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 Prolines on Conformation within Peptides

Anil K. Pandey, **Devan Naduthambi**, **Krista M. Thomas**, and **Neal J. Zondlo**[*]
Department of Chemistry and Biochemistry, University of Delaware, Newark DE 19716

Abstract

Functionalized proline residues have diverse applications. Herein we describe a practical approach, proline editing, for the synthesis of peptides with stereospecifically modified proline residues. Peptides are synthesized by standard solid-phase-peptide-synthesis to incorporate Fmoc-Hydroxyproline (4R-Hyp). In an automated manner, the Hyp hydroxyl is protected and the remainder of the peptide synthesized. After peptide synthesis, the Hyp protecting group is orthogonally removed and Hyp selectively modified to generate substituted proline amino acids, with the peptide main chain functioning to "protect" the proline amino and carboxyl groups. In a model tetrapeptide (Ac-TYPN-NH$_2$), 4R-Hyp was stereospecifically converted to 122 different 4-substituted prolyl amino acids, with 4R or 4S stereochemistry, via Mitsunobu, oxidation, reduction, acylation, and substitution reactions. 4-Substituted prolines synthesized via proline editing include incorporated structured amino acid mimetics (Cys, Asp/Glu, Phe, Lys, Arg, pSer/pThr), recognition motifs (biotin, RGD), electron-withdrawing groups to induce stereoelectronic effects (fluoro, nitrobenzoate), handles for heteronuclear NMR ([19]F:fluoro; pentafluorophenyl or perfluoro-*tert*-butyl ether; 4,4-difluoro; [77]SePh) and other spectroscopies (fluorescence, IR: cyanophenyl ether), leaving groups (sulfonate, halide, NHS, bromoacetate), and other reactive handles (amine, thiol, thioester, ketone, hydroxylamine, maleimide, acrylate, azide, alkene, alkyne, aryl halide, tetrazine, 1,2-aminothiol). Proline editing provides access to these proline derivatives with no solution phase synthesis. All peptides were analyzed by NMR to identify stereoelectronic and steric effects on conformation. Proline derivatives were synthesized to permit bioorthogonal conjugation reactions, including azide-alkyne, tetrazinetrans-cyclooctene, oxime, reductive amination, native chemical ligation, Suzuki, Sonogashira, cross-metathesis, and Diels-Alder reactions. These proline derivatives allowed three parallel bioorthogonal reactions to be conducted in one solution.

[*]To whom correspondence should be addressed. zondlo@udel.edu, phone: +1-302-831-0197, fax +1-302-831-6335.

Introduction

Protein activity, defined broadly, results from the geometrically controlled three-dimensional arrangement of a series of functional groups. In folded proteins, protein activity relies on secondary structure formation, the hydrophobic effect, and other non-covalent interactions to organize functional groups in a manner that allows specific recognition of substrates and binding partners and catalysis. Because functions are effected through specific structures, there has been broad interest in developing ways to stabilize protein structures. In addition, nature is limited by the genetically encoded amino acids, combined with post-translational modifications, plus metals and organic co-factors that bind these amino acids. Enhanced activity and novel functions may be incorporated into proteins by expanding beyond the array of natural amino acids, for example by introducing conformationally restrained amino acids, novel non-native functionalities, or amino acids with an enhanced hydrophobic effect.[1]

The field of protein design and engineering aims to achieve the functions of proteins in novel structures and/or the development of novel functions in proteins. The goals include achieving protein function within smaller structures, developing new structural topologies, de novo catalytic activity, and the development of hyperstable proteins.[2] The use of conformationally restricted amino acids has had particular application in achieving enhanced protein stability and in stabilizing protein recognition epitopes within small peptides, especially in medicinal chemistry applications of peptides.[3]

Proline residues are unique among the canonical amino acids, due to the conformational restraint of backbone cyclization and the presence of a tertiary amide bond.[4] These structural characteristics limit the available conformations for proline residues, with the ϕ torsion angle restricted to $-65° \pm 25°$. The conformational restriction and absence of a hydrogen bond donor results in proline being preferentially observed in certain structural contexts, including secondary structure termination, loops, turns, and polyproline helices (PPII). Proline residues are also preferentially employed in biomolecular recognition based on their distinction from other canonical amino acids, the hydrophobicity of the pyrrolidine ring, and the possibility of favorable aromatic-proline interactions.[5]

Proline has two key conformational equilibria: endo versus exo ring pucker, and trans versus cis amide bond (Figure 1).[4] Proline ring pucker correlates with protein ϕ and ψ main chain conformation, with an exo ring pucker favoring more compact conformations (PPII, α_R) and an endo ring pucker favoring more extended conformations.[6] Trans versus cis amide bond conformation defines the ω torsion angle, with the cis conformation strongly preferring the endo ring pucker. Thus, control of proline ring pucker permits control of all protein backbone torsion angles (ϕ, ψ, ω).[7]

The most abundant human protein, collagen, exhibits obligatory hydroxylation of proline at certain positions (Yaa of the XaaYaaGly collagen repeat, consensus sequence ProHypGly) to 4R-hydroxyproline (Hyp).[8] Proline hydroxylation results in a substantial enhancement of the stability of collagen, via the induction of a stereoelectronic effect.[9] This stereoelectronic effect is manifested in a preferred gauche relationship of the 4-hydroxyl substituent and the

carbon-amide bond (Figure 2, Figure 3, Figure 4). The sterically disfavored gauche conformation is preferred due to favorable hyperconjugative interactions from overlap of electron-rich σ_{C-H} orbitals and electron-deficient σ^*_{C-X} orbitals (X = amide or hydroxyl in Hyp).[10] This favorable gauche interaction leads to a preference for 4*R*-Hyp to adopt the exo ring pucker. Analogously, the non-native 4*S*-hydroxyproline (hyp) diastereomer leads to a preference for the endo ring pucker. The strength of this interaction depends on how electron-withdrawing the 4-substitutent is, with a more electron-withdrawing group leading to a greater stereoelectronic effect and a greater conformational preference for the sterically disfavored conformation. Thus, the incorporation of 4*R*-fluoroproline (Flp) in place of Hyp in collagen mimetic peptides leads to enhanced collagen triple helix stability due to the more electron-withdrawing nature of fluorine versus hydroxyl, and thus a greater preference for the exo conformation required at the Yaa position of collagen. In contrast, incorporation of non-electron-withdrawing 4*R*-methyl substituent, with a steric preference for an anti conformation, or an electron-withdrawing 4*S*-fluoro (flp) substituent, leads to a bias toward the endo ring pucker. The endo ring pucker is destabilizing at the Yaa position of collagen but stabilizing at the Xaa position, which is typically a Pro residue with an endo ring pucker and a more extended conformation.[11]

The stereospecific effect of proline 4-substitution on ring pucker is also manifested in amide cis-trans isomerism (ω torsion angle). The exo ring pucker stabilizes the trans amide bond, whereas an endo ring pucker is strongly favored in a cis amide bond. In proteins, proline amide conformation at specific residues is conserved evolutionarily.[12] Moreover, proper amide conformation is necessary for protein function. Cis-trans isomerism is often a rate-determining step in protein folding, with prolyl isomerases critical enzymes for rapid folding of proteins and for prevention of protein misfolding.[13] In addition, there are a growing number of examples of proteins with two alternating functions, one of which is mediated by having a prolyl cis amide and another with a trans amide.[14] Slow or incorrect prolyl cis-trans isomerism has also been implicated pathologically in protein misfolding. For example, overexpression of the prolyl isomerase Pin1 slows tau misfolding and neurofibrillary tangle formation in cellular models of Alzheimer's disease.[15] Strategies to stabilize cis or trans amide bonds thus can directly impact protein structure, stability, and function.

In addition to the stereoelectronic (gauche) effect on controlling ring pucker, and thus protein main chain conformation, conformations in peptides containing proline residues are affected by a favorable n→π* interaction between adjacent carbonyls (Figure 3).[9b, 16] In the n→π* interaction, which is also observed in non-proline residues,[16c] there is orbital overlap between a lone pair (n) on the carbonyl oxygen of the *i* residue with the π* orbital of the carbonyl at the carbon of the *i*+1 residue, with a near-ideal Bürgi-Dunitz trajectory internuclear angle. This n→π* interaction is strongest with an exo ring pucker on the proline residue, and fundamentally impacts the conformational preferences of proline residues, thus connecting proline ring pucker with protein secondary structure.

The ability to control protein structure through the incorporation of stereodefined substituted proline residues, either by favoring particular conformations or through enhanced hydrophobic effect, is an emerging strategy to increase protein stability and/or activity. Proline derivatives have been incorporated into collagen mimetics and other advanced

materials, in neurotransmitter receptor proteins, in enzymes, in synthetic polyproline scaffolds, and in globular proteins, in addition to many peptides with diverse applications.[9, 14d, 16a, 17] Certain proline derivatives (particularly fluoroprolines) have been incorporated into expressed proteins, either site-specifically through amber suppression/ orthogonal aminoacyl tRNA synthetases or by global incorporation at all proline amino acids in proline auxotrophic bacterial strains.[14d, 17c, 17i, j, 17n, 18] In every case, protein stability, activity, and/or function is dependent on both the identity and the stereochemistry of the substitution, allowing modulation of function in a manner that is predictable based on the known conformational preferences of proline derivatives.

The presence of proline in critical biological recognition and structural motifs has inspired the synthesis of a wide range of proline derivatives.[19] Interest in the synthesis of modified proline residues has also risen due to the reemergence and substantial advances in proline-catalyzed organic reactions.[20] There is an extensive history of development of novel proline derivatives, particularly for applications in medicinal chemistry, which has been the stimulus for the synthesis of most proline derivatives.[5d, 19, 21] Proline derivatives have also been significantly applied to the field of collagen mimetics, where multiple copies of proline derivatives are incorporated in the collagen triple helix, allowing modulation of stability of synthetic collagens.[9,11,17a-f,h,p-q]

A range of creative strategies has been developed for the synthesis of proline derivatives, both based on cyclization strategies and based on modification reactions on the commercially available and inexpensive Hyp. The structural effects of proline modification have been extensively measured and categorized based on a combination of steric effects and stereoelectronic effects. In addition to synthetic proline derivatives, it has also been observed that hydroxyproline may be natively phosphorylated (observed in rat crystallin protein and in diatom biosilica) and glycosylated (observed in plant cell walls).[22] Moreover, within bacteria, fungi, sponges, algae, and higher plants, an even wider range of proline analogues has been identified, in compounds with antibiotic, antifungal, and other activities.[19a, 19c]

While a significant number of proline derivatives have been synthesized, the application of these novel proline amino acids is restricted by the substantial solution phase synthetic effort to make each derivative as an Fmoc or Boc amino acid suitable for solid phase peptide synthesis, with typically 5-7 steps required per amino acid, including protection and deprotection strategies for the amine and carboxylic acids. For the preparation of a series of peptides, each with a different modified proline amino acid, for example to test structure-activity relationships, there is the requirement to synthesize each protected amino acid in solution, then incorporate each into a peptide, repeating all amide coupling steps N-terminal to the site of modification. While there has been an increase in the commercial availability of proline derivatives in recent years, highlighting the interest in their applications, these amino acids are generally quite expensive, other than the naturally derived 4R-Hyp.[23]

We recently communicated an approach, termed proline editing, to address the synthetic challenges of preparing peptides with proline derivatives.[24] In proline editing, the commercially available and inexpensive amino acid Fmoc-4R-Hyp is incorporated in a

peptide and orthogonally protected in an automated manner. After the peptide synthesis is completed, the hydroxyproline trityl protecting group is removed and the hydroxyl group selectively modified in an otherwise fully protected peptide on solid phase. This approach was applied in a model peptide, Ac-TYPN-NH$_2$, and was used to synthesize 4R- and 4S-fluoro, hydroxyl, and O-nitrobenzoate proline-containing peptides, as well as the 4-oxoproline and 4,4-difluoroproline derivatives. That work was subsequently applied to incorporate proline derivatives in a more complex peptide, Andersen's trp cage miniprotein, and used to modulate the stability of the trp cage in a manner that was predictable based on stabilization or destabilization of the observed exo ring pucker at residue 12 of the trp cage.[17r] This strategy fundamentally employs the peptide main chain as a "protecting group," abrogating the need to protect the amine and carboxylic acid functional groups. It is also a fundamentally diversity-oriented strategy, since a single fully synthesized peptide can be converted into a series of peptides exhibiting diverse functional groups and/or structural effects. Herein we provide a full description of the proline editing approach, considerably expanding the scope of the proline derivatives that may be incorporated into a peptide by the method and broadly examining the effects of incorporated proline derivatives on structure and in applications.

Results

Design of the model peptide

Proline editing was examined within a model tetrapeptide context[24-25] that was chosen to contain a central tyrosine-proline sequence (residues i+1 and i+2) that promotes cis amide bonds via a favorable interaction between the aromatic and proline rings (Figure 5).[5a, b, 5d, e, 5h, 26] The i (Thr) and i+3 (Asn) residues were selected to further promote cis amide bond formation.[25-26, 26c, 27] This context permits the ready examination of the structural effects of proline modification by ^{1}H NMR. The parent peptide Ac-TYPN-NH$_2$ exhibits a $K_{trans/cis}$ = 2.7 in aqueous solution at room temperature, with all amide protons resolved and $^3J_{\alpha N}$ values measurable to allow identification of the effects of proline modification on peptide main chain conformation. This model peptide has also been applied to examine aromatic effects on peptide structure, replacing tyrosine with electon-rich and electron-deficient aromatic amino acids as well as control peptides containing Ala and cyclyhexylalanine (Cha).[5h, 24, 28] In this context, it was observed that aromatic electronics correlated with cis-trans isomerism ($K_{trans/cis}$), with electron-rich aromatics favoring cis amide bond via a CH/π interaction between the aromatic ring and the prolyl ring, while electron-deficient aromatics, as well as Cha and Ala, relatively favored a trans amide bond due to the absence of a favorable CH/π interaction. By combining electronic and stereoelectronic effects, it was demonstrated that peptides containing aromatic-proline sequences could be designed to strongly promote cis (Ac-TWflpN-NH$_2$, $K_{trans/cis}$ = 0.65) or trans (Ac-T(4-NO$_2$-Phe)HypN-NH$_2$, $K_{trans/cis}$ = 20.1) amide bonds. In the current work, a series of peptides was synthesized via proline editing to examine the effects of proline substitution on cis-trans isomerism, via a combination of stereoelectronic and steric effects on proline conformation and thus consequently on peptide main chain conformation.

Synthesis of protected Hyp peptides

The central concept of proline editing is to incorporate a Hyp residue within a peptide and to subsequently modify the Hyp using stereospecific reactions at the γ-hydroxyl. Hyp is commercially available as an Fmoc amino acid either with a free hydroxyl group or protected as the *t*-butyl ether. Since *t*-butyl is a common protecting group in Fmoc solid phase peptide synthesis, and is not readily removed selectively on solid phase, proline editing was accomplished via coupling of the unprotected Fmoc-Hyp-OH via standard solid phase peptide synthesis (Scheme 1). The Hyp hydroxyl group was then trityl-Protected using trityl chloride in an automated manner, via programming as a double coupling with trityl chloride (and optionally imidazole) in place of an amino acid in the second coupling position. The trityl-Protected peptide was then subjected to standard solid phase peptide synthesis to complete the synthesis of the peptide, which proceeded cleanly. The effectiveness of this approach was demonstrated both within the model peptide and within the trp cage miniprotein.[17r, 24, 28a] After peptide synthesis was completed, the trityl group could be selectively removed with 2% TFA and subjected to reaction chemistry on the free hydroxyl.

In order to potentially allow the incorporation of multiple different proline derivatives within a peptide, we also examined alternative and orthogonal protecting group strategies for the Hyp hydroxyl.[29] Using an analogous approach, we were able to readily protect the Hyp as a silyl (TBS) ether using TBSCl with subsequent deprotection with TBAF on solid phase (Scheme 2). We also were able to protect the hydroxyl as an Alloc carbonate, although this method required elevated temperature and was not conducted as part of an automated synthetic procedure (Scheme 3). In addition, we also found that direct acylation of the Hyp hydroxyl after coupling proceeded readily as an alternative, using the orthogonally removable (NaN_3/MeOH)[30] nitrobenzoate ester or via acylation and alkylation to generate the bromoacetyl and subsequently azidoacetyl derivatives (Scheme 4). The direct acylation of Hyp approach was also applied in the synthesis of some modified trp cage miniproteins via proline editing.[17r]

Selective and stereospecific modification of Hyp residues in peptides on solid phase

After peptide synthesis and Hyp deprotection, the free Hyp-containing peptide could potentially be subjected to a series of modification reactions to control stereochemistry, to promote specific structures, and to introduce novel functional groups into peptides (Scheme 5). The reactions examined include the generation of sulfonates and other leaving groups for substitution reactions; Mitsunobu and related stereospecific inversion reactions; acylation; and oxidation. Because we sought to synthesize peptides containing any given functional group with defined stereochemistry at the 4-position, we examined the conduct of these reactions on either the 4*R* (trans relative stereochemistry, indicated with upper case in 3-letter codes and red lettering) or 4*S* (cis relative stereochemistry, indicated with lower case and blue lettering) hydroxyprolines. The coupled 4*R*-Hyp (**1**) was readily converted on solid phase to 4*S*-hyp (**4**) via Mitsunobu reaction with 4-nitrobenzoic acid and subsequent selective removal of the nitrobenzoate with NaN_3 in MeOH (Scheme 6, Figure 6), providing access to both 2*S*-hydroxyproline diastereomers (Hyp (4*R*) and hyp (4*S*)) on solid phase.[30]

Synthesis of hydoxyproline sulfonates and S$_N$2 reactions

We examined the conversion of hydroxyproline-containing peptides to incorporate sulfonate leaving groups, for modification on solid phase or in solution (Scheme 7).[19a, 21b, 21e, 21j, k, 31] Hydroxyproline residues were readily converted to include leaving groups including mesylates, tosylates, *para*-nosylates, and *ortho*-nosylates. These derivatives were stable to standard TFA cleavage conditions and isolable as the sulfonates (see below for structural analysis of the sulfonates and all derivatives). Attempts to synthesize the more reactive triflates and tresylates for solid phase S$_N$2 reactions did not succeed and resulted in formation of multiple side products (data not shown).

The application of sulfonates for the synthesis of diverse 4-substituted prolines via S$_N$2 reactions was examined with a series of nucleophiles (Scheme 8), including thiolate, azide, and iodide. These reactions were applied to generate the respective 4-substituted prolines, including the highly versatile 4-azidoprolines (4*R* (**18**) and 4*S* (**21**)) with good conversion.[17o, p, 21f, 32] In addition, the azidoproline-containing peptides were converted to the ionizable aminoprolines (**19**, **22**) and guanidinoprolines (**20**, **23**) via reduction of the azide to the amine and guanylation.[21l, 31d, 33] In general, the proline mesylates (**5**, **6**) and *para*-nosylates (**9**, **10**) were similarly effective for S$_N$2 reactions on solid phase (see HPLC chromatograms in the Supporting Information), whereas the tosylates (**7**, **8**) were less effective, resulting in poorer conversion and/or more side products.[21j, k]

Synthesis of 4*S*-substituted prolines via S$_N$2 reaction involved a straightforward 2-step protocol of conversion of Hyp to a sulfonate followed by S$_N$2 reaction on the sulfonate (Scheme 7 and Scheme 8). In contrast, synthesis of 4*R*-substituted prolines via S$_N$2 chemistry as described above involved Mitsunobu inversion with nitrobenzoic acid, azide-mediated release of the nitrobenzoate ester, sulfonylation to generate a leaving group, and then S$_N$2 reaction. While these 4-step sequences proceeded with high conversion in several cases, a shorter protocol to access 4*R*-subsituted prolines via S$_N$2 chemistry would be preferable. Alcohols can be converted to halide leaving groups with inversion via Mitsunobu-like reactions with appropriate halide sources. Therefore, as proof of principle, we investigated the synthesis of the 4*R*-thiophenyl proline-containing peptide (**16**) via direct conversion of Hyp to 4*S*-bromoproline (**25**), followed by S$_N$2 reaction with the thiophenolate (Scheme 9).[31d, 34] Both reactions proceeded with good overall conversion, providing simpler two-step access to 4*R*-substituted prolines via S$_N$2 reaction.

Fluorination reactions

Fluoroprolines have been the most broadly employed 4-substituted proline derivatives after hydroxyproline, due to the strong stereoelectronic effect induced stereospecifically by fluorine substitution, which results in strong structural preferences (Figure 2).[1l, m, 9, 16a, 17a-f, 17h-m, 35] Using the appropriate hydroxyprolines, both the 4*R* (Flp, **27**) and 4*S* (flp, **26**) fluoroproline-containing peptides were synthesized using conditions analogous to those developed for the solution phase synthesis of these amino acids (Scheme 10).[36] The flp-containing peptide is readily accessible directly in one step from Hyp (Figure 6a, entry 3). While the Flp peptide requires a three-step peptide modification from Hyp, including two stereoinversions, it nonetheless proceeds with good conversion (Figure 6a, entry 5).

Comparison of the NMR of the Flp-containing peptide synthesized via proline editing and the peptide synthesized using commercially available ($265/1 g) Fmoc-Flp-OH revealed identical NMR spectra, confirming that these reactions proceeded stereospecifically and without alpha-epimerization (see the Supporting Information for details).

Mitsunobu reactions

Mitsunobu reactions are among the most broadly employed reactions for the modification of alcohols, occurring with stereospecific inversion of stereochemistry.[37] Mitsunobu reactions require a relatively acidic proton ($pK_a < 11$) on the nucleophile, and thus were examined with carboxylic acids, phenols, thiophenols, selenophenol, phthalimide, hydroxyphthalimide, and thiolacetic acid (Scheme 11). Mitsunobu conditions were applied for the solid phase modification of both 4*R*- and 4*S*-hydroxyproline-containing peptides to generate a series of esters, ethers, and thioethers.

Mitsunobu reactions of phenol derivatives to yield aryl prolyl ethers proceeded readily, including with the modestly sterically hindered 2,6-dimethylphenol and with the electron-deficient pentafluorophenol (Scheme 11, **29**-**38**).[21b, 21i, 38] Other aryl prolyl ethers synthesized included cyanophenyl, which could be used as a spectroscopic probe (IR or fluorescence, see below), and aryl halides (I, Br) for palladium-mediated cross-coupling reactions. These phenyl ethers could also be employed to tune the recognition properties of proline-containing peptides.[21b, 21i, 38] While all other Mitsunobu reactions herein were conducted with DIAD, the Mitsunobu reactions with iodophenol and bromophenol both proceeded to higher conversion using the more reactive ADDP than they did with DIAD.

Thiophenol and selenophenol are larger chalcogen analogues of phenol. In addition to longer C-S and C-Se bonds and larger electron clouds, and thus greater sterics than oxygen, the thiophenyl and selenophenyl ethers are subject to modification by oxidation, and thus their properties tunable by oxidation. Mitsunobu reactions on solid phase with Hyp and hyp proceeded readily for 4-nitrothiophenol and selenophenol (derived from in situ reduction of diphenyl diselenide with sodium borohydride) (Scheme 11, **39**-**42**).[31b] In contrast, Mitsunobu reaction with thiophenol only proceeded to high conversion for hyp, to generate the 4*R* phenyl thioether (**16**). The 4*S* phenyl thioether (**17**) was instead synthesized by S_N2 reaction (Scheme 8), which was also alternatively used to synthesize the 4*R* phenyl thioether (Scheme 8 and Scheme 9).[21b]

4-Thioproline (4-mercaptoproline, Mpc (4*R*)/mpc (4*S*)) has recently been employed to prepare collagen mimetics that are readily modified by alkylation, oxidation, or pH-dependent structural switching.[39] 4*R*-thioproline has also been demonstrated to be highly effective in native chemical ligation reactions at proline.[40] Given the range of structural, redox, nucleophilic, and metal-binding properties of cysteines and the emerging role of cysteine oxidation (disulfide, glutathionylation, sulfenic acid, sulfinic acid, S-nitrosylation) in intracellular post-translational modifications, 4-thioproline could potentially have broad applications if there were increased accessibility.[21g, 28b, 41] 4-Thioproline-containing peptides were readily synthesized via Mitsunobu reaction with thiolacetic acid, to generate the thioesters (**43**, **45**), followed by deesterification to generate the free thiols (**44**, **46**),

which was conducted on solid phase and can also be effected in solution after TFA cleavage.[21g, 28b]

Mitsunobu reactions in general proceed poorly with amines. However, amines can be prepared via Mitsunobu reaction with phthalimide, followed by hydrazine deprotection. In addition, Mitsunobu reactions with phthalimide-protected hydroxylamines allow the introduction of aminoxy substitutents into molecules, with the possibility after hydrazine deprotection of a functional group capable of rapid chemoselective oxime formation.[42] The phthalimide and hydroxyphthalimide Mitsunobu reactions proceeded readily on solid phase on both Hyp- and hyp- containing peptides to generate the desired products (Scheme 11, **47**, **48**, **49**, **51**).[21h] The hydroxyphthalimides were also deprotected on solid phase with hydrazine to generate after TFA cleavage the aminoxy peptides (**50**, **52**). Notably, a previous synthesis of the protected Fmoc 4*R*-aminoxyproline required 9 steps of solution phase synthesis.[42b]

Halides are excellent leaving groups for substitution chemistry and also allow modulation of steric and stereoelectronic effects as a function of halogen size and electronegativity. The 4*S*-chloro, -bromo, and -iodo prolines (Scheme 11, **53**, **25**, **24**) were all synthesized cleanly under Mitsunobu conditions and were stable to TFA cleavage and under aqueous conditions (Scheme 11).[19a, 32a, 33b, 43] 4*S*-Bromoproline was employed directly on solid phase for the S_N2 reaction with thiophenol (Scheme 9, see above).[31d, 34a] 4*S*-Iodoproline was generated more cleanly under Mitsunobu conditions than by S_N2 reaction with sodium iodide and 4*R*-hydroxyproline nosylate (above, Scheme 8).

Of particular interest was the application of the Mitsunobu reaction with perfluoro-*tert*-butanol (pK_a 5) to incorporate a perfluoro-*tert*-butyl ether with either stereochemistry (**55**, **56**).[44] *tert*-Butyl groups have broad applications in medicinal chemistry and catalysis due to their sterics, hydrophobicity, and symmetry, which permits target binding with a reduced cost in conformational entropy. *tert*-Butyl groups have similar advantages in amino acids and peptides, with *tert*-leucine observed in pharmaceuticals and catalysts due to a strong conformational bias and steric effect.[45] A perfluoro-*tert*-butyl group would be expected to have enhanced steric and hydrophobic effects over a *tert*-butyl group, while also introducing nine equivalent fluorines that would provide a strong singlet signal in ^{19}F NMR, suggesting its possible use as a functional probe.[1e, 1l, m, 44b, 46] Mitsunobu reactions of perfluoro-*tert*-butanol proceeded with good conversion on solid phase to generate both 4*R*- and 4*S*-perfluoro-*tert*-butylhydroxyproline ethers (**55**, **56**) within peptides (Scheme 11). These reactions were more synthetically demanding than other Mitsunobu reactions, and proceeded substantially more effectively on lower loading resin (0.3-0.4 mmol/g) than on standard or higher loading resin (0.6-0.9 mmol/g). Other reactions that benefited from lower loading resin are specifically indicated as such in the Supporting Information.

Acylation reactions and reactions on acylates

Acylation of alcohols is a facile reaction that, combined with the large number of available carboxylic acids, could provide ready access to a wide range of substituted proline derivatives, including those that might induce significant stereoelectronic effects (due to the electron-withdrawing effects of the substituent) and those incorporated for the functional or

recognition properties of the conjugated substituent. Hydroxyproline-containing peptides were readily acylated on solid phase using DIC/DMAP, providing access to peptides containing diverse functional groups. (Scheme 12). Acylation could be conducted on both Hyp and hyp peptides; alternatively, the formation of the 4S esters of hyp can be more directly conducted by Mitsunobu reaction on Hyp with the relevant carboxylic acids (Scheme 9).

Acylation was applied to incorporate a broad range of functionalities, including modified proline residues with enhanced hydrophobic effect, proteinogenic side chain functional groups, reactive groups, and affinity ligands (Scheme 12). Hydrophobic groups added to proline include benzoates, acetates, and pivaloylate, which differ in hydrophobicity and in alkyl versus aryl functionality, and which could be used to optimize target binding (**3**, **57**-**61**).[47] Acylation was particularly effective for the incorporation of polar amino acid side chain functional groups. Functionalities added include ammonium (lysine mimetics (**62**, **63**), via β-alanine), guanidinium (arginine mimetics (**64**, **65**), via the guanidino acid[33c, 33k, 48] of β-alanine), and carboxylates (aspartic/glutamic acid mimetics (**66**-**70**), via maleic anhydride, succinic anhydride, or glutaric anhydride).[21e, 49] In addition, amino acids or peptides could be directly incorporated at the site of the hydroxyproline. This approach, via a β-alanine linker, allowed the incorporation of a cysteine residue (1,2 aminothiol functionality (**73**)), for native chemical ligation, and of an RGD peptide (**72**), for cell surface recognition.[1h, 50]

Acylation also permitted the incorporation of other reactive functional groups for subsequent further modification. Functionalities thus introduced via acylation reactions included α-bromo acetates (**74**, **75**) for S_N2 reactions, biotin (**76**, **77**) for affinity recognition, maleimide (**78**) for Michael reaction with thiols or for Diels-Alder reactions, alkyne (**79**) for Huisgen [3+2] cycloaddition, NHS ester (**80**) for reaction with amine nucleophiles, acrylate (**81**) for polymerization, and tetrazine (**82**) for tetrazine-trans-cyclooctene ligation.[51] These reactions were all conducted on solid phase, and the peptides containing reactive functionalities isolated in good yield after standard TFA cleavage/deprotection (typically using TFA/TIS/H_2O and avoiding thiol additives such as ethanedithiol that could react with these functional groups, particularly maleimide, fumarate, alkyne, and NHS esters). The acrylate ester was particularly sensitive to reaction and cleavage/deprotection conditions, requiring oxygen and hydroquinone as inhibitors of polymerization; even under these conditions the acrylate exhibited evidence of polymerization and substantially lower overall desired product formation than other reactions, though the peptide was still isolable and characterized by NMR.

S_N2 reactions are among the most versatile in their capability to stereospecifically incorporate a range of functional groups into molecules. The mesylate, tosylate, p-nosylate, o-nosylate, iodo, bromo, chloro, and bromoactetyl-substituted prolines thus represent a versatile range of derivatives for subsequent peptide modification. Bromoacetates exhibit enhanced S_N2 reactivity due to the adjacent carbonyl. In addition, substitution at the primary carbon is faster and provides a spacer element between the proline and potential conjugates. As proof of principle for the application of the hydroxyproline bromoacetate as an alternative site of modification, we examined a series of S_N2 reactions on this peptide on solid phase (Scheme 13). These reactions proceeded to high conversion with azide, alkyne,

primary amine, aniline, and thiol nucleophiles (**83**-**90**). These reactions proceeded under milder conditions than those on proline 4-sulfonates (Scheme 8), with all reactions proceeding at room temperature, in contrast to elevated temperature and/or the presence of a crown ether in reactions on the corresponding proline-4-sulfonates. In addition, some nucleophiles that did not react cleanly or at all on the proline 4-sulfonates (phenyl acetylene, allyl mercaptan) proceeded with high conversion on the bromoacetates. The azide and allyl thioether acetates also subsequently reacted more rapidly in bioorthogonal ligation reactions than the equivalent derivatives directly substituted on proline (see below), presumably due to reduced steric hindrance. In sum, these data indicate that hydroxyproline bromoacetates are an effective alternative site for modification of proline residues with nucleophiles.

Oxidation reactions

Alcohols may be oxidized under a variety of conditions to the corresponding carbonyls. Following established solution conditions for oxidation of protected hydroxyprolines,[36c] the Hyp-containing peptide was readily oxidized to 4-oxoproline (**91**) (Scheme 14).[31a] 4-Oxoproline contains an interesting ketone electrophilic handle that is not normally present in proteins. Notably, ketone-containing acetophenone and benzophenone derivatives of phenylalanine have been exploited as electrophiles in ligation reactions in proteins.[52]

Oxidation reactions could also modulate the structure of 4-substituted proline derivatives, through either steric or stereoelectronic effects. The proline thiophenyl ethers could be selectively oxidized to either the sulfoxide or sulfone derivatives (Scheme 14, **92**-**95**).[21b] The sulfoxides were obtained as an inseparable mixture of diastereomeric sulfoxides. The 4R-diastereomeric sulfoxides (**92**) were produced in an approximately 1:1 ratio and had apparently similar effects on cis/trans isomerism, though the NMR spectra exhibited distinct resonances in the amide region for the sulfoxide diastereomeric species. In contrast, the 4S-substituted thioether appeared to produce one sulfoxide diastereomer preferentially over the other (**93**), presumably due to the greater steric restraints present in the 4S diastereomer (Figure 3). Notably, the 4S-selenophenyl ether could be oxidized to generate the 3,4-dehydroproline (**96**) elimination product,[31b] which modulates the proline ring and main chain conformations and introduces an internal alkene to the peptide.[19a, 53]

Phosphorylation is a major intracellular post-translational modification of proteins.[54] Phosphorylation of proteins can lead to protein-protein interactions, or alternatively or additionally may lead to protein structural changes, particularly within natively disordered regions of proteins.[15a, 45g, 55] The TYPN peptide model system provides a context for potentially understanding the inherent stereoelectronic effects of phosphorylation, to provide a basis for understanding structural effects of protein phosphorylation. In addition, recent work has found that proteins may be phosphorylated and glycosylated on hydroxyproline residues, suggesting a richer range of post-translational modifications accessible in this residue.[22] In order to address these questions, both 4R and 4S diastereomers of hydroxyproline were modified to the phosphorylated and to the diethylphosphate triester forms (Scheme 14, **97**-**100**). Notably, phosphorylation ionization state changes as a function of pH, and therefore the structure of peptides with phosphorylated hydroxyproline could change as a function of pH.

Subjection of the 4-oxoproline-containing peptide to the fluorinating reagent DAST generated the 4,4-difluoroproline-containing peptide (**101**) (Scheme 15).[17b, 21d, 36c] 4,4-Difluoroproline is a noteworthy amino acid because it is conformationally similar to Pro, exhibiting a mixture of exo and endo ring puckers, while reducing the activation barrier for cistrans isomerism.[17c] Thus, given that cis-trans isomerism is commonly a rate-determining step in protein folding, peptides or proteins containing 4,4-difluoroproline could exhibit enhanced rates of protein folding. 4,4-Difluoroproline also could be exploited for a potentially enhanced hydrophobic effect in proteins without significantly changing the inherent conformational preferences of the protein.[1e, 1l, m] 4,4-Difluoroproline also introduces two ^{19}F atoms as potential NMR probes of protein folding and function (see below). Notably, 4,4-difluoroproline has been incorporated in expressed proteins in proline auxotrophs, although the ^{19}F NMR spectra of these proteins have not been reported.[17c, 18b]

Ketones are otherwise exploitable as chemical handles in peptides. Peptide cleavage/ deprotection of the 4-oxoproline-containing peptide in the presence of the typical thiol scavenger ethanedithiol resulted in isolation in high yield of the 4,4-dithiolane-containing peptide (**102**) (Scheme 16).[21b, 21d] Notably, the proline 4,4-dithiolane amino acid is part of the ACE inhibitor drug Spirapril.[21c] This reaction could also potentially be exploited to conjugate other thiols into 4-oxoproline-containing peptides. Alternatively, the 4-oxoproline-containing peptide (**91**) was cleanly obtained using TIS and water as scavengers.

Ketones are readily converted to amines using reductive amination. Reductive amination was accomplished on the 4-oxoproline-containing peptide on solid phase using isobutyl amine and sodium triacetoxyborohydride (Scheme 16), generating the substituted 4-aminoproline as a mixture of diastereomers (**104**, **105**), with the 4*R* diastereomer the major diastereomer (9:1 4*R*:4*S*). The stereochemistry of the major and minor diastereomers obtained by reductive amination was confirmed by the stereospecific synthesis of both diastereomers via S_N2 reactions. Alternatively, 4-aminoproline could react with an aldehyde on the solid phase to generate a substituted 4-aminoproline (**103**). Reductive amination provides a handle to incorporate both amines and the functional groups associated with the amine.

Organometallic reactions on solid phase

The derivatives described above include a range of chemical handles capable of application in organometallic reactions, including azides, alkenes, alkynes, and aryl halides. We initially examined the exploitation of these groups for peptide modification on solid phase (Scheme 17). The 4*S*-hydroxyproline bromophenyl ether was subjected to standard conditions for the Suzuki reaction, resulting in facile generation of the cross-coupled product (**106**) of reaction with the boronic acid of *p*-anisole. Similarly, the 4*S*-hydroxyproline iodophenyl ether readily underwent a Sonogashira reaction with phenyl acetylene on solid phase to generate the disubstituted alkyne (**107**). The azidoproline-containing peptides also reacted on solid phase with phenyl acetylene to form either the 4- or 5-substituted triazoles (**108**, **109**, **112**, **113**), depending on the use of copper or ruthenium catalyst.[51e, 56] The azidoproline peptides also readily underwent copper-mediated azide-alkyne coupling with 1-ocytne (**110**, **111**).

These approaches may be useful for solid phase diversification of peptide ligands in medicinal chemistry and other applications.[38]

Solution phase and bioorthogonal modification of peptides using proline derivatives

Unnatural amino acids containing unique functional groups can introduce reactivity not present within canonical amino acids.[57] Aryl halides, azides, alkynes, alkenes, 1,2-aminothiols, maleimides, tetrazines, and hydroxylamines incorporated via proline editing thus provide a diverse and complementary toolkit for the site-specific modification of peptides and proteins within defined structural contexts. In order to examine the possible application of these 4-substituted proline residues in bioorthogonal ligations, these derivatives were subjected to established solution reaction conditions.

Organometallic reactions provide unique selectivity in modification at distinct functional groups. In particular, palladium-mediated cross-coupling reactions have achieved preeminent status due to high chemoselectivity under highly diverse reaction conditions. Peptides containing proline iodophenyl and bromophenyl ethers were examined under conditions developed by Davis and Lin for aqueous Suzuki and Sonogashira cross-coupling reactions (Scheme 18).[58] These palladium-catalyzed reactions proceeded rapidly (30-45 minutes reaction at 37 °C) and cleanly to generate products (**106**, **107**) with high conversion in mild, neutral aqueous conditions. Suzuki-Miyaura reactions proceeded effectively with both the aryl bromide and the aryl iodide under these conditions, as did the palladium-catalyzed Sonogashira reaction with the aryl iodide, using the alternative dimethylamino ligand.

As expected, both the 4-azidoproline and the 4-azidoacetate of hydroxyproline reacted readily with the model alkyne phenyl acetylene in the [3+2] Huisgen azide-alkyne coupling under copper-mediated conditions (**109**, **114**).[17o, p, 32c, 59] In general, the azidoacetate reacted more rapidly than the azidoproline, likely due to reduced steric restriction, though both were effective in azide-alkyne coupling reactions.

Alkenes provide an alternative, orthogonal handle for chemical reactivity. Allylic groups substituted with chalcogens provide enhanced reactivity in cross-metathesis reactions, with reactivity Se > S > O.[60] The allyl thioether acetate of hydroxyproline reacted rapidly (30 minutes, 37 °C) in cross-metathesis with allyl alcohol in water (**115**), in a manner similar to reactions previously described with S-allyl cysteine, O-allyl tyrosine, and S-allyl-thiophenylalanine.[28b, 60]

Bioorthogonal reactions that do not require a metal catalyst have particular value in intracellular and extracellular labeling and in situations where removal of metal catalysts would be problematic.[57, 61] Four types of reactions were examined for peptide modification at proline under metal-free conditions: native chemical ligation reaction via 1,2-aminothiol (**116**), Diels-Alder reaction via maleimide (**117**), inverse electron demand tetrazine-trans-cyclooctene cycloaddition via conjugated tetrazine (**118**) (Scheme 19), and oxime reactions (Scheme 20) via 4-aminoxyproline (**119**-**122**) or via 4-oxoproline (**123**).[42, 51a-d, 62] All model reactions proceeded cleanly in high conversion. The oxime reaction with 4-oxoproline proceeded more slowly and required aniline as a nucleophilic catalyst to achieve

high conversion (**123**).[42a] The fastest reactions in this group were tetrazine-trans-cyclooctene ligations (**118**) and oxime reactions between aldehydes and 4*S*-aminoxyproline (**119**-**122**). The 4*S*-aminoxyproline reactions proceeded without added nucleophilic catalyst and allowed the rapid incorporation of diverse functional substituents, including 4-fluoro-benzaldehyde (**122**), which has been synthesized in one step in ^{18}F-labeled form and subsequently incorporated via oxime chemistry in RGD peptides, hormone peptides, the hormone protein leptin, and a ligand of the CXCR4 chemokine receptor for application in ^{18}F-positron emission tomography (PET).[63]

The range of functionalities incorporated in 4-substituted prolines provides the possibility of performing multiple bioorthogonal ligations in parallel, with potential future applications of incorporating multiple modifications within a single protein.[62b, 64] As proof of principle, peptides containing tetrazine, hydroxylamine, and maleimide functionalities were incubated in a single solution with the reactive partners trans-cyclooctene, isovaleraldehyde, and cyclopentadiene. These reactive functionalities ligated selectively and rapidly (30 minutes) in aqueous solution in good yield with their reactive partners in the presence of the alternative reactive groups. Notably, all of these reactions are metal-free. These data suggest that substituted prolines could be broadly applied as sites of multiple selective ligations within peptides and proteins. In this context, it is noteworthy that proline residues commonly occur in loops and turns and in natively disordered sequences, and are thus often at or near the surface or proteins.[5g, 65]

Structural analysis of the stereospecific effects of proline substitution in model peptides

Stereospecific substitutions on proline residues can significantly impact peptide and protein conformation (Figures 2, 3, 4). 4*R*-Hydroxylation (Hyp) of collagen at the Yaa position is obligatory for collagen stability and function, and 4*R*-fluorination (Flp) provides enhanced stability in collagen model peptides. In contrast, 4*S*-hydroxylation (hyp) or 4*S*-fluorination (flp) at the same position dramatically destabilizes collagen model peptides, due to the stereoelectronic preference for an endo ring pucker in these derivatives.[8-9, 17a, 17d-f, 17h] The magnitude of stereoelectronic effects on conformation allows modulation of the stability of collagen and other proline-containing proteins. However, stereoelectronic effects can be counteracted by steric effects. For example, 4*S*-methyl proline may substitute effectively at the Yaa position of collagen model peptides because this substitution leads to a steric preference for an anti relationship between the methyl substituent and the amide about the Cγ-Cδ bond, and thus a preference for an exo ring pucker (Figure 2).[7a, 11, 66] Thus, the interplay between steric and stereoelectronic effects provides alternative approaches to tune proline side chain and main chain conformation.

The derivatives described above represent a continuum of proline substitutions, representing both steric and stereoelectronic effects modulated by substitution, with both relative stereochemistries. Numerous proline derivatives have been analyzed for their conformational effects as the methyl esters of acetylated amino acids.[8-9,11,17a-c,o,39,43,47,66,70] Many of these substituted prolines have also been examined in collagen model peptides, where multiple copies of the proline derivative are incorporated in the collagen triple helix (3-30 proline substitutions, depending on single site (host-guest)

modification (three copies via one copy in each strand of the triple helix) or global substitution at the Xaa or Yaa site in collagen model peptides ((ProHypGly)$_n$, n = 7-10) (3 × n = 21-30 substituted prolines)). The Ac-TYPN-NH$_2$ peptide is an intermediate structural context, in which substitutions can impact the conformation of adjacent amino acids without impacting tertiary structure, and thus complements the limiting cases of isolated amino acids (Ac-*Prox*-OMe) and the tertiary structure of a collagen triple helix (e.g. (Pro*Prox*Gly)$_{7-10}$ or (*Prox*ProGly)$_{7-10}$, where *Prox* = 4-substituted proline). These peptides were analyzed by NMR spectroscopy (Figure 7, Table 1, Table 2, Table 3, and Supporting Information) to provide a readout of stereoelectronic versus steric effects for proline substituents, revealing additional context on the conformational effects of substitution for established proline derivatives and new data on the conformational effects of proline substitution for a wide range of previously unsynthesized or unstudied proline derivatives.

In addition to providing data on the conformational effects of proline substitution, the NMR spectroscopy also confirms that the reactions proceeded stereospecifically, as expected. As noted above, one method used to confirm that sequences of multiple reactions proceeded stereospecifically was the conversion of Ac-TYHypN-NH$_2$ to Ac-TYhypN-NH$_2$ (stereospecific Mitsunobu inversion) and the subsequent modification of the hyp-containing peptide with DAST on solid phase to generate Flp (Scheme 6, Scheme 10). The NMR spectrum of the peptide containing Flp synthesized via proline editing was identical to that of the peptide containing Flp synthesized using commercial Fmoc-Flp-OH, confirming that the two sequential stereospecific inversions had proceeded with expected stereochemistry and more generally indicating that all peptides derived from both Hyp and hyp were originating from the expected stereochemistry.

In general, the NMR spectra of peptides with the same stereochemistry and similar electronics were relatively comparable in global appearance, while peptides with the same substituent but opposite stereochemistry were divergent in appearance, indicating that the main effects of modification were determined by the stereochemistry and electronics of the 4-substitutent (Figure 7 and Supporting Information). However, one potential alternative side reaction of proline modification chemistry is epimerization at the alpha carbon. Thus, in addition to the 2*S*,4*R* (Hyp) and 2*S*,4*S* (hyp) hydroxyproline derivatives, the peptides with 2*R*,4*S* (ᴅ-hyp) and 2*R*,4*R* (ᴅ-hyp) hydroxyprolines were synthesized and analyzed by NMR spectroscopy, in order to identify the NMR signatures of peptides with alpha epimerization. These peptides were synthesized via coupling of Fmoc-ᴅ-hyp-OH and Fmoc-ᴅ-hyp-OH, which were prepared by solution phase synthesis using controlled epimerization of Hyp, subsequent modification by standard solution phase synthetic methods, and full characterization of the amino acids.[17g, 67] Comparison of the Ac-TY*Prox*N-NH$_2$ peptides containing all four hydroxyproline stereoisomers indicates distinct NMR spectral signatures for all four stereoisomers (Figure 8). In particular, the NMR signatures of peptides with 2*R* stereochemistry (ᴅ-hyp (**125**) and ᴅ-hyp (**126**)) are easily distinguished from those containing 2*S* stereochemistry, as would be expected based on the substantial literature describing the large conformational effects of replacement of ʟ- with ᴅ- amino acids in peptides. For proline these conformational effects are particularly significant because of the obligatory switch of proline from the left side of the Ramachandran plot to the right side of the Ramachandran

plot. Interestingly, the peptide containing D-hyp exhibited expected cis-trans isomerism about the Tyr-D-hyp bond, whereas the peptide containing D-hyp exhibited no evidence of cis amide bond, with the peptides exhibiting very different NMR spectra from each other. Given the number of peptides employing D-Pro-Gly, D-Pro-L-Pro, and other D-Pro sequences in turns and in cyclic peptides, these data suggest substantial potential applications in the use of stereoelectronic effects to control conformation in heterochiral peptides.[1k, 2h, 3d, 20i, j, 68]

Global analysis of the peptides allows the comparison of the effects of proline substitution on cis-trans isomerism in Ac-TY*Prox*N-NH$_2$ peptides as a function of the stereochemistry and the nature of the substituent (Figure 7, Table 1, Table 2, Table 3). Three key parameters were examined initially: $K_{trans/cis}$, the equilibrium constant of trans to cis prolyl amide bond (Figure 1); the free energy of proline modification on cis-trans isomerism relative to proline ($\Delta\Delta G_{trans/cis} = \Delta G_{trans/cis}$ (peptide) $- \Delta G_{trans/cis}$ (Pro)); and the effect of proline modification on the backbone conformation at Tyr in the cis conformation ($^3J_{\alpha N}$ = coupling constant between H^N and H$_\alpha$, which can be correlated to ϕ via a parametrized Karplus relationship, with smaller values indicating more compact conformations).[69] We also noted specific changes in Thr Hγ-methyl and Asn H^N chemical shifts in peptides as a function of proline substituent electronics and stereochemistry, indicating that proline 4-substitution changed the conformation of the entire peptide, and that the effects of proline modification should be considered within a broader structural context than cis-trans isomerism.[7a]

In total, 109 different 4-substituted prolines were analyzed by NMR within this peptide context (Figure 7, Table 1, Table 2, Table 3, and Supporting Information). For 41 substituents, data were collected on both the 4*R*- and the 4*S*-substituted variants of the functional group, and the free energy difference of conformational effects of the stereoisomeric substitutions on cistrans isomerism calculated (Table 2). Comparison of the NMR data of the peptides revealed three general trends. First, for electron-withdrawing 4-substitutents, the Ac-TY*Prox*N-NH$_2$ model peptide context provides a basis for the comparison of stereoelectronic versus steric effects. Based on these data, benzoates, substituted acetates, and fluorine had the largest stereoelectronic effects, with aryl ethers, hydroxyl, azide, and sulfonates displaying relatively smaller overall effects on cis-trans isomerism. Second, increasing the size of the atom connected to the proline resulted in reduction or elimination of the stereoelectronic effect as a function of steric size and/or reduced electronegativity, suggesting that steric effects were becoming dominant in determining conformation. Third, the effects of a substitution on conformation could be modulated by changes in protonation state, oxidation, or other chemical reactions (e.g. cycloaddition, oxime formation, or cross-coupling reaction).

Protonated-aminoxy, phthalimide-protected aminoxy, isobutylaminoacetoxy, benzoates, and fluoro were the most potent substitutions to change peptide conformation. In all of these cases, either a highly electron-withdrawing atom is directly connected to the proline ring, or an electron-withdrawing atom bound to additional electron-withdrawing atoms (e.g. nitrogen, carbonyl) was attached to the ring, enhancing the stabilization that could be provided by hyperconjugation (Figure 2) and thus enhancing the conformational bias induced by the substituent. Lesser effects were observed with hydroxyl, aryl ethers, sulfonates, and azide, with the effects of azide modification similar to those of hydroxyl.

These differences could be due to reduced inherent stereoelectronic effect, and/or due to steric interaction of these functional groups with the adjacent peptide backbone or side chains leading to an increased preference for the less sterically congested anti conformation, effects that would not be observed within acetylated amino acid methyl esters lacking this extended context.

Effects of proline substitution within the Ac-TY*Prox*N-NH$_2$ context could also potentially be influenced by interaction of the proline substituent with the tyrosine aromatic ring. This prospective interaction is more likely with the 4*R* substituent, which is on the same face of proline as that with which tyrosine interacts via the proline Hα in an aromatic-proline interaction (Figure 5). Therefore, in order to more thoroughly quantify stereoelectronic effects of peptides, for a representative subset of substitutions the Ac-TA*Prox*N-NH$_2$ peptides (**127**-**149**) were also synthesized and analyzed by NMR (Table 4, Table 5, and Supporting Information).[67b] These peptides lack a tyrosine ring that can interact with the proline. In general, proline editing reactions on Ac-TA*Prox*N-NH$_2$ peptides proceeded more readily (shorter reaction times, lower temperatures, and/or higher conversions) than those on Ac-TY*Prox*N-NH$_2$ peptides, presumably due to aromatic-proline interactions suppressing proline accessibility, suggesting that Ac-TY*Prox*N-NH$_2$ peptides, with sterically demanding groups both N-terminal (TyrO*t*Bu) and C-terminal (Asn(Trt)) to proline, are an appropriate test of the generality of proline editing.

The data on Ac-TA*Prox*N-NH$_2$ peptides broadly corroborated that of Ac-TY*Prox*N-NH$_2$ peptides. The order of stereoelectronic effects F ≥benzoates > OH ∼ N$_3$ ≥sulfonates was observed in both series. Benzoates exhibited lower overall effects on cis-trans isomerism in Ac-TA*Prox*N-NH$_2$ peptides than they did in Ac-TY*Prox*N-NH$_2$ peptides, although in 4*S* peptides the effects of benzoates were still greater than those of fluorine in Ac-TA*Prox*N-NH$_2$ peptides. Interestingly, even in Ac-TA*Prox*N-NH$_2$ peptides, which lack the possibility of an aromatic-proline interaction, the $^3J_{\alpha N}$ of Ala$_{cis}$ less than 5 Hz in 4*S*-substituted prolines indicated significant conformational restriction in the main chain (ϕ torsion angle) at alanine, consistent with type VIa1 β-turn formation. These data suggest that, within the context of a cis amide bond, an endo ring pucker more effectively promotes type VIa1 β-turns than an exo ring pucker does. Similarly, Asn$_{trans}$ also exhibited a chemical shift that was dependent on the stereochemistry and electronics of substitution (more downfield for electron-withdrawing 4*R* substitutents, more upfield for electron-withdrawing 4*S* substitutents), as it did in Ac-TY*Prox*N-NH$_2$ peptides, suggesting that these backbone effects were inherent to proline conformational restriction via 4-substitution.

The effects of sterics on switching proline conformational preferences were readily observed when the atom directly attached to the proline ring was substituted with larger atoms or was multiply substituted (Figure 7c). The 4*S*-N-phthalimide-substituted proline exhibited a strong preference for trans amide bond, despite the electronic effect of two carbonyls bound to the nitrogen. The trisubstituted nitrogen is apparently highly sterically demanding, leading to a conformational preference for trans amide bond despite the electron-withdrawing nature of the phthalimide. Similarly, replacement of an oxygen in an aryl ether with sulfur neutralized the conformational preferences of both 4*R* and 4*S* substitution, and substitution with selenium reversed the conformational preferences, with the 4*S*-selenophenyl ether

exhibiting a higher population of trans amide bond than the 4*R*-selenophenyl ether. Likewise, for mercaptoproline, the 4*S* configuration had a greater preference for trans amide bond than the 4*R* configuration, both in thiol and thiolate protonation states (Figure 7d, Table 3), as has been seen previously.[39] In addition, the 4*S*-iodoproline had a higher $K_{\text{trans/cis}}$ than the bromo, chloro, or fluoro derivatives, consistent with reduced stereoelectronic and increased steric effects for the iodide compared to the smaller halides.

Oxidation of the thiophenyl ether also resulted in an increase in sterically directed conformational effects (Figure 7d). For the thiophenyl ether, the 4*S* and 4*R* configurations exhibited similar $K_{\text{trans/cis}}$. Oxidation of the thioether to the sulfoxide led to similar conformational preferences, despite the presence of an additional electron-withdrawing oxygen and increased positive charge on sulfur that would be expected to enhance a stereoelectronic effect. Further oxidation to the sulfone led to a clear sterically dependent preference of 4*S* for trans conformation and 4*R* for cis conformation relative to proline, despite the even greater electron-withdrawing effects of the sulfone.[28b]

Two ionizable proline derivatives exhibited particularly noteworthy pH-dependent conformational preferences. First, 4-aminoproline exhibited non-standard conformational preferences when protonated, with the 4*S* substitution having a greater trans preference than the 4*R* substitution, despite an ammonium having a stronger stereoelectronic effect than an amine.[33d, 33f] These effects of anomalously high trans amide bond stabilization from an endo-favoring residue have been previously described for protonated 4*S*-aminoproline in both acetylated amino acid methyl esters and in polyproline helices.[33i, 70] This effect has been described as due to hydrogen bonding between the ammonium and the proline backbone carbonyl (here, the carbonyl conjugated to the Asn amide) for the protonated 4*S*-aminoproline, an interaction not possible for the 4*R*-aminoproline. Indeed, the chemical shift of Asn$_{\text{trans}}$ is highly aberrant for protonated 4*S*-aminoproline, consistent with a hydrogen bond between the ammonium and the carbonyl, with $\delta = 8.93$ ppm for protonated 4*S*-aminoproline compared to $\delta = 8.08$ ppm for hyp, $\delta = 8.18$ for flp, and $\delta = 7.91$ for 4*S*-azidoproline, typical values for electron-withdrawing 4*S*-substituents (Table 1, Table 3 Figure 6c). Interestingly, although the acetamide of 4-aminoproline also showed evidence of this intramolecular hydrogen bonding in the acetylated amino acid methyl esters, in our work herein the 4*S*-guanidinium-substituted proline did not exhibit this same effect, despite similarly favorable geometry and electrostatics.

Protein phosphorylation is an important post-translational modification whose structural effects are not well understood. Bielska and Zondlo observed that serine/threonine phosphorylation of peptides from the tau proline-rich domain induced a conformational change to polyproline helix, but the structural basis for this conformational change was not described.[45g, h] Phosphorylation could introduce conformational restriction on proline through steric effects, stereoelectronic effects, or through attractive or repulsive interactions with the peptide backbone. The TY*Prox*N and TA*Prox*N peptide contexts could be used to better understand the nature of the effects of phosphorylation on protein structure. Comparison of the phosphorylated hydroxyproline peptides with the non-phosphorylated hydroxyproline peptides suggested that phosphorylation indeed may induce a stereoelectronic effect, with a greater effect in the monoanionic form than the dianionic

form. These data are consistent with the idea that phosphorylation leads to conformational restriction through a combination of stereoelectronic effects and either repulsive or attractive backbone interactions, as the largest structural effects of phosphorylation are in the dianionic form. Interestingly, the neutral diethylphosphate exhibited a modestly reduced apparent stereoelectronic effect compared to the monoanionic or dianionic phosphates or to hydroxyproline, indicating that the diethyl phosphate is a more sterically demanding substituent.

Beyond these initial parameters on cis-trans isomerism, coupling constants, and amide proton chemical shift, the conformational preferences for a subset of peptides were evaluated more thoroughly with additional information provided by TOCSY data. Analysis of proline side chain chemical shifts revealed apparent signatures for proline ring pucker that are consistent with the effects on $K_{trans/cis}$ described above. For example, substitutions with larger $K_{trans/cis}$ than Pro also had a larger $\Delta\delta$ of the diastereotopic proline beta protons, whether the effect was due to an electron-withdrawing 4*R* substitutent or to a sterically demanding 4*S* substituent. In contrast, the opposing modifications that led to smaller $K_{trans/cis}$ than Pro (4*S*-electron-withdrawing, 4*R*-sterically demanding) had smaller $\Delta\delta$ of the diastereotopic beta protons. These trends were observed in both *TYProxN* and *TAProxN* peptides. A similar dichotomy was observed in the $\Delta\delta$ of the diastereotopic proline gamma protons, with 4*S*-electron-withdrawing and 4*R*-sterically demanding substituents (lower $K_{trans/cis}$) exhibiting a larger $\Delta\delta$ of the diastereotopic gamma protons, while the opposite substitutions (higher $K_{trans/cis}$) displayed a smaller $\Delta\delta$. Additional trends consistent with conformational preferences were observed in the N-terminal acetyl protons and the Thr methyl protons. Changes in ring pucker are also supported by changes in coupling patterns of Hβ and Hγ protons (see the Supporting Information for details). In total, the NMR data indicate that the effects of proline substitution are translated over the structure of the entire peptide as a function of the stereochemistry, stereoelectronic effects, and sterics of the 4-substituent.

To further examine the effects of proline substitution on peptide structure, a series of peptides containing *4R-, 4S-* and 4,4-disubstituted prolines was examined by circular dichroism (Figure 9). The circular dichroism signatures of the peptides will predominantly represent the conformation of peptides with the major trans amide bond, particularly for peptides with residues with a strong preference for an exo ring pucker *(4R* stereochemistry). The CD data were consistent with NMR data indicating that proline substitution modulates peptide main chain conformation as a function of both stereochemistry and electronics of substitution. In particular, proline *4R* and *4S* substitution led to conformational effects that were divergent in opposite directions from proline and with the stereoelectronic effects of azidoproline more modest in a peptide context than those of hydroxyproline, fluoroproline, or acetylated hydroxyproline.

Interestingly, 4,4-disubstitutued prolines, whose $K_{trans/cis}$ values were similar to proline herein and previously in acetylated amino acid methyl esters, exhibited CD spectra in peptides that were substantially different from each other and/or proline.[17b, c] 4,4-Difluoroproline (**101**) in a peptide exhibited a CD spectrum similar to proline (Figures 9a and 9g), suggestive of its potential application at the site of proline with minimal

thermodynamic effect on structure. In contrast, peptides containing 4-oxoproline (**91**) (Figure 9f) and proline-4,4-dithiolane (**102**) (Figure 9h) exhibited CD spectra that were substantially different from those of proline, as was also seen in their NMR spectra (Figure 7b). The 4,4-dithiolane in particular exhibited both unique NMR and CD spectra that were distinct from those of any other peptide examined in this study, and that in particular were very different from that of 4-oxoproline, presumably due to the greater steric demand of the 4,4-dithiolane. These data suggest that potentially distinct conformational effects may be induced with disubstituted prolines in peptides.

Proline derivatives as handles for IR, fluorescence, and ^{19}F and ^{77}Se NMR spectroscopy

Proline editing is capable of incorporating a range of functional groups into peptides, including derivatives that can serve as spectroscopic probes. The cyanophenyl ether derivative synthesized above includes an aryl nitrile, which could potentially be employed as a probe for either infrared or fluorescence spectroscopy, as has been described for 4-cyanophenylalanine.[71] Analysis of the IR spectrum (Figure 10) of the peptide Ac-TAP(4*S*-OPhCN)N-NH$_2$(**145**) in H$_2$O revealed a strong IR absorbance at 2233 cm^{-1}, a spectroscopically quiet region in the IR of proteins, with resultant high signal to noise. These data suggest that cyanophenyl hydroxyprolines may be useful infrared probes via site-specific incorporation at a proline residue.

Due to extended conjugation of the nitrile, cyanophenyl ethers also exhibit strong fluorescence. In order to understand the relationship of peptide structure to fluorescence, the fluorescence spectra of the peptides Ac-TYP(4*R*-OPhCN)N-NH$_2$(**35**), Ac-TYP(4*S*-OPhCN)N-NH$_2$(**36**), and Ac-TAP(4*S*-OPhCN)N-NH$_2$(**145**) were examined (Figure 11 and Supporting Information) and compared to the control peptide Ac-TYPN-NH$_2$. The fluorescence excitation and emission spectra of cyanophenylhydroxyproline-containing peptides reveal a strong absorbance with a broad excitation maximum λ_{max} = 262 nm and fluorescence emission at λ_{max} = 295 nm. Comparison of the data of tyrosine- versus alanine-containing peptides indicates evidence of fluorescence resonance energy transfer or quenching between cyanophenylhydroxyproline and tyrosine: the tyrosine-containing peptides exhibited both substantially reduced fluorescence compared to the alanine-containing peptide and stereochemistry-dependent fluorescence intensities. Comparing Ac-TAP(4*S*-OPhCN)N-NH$_2$ and Ac-TYPN-NH$_2$, cyanophenyl hydroxyproline had 7.5-fold greater fluorescence intensity than tyrosine with excitation at 265 nm. Based on its fluorescence emission wavelength, cyanophenyl hydroxyprolines are expected to be appropriate FRET donors for tryptophan, as has been described for cyanophenylalanine.[72]

Proline editing was employed to incorporate a range of substitutions into peptides. Among these modifications were derivatives that incorporate spin-1/2 nuclei that could be used as NMR probes, specifically probes that might be exploited site-specifically without background from natural sources.[73] ^{19}F NMR spectra were collected for fluorinated proline peptides (Figure 12), including those containing flp (**26**), Flp (**27**), 4,4-difluoroproline (**101**, **134**), pentafluorophenyl ethers (**33**, **34**), and perfluoro-*tert*-butyl ethers (**55**, **56**). ^{19}F NMR spectra are characterized in general by excellent chemical shift dispersion, a feature which was observed in fluorinated peptides, with an overall chemical shift range of over 100 ppm

in the fluorinated peptides: δ from −70 ppm (perfluoro-*tert*-butyl ethers) to −178 ppm (fluoroprolines), with difluoroprolines (∼ −100 ppm) and aryl fluorines (∼ −160 ppm) intermediate (Table 6). The ^{19}F NMR spectra corroborated the $K_{trans/cis}$ values observed by ^{1}H NMR. The perfluoro-*tert*-butyl hydroxyproline peptides were noteworthy for the fluorine signal existing as a sharp singlet, with 9 equivalent fluorines, suggesting potential applications as a magnetic resonance probe with high signal to noise.[44b, 46]

Interestingly, in the 4,4-difluoroproline-containing peptides (Figure 12cd), the diastereotopic fluorines exhibited almost identical chemical shifts when in the trans conformation, appearing as a pseudo-singlet for Ac-TYP(4,4-F$_2$)N-NH$_2$ (**101**, Figure 12c) due to Δδ < $^2J_{FF}$ and exhibiting strong second order effects due to similar chemical shifts (Δδ ∼ $^2J_{FF}$) in Ac-TAP(4,4-F$_2$)N-NH$_2$ (**134**, Figure 12d). In contrast, the diastereotopic fluorines exhibited substantial chemical shift dispersion (Δδ = 8-10 ppm) and $^2J_{FF}$ (= 252 Hz) coupling of the geminal fluorines in the cis conformation. This amide conformation-dependent chemical shift dispersion was also observed in Ac-TAP(4,4-F$_2$)N-NH$_2$, indicating that the dispersion was not due to interaction with the tyrosine aromatic ring but was an inherent effect dependent on amide trans versus cis conformation. These results suggest that 4,4-difluoroproline could potentially be a useful magnetic probe of prolyl amide cis-trans isomerism, which is difficult to easily identify within proteins.[74] Notably, 4,4-difluoroproline may be incorporated in expressed proteins in proline auxotrophic bacteria, although curiously no ^{19}F NMR spectrum has been reported for these proteins.[17c, 18b]

Selenium is a redox-sensitive trace element present in proteins as selenocysteine and selenomethionine.[75] ^{77}Se NMR has emerged as a useful spectroscopic probe due to the rare nature of selenium in proteins, its ready substitution for sulfur, and its large chemical shift dispersion that is highly sensitive to oxidation state. These properties result in information-rich NMR spectra with good signal to noise due to the absence of background ^{77}Se signals, despite only 7.6% natural abundance of ^{77}Se and a relatively low gyromagnetic ratio.[76] To examine the possible application of selenoproline analogues as NMR probes, the ^{77}Se NMR spectrum of the selenophenyl proline peptide Ac-TYP(4*R*-SePh)N-NH$_2$(**41**) was collected (Figure 13). The proton-decoupled data exhibit a sharp singlet at 299.9 ppm, compared to a chemical shift of 462.8 ppm for the reference compound PhSeSePh, consistent with large chemical shift dispersion in ^{77}Se NMR that makes it useful for probing oxidation and biological interactions.

Discussion

We have described a practical approach to the synthesis of peptides with proline residues stereospecifically modified at the 4-position with diverse functionalities and structural preferences. Proline editing involves the incorporation of the inexpensive amino acid Fmoc-4*R*-Hydroxyproline by standard solid phase synthesis, protection *in situ* in an automated manner on solid phase, and subsequent synthesis of the rest of the peptide. The fully synthesized peptide is then selectively deprotected at the hydroxyproline, and the hydroxyproline selectively and stereospecifically modified via reactions including Mitsunobu, sulfonylation, S$_N$2, acylation, and oxidation. The critical difference between solid phase modification via proline editing and solution phase reactions is the elimination

of the cumbersome need for protecting group chemistry and manipulation on the amine and carboxylic acid functional groups, with the peptide main chain serving as the "protecting group." In addition, because the modification reactions are performed on solid phase, the reactions are conducted with excess reagents that are washed away after each synthetic step, without the usual need in solution synthesis for the separation, isolation, and purification after each reaction step. In proline editing, the only purification step is standard HPLC purification required after any peptide synthesis. This practical aspect allows the rapid conduct of multistep synthetic sequences compared to similar sequences by solution phase chemistry. Moreover, typically solution syntheses of modified prolines are conducted on the more stable Boc- or Cbz-protected amines, then converted to Fmoc protection for solid phase peptide synthesis. Combining the practical advantages of solid phase synthesis with the lack of protection/deprotection/reprotection steps involving the amine and carboxylic acid, proline editing synthetic sequences are both fewer steps and more easily conducted than the sequences to prepare the equivalent Fmoc amino acids.

Proline editing provides access to a wide range of 4-substituted prolines with diverse substitution and control of stereochemistry, with both 4R and 4S substitution demonstrated for most proline derivatives. In total, peptides with 123 different 4-substituted proline amino acids were synthesized, of which 56 were novel amino acids never previously described, and many more were amino acids synthesized and used only once or just a small number of times. In contrast to proline editing, synthesis of each proline amino acid as an Fmoc amino acid for application in peptide synthesis would involve 5-9 steps of solution phase synthesis per amino acid, and would also involve the repetitive solid phase synthesis of each peptide with a different proline amino acid separately. In contrast, proline editing is both practical, with no solution phase synthesis, and diversity oriented, with a single fully synthesized peptide capable of being modified to generate a series of peptides with divergent proline substituents. Notably, the approach is general for several classes of reactions, and thus could be used to incorporate a broad range of additional functionalities not described in this work.

Functionalities incorporated include electrophiles (sulfonates, halides, ketone, maleimide, NHS ester, bromoacetate), nucleophiles (amine, thiol, aminoxy, 1,2-aminothiol), bioorthogonal functionalities (azide, alkyne, alkene, aryl halides, tetrazine), entities for polymerization in biomaterials (acrylate), affinity groups (biotin, RGD), fluorescence and IR spectroscopic probes (aryl nitrile, azide), NMR probes (^{77}Se, ^{19}F), hydrophobic groups, and native amino acid functional groups. Moreover, the reactive functional groups could be modified on solid phase and in neutral aqueous solution to further diversify the peptides, using reactions including azidealkyne coupling, Sonogashira and Suzuki-Miyaura cross-couplings, Diels-Alder reaction, Grubbs olefin cross-metathesis, native chemical ligation, reductive amination, tetrazine-transcyclooctene ligation, oxime formation, and reductive amination. Combining these reactions with a wide array of available reactive partners provides substantial opportunities for even broader functionalization of peptides containing these amino acids (e.g. rapid solution incorporation of furan, alkene, and fluorine functional groups demonstrated in oxime reactions in solution). Most notably, the range of functional groups incorporated in 4-substituted prolines included several different derivatives whose reactivities were orthogonal to one another, and thus which were capable of selective reaction in the presence of other functional groups capable of alternative selective reaction.

This work suggests the possibility of modifying a peptide with multiple functional groups capable of independent modification, and thus the possibility of multiple functionalizations of a single peptide in aqueous solution. As proof of principle of this concept, three parallel bioorthogonal reactions were conducted in aqueous solution using three peptides containing different functional groups in the presence of three different reactive small molecules, with rapid (30 minutes) and selective modification of all three peptides observed with excellent conversion.

4-Substitution of proline residues induces side chain conformational preferences that bias peptide and protein main chain conformation, in a manner dependent on the stereochemistry of the substitution and on the sterics versus electronics of the substitution. Specifically, proline 4R-substitution with electron-withdrawing groups or 4S-substitution with sterically demanding groups leads to a preference for the proline exo ring pucker, which leads to a stronger preference for trans amide bond and for more compact ϕ, ψ torsion angles (α_R, PPII). In contrast, proline 4S-substitution with electron-withdrawing groups or 4R-substitution with sterically demanding groups leads to a preference for the proline endo ring pucker, which leads to a greater relative preference for a cis amide bond or for more extended ϕ, ψ torsion angles. Due to these strong conformational biases, a variety of 4-substituted prolines have been broadly used to control peptide and protein structure, particularly within collagen mimetics but also in medicinal chemistry, biomaterials, and globular proteins. Proline editing was employed to incorporate 4-substituted prolines with divergent steric and electronic properties of substitution, and thus with a wide range of conformational properties. The effects of 4-susbstitution were quantified for 109 proline derivatives within a peptide context, which differs in context from simple acetylated amino acid methyl esters, which lack adjacent amino acids whose conformations could affect or be affected by the proline derivatives, and from the tertiary structure of collagen triple helices. NMR analysis of these proline derivatives revealed a continuum of conformational effects for the 4-substituted prolines described herein, and most notably indicated that the effects of proline modification were propagated beyond the proline amino acid to the entire tetrapeptide, with structural differences observed in all amino acids in the Ac-TY*Prox*N-NH$_2$ peptides that were a function of stereochemistry, sterics, and electronics of substitution. This work specifically allowed quantification of the structural effects of a wide range of proline 4-substitutents, including a majority that have not been previously examined, with 1.3 kcal/mol tunability in the free energy of cis-trans isomerism observed (9-fold change in $K_{\text{trans/cis}}$) in the proline derivatives studied and in general distinct structures observed for the 4R- versus 4S-substituted diastereomers of the peptides in all cases.

Among the derivatives synthesized were modified proline residues that contain amino acid functional groups and could function as structurally constrained mimics of the encoded amino acids Cys, Asp, Glu, Phe, Lys, Arg, Ser, and Thr, as well as the post-translationally modified amino acids phosphoserine and phosphothreonine (Figure 14).[21b, 21e, 21l, 31d, 33b, c, 77] Conformationally constrained amino acids and amino acid mimetics have a long history in medicinal chemistry to identify ligand binding conformations and to optimize ligands for high selectivity for targets, and have been also applied in protein and materials design to tune protein stability and function. The canonical amino acids have multiple available conformations, considering both ϕ, ψ space and χ space,

which is important to their plasticity to adopt different structures. In contrast, specific target binding usually involves the conformational entropy cost of adopting defined conformations. Limiting the available main chain and side chain conformations via conformationally restricted amino acid mimetics provides the possibility of tailored, structurally defined functions. This is particularly true given the greater flexibility of non-proline amino acids and the ability of proline-substituted amino acids to adopt most of the major protein secondary structure conformations, including α-helix and PPII via exo-favoring residues or more extended conformations via endo-favoring residues. Conformations where the main chain amide is a hydrogen bond donor are obviously not accessible via these proline-based amino acid mimetics, although this limitation is less significant in view of the critical role of loops and turns in the conformations of bound ligands, particularly those bound to GPCRs and other receptors.[68d, 78] Indeed, much of the original work to synthesize novel 4-substituted proline amino acids was conducted to increase ligand binding to GPCRs, integrin receptors, enzymes, and other proteins. Optimization of χ space in particular has proven fruitful in controlling ligand conformation and binding, due to the coupling of side chain conformation to main chain conformation.[3a-c] Therefore, numerous examples have been developed of conformationally restricted versions of the canonical amino acids, including a substantial number of proline-derived amino acid mimetics, and these amino acid mimetics have proven highly valuable in diverse applications, particularly in medicinal chemistry.[5d, 21b, 21d-g, 21i, j, 21l, 31c, d, 33b-k, 48b, c, 77, 79] The 4R- and 4S-substitituted proline amino acid derivatives described herein provide the ability to broadly functionalize proline beyond the minimal natural proline modifications, and to easily incorporate proteinogenic functional groups within structured contexts.

The key limitations of proline editing are the dependence on 4-substitution (due to the availability of commercial amino acid) and the dependence for direct modification of the proline ring on reactivity at an unactivated secondary carbon. Thus, while numerous S_N2 and Mitsunobu reactions were achieved successfully, several nucleophiles exhibited no significant product formation, including most significantly carbon nucleophiles (cyanide, alkyne, malonate) that were unreactive with sulfonates on solid phase. These functionalities potentially could be incorporated via the more activated bromoacetate electrophile, as was done using phenyl acetylene as a nucleophile in an S_N2 reaction, and was done under milder conditions for sodium azide, thiophenol, and isobutylamine as S_N2 nucleophiles. However, the ester linkage thus employed, while highly general, could be limiting for some applications, most notably situations in which the peptide would be exposed to esterases. The amide linkage with 4-aminoproline, demonstrated with the biotin derivative, could potentially be an alternative in these cases. Furthermore, while all of these reactions were conducted on functionally rich and relatively sterically hindered peptides (Ac-TY*Prox*N-NH$_2$, containing protected alcohol, phenol, and side chain (trityl-)amide functional groups and an aromatic-proline interaction that is persistent in organic solvents; and, in some cases, on the trp cage), obviously some reactions (most notably oxidation reactions) will be incompatible with certain functional groups (Met, Cys, and potentially Trp) in peptides.

Conclusion

The work herein describes proline editing, a general and highly practical approach to the synthesis of peptides modified at the 4-position of proline to incorporate a broad range of functional groups to control peptide structure, reactivity, and function. Potential applications of these functional groups were demonstrated in rapid, high yielding, and orthogonal modifications of peptides on solid phase and in neutral aqueous solution. Proline editing was also applied to broadly interrogate the effects of 4*R*- and 4*S*-substitutions on peptide structure within a peptide context. We expect this work to have broad applications in new approaches to synthesize structured functional peptides, to understand steric and stereoelectronic effects on peptide and protein structure, and to incorporate novel functionalities into peptides and proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

(1). (a) Young TS, Schultz PG. J Biol Chem. 2010; 285:11039. [PubMed: 20147747] (b) Liu CC, Schultz PG. Ann Rev Biochem. 2010; 79:413. [PubMed: 20307192] (c) Wu X, Schultz PG. J Am Chem Soc. 2009; 131:12497. [PubMed: 19689159] (d) Link AJ, Mock ML, Tirrell DA. Curr Opin Biotech. 2003; 14:603. [PubMed: 14662389] (e) Yoder NC, Kumar K. Chem Soc Rev. 2002; 31:335. [PubMed: 12491748] (f) DeGrado WF, Summa CM, Pavone V, Nastri F, Lombardi A. Ann Rev Biochem. 1999; 68:779. [PubMed: 10872466] (g) Neumann H, Hazen JL, Weinstein J, Mehl RA, Chin JW. J Am Chem Soc. 2008; 130:4028. [PubMed: 18321101] (h) Nguyen DP, Elliott T, Holt M, Muir TW, Chin JW. J Am Chem Soc. 2011; 133:11418. [PubMed: 21736333] (i) Johnson JA, Lu YY, Van Deventer JA, Tirrell DA. Curr Opin Chem Biol. 2010; 14:774. [PubMed: 21071259] (j) Chiu HP, Kokona B, Fairman R, Cheng RP. J Am Chem Soc. 2009; 131:13192. [PubMed: 19711980] (k) Struthers MD, Cheng RP, Imperiali B. Science. 1996; 271:342. [PubMed: 8553067] (l) Merkel L, Budisa N. Org Biomol Chem. 2012; 10:7241. [PubMed: 22890839] (m) Salwiczek M, Nyakatura EK, Gerling UIM, Ye SJ, Koksch B. Chem Soc Rev. 2012; 41:2135. [PubMed: 22130572]

(2). (a) Jiang L, Althoff EA, Clemente FR, Doyle L, Rothlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF, Hilvert D, Houk KN, Stoddard BL, Baker D. Science. 2008; 319:1387. [PubMed: 18323453] (b) Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Science. 2003; 302:1364. [PubMed: 14631033] (c) Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, Baker D. Nature. 2008; 453:190. [PubMed: 18354394] (d) Pochan DJ, Schneider JP, Kretsinger J, Ozbas B, Rajagopal K, Haines L. J Am Chem Soc. 2003; 125:11802. [PubMed: 14505386] (e) Dalal S, Balasubramanian S, Regan L. Nat Struct Biol. 1997; 4:548. [PubMed: 9228947] (f) Zondlo NJ, Schepartz A. J Am Chem Soc. 1999; 121:6938. (g) Neidigh JW, Fesinmeyer RM, Andersen NH. Nature Struct Biol. 2002; 9:425. [PubMed: 11979279] (h) am Ende CW, Meng HY, Ye M, Pandey AK, Zondlo NJ. ChemBioChem. 2010; 11:1738. [PubMed: 20623571] (i) Zondlo SC, Gao F, Zondlo NJ. J Am Chem Soc. 2010; 132:5619. [PubMed: 20361796]

(3). (a) Hruby VJ, Li G, Haskell-Luevano C, Shenderovich M. Biopolymers. 1997; 43:219. [PubMed: 9277134] (b) Cowell SM, Lee YS, Cain JP, Hruby VJ. Curr Med Chem. 2004; 11:2785. [PubMed: 15544476] (c) Hruby VJ. Acc Chem Res. 2001; 34:389. [PubMed: 11352717] (d) Anil B, Song B, Tang Y, Raleigh DP. J Am Chem Soc. 2004; 126:13194. [PubMed: 15479052] (e) Chapman RN, Dimartino G, Arora PS. J Am Chem Soc. 2004; 126:12252. [PubMed: 15453743] (f) Schafmeister CE, Po J, Verdine GL. J Am Chem Soc. 2000; 122:5891.(g) Shepherd NE, Hoang HN, Abbenante G, Fairlie DP. J Am Chem Soc. 2005; 127:2974. [PubMed: 15740134]

(4). Milner-White EJ, Bell LH, Maccallum PH. J Mol Biol. 1992; 228:725. [PubMed: 1469711]

(5). (a) Grathwohl C, Wüthrich K. Biopolymers. 1976; 15:2025. [PubMed: 963241] (b) Wu WJ, Raleigh DP. Biopolymers. 1998; 45:381. [PubMed: 9530015] (c) Pal D, Chakrabarti P. J Mol Biol. 1999; 294:271. [PubMed: 10556045] (d) Halab L, Lubell WD. J Am Chem Soc. 2002; 124:2474. [PubMed: 11890796] (e) Taylor CM, Hardré R, Edwards PJB, Park JH. Org Lett. 2003; 5:4413. [PubMed: 14602013] (f) Zarrinpar A, Bhattacharyya RP, Lim WA. Sci STKE. 2003; 2003:re8. [PubMed: 12709533] (g) Bhattacharyya R, Chakrabarti P. J Mol Biol. 2003; 331:925. [PubMed: 12909019] (h) Zondlo NJ. Acc Chem Res. 2012 A.S.A.P. 10.1021/ar300087y

(6). (a) Dunbrack RL Jr, Karplus M. J Mol Biol. 1993; 230:543. [PubMed: 8464064] (b) Lovell SC, Word JM, Richardson JS, Richardson DC. Proteins. 2000; 40:389. [PubMed: 10861930]

(7). (a) Shoulders MD, Satyshur KA, Forest KT, Raines RT. Proc Natl Acad Sci USA. 2010; 107:559. [PubMed: 20080719] (b) Vitagliano L, Berisio R, Mastrangelo A, Mazzarella L, Zagari A. Protein Sci. 2001; 10:2627. [PubMed: 11714932]

(8). Jenkins CL, Raines RT. Nat Prod Rep. 2002; 19:49. [PubMed: 11902439]

(9). (a) Eberhardt ES, Panasik N Jr, Raines RT. J Am Chem Soc. 1996; 118:12261. [PubMed: 21451735] (b) Bretscher LE, Jenkins CL, Taylor KM, DeRider ML, Raines RT. J Am Chem Soc. 2001; 123:777. [PubMed: 11456609] (c) DeRider ML, Wilkens SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT, Markley JL. J Am Chem Soc. 2002; 124:2497. [PubMed: 11890798] (d) Hodges JA, Raines RT. J Am Chem Soc. 2003; 125:9262. [PubMed: 12889933]

(10). Alabugin IV, Zeidan TA. J Am Chem Soc. 2002; 124:3175. [PubMed: 11902907]

(11). Shoulders MD, Hodges JA, Raines RT. J Am Chem Soc. 2006; 128:8112. [PubMed: 16787056]

(12). Lorenzen S, Peters B, Goede A, Preissner R, Frommel C. Proteins: Struct, Funct, Bioinf. 2005; 58:589.

(13). (a) Schmid FX. Adv Protein Chem. 2002; 59:243. [PubMed: 11868274] (b) Dugave C, Demange L. Chem Rev. 2003; 103:2475. [PubMed: 12848578]

(14). (a) Brazin KN, Mallis RJ, Fulton DB, Andreotti AH. Proc Natl Acad Sci USA. 2002; 99:1899. [PubMed: 11830645] (b) Andreotti AH. Biochemistry. 2003; 42:9515. [PubMed: 12911293] (c) Eckert B, Martin A, Balbach J, Schmid FX. Nature Struct Mol Biol. 2005; 12:619. [PubMed: 15937494] (d) Lummis SCR, Beene DL, Lee LW, Lester HA, Broadhurst RW, Dougherty DA. Nature. 2005; 438:248. [PubMed: 16281040]

(15). (a) Yaffe MB, Schutkowski M, Shen MH, Zhou XZ, Stukenberg PT, Rahfeld JU, Xu J, Kuang J, Kirschner MW, Fischer G, Cantley LC, Lu KP. Science. 1997; 278:1957. [PubMed: 9395400] (b) Lu PJ, Wulf G, Zhou XZ, Davies P, Lu KP. Nature. 1999; 399:784. [PubMed: 10391244]

(16). (a) Hinderaker MP, Raines RT. Protein Sci. 2003; 12:1188. [PubMed: 12761389] (b) Hodges JA, Raines RT. Org Lett. 2006; 8:4695. [PubMed: 17020280] (c) Bartlett GJ, Choudhary A, Raines RT, Woolfson DN. Nat Chem Biol. 2010; 6:615. [PubMed: 20622857] (d) Zondlo NJ. Nat Chem Biol. 2010; 6:567. [PubMed: 20644541] (e) Krow GR, Shoulders MD, Edupuganti R, Gandla D, Yu F, Sonnet PE, Sender M, Choudhary A, DeBrosse C, Ross CW, Carroll P, Raines RT. J Org Chem. 2012; 77:5331. [PubMed: 22607128]

(17). (a) Holmgren SK, Taylor KM, Bretscher LE, Raines RT. Nature. 1998; 392:666. [PubMed: 9565027] (b) Shoulders MD, Kamer KJ, Raines RT. Bioorg Med Chem Lett. 2009; 19:3859. [PubMed: 19423349] (c) Renner C, Alefelder S, Bae JH, Budisa N, Huber R, Moroder L. Angew Chem Int Ed. 2001; 40:923.(d) Barth D, Milbrandt AG, Renner C, Moroder L. ChemBioChem. 2004; 5:79. [PubMed: 14695516] (e) Malkar ND, Lauer-Fields JL, Borgia JA, Fields GB. Biochemistry. 2002; 41:6054. [PubMed: 11994000] (f) Persikov AV, Ramshaw JAM, Kirkpatrick A, Brodsky B. J Am Chem Soc. 2003; 125:11500. [PubMed: 13129344] (g) Levins CG, Schafmeister CE. J Am Chem Soc. 2003; 125:4702. [PubMed: 12696876] (h) Doi M, Nishi

Y, Uchiyama S, Nishiuchi Y, Nakazawa T, Ohkubo T, Kobayashi Y. J Am Chem Soc. 2003; 125:9922. [PubMed: 12914445] (i) Kim W, McMillan RA, Snyder JP, Conticello VP. J Am Chem Soc. 2005; 127:18121. [PubMed: 16366565] (j) Kim W, Hardcastle KI, Conticello VP. Angew Chem, Int Ed. 2006; 45:8141.(k) Crespo MD, Rubini M. PLoS One. 2011; 6:e19425. [PubMed: 21625626] (l) Holzberger B, Marx A. J Am Chem Soc. 2010; 132:15708. [PubMed: 20961065] (m) Zheng TY, Lin YJ, Horng JC. Biochemistry. 2010; 49:4255. [PubMed: 20405858] (n) Limapichat W, Lester HA, Dougherty DA. J Biol Chem. 2010; 285:8976. [PubMed: 20068044] (o) Kumin M, Sonntag LS, Wennemers H. J Am Chem Soc. 2007; 129:466. [PubMed: 17226990] (p) Erdmann RS, Wennemers H. J Am Chem Soc. 2010; 132:13957. [PubMed: 20849115] (q) Arnold U, Hinderaker MP, Köditz J, Golbik R, Ulbrich-Hofmann R, Raines RT. J Am Chem Soc. 2003; 125:7500. [PubMed: 12812474] (r) Naduthambi D, Zondlo NJ. J Am Chem Soc. 2006; 128:12430. [PubMed: 16984189]

(18). (a) Buechter DD, Paolella DN, Leslie BS, Brown MS, Mehos KA, Gruskin EA. J Biol Chem. 2003; 278:645. [PubMed: 12399455] (b) Kim WY, George A, Evans M, Conticello VP. ChemBioChem. 2004; 5:928. [PubMed: 15239049] (c) Merkel L, Schauer M, Antranikian G, Budisa N. ChemBioChem. 2010; 11:1505. [PubMed: 20572253]

(19). (a) Mauger AB, Witkop B. Chem Rev. 1966; 66:47.(b) Remuzon P. Tetrahedron. 1996; 52:13803.(c) Mauger AB. J Nat Prod. 1996; 59:1205. [PubMed: 8988607]

(20). (a) Eder U, Sauer G, Weichert R. Angew Chem Int Ed. 1971; 10:496.(b) Hajos ZG, Parrish DR. J Org Chem. 1974; 39:1615.(c) List B, Lerner RA, Barbas CF. J Am Chem Soc. 2000; 122:2395. (d) List B. Tetrahedron. 2002; 58:5573.(e) Bock DA, Lehmann CW, List B. Proc Natl Acad Sci USA. 2010; 107:20636. [PubMed: 21068369] (f) Quintard A, Langlois JB, Emery D, Mareda J, Guenee L, Alexakis A. Chem Eur J. 2011; 17:13433. [PubMed: 22038655] (g) Schmid MB, Zeitler K, Gschwind RM. Chem Sci. 2011; 2:1793.(h) Tanzer EM, Zimmer LE, Schweizer WB, Gilmour R. Chem Eur J. 2012; 18:11334. [PubMed: 22829538] (i) Sculimbrene BR, Morgan AJ, Miller SJ. J Am Chem Soc. 2002; 124:11653. [PubMed: 12296730] (j) Fiori KW, Puchlopek ALA, Miller SJ. Nat Chem. 2009; 1:630. [PubMed: 20161563]

(21). (a) Flippenanderson JL, Gilardi R, Karle IL, Frey MH, Opella SJ, Gierasch LM, Goodman M, Madison V, Delaney NG. J Am Chem Soc. 1983; 105:6609.(b) Krapcho J, Turk C, Cushman DW, Powell JR, Deforrest JM, Spitzmiller ER, Karanewsky DS, Duggan M, Rovnyak G, Schwartz J, Natarajan S, Godfrey JD, Ryono DE, Neubeck R, Atwal KS, Petrillo EW. J Med Chem. 1988; 31:1148. [PubMed: 2836590] (c) Smith EM, Swiss GF, Neustadt BR, McNamara P, Gold EH, Sybertz EJ, Baum T. J Med Chem. 1989; 32:1600. [PubMed: 2544729] (d) Karanewsky DS, Badia MC, Cushman DW, Deforrest JM, Dejneka T, Lee VG, Loots MJ, Petrillo EW. J Med Chem. 1990; 33:1459. [PubMed: 2329568] (e) Bridges RJ, Stanley MS, Anderson MW, Cotman CW, Chamberlin AR. J Med Chem. 1991; 34:717. [PubMed: 1671706] (f) Bhagwat SS, Fink CA, Gude C, Chan K, Qiao Y, Sakane Y, Berry C, Ghai RD. Bioorg Med Chem Lett. 1994; 4:2673.(g) Kolodziej SA, Nikiforovich GV, Skeean R, Lignon MF, Martinez J, Marshall GR. J Med Chem. 1995; 38:137. [PubMed: 7837225] (h) Bisang C, Weber C, Robinson JA. Helv Chim Acta. 1996; 79:1825.(i) Bellier B, McCortTranchepain I, Ducos B, Danascimento S, Meudal H, Noble F, Garbay C, Roques BP. J Med Chem. 1997; 40:3947. [PubMed: 9397175] (j) Lerner C, Siegrist R, Schweizer E, Diederich F, Gramlich V, Jakob-Roetne R, Zurcher G, Borroni E. Helv Chim Acta. 2003; 86:1045.(k) Lafrance D, Caron S. Org Process Res Dev. 2012; 16:409.(l) Qu HC, Cai MY, Mayorov AV, Grieco P, Zingsheim M, Trivedi D, Hruby VJ. J Med Chem. 2009; 52:3627. [PubMed: 19473029]

(22). (a) Poulsen N, Sumper M, Kröger N. Proc Natl Acad Sci USA. 2003; 100:12075. [PubMed: 14507995] (b) Kühlberg A, Haid M, Metzger S. J Biol Chem. 2010; 285:31484. [PubMed: 20682783] (c) Velasquez SM, Ricardi MM, Dorosz JG, Fernandez PV, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M, Verli H, Pauly M, Bacic A, Olsen CE, Ulvskov P, Petersen BL, Somerville C, Iusem ND, Estevez JM. Science. 2011; 332:1401. [PubMed: 21680836] (d) Taylor CM, Karunaratne CV, Xie N. Glycobiology. 2012; 22:757. [PubMed: 22190471]

(23). Fmoc-Hyp-OH is $135/25 g (Chem Impex). In contrast, other 4-substituted proline analogs that are commercially available are substantialy more expensive e.g. Fmoc-(2S,4R)-Flp (1 g, $ 337, Bachem), Fmoc-(2S,4S)-flp (1 g, $ 337, Bachem), Fmoc-L-(3,4-dehydroproline) (1 g, $ 470,

Chem Impex), Fmoc-L-4,4-F$_2$-proline (1 g, $ 524, Polypeptide Group), and Fmoc-(2S,4R)-proline(4-OPh) (1 g, $ 458, Polypeptide Group)

(24). Thomas KM, Naduthambi D, Tririya G, Zondlo NJ. Org Lett. 2005; 7:2397. [PubMed: 15932207]

(25). Meng HY, Thomas KM, Lee AE, Zondlo NJ. Biopolymers (Peptide Sci). 2006; 84:192.

(26). (a) Dyson HJ, Rance M, Houghten RA, Lerner RA, Wright PE. J Mol Biol. 1988; 201:161. [PubMed: 2843644] (b) Yao J, Dyson HJ, Wright PE. J Mol Biol. 1994; 243:754. [PubMed: 7966294] (c) Yao J, Feher VA, Espejo BF, Reymond MT, Wright PE, Dyson HJ. J Mol Biol. 1994; 243:736. [PubMed: 7966293] (d) Reimer U, Scherer G, Drewello M, Kruber S, Schutkowski M, Fischer G. J Mol Biol. 1998; 279:449. [PubMed: 9642049]

(27). Hutchinson EG, Thornton JM. Protein Sci. 1994; 3:2207. [PubMed: 7756980]

(28). (a) Thomas KM, Naduthambi D, Zondlo NJ. J Am Chem Soc. 2006; 128:2216. [PubMed: 16478167] (b) Forbes CR, Zondlo NJ. Org Lett. 2012; 14:464. [PubMed: 22224916]

(29). Isidro-Llobet A, Alvarez M, Albericio F. Chem Rev. 2009; 109:2455. [PubMed: 19364121]

(30). Gomez-Vidal JA, Forrester MT, Silverman RB. Org Lett. 2001; 3:2477. [PubMed: 11483039]

31. (a) Patchett AA, Witkop B. J Am Chem Soc. 1957; 79:185.(b) Rüeger H, Benn MH. Can J Chem. 1982; 60:2918.(c) Natchus MG, Bookland RG, De B, Almstead NG, Pikul S, Janusz MJ, Heitmeyer SA, Hookfin EB, Hsieh LC, Dowty ME, Dietsch CR, Patel VS, Garver SM, Gu F, Pokross ME, Mieling GE, Baker TR, Foltz DJ, Peng SX, Bornes DM, Strojnowski MJ, Taiwo YO. J Med Chem. 2000; 43:4948. [PubMed: 11150165] (d) Tamaki M, Han GX, Hruby VJ. J Org Chem. 2001; 66:1038. [PubMed: 11430070]

(32). (a) Andreatta RH, Nair V, Robertson AV, Simpson WR. J Aust J Chem. 1967; 20:1493.(b) Gomez-Vidal JA, Silverman RB. Org Lett. 2001; 3:2481. [PubMed: 11483040] (c) Sonntag LS, Schweizer S, Ochsenfeld C, Wennemers H. J Am Chem Soc. 2006; 128:14697. [PubMed: 17090057]

(33). (a) Bartra M, Romea P, Urpí F, Vilarrasa J. Tetrahedron. 1990; 46:587.(b) Webb TR, Eigenbrot C. J Org Chem. 1991; 56:3009.(c) Zhang ZY, Van Aerschot A, Hendrix C, Busson R, David F, Sandra P, Herdewijn P. Tetrahedron. 2000; 56:2513.(d) Babu IR, Ganesh KN. J Am Chem Soc. 2001; 123:2079. [PubMed: 11456840] (e) Crespo L, Sanclimens G, Montaner B, Pérez-Tomás R, Royo M, Pons M, Albericio F, Giralt E. J Am Chem Soc. 2002; 124:8876. [PubMed: 12137542] (f) Umashankara M, Babu IR, Ganesh KN. Chem Commun. 2003:2606.(g) Farrera-Sinfreu J, Giralt E, Castel S, Albericio F, Royo M. J Am Chem Soc. 2005; 127:9459. [PubMed: 15984873] (h) Sanclimens G, Crespo L, Giralt E, Albericio F, Royo M. J Org Chem. 2005; 70:6274. [PubMed: 16050687] (i) Sonar MV, Ganesh KN. Org Lett. 2010; 12:5390. [PubMed: 21049969] (j) Nanda M, Ganesh KN. J Org Chem. 2012; 77:4131. [PubMed: 22428628] (k) Fillon YA, Anderson JP, Chmielewski J. J Am Chem Soc. 2005; 127:11798. [PubMed: 16104758]

(34). (a) Koskinen AMP, Helaja J, Kumpulainen ETT, Koivisto J, Mansikkamaki H, Rissanen K. J Org Chem. 2005; 70:6447. [PubMed: 16050708] (b) Attempts to directly synthesize the mesylate with inversion via Mitsunobu reactions with methane sulfonic acid were unsuccessful.

(35). (a) Gerig JT, McLeod RS. J Am Chem Soc. 1973; 95:5725. [PubMed: 4733811] (b) Muller K, Faeh C, Diederich F. Science. 2007; 317:1881. [PubMed: 17901324]

(36). (a) Hudlicky M, Merola JS. Tet Lett. 1990; 31:7403.(b) Avent AG, Bowler AN, Doyle PM, Marchand CM, Young DW. Tet Lett. 1992; 33:1509.(c) Demange L, Ménez A, Dugave C. Tet Lett. 1998; 39:1169.(d) Chorghade MS, Mohapatra DK, Sahoo G, Gurjar MK, Mandlecha MV, Bhoite N, Moghe S, Raines RT. J Fluorine Chem. 2008; 129:781.(e) Doi M, Nishi Y, Kiritoshi N, Iwata T, Nago M, Nakano H, Uchiyama S, Nakazawa T, Wakamiya T, Kobayashi Y. Tetrahedron. 2002; 58:8453.

(37). (a) Mitsunobu O. Synthesis. 1981:1.b Hughes DL. Org Prep Proced Int. 1996; 28:127.(c) Swamy KCK, Kumar NNB, Balaraman E, Kumar K. Chem Rev. 2009; 109:2551. [PubMed: 19382806] (d) Wisniewski K, Koldziejczyk AS, Falkiewicz B. J Pept Sci. 1998; 4:1. [PubMed: 9523751]

(38). Vergnon AL, Pottorf RS, Player MR. J Comb Chem. 2004; 6:91. [PubMed: 14714991]

(39). Cadamuro SA, Reichold R, Kusebauch U, Musiol HJ, Renner C, Tavan P, Moroder L. Angew Chem Int Ed. 2008; 47:2143.

(40). (a) Shang S, Tan Z, Dong S, Danishefsky SJ. J Am Chem Soc. 2011; 133:10784. [PubMed: 21696220] (b) Townsend SD, Tan ZP, Dong SW, Shang SY, Brailsford JA, Danishefsky SJ. J Am Chem Soc. 2012; 134:3912. [PubMed: 22332757]

(41). (a) Verbiscar AJ, Witkop B. J Org Chem. 1970; 35:1924. [PubMed: 5446984] (b) Eswarakrishnan V, Field L. J Org Chem. 1981; 46:4182.(c) Kemp DS, Curran TP, Davis WM, Boyd JG, Muendel C. J Org Chem. 1991; 56:6672.(d) Reddie KG, Carroll KS. Curr Opin Chem Biol. 2008; 12:746. [PubMed: 18804173]

(42). (a) Dirksen A, Hackeng TM, Dawson PE. Angew Chem Int Ed. 2006; 45:7581.(b) Liu F, Stephen AG, Fisher RJ, Burke TR Jr. Bioorg Med Chem Lett. 2008; 18:1096. [PubMed: 18083557] (c) Liu F, Stephen AG, Waheed AA, Aman MJ, Freed EO, Fisher RJ, Burke TR Jr. ChemBioChem. 2008; 9:2000. [PubMed: 18655064]

(43). Shoulders MD, Guzei IA, Raines RT. Biopolymers. 2008; 89:443. [PubMed: 17937398]

(44). (a) Sebesta DP, Orourke SS, Pieken WA. J Org Chem. 1996; 61:361.(b) Jiang ZX, Yu YB. J Org Chem. 2007; 72:1464. [PubMed: 17243713]

(45). (a) Bisel P, Al-Momani L, Muller M. Org Biomol Chem. 2008; 6:2655. [PubMed: 18633519] (b) Lyu PC, Sherman JC, Chen A, Kallenbach NR. Proc Natl Acad Sci USA. 1991; 88:5317. [PubMed: 2052608] (c) Cornish VW, Kaplan MI, Veenstra DL, Kollman PA, Schultz PG. Biochemistry. 1994; 33:12022. [PubMed: 7918421] (d) Sigman MS, Jacobsen EN. J Am Chem Soc. 1998; 120:4901.(e) Guerin DJ, Miller SJ. J Am Chem Soc. 2002; 124:2134. [PubMed: 11878965] (f) Formaggio F, Baldini C, Moretto V, Crisma M, Kaptein B, Broxterman QB, Toniolo C. Chem Eur J. 2005; 11:2395. [PubMed: 15669065] (g) Bielska AA, Zondlo NJ. Biochemistry. 2006; 45:5527. [PubMed: 16634634] (h) Brown AM, Zondlo NJ. Biochemistry. 2012; 51:5041. [PubMed: 22667692]

(46). (a) Dalvit C. Prog Nucl Magn Reson Spectrosc. 2007; 51:243.(b) Papeo G, Giordano P, Brasca MG, Buzzo F, Caronni D, Ciprandi F, Mongelli N, Veronesi M, Vulpetti A, Dalvit C. J Am Chem Soc. 2007; 129:5665. [PubMed: 17417847]

(47). Jenkins CL, McCloskey AI, Guzei IA, Eberhardt ES, Raines RT. Biopolymers. 2005; 80:1. [PubMed: 15558658]

(48). (a) Yong YF, Kowalski JA, Lipton MA. J Org Chem. 1997; 62:1540.(b) Peterlin-Masic L, Kikelj D. Tetrahedron. 2001; 57:7073.(c) Balakrishnan S, Scheuermann MJ, Zondlo NJ. ChemBioChem. 2012; 13:259. [PubMed: 22213184] (d) Balakrishnan S, Zhao C, Zondlo NJ. J Org Chem. 2007; 72:9834. [PubMed: 17979289]

(49). Lee J, Chmielewski J. Chem Biol Drug Des. 2010; 75:161. [PubMed: 20028399]

(50). (a) Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Science. 1994; 266:776. [PubMed: 7973629] (b) Ruoslahti E. Ann Rev Cell Devel Biol. 1996; 12:697. [PubMed: 8970741]

(51). (a) Blackman ML, Royzen M, Fox JM. J Am Chem Soc. 2008; 130:13518. [PubMed: 18798613] (b) Selvaraj R, Liu SL, Hassink M, Huang CW, Yap LP, Park R, Fox JM, Li ZB, Conti PS. Bioorg Med Chem Lett. 2011; 21:5011. [PubMed: 21601452] (c) Taylor MT, Blackman ML, Dmitrenko O, Fox JM. J Am Chem Soc. 2011; 133:9646. [PubMed: 21599005] (d) Seitchik JL, Peeler JC, Taylor MT, Blackman ML, Rhoads TW, Cooley RB, Refakis C, Fox JM, Mehl RA. J Am Chem Soc. 2012; 134:2898. [PubMed: 22283158] (e) Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG. J Am Chem Soc. 2003; 125:3192. [PubMed: 12630856] (f) Oh S, Jung JC, Avery MAZ. Naturforsch, B: J Chem Sci. 2008; 63b:210.(g) Kristensen TE, Hansen FK, Hansen T. Eur J Org Chem. 2009:387.

(52). Cornish VW, Hahn KM, Schultz PG. J Am Chem Soc. 1996; 118:8150.

(53). (a) Benedetti E, Diblasio B, Pavone V, Pedone C, Felix A, Goodman M. Biopolymers. 1981; 20:283.(b) Kang YK, Park HS. Biopolymers. 2009; 92:387. [PubMed: 19373924]

(54). Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. Science. 2002; 298:1912. [PubMed: 12471243]

(55). (a) Johnson LN, Barford D. Annu Rev Biophys Biomol Struct. 1993; 22:199. [PubMed: 8347989] (b) Szilak L, Moitra J, Krylov D, Vinson C. Nature Struct Biol. 1997; 4:112. [PubMed: 9033589] (c) Andrew CD, Warwicker J, Jones GR, Doig AJ. Biochemistry. 2002; 41:1897. [PubMed: 11827536] (d) Nagar B, Hantschel O, Young MA, Scheffzek K, Veach D, Bornmann V, Clarkson B, Superti-Furga G, Kuriyan J. Cell. 2003; 112:859. [PubMed: 12654251] (e) Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK.

Nucleic Acids Res. 2004; 32:1037. [PubMed: 14960716] (f) Balakrishnan S, Zondlo NJ. J Am Chem Soc. 2006; 128:5590. [PubMed: 16637598]

(56). Zhang L, Chen X, Xue P, Sun HHY, Williams ID, Sharpless KB, Fokin VV, Jia G. J Am Chem Soc. 2005; 127:15998. [PubMed: 16287266]

(57). (a) Sletten EM, Bertozzi CR. Angew Chem Int Ed. 2009; 48:6974.(b) Lim RKV, Lin Q. Chem Commun. 2010; 46:1589.

(58). (a) Chalker JM, Wood CSC, Davis BG. J Am Chem Soc. 2009; 131:16346. [PubMed: 19852502] (b) Spicer CD, Davis BG. Chem Commun. 2011; 47:1698.(c) Spicer CD, Triemer T, Davis BG. J Am Chem Soc. 2012; 134:800. [PubMed: 22175226] (d) Li N, Lim RKV, Edwardraja S, Lin Q. J Am Chem Soc. 2011; 133:15316. [PubMed: 21899368]

(59). (a) Gopi HN, Tirupula KC, Baxter S, Ajith S, Chaiken IM. ChemMedChem. 2006; 1:54. [PubMed: 16892335] (b) Lee KK, Park KH, Joo C, Kwon HJ, Jeon J, Jung HI, Park S, Han H, Cho M. J Phys Chem B. 2012; 116:5097. [PubMed: 22512849] (c) Bonger KM, Kapoerchan VV, Grotenbreg GM, van Koppen CJ, Timmers CM, van der Marel GA, Overkleeft HS. Org Biomol Chem. 2010; 8:1881. [PubMed: 20449493]

(60). (a) Lin YA, Chalker JM, Floyd N, Bernardes GJL, Davis BG. J Am Chem Soc. 2008; 130:9642. [PubMed: 18593118] (b) Lin YYA, Chalker JM, Davis BG. ChemBioChem. 2009; 10:959. [PubMed: 19343741] (c) Lin YA, Chalker JM, Davis BG. J Am Chem Soc. 2010; 132:16805. [PubMed: 21050005]

(61). Dirksen A, Dawson PE. Curr Opin Chem Biol. 2008; 12:760. [PubMed: 19058994]

(62). (a) Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, Fox JM, Chin JW. J Am Chem Soc. 2012; 134:10317. [PubMed: 22694658] (b) Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. J Am Chem Soc. 2012; 134:792. [PubMed: 22176354] (c) Lang K, Davis L, Torres-Kolbus J, Chou CJ, Deiters A, Chin JW. NatChem. 2012; 4:298.

(63). (a) Poethko T, Schottelius M, Thumshirn G, Hersel U, Herz M, Henriksen G, Kessler H, Schwaiger M, Wester HJ. J Nucl Med. 2004; 45:892. [PubMed: 15136641] (b) Flavell RR, Kothari P, Bar-Dagan M, Synan M, Vallabhajosula S, Friedman JM, Muir TW, Ceccarini G. J Am Chem Soc. 2008; 130:9106. [PubMed: 18570424] (c) Åberg O, Pisaneschi F, Smith G, Nguyen QD, Stevens E, Aboagye EO. J Fluorine Chem. 2012; 135:200.

(64). (a) Ding H, Shigenaga A, Sato K, Morishita K, Otaka A. Org Lett. 2011; 13:5588. [PubMed: 21916452] (b) Wang YS, Fang XQ, Wallace AL, Wu B, Liu WSR. J Am Chem Soc. 2012; 134:2950. [PubMed: 22289053]

(65). (a) MacArthur MW, Thornton JM. J Mol Biol. 1991; 218:397. [PubMed: 2010917] (b) Kay BK, Williamson MP, Sudol P. FASEB J. 2000; 14:231. [PubMed: 10657980]

(66). Hodges JA, Raines RT. J Am Chem Soc. 2005; 127:15923. [PubMed: 16277536]

(67). (a) Baker GL, Fritschel SJ, Stille JR, Stille JK. J Org Chem. 1981; 46:2954.(b) See the Supporting Information for details.

(68). (a) Ciarkowski J, Gdaniec M, Kolodziejczyk A, Liberek B, Borremans FAM, Anteunis MJO. Int J Pept Prot Res. 1990; 36:285.(b) Bean JW, Kopple KD, Peishoff CE. J Am Chem Soc. 1992; 114:5328.(c) Stanger HE, Gellman SH. J Am Chem Soc. 1998; 120:4236.(d) Favre M, Moehle K, Jiang LY, Pfeiffer B, Robinson JA. J Am Chem Soc. 1999; 121:2679.(e) Schneider JP, Pochan DJ, Ozbas B, Rajagopal K, Pakstis L, Kretsinger J. J Am Chem Soc. 2002; 124:15030. [PubMed: 12475347]

(69). Vuister GW, Bax A. J Am Chem Soc. 1993; 115:7772.

(70). Kuemin M, Nagel YA, Schweizer S, Monnard FW, Ochsenfeld C, Wennemers H. Angew Chem Int Ed. 2010; 49:6324.

(71). (a) Getahun Z, Huang CY, Wang T, De Leon B, DeGrado WF, Gai F. J Am Chem Soc. 2003; 125:405. [PubMed: 12517152] (b) Aprilakis KN, Taskent H, Raleighj DP. Biochemistry. 2007; 46:12308. [PubMed: 17924662] (c) Schultz KC, Supekova L, Ryu Y, Xie J, Perera R, Schultz PG. J Am Chem Soc. 2006; 128:13984. [PubMed: 17061854] (d) Taskent-Sezgin H, Chung J, Patsalo V, Miyake-Stoner SJ, Miller AM, Brewer SH, Mehl RA, Green DF, Raleigh DP, Carrico I. Biochemistry. 2009; 48:9040. [PubMed: 19658436] (e) Zimmermann J, Thielges MC, Seo YJ, Dawson PE, Romesberg FE. Angew Chem Int Ed. 2011; 50:8333.

(72). (a) Serrano AL, Troxler T, Tucker MJ, Gai F. Chem Phys Lett. 2010; 487:303. [PubMed: 20419080] (b) Taskent-Sezgin H, Marek P, Thomas R, Goldberg D, Chung J, Carrico I, Raleigh DP. Biochemistry. 2010; 49:6290. [PubMed: 20565125]

(73). (a) Ruiz-Cabello J, Barnett BP, Bottomley PA, Bulte JWM. NMR in Biomedicine. 2011; 24:114. [PubMed: 20842758] (b) Mykhailiuk PK, Afonin S, Palamarchuk GV, Shishkin OV, Ulrich AS, Komarov IV. Angew Chem Int Ed. 2008; 47:5765.(c) Wadhwani P, Burck J, Strandberg E, Mink C, Afonin S, Ulrich AS. J Am Chem Soc. 2008; 130:16515. [PubMed: 19554688] (d) Li C, Wang GF, Wang Y, Creager-Allen R, Lutz EA, Scronce H, Slade KM, Ruf RAS, Mehl RA, Pielak GJ. J Am Chem Soc. 2010; 132:321. [PubMed: 20050707]

(74). Thomas CA, Talatya ER, Bann JG. Chem Comm. 2009:3366. [PubMed: 19503872]

(75). (a) Hawkes WC, Alkan Z. Biol Trace Elem Res. 2010; 134:235. [PubMed: 20306235] (b) Lobanov AV, Hatfield DL, Gladyshev VN. Biochim Biophys Acta, Gen Subj. 2009; 1790:1424. (c) Papp LV, Lu J, Holmgren A, Khanna KK. Antioxid Redox Signaling. 2007; 9:775.

(76). (a) Duddeck H. Se-77 Nmr Spectroscopy and Its Applications in Chemistry. Advances in Solid State Nmr Studies of Materials and Polymers: A Special Volume Dedicated to Isao Ando. 2004; 52:105.(b) Mobli M, de Araujo AD, Lambert LK, Pierens GK, Windley MJ, Nicholson GM, Alewood PF, King GE. Angew Chem, Int Ed. 2009; 48:9312.(c) Duddeck H. Prog Nucl Magn Reson Spectrosc. 1995; 27:1.

(77). (a) Wang Q, Sasaki NA, Potier P. Tetrahedron. 1998; 54:15759.(b) Ganorkar R, Natarajan A, Mamai A, Madalengoitia JS. J Org Chem. 2006; 71:5004. [PubMed: 16776535]

(78). (a) Kee S, Jois SDS. Current Pharm Design. 2003; 9:1209.(b) Tyndall JDA, Pfeiffer B, Abbenante G, Fairlie DP. Chem Rev. 2005; 105:793. [PubMed: 15755077]

(79). (a) Jeannotte G, Lubell WD. J Am Chem Soc. 2004; 126:14334. [PubMed: 15521735] (b) Geisler I, Chmielewski J. Bioorg Med Chem Lett. 2007; 17:2765. [PubMed: 17363245] (c) Farrera-Sinfreu J, Zaccaro L, Vidal D, Salvatella X, Giralt E, Pons M, Albericio F, Royo M. J Am Chem Soc. 2004; 126:6048. [PubMed: 15137770] (d) Hanessian S, McNaughtonSmith G, Lombart HG, Lubell WD. Tetrahedron. 1997; 53:12789.(e) Melendez RE, Lubell WD. J Am Chem Soc. 2004; 126:6759. [PubMed: 15161304] (f) Del Valle JR, Goodman M. J Org Chem. 2003; 68:3923. [PubMed: 12737573] (g) Wittelsberger A, Keller M, Scarpellino L, Patiny L, Acha-Orbea H, Mutter M. Angew Chem Int Ed. 2000; 39:1111.(h) Guitot K, Larregola M, Pradhan TK, Vasse JL, Lavielle S, Bertus P, Szymoniak J, Lequin O, Karoyan P. ChemBioChem. 2011; 12:1039. [PubMed: 21448869]

Abbreviations

ACE	angiotensin converting enzyme
ADDP	azodicarboxylic acid dipiperidide
Alloc	allyloxycarbonyl
AllocCl	allylchloroformate
α_R	right-handed α-helix conformation
Azp	4*R*-azidoproline
azp	4*S*-azidoproline
Boc	*tert*-butoxycarbonyl
15-C-5	15-crown-5
DAST	diethylaminosulfur trifluoride
DIAD	diisopropylazodicarboxylate

DIC	N,N-diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMAP	4-dimethylamino pyridine
EWG	electron-withdrawing group
Flp	4*R*–fluoroproline
flp	4*S*–fluoroproline
Fmoc	9-fluorenylmethoxycarbonyl
HBTU	O-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
Hyp	4*R*–hydroxyproline
hyp	4*S*–hydroxyproline
Ms	methylsulfonyl
Mst	2,4,6,-trimethylbenzenesulfonyl
NHS	N-hydroxysuccinimide
Ns	4-nitrobenzenesulfonyl
Nu	nucleophile
OBzCF$_3$	4-trifluoromethyl benzoate ester
OBzNO$_2$	4-nitrobenzoate ester
OMst	2,4,6–trimethyl phenylsulfonate
OPhBr	4-bromophenyl ether
OPhCN	4-cyanophenyl ether
OPhI	4-iodophenyl ether
OMs	methanesulfonate
ONs	nitrobenzenesulfonate
OTs	toluenesulfonate
PDC	pyridinium dichromate
Ph	phenyl
pivaloyl	trimethyl acetyl
PPII	type II polyproline helix
Prox	4-substituted proline
SAc	thioacetyl
SPhNO$_2$	4-nitrophenyl thioether
TBAF	tetra-n-butylammonium fluoride

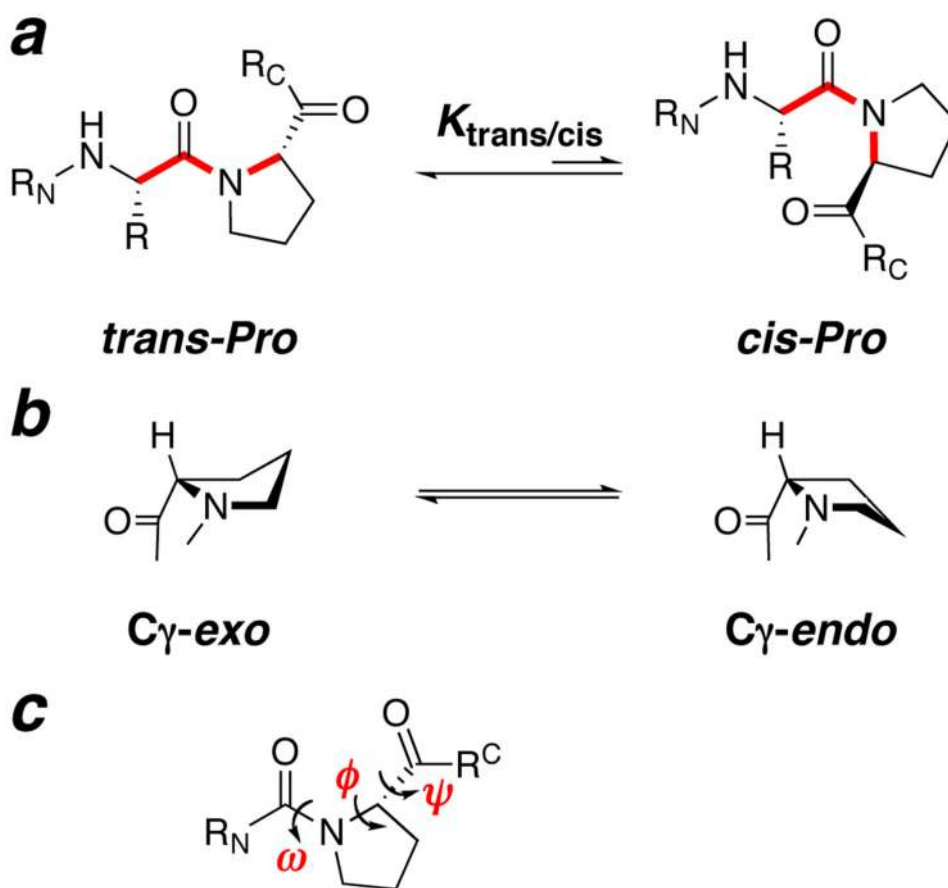|	
TBS	*tert*-butyldimethylsilyl
TBSCl	*tert*-butyldimethylsilyl chloride
TES	triethylsilane
TFA	trifluoroacetic acid
TIS	triisopropylsilane
Trt	triphenylmethyl (trityl)
Ts	4-methylbenzenesulfonyl

Figure 1.
Proline conformational equilibria. (a) Cis-trans isomerism of the prolyl amide bond. In proteins, 95% of prolyl amide bonds are trans, 5% are cis. The trans and cis amide bond are in slow exchange on the NMR timescale, so the relative populations ($K_{trans/cis}$ = [peptide with trans amide bond]/[peptide with cis amide bond]) are determined via quantification of their distinct NMR resonances. (b) Exo and endo ring puckers of proline. Exo and endo ring puckers are in rapid equilibrium. (c) Definition of main chain torsion angles in proline.
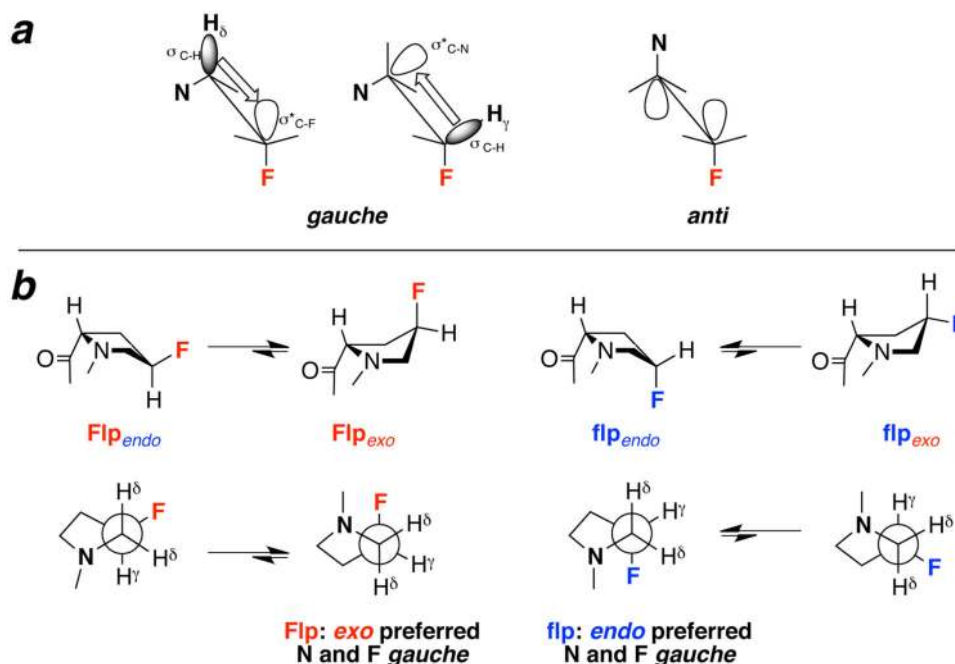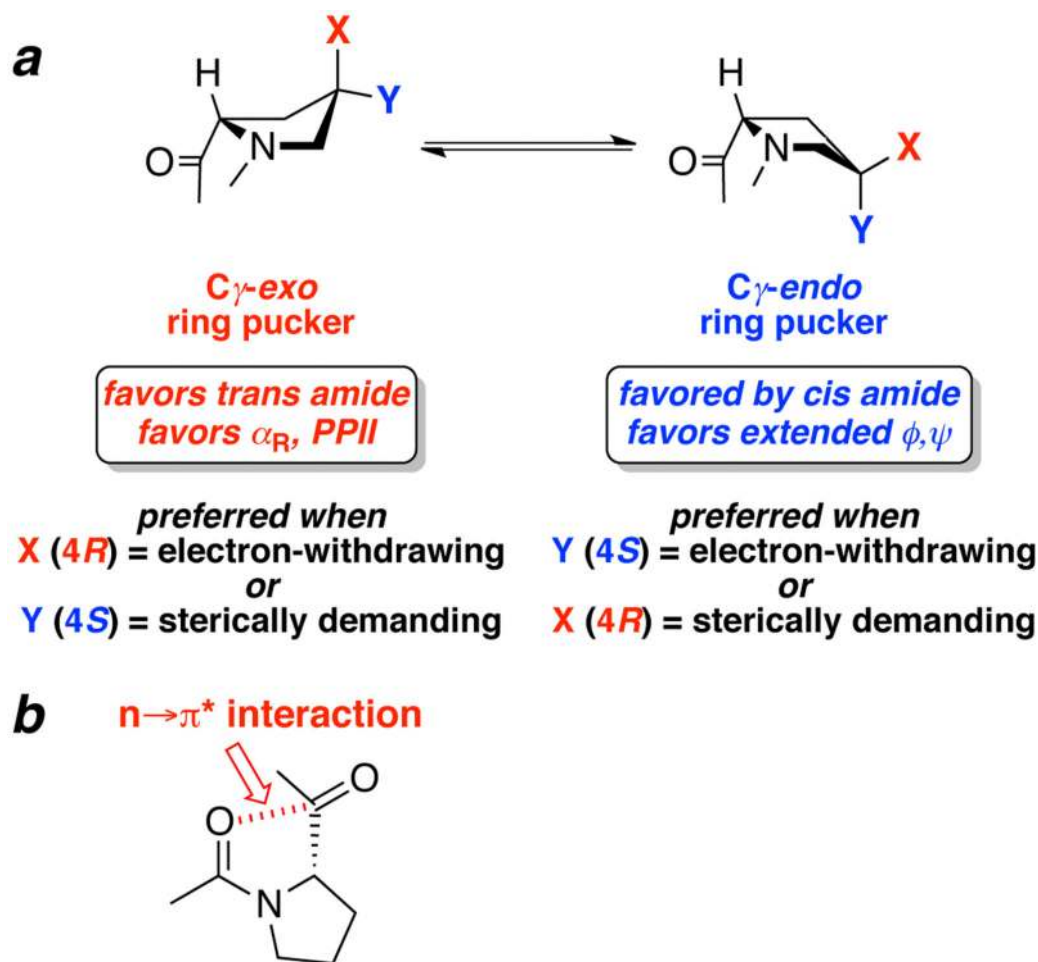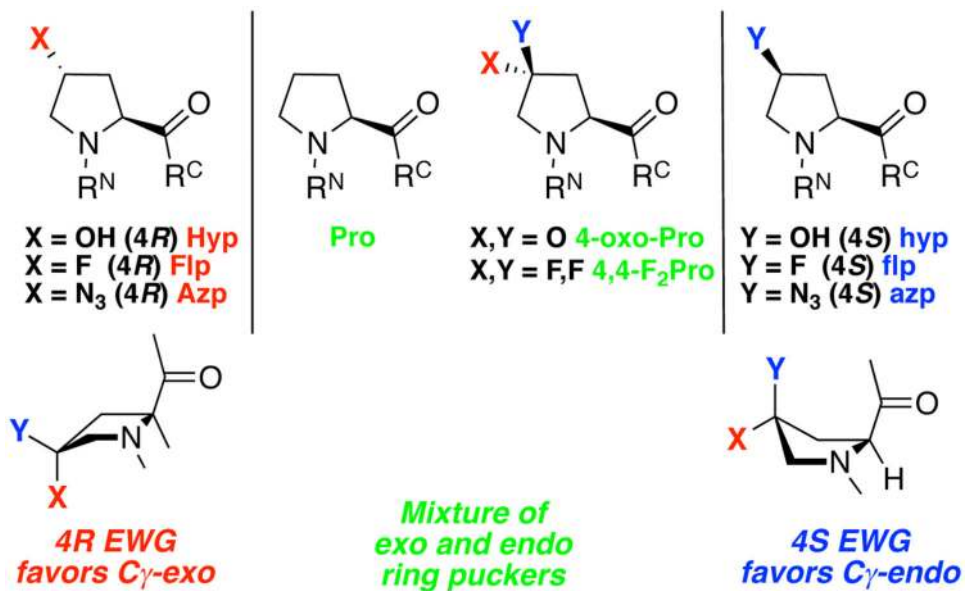
Figure 2.
Stereoelectronic effects in 4-substituted prolines lead to a preference in the Cγ-CS bond for a gauche relationship between the amide and electron-withdrawing 4-substituents. (a) In 4R-fluoroproline (Flp) and 4S-fluoroproline (flp), two major hyperconjugative interactions stabilize the *gauche* conformation: orbital overlap between the electron-rich C-H$_\delta$ bond (σ) and the electron-deficient C-F bond (σ*) (left), and orbital overlap between the electron-rich C-H, bond and the electron-deficient C-N bond (middle). When N and F are *anti*, neither stabilizing interaction is possible (right). Additional hyperconjugative stabilization is provided by an anti arrangement between a beta proton and the fluorine when Flp is exo or when flp is endo. (b) Manifestation of gauche effects on ring pucker. With electron-withdrawing substituents, a gauche conformation with 4R-substitution leads to a strong preference for exo ring pucker, while a gauche conformation with 4S-substitution leads to a strong preference for endo ring pucker. The exo/endo preference thus depends on both the stereochemistry of the substitution and how electron-withdrawing the substituent is. For sterically demanding and/or non-electron-withdrawing substituents, the reverse preferences will be observed due to a steric preference for anti over gauche conformation. Newman projections are simplified to show idealized torsion angles.

Figure 3.
Effects of proline ring pucker on main chain ϕ, ψ, and ω torsion angles. (a) 4-Substituted proline residues can bias proline ring pucker by stereoelectronic effects or by steric effects. (b) An n→π* interaction between the carbonyl lone pair (n) of residue i and the π* orbital of the carbonyl on the subsequent residue (i+1) provides local organization of the protein main chain, including stabilization of the trans amide bond and the α_R and PPII conformations. This interaction is preferential with the *Cγ-exo* ring pucker of proline and is only possible with a trans amide bond.

Figure 4.
Most-studied 4-substituted proline derivatives, three letter codes, and their conformational preferences. Red color and upper case 3-letter code indicates trans relative stereochemistry (4R substitution) of the 4-substituent and the carbonyl. Blue color and lower case 3-letter code indicates cis relative stereochemistry (4S) of the 4-substitutent and the carbonyl. Green color indicates 4,4-disubstituted prolines. EWG = electron-withdrawing group. Non-electron-withdrawing or sterically demanding substituents have opposite conformational preferences: 4R-substituted methylproline (X = CH$_3$) and 4R-mercaptoproline (X = SH) favor Cγ-endo ring pucker, while 4S-substituted-methylproline and 4S-mercaptoproline favor Cγ-exoring pucker.[7a, 11, 39]
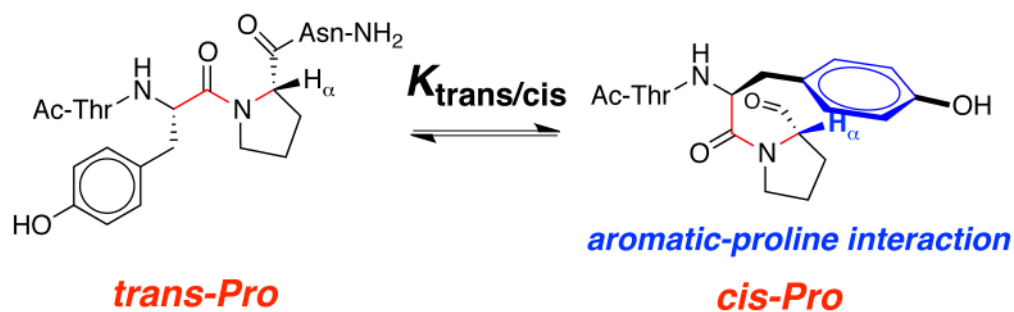
Figure 5.
Aromatic-proline CH/π interactions between H_α (partially positive due to adjacent electron-withdrawing groups) and the negatively charged aromatic π face leads to an increased preference for cis amide bond in aromatic-proline sequences such as the model peptide sequence Ac-TYPN-NH$_2$.
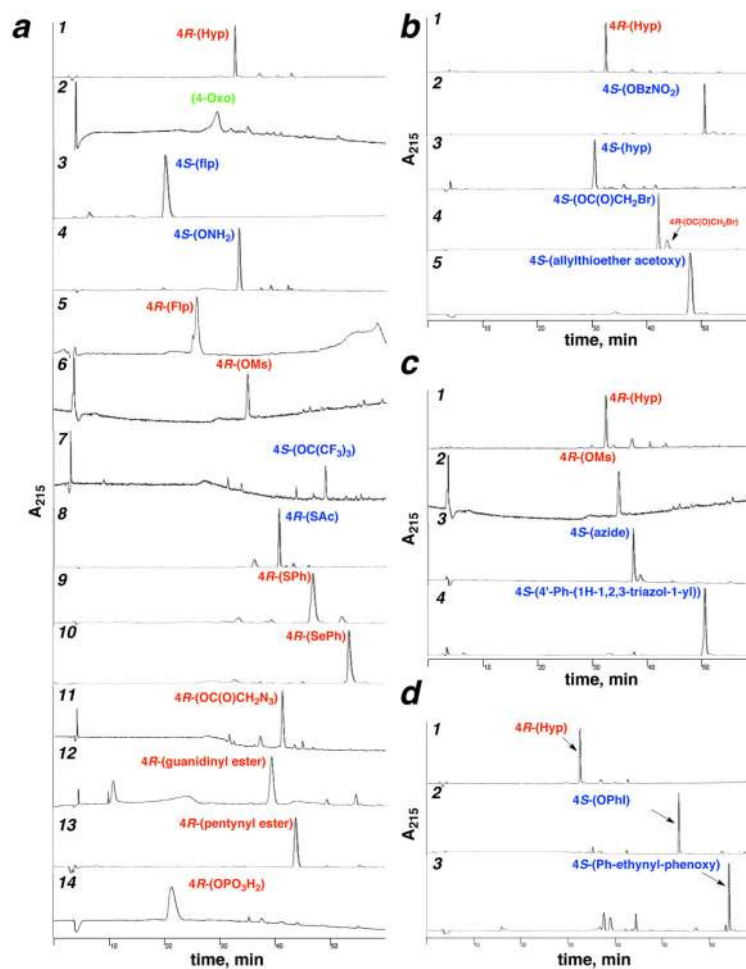
Figure 6.
Representative crude HPLC chromatograms of proline editing reactions on solid phase, including inversion of hydroxyproline stereochemistry, sulfonylation, S_N2, Mitsunobu, acylation, oxidation, and solid phase organometallic reactions. HPLC chromatograms for the synthesis of all peptides are in the Supporting Information (Part 2). (a) Representative reaction products from Hyp (**1**), including oxidation (entry 2 (**91**)), DAST reaction (entries 3 (**26**) and 5 (**27**) (via hyp)), Mitsunobu reaction (entries 4 (**52**), 7 (**56**)), sulfonation (entry 6 (**5**)), 3-step sequences via Mitsunobu reaction on hyp (entries 8 (**43**), 9 (**16**), 10 (**41**)), acylation plus S_N2 (entry 11 (**83**)), acylation (entries 12 (**64**), 13 (**79**)), and phosphorylation (entry 14 (**97**)). Entry 5 (Flp) includes flp as a minor impurity in the chromatogram due to incomplete Mitsunobu reaction to form the hyp starting material for the DAST reaction. (b) Synthesis of 4*S*-acetoxy allylthioether (**90**) from Hyp (**1**) via Mitsunobu inversion to the hyp nitrobenzoate (**2**), azidemediated deesterification to hyp (**4**), acetylation to the bromoacetate (**75**), and S_N2 reaction with allyl mercaptan. (c) Synthesis of the 4*S*-azide-alkyne cycloaddition product (**109**) from Hyp (**1**) via mesylation (**5**), S_N2 reaction with sodium azide (**21**), and copper-mediated Huisgen cycloaddition. (d) Synthesis of the 4*S*-Sonogashira product (**107**) from Hyp (**1**) via Mitsunobu reaction with 4-iodophenol (**37**) and Pd-mediated cross-coupling with phenylacetylene.
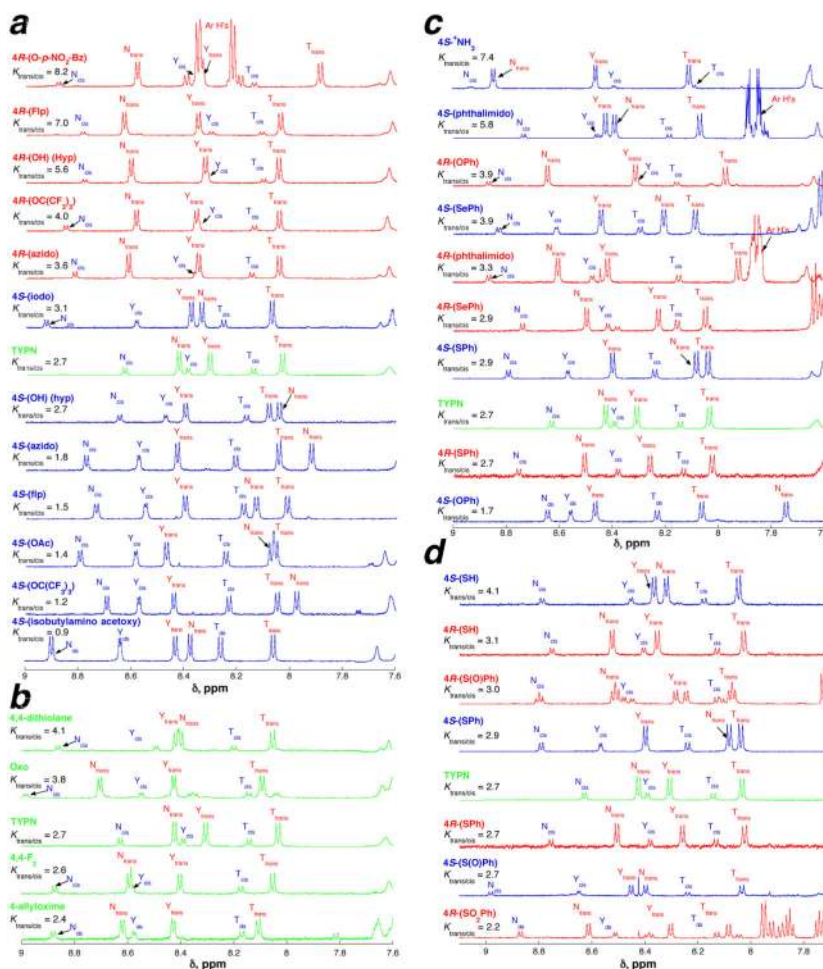
Figure 7.
Representative ^{1}H NMR spectra (5 mM phosphate (pH 4 unless an ionizable proton is present), 25 mM NaCl, 90% H_2O/10% D_2O) of Ac-TYP(4-substituted)N-NH$_2$ peptides with 4-substituted prolines, with species with trans and cis amide bonds and $K_{trans/cis}$ indicated. The largest differences between 4R- versus 4S-substituted peptides are observed in $K_{trans/cis}$, the δ of Tyr$_{cis}$ and Asn$_{trans}$, and the $^3J_{\alpha N}$ of Tyr$_{cis}$. These effects are modulated as a function of the nature of the substituent with the 4R or 4S series. (a) Stereoelectronic effects of 4-substituents. (b) Disubstituted prolines. The NMR spectrum of 4-oxo-proline also includes the hydrate, which is in equilibrium with the ketone ($K_{hydrate/ketone}$ = 0.15).[67b] (c) Steric effects of 4-substituted prolines and the interplay of steric versus stereoelectronic effects in chalcogen-substituted prolines. (d) Steric effects versus stereoelectronic effects as a result of sulfur oxidation. NMR spectra for all peptides are in the Supporting Information (part 1: amide regions of the NMR spectra for all peptides; part 3: full NMR spectra for all peptides).
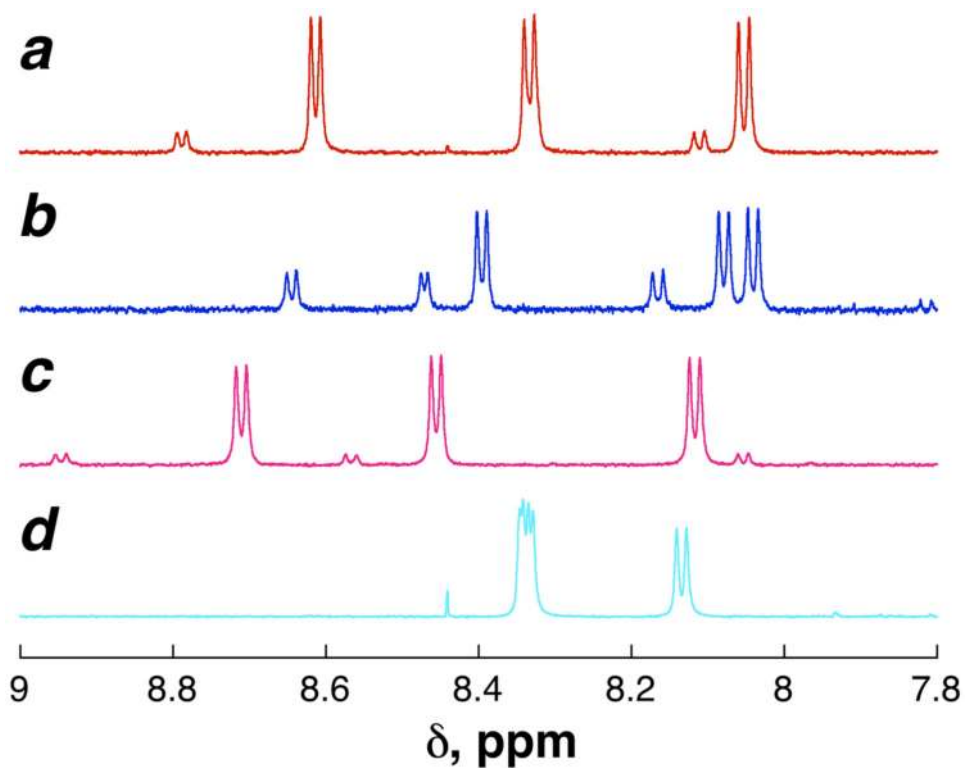
Figure 8.
^{1}H NMR spectra (amide region) of the four 4-hydroxyproline Ac-TY-4-hydroxyproline-N-NH₂ diastereomers. (a) 2*S*,4*R*-hydroxyproline (*trans*-hydroxyproline, Hyp) (**1**) (red). (b) 2*S*,4*S*-hydroxyproline (*cis*-hydroxyproline, hyp) (**4**) (blue); (c) 2*R*,4*S*-hydroxyproline (ᴅ-*trans*-hydroxyproline, ᴅ-hyp) (**125**) (magenta); (d) 2*R*,4*R*-hydroxyproline (ᴅ-*cis*-hydroxyproline, ᴅ-hyp) (**126**) (cyan). Full NMR spectra and TOCSY-derived amide resonance assignments are in the Supporting Information.
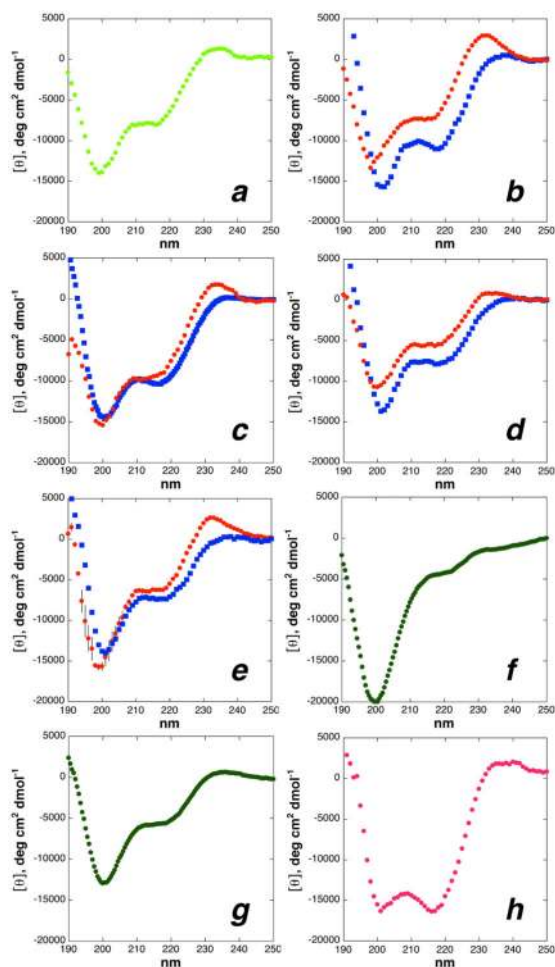
Figure 9.
CD spectra of TYXN peptides indicating the effect of proline substitutions on modulating peptide conformation. (a) Ac-TYPN-NH$_2$; (b) Ac-TYHypN-NH$_2$(**1**) (red circles) and Ac-TYhypN-NH$_2$(**4**) (blue squares); (c) Ac-TYFlpN-NH$_2$(**27**) (red circles) and Ac-TYflpN-NH$_2$(**26**) (blue squares); (d) Ac-TYP(4R-azido)N-NH$_2$(**18**) (red circles) and Ac-TYP(4S-azido)N-NH$_2$(**21**) (blue squares); (e) Ac-TYP(4R-OAc)N-NH$_2$(**58**) (red circles) and Ac-TYP(4S-OAc)N-NH$_2$(**59**) (blue squares); (f) Ac-TYP(4-oxo)N-NH$_2$(**91**); (g) Ac-TYP(4,4-F$_2$)N-NH$_2$(**101**) and (h) Ac-TYP(4,4-dithiolane)N-NH$_2$(**102**). CD data were collected in water with 25 mM KF and 20 mM phosphate pH 7.0. CD data were background corrected but were not smoothed. CD spectra with error bars are shown in the Supporting Information.
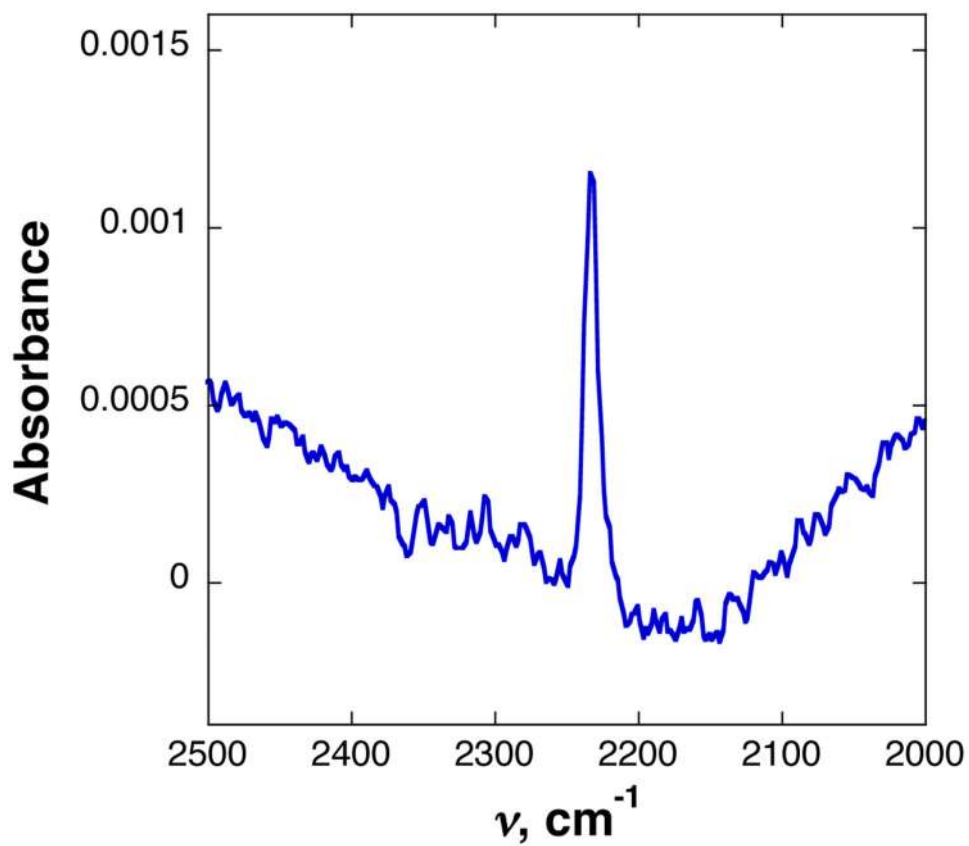
Figure 10.
FTIR spectrum of Ac-TAP(4S-OPhCN)N-NH$_2$(**145**) in H$_2$O (ν = 2233 cm^{-1}). The full FTIR spectrum is in the Supporting Information.
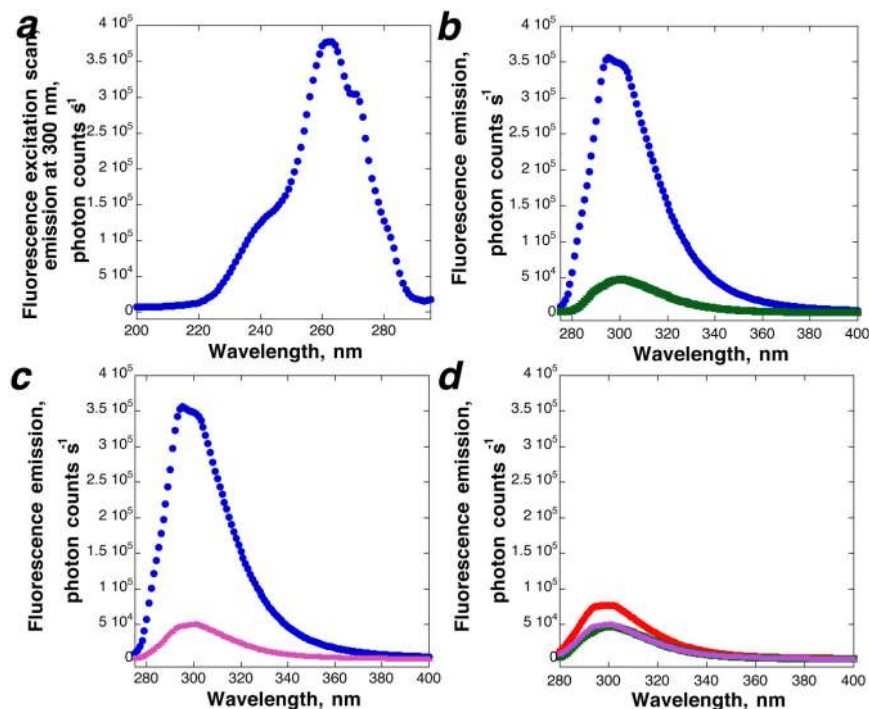
Figure 11.
Fluorescence of peptides in 5 mM HEPES buffer pH 7.5 with 100 mM NaCl and 2 mM MgCl$_2$. Data were collected with band widths of 3 nm. (a) Fluorescence excitation spectrum of Ac-TAP(4S-OPhCN)N-NH$_2$(**145**) with detection of emission at 300 nm. (b) Fluorescence emission spectra of Ac-TAP(4S-OPhCN)N-NH$_2$ (blue circles) and Ac-TYPN-NH$_2$ (green squares) (50 μM peptide) with excitation at 265 nm. (c) Fluorescence emission spectra of Ac-TAP(4S-OPhCN)N-NH$_2$ (blue circles) and Ac-TYP(4S-OPhCN)N-NH$_2$(**36**) (magenta triangles) with excitation at 265 nm, showing the effect of tyrosine in quenching fluorescence. (d) Fluorescence emission spectra of Ac-TYP(4R-OPhCN)N-NH$_2$(**35**) (red circles), Ac-TYP(4S-OPhCN)N-NH$_2$(**36**) (magenta triangles), and Ac-TYPN-NH$_2$ (green squares). The fluorescence emission spectra of these peptides, with excitation at 265 nm and at 280 nm, with error bars, are in the Supporting Information.
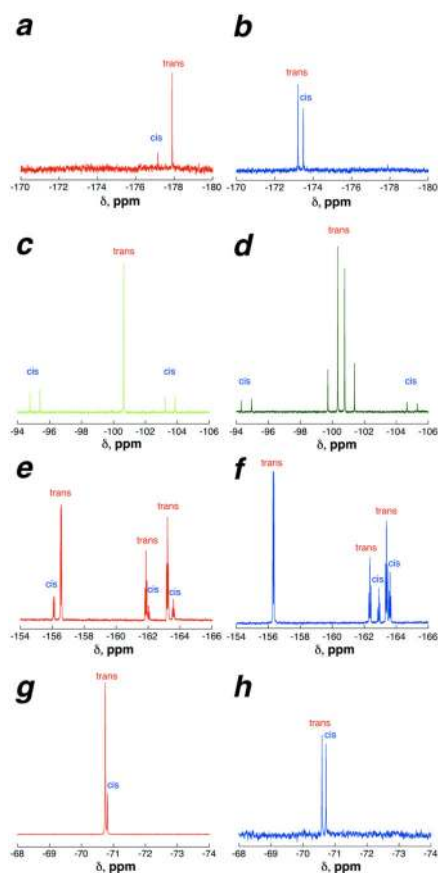
Figure 12.
^{19}F NMR spectra of peptides in 5 mM phosphate buffer pH 4 with 25 mM NaCl in 90% H_2O/10% D_2O at 23 °C. The data were collected with ^{1}H decoupling. (a) Ac-TYFlpN-NH_2(**27**); (b) Ac-TYflpN-NH_2(**26**); (c) Ac-TYP(4,4-F_2)N-NH_2(**101**); (d) Ac-TAP(4,4-F_2)N-NH_2(**134**); (e) Ac-TYP(4R-O-C_6F_5)N-NH_2(**33**); (f) Ac-TYP(4S-O-C_6F_5)N-NH_2(**34**); (g) Ac-TYP(4R-OC(CF$_3$)$_3$)N-NH_2(**55**); (h) Ac-TYP(4S-OC(CF$_3$)$_3$)N-NH_2(**56**).
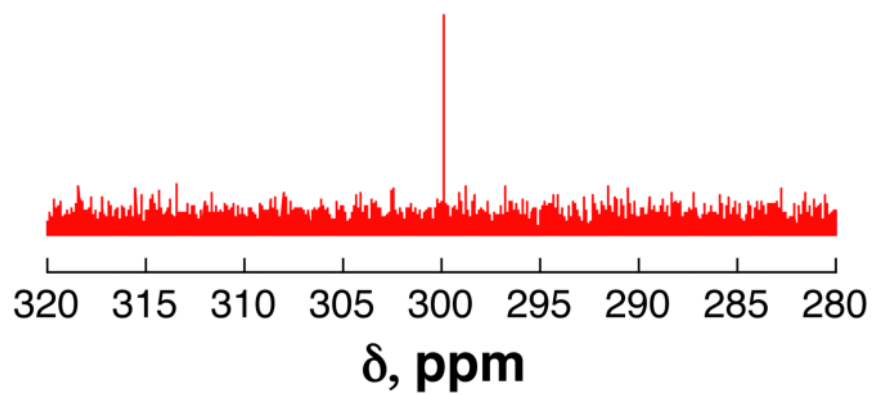
Figure 13.
^{77}Se NMR spectrum of Ac-TYP(4*R*-SePh)N-NH$_2$(**41**) (2 mM) (δ = 299.9 ppm) at 23 °C in D$_2$O. The data were collected with ^{1}H decoupling. The full spectrum is in the Supporting Information.
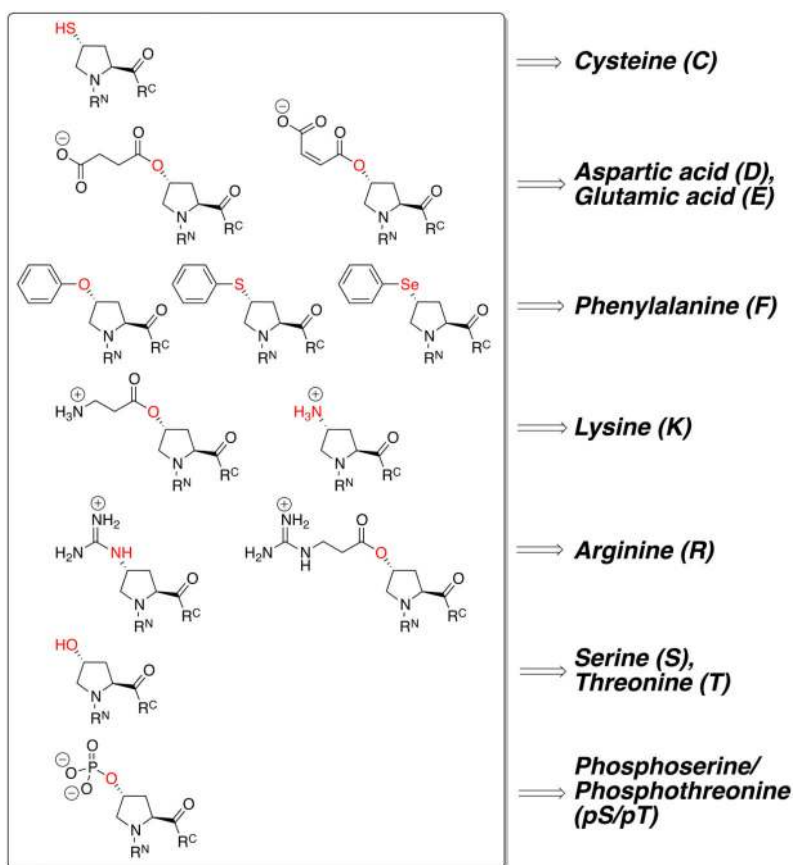
Figure 14.
Structured amino acid mimetics. Peptides with all functionalities herein were synthesized with both 4*R* (shown) and 4*S* stereochemistries. NMR data indicating the structural differences between 4*R* and 4*S* substitution are in Table 1 and the Supporting Information.
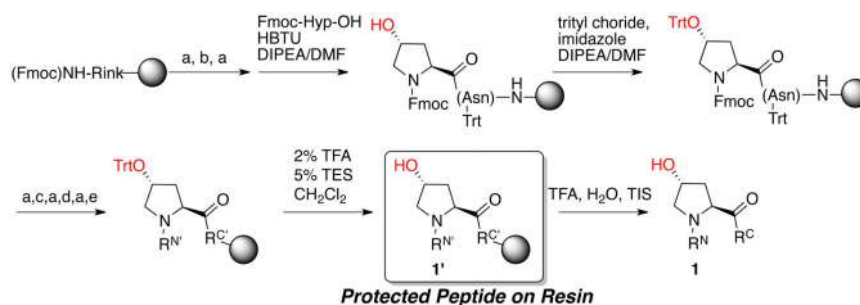
Scheme 1.
Proline editing general approach: automated synthesis of the peptide Ac-TYHypN-NH$_2$(**1**) via trityl hydroxyl protection.[a]

[a] (a) 20% piperidine/DMF; (b) Fmoc-Asn(Trt)-OH, HBTU, DIPEA/DMF; (c) Fmoc-Tyr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (d) Fmoc-Thr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (e) 10% Ac$_2$O/pyridine. R$^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, R$^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$. For all peptides in all schemes, the prime designation (e.g. **1′**, R$^{N\prime}$, R$^{C\prime}$) indicates the protected peptide on solid phase, whereas the absence of a prime designation indicates free peptide in solution.
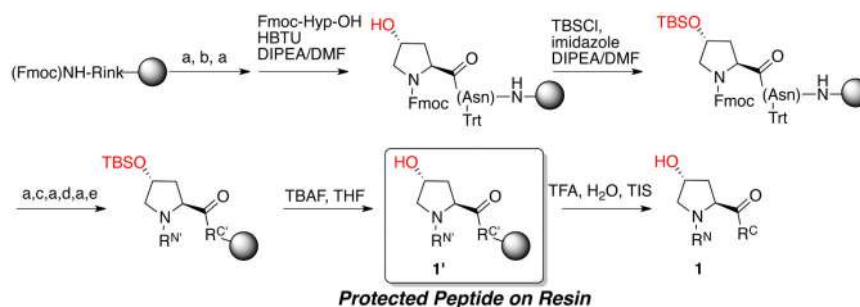
Scheme 2.

Proline editing alternative approach: automated synthesis of the peptide Ac-TYHypN-NH$_2$(**1**) via TBS hydroxyl protection.[a]

[a] (a) 20% piperidine/DMF; (b) Fmoc-Asn(Trt)-OH, HBTU, DIPEA/DMF; (c) Fmoc-Tyr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (d) Fmoc-Thr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (e) 10% Ac$_2$O/pyridine. R$^{N'}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, R$^{C'}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$.
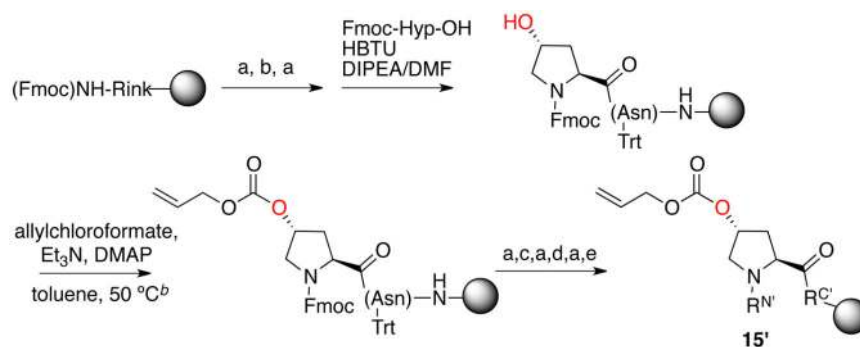
Scheme 3.

Proline editing alternative approach: synthesis of the peptide Ac-TYP(4*R*-O-Alloc)N-NH$_2$(**15**) via Alloc protection and direct acylation. The Alloc group was stable to TFA cleavage/deprotection. [a]

[a] (a) 20% piperidine/DMF; (b) Fmoc-Asn(Trt)-OH, HBTU, DIPEA/DMF; (c) Fmoc-Tyr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (d) Fmoc-Thr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (e) 10% Ac$_2$O/pyridine. $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$

[b] Reaction was performed in a glass vial manually.
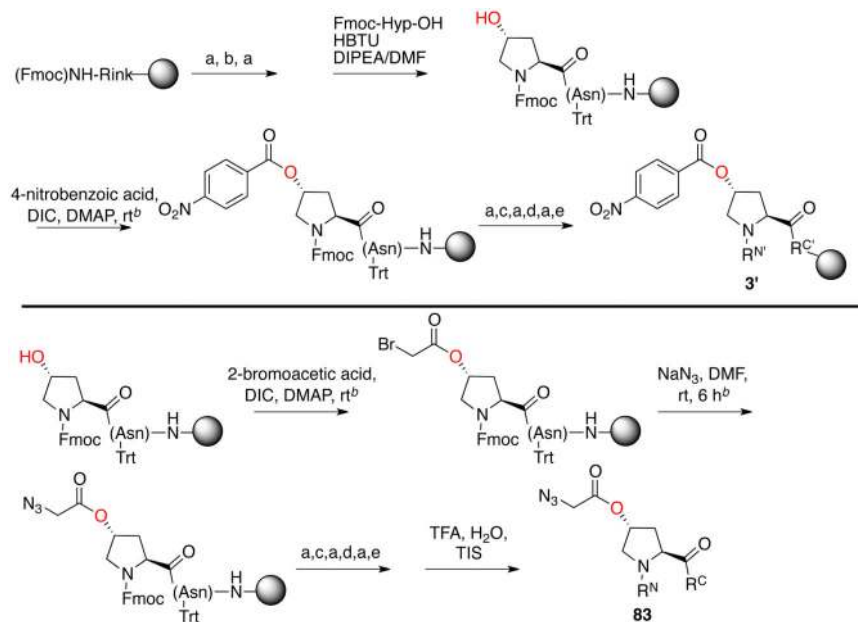
Scheme 4.

Proline editing alternative approach: direct modification immediately after coupling of Fmoc-Hyp-OH. Synthesis of the peptides Ac-TYP(4*R*-OBzNO$_2$)N-NH$_2$(**3**) (top) and Ac-TYP(4*R*-azidoacetate)N-NH$_2$(**83**) via direct acylation.[a]

[a] (a) 20% piperidine/DMF; (b) Fmoc-Asn(Trt)-OH, HBTU, DIPEA/DMF; (c) Fmoc-Tyr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (d) Fmoc-Thr(O*t*Bu)-OH, HBTU, DIPEA/DMF; (e) 10% Ac$_2$O/pyridine. R$^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, R$^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$

[b] Reaction was performed manually in a disposable fritted tube.
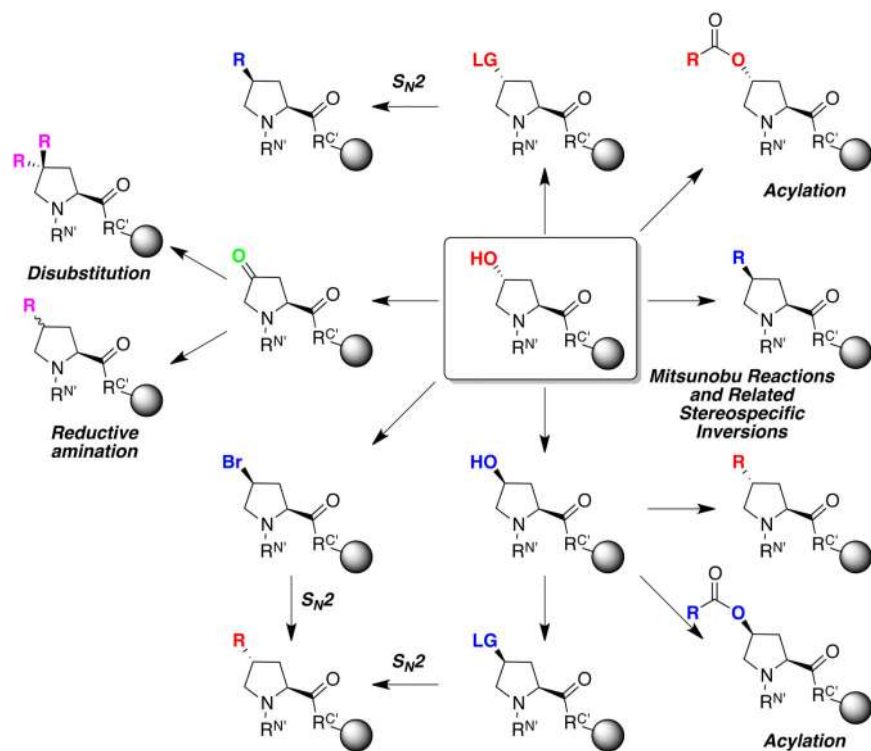
Scheme 5.
Proline editing to introduce functional groups with defined stereochemistry into peptides synthesized with commercially available and inexpensive Fmoc-Hyp-OH.[a]

[a] LG = leaving group. R$^{N'}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, R$^{C'}$ = -Asn(Trt)-NHRink-Resin.
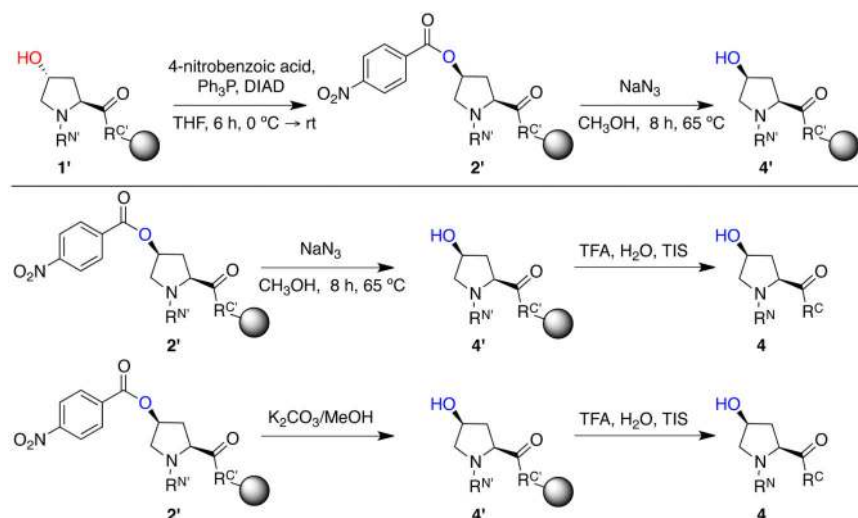
Scheme 6.

Mitsunobu inversion on solid phase to convert Hyp (4*R*) (**1′**) to hyp (4*S*) (**4′**). The nitrobenzoate may be deprotected selectively on solid phase with NaN_3/MeOH[30] or non-selectively with LiOH or K_2CO_3/MeOH.[a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH_2.
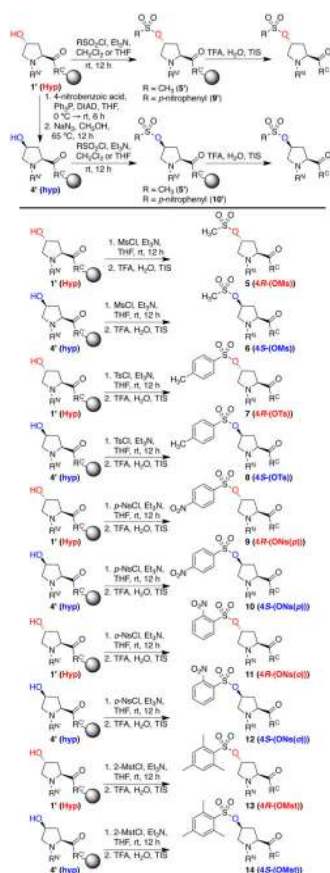
Scheme 7.
Sulfonylation reactions on solid phase. Sulfonates were subsequently modified on solid phase (Scheme 8) or were subjected to TFA cleavage/deprotection and purified and analyzed in aqueous solution.[a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$.
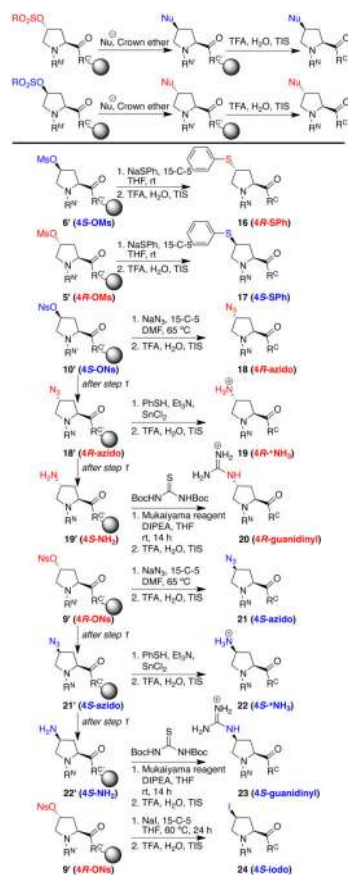
Scheme 8.
Solid phase S_N2 reactions on sulfonates and subsequent reactions on S_N2 products.[a]
[a] R^N = Ac-Thr(O tBu)-Tyr(O tBu)-, R^C = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn- NH_2.
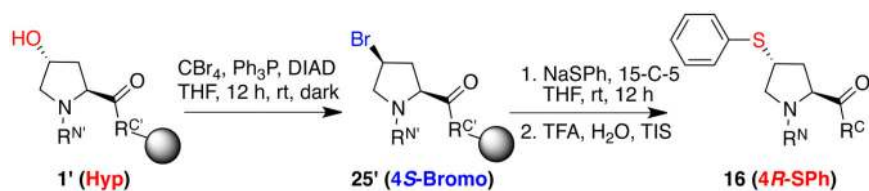
Scheme 9.

Alternative solid phase synthesis of 4*R*-substituted proline via an S_N2 reaction on 4*S*-bromoproline.[a]

[a] $R^{N'}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C'}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.
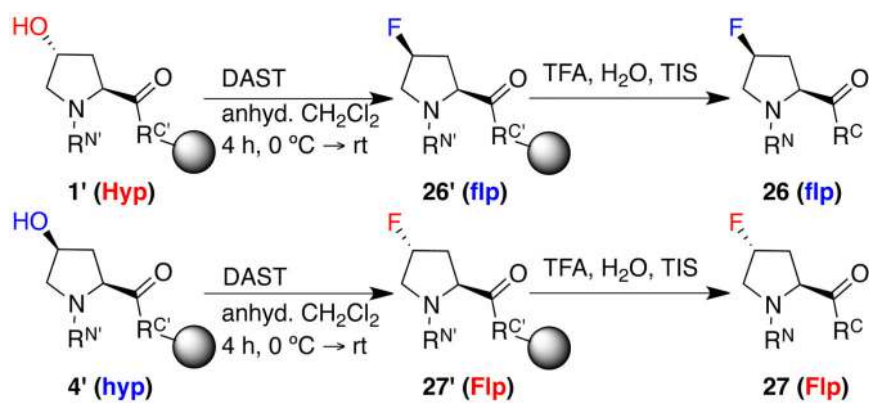
Scheme 10.
Solid phase fluorination of Hyp and hyp to generate flp and Flp. [a]
[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.

Scheme 11.
Mitsunobu reactions on solid phase.[a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.

Scheme 12.
Acylation reactions on solid phase. [a]

[a] $R^{N'}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C'}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.
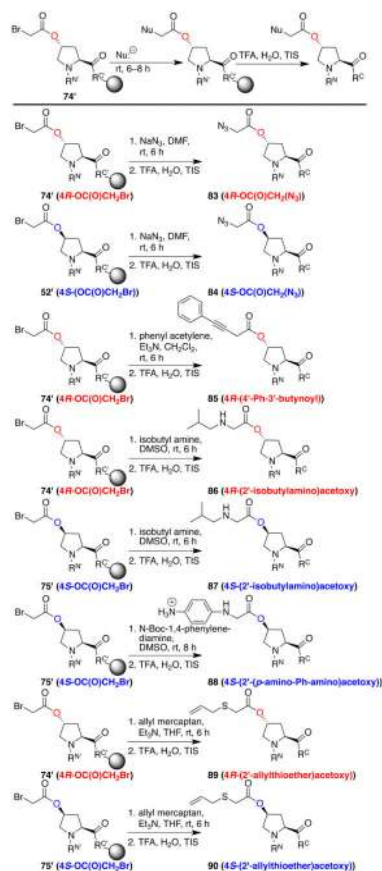
Scheme 13.
Substitution reactions on 4*R*- or 4*S*-hydroxyproline bromoacetate on solid phase. [a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$.
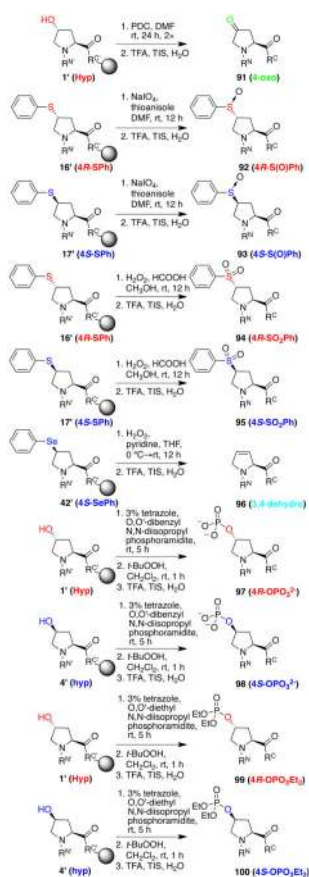
Scheme 14.
Oxidation reactions on 4-substituted prolines in TYXN peptides.[a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.
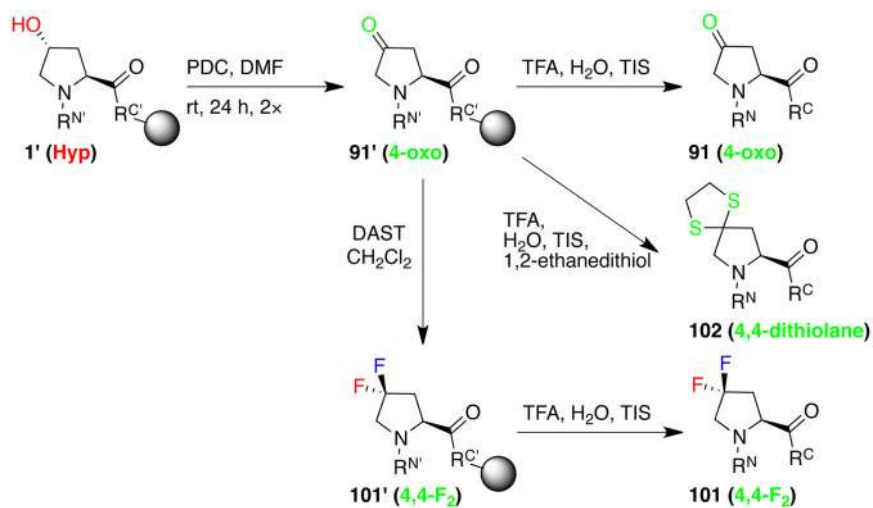
Scheme 15.
Synthesis of disubstituted prolines.[a]

[a] $R^{N'}$ = Ac-Thr(OtBu)-Tyr(OtBu)-, $R^{C'}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.
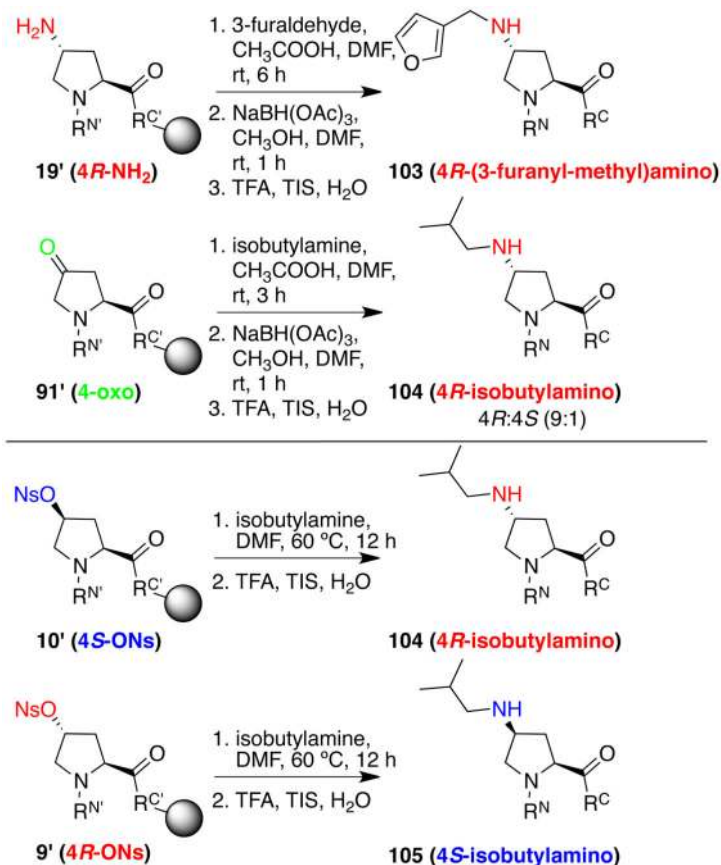
Scheme 16.
Reductive amination on 4*R*-aminoproline and 4-oxoproline. The stereochemistry of the oxoproline reductive amination products, suggested by ^{1}H NMR spectra to indicate 4*R* (major) versus 4*S* (minor) stereochemistry, was confirmed by S_N2 reactions of isobutyl amine on the respective nosylates. [a]

[a] $R^{N\prime}$ = Ac-Thr(O*t*Bu)-Tyr(O*t*Bu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = - Asn-NH$_2$.
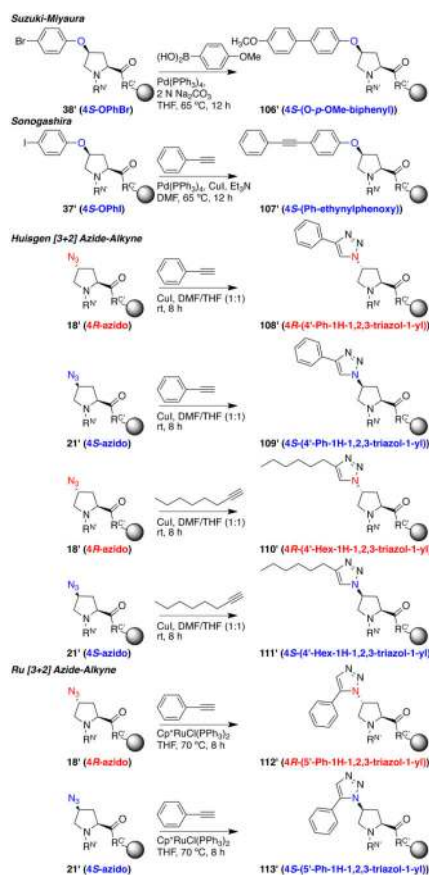
Scheme 17.
Organometallic reactions on solid phase. Peptides were subjected to TFA cleavage/deprotection, purified, and characterized. Chromatograms for the multistep synthesis of **107** are in Figure 6d. Chromatograms for the multistep synthesis of **109** are in Figure 6c[a]
[a] $R^{N'}$ = Ac-Thr(OtBu)-Tyr(OtBu)-, $R^{C'}$ = -Asn(Trt)-NHRink-Resin. R^N = Ac-Thr-Tyr-, R^C = -Asn-NH$_2$.
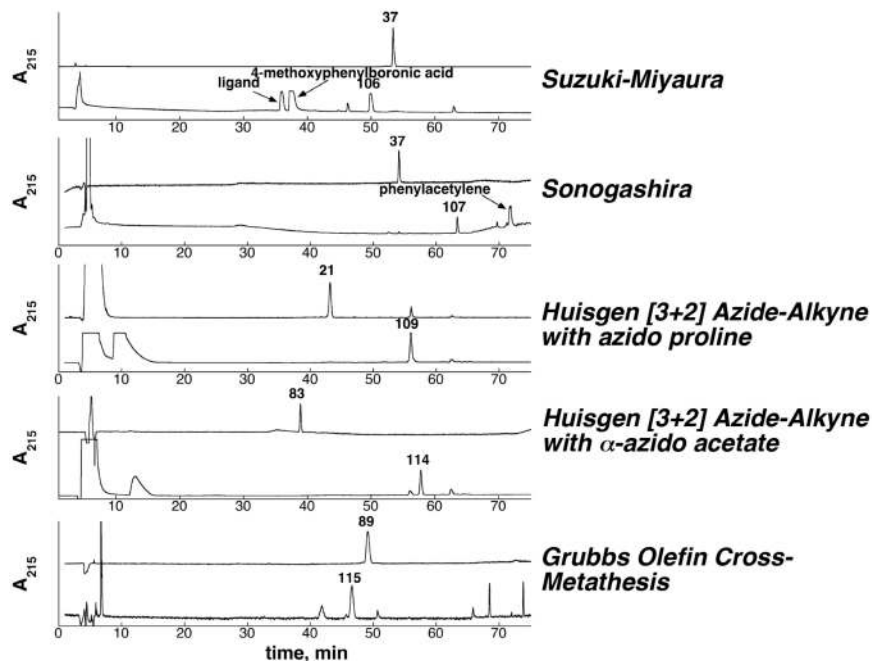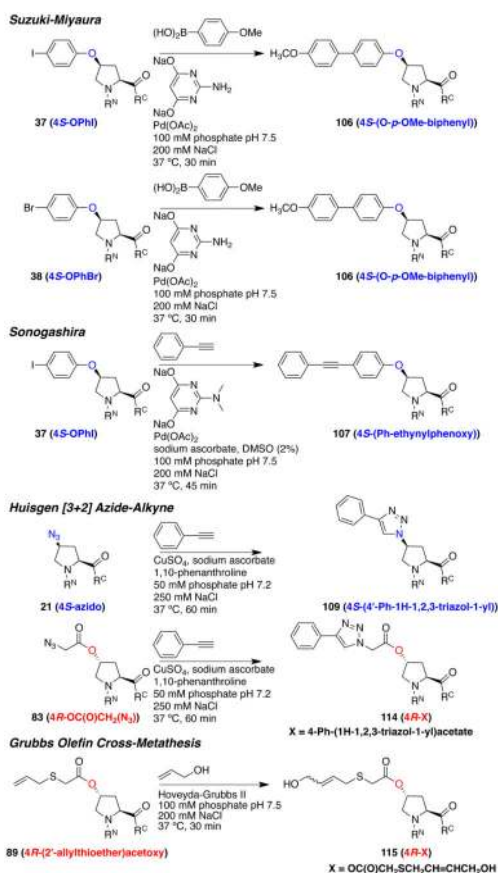
Scheme 18.

Organometallic reactions on peptides in aqueous solution. Bottom: HPLC chromatograms of reactions. For each, top panel: starting material used for that reaction, bottom panel: crude HPLC of reaction products. [a]

[a] $R^{N\prime}$ = Ac-Thr(OtBu)-Tyr(OtBu)-, $R^{C\prime}$ = -Asn(Trt)-NHRink-Resin. R^{N} = Ac-Thr-Tyr-, R^{C} = - Asn-NH$_2$.
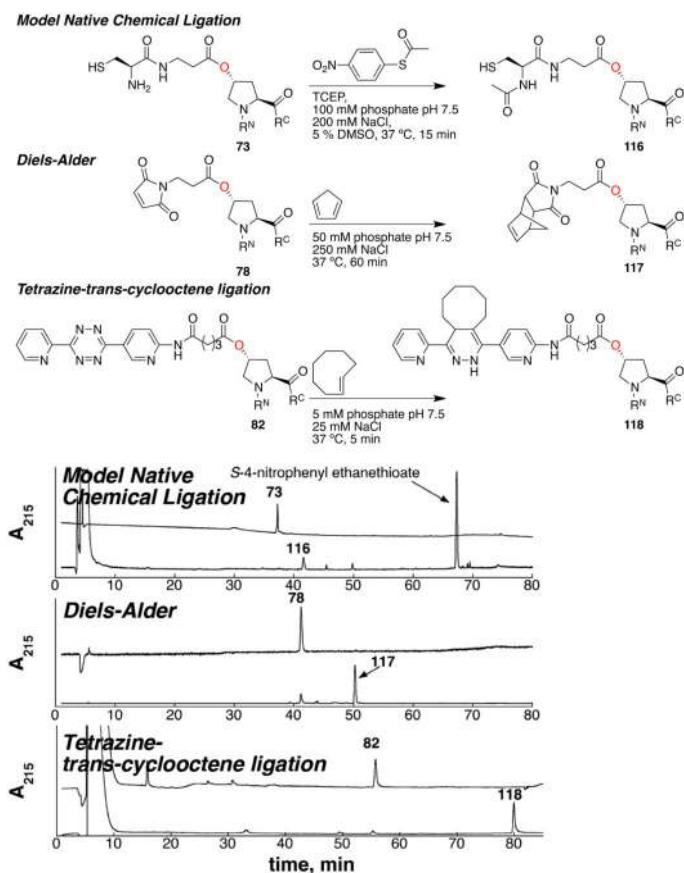
Scheme 19.
Bioorthogonal ligation reactions in aqueous solution. Bottom: HPLC chromatograms of reactions. For each, top panel: starting material used for that reaction, bottom panel: crude HPLC of reaction products. For **116** the amide product was distinguished from the thioester product based on the presence a new amide proton in the NMR spectrum. See the Supporting Information for details. Compound **117** could be the Diels-Alder endo and/or exo diastereomeric products (Diels-Alder stereochemistry not determined). Compound **118** results from isomerization of an initial 1,3-diene product.[51a] R^N = Ac-Thr-Tyr-, R^C = -Asn-NH_2.
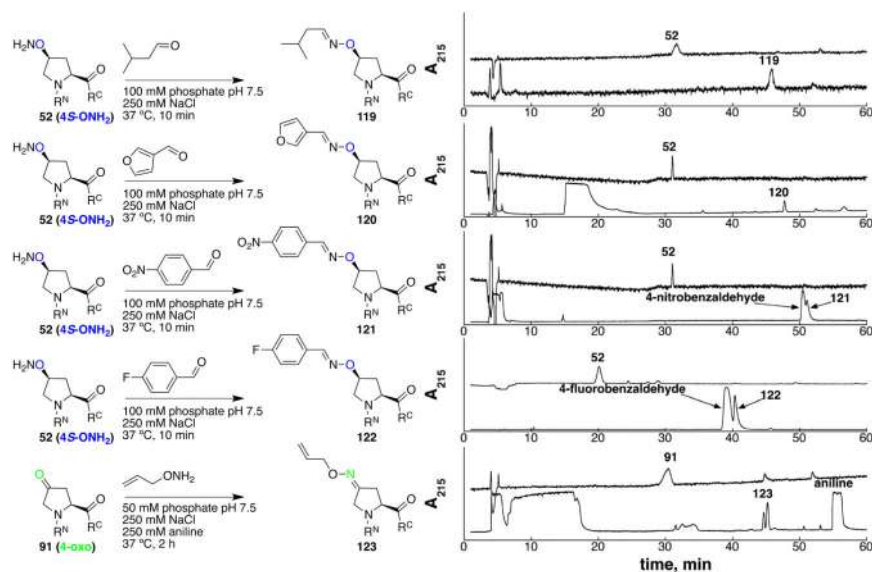
Scheme 20.

Oxime reactions in aqueous solution. Right: HPLC chromatograms of reactions. For each, top panel: starting material used for that reaction, bottom panel: crude HPLC of reaction products. The reaction with oxoproline to synthesize **123** generates two products that were indistinguishable by NMR and mass spectrometry, presumably the *E* and *Z* oximes. $R^N =$ Ac-Thr-Tyr-, $R^C =$ -Asn-NH$_2$.
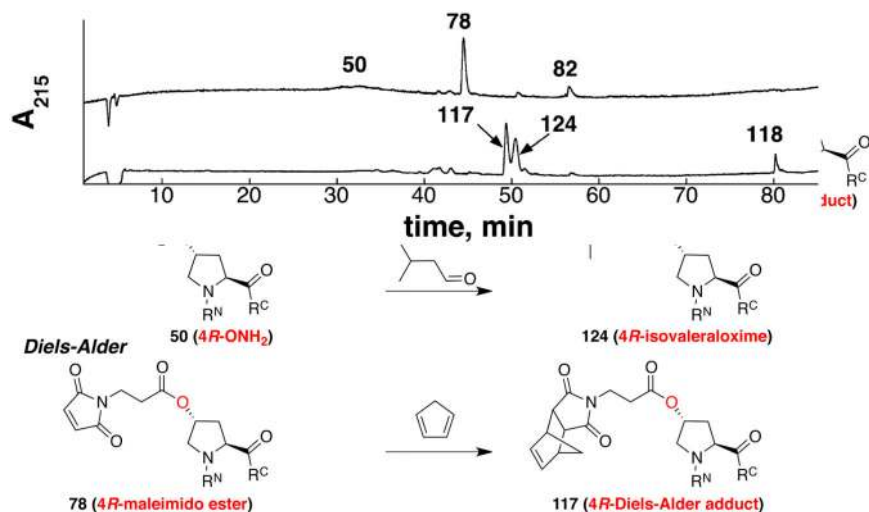
Scheme 21.

Three parallel bioorthogonal reactions in one solution.[a, b]

[a] 50 μM **50**, 50 μM **78**, and 10 μM **82** in 5 mM phophate pH 7.5 with 25 mM NaCl at 37 °C in the presence of 50 mM each 3-methylbutyraldehyde and trans-cyclooctene and 40 mM cyclopentadiene. HPLC, top: starting material peptides; bottom: crude HPLC chromatogram after 30 minutes. All products were confirmed by isolation, mass spectrometry, and comparison to authentic material synthesized independently.

[b] R^N = Ac-Thr-Tyr-, R^C = -Asn-NH_2.

Table 1

$K_{trans/cis}$ and NMR data for Ac-TYXN-NH$_2$ peptides in aqueous solution (25 mM NaCl, 5 mM phosphate (pH 4 unless otherwise indicated), 90% H$_2$O/10% D$_2$O) at 25 °C. $K_{trans/cis}$ = ratio of peptide with trans amide bond to peptide with cis amide bond as determined by NMR. $\Delta G_{trans/cis}$ = −RT ln $K_{trans/cis}$. $\Delta\Delta G_{trans/cis}$ = $\Delta G_{trans/cis}$ (peptide) − $\Delta G_{trans/cis}$ (Pro). n.d. = not determined due to spectral overlap. $^3J_{\alpha N}$ = coupling constant between H^N and H$_\alpha$, which can be correlated to ϕ via a parametrized Karplus relationship.[69] All NMR spectra are in the Supporting Information (amide region, Part 1; full NMR spectra, Part 3).

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4*R*-OBzNO$_2$ (**3**)	8.2	−1.25	−0.66	n.d.	n.d.	8.57
4*R*-OBzCF$_3$ (**57**)	8.2	−1.25	−0.66	n.d.	n.d.	8.58
4*S*-$^+$NH$_3$ (**22**)	7.5	−1.19	−0.61	6.0	8.43	8.92
4*R*-O(2,6-CH$_3$)$_2$Ph (**31**)	7.0	−1.15	−0.56	7.7	8.26	8.73
4*R*-Flp (**27**)	7.0	−1.15	−0.56	7.7	8.30	8.63
4*R*-Diels-Alder adduct(cyclopentadiene) (**117**)	6.7	−1.13	−0.54	n.d.	8.41	8.58
4*R*-β-Alanyl ester (**62**)	6.6	−1.12	−0.53	7.6	8.36	8.07
4*R*-N-biotinyl (**77**)	6.5	−1.11	−0.52	6.6	8.45	8.59
4*R*-ONH$_3$$^+$ (**50**)	6.4	−1.10	−0.51	n.d.	8.33	8.61
4*R*-OPO$_3$H (**97**)	6.3	−1.09	−0.50	n.d.	8.31	8.67
4*S*-phthalimide (**48**)	5.8	−1.04	−0.45	6.5	8.45	8.42
4*R*-OC(O)CH$_2$Br (**74**)	5.8	−1.04	−0.45	7.7	8.33	8.60
4*R*-pentynyl ester (**79**)	5.8	−1.04	−0.45	7.8	8.38	8.66

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4R-β-Ala-guanidinyl ester (64)	5.8	-1.04	-0.45	n.d.	8.30	8.54
3,4-dehydro (96)	5.7	-1.03	-0.44	n.d.	8.35	8.54
4R-succinyl ester (66)	5.7	-1.03	-0.44	n.d.	8.32	8.72
4R-Hyp (1)	5.6	-1.02	-0.43	7.5	8.34	8.59
4R-isobutylaminoacetoxy (86)	5.3	-0.99	-0.40	7.3	8.35	8.59
4R-maleinyl ester (68)	5.1	-0.96	-0.38	n.d.	8.32	8.80
4R-isobutylamino (104)	5.1	-0.96	-0.38	7.4	8.47	8.60
4R-OAc (58)	5.1	-0.96	-0.38	n.d.	8.38	8.66
4R-Alloc (15)	5.1	-0.96	-0.38	n.d.	8.35	8.64
4R-maleimido ester (78)	5.1	-0.96	-0.38	7.2	8.34	8.61
4R-(4'-Ph-3'-butynoyl) (85)	5.0	-0.95	-0.36	7.4	8.38	8.62
4R-OC$_6$F$_5$ (33)	4.9	-0.94	-0.35	n.d.	8.31	8.71
4R-hydroxyphthalimide (49)	4.7	-0.92	-0.33	6.0	8.30	8.73
4R-O-pivaloyl (60)	4.7	-0.92	-0.33	7.3	8.33	8.59
4R-OC(O)CH$_2$N$_3$ (83)	4.6	-0.90	-0.32	7.2	8.34	8.59
4R-acrylic ester (81)	4.5	-0.89	-0.30	n.d.	8.38	8.61
4R-(5'-Ph$_3$-(1H-1,2,3-triazol-1-yl)) (112)	4.5	-0.89	-0.30	7.2	8.33	8.67
4R-O-biotinyl (76)	4.4	-0.88	-0.29	7.4	8.31	8.57

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4*R*-$^+$NH$_3$ (19)	4.3	-0.86	-0.28	6.0	8.47	8.63
4*R*-fluorobenzaldehyde-oxime (122a)	4.3	-0.86	-0.28	n.d.	8.25	8.57
4*R*-s-tetrazine (82)	4.2	-0.85	-0.26	n.d.	8.34	8.61
4*R*-(4'-Ph-(1H-1,2,3-triazol-1-yl)) (108)	4.2	-0.85	-0.26	7.9	8.56	8.77
4*S*-SH (46)	4.1	-0.84	-0.25	6.0	8.46	8.33
4*S*-nitrobenzaldehyde-oxime (121)	4.1	-0.84	-0.25	4.9	8.96	8.26
4*R*-allylthioetheracetoxy (89)	4.1	-0.84	-0.25	7.2	8.31	8.61
4*R*-OMs (5)	4.1	-0.84	-0.25	n.d.	8.36	8.66
4,4-dithiolane (102)	4.1	-0.84	-0.25	6.1	8.48	8.41
4*R*-OC(CF$_3$)$_3$ (55)	4.0	-0.82	-0.23	n.d.	8.35	8.58
4*R*-OPh (29)	3.9	-0.81	-0.22	8.2	8.15	8.64
4*S*-SePh (42)	3.9	-0.81	-0.22	4.9	8.61	8.20
4*S*-(4'-Ph-(1H-1,2,3-triazol-1-yl)) (109)	3.9	-0.81	-0.22	5.2	8.55	8.14
4-oxo (91)	3.8	-0.79	-0.20	5.7	8.56	8.72
4*R*-OPO$_3$Et$_2$ (100)	3.8	-0.79	-0.20	n.d.	8.33	8.67
4*R*-OMst (13)	3.7	-0.77	-0.19	6.6	8.33	8.61
4*R*-azido (18)	3.6	-0.76	-0.17	n.d.	8.32	8.60
4*S*-(4'-hexyl-(1H-1,2,3-triazol-1-yl)) (111)	3.6	-0.76	-0.17	5.6	8.53	8.19

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4S-Grubbs metathesis (115)	3.5	-0.74	-0.15	7.8	8.33	8.30
4R-ONs (9)	3.5	-0.74	-0.15	n.d.	8.32	8.62
4R-OPhCN (35)	3.4	-0.72	-0.14	n.d.	8.32	8.64
4R-guanidinyl (20)	3.4	-0.72	-0.14	n.d.	8.38	8.66
4R-phthalimide (47)	3.3	-0.71	-0.12	7.6	8.44	8.60
4R-OTs (7)	3.2	-0.69	-0.10	n.d.	8.30	8.58
4R-glutaryl ester (70)	3.2	-0.69	-0.10	n.d.	8.31	8.61
4S-SO$_2$Ph (95)	3.2	-0.69	-0.10	5.2	8.54	8.39
4R-SH (44)	3.1	-0.67	-0.08	6.8	8.40	8.52
4S-SPhNO$_2$ (40)	3.1	-0.67	-0.08	4.2	8.62	7.98
4S-iodo (24)	3.1	-0.67	-0.08	4.2	8.58	8.32
4S-SAc (45)	3.0	-0.65	-0.06	5.0	8.54	8.08
4R-S(O)Ph (92)	3.0	-0.65	-0.06	n.d.	8.53	8.37
4R-SePh (41)	2.9	-0.63	-0.04	6.2	8.40	8.48
4S-SPh (17)	2.9	-0.63	-0.04	4.0	8.56	8.03

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4R-SPhNO$_2$ (39)	2.8	-0.61	-0.02	6.6	8.40	8.57
4S-(5'-Ph-(1H-1,2,3-triazol-1-yl)) (113)	2.8	-0.61	-0.02	5.3	8.56	8.10
4R-(4'-Hex-(1H-1,2,3-triazol-1-yl)) (110)	2.8	-0.61	-0.02	6.9	8.33	8.68
H (Pro)	2.7	-0.59	0.00	6.1	8.37	8.40
4S-β-Ala-guanidinyl ester (64)	2.7	-0.59	0.00	n.d.	8.36	8.67
4R-SAc (43)	2.7	-0.59	0.00	6.4	8.42	8.52
4R-SPh (16)	2.7	-0.59	0.00	6.2	8.40	8.52
4S-S(O)Ph (93)	2.7	-0.59	0.00	4.2	8.65	8.40
4S-hyp (4)	2.7	-0.59	0.00	4.9	8.49	8.08
4,4-F$_2$ (101)	2.6	-0.57	0.02	n.d.	8.56	8.72
4-allyl oxime (123)	2.4	-0.52	0.07	7.2	8.42	8.57
4S-(2'-allylthioether)acetoxy (90)	2.4	-0.52	0.07	5.6	8.51	8.10
4S-OTs (8)	2.4	-0.52	0.07	4.1	8.55	7.74
4S-OPO$_3$H (98)	2.3	-0.49	0.09	5.3	8.49	8.08
4S-maleinyl ester (69)	2.2	-0.47	0.12	5.4	8.56	8.01
4R-SO$_2$Ph (94)	2.2	-0.47	0.12	5.5	8.52	8.61

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4S-chloro (53)	2.1	-0.44	0.15	4.6	8.60	8.04
4S-OPhI (37)	2.1	-0.44	0.15	5.4	8.58	7.72
4S-(O(2,6-CH$_3$)$_2$Ph) (32)	2.1	-0.44	0.15	5.6	8.48	8.02
4S-bromo (25)	2.0	-0.41	0.18	4.2	8.60	8.06
4S-ONs (10)	2.0	-0.41	0.18	4.5	8.58	7.97
4S-guanidinyl (23)	2.0	-0.41	0.18	5.4	8.58	8.27
4S-Ph-ethynylphenoxy (107)	1.9	-0.38	0.21	4.8	8.54	7.72
4S-OC$_6$F$_5$ (34)	1.9	-0.38	0.21	4.4	8.56	7.70
4S-OMst (14)	1.9	-0.38	0.21	5.4	8.53	8.18
4S-OBzCF$_3$ (28)	1.9	-0.38	0.21	5.4	8.54	8.09
4S-O-p-methoxybiphenyl (106)	1.8	-0.35	0.24	4.8	8.57	7.77
4S-OPhBr (38)	1.8	-0.35	0.24	4.8	8.58	7.73
4S-OPO$_3$Et$_2$ (100)	1.8	-0.35	0.24	5.4	8.56	8.05
4S-OBzNO$_2$ (2)	1.8	-0.35	0.24	5.3	8.56	8.17
4S-azido (21)	1.8	-0.35	0.24	4.2	8.57	7.91
4S-succinyl ester (67)	1.8	-0.35	0.24	4.1	8.57	8.07

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4S-OPhCN (36)	1.7	-0.31	0.27	4.7	8.60	7.75
4S-OPh (30)	1.7	-0.31	0.27	4.6	8.56	8.10
4S-fluorobenzaldehyde-oxime (122)	1.5	-0.24	0.35	4.8	8.56	7.72
4S-flp (26)	1.5	-0.24	0.35	4.9	8.58	8.16
4S-OMs (6)	1.5	-0.24	0.35	4.8	8.62	8.07
4S-allylthioetheracetoxy (90)	1.4	-0.20	0.39	4.2	8.59	8.07
4S-nitrobenzaldehyde-oxime (121)	1.4	-0.20	0.39	5.4	8.49	7.91
4S-OAc (59)	1.4	-0.20	0.39	4.8	8.59	8.06
4S-β-Alanyl ester (63)	1.4	-0.20	0.39	5.2	8.57	8.31
4S-OC(O)CH$_2$N$_3$ (84)	1.3	-0.16	0.43	3.7	8.58	8.06
4S-OC(O)CH$_2$Br (75)	1.3	-0.16	0.43	4.5	8.60	8.06
4S-(2'-(4'-aminophenylamino)acetoxy (88)	1.3	-0.16	0.43	4.3	8.60	8.21
4S-O-pivaloyl (61)	1.3	-0.16	0.43	4.4	8.53	8.18
4S-OC(CF$_3$)$_3$ (56)	1.2	-0.11	0.48	4.7	8.59	8.24
4S-hydroxyphthalimide (51)	1.0	0.00	0.59	6.0	8.47	8.05
4S-ONH$_3^+$ (52)	1.0	0.00	0.59	4.4	8.57	8.05

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$
4S-isobutylaminoacetoxy (87)	0.9	0.06	0.65	4.2	8.62	8.24
(2R,4R)-D-hyp (126)	all trans	n.d.	n.d.	2.5	8.34	8.33
(2R,4S)-D-Hyp (125)	7.2	-1.17	-0.58	8.4	8.53	8.67

Table 2

Overall functional group effects on cis-trans isomerism in Ac-TYXN-NH$_2$ peptides. $\Delta\Delta G = \Delta G_{trans/cis}(4R$ stereoisomer) $- \Delta G_{trans/cis}(4S$ stereoisomer), with negative ΔG consistent with the magnitude of stereoelectronic effects and positive ΔG indicating steric effects dominating over stereoelectronic effects. By definition, for Pro ($4R = 4S =$ H), $\Delta\Delta G = 0.00$.

Ac-TYP(4-X)N-NH$_2$ X =	$\Delta\Delta G_{(4R-4S)}$ kcal mol^{-1}
ONH$_3^+$	-1.13
isobutylaminoacetoxy	-0.93
β-Alanyl ester	-0.92
F	-0.91
OBzNO$_2$	-0.90
OC(O)CH$_2$Br	-0.88
OBzCF$_3$	-0.87
hydroxyphthalimide	-0.81
OAc	-0.76
O-pivaloyl	-0.76
OC(O)CH$_2$N$_3$	-0.74
O-(2,6-(CH$_3$)$_2$Ph)	-0.71
OC(CF$_3$)$_3$	-0.71
succinyl ester	-0.68
allylthioetheracetoxy	-0.64
OPO$_3$H$^-$	-0.60
OMs	-0.60
OC$_6$F$_5$	-0.56
OPh	-0.50
maleinyl ester	-0.49
β-Ala-guanidinyl ester	-0.45
OPO$_3$Et$_2$	-0.44
OH	-0.43
azido	-0.41
OPhCN	-0.41
OMst	-0.39
ONs	-0.33
guanidinyl	-0.31
5′-Ph-(1H-1,2,3-triazol-1-yl)	-0.28
OTs	-0.17
S(O)Ph	-0.06
4′-Ph-(1H-1,2,3-triazol-1-yl)	-0.04
SPh	0.04
SPhNO$_2$	0.06

Ac-TYP(4-X)N-NH$_2$ X =	$\Delta\Delta G_{(4R\text{-}4S)}$ kcal mol^{-1}
SAc	0.06
4′-hexyl-(1H-1,2,3-triazol-1-yl)	0.15
SH	0.17
SePh	0.18
SO$_2$Ph	0.22
phthalimide	0.33
$^+$NH3	0.33

Table 3

NMR data for ionizable derivatives as a function of pH and protonation state. Expected major protonation state at the given pH is indicated.

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG kcal mol^{-1}	$\Delta\Delta G$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4R-$^+$NH$_3$ (**19**)	4.3	-0.86	-0.28	7.0	8.46	8.62	2.0
4S-$^+$NH$_3$ (**22**)	7.5	-1.19	-0.60	6.0	8.43	8.92	2.0
4R-$^+$NH$_3$ (**19**)	4.2	-0.85	-0.26	5.4	8.47	8.63	4.0
4S-$^+$NH$_3$ (**22**)	7.4	-1.18	-0.60	6.6	8.41	8.91	4.0
4R-$^+$NH$_3$ (**19**)	4.0	-0.82	-0.23	6.6	8.44	8.61	6.3
4S-$^+$NH$_3$ (**22**)	7.3	-1.18	-0.59	6.1	8.45	8.93	6.3
4R-$^+$NH$_3$ (**19**)	3.9	-0.81	-0.22	n.d.	n.d.	n.d.	7.6
4S-$^+$NH$_3$ (**22**)	6.8	-1.13	-0.55	n.d.	n.d.	n.d.	7.6
4R-NH$_2$ (**19**)	4.1	-0.84	-0.25	n.d.	n.d.	n.d.	8.5
4S-NH$_2$ (**22**)	6.2	-1.08	-0.49	n.d.	n.d.	n.d.	8.5
Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG kcal mol^{-1}	$\Delta\Delta G$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4R-ONH$_3^+$ (**50**)	6.4	-1.10	-0.51	n.d.	8.32	8.57	2.0

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG kcal mol^{-1}	$\Delta\Delta G$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4*S*-ONH$_3^+$ (52)	1.0	0.00	0.59	4.4	8.57	8.05	2.0
4*R*-ONH$_3^+$ (50)	5.9	-1.05	-0.46	n.d.	8.32	8.61	4.0
4*S*-ONH$_3^+$ (52)	1.4	-0.20	0.39	4.4	8.58	7.83	4.0
4*R*-ONH$_2$ (50)	5.5	-1.01	-0.42	n.d.	8.33	8.61	5.6
4*S*-ONH$_2$ (52)	1.4	-0.20	0.39	4.4	8.58	7.83	5.6
4*R*-ONH$_2$ (50)	3.2	-0.69	-0.10	n.d.	n.d.	n.d.	7.6
4*S*-ONH$_2$ (52)	1.5	-0.24	0.35	n.d.	n.d.	n.d.	7.6
4*R*-ONH$_2$ (50)	2.7	-0.59	0.00	n.d.	n.d.	n.d.	10.0
4*S*-ONH$_2$ (52)	1.4	-0.20	0.39	n.d.	n.d.	n.d.	10.0

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG kcal mol^{-1}	$\Delta\Delta G$ kcal mol^{-1}	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4*R*-OPO$_3$H (97)	6.3	-1.09	-0.50	n.d.	8.31	8.67	4.0
4*S*-OPO$_3$H$^-$ (98)	2.3	-0.49	0.10	5.3	8.49	8.08	4.0
4*R*-OPO$_3$H (97)	6.0	-1.06	-0.47	n.d.	8.30	8.66	5.6
4*S*-OPO$_3$H$^-$ (98)	2.2	-0.47	0.12	6.0	8.47	8.14	5.6

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG (kcal mol^{-1})	$\Delta\Delta G$ (kcal mol^{-1})	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4R-OPO$_3$H (97)	5.9	-1.05	-0.46	7.9	8.20	8.68	6.8
4S-OPO$_3$H$^-$ (98)	2.9	-0.63	-0.04	6.6	8.45	8.21	6.8
4R-OPO$_3^{2-}$ (97)	6.7	-1.13	-0.54	n.d.	n.d.	n.d.	8.5
4S-OPO$_3^{2-}$ (98)	2.9	-0.63	-0.04	n.d.	n.d.	n.d.	8.5

Ac-TYP(X)N-NH$_2$ X =	$K_{trans/cis}$	ΔG (kcal mol^{-1})	$\Delta\Delta G$ (kcal mol^{-1})	$^3J_{\alpha N}$ Tyr$_{cis}$	δ, H^N Tyr$_{cis}$	δ, H^N Asn$_{trans}$	pH
4R-SH (44)	3.1	-0.67	-0.08	6.8	8.40	8.52	4.0
4S-SH (46)	4.1	-0.84	-0.25	6.0	8.46	8.33	4.0
4R-SH (44)	3.2	-0.69	-0.10	7.1	8.41	8.52	6.8
4S-SH (46)	4.6	-0.90	-0.32	7.0	8.46	8.33	6.8
4R-SH (44)	2.8	-0.61	-0.02	-	-	-	7.6
4S-SH (46)	4.4	-0.88	-0.29	-	-	-	7.6
4R-S$^-$ (44)	3.0	-0.65	-0.06	-	-	-	8.5
4S-S$^-$ (46)	4.4	-0.88	-0.29	-	-	-	8.5

Table 4

$K_{trans/cis}$ for Ac-TAXN-NH$_2$ peptides. $\Delta G = -RT \ln K_{trans/cis}$. $\Delta\Delta G_{trans/cis} = \Delta G_{trans/cis}$ (peptide) $- \Delta G_{trans/cis}$ (Pro). n.d. = not determined due to spectral overlap. $^3J_{\alpha N}$ = coupling constant between H^N and H$_\alpha$, which can be correlated to ϕ via a parametrized Karplus relationship.

Ac-TAP(4-X)N-NH$_2$X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol^{-1}	$\Delta\Delta G_{trans/cis}$ kcal mol^{-1}	$^3J_{\alpha N}$ Ala$_{cis}$	δ, H^N Ala$_{cis}$	δ, H^N Asn$_{trans}$
Flp (132)	23.8	-1.88	-0.48	7.2	8.31	8.68
Hyp (127)	20.8	-1.80	-0.40	6.6	8.60	8.65
4R-OBzNO$_2$ (129)	19.0	-1.74	-0.34	n.d.	n.d.	8.72
4R-OPO$_3$H (146)	18.2	-1.72	-0.32	5.9	8.32	8.69
4R-azido (139)	16.3	-1.65	-0.25	n.d.	n.d.	8.67
4R-OBzCF$_3$ (135)	15.1	-1.61	-0.21	n.d.	8.89	8.68
4R-ONs (143)	11.6	-1.45	-0.05	6.6	8.30	8.68
4R-OMs (137)	11.6	-1.45	-0.05	n.d.	n.d.	8.89
4R-OPO$_3$Et$_2$ (148)	11.1	-1.42	-0.02	5.7	8.36	8.71
H (Pro)	*10.7*	*-1.40*	*0.00*	*5.4*	*8.32*	*8.45*
4-oxo (131)	10.1	-1.37	0.03	6.0	8.50	8.61
4R-OTs (141)	9.2	-1.31	0.09	6.8	8.31	8.67
4,4-F$_2$ (134)	8.5	-1.27	0.13	n.d.	8.53	8.69

Ac-TAP(4-X)N-NH₂X =	$K_{trans/cis}$	$\Delta G_{trans/cis}$ kcal mol⁻¹	$\Delta\Delta G_{trans/cis}$ kcal mol⁻¹	$^3J_{\alpha N}$ Ala$_{cis}$	δ, Hᴺ Ala$_{cis}$	δ, Hᴺ Asn$_{trans}$
4S-OPO₃H (147)	8.5	-1.27	0.13	4.8	8.42	8.17
hyp (130)	8.2	-1.24	0.16	5.4	8.44	8.18
4S-azido (140)	6.5	-1.11	0.29	6.2	8.41	8.28
4S-OPO₃Et₂ (149)	5.8	-1.04	0.36	4.7	8.49	8.22
flp (133)	5.5	-1.01	0.39	4.8	8.53	8.34
4S-OBzCF₃ (136)	5.1	-0.96	0.44	4.4	8.49	8.33
4S-OTs (142)	4.7	-0.92	0.48	4.7	8.46	8.19
4S-ONs (144)	4.7	-0.92	0.48	n.d.	n.d.	8.26
4S-OBzNO₂ (128)	4.7	-0.92	0.48	4.8	8.49	8.37
4S-OPhCN (145)	4.4	-0.88	0.52	n.d.	n.d.	8.21
4S-OMs (138)	4.3	-0.86	0.54	n.d.	8.73	8.44

Table 5

Overall functional group effects on cis-trans isomerism in Ac-TAXN-NH$_2$ peptides. $\Delta\Delta G = \Delta G_{\text{trans/cis}}(4R$ stereoisomer$) - \Delta G_{\text{trans/cis}}(4S$ stereoisomer$)$, with negative ΔG consistent with the magnitude of stereoelectronic effects and positive indicating steric effects dominating over stereoelectronic effects. By definition, for Pro ($4R = 4S = $ H), $\Delta\Delta G = 0.00$.

Ac-TAP(4-X)N-NH$_2$ X =	$\Delta\Delta G_{(4R\text{-}4S)}$ kcal mol^{-1}
F	-0.87
OBzNO$_2$	-0.82
OBzCF$_3$	-0.65
OMs	-0.59
OH	-0.56
azido	-0.54
ONs	-0.53
OPO$_3$H$^-$	-0.45
OTs	-0.39
OPO$_3$Et$_2$	-0.38

Table 6

^{19}F NMR chemical shifts. Chemical shifts are listed in order from upfield to downfield based on δ in the trans conformation. n.d. = not determined due to spectral overlap.

Peptide	$δ_{trans}$, ppm	Multiplicity (J, Hz)	$δ_{cis}$, ppm	Multiplicity (J, Hz)
Ac-TYFlpN-NH$_2$ (**27**)	−177.1	s	−177.9	s
Ac-TYflpN-NH$_2$ (**26**)	−173.2	s	−173.5	s
Ac-TYP(4R-OC$_6$F$_5$)N-NH$_2$ (m) (**33**)	−163.2	dd (24, 24)	−163.6	dd (20, 20)
Ac-TYP(4R-OC$_6$F$_5$)N-NH$_2$ (p) (**33**)	−161.9	t (24)	−162.1	t (20)
Ac-TYP(4R-OC$_6$F$_5$)N-NH$_2$ (o) (**33**)	−156.5	d (24)	−156.1	d (20)
Ac-TYP(4S-OC$_6$F$_5$)N-NH$_2$ (m) (**34**)	−163.4	dd (24, 24)	−163.6	dd (24, 24)
Ac-TYP(4S-OC$_6$F$_5$)N-NH$_2$ (o) (**34**)	−162.4	t (24)	−162.9	t (24)
Ac-TYP(4S-OC$_6$F$_5$)N-NH$_2$ (o) (**34**)	−156.3	d (24)	−156.3	n.d.
Ac-TYP(4,4-F$_2$)N-NH$_2$ (**101**)	−100.6	s	−103.2, −103.9	d (252)
Ac-TAP(4,4-F$_2$)N-NH$_2$ (**130**)	−100.8, −101.4	d (248)	−94.8, −95.4	d (252)
			−104.7, −105.3	d (252)
	−99.7, −100.3	d (248)	−94.3, −94.9	d (252)
Ac-TYP(4R-OC(CF$_3$)$_3$)N-NH$_2$ (**55**)	−70.7	s	−70.8	s
Ac-TYP(4S-OC(CF$_3$)$_3$)N-NH$_2$ (**56**)	−70.0	s	−70.7	s