

N-Terminal Proline Editing for the Synthesis of Peptides with
Mercaptoproline and Selenoproline:
Mechanistic Insights Lead to Greater Efficiency in Proline Native
Chemical Ligation

Brice A. Ludwig, Christina R. Forbes, and Neal J. Zondlo*

Department of Chemistry and Biochemistry

University of Delaware

Newark, DE 19716

United States

* To whom correspondence should be addressed. email: zondlo@udel.edu, phone: +1-302-831-0197.

Abstract

Native chemical ligation (NCL) at proline has been limited by cost and synthetic access. In addition, prior examples of NCL using mercaptoproline have exhibited stalling of the reaction after thioester exchange, due to inefficient $S \rightarrow N$ acyl transfer. Herein, we develop methods, using inexpensive Boc-4*R*-hydroxyproline, for the solid-phase synthesis of peptides containing N-terminal 4*R*-mercaptoproline and 4*R*-selenoproline. The synthesis proceeds via proline editing on the N-terminus of fully synthesized peptides on the solid phase, converting an N-terminal Boc-4*R*-hydroxyproline to the 4*S*-bromoproline, followed by S_N2 reaction with potassium thioacetate or selenobenzoic acid. After cleavage from the resin and deprotection, peptides with functionalized N-terminal proline amino acids were obtained. NCL reactions with mercaptoproline proceeded slowly under standard NCL conditions, with the S -acyl transthioesterification intermediate observed as a major species. Computational investigations indicated that the bicyclic intermediates and transition states for $S \rightarrow N$ acyl transfer are sufficiently low in energy (10-15 kcal mol⁻¹ above starting material) that ring strain cannot explain slow $S \rightarrow N$ acyl transfer. Instead, the bicyclic zwitterionic tetrahedral intermediate has a low barrier for reversion to the S -acyl intermediate, causing reversion to the thioester (reverse reaction) to occur preferentially over elimination to generate the amide (forward reaction). We hypothesized that a buffer capable of general acid and/or general base catalysis could promote $S \rightarrow N$ acyl transfer, and thus achieve greater efficiency in proline NCL. In the presence of 2 M imidazole at pH 6.8, NCL with mercaptoproline proceeded efficiently to generate the peptide with a native amide bond. NCL with selenoproline also proceeded efficiently to generate the desired products when a thiophenol thioester was employed as a ligation partner. After desulfurization or deselenization, the products obtained were identical to those synthesized directly, confirming that the solid-phase proline editing reactions proceeded stereospecifically and without epimerization.

Introduction

Native chemical ligation (NCL) and its variant expressed protein ligation (EPL) allow the synthesis of large, complex proteins with non-native and non-encodable functionalities that are incorporated via synthetically prepared segments.¹⁻⁷ NCL was originally described, and is still most widely utilized,⁸ via peptides and proteins with an N-terminal cysteine, which forms an amide bond upon reaction with peptides and proteins with C-terminal thioesters, via thioester exchange (tranthioesterification) followed by *S*→*N* acyl transfer (Figure 1). The scope of NCL with other N-terminal residues has been substantially expanded via the incorporation of thiols and selenols at other native amino acids (as amino acid surrogates), followed by desulfurization or deselenization.⁹⁻¹⁶

Proline is a critical residue in turn and loop regions of proteins, as well as within intrinsically disordered proteins (IDPs) including proline-rich domains of proteins.¹⁷⁻²¹ Proline thus represents an attractive amino acid for NCL, as many of the ligation sites will be solvent-exposed even after protein folding. NCL using 4*R*-mercaptoproline and 4*R*-selenoproline (Figure 1b) as a proline surrogate has been described in work by Danishefsky, Otaka, Payne, and others.²²⁻²⁵ Alternatively, a C-terminal proline thioester could be applied for NCL at junctions with proline. However, proline C-terminal thioesters typically exhibit low reactivity due to the thioester being the electron acceptor of an *n*→ π^* interaction with the pre-prolyl carbonyl, which slows the rate of thioester exchange.^{25,26} This limitation can be overcome using the proline surrogate 4*S*-mercaptoproline at the C-terminal thioester, in which the 4*S*-thiolate internally activates the proline carbonyl thioester to generate a bicyclic intermediate that accelerates thioester exchange with the coupling thiol partner.^{25,27-29} However, the cost of 4*S*-mercaptoproline (\$400/g as Fmoc-4*S*-mcp(Trt)-OH, \$245/mmol) has, to date, limited its applications.

For NCL with an N-terminal mercaptoproline, the 4*R* stereochemistry at proline is required for proline NCL to proceed, due to a lack of functional group proximity that prevents *S*→*N* acyl transfer with 4*S*-mercaptoproline.^{22,23,25} 4*R*-Mercaptoproline has been applied in NCL reactions to synthesize cyclic peptides; a polyglycosylated 88-amino acid fragment of erythropoietin; and *O*-GlcNAcylated heat shock proteins, in work that demonstrated that *O*-GlcNAcylation inhibits amyloidosis of A β and α -synuclein.³⁰⁻³² Mercaptoproline stereoisomers have also been incorporated in conformationally constrained cyclic peptides that function as protein inhibitors or as α -helical capping motifs, as well as in other applications in medicinal chemistry, polymers, and materials science.³³⁻⁴²

However, two key factors have limited the application of NCL at proline. First, while Fmoc-4*R*-mercaptoproline (ProSH, Mcp) is commercially available with trityl protection on the thiol, the cost of this amino acid is high (\$400/g, \$245/mmol), limiting its use, especially in exploratory applications to identify appropriate junctions for NCL in chemical protein synthesis. Alternatively, selenoproline has potentially significant advantages in NCL, including the possibility of deselenization in the presence of other sulfur-containing amino acids and its application in syntheses involving multiple NCL reactions. The combination of 4*R*-selenoproline with selenoesters allows exquisite specificity and increases in rates of NCL reactions, via diselenide-selenoester ligation.²⁵ However, there is no commercially available source of a 4*R*-selenoproline. The syntheses of suitably chalcogen-protected Boc-4*R*-mercaptoproline or Boc-4*R*-selenoproline (Figure 2) require 7 steps in 25% overall yield; or 4 or 5 steps, in 56% overall yield (as the diprolyl diselenide, which reduces the yield in subsequent amide coupling reactions) or in 26% overall yield (as the PMB-protected selenoether), respectively, each from inexpensive Boc-4*R*-hydroxyproline (Boc-Hyp-OH) or 4*R*-hydroxyproline (Hyp-OH).^{23-25,35,43,44} The requirement

of solution-phase synthesis of the amino acid, the length of the synthetic sequences, and the number of purification steps have resulted in these amino acids having been only rarely applied in published examples of NCL.

Applications of NCL at mercaptoproline have been further hindered by a problem that limits reaction scope: inefficient $S \rightarrow N$ acyl transfer. In Danishefsky's initial report, the authors described the presence of significant amounts of the S -acyl intermediate after transthioesterification.²² This thioester intermediate, because it lacks the key new amide bond, is subject to reversion to the mercaptoproline starting material via reaction with thiols in solution. NCL using mercaptoproline was also limited to reaction at thioesters with less sterically congested amino acids: Otaka reported slow NCL reactions even at Ala thioesters, and Danishefsky indicated very low yields (< 15%) at Val thioesters, due to the accumulation of the S -acyl intermediates (the species immediately prior to $S \rightarrow N$ acyl transfer).^{22,23} Given that typical NCL reactions are conducted with an excess concentration of thiol additives, in order to maintain the reactive thiol in the reduced state and to promote transthioesterification, the slow conversion of the thioester into the amide bond ($S \rightarrow N$ acyl transfer) might substantially limit the use of mercaptoproline in NCL. Herein, we seek to examine methods to increase the efficiency of $S \rightarrow N$ acyl transfer at mercaptoproline based on mechanistic insights.

We also seek to explore strategies for the more practical synthesis of peptides with N-terminal mercaptoproline and N-terminal selenoproline. We previously developed a method, termed proline editing, for the practical synthesis of peptides with chemically diverse 4-substituted prolines.⁴⁵⁻⁴⁷ Proline editing, as originally described, proceeds via the incorporation of inexpensive Fmoc-4*R*-hydroxyproline (Fmoc-Hyp-OH) into internal positions in peptides on solid phase, *in situ* protection of the hydroxyl group (as an orthogonally removable trityl or silyl

ether, Alloc carbonate, or nitrobenzoate ester) after amide coupling, and completion of the synthesis of the peptide. After synthesis of the full peptide, the Hyp hydroxyl is selectively deprotected, while the rest of the peptide on solid phase remains protected. The hydroxyl can then be converted via stereospecific reactions on the solid phase into a variety of 4*R*- or 4*S*-substituted prolines. After synthesis to generate the modified proline, standard peptide cleavage and deprotection yield the product peptide, with no solution-phase purification and only a single HPLC purification, as is required for all peptides prepared by solid-phase peptide synthesis. This practical method was applied to convert a single protected peptide on solid phase into 122 peptides with different combinations of substitution and stereochemistry on the proline.⁴⁷

In our work on proline editing, we synthesized peptides that had proline 4*R*-substitution with sulfur and selenium.⁴⁷ These reactions proceeded via a Mitsunobu reaction on the 4*R*-hydroxyproline with nitrobenzoic acid to generate the 4*S*-nitrobenzoate, deprotection of the nitrobenzoate with NaN₃ to generate the 4*S*-hydroxyproline, and then a second Mitsunobu reaction using a sulfur or selenium nucleophile, which regenerates the 4*R*-stereochemistry. This sequence on the solid phase proceeds efficiently because no purification steps are involved, allowing reactions to be conducted with excess reagents and removal of reagents by filtration at the end of each step. However, we also demonstrated the solid-phase Appel reaction on these peptides, generating the 4*S*-bromoproline, which is positioned for direct modification with inversion to the 4*R* derivative via an S_N2 reaction, avoiding the nitrobenzoate deprotection step.⁴⁷ A strategy for the synthesis of peptides with N-terminal 4*R*-mercaptoproline or 4*R*-selenoproline will be examined using this approach.

Results and Discussion

Synthesis of peptides with N-terminal 4R-mercaptoproline and 4R-selenoproline. We envisioned that peptides with an N-terminal mercaptoproline or N-terminal selenoproline could be synthesized via solid-phase modification of an N-terminal 4R-hydroxyproline (Figure 3). To test this possibility, the standard NCL model peptide XRAFS-NH₂ was synthesized by solid-phase peptide synthesis. Here, inexpensive Boc-Hyp-OH (\$0.60/g, \$0.14/mmol) was used as the N-terminal amino acid (X) in Fmoc solid-phase peptide synthesis, since the Boc group would be removed during standard TFA cleavage of peptides from resin and deprotection, generating an N-terminal aminothiols (free N-terminus and side-chain thiol) appropriate for NCL. The 4R-hydroxyproline was converted with inversion of stereochemistry to the 4S-bromoproline via a solid-phase Appel reaction (Figure 3b).⁴⁷⁻⁵³ This product was then converted to the protected 4R-mercaptoproline via S_N2 reaction with potassium thioacetate (KSAc) in the presence of 18-crown-6. Reaction with KSAc proceeded more efficiently than using sodium hydrosulfide in the presence of 15-crown-5, which generated significant quantities of the dehydropyran elimination side product (Figure S10). Thioester deprotection was conducted on the solid phase, in order to prevent proline N-acylation via S→N acyl transfer from the thioacetate upon peptide cleavage/deprotection. Finally, standard TFA cleavage/deprotection yielded the peptide with an N-terminal 4R-mercaptoproline.

The peptide with an N-terminal 4R-selenoproline was synthesized similarly (Figure 3cd). The 4S-bromoproline on solid phase was subjected to S_N2 reaction with selenobenzoic acid.^{24,54} Hydrolysis of the selenoester was conducted with LiOH in the presence of benzyl mercaptan (BnSH), which generated, after standard TFA cleavage and deprotection, the peptide with proline 4R-substituted with selenium, as the mixed selenosulfide. The peptide synthesis

proceeded with high overall conversion to generate this peptide, which was isolated in higher yield as the selenosulfide than as a diselenide.

Initial attempts at NCL reactions with 4R-mercaptoproline. The peptide with an N-terminal 4R-mercaptoproline was subjected to native chemical ligation conditions with the model peptides Ac-LYR₄AA-SBn and Ac-LYR₄AA-SPh (Figure 4 and Supporting Information). As had been observed previously in NCL reactions with mercaptoproline,²²⁻²⁴ the reactions proceeded sluggishly, generating the transthioesterification intermediate in significant quantities, with little or no product observed that had undergone *S*→*N* acyl transfer to generate the stable amide bond. Importantly, this intermediate thioester is subject to rapid thioester exchange in the presence of thiols in solution. The thioester (isomeric and identical in mass to the desired product with an amide at proline) can be isolated and identified as containing a thioester, and not a new amide bond, via its reaction with external thiols (e.g. added benzylmercaptan or thiophenol), which regenerates the two component starting material peptide fragments.

Computational analysis of plausible mechanisms of NCL with 4R-mercaptoproline. We hypothesized that the slow *S*→*N* acyl transfer could potentially be due to strain in the bicyclic zwitterionic tetrahedral intermediate that results from intramolecular nucleophilic attack of the proline nitrogen on the thioester. To assess this possibility, we conducted DFT calculations on these structures (Figure 5). Inconsistent with this hypothesis, we found that the bicyclic zwitterionic tetrahedral intermediate was (depending on computational methods used) only 4-11 kcal mol⁻¹ higher in energy than the *S*-acyl starting material, indicating that the ring strain in the bicyclic tetrahedral intermediate was relatively modest. In order to better understand the mechanism of *S*→*N* acyl transfer, we therefore conducted calculations on the various intermediates in the full pathway from *S*-acyl proline to the product with an amide bond, with the

aim of devising a strategy for overcoming the stalling at the *S*-acyl intermediate observed experimentally (both herein and previously). We also calculated transition state structures and energies for key steps in the pathway.

These calculations (Figure 5) confirmed that the overall energetics of *S*→*N* acyl transfer (downhill by 16–17 kcal mol⁻¹) were highly favorable. In addition, the inherent activation barriers for these reactions (7–17 kcal mol⁻¹ above the starting material, which all correspond to first-order reaction half-lives of under a second at room temperature) suggest that the *S*-acyl intermediate should readily react to generate the more stable amide product. Why is this result not observed experimentally?

The initially generated zwitterionic tetrahedral intermediate could take three distinct pathways for elimination of the thiolate to generate an amide bond (Figure 5): via an anionic tetrahedral intermediate (anionic pathway), directly via the zwitterionic tetrahedral intermediate (zwitterionic pathway), or via a neutral tetrahedral intermediate (neutral pathway). Each pathway was considered individually.

In the anionic pathway, deprotonation of the proline nitrogen of the zwitterionic tetrahedral intermediate would generate an anionic tetrahedral intermediate (anionic on O, neutral on N) that upon elimination would yield a neutral amide. The anionic tetrahedral intermediate would be modestly uphill in energy from the zwitterionic tetrahedral intermediate, requiring a proton transfer to generate a higher energy intermediate (based on typical ammonium pK_a values). We were unable to identify an energy-minimum structure for this putative anionic tetrahedral intermediate: in all attempts, the molecule spontaneously eliminated the thiolate to generate the amide bond, with no discernable activation barrier. Thus, we conclude that if

deprotonation of the proline nitrogen occurs from the zwitterion, the amide product will be obtained via a highly favorable elimination reaction.

The zwitterionic tetrahedral intermediate can eliminate directly to generate an amide that is protonated on the nitrogen. While that species, if generated, would be expected to be rapidly deprotonated, the initial cationic elimination product was found to be uphill energetically from the zwitterionic tetrahedral intermediate. In addition, the transition state for this reaction was calculated to be significantly higher in energy than that for reversion of the zwitterionic tetrahedral intermediate back to the *S*-acyl starting material.

Finally, the neutral tetrahedral intermediate (neutral on both O and N) was considered. This intermediate would result from proton transfers on the zwitterion to the O and from the N. While the neutral tetrahedral intermediate is more stable than the zwitterionic tetrahedral intermediate, it is less stable than the cationic tetrahedral intermediate at pH 7, based on the expected pK_a values of the alcohol and the ammonium. The cationic tetrahedral intermediate is also modestly more stable than the *S*-acyl starting material. However, elimination from the cationic tetrahedral intermediate is energetically implausible, as it would result in an amide that is cationic on both the nitrogen and the oxygen, which would be extremely unfavorable. In contrast, elimination from the neutral tetrahedral intermediate would generate an amide that is only protonated (cationic) on the carbonyl oxygen, which would convert to the final neutral amide rapidly via proton transfer. This elimination reaction is modestly uphill (+2 kcal mol⁻¹ from the neutral tetrahedral intermediate, +7 kcal mol⁻¹ from the cationic tetrahedral intermediate) prior to the proton transfer to generate the neutral amide.

Each of the three plausible elimination pathways had a calculated activation barrier that was 15–16 kcal mol⁻¹ greater than the energy of the *S*-acyl starting material, which indicates that

these reactions could occur rapidly at room temperature. These computational results thus suggest that the *S*-acyl intermediate *inherently* should not significantly accumulate, in contrast to experimental observations. So why, experimentally, was more *S*-acyl intermediate observed than the lower-energy amide (*N*-acyl) final product?

The *S*→*N* acyl transfer is thermodynamically favorable, and can proceed with individual barriers that should allow rapid reaction at room temperature. However, the calculations suggested that the inherent kinetic problem in this reaction is the instability of the zwitterionic tetrahedral intermediate: the reaction generating the zwitterion exhibits a very late transition state, with a very small barrier of 1–3 kcal mol⁻¹ to revert from the zwitterion back to the *S*-acyl starting material, with the exact barrier dependent on the computational methods employed. This barrier is similar to or smaller than barriers for proton transfer, and is significantly less than the barrier for elimination via any of the three plausible pathways. In addition, the reversion reaction is unimolecular, while any proton transfer is bimolecular. Thus, the calculations suggest that the key kinetic problem of *S*→*N* acyl transfer is the instability of the zwitterionic tetrahedral intermediate, and its strong kinetic propensity to revert back to the *S*-acyl starting material due to the very small barriers for that reversion reaction. This propensity for reversion would be expected to be even greater with a more sterically demanding amino acid (e.g. Val) at the thioester, which would be expected to have a higher energy of the zwitterionic tetrahedral intermediate, and thus (by the Hammond postulate) have an even later transition state (i.e. an even smaller barrier to reversion). These predictions are consistent with prior experimental data indicating very low yields of mercaptoproline NCL at Val thioesters, and slow reactions even at Ala thioesters.²²⁻²⁴

Based on these computational insights, we hypothesized that $S \rightarrow N$ acyl transfer might be accelerated via general base and/or general acid catalysis. In a plausible general acid catalysis mechanism (e.g. with imidazoleH⁺), the nucleophilic attack of the proline nitrogen would occur concomitantly with proton transfer to the S -acyl carbonyl, and therefore generate the cationic tetrahedral intermediate directly, bypassing the metastable zwitterionic tetrahedral intermediate. In a plausible general base catalysis mechanism (e.g. with imidazole), the nucleophilic proline nitrogen could be deprotonated simultaneously with nucleophilic attack on the carbonyl, thus generating the anionic tetrahedral intermediate that spontaneously eliminates. Alternatively, general base catalysis could reduce the barrier for elimination from either the neutral tetrahedral intermediate or the zwitterionic tetrahedral intermediate, via oxygen or nitrogen deprotonation, respectively, that is simultaneous with elimination, generating the neutral amide directly in a highly thermodynamically favorable downhill reaction, and avoiding higher energy cationic amide intermediates. In addition, mechanisms that combine general acid and general base catalysis could be envisioned.

Finally, a general acid or general base catalyst could function solely to accelerate proton transfer(s) from the zwitterionic tetrahedral intermediate, as was proposed by Barnett and Jencks in the analogous reaction of $S \rightarrow N$ acyl transfer of S -acetylmercaptoethylamine.^{6,55} In that reaction, the zwitterionic tetrahedral intermediate was metastable and prone to reversion to the neutral (S -acetyl and neutral amine, AcS-CH₂-CH₂-NH₂) starting material. Proton transfers from the zwitterionic tetrahedral intermediate to the neutral tetrahedral intermediate were proposed to be the rate-limiting step, with rate acceleration promoted by general acid catalysis (or, for HCO₃⁻, a proposed dual general acid-general base catalysis that results in both proton transfers occurring in a single step). Notably, Seitz and coworkers recently developed auxiliaries for NCL

reactions in which a pyridyl base on the auxiliary accelerated $S \rightarrow N$ acyl transfer, consistent with proton transfer from the zwitterionic tetrahedral intermediate also being the rate-determining step in amide bond formation in that reaction.⁵⁶

Optimized NCL reactions with 4R-mercaptoproline. Therefore, we examined the possibility that general acid and/or general base catalysis could promote $S \rightarrow N$ acyl transfer, by re-examining the NCL reaction as a function of buffer identity and buffer concentration. We examined 0.2 M and/or 2 M concentrations of NaOAc/AcOH, Na₂HPO₄/NaH₂PO₄, imidazole/imidazoleH⁺, and NaCO₃H/Na₂CO₃ buffers (Figures S63–S79). The most efficient reaction conditions were in the presence of 2 M imidazole/imidazoleH⁺ at pH 6.8 (i.e. near the p*K*_a of imidazolium) (Figure 6a-c), which resulted in the generation of the amide NCL product with high conversion within a reasonable timeframe, with little accumulation of the *S*-acyl intermediate. The lowest conversion to amide product was observed with the acetate buffer, while the bicarbonate buffer exhibited modest conversion and more significant levels of thioester hydrolysis. The observation of the conversion to amide depending on both the concentration of the buffer and the identity of the buffer is consistent with a role for general acid and/or general base catalysis in achieving high conversion to the amide product.

While imidazole at high concentrations (≥ 2.5 M) can accelerate cysteine NCL via the formation of an *N*-acyl imidazole intermediate that is more reactive to thiolates than a thioester,⁵⁷⁻⁵⁹ this mechanism would be expected to increase the rate of formation of the transthioesterification intermediate, but have no impact on the rate of $S \rightarrow N$ acyl transfer. Indeed, imidazole could potentially directly react with the transthioesterification intermediate (reverse reaction), leading to an acyl imidazole intermediate that would not be able to undergo

intramolecular *S*→*N* acyl transfer, resulting in slower *S*→*N* acyl transfer, which is inconsistent with the experimental results.

Using this optimized approach, we also observed efficient NCL at Val thioesters (Figure 6d-g and Figure S49). At 4 h, both *S*-acyl (thioester) and *N*-acyl (amide) products were present, along with unreacted starting material. At 24 h, high conversion was observed from the peptide containing mercaptoproline into the *N*-acyl amide product, with little observed *S*-acyl thioester intermediate. These results are in contrast to the very low yields of amide product observed by Danishefsky using standard NCL conditions.²⁴ NCL reactions with 2 M imidazole were also compatible with 6 M guanidine•HCl, which is commonly used for NCL reactions in order to solubilize protein fragments and assure accessibility of their thioester and aminothiol functional groups (Figures S77–S78). Combined with the computational results, these data suggest that the propensity of the zwitterionic tetrahedral intermediate to revert to the *S*-acyl starting material is the key limiting step in *S*→*N* acyl transfer for proline NCL. These computational insights and experimental results should be transferable to other applications of NCL at proline, broadening its use in chemical protein synthesis and protein semi-synthesis.

We sought to assess the generality of this approach for the synthesis of longer peptides with 4*R*-mercaptoproline and their application to NCL. Hyperphosphorylation in the tau proline-rich domain is associated with structural changes in peptides and proteins and with increased tau aggregation and neurodegeneration in Alzheimer's disease and in animal models thereof.⁶⁰⁻⁶⁵ In contrast, O-GlcNAcylation of tau is associated with protection against aggregation.⁶⁶⁻⁶⁸ Thus, the tau proline-rich domain makes it an attractive target for NCL reactions to understand the code of effects of tau post-translational modifications on structure and function.⁶⁹⁻⁷¹ The proline rich-domain of tau lacks a cysteine residue, and has only one alanine residue, near its C-terminus, and

thus lacks ideal sites for NCL with Cys. Proline residues represent the most attractive sites for NCL reactions here. Toward the goal of understanding the PTM code in the tau proline-rich domain, we synthesized two fragments of the tau proline-rich domain, tau₁₉₅₋₂₁₂-SBn and tau₂₁₃₋₂₃₈. The N-terminal residue Pro₂₁₃ residue of tau₂₁₃₋₂₃₈ was incorporated as Boc-Hyp. This peptide was subjected to solid-phase Appel reaction, and that product was subjected to solid-phase S_N2 reaction with KSAc in the presence of 18-crown-6 (Figures S85-S87). HPLC analysis indicated that both of these reactions proceeded with high conversion to product. The product peptide with an N-terminal 4*R*-mercaptoproline was then subjected to the optimized NCL reaction conditions with tau₁₉₅₋₂₁₂-SBn, generating the product in good conversion in 3 hours (Figure 7).^{65,72-77} These results confirm that the solid-phase proline editing protocol is compatible with larger peptides, and that the optimized mercaptoproline NCL reaction conditions with 2 M imidazole are applicable to the synthesis of larger peptides.

*Computational and experimental analysis of NCL reactions with 4*R*-selenoproline* We also investigated the pathways for NCL of selenoproline, using similar computational approaches (Figure 8). Here, we found that the zwitterionic tetrahedral intermediate was substantially more stable than that of mercaptoproline, presumably due to the longer C–Se bonds (compared to C–S bonds) reducing ring strain in the bicyclic tetrahedral intermediate. In this case, the barrier for formation of the zwitterionic tetrahedral intermediate was found to be quite modest, suggesting the *Se*→*N* acyl transfer should occur more readily than that for *S*→*N* acyl transfer. While the barrier for reversion was also small for the selenopeptide zwitterionic tetrahedral intermediate, the overall energetics via the reaction intermediates were also substantially more favorable, consistent with Payne's success in selenoproline NCL in divergent applications, including the synthesis of Pro-Pro junctions using selenoester-diselenide ligation.²⁵

The primary challenges of selenoproline NCL using standard conditions with thioesters or added thiols are (1) the greater stability of thioesters compared to selenoesters, which reduces the rate of formation of and the stability of selenoesters prepared via thioesters; and (2) the significantly greater sensitivity of selenols to oxidation compared to thiols. Thus, standard approaches of NCL using excess thiol additives^{78,79}, in conditions where reduction of diselenide is particularly important, present a potentially significant kinetic liability for selenoproline NCL, via the promotion of more stable thioesters over less stable selenoesters. Indeed, we observed low rates of NCL product formation in reactions using high concentrations of MPAA or other thiols that are commonly added in order to keep the selenol in the reduced state (Figure S20).^{13,78} We identified optimized conditions (Figure 9) for selenoproline NCL (1) using TCEP in the presence of ascorbate in order to reduce diselenides to selenols (with ascorbate present to prevent deselenization)⁸⁰; (2) avoiding added thiols; and (3) using thiophenol-derived thioesters, in place of typical benzyl thioesters, which are more stable than the thiophenolates and thus result in less efficient conversion of thioesters into selenoesters.^{25,78} The reactions were also conducted at pH 6.5, in order to slow hydrolysis of the phenyl thioester while maintaining the selenoate in an anionic form (to promote selenoester formation) and maintaining a sufficient fraction of the proline amine in the neutral form (to promote *Se*→*N* acyl transfer).^{81,82} Using these optimized conditions, selenoproline NCL proceeded efficiently and in high yield, generating a mixture of selenol and diselenide products.

Mercaptoproline and selenoproline within peptides are directly useful in the synthesis of peptides with constrained cyclic structures, via amide cyclization, via disulfides or diselenides, or via alkylation of these products, for either modification or cyclization.^{33,34} Because of the structural utility of these amino acids, we used DFT calculations to examine their conformational

preferences, in both the neutral (thiol or selenol) and anionic (thiolate or selenoate) ionization states (Figures 10 and 11). Previous work has demonstrated that thiols and selenols exhibit only a modest stereoelectronic effect in 4-substituted prolines, with the steric effect of the larger chalcogens counterbalancing their weaker stereoelectronic preference.^{47,83} Using models of Ac-4*R*-mercaptoproline-NHMe and Ac-4*R*-selenoproline-NHMe, we observed that these amino acids exhibit only a slight preference for the proline *exo* ring pucker, in contrast to the strong *exo* ring pucker preference of 4*R*-hydroxyproline or 4*R*-fluoroproline,⁸⁴ consistent with prior experimental and computational results indicating minimal *exo/endo* preference for 4*R*-mercaptoproline.⁸³ The structures with sulfur or selenium and an *exo* ring pucker also exhibited more extended values of ϕ and weaker intercarbonyl $n \rightarrow \pi^*$ interactions than those in hydroxyprolines. In the anionic ionization states, the steric preferences for an *endo* ring pucker modestly predominated for both mercaptoproline and selenoproline, although in both ionization states there was relatively little *exo/endo* preference. The selenoproline also exhibited an only modest *trans/cis* amide preference in these calculations. These results suggest that mercaptoproline and selenoproline should both be applicable in diverse structural contexts in peptides and proteins due to their modest conformational preferences.

Determination that synthesis and NCL of peptides with 4R-mercaptoproline and 4R-selenoproline proceeded stereospecifically. Finally, in order to confirm that the syntheses of these amino acids on solid phase proceeded stereospecifically, and without epimerization at proline H α , the products of NCL reactions with model peptides were subjected to desulfurization or deselenization using standard conditions, in order to generate peptides with native proline residues at the amide bond junction. The desulfurization and deselenization products were then compared, by HPLC co-injection (Figure 12) and by comparison of the NMR spectra (Figure 13;

superposition of TOCSY spectra: Figures S35-S38, S45-S48), with the equivalent peptide synthesized directly by solid-phase peptide synthesis. These data indicate that the peptide products obtained via NCL plus desulfurization/deselenization to generate proline or via direct synthesis with proline were identical. These results suggest that the N-terminal proline editing approach is a practical, alternative approach to the synthesis of peptides with N-terminal 4*R*-mercaptoproline or 4*R*-selenoproline for application in proline NCL.

Conclusion

The results herein substantially expand the scope and potential applications of proline NCL. First, we developed practical methods for the solid-phase synthesis of peptides with an N-terminal 4*R*-mercaptoproline or 4*R*-selenoproline. These peptides were synthesized using inexpensive Boc-4*R*-hydroxyproline-OH as a starting material, with all reactions conducted via solid-phase modification chemistry (N-terminal proline editing), eliminating the need for purification and isolation of the intermediates in the synthesis of these amino acids. The only purification step required in this method is standard purification of the final peptide. Similar methods should be applicable to synthesize peptides with C-terminal 4*S*-mercaptoproline, which greatly accelerates NCL at proline thioesters,^{27,28} as well as peptides with internal 4*R*- and 4*S*-mercaptoproline and selenoproline. These approaches should broaden the scope of applications of these amino acids, and in particular should encourage exploratory applications of proline NCL.

In addition, we conducted detailed computational analysis of plausible mechanistic pathways in NCL with 4*R*-mercaptoproline, in order to determine the basis for slow *S*→*N* acyl transfer observed previously and herein. We identified that the primary limitation of the *S*→*N*

acyl transfer reaction is not the inherent barriers for the reaction, nor the energies of the intermediates, but is instead the kinetic instability of the initially generated zwitterionic tetrahedral intermediate, which is highly prone to reversion to the *S*-acyl intermediate that is experimentally observed as a major product of the NCL reaction. Based on the computational analysis of mechanistic pathways, we identified that a high concentration of a buffer (imidazoleH⁺/imidazole) capable of functioning as a general acid and/or a general base catalyst greatly accelerates *S*→*N* acyl transfer, generating the final NCL product with an amide bond in high chemical yield, with no *S*-acyl intermediate observed at the end of the reaction. These results, by identifying the key limiting step in achieving high yield, greatly expand the scope of NCL using 4*R*-mercaptoproline, including reactions at more sterically hindered thioesters, and should be broadly applicable in improving yields in complex applications of NCL and EPL using mercaptoproline.

Experimental

Peptide synthesis. Peptides were synthesized by standard Fmoc solid-phase peptide synthesis. All peptides were synthesized on Rink amide resin, except for peptides containing a C-terminal hydrazide. These peptides, precursors to peptides with C-terminal thioesters, were synthesized on 2-chlorotrityl chloride resin that was treated with 5% hydrazine in DMF for 45 minutes twice prior to peptide synthesis. Peptides containing C-terminal hydrazides were acetylated at the N-terminus. All peptides were subjected to cleavage from the resin, deprotection, and purification by standard methods.

The purified peptides containing C-terminal hydrazides were subjected to oxidation of the hydrazide in solution phase using 0.45 mM NaNO₂ in 100 mM phosphate buffer pH 3 at 0 °C

for 15 minutes, after which excess benzyl mercaptan or thiophenol was added (depending on the identity of the desired thioester).⁸⁵ The pH of the solution was then increased to 6.5–7.0, and the peptides were allowed to react for 3 hours before purification.

Modification of N-terminal proline residues in peptides was conducted on solid phase, with Boc-4*R*-hydroxyproline (Hyp) as the N-terminal residue. For synthesis of both 4*R*-mercaptoproline and 4*R*-selenoproline, the N-terminal Boc-Hyp (10 mg of resin, 27 μ mol of peptide on resin) was first converted to the 4*S*-bromoproline via an Appel reaction on solid phase, conducted as a reaction with 1 mmol carbon tetrabromide (CBr₄), 1 mmol triphenylphosphine (Ph₃P), and (optionally) 1 mmol DIAD in 2 mL THF for 24 hours, with the reaction allowed to proceed twice, replenishing reagents between reactions. CBr₄ and Ph₃P were dried by vacuum for at least 12 hours immediately prior to the reactions. In model peptides, the Appel reaction proceeded in similar yields with or without DIAD (Figure S7–S8). In larger peptides, higher yields were observed conducting the reaction without DIAD (Figure S78, S85).

Within peptides, N-terminal 4*R*-mercaptoproline was synthesized from 4*S*-bromoproline on solid phase by allowing the peptide on resin (10 mg of resin, 27 μ mol) to react with potassium thioacetate (0.1 mmol) and 18-crown-6 (0.1 mmol) in anhydrous DMF. The peptide was then subjected to thiolysis on solid phase using 2-aminoethanethiol (0.3 mmol), thiophenol (0.3 mmol), and sodium borohydride (0.3 mmol) in 3 mL MeOH for 90 minutes, with the thiolysis reaction conducted twice, replenishing reagents between reactions.

Within peptides, N-terminal 4*R*-selenoproline was synthesized from 4*S*-bromoproline on solid phase by first synthesizing selenobenzoic acid *in situ*, by allowing Woollins' reagent (0.82 mmol) to react with benzoic acid (2.45 mmol) in 1 mL toluene at reflux for 2 hours under N₂ gas.²⁴ The solution was then allowed to cool to 50 °C, the resin containing the peptide with N-

terminal Boc-4*S*-bromoproline (100 mg of resin, 270 μ mol), DIPEA (500 μ L, 2.85 mmol), and 1 mL anhydrous DMF were added, and the reaction was allowed to proceed for 6 hours. The N-terminal 4*R*-selenoproline was then protected as the benzyl selenosulfide by allowing the peptide on resin to react with benzyl mercaptan (170 μ mol) and lithium hydroxide (5.85 mmol) in 3 mL anhydrous THF for 3 hours.

Details of synthetic procedures, purification, and characterization for all peptides are in the Supporting Information.

Computational chemistry. Calculations were conducted with Gaussian 09.⁸⁶ Geometry optimization was conducted using initial structures developed previously⁸⁷ that were modified *in silico*. Geometry optimization was conducted iteratively, with final geometry optimization conducted using the M06-2X DFT functional and the Def2TZVP basis set in implicit water (IEFPCM method).⁸⁸⁻⁹⁰ For 4*R*-mercaptoproline (Ac-ProSH-NHMe) and 4*R*-selenoproline (Ac-ProSeH-NHMe), geometry optimization was conducted as a function of proline amide rotamer (*trans* versus *cis*), proline ring pucker (*exo* versus *endo*), region of the Ramachandran plot (PPII versus α R/ δ), and side chain ionization state (neutral thiol/selenol versus thiolate/selenoate). The energies of each combination of conformation and ionization state were determined using the MP2 method and the Def2TZVP basis set in implicit water (IEFPCM method).⁹¹

In order to understand the mechanisms of *S* $\rightarrow$ *N* acyl transfer and *Se* $\rightarrow$ *N* acyl transfer, a series of structures derived from H-ProSAc-NHMe and H-ProSeAc-NHMe was developed via iterative geometry optimization calculations, with final geometry optimization via the M06-2X method and the Def2TZVP basis set in implicit water (IEFPCM). For mercaptoproline, structures were developed both with (*R*) stereochemistry at the carbonyl-derived carbon of the tetrahedral intermediate (which results in a *trans* amide bond after elimination), and with (*S*)

stereochemistry at the carbonyl-derived carbon of the tetrahedral intermediate (which results in a *cis* amide bond after elimination), in order to examine the pathways from stereoisomeric intermediates resulting from attack of the proline nitrogen lone pair on both the *re* and *si* faces of the thioacetate carbonyl. Transition states were identified using initial scans of C–N and/or C–S bond length in addition to the carbonyl or elimination from the tetrahedral intermediates, in order to locate structures near the transition state (structures with energy local maxima and minimal errors). These initial geometries were then subjected to transition state optimization within Gaussian09.

All final local energy minimum structures generated using the M06-2X method and the Def2TZVP basis set were analyzed by frequency calculations, using the same combination of method, basis set, and solvent used in the original geometry optimization calculations. All local energy minima exhibited no negative (imaginary) frequencies unless otherwise indicated. In all cases, the analysis of the one imaginary frequency in transition states was consistent with these structures being transition states of the relevant reaction.

In order to more thoroughly investigate all intermediates and all key transition states in the pathways, the mercaptoproline structures were then selectively subjected to further geometry optimization calculations using different combinations of DFT functionals (M06-2X, M11, and ω B97-xD), basis sets (Def2TZVP and 6-311++G(d,p)), and water solvation models (IEFPCM and SMD).⁹²⁻⁹⁴ In addition, geometry optimization calculations were conducted on energy minima with the MP2 method and the Def2TZVP basis set and either the IEFPCM or SMD solvation model. Further details and analysis are in the Supporting Information.

Optimized NCL reactions. Peptides containing N-terminal 4*R*-mercaptoproline (1.5–2.0 mM) and peptides containing a C-terminal thioester (1.5–2.0 mM) were allowed to incubate at

23 °C in 30 μ L of a solution containing 2 M imidazole and 20 mM TCEP at pH 6.8, which had been degassed by the freeze-pump-thaw method immediately prior to use. The headspace of the tube was replaced with N₂ gas, and the tube was sealed with PTFE tape. Aliquots of the solution were removed and monitored via HPLC in order to monitor reaction progress. Details are in the Supporting Information.

Desulfurization and deselenization. Desulfurization was conducted by allowing the purified ligation product peptide containing mercaptoproline to react with sodium borohydride (1.3 M) and nickel(II) chloride hexahydrate (9 mM) in 20 μ L of 100 mM phosphate buffer pH 7.2 for 20 minutes at 0 °C in a microcentrifuge tube. The solution was allowed to vent by piercing a needle through the top of the tube. The reaction was quenched by adding 500 μ L of 100 mM phosphate buffer pH 7.2 containing EDTA (10 mM final concentration) and allowing the resultant solution to incubate for 30 minutes at 37 °C, followed by peptide purification and analysis.

Deselenization was conducted by allowing the purified selenoproline ligation product peptide to dissolve in 100 mM phosphate buffer pH 5 containing 250 mM TCEP, which was degassed by the freeze-pump-thaw method immediately prior to use. The headspace of the microcentrifuge tube containing the reaction was replaced with N₂ gas, the tube was sealed with PTFE tape, and the solution was allowed to incubate at 37 °C for 12 hours, followed by peptide purification and analysis.

Supporting Information Available

Details of peptide synthesis, purification, and characterization; NMR spectra comparing peptides synthesized by NCL to peptides prepared via direct chemical synthesis; additional

details of calculations, including an analysis of energetics of NCL reactions as a function of DFT functional, basis set, and solvation model, as well as coordinates of geometry-optimized structures; and a comparison of NCL reactions as a function of buffer identity and buffer concentration. This material is available free of charge via the journal web site.

Acknowledgements

We thank NIH (GM93225), the DOD CDMRP PRARP program (AZ140115), and NSF (CHE-2004110, CHE-1412978, and BIO-1616490) for funding. Instrumentation support was provided by NIH (GM110758) and NSF (CHE-1229234).

Figures

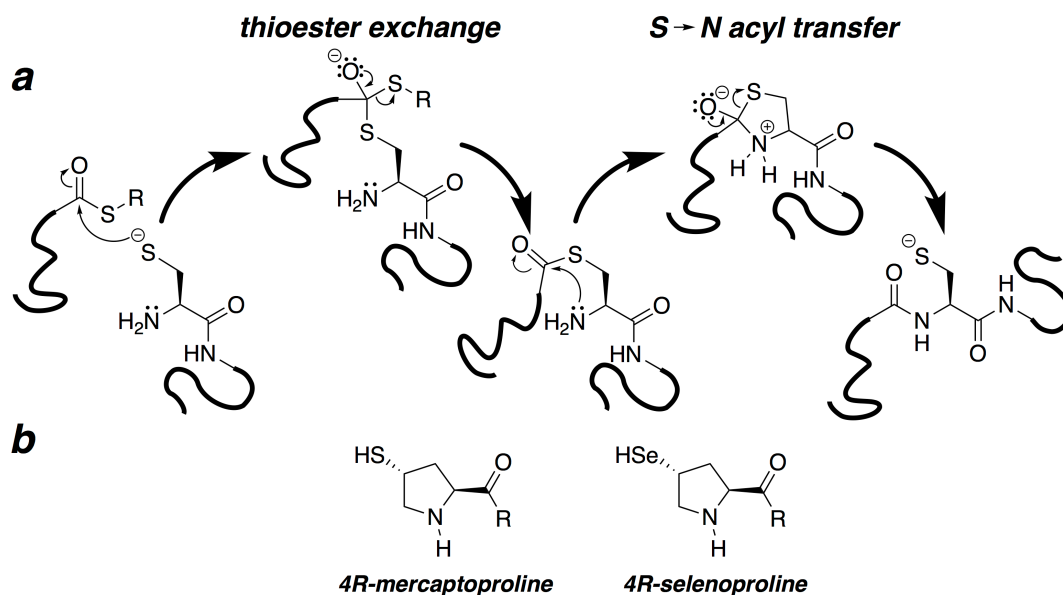


Figure 1. (a) General mechanism of native chemical ligation (NCL) via cysteine. (b) 4*R*-Mercaptoproline and 4*R*-selenoproline can be used in native chemical ligation of peptides. These proline derivatives can be readily desulfurized or deselenized, respectively, to result in a native proline within the product peptide. 4*S*-Substituted proline derivatives do not proceed beyond thioester exchange due to their preference for the *endo* ring pucker, which hinders nucleophilic attack by the amine on the *S*-acyl carbonyl to effect *S* → *N* acyl transfer.

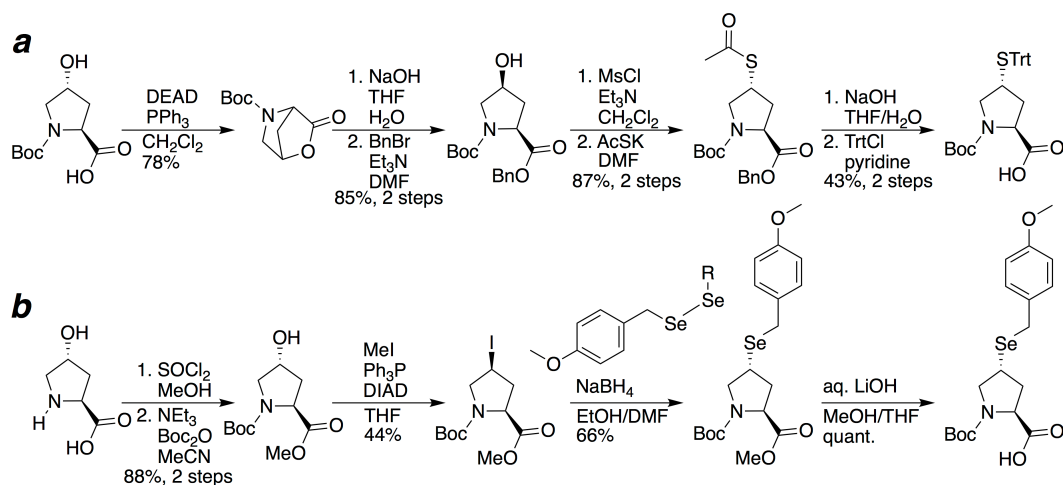


Figure 2. Solution-phase synthesis of 4-substituted proline derivatives previously described for use in NCL. (a) The published solution-phase syntheses of Boc-4*R*-mercaptoproline (25% overall yield)²³ and (b) Boc-4*R*-selenoproline (26% overall yield)²⁵. Both syntheses require several purification steps to prepare the protected amino acid for solid-phase peptide synthesis.

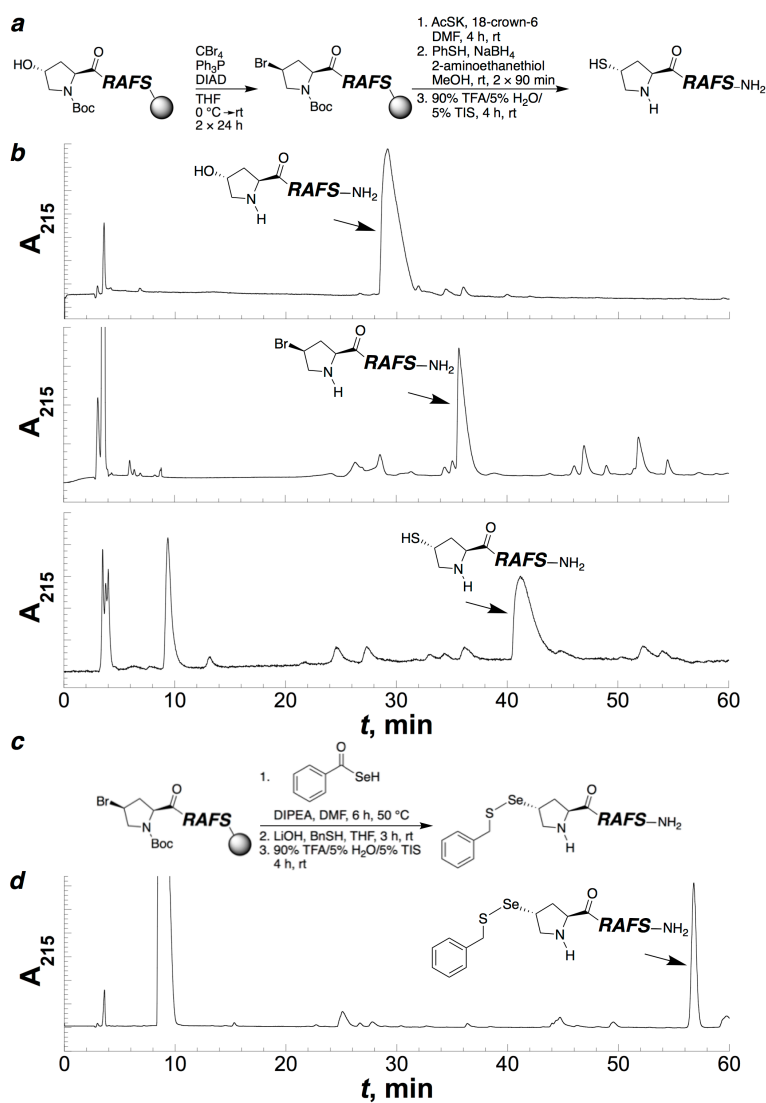


Figure 3. Synthesis of 4*R*-substituted proline derivatives on solid phase starting from Boc-4*R*-hydroxyproline (X) on the N-terminus of the model peptide XRAFS-NH₂. (a) Reaction scheme and (b) crude analytical HPLC chromatograms of the starting material and products of each step of the solid-phase synthesis of 4*R*-mercaptoproline, with each product shown following cleavage from resin and deprotection (90% TFA/5% H₂O/5% TIS, 4 h). The chromatogram for the 4*S*-bromoproline intermediate exhibits additional side products due to elimination or substitution during cleavage from the resin and deprotection, as a result of the relative instability of the alkyl bromide during these reactions.⁴⁷ The solid-phase Appel reaction can also be conducted without DIAD (Figure S8). (c) Reaction scheme and (d) crude analytical HPLC chromatograms of the products of the solid-phase synthesis of 4*R*-selenoproline benzyl selenosulfide following cleavage from resin and deprotection (90% TFA/5% H₂O/5% TIS, 4 h). Chromatography for each peptide was conducted using a linear gradient of 0-30% buffer B (20% H₂O, 80% MeCN, 0.05% TFA) in buffer A (98% H₂O, 2% MeCN, 0.06% TFA) over 60 minutes on an analytical C18 column.

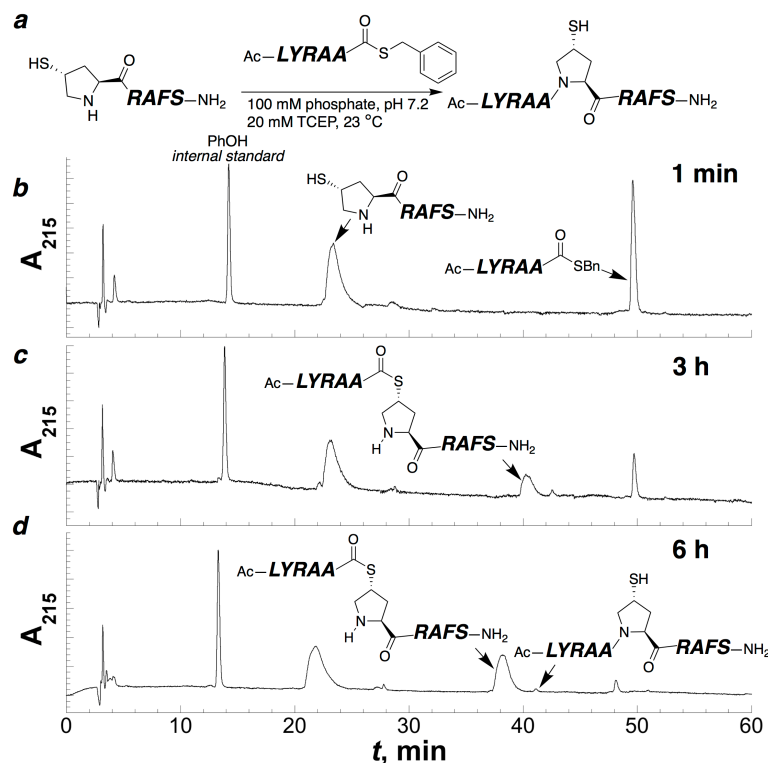


Figure 4. Native chemical ligation (a) scheme and (b-d) crude HPLC chromatograms for the ligation reaction between the peptides (4*R*-mercaptoproline)-RAFS and Ac-LYRAA-SBn, with the reaction products shown after (b) 1 minute, (c) 3 hours, and (d) 6 hours in 100 mM phosphate buffer pH 7.2. Chromatography was conducted using a linear gradient of 0-60% buffer B in buffer A over 60 minutes on an analytical C18 column. After 3 hours, the major ligation product remained stalled as the *S*-acyl intermediate. The identity of the unarranged thioester intermediate was verified by the addition of excess benzyl mercaptan to the isolated peak, which resulted in reversion to the original component peptides. Similar results were obtained with phenyl thioesters, and at different values of pH, with the proline *S*-acyl intermediate predominant over the desired amide product.

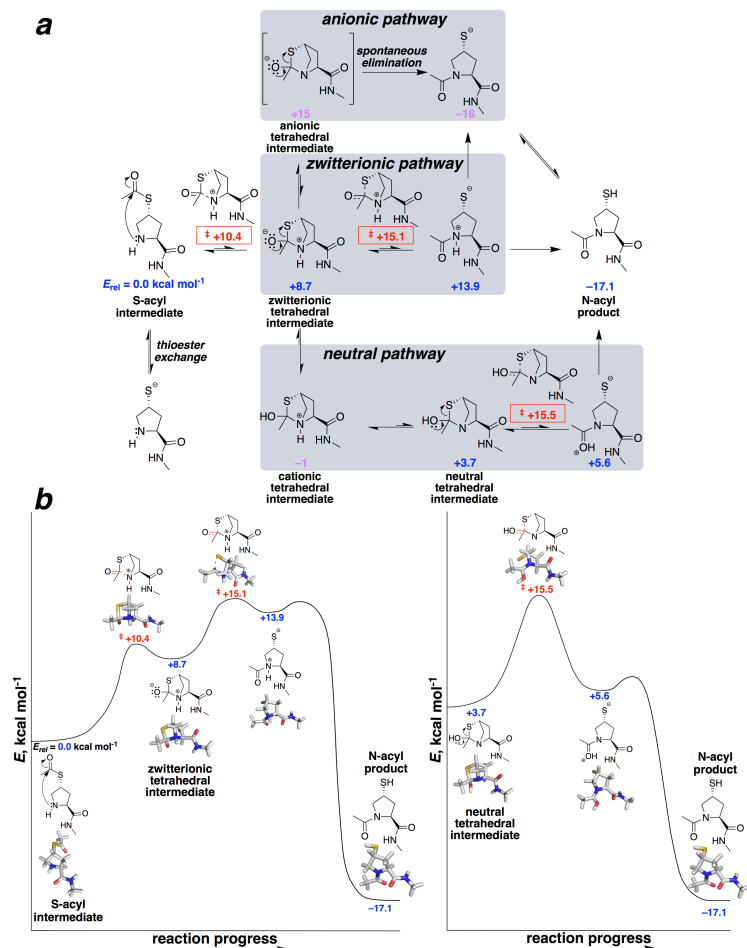


Figure 5. (a) Computational investigation of the mechanism of *S*→*N* acyl transfer for 4*R*-mercaptoproline. (b) Energy diagrams of the key intermediates and transition states in the proposed mechanism. Blue (overall neutral) and magenta (overall -1 or overall +1 charge) energies are colored corresponding to states whose calculated energies are directly comparable. For species that have an overall charge, energies are estimated based on expected populations at pH 7 compared to species that are overall neutral, based on typical pK_a values of the relevant functional groups. Numbers in red indicate transition state energies. The bracketed anionic tetrahedral intermediate structure is not a stable species computationally, and spontaneously generates the *N*-acyl product during geometry optimization, indicating at most a minimal barrier to elimination upon deprotonation of the zwitterion. *S*→*N* acyl transfer can proceed via nucleophilic attack on either the *re* or the *si* face of the carbonyl; thus, either an (*R*) or (*S*) stereoisomer may be present at the carbon derived from the acyl carbonyl. The (*R*) stereoisomers shown in tetrahedral intermediates herein exhibit lower energies and result in a *trans* amide bond upon elimination, while the (*S*) stereoisomers result in a *cis* amide bond. The structures for the (*S*)-acyl-derived stereoisomeric pathway and a summary of all calculated energies are in Figure S55 and Table S2. Geometry optimization was conducted using multiple combinations of DFT functionals, basis sets, and solvation models, in order to obtain the most accurate approximation of energies of species and energy barriers between intermediates (see Table S3 for a summary). Geometry-optimized structures and energies shown in this Figure were obtained using the ω B97-xD functional and the 6-311++G(d,p) basis set with the SMD water solvation model.

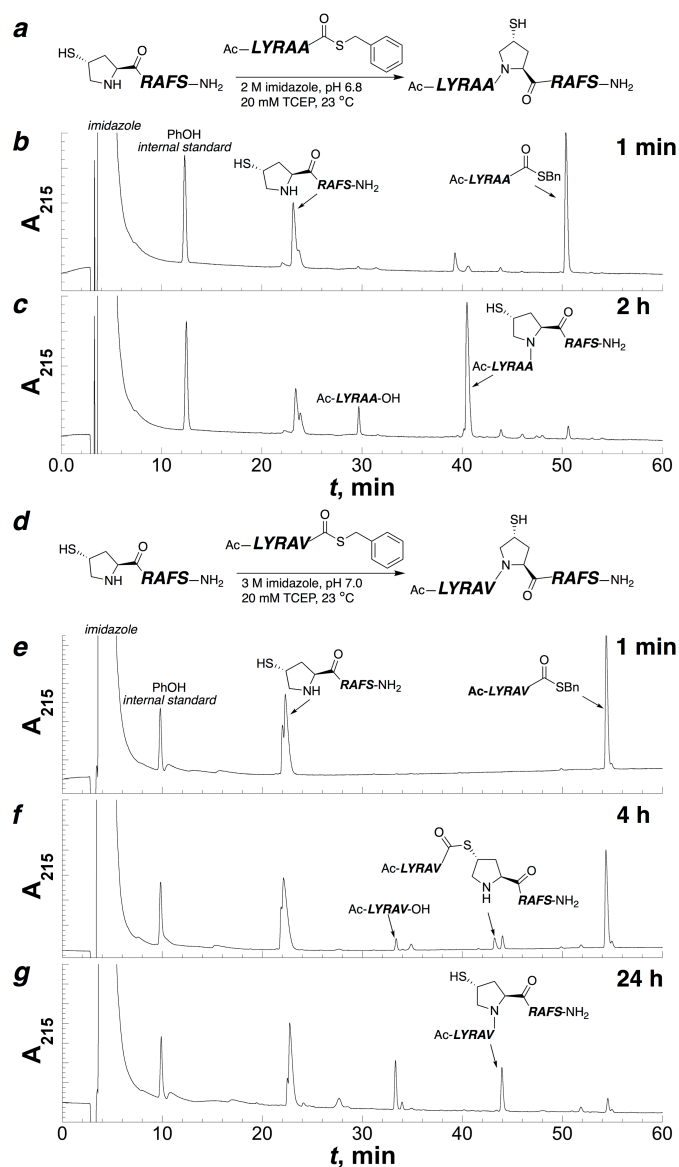


Figure 6. Native chemical ligation (a) scheme and (b, c) crude HPLC chromatograms for the ligation reaction between the peptides (4*R*-mercaptoproline)-RAFS and Ac-LYRAA-SBn, with the reaction analyzed (b) within 1 minute of sample preparation and (c) after 2 hours in the optimized reaction conditions. Native chemical ligation (d) scheme and (e-g) crude HPLC chromatograms for the ligation reaction between the peptides (4*R*-mercaptoproline)-RAFS and Ac-LYRAV-SBn, with the reaction analyzed (e) within 1 minute of sample preparation, (f) after 4 hours, and (g) after 24 hours in the optimized reaction conditions. At 4 hours, similar quantities of *S*-acyl (thioester) intermediate and *N*-acyl (amide) product are observed, while at 24 hours there is minimal *S*-acyl product present. Analysis of additional timepoints between 4 and 24 hours is in the Supporting Information (Figure S49). The longer reaction time for Val thioesters also resulted in greater hydrolysis of the thioester. Chromatography was conducted using a linear gradient of 0-60% buffer B in buffer A over 60 minutes on an analytical C18 column.

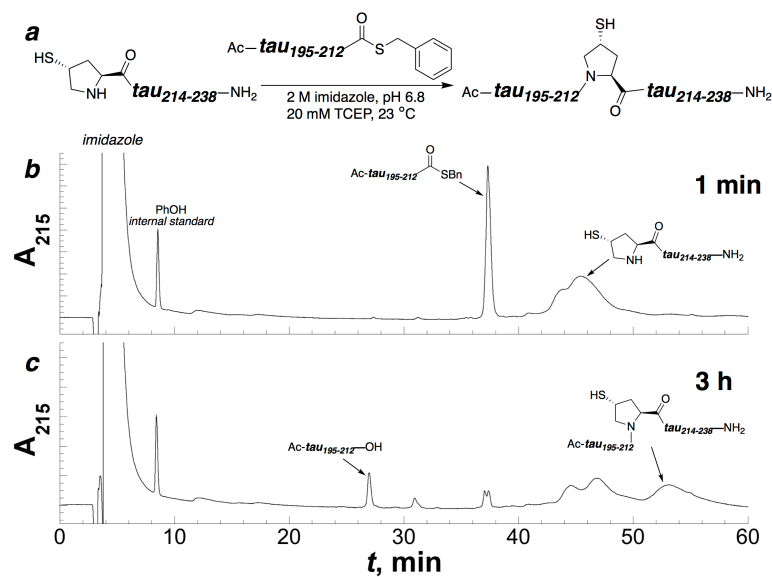


Figure 7. Native chemical ligation (a) scheme and (b, c) crude HPLC chromatograms for the ligation reaction between the peptides $\text{tau}_{213-238}$ (containing N-terminal 4*R*-mercaptoproline) and $\text{tau}_{195-212}$ (T212S), with the reaction analyzed (b) within 1 minute of sample preparation and (c) after 3 hours. Chromatography was conducted using a linear gradient of 0-60% buffer B in buffer A over 60 minutes on an analytical C18 column. Broadening of the peaks corresponding to the $\text{tau}_{214-238}$ and $\text{tau}_{195-238}$ peptides occurs as a result of these sequences containing multiple proline residues, with peak broadening and multiple peaks observed due to heterogeneous proline *cis-trans* isomerism.^{65,72-77}

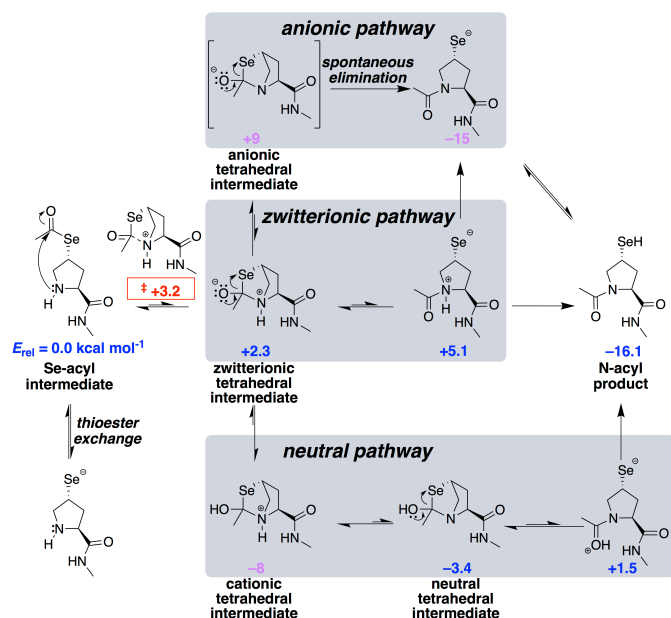


Figure 8. Computational investigation of the possible mechanisms of $Se \rightarrow N$ acyl transfer for 4*R*-selenoproline. Blue (overall neutral) and magenta (overall -1 or overall +1 charge) energies are colored corresponding to states whose calculated energies are directly comparable, with energies of charged species estimated at pH 7 relative to neutral intermediates, using typical pK_a values of the relevant groups. Numbers in red indicate transition state energies. The bracketed anionic tetrahedral intermediate structure is not a stable species computationally, and spontaneously generates the *N*-acyl product during geometry optimization. $Se \rightarrow N$ acyl transfer can proceed via nucleophilic attack on either the *re* or the *si* face of the carbonyl; thus, in the tetrahedral intermediates, either an (*R*) or (*S*) stereoisomer may be present at the carbon derived from the acyl carbonyl. The (*R*) stereoisomers shown exhibit lower energies and result in a *trans* amide bond upon elimination, while the (*S*) stereoisomers generate a *cis* amide bond. The structures for the (*S*)-acyl-derived stereoisomeric pathway and all calculated energies are in Figure S60 and Table S4. Geometry optimization calculations were conducted using the M06-2X DFT functional and Def2TZVP basis set in implicit water (IEFPCM method). Energies of all structures were determined using the MP2 method and the Def2TZVP basis set in implicit water (IEFPCM method).

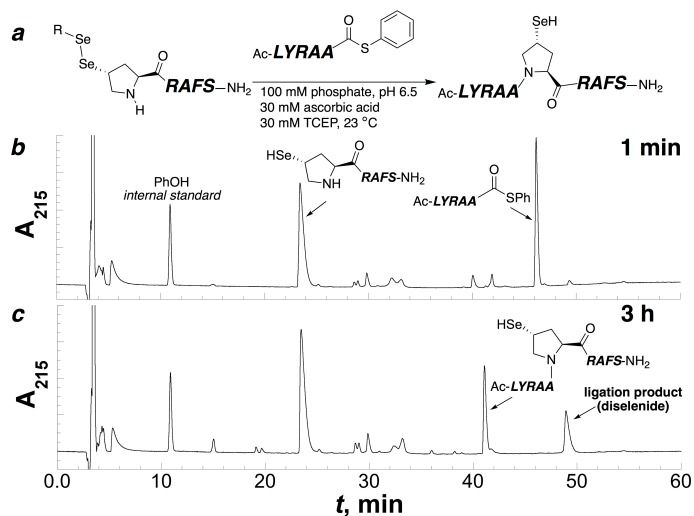


Figure 9. Native chemical ligation (a) scheme and (b, c) crude HPLC chromatograms for the ligation reaction between the peptides (4*R*-selenopropionamide)-RAFS and Ac-LYRAA-SPh, with the reaction analyzed at (b) initial sample preparation (≤ 1 min), and (c) after 3 hours. Chromatography was conducted using a linear gradient of 0-60% buffer B in buffer A over 60 minutes on a C18 analytical column. The peaks at 41 minutes and 49 minutes represent selenol monomer and diselenide dimer of the ligation product, respectively, as confirmed by mass spectrometry and co-injection.

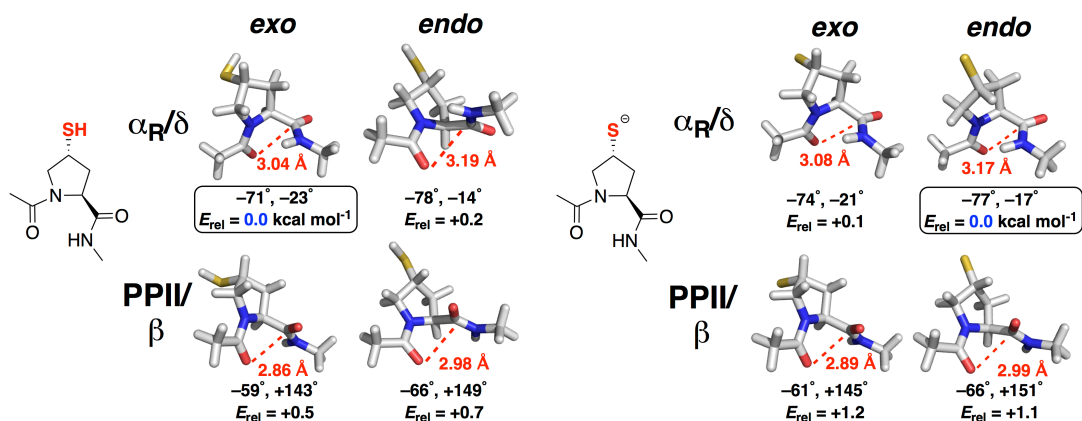


Figure 10. Summary of the conformational preferences of 4*R*-mercaptoproline as a function of thiol ionization state, as determined by computational investigations. Conformations with *cis*-proline were also examined, and are reported in the Supporting Information (Figure S51). The thiol (left) form of 4*R*-mercaptoproline very modestly prefers an *exo* ring pucker, while the thiolate (right) slightly prefers an *endo* ring pucker. Boxes indicate the conformations with the lowest relative energies for each ionization state. Torsion angles indicated are for ϕ and ψ , respectively. Red dashed lines indicate measured intercarbonyl distances (d , $O_i \cdots C_{i+1}$), with distances substantially below the sum of the van der Waals radii of oxygen and carbon (3.22 Å) indicating an $n \rightarrow \pi^*$ interaction. Geometry optimization calculations were conducted using the M06-2X DFT functional and Def2TZVP basis set in implicit water (IEFPCM method). Energies (in kcal mol⁻¹) were determined using the MP2 method and the Def2TZVP basis set in implicit water (IEFPCM method).

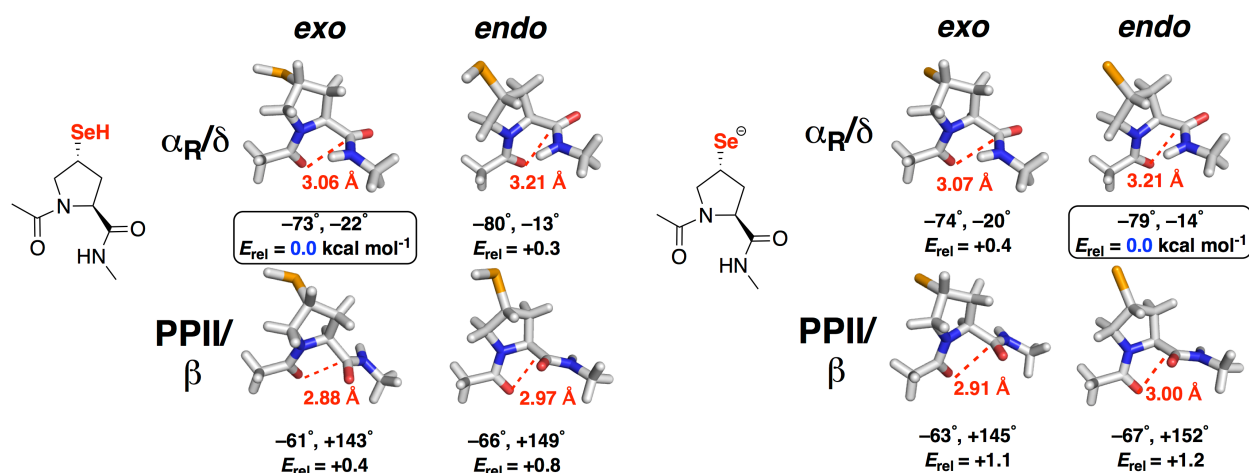


Figure 11. Summary of the conformational preferences of 4*R*-selenoproline as a function of selenol ionization state, as determined by computational investigations. Conformations with *cis*-proline were also examined, and are reported in the Supporting Information (Figure S53). The selenol (left) form of 4*R*-selenoproline very modestly prefers an *exo* ring pucker, while the selenoate (right) slightly prefers an *endo* ring pucker. Boxes indicate the conformations with the lowest relative energies for each ionization state. Torsion angles indicated are for ϕ and ψ , respectively. Red dashed lines indicate measured intercarbonyl distances (d , $O_i \cdots C_{i+1}$), with distances substantially below the sum of the van der Waals radii of oxygen and carbon (3.22 Å) indicating an $n \rightarrow \pi^*$ interaction. Final geometry optimization calculations were conducted using the M06-2X DFT functional and Def2TZVP basis set in implicit water (IEFPCM method). Final energies were determined using the MP2 method and the Def2TZVP basis set in implicit water (IEFPCM method).

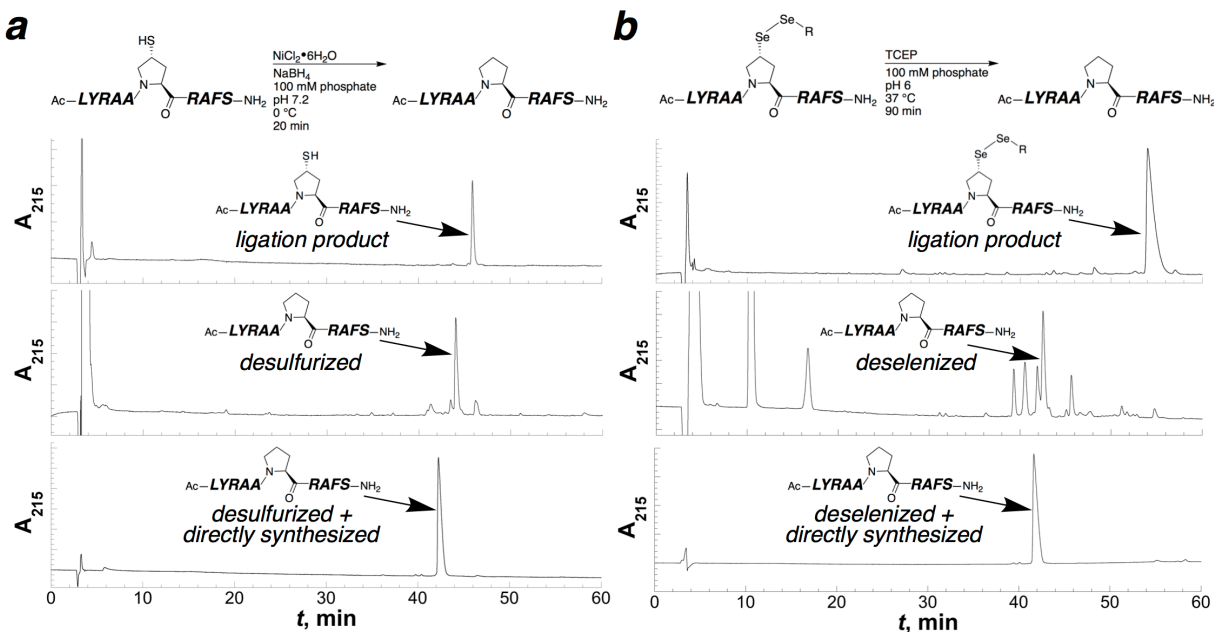


Figure 12. Desulfurization and deselenization of ligation products. HPLC chromatograms of the product of ligation prior to and following (a) desulfurization, as well as prior to and following (b) deselenization. For both reactions, the major peaks at 43 minutes were identified as the dechalcogenated product. A side product of the deselenization reaction was identified as Ac-LYRAA-Hyp-RAFS-NH₂ (Figures S25-S26), which results from residual oxygen within the solution.⁸⁰ A co-injection of the desulfurized or deselenized peptide with the directly synthesized peptide Ac-LYRAAPRAFS-NH₂ (bottom panels) was also conducted to confirm that each product is identical to the native peptide. Chromatography was conducted on an analytical C18 column using a linear gradient of 0-50% buffer B in buffer A over 60 minutes.

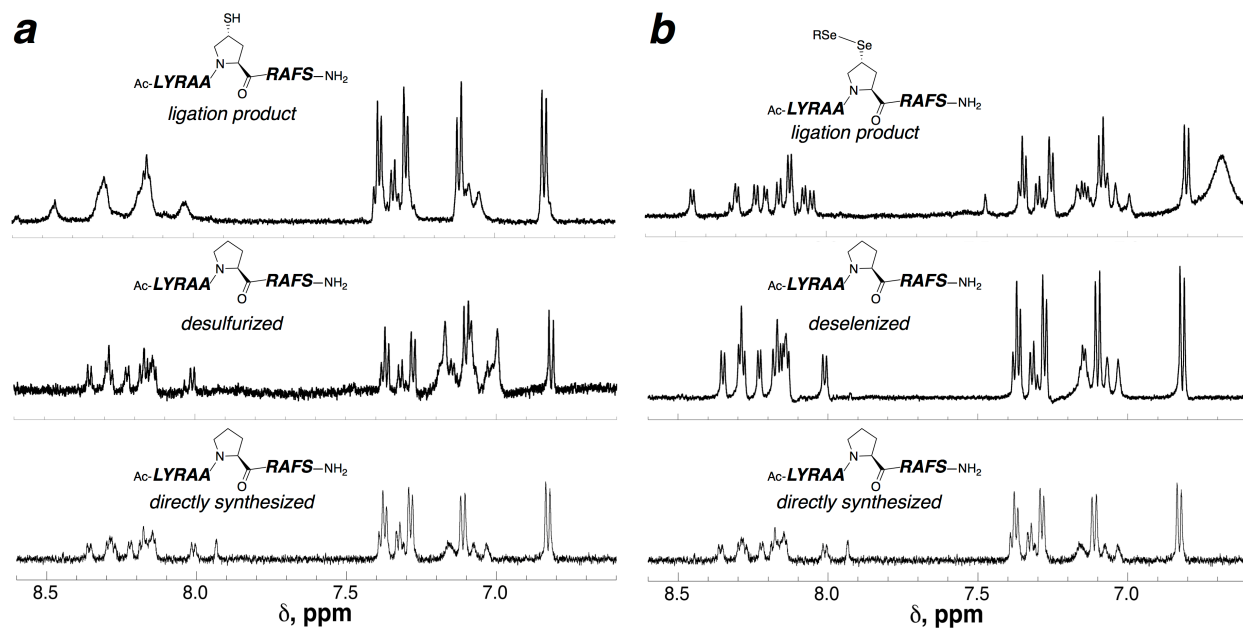


Figure 13. Amide-aromatic regions of the ^1H NMR spectra of the (a) mercaptoproline ligation product and (b) selenoproline ligation product, in comparison with the desulfurized and deselenized products, respectively, and the directly synthesized peptide Ac-LYRAAPRAFS- NH_2 in 90% H_2O /10% D_2O with 5 mM phosphate buffer pH 4, 25 mM NaCl, 0.1 mM TCEP, and 0.1 mM TSP. Additional comparisons, including superpositions of the TOCSY spectra of the desulfurized or deselenized product and the directly synthesized product, are in the Supporting Information (Figures S29–S38 and Figures S41–48).

References

- (1) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266*, 776–779.
- (2) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705-6710.
- (3) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338-351.
- (4) Malins, L. R.; Payne, R. J. Recent extensions to native chemical ligation for the chemical synthesis of peptides and proteins. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70-78.
- (5) Bondalapati, S.; Jbara, M.; Brik, A. Expanding the chemical toolbox for the synthesis of large and uniquely modified proteins. *Nat. Chem.* **2016**, *8*, 407-418.
- (6) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoet, M.; Monbaliu, J. C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **2019**, *119*, 7328-7443.
- (7) Thompson, R. E.; Muir, T. W. Chemoenzymatic Semisynthesis of Proteins. *Chem. Rev.* **2020**, *120*, 3051-3126.
- (8) Agouridas, V.; El Mahdi, O.; Cargoet, M.; Melnyk, O. A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorg. Med. Chem.* **2017**, *25*, 4938-4945.
- (9) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.
- (10) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.
- (11) Rohde, H.; Seitz, O. Ligation-Desulfurization: A Powerful Combination in the Synthesis of Peptides and Glycopeptides. *Biopolymers* **2010**, *94*, 551-559.
- (12) Dawson, P. E. Native Chemical Ligation Combined with Desulfurization and Deselenization: A General Strategy for Chemical Protein Synthesis. *Isr. J. Chem.* **2011**, *51*, 862-867.
- (13) Mitchell, N. J.; Malins, L. R.; Liu, X. Y.; Thompson, R. E.; Chan, B.; Radom, L.; Payne, R. J. Rapid Additive-Free Selenocystine-Selenoester Peptide Ligation. *J. Am. Chem. Soc.* **2015**, *137*, 14011-14014.
- (14) Conibear, A. C.; Watson, E. E.; Payne, R. J.; Becker, C. F. W. Native chemical ligation in protein synthesis and semi-synthesis. *Chem. Soc. Rev.* **2018**, *47*, 9046-9068.
- (15) Kulkarni, S. S.; Sayers, J.; Premjee, B.; Payne, R. J. Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat. Rev. Chem.* **2018**, *2*, article 0122.
- (16) Guan, I. V. Y.; Williams, K.; Liu, J. S. T.; Liu, X. Y. Synthetic Thiol and Selenol Derived Amino Acids for Expanding the Scope of Chemical Protein Synthesis. *Frontiers Chem.* **2022**, *9*, 826764.
- (17) MacArthur, M. W.; Thornton, J. M. Influence of proline residues on protein conformation. *J. Mol. Biol.* **1991**, *218*, 397-412.
- (18) Williamson, M. P. The Structure and Function of Proline-Rich Regions in Proteins. *Biochem. J.* **1994**, *297*, 249-260.

- (19) Reiersen, H.; Rees, A. R. The hunchback and its neighbours: proline as an environmental modulator. *Trends Biochem. Sci.* **2001**, *26*, 679-684.
- (20) Bhattacharyya, R.; Chakrabarti, P. Stereospecific Interactions of Proline Residues in Protein Structures and Complexes. *J. Mol. Biol.* **2003**, *331*, 925-940.
- (21) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* **2005**, *18*, 343-384.
- (22) Shang, S.; Tan, Z.; Dong, S.; Danishefsky, S. J. An Advance in Proline Ligation. *J. Am. Chem. Soc.* **2011**, *133*, 10784-10786.
- (23) Ding, H.; Shigenaga, A.; Sato, K.; Morishita, K.; Otaka, A. Dual Kinetically Controlled Native Chemical Ligation Using a Combination of Sulfanylproline and Sulfanylethylanilide Peptide. *Org. Lett.* **2011**, *13*, 5588-5591.
- (24) Townsend, S. D.; Tan, Z. P.; Dong, S. W.; Shang, S. Y.; Brailsford, J. A.; Danishefsky, S. J. Advances in Proline Ligation. *J. Am. Chem. Soc.* **2012**, *134*, 3912-3916.
- (25) Sayers, J.; Karpati, P. M. T.; Mitchell, N. J.; Goldys, A. M.; Kwong, S. M.; Firth, N.; Chan, B.; Payne, R. J. Construction of Challenging Proline-Proline Junctions via Diselenide-Selenoester Ligation Chemistry. *J. Am. Chem. Soc.* **2018**, *140*, 13327-13334.
- (26) Pollock, S. B.; Kent, S. B. H. An investigation into the origin of the dramatically reduced reactivity of peptide-prolyl-thioesters in native chemical ligation. *Chem. Commun.* **2011**, *47*, 2342-2344.
- (27) Nakamura, T.; Shigenaga, A.; Sato, K.; Tsuda, Y.; Sakamoto, K.; Otaka, A. Examination of native chemical ligation using peptidyl prolyl thioesters. *Chem. Commun.* **2014**, *50*, 58-60.
- (28) Gui, Y.; Qiu, L. Q.; Li, Y. H.; Li, H. X.; Dong, S. W. Internal Activation of Peptidyl Prolyl Thioesters in Native Chemical Ligation. *J. Am. Chem. Soc.* **2016**, *138*, 4890-4899.
- (29) Shah, M. I. A.; Xu, Z. Y.; Liu, L.; Jiang, Y. Y.; Shi, J. Mechanism for the enhanced reactivity of 4-mercaptoprolyl thioesters in native chemical ligation. *RSC Adv.* **2016**, *6*, 68312-68321.
- (30) Dong, S.; Shang, S.; Tan, Z.; Danishefsky, S. J. Toward Homogeneous Erythropoietin: Application of Metal-Free Dethylation in the Chemical Synthesis of the Ala79-Arg166 Glycopeptide Domain. *Isr. J. Chem.* **2011**, *51*, 968-976.
- (31) Hosseini, A. S.; Wang, W. J.; Haeffner, F.; Gao, J. M. Metal-Assisted Folding of Prolinomycin Allows Facile Design of Functional Peptides. *ChemBioChem* **2017**, *18*, 479-482.
- (32) Balana, A. T.; Levine, P. M.; Craven, T. W.; Mukherjee, S.; Pedowitz, N. J.; Moon, S. P.; Takahashi, T. T.; Becker, C. F. W.; Baker, D.; Pratt, M. R. O-GlcNAc modification of small heat shock proteins enhances their anti-amyloid chaperone activity. *Nature Chem.* **2021**, *13*, 441-450.
- (33) Pace, J. R.; Lampkin, B. J.; Abakah, C.; Moyer, A.; Miao, J. Y.; Deprey, K.; Cerulli, R. A.; Lin, Y. S.; Baleja, J. D.; Baker, D.; Kritzer, J. A. Stapled beta-Hairpins Featuring 4-Mercaptoproline. *J. Am. Chem. Soc.* **2021**, *143*, 15039-15044.
- (34) Brown, H.; Chung, M.; Ueffing, A.; Batistatou, N.; Tsang, T.; Doskocil, S.; Mao, W. Q.; Willbold, D.; Jr, R. C. B.; Lu, Z.; Weiergraeber, O. H.; Kritzer, J. A. Structure-Based Design of Stapled Peptides That Bind GABARAP and Inhibit Autophagy. *J. Am. Chem. Soc.* **2022**, *144*, 14687-14697.
- (35) Matinkhoo, K.; Wong, A.; Hambira, C. M.; Kato, B.; Wei, C.; Muller, C.; Hechler, T.; Braun, A.; Gallo, F.; Pahl, A.; Perrin, D. M. Design, Synthesis, and Biochemical

Evaluation of Alpha-Amanitin Derivatives Containing Analogs of the trans-Hydroxyproline Residue for Potential Use in Antibody-Drug Conjugates. *Chem. Eur. J.* **2021**, *27*, 10282-10292.

(36) Kemp, D. S.; Curran, T. P.; Davis, W. M.; Boyd, J. G.; Muendel, C. Studies of N-Terminal Templates for Alpha-Helix Formation - Synthesis and Conformational-Analysis of (2s,5s,8s,11s)-1-Acetyl-1,4-Diaza-3-Keto-5-Carboxy-10-Thiatricyclo[2.8.1 .0(4,8)]Tridecane (Ac-Hel1-Oh). *J. Org. Chem.* **1991**, *56*, 6672-6682.

(37) Kataoka, T.; Beusen, D. D.; Clark, J. D.; Yodo, M.; Marshall, G. R. The Utility of Side-Chain Cyclization in Determining the Receptor-Bound Conformation of Peptides - Cyclic Tripeptides and Angiotensin-II. *Biopolymers* **1992**, *32*, 1519-1533.

(38) Plucinska, K.; Kataoka, T.; Yodo, M.; Cody, W. L.; He, J. X.; Humblet, C.; Lu, G. H.; Lunney, E.; Major, T. C.; Panek, R. L.; Schelkun, P.; Skeeane, R.; Marshall, G. R. Multiple Binding Modes for the Receptor-Bound Conformations of Cyclic Aii Agonists. *J. Med. Chem.* **1993**, *36*, 1902-1913.

(39) Kemp, D. S.; Allen, T. J.; Oslick, S. L. The Energetics of Helix Formation by Short Templated Peptides in Aqueous Solution. 1. Characterization of the Reporting Helical Template Ac-Hel₁. *J. Am. Chem. Soc.* **1995**, *117*, 6641-6657.

(40) Lewis, A.; Wilkie, J.; Rutherford, T. J.; Gani, D. Design, construction and properties of peptide N-terminal cap templates devised to initiate alpha-helices. Part 2. Caps derived from N-[(2S)-2-chloropropionyl]-(2S)-Pro-(2S)-Pro-(2S,4S)-4-thioPro-OMe. *J. Chem. Soc. Perkin 1* **1998**, 3777-3793.

(41) Gutierrez, A.; Cativiela, C.; Laguna, A.; Gimeno, M. C. Bioactive gold(I) complexes with 4-mercaptoproline derivatives. *Dalton Trans.* **2016**, *45*, 13483-13490.

(42) Yuan, J. S.; Xiong, W.; Zhou, X. H.; Zhang, Y.; Shi, D.; Li, Z. C.; Lu, H. 4-Hydroxyproline-Derived Sustainable Polythioesters: Controlled Ring-Opening Polymerization, Complete Recyclability, and Facile Functionalization. *J. Am. Chem. Soc.* **2019**, *141*, 4928-4935.

(43) Verbiscar, A. J.; Witkop, B. Synthesis of Cis-4-Mercapto-L-Proline and Trans-4-Mercapto-L-Proline Derivatives. *J. Org. Chem.* **1970**, *35*, 1924-1927.

(44) Eswarakrishnan, V.; Field, L. Sulfinic Acids and Related-Compounds .13. Unsymmetrical Disulfides Based on Methyl 4-Mercaptobutanesulfinate and 4(S)-Mercaptoproline or 4(R)-Mercaptoproline. *J. Org. Chem.* **1981**, *46*, 4182-4187.

(45) Thomas, K. M.; Naduthambi, D.; Tririya, G.; Zondlo, N. J. Proline Editing: A Divergent Strategy for the Synthesis of Conformationally Diverse Peptides. *Org. Lett.* **2005**, *7*, 2397-2400.

(46) Naduthambi, D.; Zondlo, N. J. Stereoelectronic tuning of the structure and stability of the trp cage miniprotein. *J. Am. Chem. Soc.* **2006**, *128*, 12430-12431.

(47) Pandey, A. K.; Naduthambi, D.; Thomas, K. M.; Zondlo, N. J. Proline Editing: A General and Practical Approach to the Synthesis of Functionally and Structurally Diverse Peptides. Analysis of Steric versus Stereoelectronic Effects of 4-Substituted Proline on Conformation within Peptides. *J. Am. Chem. Soc.* **2013**, *135*, 4333-4363.

(48) Mauger, A. B.; Witkop, B. Analogs and Homologs of Proline and Hydroxyproline. *Chem. Rev.* **1966**, *66*, 47-86.

(49) Andreatta, R. H.; Nair, V.; Robertson, A. V.; Simpson, W. R. J. Synthesis of Cis and Trans Isomers of 4-Chloro-L-Proline 4-Bromo-L-Proline and 4-Amino-L-Proline. *Aust. J. Chem.* **1967**, *20*, 1493-1509.

(50) Tamaki, M.; Han, G. X.; Hruby, V. J. Practical and efficient synthesis of orthogonally protected constrained 4-guanidinoproline. *J. Org. Chem.* **2001**, *66*, 1038-1042.

- (51) Hooz, J.; Gilani, S. S. H. A Rapid Mild Procedure for Preparation of Alkyl Chlorides and Bromides. *Can. J. Chem.* **1968**, *46*, 86-87.
- (52) Verheyden, J. P.; Moffatt, J. G. Halo Sugar Nucleosides .3. Reactions for Chlorination and Bromination of Nucleoside Hydroxyl Groups. *J. Org. Chem.* **1972**, *37*, 2289-2299.
- (53) Appel, R. Tertiary Phosphane-Tetrachloromethane, a Versatile Reagent for Chlorination, Dehydration, and P-N Linkage. *Angew. Chem., Int. Ed.* **1975**, *14*, 801-811.
- (54) Knapp, S.; Darout, E. New Reactions of Selenocarboxylates. *Org. Lett.* **2005**, *7*, 203-206.
- (55) Barnett, R. E.; Jencks, W. P. Diffusion-Controlled Proton Transfer in Intramolecular Thiol Ester Aminolysis and Thiazoline Hydrolysis. *J. Am. Chem. Soc.* **1969**, *91*, 2358-2369.
- (56) Fuchs, O.; Trunschke, S.; Hanebrink, H.; Reimann, M.; Seitz, O. Enabling Cysteine-Free Native Chemical Ligation at Challenging Junctions with a Ligation Auxiliary Capable of Base Catalysis. *Angew. Chem., Int. Ed.* **2021**, *60*, 19483-19490.
- (57) Sakamoto, K.; Tsuda, S.; Mochizuki, M.; Nohara, Y.; Nishio, H.; Yoshiya, T. Imidazole-Aided Native Chemical Ligation: Imidazole as a One-Pot Desulfurization-Amenable Non-Thiol-Type Alternative to 4-Mercaptophenylacetic Acid. *Chem. Eur. J.* **2016**, *22*, 17940-17944.
- (58) Mino, T.; Sakamoto, S.; Hamachi, I. Recent applications of N-acyl imidazole chemistry in chemical biology. *Biosci. Biotechnol. Biochem.* **2021**, *85*, 53-60.
- (59) Naudin, E. A.; McEwen, A. G.; Tan, S. K.; Poussin-Courmontagne, P.; Schmitt, J. L.; Birck, C.; DeGrado, W. F.; Torbeev, V. Acyl Transfer Catalytic Activity in De Novo Designed Protein with N-Terminus of alpha-Helix As Oxyanion-Binding Site. *J. Am. Chem. Soc.* **2021**, *143*, 3330-3339.
- (60) Augustinack, J. C.; Schneider, A.; Mandelkow, E. M.; Hyman, B. T. Specific tau phosphorylation sites correlate with severity of neuronal cytopathology in Alzheimer's disease. *Acta Neuropathologica* **2002**, *103*, 26-35.
- (61) Jeganathan, S.; von Bergen, M.; Brutlach, H.; Steinhoff, H.-J.; Mandelkow, E. Global hairpin folding of tau in solution. *Biochemistry* **2006**, *45*, 2283-2293.
- (62) Jeganathan, S.; Hascher, A.; Chinnathambi, S.; Biernat, J.; Mandelkow, E. M.; Mandelkow, E. Proline-directed Pseudo-phosphorylation at AT8 and PHF1 Epitopes Induces a Compaction of the Paperclip Folding of Tau and Generates a Pathological (MC-1) Conformation. *J. Biol. Chem.* **2008**, *283*, 32066-32076.
- (63) Despres, C.; Byrne, C.; Qi, H.; Cantrelle, F. X.; Huvent, I.; Chambraud, B.; Baulieu, E. E.; Jacquot, Y.; Landrieu, I.; Lippens, G.; Smet-Nocca, C. Identification of the Tau phosphorylation pattern that drives its aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 9080-9085.
- (64) Bielska, A. A.; Zondlo, N. J. Hyperphosphorylation of tau induces local polyproline II helix. *Biochemistry* **2006**, *45*, 5527-5537.
- (65) Brister, M. A.; Pandey, A. K.; Bielska, A. A.; Zondlo, N. J. OGlcNAcylation and Phosphorylation Have Opposing Structural Effects in tau: Phosphothreonine Induces Particular Conformational Order. *J. Am. Chem. Soc.* **2014**, *136*, 3803-3816.
- (66) Liu, F.; Iqbal, K.; Grundke-Iqbal, I.; Hart, G. W.; Gong, C. X. O-GlcNAcylation regulates phosphorylation of tau: A mechanism involved in Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10804-10809.

- (67) Smet-Nocca, C.; Broncel, M.; Wieruszeski, J. M.; Tokarski, C.; Hanouille, X.; Leroy, A.; Landrieu, I.; Rolando, C.; Lippens, G.; Hackenberger, C. P. R. Identification of O-GlcNAc sites within peptides of the Tau protein and their impact on phosphorylation. *Mol. Biosyst.* **2011**, *7*, 1420-1429.
- (68) Yuzwa, S. A.; Shan, X. Y.; Macauley, M. S.; Clark, T.; Skorobogatko, Y.; Vosseller, K.; Vocadlo, D. J. Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation. *Nat. Chem. Biol.* **2012**, *8*, 393-399.
- (69) Broncel, M.; Krause, E.; Schwarzer, D.; Hackenberger, C. P. R. The Alzheimer's Disease Related Tau Protein as a New Target for Chemical Protein Engineering. *Chem.-Eur. J.* **2012**, *18*, 2488-2492.
- (70) Schwagerus, S.; Reimann, O.; Despres, C.; Smet-Nocca, C.; Hackenberger, C. P. R. Semi-synthesis of a tag-free O-GlcNAcylated tau protein by sequential chemoselective ligation. *J. Pept. Sci.* **2016**, *22*, 327-333.
- (71) Cantrelle, F. X.; Loyens, A.; Trivelli, X.; Reimann, O.; Despres, C.; Gandhi, N. S.; Hackenberger, C. P. R.; Landrieu, I.; Smet-Nocca, C. Phosphorylation and O-GlcNAcylation of the PHF-1 Epitope of Tau Protein Induce Local Conformational Changes of the C-Terminus and Modulate Tau Self-Assembly Into Fibrillar Aggregates. *Front. Mol. Neurosci.* **2021**, *14*, 661368.
- (72) Melander, W. R.; Jacobson, J.; Horvath, C. Effect of Molecular-Structure and Conformational Change of Proline-Containing Dipeptides in Reversed-Phase Chromatography. *J. Chromatogr.* **1982**, *234*, 269-276.
- (73) Gesquiere, J. C.; Diesis, E.; Cung, M. T.; Tartar, A. Slow Isomerization of Some Proline-Containing Peptides Inducing Peak Splitting during Reversed-Phase High-Performance Liquid-Chromatography. *J. Chromatogr.* **1989**, *478*, 121-129.
- (74) Oneal, K. D.; Chari, M. V.; McDonald, C. H.; Cook, R. G.; YuLee, L. Y.; Morrisett, J. D.; Shearer, W. T. Multiple cis-trans conformers of the prolactin receptor proline-rich motif (PRM) peptide detected by reverse-phase HPLC, CD and NMR spectroscopy. *Biochem. J.* **1996**, *315*, 833-844.
- (75) Glover, M. S.; Shi, L. Q.; Fuller, D. R.; Arnold, R. J.; Radivojac, P.; Clemmer, D. E. On the Split Personality of Penultimate Proline. *J. Am. Soc. Mass. Spec.* **2015**, *26*, 444-452.
- (76) Ahuja, P.; Cantrelle, F. X.; Huvent, I.; Hanouille, X.; Lopez, J.; Smet, C.; Wieruszeski, J. M.; Landrieu, I.; Lippens, G. Proline Conformation in a Functional Tau Fragment. *J. Mol. Biol.* **2016**, *428*, 79-91.
- (77) Urbanek, A.; Popovic, M.; Elena-Real, C. A.; Morato, A.; Estana, A.; Fournet, A.; Allemand, F.; Gil, A. M.; Cativiela, C.; Cortes, J.; Jimenez, A. I.; Sibille, N.; Bernado, P. Evidence of the Reduced Abundance of Proline cis Conformation in Protein Poly Proline Tracts. *J. Am. Chem. Soc.* **2020**, *142*, 7976-7986.
- (78) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* **1997**, *119*, 4325-4329.
- (79) Johnson, E. C. B.; Kent, S. B. H. Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* **2006**, *128*, 6640-6646.
- (80) Dery, S.; Reddy, P. S.; Dery, L.; Mousa, R.; Dardashti, R. N.; Metanis, N. Insights into the deselenization of selenocysteine into alanine and serine. *Chem. Sci.* **2015**, *6*, 6207-6212.

- (81) Gieselman, M. D.; Xie, L. L.; van der Donk, W. A. Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.* **2001**, *3*, 1331-1334.
- (82) Hondal, R. J.; Nilsson, B. L.; Raines, R. T. Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **2001**, *123*, 5140-5141.
- (83) Cadamuro, S. A.; Reichold, R.; Kusebauch, U.; Musiol, H.-J.; Renner, C.; Tavan, P.; Moroder, L. Conformational properties of 4-mercaptoproline and related derivatives. *Angew. Chem., Int. Ed.* **2008**, *47*, 2143-2146.
- (84) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505.
- (85) Zheng, J.-S.; Tang, S.; Qi, Y.-K.; Wang, Z.-P.; Liu, L. Chemical synthesis of proteins using peptide hydrazines as thioester surrogates. *Nature Protocols* **2013**, *8*, 2483-2485.
- (86) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.: Gaussian 09, Revision D.01. Gaussian, Inc.: Wallingford, CT, 2013.
- (87) Wenzell, N. A.; Ganguly, H. K.; Pandey, A. K.; Bhatt, M. R.; Yap, G. P. A.; Zondlo, N. J. Electronic and steric control of $n \rightarrow \pi^*$ interactions via N-capping: stabilization of the α -helix conformation without a hydrogen bond. *ChemBioChem* **2019**, *20*, 963-967.
- (88) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
- (89) Dunning, T. H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations. 1. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007-1023.
- (90) Tomasi, J.; Mennucci, B.; Cancès, E. The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211-226.
- (91) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275-280.
- (92) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810-2817.
- (93) Chai, J. D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615-6620.
- (94) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.

Graphical Table of Contents

