected peptide, purification, cyclization, side-chain deprotections, and another purification. To simplify the workflow and to examine the utility of KAHA ligation for preparing peptide macrocycles, we joined forces with the Swiss biotech company Polyphor. Working with Dr. Gildas Deniau and Dr. Anatol Luther, Florian took advantage of the new α-ketoacid-forming linkers to devise a method to form peptide macrocycles directly upon resin cleavage.[17] The cleaved peptide, containing both an unprotected α-ketoacid and a free oxaproline, is diluted with aqueous buffer to effect cyclization, followed by switch to basic pH to induce O->N acyl shift (Scheme 5). This approach works at higher concentration, requires fewer steps, and often gives superior results compared to cyclization of side chain protected peptides prepared from chlorotrityl resins, especially for unconstrained examples.

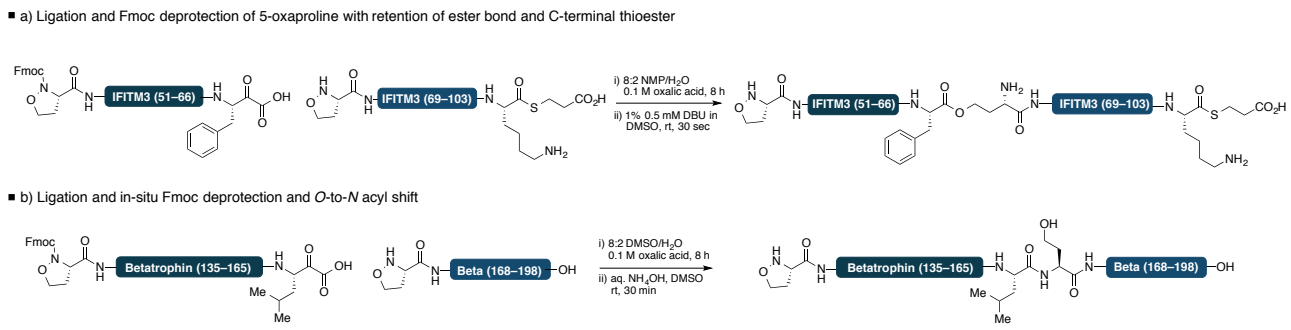


Scheme 5. Example of cyclic peptide synthesis with the KAHA ligation. The linear peptides are prepared using standard Fmoc-SPPS. Following resin cleavage, the unprotected peptides undergo chemoselective cyclization under aqueous conditions, followed by treatment with aqueous ammonia to induce O-to-N acyl shift of the intermediate *depsi*-peptide.

7. New Protecting Groups For Sequential KAHA ligations

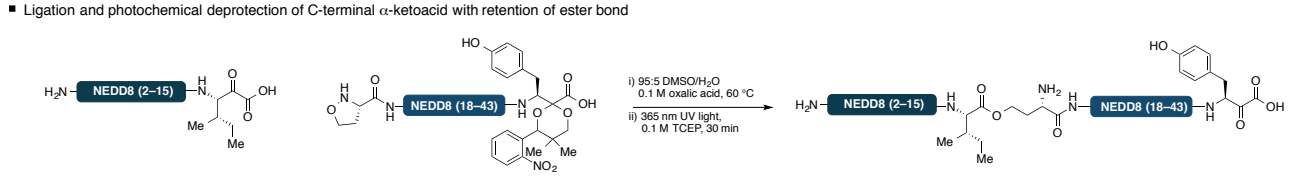
Of course, we were not content to just make small proteins by one KAHA ligation. Sequential ligation strategies require that at least one of the key functional groups be masked during the ligation and then readily revealed from the large, unprotected product. The ability to conduct KAHA ligation under acidic conditions suggested that Fmoc-protected Opr could serve as a suitable masking group, provided that we could remove it on the fully unprotected peptides formed from the ligation.

Somewhat to our surprise, this worked well.[18] Base deprotections on unprotected peptides can be problematic, but by treatment with Hünig's base or DBU in DMSO, we identified conditions that could reliably remove the Fmoc-group from an N-terminal Fmoc-Oxaproline. These conditions could chemoselectively remove the Fmoc group while leaving the ester product of KAHA ligations intact – a critical step for accomplishing the syntheses of some of the more challenging hydrophobic proteins (Scheme 6a). Alternatively, we could perform a one-step Fmoc-deprotection and acyl shift, often on the ligation mixtures directly. This proved to be the method of choice for some of the more soluble protein targets, such as betatrophin (Scheme 6b).



Scheme 6. Various conditions for Fmoc-deprotections of ligated peptide segments, with or without concomitant *O*-to-*N* acyl shift.

To increase convergence and flexibility in protein synthesis with the KAHA ligation, we required a method to mask the C-terminal α-ketoacid during the ligation. Our previously reported cyanosulfurylides can serve this role, but the oxidative conversion to the α-ketoacid is not compatible with unprotected Cys or Met residues. We therefore required a protecting group that would be stable to both acidic and basic conditions, yet still be removable in the presence of a multitude of unprotected side chains. Dr. Frédéric Thuaud came up with the idea of a photolabile protecting group, based on the approach developed for the traceless synthesis of peptide α-ketoacids.[19] By replacing the *para*-methoxyphenyl moiety of our masking group with an ortho-nitrophenyl group, Frédéric invented a protecting group that survived both Fmoc-SPPS and resin cleavage with TFA. Despite protests from the lab head that photolytic deprotection would not work on larger substrates, Fred proceeded to demonstrate that it actually works quite well by simply using a handheld UV lamp (Scheme 7). It was so convenient that it was quickly adopted by the rest of the KAHA team, who began improving upon it with enantiopure versions and the synthesis of photomasked oxaproline derivatives. The main advantage of these photoprotected groups is simple deprotection directly after ligation, which also opens the possibility of one-pot, multi-segment ligations.

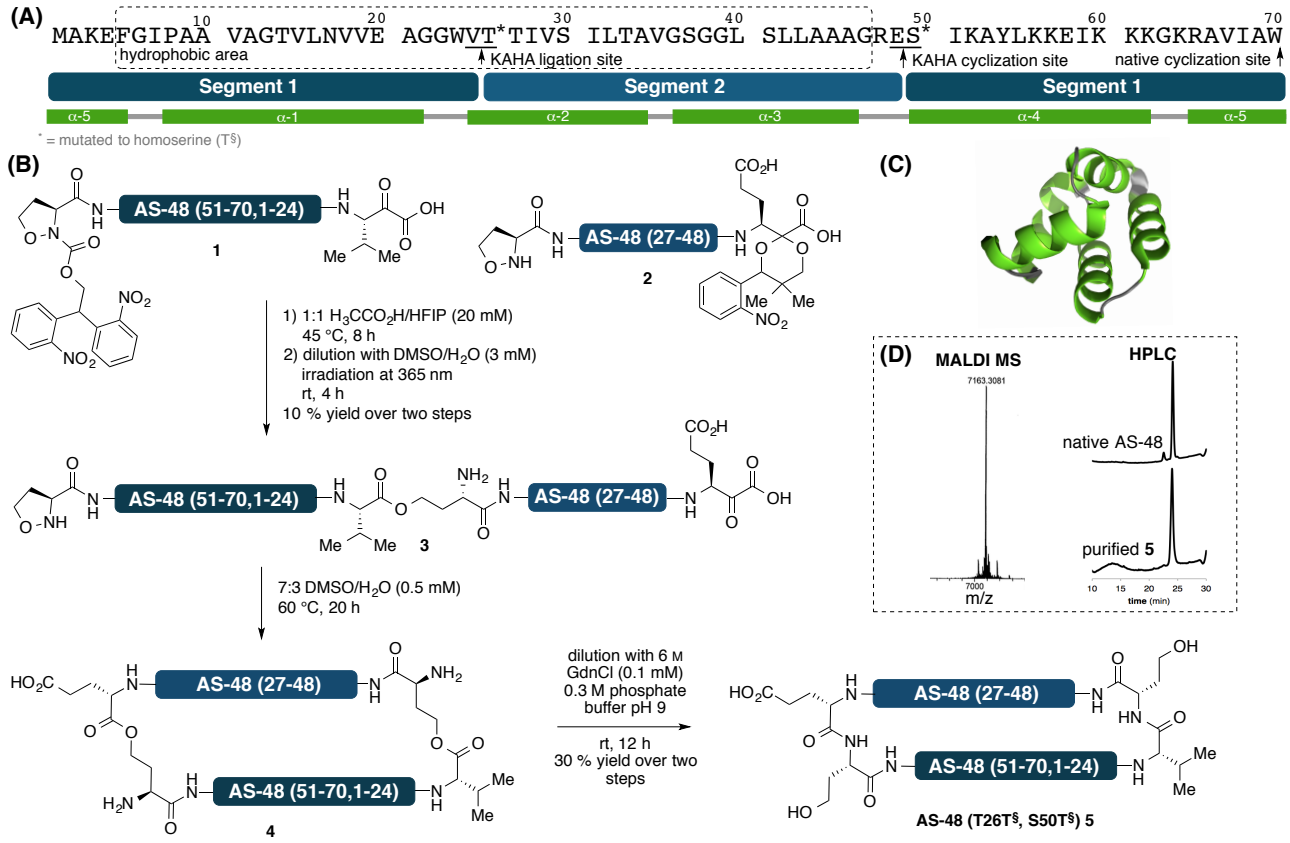


Scheme 7. Photochemical deprotection of C-terminal peptide α-ketoacids after KAHA ligation. This convenient protocol can be used for sequential and one-pot ligation strategies.

8. Synthesis of the Cyclic Protein AS-48

For at least one project, the photoprotected functional groups were essential for success. The head-to-tail cyclized protein AS-48 is the largest naturally occurring peptide macrocycle.[20] It is an amphiphilic, membrane-associate protein containing five α-helical sequences, making portions of

this molecule extremely hydrophobic and difficult to synthesize. In 2015, Tam successfully prepared it by an impressive enzyme-mediated cyclization.[21] Interestingly, he found the linear peptide needed to first be folded prior to cyclization – direct, enzymatic cyclization of the unfolded peptide did not proceed.



Scheme 8. Synthesis of AS-48 by KAHA ligation of two segments. (A) Sequence of AS-48 with division into two peptide segments. The dashed box indicates an area of mainly hydrophobic residues. α-helices are depicted in green, loop and coil regions in gray. The asterisks (*) indicate residues mutated to Hse in the synthesis by KAHA ligation. (B) Synthesis by a two-segment strategy with photoprotected 5-oxaproline segment 2 and photoprotected α−ketoacid segment 1. (C) Three dimensional structure of native AS-48. (D) MALDI mass spectrum and analytical HPLC traces at 220 nm of purified cyclic AS-48 and comparison with HPLC trace of native AS-48.

The large size, biological activity, and challenge of preparing amphiphilic peptides made AS-48 an attractive synthetic target by KAHA ligation. Florian began to synthesize the requisite fragments, but soon found that these sequences were incredibly difficult to prepare and failed unless a side-chain solubilization tag was employed. Although ultimately successful, this led to a long, low-yielding synthesis that did not deliver enough material. By recognizing that this cyclic protein can be separated into any two segments, Florian designed a new route that, while still challenging, allowed him to prepare the two segments without the need for solubilizing tags (Scheme 8). Photolabile protecting groups were installed onto both segments, which were joined by KAHA ligation in AcOH/HFIP – a modification necessary to solubilize the difficult segments and promote their union – to give the ester product.[22] Photochemical removal of the protecting groups followed

by cyclization of the unfolded protein gave the bis-depsipeptide in decent yield. Upon exposure to folding conditions, this peptide underwent two acyl shifts – one fast, one slow – to give the folded, cyclic peptide with a 11-minute shift on the HPLC! Isolation of this product, followed by further folding, gave synthetic AS-48 that showed nearly identical CD spectra and biological activity as a sample of the natural product kindly provided to us by Prof. Mercedes Maqueda.

9. The "Homoserine" Problem and Alternative Hydroxylamines

As detailed in the subsequent section, protein synthesis using KAHA ligation with oxaproline works extremely well. Preparation of the peptide segments bearing the α-ketoacids or oxaproline – in both free and unprotected form – is now routine and traceless. The ligations proceed in a combination of organic solvent (DMSO or NMP) and aqueous acid (usually oxalic acid), conditions that excel at solubilizing even hydrophobic segments. We typically obtain 50-60% yields of the ligation products after isolation by preparative HPLC. The formation of the ester, rather than the amide, as the primary product is more often a blessing than a curse. Just one thing keeps it from being perfect – the introduction of the non-canonical residue homoserine at the ligation site.

There are two schools of thought. The first insists that we must always strive to prepare the natural proteins, with the identical sequences to the natural ones. Chemical synthesis is a welcome and necessary approach to preparing proteins with posttranslational modifications or unnatural residues, but preferably not at the ligation site. The second recognizes that mutation of most natural residues to a homoserine is unlikely to dramatically affect the biologically activity of protein, particularly for applications in chemical biology. We currently follow the second school of thought and routinely mutate Ser, Thr, Asn, Asp, Met, Glu and other residues in proteins without noticing effects on structure or biological activity (Figure 2). At the same time, we have pursued the design and synthesis of monomers that afford canonical residues at the ligation site.

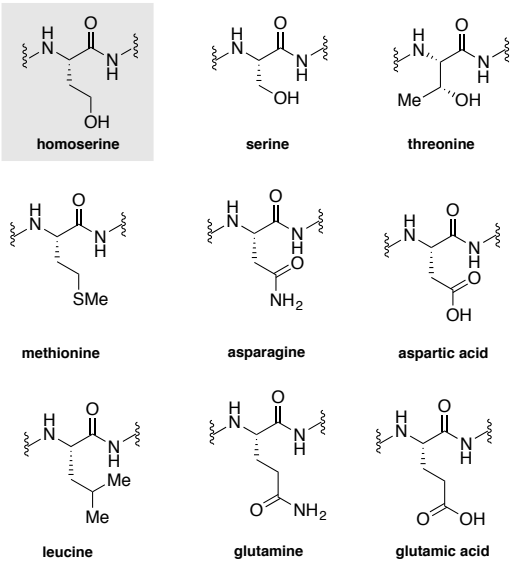
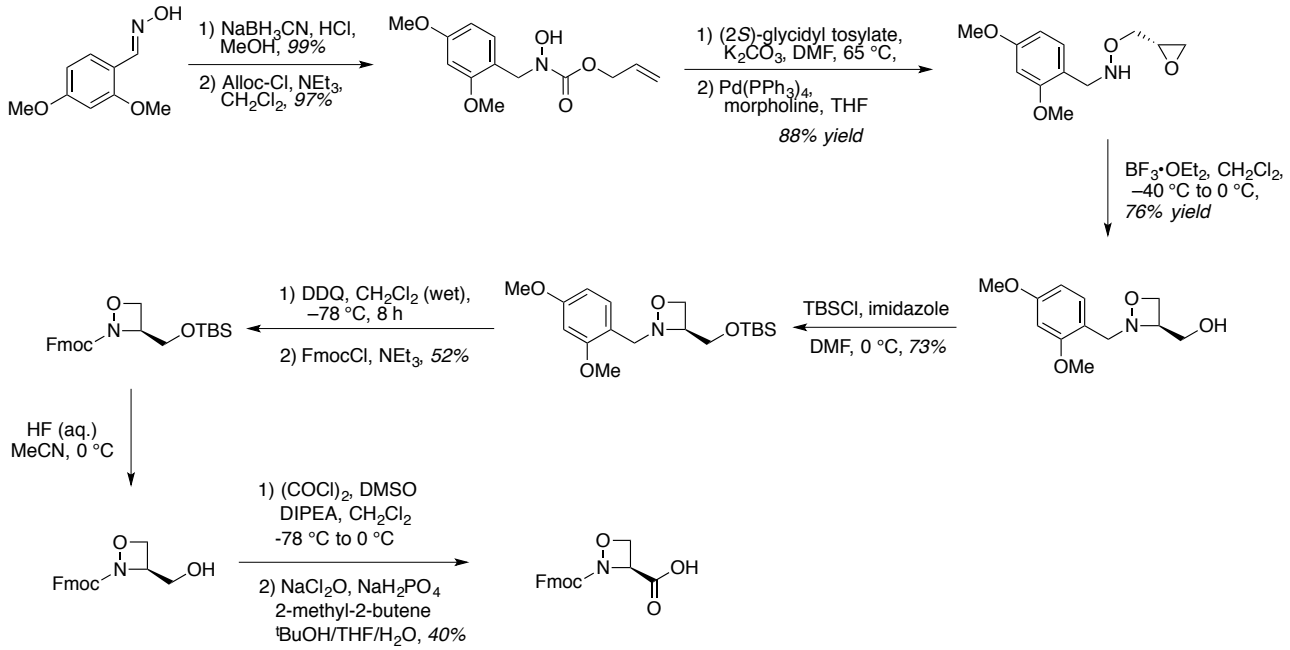


Figure 2. Typical amino acids mutated to homoserine during the course of protein synthesis with 5-oxaproline.

9. 1 Serine-forming Oxazetidines

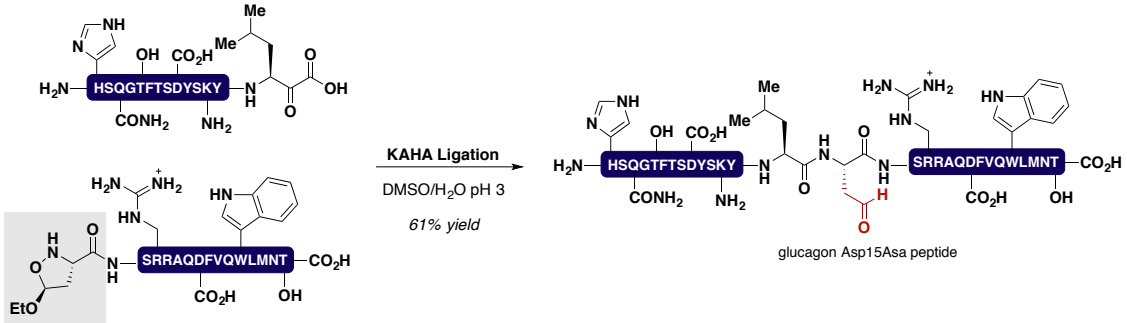
The most obvious candidate for a KAHA ligation monomer that affords a natural residue at the ligation site is an oxazetidine – a four membered hydroxylamine. These structures were, however, almost completely unknown and calculated to be highly strained.[23] Despite having little information about their stability or performance in the KAHA ligation, Ivano Pusterla bravely set out on what became a nearly four year quest to prepare and utilize the serine-forming oxazetidine. Space precludes a full account of the struggles and frustrations of his ultimately successful synthesis of enantiopure (*S*)-Fmoc-oxazetidine and its application to the synthesis of the protein S100A4, but his overall route is shown in Scheme 9.[24] Amazingly, the Fmoc-protected oxazetidine is completely stable under strong acid and other harsh conditions; the unprotected form, unfortunately, decomposes and needs to be used immediately. This has deterred further applications, but we now have alternative threonine and serine-forming monomers that appear to be much more stable.



Scheme 9. Synthesis of (*S*)-Fmoc-oxazetidine.

9.2 Aldehyde-forming Monomer

In parallel, PhD student Claudia Murar embraced the philosophy that ligation monomers that introduce unique chemical handles for late-stage functionalization would be powerful methods and devised a monomer that results in the formation on an aspartic acid semi-aldehyde residue at the ligation site (Scheme 10).[25] This strategy, which works just as well as our standard (*S*)-5-oxaproline residue, was applied to aldehyde-containing forms of glucagon and ubiquitin.



Scheme 10. Synthesis of aspartyl semi-aldehyde containing peptides by KAHA ligation with ethoxy-oxaproline peptides.

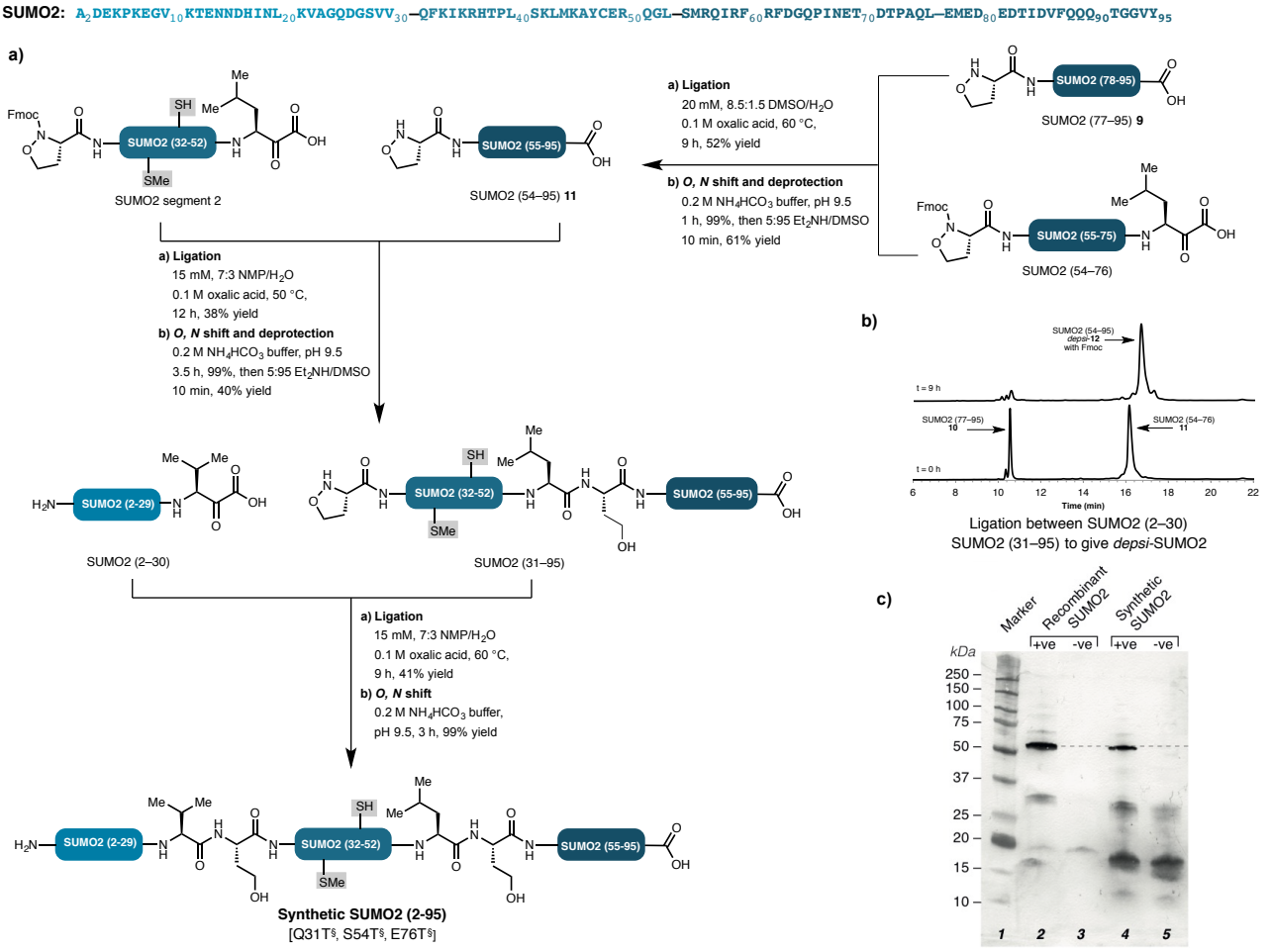
10. Protein Synthesis with the KAHA Ligation

From mid-2013, we were finally ready to get serious about protein synthesis. We had established all the requisite building blocks and developed approaches for iterative segment ligations. The question remained, however, of whether or not KAHA ligation was really suited for the synthesis of

proteins >100 residues. We selected four proteins as synthetic targets: SUMO 2/3, IFITM3, □□□□trophin, and nitrophorin 4 and detail their synthesis below. In the mean time, many other proteins, including irisin, NEDD8, Unagi Fluorescent Protein, interleukin-2, and UBC9 have been prepared by KAHA ligation.

10.1 SUMO 2/3

Early in our efforts, we chose the modifier protein SUMO 2/3 as a synthetic target, for two primary reasons. First, it contains a Cys residue near the middle of the sequence, requiring us to have a non-oxidative route to the central α-ketoacids. Second, a number of easy and well-established assays would allow us to examine the effect of mutating two or three residues to homoserines.

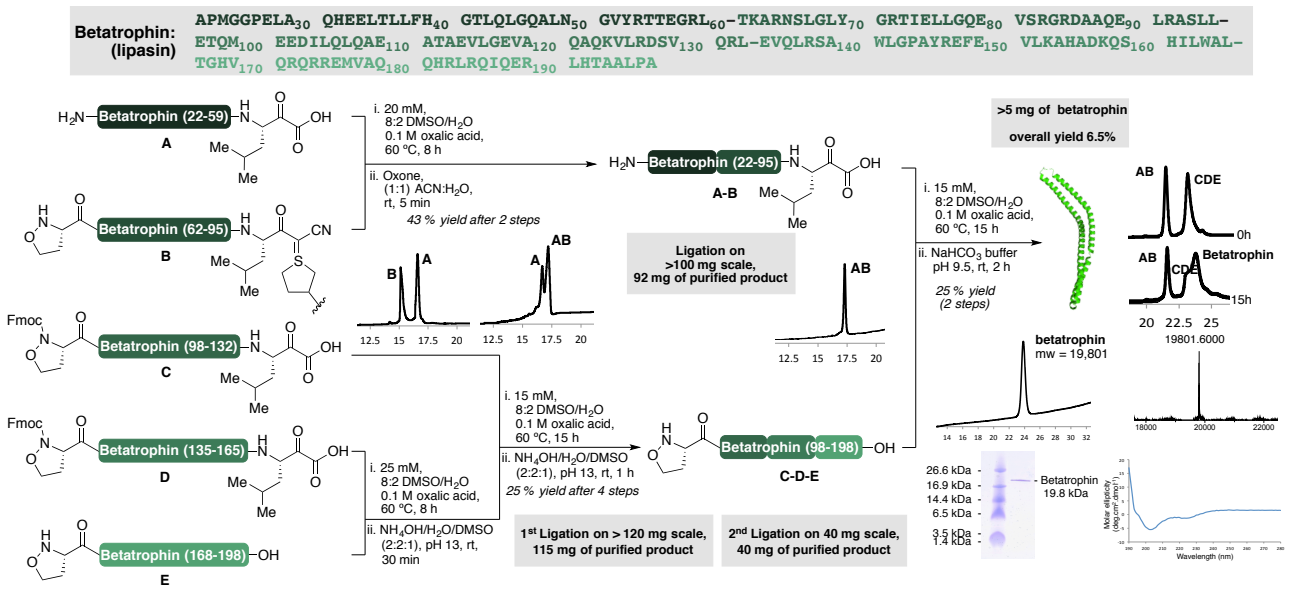


Scheme 11. Chemical synthesis and biochemical characterization of SUMO 2 prepared by KAHA ligation. a) Synthesis of SUMO2 by a four segment, three ligaiton strategy. b) RP-HPLC analysis of the final ligation. c) Biochemical characterization of synthetic SUMO2 in a protein conjugation assay.

The synthesis itself, shown in Scheme 11, was relatively straightforward and delivers milligram quantities of the SUMO protein, often with dyes, tags or other modifications for ongoing chemical biology applications.

10.2 Betatrophin

The high profile report in 2013 of a protein hormone, named betatrophin that promoted the growth of β-cells attracted worldwide attention, including ours.[26] This 179-residue protein contained no Cys residues and, from its primary sequence, seemed unlikely to have a discrete fold. It looked, to us, like an ideal target for chemical synthesis and, eventually, for introducing synthetic modifications to develop more stable and active variants for pharmaceutical applications. Two first year graduate students, Claudia Murar and Thibault Harmand, took on the synthesis of betatrophin as a "side project". In a relatively short time, they were able to prepare the key segments and deliver the first amounts of this protein. Some minor modifications to the order of the ligations and protecting group strategy lead to an easy and effective synthesis that delivered milligram quantities of pure betatrophin from modest amounts of the starting peptide segments (Scheme 12).



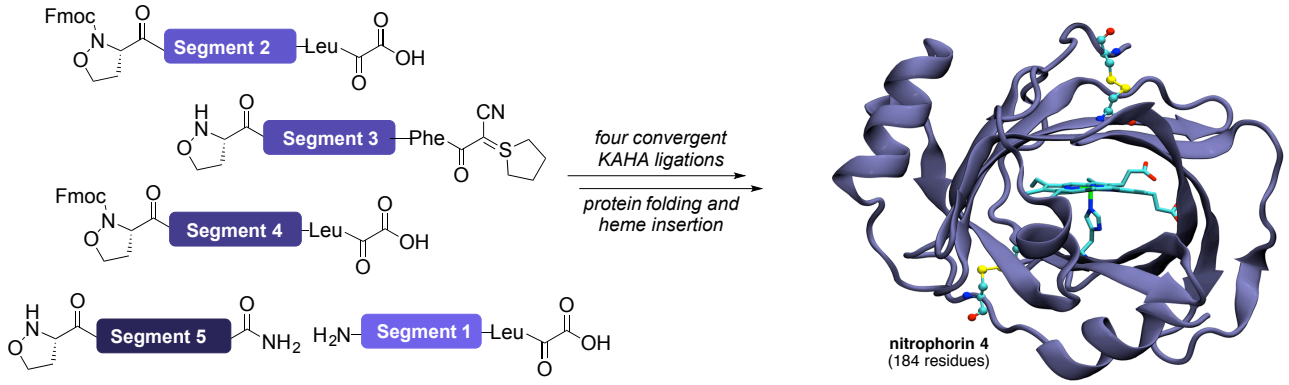
Scheme 12. Chemical synthesis of betatrophin by convergent KAHA ligations of five segments.

Our excitement at this facile and quite high yielding synthesis was tempered as report after report emerged questioning the proposed biologically activity of betatrophin. Although we had more than 10 milligrams of the synthetic protein in hand, we lacked any working assay to confirm its biological activity. The original report was eventually retracted,[27] diminishing considerably the value of our synthetic protein! All was not lost, however, and we were able to report our synthesis as a detailed account of protein synthesis with the KAHA ligation in *Nature Protocols*.[28]

10.3 Nitrophorin 4 (NP4)

With the potential disadvantages of introducing homoserine residues into our synthetic proteins always at the back of our mind, we search for a protein with known structure and known sensitivity

to mutations to examine their effect. We eventually selected nitrophorin 4, an unusual heme-binding lipocalin isolated from the saliva of the blood sucking insect *Rhodnius prolixus*.[29] This protein can be expressed and refolded as its apo protein, to which heme can be added and its NO binding affinity and reduction ability assayed. Several reports noted that point mutations resulted in loss of the NO dismutase activity. We elected to prepare this protein from five segments and chose four different amino acids to mutate to homoserines in the 184 residue sequence: Ala45Hse, Gln75Hse, Thr108Hse, and Asn138Hse (Scheme 13). After a few bumps – including the important lesson that Asp-Pro junctions are not always stable to HPLC conditions[30] – we succeed in preparing the full protein, folding it, and demonstrating that its heme binding and NO dismutase activity were essentially identical to the expressed, wild-type protein.[31] This study gave us the confidence that, for most application, a few homoserine residues were unlikely to affect the biological activity of most proteins – although we are sure there will be some exceptions.



Scheme 13. Synthesis of the 20 kDa protein nitrophorin 4 by KAHA ligation. Despite mutating four residues to homoserines, the synthetic protein was properly folded and shown to bind heme and exhibit NO dismutase activity.

11. ITbM and Cloned Lab

The synthesis of nitrophorin 4 was done by a special group of postdocs, who were the pioneers in establishing our “satellite” lab in Nagoya, Japan. In mid-2012, Prof. Kenichro Itami approached me about joining an application for a World Premiere Institute that combined synthetic chemistry and plant and animal biology. When it was funded later that year and incorporated as the “Institute of Transformative bio-Molecules”, we were suddenly tasked with setting up and staffing an entirely new lab dedicated to peptide and protein synthesis – all within a relatively short time. The first and most stressful step was identifying and hiring a “team leader” for the Nagoya lab, and we were fortunate to recruit Dr. Shunsuke Oishi for this important role.

As any peptide chemist knows, we are often at the mercy of our peptide synthesizers and HPLCs. We had our own growing pains establishing our facilities in Zürich, and we were reluctant to repeat this experience. Instead, we established a “clone lab” with the same instruments, software, and workflows in both Zürich and Nagoya. In addition to constant videoconferences, student

exchanges, and an annual retreat of all group members in the Swiss mountains, the two sites are well connected and brings an unique aspect to our already exceptionally international research group. Under Dr. Oishi's direction, we established the "Peptide and Protein Center" at ITbM, which now produces and hundreds of peptides a year for our biological collaborators and work on cutting edge chemical biology and imaging at ITbM.

12. State of the Ligation

Twelve years after the first successful KAHA ligation, we have finally almost reached our initial goal: a robust, easy to execute, Fmoc-compatible, ligation of unprotected peptide segments for protein synthesis. Although the key linkers for peptide synthesis are not yet commercially available, they are easily prepared and provide direct, traceless access to either unprotected or masked peptide α-ketoacids. The key hydroxylamine building block, (*S*)-5-oxaproline, can be prepared on hundred-gram scale by a straightforward route. So long as one is willing to accept the introduction of a homoserine residue at the ligation site, the current implementation of KAHA ligation offers simple ligation conditions, straightforward masking of either the C- or N-terminus, acidic conditions that excel at solubilizing even difficult peptides, and compatibility with many posttranslational modification. Furthermore, in unpublished results we have established that KAHA ligation is compatible with NCL, making possible one-pot ligations of multiple segments. In addition to the protein target discussed here, many others – both published and unpublished – have been completed using KAHA ligaiton.

Chemical protein synthesis – whether with NCL, KAHA, or other methods – is reaching its intrinsic limits. It's becoming increasingly easier to introduce and mask the key functional groups for both NCL and KAHA ligation – the challenge and throughput limiting step is the time it takes to prepare and purify long peptide segments. Continued advances in the synthesis of modified peptide segment, particularly on larger scale, are therefore essential for the future of the field. As ligating larger molecules becomes difficult and neither KAHA or NCL is well-suited to ligations of folded protein domains, new ligations reactions or enzyme-mediate approaches will be needed to push the field beyond the current achievements of approximately 300 residue proteins. Much of our efforts in reaction development are focused on this challenge[32] – with the hope that the next decade will see yet more advanced in the construction of synthetic proteins.

13. Author Information

Corresponding Author

To whom correspondence should be addressed. **bode@org.chem.ethz.ch**

Biographical Information

Jeffrey W. Bode is Professor of Synthetic Organic Chemistry at ETH Zürich. In addition, he serves as an Executive Editor for the *Encyclopedia of Reagents for Organic Synthesis*, co-Editor in Chief of *Helvetica Chimica Acta*, and a Principal Investigator at the *Institute of Transformative bio-Molecules (ITbM) at* Nagoya University in Japan. His research group focuses on the development of new reactions and chemical biology.

Funding Sources

This work was supported by the Swiss National Science Foundation (169451).

Notes

J.W.B. is named as an inventor on a patent covering the use of the KAHA ligation for peptide and protein synthesis has been granted to the University of California at Santa Barbara.

Acknowledgements

J.W.B thanks the talented and determine group of graduate students and postdocs who have developed the KAHA ligation from a novely amide-bond forming reaction into a powerful platform for chemical protein synthesis. Notable contributions from Ryan Fox, Lei Ju, Alex Lippert, Takeo Fukuzumi, Irene Medina, Melissa Flores, Tetsuo Narumi, Jian Wu, Javier Ruiz, Ayo Ogunkoya, Thomas Wucherpfenning, Felix Limberger, Thibault Harmand, Claudia Murar, Frederic Thuaud, Ivano Pusterla, Chunmao He, Sameer Kulkarni, Sunshuke Oishi, Yoko Oishi, Simon Baldauf, Gabor Boross, Tz-Li Chen, André Zwicky, Esmail Taghizadeh and many others are greatly acknowledged. Special thanks goes to Vijaya Pattabirman for long term support and development of the KAHA ligation and careful proofreading of this manuscript.

15. References

1 Bode, J. W.; Fox, R. M.; Baucom, K. D. Decarboxylative Condensations of α-Ketoacids and N-Alkylhydroxylamines: A New Amide Ligation Reaction. *Angew. Chem. Int. Ed.* **2006**, *45*, 1248–1252.

2 Ju, L.; Lippert, A. L.; Bode, J. W. Stereoretentive Synthesis for Chemoselective Amide-Forming Ligation of C-Terminal Peptide α-Ketoacids. *J. Am. Chem. Soc.* **2008,** *130*, 4253–4255.

3 Ju, L.; Bode, J. W. A General Strategy for the Preparation of C-Terminal Peptide α-Ketoacids by Solid Phase Peptide Synthesis. *Organic and Biomolecular Chem.* **2009**, *7*, 2259–2264.

4 Medina, S. I.; Wu, J.; Bode, J. W. Nitrone Protecting Groups for Enantiopure *N*-hydroxyamino acids: Synthesis of N-terminal Peptide Hydroxylamines for Chemoselective Ligations. *Org. Biomol. Chem.* **2010**, *8*, 3405–3417.

5 Pusterla, I; Bode, J. W. The Mechanism of the α-Ketoacid–Hydroxylamine (KAHA) Amide Forming Ligation. *Angew. Chem. Int. Ed.* **2012**, *51*, 513–516.

6 Knowles, D. A.; Mathews, C. H.; Tomkinson, N. C. O. Oxidation of Primary Amines to Ketones. *Synlett* **2008**, *18*, 2769–2772.

7 Carrillo, N.; Davalos, E. A.; Russak, J. A.; Bode, J. W. Iterative, Aqueous Synthesis of β3-Oligopeptides without Coupling Reagents. *J. Am. Chem. Soc.* **2006**, *128*, 1452–1453.

8 Vasella, A.; Voeffray, R. Asymmetric synthesis of a new proline analogue. *J. Chem. Soc., Chem. Commun.* **1981**, 97–98.

9 Pearce, M. J.; Mintseris, J.; Ferreyra, J.; Gygi, S. P.; K.; Darwin, K. H. Ubiquitin-Like Protein Involved in the Proteasome Pathway of *Mycobacterium tuberculosis. Science* **2008**, *322*, 1104–1107.

10 Pattabiraman, V. R.; Ogunkoya, A. O.; Bode, J. W. Chemical Protein Synthesis by Chemoselective α-Ketoacid–Hydroxylamine (KAHA) Ligations with 5-Oxaproline. *Angew. Chem. Int. Ed.* **2012**, *51*, 5114–5118.

11 Wucherpfennig, T. G.; Rohrbacher, F.; Pattabiraman, V. R; Bode, J. W. Formation and Rearrangement of Homoserine Depsipeptides and Depsiproteins in the KAHA Ligation with 5-Oxaproline. *Angew. Chem. Int. Ed.* **2014**, *53*, 12244–12247.

12 Sohma, Y.; Sasaki, M.; Hayashi, Y.; Kimura, T.; Kiso, Y. Novel and Efficient Synthesis of Difficult Sequence-Containing Peptides through O–N Intramolecular Acyl Migration Reaction of O-Acyl Isopeptides. *Chem. Commun.* **2004**, *1*, 124–125.

13 Rohrbacher, F.; Baldauf, S.; Wucherpfennig, T. G.; Bode, J. W. Product Selectivity in KAHA Ligations: Ester vs. Amide Formation with Cyclic Hydroxylamines. *Synlett.* **2017**, *28*, doi: 10.1055/s-0036-1588480

14 Dixon, W. J.; Hibbert, F. Formation of a Cyclic Tetrahedral Intermediate by the Addition of Water to 2-Methyl-4H- 3,1-Benzoxazine Followed by Ring Opening to 2-Aminobenzyl Acetate and 2-Acetylaminobenzyl Alcohol; pH-Dependence of Rate of Reaction and Product Ratio. *Can. J. Chem.* **1999**, *77* (5–6), 1035–1041.

15 Wucherpfennig, T. G.; Pattabiraman, V. R; Limberg, F.; Ruiz-Rodriguez, J.; Bode, J. W. Traceless Preparation of *C*-terminal □-Ketoacids for Chemical Protein Synthesis with the KAHA Ligation: Synthesis of SUMO2/3. *Angew. Chem. Int. Ed.* **2014**, *53*, 12448–12452.

16 Thuaud, F.; Rohrbacher, F.; Zwicky, A.; Bode, J. W.; Incorporation of Acid-Labile Masking Groups for the Traceless Synthesis of C-Terminal Peptide α-Ketoacids. *Org. Lett.* **2016**, *18*, 3670–3673.

17 Rohrbacher, F.; Deniau, G.; Luther, A.; Bode, J. W. Spontaneous Head-to-Tail Cyclization of Unprotected Linear Peptides with the KAHA Ligation. *Chem. Sci*. **2015**, *6*, 4889–4896.

18 Ogunkoya, A.; Pattabiramin, V.; Bode, J. W. Sequential α-Ketoacid-Hydroxylamine (KAHA) Ligations: Synthesis of C-Terminal Variants of the Modifier Protein UFM1. *Angew. Chem. Int. Ed.* **2012**, *51*, 9693–9697.

19 Thuaud, F.; Rohrbacher, F.; Zwicky, A.; Bode, J. W. Photoprotected Peptide α-Ketoacids and Hydroxylamines for Iterative and One-Pot KAHA Ligations: Synthesis of NEDD8. *Helv. Chim. Acta.* **2016**, *99*, 868–894.

20 Cobos, E. S.; Filimonov, V. V.; Gálvez, A.; Maqueda, M.; Valdívia, E.; Martínez, J. C.; Mateo, P. L. AS-48: A Circular Protein with an Extremely Stable Globular Structure. *FEBS Lett.* **2001**, *505* (3), 379–382.

21 Hemu, X.; Qiu, Y.; Nguyen, G. K. T.; Tam, J. P. Total Synthesis of Circular Bacteriocins by Butelase 1. *J. Am. Chem. Soc.* **2016**, *138*, 6968–6971.

22 Rohrbacher, F. Zwicky, A.; Bode, J. W. Chemical Synthesis of a Homoserine-Mutant of the Antibacterial, Head-to-Tail Cyclized Protein AS-48 by α-Ketoacid–Hydroxylamine (KAHA) Ligation, *Chem. Sci.* **2017**, *8*, 4051–4055.

23 Magers, D. H.; Davis, S. R. Ring strain in the oxazetidines. *J. Mol. Struc. Theochem.* **1999**, *487*, 205–210.

24 Pusterla, I.; Bode, J. W. Synthesis of an Oxazetidine Amino Acid for Rapid, Serine-Forming Protein Ligations, *Nature Chemistry* **2015**, *7*, 668–672.

25 Murar, C. E.; Thuaud, F.; Bode, J. W. KAHA Ligations that form Aspartyl Aldehyde Residues as Synthetic Handles for Protein Modification and Purification. *J. Am. Chem. Soc.* **2014,** *136*, 18140–18144.

26 Yi, P.; Park, J.S.; Melton, D.A. Betatrophin: a Hormone that Controls Pancreatic Beta Cell Proliferation. *Cell* **2013**, *153*, 747–758.

27 Yi, P.; Park, J.S.; Melton, D.A. Retraction Notice to: Betatrophin: A Hormone that Controls Pancreatic b Cell Proliferation. http://dx.doi.org/10.1016/j.cell.2016.12.017

28 Harmand, T. J.; Murar, C. E.; Bode, J. W.; Protein Chemical Synthesis by α-Ketoacid-Hydroxylamine Ligation. *Nature Protocols* **2016**, *11*, 1130–1147.

29 Walker, F. A. Nitric Oxide Interaction with Insect Nitrophorins and Thoughts on the Electron Configuration of the $\{FeNO\}_6$ Complex. *J. Inorg. Biochem.* **2005**, *99*, 216–236.

30 Marcus, F. Preferential Cleavage at Aspartyl-prolyl Peptide Bonds in Dilute Acid. *Int. J. Pept. Protein Res.* **1985**, *25*, 542–546.

31 He, C., Kulkarni, S., Thuaud, F, Bode, J. W. Chemical Synthesis of the 20 kDa Heme Protein Nitrophorin 4 by α-Ketoacid-Hydroxylamine (KAHA) Ligation. *Angew. Chem. Int. Ed.* **2015**, *54,* 12996–13001.

32 Saito, F.; Noda, H.; Bode, J.W. Evaluation and second-order rate constants of chemoselective ligation reactions for stoichiometric conjugations of peptide substrates in water. *ACS Chemical Biology* **2015**, *10,* 1026–1033.