# Tunable Control of Polyproline Helix (PPII) Structure via Aromatic Electronic Effects: An Electronic Switch of Polyproline Helix

Anil K. Pandey, Krista M. Thomas, Christina R. Forbes, and Neal J. Zondlo*

Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716, United States

S *Supporting Information*

ABSTRACT: Aromatic rings exhibit defined interactions via the unique aromatic π face. Aromatic amino acids interact favorably with proline residues via both the hydrophobic effect and aromatic−proline interactions, C−H/π interactions between the aromatic π face and proline ring C−H bonds. The canonical aromatic amino acids Trp, Tyr, and Phe strongly disfavor a polyproline helix (PPII) when they are present in proline-rich sequences because of the large populations of *cis* amide bonds induced by favorable aromatic−proline interactions (aromatic−*cis*-proline and proline−*cis*-proline−aromatic interactions). We demonstrate the ability to tune polyproline helix conformation and *cis*−*trans* isomerism in proline-rich sequences using aromatic electronic effects. Electron-rich aromatic residues strongly disfavor polyproline helix and exhibit large populations of *cis* amide bonds, while electron-poor aromatic residues exhibit small populations of *cis* amide bonds and favor polyproline helix. 4-Aminophenylalanine is a pH-dependent electronic switch of polyproline helix, with *cis* amide bonds favored as the electron-donating amine, but *trans* amide bonds and polyproline helix preferred as the electron-withdrawing ammonium. Peptides with block proline−aromatic PPXPPXPPXPP sequences exhibited electronically switchable pH-dependent structures. Electron-poor aromatic amino acids provide special capabilities to integrate aromatic residues into polyproline helices and to serve as the basis of aromatic electronic switches to change structure.
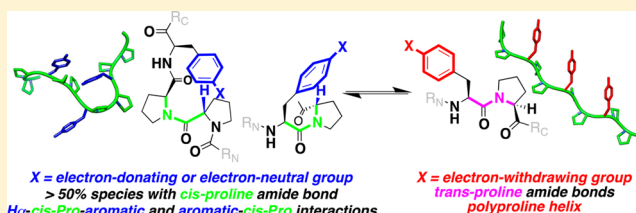
Aromatic amino acids play distinct structural and functional roles, because of the combination of hydrophobicity with the unique interactions possible via the negatively charged aromatic faces and the positively charged aromatic edges.[1−5] Aromatic quadrupoles, in contrast to simple aliphatic hydrophobic groups (e.g., cyclohexyl or *tert*-butyl), provide the possibility of specific structural interactions and well-defined geometric orientations for aromatic rings that are distinct from those possible solely via the hydrophobic effect.

Aromatic residues exhibit special interactions with proline residues, which promote *cis* amide bonds via local aromatic−proline interactions (either aromatic−*cis*-proline sequences or Hα-*cis*-proline−aromatic sequences) (Figure 1).[6−26] In the Protein Data Bank (PDB) and in model peptides, proline−proline and aromatic−proline sequences are the most likely to adopt a *cis* amide bond. In aromatic−proline sequences, the population of *cis* amide bonds correlates with aromatic electronics, consistent with a C−H/π interaction in which the partial positive charge on the hydrogen of a polarized proline C−H bond (i.e., Hα or Hδ, adjacent to the electron-withdrawing amide carbonyl and/or amide nitrogen) interacts with the negatively charged face of the aromatic ring to stabilize the *cis* amide bond, in a manner comparable to a classical cation−π interaction.[15,22,24,25,27−29] The aromatic−proline interaction also could potentially, with appropriate geometry, be additionally stabilized by interactions between the aromatic π orbitals and the σ^* orbital of the C−H bond, which is consistent with the observation of sub-van der Waals distances in a significant number of C−H/π interactions, including those

in small molecules, peptides, and proteins.[27,28,30−35] Thus, the aromatic−proline interaction can be considered electrostatic ($\delta^-_{aromatic} \cdot \delta^+_{proline}$) and/or potentially stereoelectronic ($\pi_{aromatic} \rightarrow \sigma^*_{proline\ C−H}$) in nature, in addition to contributions from the hydrophobic effect. The strength of a C−H/π interaction is dependent on the electronics of the aromatic system, as is typical for cation−π and polar X−H/π interactions, with stronger interactions observed for more electron-rich aromatics and weaker interactions for electron-poor aromatics.[3,27,29,36,37]

Proline-rich domains, as defined in ref 38, are among the most common domains in eukaryotes. Proline-rich domains have central functional roles, including in protein−protein interactions, in linking globular or other functional domains, and in responsiveness to phosphorylation.[39−43] Proline-rich domains are considered to be intrinsically disordered, rendering them resistant to structural analysis by X-ray crystallography. To understand protein structure within proline-rich domains, we recently analyzed the propensities of the 20 canonical amino acids to adopt the polyproline helix (PPII) conformation, via a host−guest model system that was designed to understand the conformational preferences in a typical proline-rich sequence (Ac-GPPXPPGY-NH$_2$, where X is any canonical amino acid).[23] In that work, we found that the aromatic residues Phe, Tyr, and Trp strongly disfavor polyproline helix when they are present in
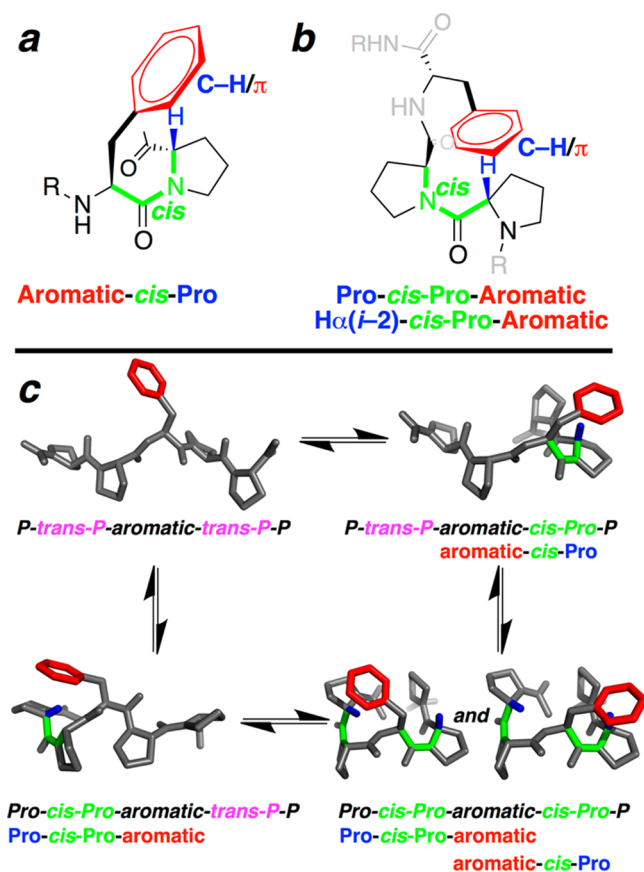
Figure 1. Aromatic−proline interactions that favor *cis* amide bonds in proline-rich sequences. (a) Aromatic−*cis*-proline interaction. (b) Hα-*cis*-proline−aromatic interaction. Hα (*i* − 2)-*cis*-proline−aromatic interactions may occur with the Hα of any residue two residues prior to (*i* − 2 to) the aromatic residue but are most favorable with proline.[20] (c) Conformational heterogeneity in proline-rich sequences (PPFPP) due to multiple aromatic−proline interactions. Additional *cis*−*trans* isomerism is possible at other X-Pro sequences.

proline-rich sequences, via the induction of very large populations of *cis* amide bonds (45−60% of all species contained at least one *cis* amide, compared to ≤10% of species with a *cis* amide bond for nonaromatic residues X). The very large population of *cis* amide bonds for peptides with aromatic residues in this context, which inherently prevents polyproline II helix formation in the residues involved in the *cis* amide bond, was primarily due to substantial induction of *cis* amide bonds at both the Pro2−*cis*-Pro3−aromatic4 and the aromatic4−*cis*-Pro5 sequences (Figure 1). In addition, bioinformatics analysis of proline-rich sequences from diverse eukaryotes (humans, mice, *Drosophila*, *Caenorhabditis elegans*, and *Arabidopsis*) demonstrated very strong biases against the incorporation of aromatic residues in proline-rich sequences, presumably because of substantial conformational heterogeneity that would result from multiple *cis*−*trans* isomerism events. Notably, collagen sequences (containing ProHypGly consensus repeats that assemble into a triple helix of polyproline helices) also exhibit very strong preferences against replacement of either the Pro or Hyp residues with aromatic amino acids, presumably because of similar problems of conformational heterogeneity and resultant slow kinetics in collagen folding and assembly.[44] Collectively, these data suggest that the very poor PPII propensity of aromatic residues in proline-rich

sequences, in contrast to non-proline-rich sequences, is due to the special interactions of the proline residues with the aromatic rings inducing multiple *cis* amide bonds and non-PPII structure (Figure 1c).[45−50]

The polyproline helix is a fundamental secondary structure of proteins that is widely employed in molecular design because of its rigidity and lack of dependence on hydrogen bonding.[45,46,51−63] Notably, both aromatic residues and polyproline helices are, separately, broadly employed in molecular recognition.[1,2,5,39] The incorporation of aromatic amino acids within polyproline helices thus could be exploited in molecular recognition and biomolecular design. In peptides with aromatic−proline dipeptide sequences, the population of *trans* amide bonds can be increased using electron-poor aromatics, which weakens the aromatic−proline interaction. These data suggest that electron-deficient aromatic amino acids in proline-rich sequences could potentially promote polyproline helix. Herein, we examine the ability to control structure in proline-rich peptides using aromatic electronic effects.

■ EXPERIMENTAL PROCEDURES

Peptide Synthesis. Peptides were synthesized via standard solid-phase peptide synthesis and purified to homogeneity via high-performance liquid chromatography (HPLC), as determined by the presence of a single peak on HPLC re-injection (>95% purity). Peptide synthesis and characterization details are given in the Supporting Information.

Circular Dichroism (CD). CD experiments were conducted on a Jasco J-810 spectropolarimeter. Unless otherwise indicated, peptides were analyzed at 25 °C in aqueous buffer containing 5 mM phosphate and 25 mM KF. CD data are the average of at least three independent trials. Data are background-corrected but not smoothed. Error bars indicate the standard error.

Nuclear Magnetic Resonance (NMR) Spectroscopy. NMR experiments were conducted at 23 °C in 90% H_2O/10% D_2O containing 5 mM phosphate buffer (pH 4 or as indicated) and 25 mM NaCl, unless otherwise indicated. NMR experiments were conducted using watergate water suppression. Populations of species with all-*trans* amide bonds versus species with one or more *cis* amide bonds were determined via integration of all peaks associated with a resonance in the one-dimensional spectrum using assignments from TOCSY spectra or by analogy. Each calculated % *cis* (= 100 − % of the all-*trans* species = 100 × (sum of all species containing at least one *cis* amide bond divided by the total of all species, including the all-*trans* species)) involved integration of multiple sets of resonances to confirm the conformational identity. In peptides, % *cis* includes all populations of peptide species containing at least one *cis* amide bond, compared to the peptide species containing all-*trans* amide bonds. The error in each % *cis* is estimated to be ≤±3 percentage points.

■ RESULTS AND DISCUSSION

To test whether the electronics of aromatic amino acids could be used to control *cis*−*trans* isomerism and polyproline helix structure within proline-rich sequences, a series of Ac-GPPXPPGY-NH₂ peptides was synthesized, where X is one of a series of electron-rich to electron-poor aromatic amino acids. Peptides were synthesized using both commercially available amino acids, as well as derivatives of 4-thiophenylalanine, which were synthesized from the peptide containing 4-

iodophenylalanine and a copper-mediated cross-coupling reaction of thiolacetic acid with the fully synthesized peptide on the solid phase. Thiolysis of the resultant peptide yielded 4-thiophenylalanine.[24]

All peptides were analyzed by NMR spectroscopy, and the populations of species with all-*trans* amide bonds versus those with one or more *cis* amide bonds were quantified (Figure 2 and Table 1). As expected, the peptide containing the electron-donating amino substituent (4-NH_2-Phe) exhibited large populations of species with *cis* amide bonds, similar to those observed with Tyr. In contrast, increasingly electron-withdrawing substituents exhibited reduced populations of *cis* amide bonds, in a manner concomitant with aromatic electronics. These data are consistent with the interpretation that large populations of *cis* amide bonds for aromatic amino acids in proline-rich sequences are due to C−H/π interactions, which are weaker with less electron-rich aromatics. As expected for a π-facial interaction,[3,64−66] the effects observed were broadly independent of the position on the aromatic ring (2- vs 3- vs 4-substituted). These data demonstrate the ability to modulate *cis*−*trans* isomerism within a proline-rich sequence using aromatic electronics. Interestingly, by NMR, significant upfield shifts were observed in a subset of proline resonances (Figures S14−S24 of the Supporting Information), compared to the peptides for which X is alanine, consistent with aromatic−proline interactions that place these proline hydrogens near the face of the aromatic ring. Moreover, the extent of upfield proline chemical shifts correlated with aromatic electronics, with the largest effects in the X = tryptophan peptide, the smallest effects in the X = pentafluorophenylalanine peptide, and intermediate effects in peptides with other aromatic amino acids. Notably, the $^3J_{\alpha N}$ coupling constant, which correlates with the backbone torsion angle ϕ,[67] of the aromatic guest residue (X) in the species with all-*trans* amide bonds was 6.9−7.9 Hz. These data indicate the compatibility of aromatic residues with polyproline helix when in the all-*trans* amide conformation, as has been observed previously for aromatic residues in glycine-rich sequences.[45−50]

Among the aromatic amino acids examined, the smallest populations of *cis* amide bonds were observed for peptides with electron-withdrawing substituents, including 4-nitrophenylalanine, 4-thioacetylphenylalanine, protonated 4-aminophenylalanine, and pentafluorophenylalanine. In contrast to most aromatic residues, the peptide with pentafluorophenylalanine exhibited only one major species by NMR, with populations of species with *cis* amide bonds that were close to the small populations of *cis* amide bonds observed in peptides with aliphatic amino acids (≤10% *cis*).[23] To identify the compatibility of electron-poor aromatic amino acids with PPII, the peptide with pentafluorophenylalanine was examined by CD. The CD spectrum of this peptide exhibited a significant positive band with a λ_{max} of 223 nm and a larger minimum at a λ_{min} of 204 nm, typical of PPII observed in peptides and proteins,[23,45,68−71] indicating that the peptide with pentafluorophenylalanine significantly adopts PPII (Figure 3).

The peptides with thioacetate/thiol/thiolate substituents (4-thiophenylalanine, pK_a = 6.4[24]) or ammonium/amine substituents [conjugate acid of aniline, pK_a = 4.9; measured pK_a within peptide of 4.7 ± 0.2 (Figure S7 of the Supporting Information)] exhibited NMR spectra that indicated that peptide structure could be controlled by chemical reactions (e.g., thiolysis of a thioester) or solution acidity. To examine the use of pH as an electronic switch of structure, the peptide
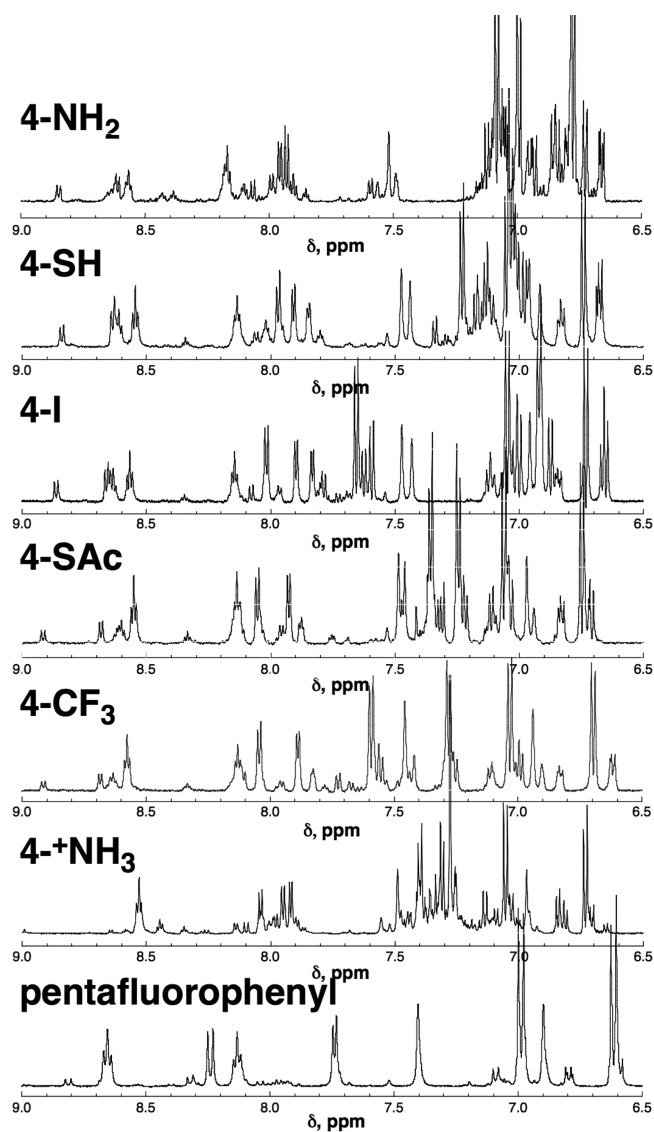

Figure 2. ^1H NMR spectra (amide−aromatic region) of Ac-GPPXPPGY-NH_2 peptides, where X is phenylalanine substituted as indicated. NMR spectra for peptides for which X is Phe, Tyr, or Trp are shown in ref 23. All other NMR spectra are given in the Supporting Information. In the peptide with pentafluorophenylalanine (bottom), the peak assignments for the major (all-*trans* amide bonds) species are δ 8.66 (Gly7 amide), 8.24 (F_5-Phe amide), 8.13 (Gly1 amide), 7.74 (Tyr8 amide), 7.40 (-$CONH_2$), 6.99 (Tyr aromatic), 6.90 (-$CONH_2$), and 6.62 (Tyr aromatic). Analogous resonance assignments apply for the all-*trans* amide species in other peptides. For species with *cis* amide bonds, glycine amides are readily identified by their pseudotriplet (doublet of doublet) coupling patterns, while C-terminal carboxamides are readily identified as singlets, with the singlet at ~7.40 ppm being a relatively isolated resonance that in most cases can be used to determine the number and relative populations of the species present. The glycine amide peaks downfield of the all-*trans* species at ~8.5 ppm also are diagnostic of the relative populations of the species with *cis* amide bonds. Control experiments with Pro3 or Pro5 replaced with Leu or Hyp confirmed that the major *cis*−*trans* isomerization events were those indicated in Figure 1.

Ac-GPPXPPGY-NH_2 (X = 4-aminophenylalanine) was analyzed by CD in both the ammonium (pH 3) and amine (pH 7) protonation states (Figure 3). At pH 7 (Ar-NH_2 protonation state), this peptide exhibited a diffuse CD spectrum with a positive band with a λ_{max} of 240 nm, substantially red-shifted

Table 1. Summary of ^{1}H NMR Data for All Ac-GPPXPPGY-NH$_2$ Peptides^a

X	% cis	$K_{trans/\sum cis}$
Trp	63	0.59
4-OH-Phe (Tyr)	58	0.74
4-S$^-$-Phe	57	0.77
4-NH$_2$-Phe	56	0.80
4-SH-Phe	56	0.80
Phe (4-H)	51	0.97
3,4-di-F-Phe	49	1.1
4-OAc-Phe (TyrAc)	49	1.1
2-F-Phe	46	1.2
4-I-Phe	46	1.2
3-F-Phe	45	1.2
3-Cl-Phe	45	1.2
4-S-(2-NO$_2$Bn)-Phe	45	1.2
3,4-di-Cl-Phe	41	1.4
3,5-di-F-Phe	41	1.4
4-F-Phe	40	1.5
3-NO$_2$-Phe	36	1.8
4-CF$_3$-Phe	34	2.0
4-NO$_2$-Phe	32	2.1
4-SAc-Phe	31	2.2
4-$^+$NH$_3$-Phe	17	4.9
pentafluoro-Phe	16	5.2
His(H$^+$)	8	12.3

^{a}% cis is the percent of the total peptide species containing at least one cis amide bond. $\sum cis$ is the sum of the population of all species containing at least one cis amide bond. $K_{trans/\sum cis}$ is the ratio of the population of the species will all-trans amide bonds divided by the sum of the populations of all species containing at least one cis amide bond.
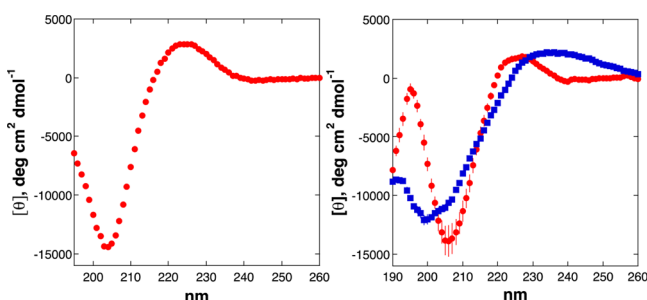

Figure 3. CD spectra of Ac-GPPXPPGY-NH$_2$ peptides. Left: X = pentafluorophenylalanine. Right: X = 4-NH$_2$-phenylalanine (blue squares, pH 7) and X = 4-$^+$NH$_3$-phenylalanine (red circles, pH 3). Peptides for which X is Phe, Tyr, and Trp exhibited no difference in CD spectra at pH 3 vs pH 7 (Figure S1 of the Supporting Information). CD spectra of peptides for which X is one of the 20 canonical amino acids, including PPII-favoring Pro, Leu, and Ala and PPII-disfavoring Thr, Ile, and Val, are shown in ref 23. PPII in this context is indicated by the presence of a positive band at a λ_{max} of ~228 nm, with red-shifted and smaller maxima indicating lower PPII content.[23,45,68−71]

from that of an ideal PPII, and also consistent with a strong aromatic contribution to CD.[23,45,70,72,73] In contrast, at pH 3 (Ar-$^+$NH$_3$ protonation state), where small populations of cis amide bonds were observed, this peptide exhibited a CD signature consistent with a stable polyproline helix (λ_{max} = 229 nm; λ_{min} = 204 nm), similar to that observed for the peptide with pentafluorophenylalanine or peptides with nonaromatic guest residues.[23]

These data indicate that aromatic residues in proline-rich peptides can be used to switch conformational ensembles from species with large populations of cis amide bonds, with an inherently more compact conformation, to species with predominantly trans amide bonds and a relatively more extended polyproline helix conformation. PPII structures are stable in proline-rich sequences despite the absence of hydrogen bonds, because of the presence of n → π^* interactions between adjacent residues and the absence of competing hydrogen-bonded structures.[59,61,74−78] We sought to examine whether alternating block sequences of proline and aromatic residues with two electronically distinct states could be used to develop electronic switches of peptide structure. Peptides were synthesized with one (Ac-PPXPPGY-NH$_2$), two (Ac-PPXPPXPPGY-NH$_2$), or three (Ac-PPXPPXPPXPPGY-NH$_2$ and Ac-PPXPPXPPX-NH$_2$) 4-aminophenylalanine residues embedded within proline-rich sequences. As a secondary structure, polyproline helix exhibits a 3.1 Å rise/residue. Thus, these peptides, if in a polyproline helix throughout their proline−aromatic sequence, would extend 16, 25, and 34 Å, respectively, with all aromatic residues along one PPII-helical face.

These peptides were analyzed by CD and NMR in both protonation states (Figure 4). At pH 7 (Ar-NH$_2$ protonation state), by CD all peptides exhibited broad bands with maxima at ~240 nm, similar to that observed in the Ac-GPPXPPGY-NH$_2$ model peptide containing 4-NH$_2$-Phe. By NMR, high populations of cis amide bond and exceptional conformational hetereogeneity were observed in all peptides with 4-NH$_2$−Phe, consistent with this residue inducing substantial populations of cis amide bond around each aromatic residue via multiple aromatic-proline interactions. In contrast, at pH 3 (Ar-$^+$NH$_3$ protonation state), by CD all peptides exhibited a positive band at 225−230 nm, the loss of the intense band at 240 nm, and a more defined minimum at 205−210 nm, consistent with typical PPII CD spectra. NMR confirmed a substantially reduced population of cis amide bonds in these peptides, despite two, four, and six (for peptides with one, two, and three PPX sequences, respectively) potential cis-proline−aromatic or aromatic−cis-proline interactions in these peptides (pages S25−S27 and S34−S36 and Figures S25−S27 of the Supporting Information).

In the simplified NMR spectra of the Ac-PPXPPGY-NH$_2$ peptides (Figure 4a,b), the equilibrium between the all-trans conformation and conformations with each of the individual cis amide bonds [aromatic−cis-Pro and Pro−cis-Pro−aromatic (Figure 1c)] could be determined. In the amine (Ar-NH$_2$) protonation state, the individual equilibria of aromatic−proline interactions exhibited $K_{trans/cis}$ = 2.5 and 1.6. In contrast, in the ammonium (Ar-$^+$NH$_3$) protonation state, $K_{trans/cis}$ = 8.1 and 6.1 were observed, representing $\Delta\Delta G$ = −0.66 and −0.75 kcal mol^{-1}, respectively, of increased preference for trans amide at each of these prolines in the peptides with an electron-withdrawing ammonium compared to an electron-donating amine. These peptides also exhibited temperature-dependent CD spectra that were consistent with the presence of PPII as the major species in the ammonium protonation state, but non-PPII structures as the major species in the amine protonation state (Figures S8−S10 of the Supporting Information). In total, these data demonstrate that aromatic electronic properties can be used to switch cis−trans isomerism in proline-rich sequences and polyproline helix conformation. Given the range of potentially electronically tunable and conditionally responsive
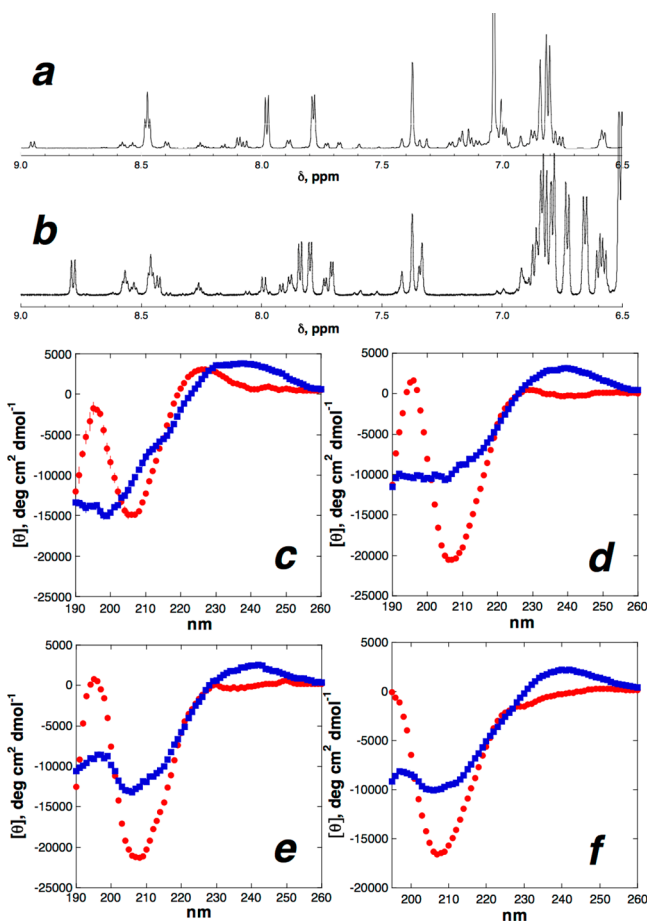
Figure 4. (a and b) NMR spectra of Ac-PPXPPGY-NH$_2$ (X = 4-aminophenylalanine) at (a) pH 3 and (b) pH 7 at 4 °C. In these simplified NMR spectra, four prominent species are observed, including all-*trans* amide bonds, aromatic−*cis*-Pro, Pro−*cis*-Pro−aromatic, and presumably the peptide with both aromatic−*cis*-Pro and Pro−*cis*-Pro−aromatic species. (c−f) CD spectra of Ac-(PPX)$_n$PPGY-NH$_2$ [(c) n = 1; (d) n = 2; (e) n = 3] and (f) Ac-PPXPPPXPPX-NH$_2$ peptides, X = 4-aminophenylalanine, at pH 3 (R-$^+$NH$_3$, red circles) and pH 7 (R-NH$_2$, blue squares). These peptides all exhibit one major species by NMR at pH 3, but a mixture of species at pH 7 (Supporting Information). Additional comparative CD spectra are given in Figures S2−S5 of the Supporting Information. pH-dependent CD spectra for Ac-PPXPPGY-NH$_2$ and Ac-(PPX)$_3$P-PGY-NH$_2$ are shown in Figure S6 of the Supporting Information. Temperature-dependent CD data for Ac-PPXPPGY-NH$_2$ at pH 3 and 7 are shown in Figure S8 of the Supporting Information. Temperature-dependent CD data for Ac-(PPX)$_3$PPGY-NH$_2$ at pH 3 and 7 are shown in Figure S9 of the Supporting Information. Temperature-dependent CD data for Ac-GPPPPPGY-NH$_2$ are given in ref 23.

aromatic amino acids, and the large distances that could be spanned by polyproline helices, including within collagen triple helices, these data suggest the application of aromatic electronics in proline-rich sequences as an intriguing strategy in macromolecular control.

We have previously demonstrated, in model peptides and in peptides derived from the microtubule-binding protein tau, that serine/threonine phosphorylation can induce a reversible conformational switch to PPII.[79,80] The basis of this switch is the modification of a residue with poor PPII propensity (Ser or Thr) to a residue with substantially greater PPII propensity (phosphoserine or phosphothreonine). Herein, we demonstrate the ability to use aromatic electronics to control

polyproline helix conformation.[81,82] Electron-rich aromatic amino acids preceding or following proline residues strongly favor *cis*-proline amide bonds, and thus strongly disfavor polyproline helix, because of multiple favorable proline−aromatic C−H/π interactions. In contrast, in electron-poor aromatic residues, the reduced electron density in the π face (reducing the electrostatic driving force of a C−H/π interaction) significantly reduces the population of *cis*-proline amide bonds, resulting in substantially greater polyproline helix.

C−H/π interactions are typically observed when polarized C−H bonds interact with aromatic rings, with a favorable interaction between δ^- of the aromatic ring and δ^+ of the polarized C−H bond. C−H/π interactions are significantly electrostatic in nature, with substantial additional contributions from the hydrophobic effect and from dispersion forces.[30,32,33,36,83−85] The significantly electrostatic nature of C−H/π interactions is consistent with their geometric dependence, their common orientation toward the centroid of the aromatic ring, and the observation of stronger C−H/π interactions with more acidic hydrogens. Interestingly, surveys of the CSD and the PDB have revealed that many C−H/π interactions exhibit distances below the sum of the van der Waals radii (nonbonded H···C distances of <2.90 Å), and also orientations of the C−H bond away from the centroid of the aromatic ring.[15,27,28,30−33] The observation of sub-van der Waals distances is consistent with a potential additional role of stereoelectronic effects in stabilizing C−H/π interactions, with appropriate overlap of the donor π orbitals with the σ* orbital of the C−H bond. These orbital interactions are analogous to stereoelectronic effects observed in C−H···O interactions, with C−H σ* acceptors to an O lone pair donor.[35] Short C···H distances are found in both aromatic−*cis*-proline interactions and Hα-*cis*-proline−aromatic interactions (Figure 5;[86,87] see
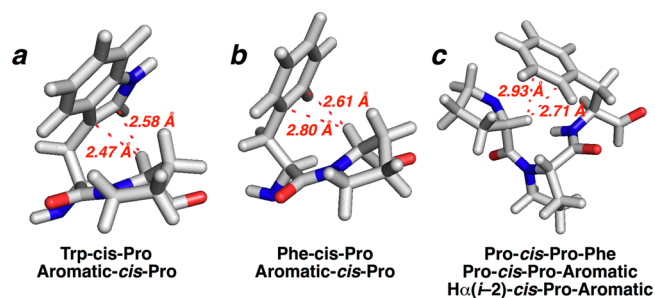
Figure 5. Aromatic−proline interactions with short inter-residue C−H distances. Distances shown are between the indicated Pro Hα and the indicated carbons of the aromatic ring, with distances < 2.90 Å being less than the sum of the van der Waals radii of C and H. Hydrogens were added in Pymol. (a) Residues 191 and 192 of PDB entry 3wfh (sequence WP, 1.90 Å resolution, anti-prostaglandin E2 Fab). (b) Residues 71 and 72 of PDB entry 1h4i (sequence FP, 1.94 Å resolution, *Methylobacterium extorquens* methanol dehydrogenase).[87] (c) Residues 79−81 of PDB entry 1aoc (sequence PPF, 2.00 Å resolution, horseshoe crab coagulogen).[86]

refs 15, 27, 28, 30−33, and 88 for additional examples). Notably, in these structures, the C−H bond is often oriented away from the centroid of the aromatic ring, positioning the C−H σ* orbital near the aromatic π HOMO. These distances and geometries, which are divergent from described idealized C−H/π interaction geometries based on optimized electrostatic interactions, are consistent with a potential role for

stereoelectronic effects in stabilizing aromatic—proline C—H/π interactions.

CONCLUSION

Electron-rich aromatic amino acids strongly disfavor polyproline helix in proline-rich sequences because of *cis*—*trans* isomerism resulting from multiple aromatic—proline interactions. In contrast, electron-poor aromatic amino acids disfavor aromatic—proline interactions and favor polyproline helix. Electron-poor aromatic amino acids provide special capabilities to integrate aromatic residues into polyproline helices and to serve as the basis of aromatic electronic switches to change structure. The use of aromatic electronics to tune polyproline helix conformation could have broad applications in biomolecular design, medicinal chemistry, biomaterials, and engineering.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Peptide synthesis and characterization data, additional temperature-dependent and pH-dependent CD data, full NMR spectra for all peptides, and TOCSY spectra of selected peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: zondlo@udel.edu. Phone: +1 (302) 831-0197. Fax: +1 (302) 831-6335.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Meyer, E. A., Castellano, R. K., and Diederich, F. (2003) Interactions with Aromatic Rings in Chemical and Biological Recognition. *Angew. Chem., Int. Ed. 42*, 1210−1250.

(2) Salonen, L. M., Ellermann, M., and Diederich, F. (2011) Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. *Angew. Chem., Int. Ed. 50*, 4808−4842.

(3) Ma, J. C., and Dougherty, D. A. (1997) The Cation−π Interaction. *Chem. Rev. 97*, 1303−1324.

(4) Hunter, C. A., Lawson, K. R., Perkins, J., and Urch, C. J. (2001) Aromatic interactions. *J. Chem. Soc., Perkin Trans. 2*, 651−669.

(5) Koide, S., and Sidhu, S. S. (2009) The Importance of Being Tyrosine: Lessons in Molecular Recognition from Minimalist Synthetic Binding Proteins. *ACS Chem. Biol. 4*, 325−334.

(6) Grathwohl, C., and Wüthrich, K. (1976) X-Pro Peptide Bond as an NMR probe for conformational studies of flexible linear peptides. *Biopolymers 15*, 2025−2041.

(7) Montelione, G. T., Arnold, E., Meinwald, Y. C., Stimson, E. R., Denton, J. B., Huang, S. G., Clardy, J., and Scheraga, H. A. (1984) Chain-Folding Initiation Structures in Ribonuclease-a: Conformational Analysis of *trans*-Ac-Asn-Pro-Tyr-NHMe and *trans*-Ac-Tyr-Pro-Asn-NHMe in Water and in the Solid-State. *J. Am. Chem. Soc. 106*, 7946−7958.

(8) Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) Folding of Immunogenic Peptide-Fragments of Proteins in Water Solution. 1. Sequence Requirements for the Formation of a Reverse Turn. *J. Mol. Biol. 201*, 161−200.

(9) Stewart, D. E., Sarkar, A., and Wampler, J. E. (1990) Occurrence and role of cis peptide-bonds in protein structures. *J. Mol. Biol. 214*, 253−260.

(10) Poznanski, J., Ejchart, A., Wierzchowski, K. L., and Ciurak, M. (1993) ^{1}H- and ^{13}C-NMR investigations on cis-trans isomerization of proline peptide bonds and conformation of aromatic side chains in H-Trp-(Pro)$_n$-Tyr-OH peptides. *Biopolymers 33*, 781−795.

(11) Yao, J., Feher, V. A., Espejo, B. F., Reymond, M. T., Wright, P. E., and Dyson, H. J. (1994) Stabilization of a Type-VI Turn in a Family of Linear Peptides in Water Solution. *J. Mol. Biol. 243*, 736−753.

(12) Kemmink, J., and Creighton, T. E. (1995) The Physical Properties of Local Interactions of Tyrosine Residues in Peptides and Unfolded Proteins. *J. Mol. Biol. 245*, 251−260.

(13) Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl cis/trans isomerization. *J. Mol. Biol. 279*, 449−460.

(14) Wu, W.-J., and Raleigh, D. P. (1998) Local Control of Peptide Conformation: Stabilization of cis Proline Peptide Bonds by Aromatic Proline Interactions. *Biopolymers 45*, 381−394.

(15) Pal, D., and Chakrabarti, P. (1999) Cis Peptide Bonds in Proteins: Residues Involved, their Conformations, Interactions and Locations. *J. Mol. Biol. 294*, 271−288.

(16) Nardi, F., Kemmink, J., Sattler, M., and Wade, R. C. (2000) The cisproline(i-1)-aromatic(i) interaction: Folding of the Ala-cisPro-Tyr peptide characterized by NMR and theoretical approaches. *J. Biomol. NMR 17*, 63−77.

(17) Halab, L., and Lubell, W. D. (2002) Effect of sequence on peptide geometry in 5-tert-butylprolyl type VI β-turn mimics. *J. Am. Chem. Soc. 124*, 2474−2484.

(18) Taylor, C. M., Hardré, R., Edwards, P. J. B., and Park, J. H. (2003) Factors affecting conformation in proline-containing peptides. *Org. Lett. 5*, 4413−4416.

(19) Ganguly, H. K., Majumder, B., Chattopadhyay, S., Chakrabarti, P., and Basu, G. (2012) Direct Evidence for CH···π Interaction Mediated Stabilization of Pro-cisPro Bond in Peptides with Pro-Pro-Aromatic Motifs. *J. Am. Chem. Soc. 134*, 4661−4669.

(20) Ganguly, H. K., Kaur, H., and Basu, G. (2013) Local Control of cis-Peptidyl—Prolyl Bonds Mediated by CH···π Interactions: The Xaa-Pro-Tyr Motif. *Biochemistry 52*, 6348−6357.

(21) Meng, H. Y., Thomas, K. M., Lee, A. E., and Zondlo, N. J. (2006) Effects of *i* and *i*+3 Residue Identity on Cis-Trans Isomerism of the Aromatic$_{i+1}$-Prolyl$_{i+2}$ Amide Bond: Implications for Type VI β-turn Formation. *Biopolymers 84*, 192−204.

(22) Thomas, K. M., Naduthambi, D., and Zondlo, N. J. (2006) Electronic control of amide cis-trans isomerism via the aromatic-prolyl interaction. *J. Am. Chem. Soc. 128*, 2216−2217.

(23) Brown, A. M., and Zondlo, N. J. (2012) A Propensity Scale for Type II Polyproline Helices (PPII): Aromatic Amino Acids in Proline-Rich Sequences Strongly Disfavor PPII Due to Proline-Aromatic Interactions. *Biochemistry 51*, 5041−5051.

(24) Forbes, C. R., and Zondlo, N. J. (2012) Synthesis of Thiophenylalanine-Containing Peptides via Cu(I)-mediated Cross-Coupling. *Org. Lett. 14*, 464−467.

(25) Zondlo, N. J. (2013) Aromatic-Proline Interactions: Electronically Tunable CH/π Interactions. *Acc. Chem. Res. 46*, 1039−1049.

(26) Greenwood, A. I., Kwon, J., and Nicholson, L. K. (2014) Isomerase-Catalyzed Binding of Interleukin-1 Receptor-Associated Kinase 1 to the EVH1 Domain of Vasodilator-Stimulated Phosphoprotein. *Biochemistry 53*, 3593−3607.

(27) Brandl, M., Weiss, M. S., Jabs, A., Sühnel, J., and Hilgenfeld, R. (2001) C-H···π-Interactions in Proteins. *J. Mol. Biol. 307*, 357−377.

(28) Bhattacharyya, R., and Chakrabarti, P. (2003) Stereospecific Interactions of Proline Residues in Protein Structures and Complexes. *J. Mol. Biol. 331*, 925−940.

(29) Saggu, M., Levinson, N. M., and Boxer, S. G. (2012) Experimental Quantification of Electrostatics in X−H···π Hydrogen Bonds. *J. Am. Chem. Soc. 134*, 18986−18997.

(30) Umezawa, Y., Tsuboyama, S., Honda, K., Uzawa, J., and Nishio, M. (1998) CH/π interaction in the crystal structure of organic compounds. A database study. *Bull. Chem. Soc. Jpn. 71*, 1207−1213.

(31) Umezawa, Y., Tsuboyama, S., Takahashi, H., Uzawa, J., and Nishio, M. (1999) CH/π interaction in the conformation of peptides. A database study. *Bioorg. Med. Chem. 7*, 2021−2026.

(32) Takahashi, O., Kohno, Y., Iwasaki, S., Saito, K., Iwaoka, M., Tomoda, S., Umezawa, Y., Tsuboyama, S., and Nishio, M. (2001) Hydrogen-bond-like nature of the CH/π interaction as evidenced by crystallographic database analyses and ab initio molecular orbital calculations. *Bull. Chem. Soc. Jpn. 74*, 2421−2430.

(33) Nishio, M., Umezawa, Y., Fantini, J., Weiss, M. S., and Chakrabarti, P. (2014) CH-π hydrogen bonds in biological macro-molecules. *Phys. Chem. Chem. Phys. 16*, 12648−12683.

(34) Kumar, M., and Balaji, P. V. (2014) C-H···π interactions in proteins: Prevalence, pattern of occurrence, residue propensities, location, and contribution to protein stability. *J. Mol. Model. 20*, 2136.

(35) Jones, C. R., Baruah, P. K., Thompson, A. L., Scheiner, S., and Smith, M. D. (2012) Can a C−H···O Interaction Be a Determinant of Conformation? *J. Am. Chem. Soc. 134*, 12064−12071.

(36) Nishio, M. (2004) CH/π hydrogen bonds in crystals. *CrystEngComm 6*, 130−158.

(37) Tatko, C. D., and Waters, M. L. (2004) Comparison of C-H···π and hydrophobic interactions in a β-hairpin peptide: Impact on stability and specificity. *J. Am. Chem. Soc. 126*, 2028−2034.

(38) Rubin, G. M., Yandell, M. D., Wortman, J. R., Miklos, G. L. G., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S. B., Gong, F. C., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J. Y., Li, Z. Y., Hynes, R. O., Jones, S. J. M., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. Q. H., Zhong, F., Zhong, W. Y., Gibbs, R., Venter, J. C., Adams, M. D., and Lewis, S. (2000) Comparative Genomics of the Eukaryotes. *Science 287*, 2204−2215.

(39) Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003) The Structure and Function of Proline Recognition Domains. *Sci. STKE 2003*, re8.

(40) Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function. *Biochemistry 41*, 6573−6582.

(41) Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., and Dunker, A. K. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res. 32*, 1037−1049.

(42) Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2005) Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit. 18*, 343−384.

(43) Elam, W. A., Schrank, T. P., Campagnolo, A. J., and Hilser, V. J. (2013) Evolutionary conservation of the polyproline II conformation surrounding intrinsically disordered phosphorylation sites. *Protein Sci. 22*, 405−417.

(44) Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2000) Amino acid propensities for the collagen triple-helix. *Biochemistry 39*, 14960−14967.

(45) Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls, J. E., and Creamer, T. P. (2003) Host-Guest Scale of Left-Handed Polyproline II Helix Formation. *Proteins 53*, 68−75.

(46) Shi, Z., Chen, K., Liu, Z., Ng, A., Bracken, W. C., and Kallenbach, N. R. (2005) Polyproline II propensities from GGXGG peptides reveal an anticorrelation with β-sheet scales. *Proc. Natl. Acad. Sci. U.S.A. 102*, 17964−17968.

(47) Jha, A. K., Colubri, A., Zaman, M. H., Koide, S., Sosnick, T. R., and Freed, K. F. (2005) Helix, sheet, and polyproline II frequencies and strong nearest neighbor effects in a restricted coil library. *Biochemistry 44*, 9691−9702.

(48) Grdadolnik, J., Mohacek-Grosev, V., Baldwin, R. L., and Avbelj, F. (2011) Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proc. Natl. Acad. Sci. U.S.A. 108*, 1794−1798.

(49) Moradi, M., Babin, V., Sagui, C., and Roland, C. (2011) A Statistical Analysis of the PPII Propensity of Amino Acid Guests in Proline-Rich Peptides. *Biophys. J. 100*, 1083−1093.

(50) Toal, S. E., Verbaro, D. J., and Schweitzer-Stenner, R. (2014) Role of Enthalpy-Entropy Compensation in Determining the Conformational Propensities of Amino Acid Residues in Unfolded Peptides. *J. Phys. Chem. B 118*, 1309−1318.

(51) Schimmel, P. R., and Flory, P. J. (1967) Conformational Energy and Configurational Statistics of Poly-L-Proline. *Proc. Natl. Acad. Sci. U.S.A. 58*, 52−59.

(52) Stryer, L., and Haugland, R. P. (1967) Energy transfer: A spectroscopic ruler. *Proc. Natl. Acad. Sci. U.S.A. 58*, 719−726.

(53) Adzhubei, A. A., and Sternberg, M. J. E. (1993) Left-Handed Polyproline-II Helices Commonly Occur in Globular-Proteins. *J. Mol. Biol. 229*, 472−493.

(54) Crespo, L., Sanclimens, G., Montaner, B., Perez-Tomas, R., Royo, M., Pons, M., Albericio, F., and Giralt, E. (2002) Peptide dendrimers based on polyproline helices. *J. Am. Chem. Soc. 124*, 8876−8883.

(55) Shi, Z. S., Olson, C. A., Rose, G. D., Baldwin, R. L., and Kallenbach, N. R. (2002) Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl. Acad. Sci. U.S.A. 99*, 9190−9195.

(56) Lam, S. L., and Hsu, V. L. (2003) NMR identification of left-handed polyproline type II helices. *Biopolymers 69*, 270−281.

(57) Fillon, Y. A., Anderson, J. P., and Chmielewski, J. (2005) Cell penetrating agents based on a polyproline helix scaffold. *J. Am. Chem. Soc. 127*, 11798−11803.

(58) Sanclimens, G., Crespo, L., Giralt, E., Albericio, F., and Royo, M. (2005) Preparation of de novo globular proteins based on proline dendrimers. *J. Org. Chem. 70*, 6274−6281.

(59) Horng, J. C., and Raines, R. T. (2006) Stereoelectronic effects on polyproline conformation. *Protein Sci. 15*, 74−83.

(60) Sahoo, H., Roccatano, D., Hennig, A., and Nau, W. M. (2007) A 10-Å Spectroscopic Ruler Applied to Short Polyprolines. *J. Am. Chem. Soc. 129*, 9762−9772.

(61) Kumin, M., Sonntag, L. S., and Wennemers, H. (2007) Azidoproline containing helices: Stabilization of the polyproline II structure by a functionalizable group. *J. Am. Chem. Soc. 129*, 466−467.

(62) Bartlett, G. J., Choudhary, A., Raines, R. T., and Woolfson, D. N. (2010) n → π* interactions in proteins. *Nat. Chem. Biol. 6*, 615−620.

(63) Adzhubei, A. A., Sternberg, M. J. E., and Makarov, A. A. (2013) Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol. 425*, 2100−2132.

(64) Cockroft, S. L., Hunter, C. A., Lawson, K. R., Perkins, J., and Urch, C. J. (2005) Electrostatic Control of Aromatic Stacking Interactions. *J. Am. Chem. Soc. 127*, 8594−8595.

(65) Lummis, S. C. R., Beene, D. L., Lee, L. W., Lester, H. A., Broadhurst, R. W., and Dougherty, D. A. (2005) Cis-trans isomer-ization at a proline opens the pore of a neurotransmitter-gated channel. *Nature 438*, 248−252.

(66) Limapichat, W., Lester, H. A., and Dougherty, D. A. (2010) Chemical Scale Studies of the Phe-Pro Conserved Motif in the Cys Loop of Cys Loop Receptors. *J. Biol. Chem. 285*, 8976−8984.

(67) Vuister, G. W., and Bax, A. (1993) Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond $J_{HNH\alpha}$ Coupling Constants in ^{15}N-enriched Proteins. *J. Am. Chem. Soc. 115*, 7772−7777.

(68) Sreerama, N., and Woody, R. W. (1994) Poly(Pro)II Helices in Globular Proteins: Identification and Circular Dichroic Anaysis. *Biochemistry 33*, 10022−10025.

(69) Bochicchio, B., and Tamburro, A. M. (2002) Polyproline II structure in proteins: Identification by chiroptical spectroscopies, stability, and functions. *Chirality 14*, 782−792.

(70) Woody, R. W. (2009) Circular Dichroism Spectrum of Peptides in the Poly(Pro)II Conformation. *J. Am. Chem. Soc. 131*, 8234−8245.

(71) Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A. F., and Creamer, T. P. (2001) Host-guest study of left-handed polyproline II helix formation. *Biochemistry 40*, 14376−14383.

(72) Woody, R. W. (1994) Contributions of Tryptophan Side-chains to the Far-Ultraviolet Circular-Dichroism of Proteins. *Eur. Biophys. J. 23*, 253−262.

(73) Sreerama, N., Manning, M. C., Powers, M. E., Zhang, J. X., Goldenberg, D. P., and Woody, R. W. (1999) Tyrosine, phenylalanine, and disulfide contributions to the circular dichroism of proteins: Circular dichroism spectra of wild-type and mutant bovine pancreatic trypsin inhibitor. *Biochemistry 38*, 10814−10822.

(74) Hinderaker, M. P., and Raines, R. T. (2003) An electronic effect on protein structure. *Protein Sci. 12*, 1188−1194.

(75) Hodges, J. A., and Raines, R. T. (2006) Energetics of an n → π^* interaction that impacts protein structure. *Org. Lett. 8*, 4695−4697.

(76) Gorres, K. L., and Raines, R. T. (2010) Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol. 45*, 106−124.

(77) Chiang, Y. C., Lin, Y. J., and Horng, J. C. (2009) Stereoelectronic effects on the transition barrier of polyproline conformational interconversion. *Protein Sci. 18*, 1967−1977.

(78) Kuemin, M., Nagel, Y. A., Schweizer, S., Monnard, F. W., Ochsenfeld, C., and Wennemers, H. (2010) Tuning the cis/trans Conformer Ratio of Xaa-Pro Amide Bonds by Intramolecular Hydrogen Bonds: The Effect on PPII Helix Stability. *Angew. Chem., Int. Ed. 49*, 6324−6327.

(79) Bielska, A. A., and Zondlo, N. J. (2006) Hyperphosphorylation of tau induces local polyproline II helix. *Biochemistry 45*, 5527−5537.

(80) Brister, M. A., Pandey, A. K., Bielska, A. A., and Zondlo, N. J. (2014) OGlcNAcylation and Phosphorylation Have Opposing Structural Effects in tau: Phosphothreonine Induces Particular Conformational Order. *J. Am. Chem. Soc. 136*, 3803−3816.

(81) Umashankara, M., Babu, I. R., and Ganesh, K. N. (2003) Two prolines with a difference: Contrasting stereoelectronic effects of 4R/S-aminoproline on triplex stability in collagen peptides [Pro(X)-Pro(Y)-Gly]$_n$. *Chem. Commun.*, 2606−2607.

(82) Sonar, M. V., and Ganesh, K. N. (2010) Water-Induced Switching of β-Structure to Polyproline II Conformation in the 4S-Aminoproline Polypeptide via H-Bond Rearrangement. *Org. Lett. 12*, 5390−5393.

(83) Nishio, M., Umezawa, Y., Honda, K., Tsuboyama, S., and Suezawa, H. (2009) CH/π hydrogen bonds in organic and organometallic chemistry. *CrystEngComm 11*, 1757−1788.

(84) Tsuzuki, S., and Fujii, A. (2008) Nature and physical origin of CH/π interaction: Significant difference from conventional hydrogen bonds. *Phys. Chem. Chem. Phys. 10*, 2584−2594.

(85) Plevin, M. J., Bryce, D. L., and Boisbouvier, J. (2010) Direct detection of CH/π interactions in proteins. *Nat. Chem. 2*, 466−471.

(86) Bergner, A., Oganessyan, V., Muta, T., Iwanaga, S., Typke, D., Huber, R., and Bode, W. (1996) Crystal structure of a coagulogen, the clotting protein from horseshoe crab: A structural homologue of nerve growth factor. *EMBO J. 15*, 6789−6797.

(87) Ghosh, M., Anthony, C., Harlos, K., Goodwin, M. G., and Blake, C. (1995) The refined structure of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens* at 1.94 Å. *Structure 3*, 177−187.

(88) Joshi, K. B., and Singh, P. (2014) L-Proline induced self-assembly of indolicidin derived palindromic tripeptide. *Tetrahedron Lett. 55*, 3534−3537.