BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Local Control of *cis*-Peptidyl−Prolyl Bonds Mediated by CH···π Interactions: The Xaa-Pro-Tyr Motif

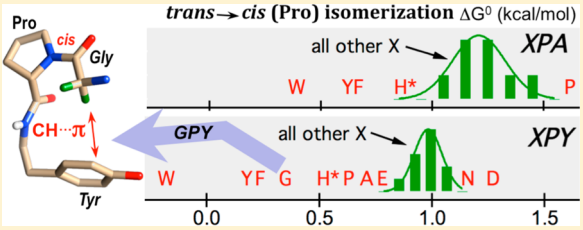
Himal K. Ganguly,[†] Hundeep Kaur,[‡] and Gautam Basu[†,*]

[†]Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata 700054, India

[‡]National Institute of Pharmaceutical Education and Research, 4, Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, India

Supporting Information

ABSTRACT: Compared to generic peptide bonds, the peptidyl−prolyl bond shows a strong propensity for the *cis* conformer. The presence of a sequence-contiguous aromatic (Aro) residue can further stabilize the *cis* conformer, as observed for the Aro-Pro motif. The *cis* propensity of the reverse sequence motif, Pro-Aro, is not so well understood, especially the effect of N-capping the Pro-Aro motif with different amino acid residues. From a comparative nuclear magnetic resonance study of two peptide series with the general sequences Ac-Xaa-Pro-Tyr-NH_2 and Ac-Xaa-Pro-Ala-NH_2, we present a relative thermodynamic scale that reflects how the nature of the Xaa side chain influences the *cis* propensity of the Xaa-Pro-Tyr motif, with Gly, Pro, and Ala at position Xaa giving the greatest enhancement of the *cis*-peptidyl−prolyl population. We also show that CH···π interaction between Xaa and Tyr is responsible for the enhanced *cis* population. However, the mere presence of the CH···π interaction does not guarantee that the peptidyl−prolyl bond will have a higher *cis* content in Xaa-Pro-Tyr than in Xaa-Pro-Ala. Xaa-dependent intramolecular interactions present in Xaa-*trans*-Pro-Tyr can nullify favorable CH···π interactions in Xaa-*cis*-Pro-Tyr. The relative *cis*-peptidyl−prolyl stabilizing propensities of Xaa (Xaa-Pro-Tyr) in proteins and in our peptide series show strong linear correlation except when Xaa is aromatic. We also explore the Xaa-Pro-Gly-Tyr sequence motif and show that mediated by a Pro-Tyr CH···π interaction, the *cis*-peptidyl−prolyl bond in the motif is stabilized when Xaa is Pro.



The *trans* isomer of a peptide bond is favored over the *cis* isomer because of the absence of unfavorable $C^{\alpha}(i)\cdots C^{\alpha}(i+1)$ steric interactions[1,2] and the presence of propitious $O(i)\cdots O(i+1)$ n → π^* interactions.[3,4] However, for peptidyl−prolyl bonds, unfavorable steric interactions, present in the *cis* and *trans* conformers, become comparable and the *trans* isomer is only marginally stable because of favorable backbone n → π^* interaction.[3] This is reflected in the observation that although ∼0.3% of all peptide bonds in known protein structures are present in the *cis* conformation, for peptidyl−proline bonds, the number increases by a whopping 10-fold (∼5%).[5−8] When an aromatic residue ($\tilde{A}$) precedes Pro, the peptidyl−proline bond exhibits an even stronger propensity to adopt the *cis* conformation (∼10%).[8−10] The *cis*-peptidyl−prolyl conformation is often functionally important and exhibits a higher degree of evolutionary conservation in homologous proteins than the *trans*-peptidyl−prolyl units.[11−16] The peptidyl−prolyl bond can also assume the *cis* conformation in a transient state that is functionally relevant.[17] The unique structural and functional features of *cis*-peptidyl−prolyl bonds emphasize the importance of understanding the underlying causes that shift the balance of the *cis*−*trans* equilibrium.

Both local and global factors can subtly control the *cis* versus *trans* fate of a peptidyl−prolyl bond in a folded protein.[5,17] However, only local factors contribute to the *cis*−*trans* equilibrium in short peptides. Therefore, conformational studies of designed peptides can yield important clues about mechanisms that stabilize the *cis*-peptidyl−prolyl bond, especially in the absence of additional tertiary interactions present in a folded protein. Studies of designed peptides containing the ***XP*** sequence motif (unless stated otherwise, all single letters appearing in motifs mentioned in this work correspond to standard amino acid nomenclature; ***X*** signifies any amino acid, while $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{N}}$, and ***Z*** signify aromatic, nonaromatic, and non-prolyl residues, respectively) have demonstrated that similar to proteins, the *cis* content of ***XP*** is higher than that of ***XZ*** even in short peptides.[9,10,18−21] Also it has been shown that peptides containing the $\tilde{\boldsymbol{A}}$***P*** motif have a higher *cis* content than peptides containing the $\tilde{\boldsymbol{N}}$***P*** motif,[9,18] an observation that is consistent with protein structural analyses. Although the trend of enhancement of *cis* content is similar in proteins and short peptides, the absolute increase is much greater in peptides,[10] highlighting nonlocal folding compulsions operative in a protein.

Short peptides containing the ***XP*** motif have been studied extensively. In an earlier study, we had focused on peptides with the ***PP*** motif and showed that the *cis* content of the ***PP*** motif is remarkably enhanced when it is C-capped with an aromatic

Received: June 25, 2013
Revised: August 13, 2013
Published: August 13, 2013

   dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

residue, leading to the establishment of a ***PPÃ*** motif.[10,19] The mechanism responsible for the enhancement was shown to be CH···π interaction, operative between the N-terminal C^{α}-H group of ***P*** and the π-cloud of ***Ã***. However, it was also shown that the N-terminal ***P*** in the ***PPÃ*** motif was dispensable because the ***APW*** motif also showed enhanced *cis* content. However, in the absence of a more comprehensive study, we could not ascertain if the ***XPÃ*** motif is equally *cis*-enhancing for all types of ***X***.[10,19] Isolated works by other groups have shown that ***APY***, ***GPY***, and ***PPY*** motifs are *cis*-enhancing.[22,23] Another study proposed that the ***XPÃ*** motif is more *cis*-enhancing when ***X*** is hydrophobic, although the conclusions were not based on a set of ***XPÑ*** control peptides.[21]

Here we present a systematic analysis of the effect of N-capping ***PY*** (representing the ***PÃ*** motif) by ***X*** from nuclear magnetic resonance (NMR)-derived *cis*–*trans* populations of the peptidyl–prolyl bond in two peptide series, ***XPY*** and ***XPA***, the latter representing the control peptide series (Tyr replaced with Ala). We demonstrate that CH···π interaction, operative between ***X*** and ***Y***, is responsible for the enhanced *cis* population of the ***XPY*** motif. Interestingly, the enhancement of the *cis* population varies with the nature of ***X***. We also show that when the side chain of ***X*** is capable of interacting with the backbone, even with CH···π interaction, the *cis* content of ***XPY*** may not be higher than that of the ***XPA*** motif. We then explore the occurrence of *cis* and *trans* ***XPÃ*** motifs in proteins and compare them with the peptide data. Finally, we discuss additional sequence motifs that can enhance the *cis* content of a peptide–prolyl bond where an aromatic residue is present in the vicinity of but not contiguous in sequence to Pro.

MATERIALS AND METHODS

Peptide Synthesis, Purification, and Characterization. A total of 40 peptides (20 in the ***XPA*** series and 20 in the ***XPY*** series, with the general formulas Ac-Xaa-Pro-Ala-NH_2 and Ac-Xaa-Pro-Tyr-NH_2, respectively) were used in this work. Of these, two peptides (***PPA*** and ***PPY***) have been synthesized previously.[10,19] All other peptides were synthesized in a stepwise manner using the standard solid phase peptide synthesis protocol using Fmoc chemistry and rink amide MBHA resin (0.39 mmol/g substitution) on an AAPPTEC 90II peptide synthesizer. The Fmoc-protected amino acid derivatives (Novabiochem) were coupled using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), hydroxybenzotriazole (HOBt), and diisopropylethylamine (DIPEA) (used as 5-, 5-, and 10-fold excesses of resin substitution, respectively), and the Fmoc protection of the α-amino group was removed by 20% (v/v) piperidine in *N*,*N*-dimethylformamide (DMF). The N-terminal acetylation of the peptide sequence was achieved with acetic acid, PyBOP, and DIPEA (used as 5-, 5-, and 10-fold excesses of resin substitution, respectively). The peptides were cleaved from the resin, and their side chain-protecting group was removed using a cocktail (85% TFA, 5% water, 5% phenol, 2.5% anisole, and 2.5% triisopropylsilane). TFA was removed by evaporation in a rotary evaporator, and the crude peptide obtained was dissolved in methanol. The peptides were purified using a reverse phase high-performance liquid chromatography system using a Phenomenex C18 column using a H_2O/CH_3OH (0 to 80% over 40 min) linear gradient containing 0.1% TFA. The purified peptides were lyophilized and then characterized by mass and ^{1}H NMR spectrometry (Tables S1–S4 of the Supporting Information).

NMR Spectroscopy. NMR experiments were performed on a Bruker Avance III 500 spectrometer. Unless stated otherwise, all samples were prepared in 20 mM phosphate buffer (pH 7 at 4 °C) containing 10% ^2H_2O and the sodium salt of 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid (internal standard). Standard protocols were used to analyze NMR spectra.[24] Water signals in ^{1}H NMR spectra were suppressed by the standard excitation sculpting procedure. Resonance assignments were achieved by analyzing TOCSY and DQF-COSY experiments, while sequential assignments were determined via ROESY experiments (mixing time of 250 ms). For each peptide (both ***XPA*** and ***XPY*** series), resonance signals corresponding to the *cis* and *trans* isomers were assigned from ROESY experiments: *cis* and *trans* isomers were assigned by the presence of $H^{\alpha}(i)\cdots H^{\alpha}(i+1)$ and $H^{\alpha}(i)\cdots H^{\delta}(i+1)$ cross-peaks in the ***XP*** moiety, respectively. The $^3J_{\alpha\beta}$ coupling constants were measured from one-dimensional ^{1}H NMR spectra. Relative populations of the isomers were estimated from the relative integrals of appropriate well-resolved resonances (amide, acetyl, side chain methyl, or aromatic protons). For temperature-dependent NMR studies (van't Hoff analysis), we equilibrated peptides for at least 45 min at a given temperature prior to recording the NMR spectra.

Database Analysis. A representative list of nonredundant protein crystallographic structures from the Protein Data Bank[25] was assembled using the culling server PISCES[26] with a sequence identity of ≤25% and a resolution of ≤2 Å. The database contained 4148 unique chains (December 30, 2012, release). After the database had been cured (ignoring multiple occurring collagen GPP and PPPP motifs, and considering only the first model of multiple models), a total of 36876 Pro residues were identified. Dihedral angles (ω) for all Pro residues were calculated (for *cis*, $-90° \le \omega \le 90°$). Appropriate propensities were calculated using the general formula

$$P^{X}_{cis} = \frac{N^{X}_{cis}(N^{all}_{cis} + N^{all}_{trans})}{N^{all}_{cis}(N^{X}_{cis} + N^{X}_{trans})} \quad (1)$$

where *P*, *N*, and X stand for the propensity, the number of occurrences, and a particular kind of amino acid (***X***) and all amino acids (20 types), respectively. *Z* values were estimated as described previously.[19]

RESULTS

cis*–*trans* Isomerization in *XPA* and *XPY* Peptides.** The central aim of this study is to probe the tandem effect of two amino acid residues, ***X (any residue) and ***Y*** (Tyr), on the stability of a *cis*-peptidyl–prolyl conformation in the ***XPY*** motif. As part of this goal, a series of peptides, with the general formula ***XPY***, was synthesized to systematically study the effect of varying the nature of ***X*** on the *cis*–*trans* equilibrium of the peptidyl–prolyl bond in ***XPY***. The relative stability mentioned here is with respect to an identical peptide where Tyr is replaced with a nonaromatic amino acid Ala. The control peptide series, with the general formula ***XPA***, was also synthesized. Subsequently, the *cis* contents of both ***XPY*** and ***XPA*** series of peptides were measured by integrating appropriate (*cis* or *trans*) resonances in the ^{1}H NMR spectra in aqueous buffer (Figure 1a,b).

The percent *cis* populations of the ***XPA*** peptide series are listed in Table 1. ***XPA*** exhibits the highest *cis* content when ***X*** is aromatic (33% for W, 24% for Y, and 23% for F). The ***X*** dependence of the *cis* content of the ***XP*** motif, without an

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348–6357

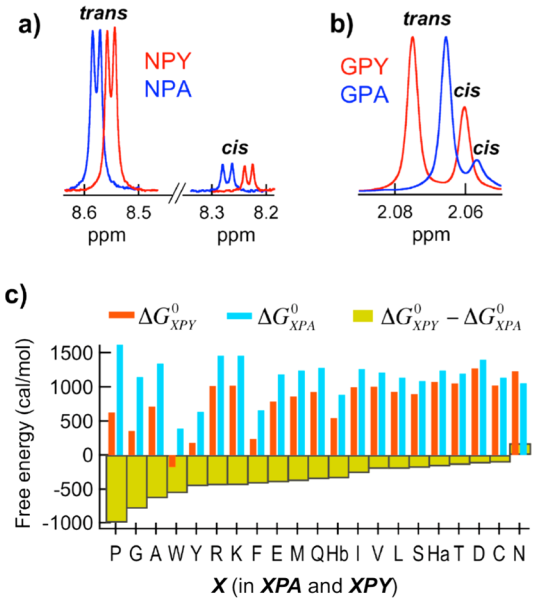


Figure 1. (a) Amide resonances of Asn in ***NPY*** and ***NPA*** and (b) acetyl resonances in ***GPY*** and ***GPA*** for the *cis*- and *trans*-peptidyl–prolyl conformations in aqueous buffer (pH 7.0) at 4 °C. (c) Standard free energies associated with *cis* → *trans* isomerization of the peptidyl–prolyl bond in ***XPA*** (ΔG^0_{XPA}) and ***XPY*** (ΔG^0_{XPY}) as a function of ***X*** (standard single-letter codes represent ***X***; Ha and Hb stand for His at pH 3.5 and 8.5, respectively). The difference ($\Delta G^0_{XPY} - \Delta G^0_{XPA}$) is also plotted for comparison.

Table 1. Percent *cis* Contents and Free Energies (*trans* → *cis*) of *XPA* and *XPY*

	% *cis* content				
X	***XPY***	***XPA***	ΔG^0_{XPY} (kcal/mol)	ΔG^0_{XPA} (kcal/mol)	$\Delta\Delta G^0$ (kcal/mol)
A	21.6	8.0	0.71	1.34	−0.63
G	34.5	11.1	0.35	1.15	−0.79
V	13.9	10.0	1.00	1.21	−0.21
I	14.1	9.2	0.99	1.26	−0.27
L	15.6	11.3	0.93	1.13	−0.20
D	9.0	7.3	1.27	1.40	−0.13
E	19.4	10.4	0.78	1.19	−0.40
N	10.9	12.6	1.16	1.07	0.09
Q	15.7	8.9	0.93	1.28	−0.36
K	13.6	6.6	1.02	1.46	−0.44
R	13.7	6.6	1.01	1.46	−0.45
M	17.4	9.5	0.86	1.24	−0.38
W	57.9	33.1	−0.18	0.39	−0.56
Y	42.0	24.0	0.18	0.63	−0.46
F	39.4	23.2	0.24	0.66	−0.42
P	24.3	5.0	0.63	1.62	−0.99
T	12.9	10.2	1.05	1.20	−0.15
S	16.4	12.2	0.90	1.09	−0.19
C	13.5	11.3	1.02	1.13	−0.11
Ha[a]	12.5	9.5	1.07	1.24	−0.17
Hb[a]	27.2	16.7	0.54	0.88	−0.34

[a] ***Ha*** and ***Hb*** stand for data at pH 3.5 and 8.5, respectively.

aromatic residue following it, has been reported for another designed peptide series with the general sequence ***AXPAK***, capped with an acetyl and an amide at the N- and C-termini, respectively.[9] In comparison to ***AXPAK***, our control peptide ***XPA*** lacks ***A*** at the N-terminus and ***K*** at the C-terminus. The *cis* contents of the ***XPA*** and ***AXPAK*** peptides are almost identical (Figure S1 of the Supporting Information), suggesting that the terminal residues (A and K) do not influence the *cis*/*trans* equilibrium of ***XPA***.

The percent *cis* populations of the ***XPY*** peptide series are also listed in Table 1. Similar to ***XPA***, when ***X*** is aromatic, ***XPY*** exhibits the highest percent *cis* content [58% for W, 42% for Y, 40% for F, and 27% for H (pH 8.5)], but there are others for which the percent *cis* content is greater than 20% (35% for G, 24% for P, and 22% for A). Interestingly, the percent *cis* content of the ***WP*** bond in ***WPY*** is more than 50%, making the *cis* isomer marginally more stable than the *trans* counterpart.

Comparison of the *cis* contents of ***XPY*** and ***XPA*** series of peptides showed that the replacement of Ala with Tyr increased the *cis* content of ***XP*** for almost all types of ***X*** (Table 1). The only exception is Asn, for which the *cis* content actually slightly decreased. A marked increase (>10%) in ***XP*** *cis* content was observed when ***X*** was Gly, Pro, Ala, Tyr, Phe, His (pH 8.5), or Trp. The effect of replacing Ala with Tyr was also analyzed by calculating the difference in free energy differences ($\Delta\Delta G^0_{A\to Y}$) for the *trans* → *cis* isomerization process of the ***XP*** peptide bond in ***XPY*** (ΔG^0_{XPY}) and ***XPA*** (ΔG^0_{XPA}) series of peptides. Along with percent *cis* contents, $\Delta\Delta G^0_{A\to Y}$ ($\Delta G^0_{XPY} - \Delta G^0_{XPA}$) values are also listed in Table 1 and plotted in Figure 1c. A negative value of $\Delta\Delta G^0_{A\to Y}$ reflects a favorable *cis* isomer upon the Ala to Tyr substitution and vice versa, while the magnitudes of $\Delta\Delta G^0_{A\to Y}$ signify the extent of favor or disfavor. Negative values for $\Delta\Delta G^0_{A\to Y}$ were observed for all amino acids except for Asn. The effect of the Ala to Tyr substitution was most pronounced in peptides for which ***X*** was Pro, Gly, or Ala ($\Delta\Delta G^0_{A\to Y}$ = −0.99, − 0.79, or −0.63 kcal/mol, respectively), while Asn at position ***X*** yielded a positive value ($\Delta\Delta G^0_{A\to Y}$ = 0.09 kcal/mol). Amino acids with hydrophobic side chains (Leu, Ile, and Val) and alcoholic groups (Thr and Ser) showed low values (−0.24 to −0.15 kcal/mol) of $\Delta\Delta G^0_{A\to Y}$, while for the remaining peptides, $\Delta\Delta G^0_{A\to Y}$ values were between −0.57 (Arg) and −0.37 (Glu) kcal/mol.

Upfield-Shifted C^α-H Resonances of *X* in *XPY* and *XPA*. In an earlier report, we showed that interaction between the C^α-H of ***P***(i) and the π-cloud of $\tilde{\mathbf{A}}(i+2)$ in ***PP$\tilde{\mathbf{A}}$*** peptide series was responsible for the enhanced stability (compared to that of ***PPA***) of the *cis*-peptidyl–prolyl bond in the ***PP$\tilde{\mathbf{A}}$*** motif.[10,19] The CH···π interaction resulted in an upfield shift of the C^α-H of ***P***(i). Because ***XPY*** showed an enhanced *cis* content irrespective of the nature of the amino acid present at position ***X***, the increment must arise because of the presence of Tyr, and like in ***PP$\tilde{\mathbf{A}}$***, any ***X***···***Y*** interaction is expected to leave its signature on the chemical shifts of ***X***. We computed $\Delta\delta$, the C^α-H chemical shift (NMR) differences of residue ***X***, between the *cis* and *trans* conformations (*cis*P − *trans*P) in ***XPA*** and ***XPY*** peptide series. As shown in Figure 2, members of the ***XPA*** peptide series generally show a small upfield shift (−0.15 ± 0.07 ppm) with the exception of G, Y, F, and P, which exhibit shifts greater than −0.2 ppm. On the other hand, the ***XPY*** peptide series shows a large upfield shift (−0.64 ± 0.25 ppm). More interestingly, unlike the case with ***XPA*** peptide series, the $\Delta\delta$ values for the ***XPY*** peptide series show a good linear correlation when plotted against ΔG^0_{XPY}. This suggests that a common mechanism must be responsible for stabilizing the *cis*-peptidyl–prolyl bond in ***XPY*** and causing the upfield-shifted $\Delta\delta$.

To focus exclusively on the effect of Tyr, $\Delta\delta$ values for ***XPA*** were subtracted from $\Delta\delta$ values for ***XPY*** to yield $\Delta\Delta\delta_{A\to Y}$

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

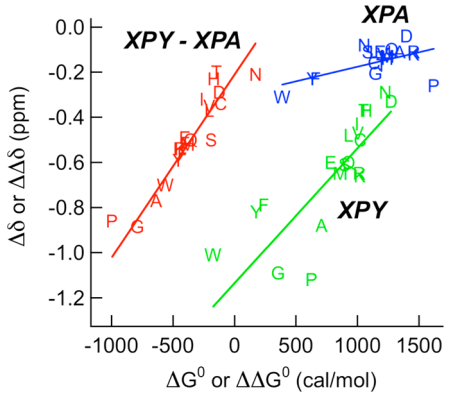


Figure 2. Correlations among the C^{α}-H chemical shift differences of ***X*** in ***XPA*** ($\Delta\delta_{XPA} = \delta^{cis}_{XPA} - \delta^{trans}_{XPA}$) vs ΔG^0_{XPA} (blue), the C^{α}-H chemical shift differences of ***X*** in ***XPY*** ($\Delta\delta_{XPY} = \delta^{cis}_{XPY} - \delta^{trans}_{XPY}$) vs ΔG^0_{XPY} (green), and differences in chemical shift differences of ***XPA*** and ***XPY*** ($\Delta\Delta\delta_{XPY} = \Delta\delta_{XPY} - \Delta\delta_{XPA}$) vs the differences in the corresponding free energies [$\Delta\Delta G^0 = \Delta G^0_{XPY} - \Delta G^0_{XPA}$ (red)]. The solid lines represent the best linear fits, and single letters signify the nature of ***X*** in ***XPA*** or ***XPY***.

values. The $\Delta\Delta\delta_{A\rightarrow Y}$ values exhibited an excellent linear correlation with the corresponding $\Delta\Delta G^0_{A\rightarrow Y}$ values (Figure 2). The correlation of the observed ring current effect ($\Delta\Delta\delta_{A\rightarrow Y}$) and enhanced *cis* content ($\Delta\Delta G^0_{A\rightarrow Y}$) indicates that C^{α}-H···Tyr(*i*, *i* + 2) interaction is dominantly responsible for the observed change in the free energy of the *cis*–*trans* isomerization. Because the proton accepting capacity of the phenolic ring of Tyr is constant throughout the ***XPY*** peptide series, the differential upfield shift of the C^{α}-H of ***X*** must originate from its proton donating capacity, which would in turn depend on its partial positive charge as well as its geometric availability to the Tyr ring.

ROE Cross-Peaks between C^{α}-H (*X*) and Aromatic Protons of *Y* in *XpY*. CH···π interaction between ***X*** and ***Y*** in conformational microstate ***XpY*** (from here on *p* stands for *cis*Pro and *P* stands for *trans*Pro) would bring the C^{α}-H of ***X*** and aromatic protons of ***Y*** into the proximity of each other, giving rise to nuclear Overhauser effect (NOE) cross-peaks in two-dimensional NMR spectra. The appearance of such cross-peaks would be direct evidence of the CH···π interaction. In a previous paper, we had reported ROE cross-peaks between the C^{α}-H of the N-terminal ***P*** and $\tilde{A}$ ring protons in peptide ***Pp$\tilde{A}$***.[10] As shown in Figure 3, ROE cross-peaks, between aromatic protons of ***Y*** (C^{δ}-H and C^{ε}-H) and the C^{α}-H of ***X***, were observed for the *cis* conformations of two peptides in the ***XPY*** peptide series, ***ApY*** and ***GpY***. No such ROE peaks were observed for the corresponding *trans* conformations, ***APY*** and ***GPY***.

Restriction of the Tyr χ^1 Angle in *XpY*. A consequence of CH···π interaction in ***XpY*** is the restriction of the side chain dihedral angle χ^1 of Tyr. Nondegenerate $^3J_{\alpha\beta}$ coupling constants (Table 2) along with nondegenerate β_A and β_B proton chemical shifts of the Tyr residue, in the *cis* isomers of all ***XPY*** peptides, indicate a restriction of its χ^1 dihedral angle. The χ^1 angle of Tyr can assume three canonical values, $g^{(-)}$ (−60°), $g^{(+)}$ (60°), and *t* (180°), but when it is involved in CH···π interaction with ***X*** in ***XpY***, as was seen for peptide ***PpY***,[10] only two states [*t* and $g^{(-)}$] are accessible. This was concluded from one large (>9 Hz) and one small (<5 Hz) $^3J_{\alpha\beta}$ coupling constant of ***Y***, a trend observed for all ***XpY*** conformers (Table 2). Complementary data from differential ROE intensities of the two C^{β}-H resonances of Tyr (in ***XpY***) with N-H and C^{α}-H [α-β_A, strong; α-β_B, weak; N-β_A, weak or absent; N-β_B, strong (Table S5 of

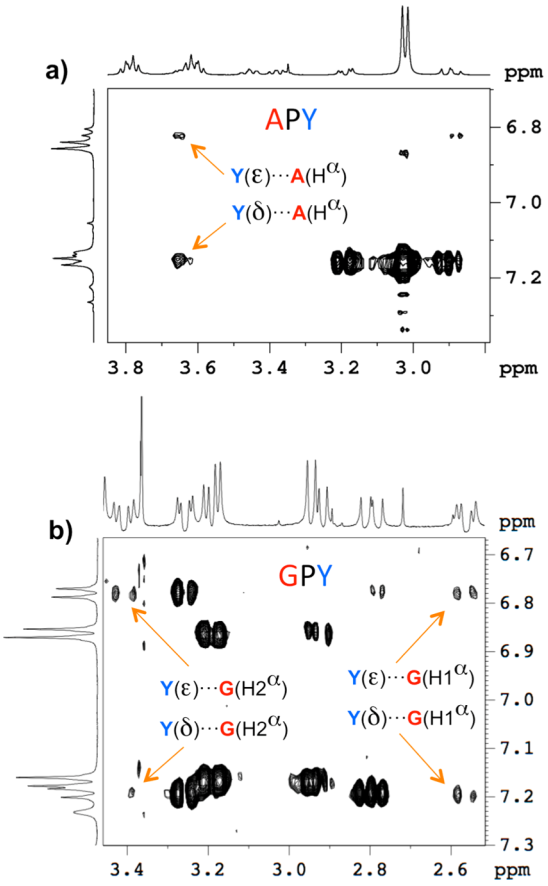


Figure 3. ROE cross-peaks between aromatic protons (C^{δ}-H and C^{ε}-H) of Tyr and (a) C^{α}-H of Ala (in ***ApY***) and (b) C^{α}-H (two sets) of Gly (in ***GpY***).

Table 2. $^3J_{\alpha\beta}$ Values and Percent Rotameric (χ^1) Populations of *Y* in *XPY* Peptide Series

	cis			*trans*		
X	$^3J_{\alpha\beta}$ (Hz)	*t*	$g^{(-)}$	$^3J_{\alpha\beta}$ (Hz)	*t*	$g^{(-)}$
A	4.7/11.8	19	84	7.6[a]	–	–
G	4.1/12.2	14	88	6.2/9.9	33	66
V	6.1/10.0	32	66	7.5/7.7	45	46
I	6.1/10.0	32	70	7.8/8.3	48	52
L	5.7/10.5	29	72	8.1[a]	–	–
D	5.8/9.8	29	66	5.7/10.6	28	73
E	5.7/10.5	29	72	7.5[a]	–	–
N	6.8/9.0	39	60	6.0/10.1	31	69
Q	5.6/10.6	27	74	7.7[a]	–	–
K	5.10/ov[b]	23	77	ov[b]	–	–
R	5.35/11.0	25	77	7.65[a]	–	–
M	5.4/10.8	26	75	7.60[a]	–	–
F	5.3/11.2	25	79	ov[b]	–	–
W	ov[b] /11.7	27	83	ov[b] /9.2	–	60
Y	5.3/ov[b]	25	75	ov[b]	–	–
T	ov[b] /8.4	34	66	7.1/8.3	42	52
S	ov/10.3	29	71	6.5/9.6	36	66

[a]Chemical shifts of both β-protons are identical. [b]Spectral overlap with other protons.

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

the Supporting Information)] further suggest that the dominant rotamer is $g^{(-)}$ ($\chi^1 = -60°$) in the *cis* conformer.

Equilibrium between Microstates (with and without X···Y interactions) in *Xp*. The entire population of ***XpY*** may or may not be involved in CH···π interaction. Instead, two microstates, one with ($\boldsymbol{XpY}^+$) and one without ($\boldsymbol{XpY}^-$) CH···π interaction, will be in equilibrium with each other with the equilibrium constant $K^+ = [\boldsymbol{XpY}^+]/[\boldsymbol{XpY}^-]$. The percent of time ***XpY*** occupies the $\boldsymbol{XpY}^+$ state can be estimated from K^+ following the work of Kemmink and Creighton.[22,23] In this method, K_{XPY}, the *trans* to *cis* equilibrium constant of ***XPY***, is expanded in terms $[\boldsymbol{XpY}^+]$ and $[\boldsymbol{XpY}^-]$, followed by the assumption that K_{XPA}, the *trans* to *cis* equilibrium constant of ***XPA*** ($K_{XPA} = [\boldsymbol{XpA}]/[\boldsymbol{XPA}]$), is equal to $[\boldsymbol{XpY}^-]/[\boldsymbol{XPY}]$. This yields a relationship among K_{XPA}, K_{XPY}, and K^+ of

$$K_{XPY} = \frac{[\boldsymbol{XpY}^-] + [\boldsymbol{XpY}^+]}{[\boldsymbol{XPY}]} = \frac{1 + [\boldsymbol{XpY}^+]/[\boldsymbol{XpY}^-]}{[\boldsymbol{XPY}]/[\boldsymbol{XpY}^-]} \approx \frac{1 + K^+}{1/K_{XPA}} \quad (2)$$

Because K_{XPA} and K_{XPY} can be estimated from the percent *cis* populations of ***XPA*** and ***XPY***, respectively (Table 1), it is straightforward to calculate K^+ (and the corresponding percent Creighton occupancy of the $\boldsymbol{XpY}^+$ state) from eq 2. As shown in Figure 4, the Creighton occupancy values of the $\boldsymbol{XpY}^+$ state

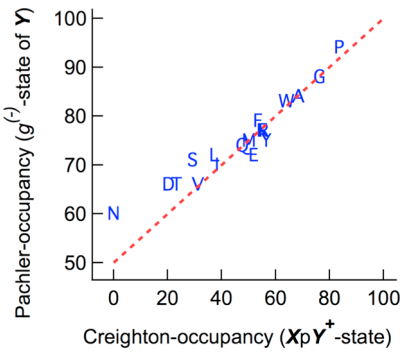


Figure 4. Correlation between Pachler occupancies [population of the $g^{(-)}$ rotamer] and Creighton occupancies (population of the $\boldsymbol{XpY}^+$ conformer) for the *cis*Pro conformers in ***XPY*** peptides. The dotted line represents a 1:1 correlation [assuming only two allowed rotameric states of ***Y*** in ***XpY***, $g^{(-)}$ and *t*].

show a wide variation (44 ± 21), with the K^+ of Ala, Gly, and Pro being the highest and that of Asn being the lowest (for Asn, K_{NPA} was larger than K_{NPY}, yielding unrealistic negative values; a Creighton occupancy of zero was assumed for $\boldsymbol{NpY}^+$).

Using a method suggested by Pachler,[27] microstates associated with the ***XpY*** conformation can also be probed using a complementary technique, where $^3J_{\alpha\beta}$ coupling constants of Tyr can be used to extract relative populations of all χ^1 rotameric states. $^3J_{\alpha\beta}$ analyses showed that only two rotameric states, $g^{(-)}$ and *t*, are possible for Tyr in ***XpY***, while analyses of relative ROE intensities showed $g^{(-)}$ to be the dominant state. As shown in Figure 4, the $g^{(-)}$ state is populated more than the *t* state for all members of the ***XPY*** series. The mean Pachler occupancy of the $g^{(-)}$ state is 75 ± 8%, with ***NpY*** having the lowest values (60%) and ***PpY*** the highest (94%).

If the $g^{(-)}$ state is indeed associated with the CH···π interaction-stabilized *cis* form of ***XPY*** ($\boldsymbol{XpY}^+$), then the Pachler occupancy of the $g^{(-)}$ state must be correlated with the Creighton occupancy of the $\boldsymbol{XpY}^+$ state. Figure 4 shows a plot of Pachler occupancies [percent of ***XpY*** associated with the Tyr $g^{(-)}$ state] versus Creigton occupancies (percent of ***XpY*** present in the $\boldsymbol{XpY}^+$ state) for all members of the ***XPY*** series (***X*** = ***C*** and ***H*** are not shown because the $^3J_{\alpha\beta}$ values could not be determined because of spectral overlap). Creigton occupancies of the $\boldsymbol{XpY}^+$ state vary from zero (no CH···π interaction) to 100 (exclusive CH···π interaction), while Pachler occupancies of the $g^{(-)}$ state vary from 50 [50% of the population in the $g^{(-)}$ state indicates no bias for the $g^{(-)}$ state] to 100 [entire population present in the $g^{(-)}$ state]. Thus, if the $g^{(-)}$ state of ***XpY*** is indeed associated with the CH···π interaction, the Pachler−Creigton data should lie on a straight line described by the relationship Pachler occupancy = Creigton occupancy/2 + 50 (dotted line in Figure 4). Indeed, this is what is observed, suggesting that the $g^{(-)}$ state of Tyr in $\boldsymbol{XpY}^+$ is associated with CH···π interaction.

Three peptides (***DPY***, ***TPY***, and ***SPY***) showed slight deviations from the dotted line in Figure 4, while ***NPY*** showed a significant deviation from the dotted line in Figure 4. This may happen if the *cis* isomers of ***XPA*** for these residues are stabilized due to some unique nature of residue ***X***, which is absent in ***XPY***. Alternatively, this can happen if the *trans* isomer of ***XPY*** is characterized by some unique interactions that are absent in *trans* isomer ***XPA***. Both would make the assumption $[\boldsymbol{XpA}]/[\boldsymbol{XPA}] \approx [\boldsymbol{XpY}^-]/[\boldsymbol{XPY}]$, leading to eq 2 being invalid. Scheraga and co-workers reported that Ac-***NPY***-NHMe is involved in an intramolecular H-bond in which the terminal NHMe is the donor.[28] For ***NPY***, we observed a ROE cross-peak between NH (terminal NH_2) and C^β-H of Asn. In addition, the temperature dependence of the terminal NH_2 group was low ($\Delta\delta/\Delta T = -3.8$ ppb/deg, as opposed to −6 ppb/deg in ***NPA***). This suggests that the side chain C═O group of Asn is involved in a H-bond with the C-terminal amide in ***NPY***. Such a H-bond is absent in ***NPA***. Neither ***NpA*** nor ***NpY*** showed any detectable hydrogen bonding signatures. Similar results were observed for ***DPY*** (terminal NH_2 $\Delta\delta/\Delta T = -3.8$ and −6.0 ppb/deg for ***DPY*** and ***DPA***, respectively). Interestingly, ***NPY*** and ***DPY*** show a dominant (~70%) χ^1 value for Tyr in the *cis* and *trans* states (Table 2). For ***SPY*** and ***TPY***, temperature-dependent terminal NH_2 chemical shift coefficients ($\Delta\delta/\Delta T = -4.5$ and −6.0 ppb/deg for ***SPY*** and ***SPA***, respectively, and −5.2 and −7.0 ppb/deg for ***TPY*** and ***TPA***, respectively) also suggest that the C-terminal NH_2 could be H-bonded in both the peptides, and only in the *trans* isomers. However, no ROE cross-peaks could be detected that would point toward a particular H-bond acceptor. It is quite likely that the side chain hydroxyl oxygen atoms in Thr and Ser act as the acceptor in these two peptides when the peptidyl−prolyl bond assumes the *trans* conformation.

Energetics of *Y*(*i* + 2)···*X*(*i*) Interaction in *XpY*. To probe the energetics of CH···π interaction in ***XpY*** conformers, we estimated the enthalpic ($\Delta H°$) and entropic ($\Delta S°$) components associated with the *trans* → *cis* equilibria in ***XPY*** and ***XPA*** peptide series from a van't Hoff analysis of temperature-dependent NMR data. The $\Delta H°$ and $\Delta S°$ values are listed in Table S6 of the Supporting Information. For the ***XPA*** series, $\Delta H°$ was positive for all peptides with a mean value of 0.92 ± 0.45 kcal/mol. This is consistent with the notion that devoid of any other interactions, especially by any aromatic group, the peptidyl−prolyl *trans* → *cis* isomerization process is associated with an unfavorable enthalpy.[29] In contrast to this, except for

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

Ile and Leu (ΔH° = 0.61 and 0.76 kcal/mol, respectively), the ΔH° values for all members of the ***XPY*** peptide series were negative or near zero, with a mean of −0.54 ± 0.88 kcal/mol.

Significant differences were also observed in the ΔS° values, between the ***XPA*** and ***XPY*** peptide series (Table 3). While the ***XPA*** peptides showed a slightly negative value of ΔS° (−0.8 ± 1.2 cal/mol/deg), the ***XPY*** peptides were associated with significantly negative values of ΔS° (−4.9 ± 2.2 cal/mol/deg). This is expected because at least three (ψ_X, ψ_P, and ϕ_Y) backbone torsion angles and one (χ^1_Y) side chain torsion angle of the *cis* isomers in ***XPY*** peptides are expected to be restricted for the peptides to exhibit CH···π interaction.

Table 3. Propensities of Xaa-*cis*Pro-Aro Motifs in the Protein Data Bank

Xaa	Ec[a]	Nc[a]	Nt[a]	propensity	*Z* value
Ala	25.62	36	304	1.4	2.06
Gly	19.67	50	211	2.54	6.85
Val	23.36	11	299	0.47	−2.56
Ile	21.25	6	276	0.28	−3.32
Leu	29.47	16	375	0.54	−2.49
Asp	18.16	11	230	0.61	−1.68
Glu	17.63	36	198	2.04	4.38
Asn	16.35	10	207	0.61	−1.57
Gln	12.81	11	159	0.86	−0.51
Lys	20.5	11	261	0.54	−2.1
Arg	15.68	11	197	0.7	−1.18
Met	5.58	2	72	0.36	−1.52
Pro	11.98	32	127	2.67	5.79
Thr	22.38	13	284	0.58	−1.99
Ser	19.44	17	241	0.87	−0.56
Trp	3.77	5	45	1.33	0.63
Tyr	13.79	22	161	1.6	2.21
Phe	11.08	16	131	1.44	1.48
Cys	5.28	2	68	0.38	−1.43
His	9.19	5	117	0.54	−1.38

[a]Ec, Nc, and Nt represent the expected occurrence of *cis* and the observed occurrence of *cis* and *trans* conformers, respectively.

The ΔH° (van't Hoff) and ΔG° (from integration of NMR peaks) values are linearly correlated (Figure 5) for both ***XPA*** (correlation coefficient of 0.56, with ***RPA***, ***TPA***, and ***EPA*** being outliers) and ***XPY*** (correlation coefficient of 0.80, with ***GPY***, ***IPY***, and ***LPY*** being outliers) peptide series, suggesting that enthalpy, whether originating from CH···π-type interaction (as in ***XPY***) or an other kind (as in ***XPA***), plays a role in stabilizing or destabilizing the *cis* conformer. To focus exclusively on the "CH···π component" of enthalpy in the ***XPY*** series, ΔH° values for the ***XPA*** series were subtracted from the ΔH° values for the ***XPY*** series to yield $\Delta\Delta H^\circ$ values. The $\Delta\Delta H^\circ$ values were all negative (−1.46 ± 0.95 kcal/mol) and exhibited a strong linear correlation (Figure 5; correlation coefficient of 0.84) with the corresponding $\Delta\Delta G^\circ$ values.

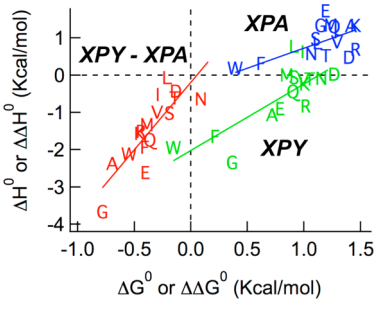


Figure 5. Correlations between standard enthalpy and standard free energy changes associated with *cis* → *trans* isomerization of the peptidyl−prolyl bond in ***XPA*** peptides [ΔH^0_{XPA} vs ΔG^0_{XPA} (blue)] and ***XPY*** peptides [ΔH^0_{XPY} vs ΔG^0_{XPY} (green)] and their differences [$\Delta\Delta H^0 = \Delta H^0_{XPY} - \Delta H^0_{XPA}$ vs $\Delta\Delta G^0 = \Delta G^0_{XPY} - \Delta G^0_{XPA}$ (red)]. The solid lines represent the best linear fits, and single letters signify the nature of ***X*** in ***XPA*** or ***XPY***.

XpÃ* Motifs in the Protein Data Bank.** Analysis of a nonredundant set of known protein structures showed that ∼4.7% of ***XP sequences are present in the *cis* conformation. The percent *cis* content increased to 7.6% when only ***XPÃ*** sequence motifs were considered, indicating that ***PÃ*** is a sequence motif that enhances the *cis* content of the peptidyl−prolyl bond. We also calculated the propensity of all 20 types of ***X*** to assume the ***XpÃ*** conformation. As shown in Table 3, ***XpÃ*** is over-represented for ***X*** = Ala, Gly, Glu, and Pro. On the other hand, Val, Ile, and Thr are under-represented. The propensity values for these seven amino acids are associated with a high level of confidence (a |*Z* scorel of >1.95 indicates a >95% level of confidence). This correlates well with the experimentally observed *cis* populations of ***XPY*** peptides, where Ala, Gly, and Pro at the ***X*** position showed considerable enhancement compared with their ***XPA*** counterparts, while Val, Ile, and Thr showed the least enhancement. The only exception is Glu. However, for Glu, even without an aromatic residue succeeding it, the *cis* propensity for Glu-Pro was reported to be high (1.4) in proteins.[10] The incompatibility between the high protein structure-derived *cis* propensity and the marginal enhancement of *cis* populations of ***EPY*** can be understood by considering the sequence and structure of four residues that follow ***Ep*** motifs in proteins. In a majority of cases, it was observed that the *cis*Pro is present at the N-terminus of a helix, where the pre-Pro-Glu side chain forms a H-bond or exhibits side chain−side chain interaction with residues (Arg, Ser, and Thr) at position *i* + 3 or *i* + 4. Our designed peptides were too short to include this effect.

Percent *cis* populations of the ***XP*** motif (with and without an aromatic residue present at its C-terminus) are plotted in Figure 6, where the abscissa and the ordinate refer to percent *cis* populations in proteins (***XPÃ*** and ***XPÑ*** motifs) and peptides (***XPY*** and ***XPA*** motifs), respectively. For the ***XPA***/***XPÑ*** motif, the percent *cis* populations are restricted to a small range, both for peptides and for proteins, for all ***X*** except when it is aromatic. The spread in the percent *cis* populations is much larger for the ***XPY***/***XPÃ*** motif with the percent *cis* populations (for nonaromatic ***X***) exhibiting a linear correlation. The linear correlation clearly indicates that, just as was observed in short peptides, the ***X***(C^α-H)···***Ã***(π) interaction also plays an important role in the *cis*−*trans* isomerization of the peptidyl−prolyl unit in protein ***XPÃ*** motifs. Interestingly, compared to the ordinate = abscissa line, the linear fit to the ***XPY***/***XPÃ*** data (excluding ***X*** = ***Ã***) is shifted upward by ∼10% along the ordinate (***XPY*** peptides), suggesting that when present in a protein, depending on the exact local sequence, the π-electrons of ***Ã*** in ***ÑPÃ*** might participate in some other local interaction, thus being unavailable for stabilizing the *cis*-peptidyl−prolyl conformation. When ***X*** itself is aromatic (***Ã*$_1$*P*$\tilde{A}_2$**), the *cis* populations in peptides are much higher than that in proteins, and not consistent with the linear correlation observed for ***ÑPY***/***ÑPÃ*** data. Because two CH···π interactions can become

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

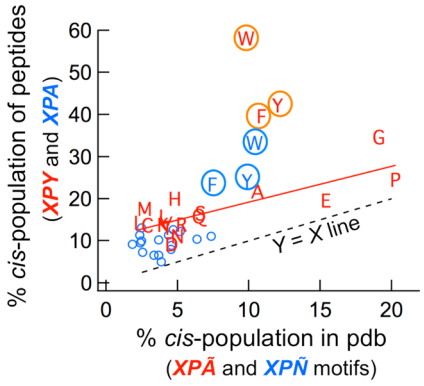


Figure 6. Correlation between the percent *cis* population of the peptidyl−prolyl bond in ***XPY*** (red) and ***XPA*** (blue) peptide series and the corresponding percent *cis* population of the peptidyl−prolyl bond in ***XPÃ*** and ***XPÑ*** motifs in known protein structures (***Ã*** for aromatic and ***Ñ*** for nonaromatic residues). The solid red line represents the best linear fit to the ***XPY***/***XPÃ*** data, and the dashed line represents a 1:1 correlation; single letters signify the nature of ***X*** in ***XPY***/***XPÃ*** data (all) and ***XPA***/***XPÑ*** data (only when ***X*** is aromatic). Blue circles represent ***XPA***/***XPÑ*** data (only when ***X*** is nonaromatic).

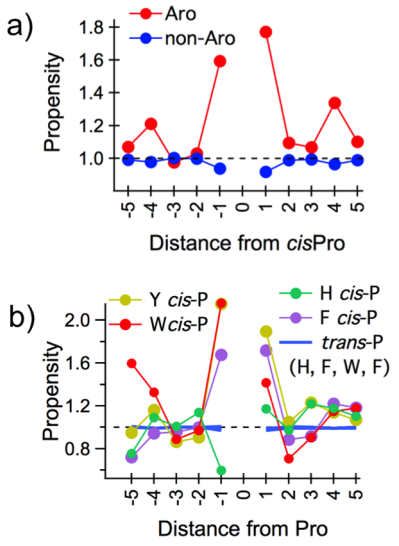


Figure 7. (a) Propensities of aromatic (red) and nonaromatic (blue) residues to appear within five residues of a *cis*Pro in known protein structures. (b) Propensities of aromatic residues to appear within five residues of a *cis*Pro [***W***, ***Y***, and ***F*** (color annotation in the figure)] and *trans*Pro [***W***, ***Y***, ***F***, and ***H*** (blue)] in known protein structures.

operative in $\tilde{A}_1P\tilde{A}_2$ motifs, $\tilde{A}_1(\pi)\cdots P(C^{\alpha}\text{-H})$ and $\tilde{A}_1(C^{\alpha}\text{-H})\cdots\tilde{A}_2(\pi)$, this suggests that while both might be operative in ***ÃPY*** peptide series, only one is commonly found in protein $\tilde{A}_1P\tilde{A}_2$ motifs.

Hint about More Sequence Motifs That May Stabilize Xaa-*cis*Pro: The XPGY Motif. Motifs that are known to stabilize the *cis*-peptidyl−prolyl bond often contain an aromatic residue, contiguous in sequence to *cis*Pro. In addition to the classic ***ÃP*** motif, this study established that the ***PÃ*** motif is another important motif. What happens if the ***Ã*** residue, unlike in ***ÃP*** or ***PÃ***, is noncontiguous to the *cis*Pro but not very far in the sequence? Because the Pro−*cis*Pro bond is substantially stabilized in ***PPÃ*** (compared to that in ***PPA***), we decided to study a related peptide pair, Ac-Pro-Pro-Gly-Tyr-NH_2 (***PPGY***) and Ac-Pro-Pro-Gly-Ala-NH_2 (***PPGA***, control peptide), and compare the Pro-*cis*Pro populations in both.

Note that in the new peptide pair, the aromatic residue (***Y***) is not contiguous to the nonterminal Pro residue and is separated by one Gly residue at its C-terminus. The measured *cis* populations were 13.1 and 8.9% for ***PPGY*** and ***PPGA***, respectively. The observed *cis* population of ***PPGA*** is higher than that of ***PPA*** (5%). This could be due to the effect of differential C-capping of ***PP*** by ***A*** and ***G***. On the other hand, the higher *cis* population of ***PPGY***, compared to that of ***PPGA***, suggests that Tyr at position $i + 2$ might play a role in enhancing the *cis* population of ***PP***. The C^{α}-H chemical shift differences ($\delta_{cis} - \delta_{trans}$) in ***PPGY*** and ***PPGA*** (−0.14 and 0.51 ppm, respectively) showed that the chemical shift of C^{α}-H of the N-terminal Pro in ***PPGY*** is significantly more upfield-shifted than that in ***PPGA***. Although preliminary, the results demonstrate that the presence of an aromatic residue, noncontiguous to a Pro residue (position $i + 2$), can enhance the population of the *cis*-peptidyl−prolyl bond and the mechanism of enhancement is Pro−Aro interaction (indicated by the upfield-shifted C^{α}-H of the N-terminal Pro).

We analyzed the propensities of aromatic and nonaromatic residues to appear around *cis*Pro (±5 residues) in proteins with known structures (Figure 7). Nonaromatic residues show no preference for appearing five residues upstream or downstream of *cis*Pro. In fact, there is a slight negative preference (propensity of <1) for nonaromatic residues to appear at *cis*Pro ± 1 positions. In contrast, aromatic residues are strongly over-represented at position ±1 and consistently over-represented in the C-terminal stretch of *cis*Pro (positions 3, 4, and 5). The strong over-representation at position ±1 is not surprising and reaffirms that ***ÃP*** and ***PÃ*** are motifs that favor *cis*-peptidyl−prolyl bonds. The over-representation in the C-terminal stretch probably indicates that the presence of aromatic amino acids at these positions might play a role in stabilizing *cis*-peptidyl−prolyl bonds. However, with a sparse data set and idiosyncrasy of individual proteins, no general conclusion could be drawn. This point needs a more careful analysis with a much larger data set.

DISCUSSION

A peptide bond can assume two conformations: the energetically favorable *trans* conformation and the less favorable *cis* conformation.[1,2] This is reflected in the *cis*−*trans* equilibrium in short peptides in solution.[9,10,18−23] The *cis*−*trans* equilibrium becomes unavailable to folded proteins that pack into crystals, and as a result, the peptide bonds in protein crystal structures are forced to adopt either the *cis* or *trans* conformation, but not an admixture of both, like in short peptides. Given this strict binary criterion, it is expected that all peptide bonds in proteins are *trans* because the *trans* population is almost always greater than 50% in short peptides in solution. However, *cis* peptide bonds do occur in proteins[5−8] and are often associated with functional significance.[11−17] It is therefore important to identify sequence motifs that stabilize *cis* peptide bonds and establish unique mechanisms behind their stabilization.

The cyclic nature of Pro side chain makes the *cis*- and *trans*-peptidyl−prolyl bonds energetically comparable. As a result, ***XP*** motifs are almost 100 times more susceptible to adopting the *cis* conformation than the corresponding non-prolyl units, in both proteins and short peptides. The presence of a sequence-contiguous aromatic residue can provide additional stability to the peptidyl−prolyl *cis* conformation. For example, when an aromatic residue precedes Pro, as in the ***ÃP*** motif, the

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

peptidyl−prolyl *cis* conformation is 5−10-fold more probable than when $\tilde{A}$ is replaced by a nonaromatic residue, in both proteins[8−10] and systematic work on short peptides.[9,10,18−23] The suggested mechanism behind this is a weak CH···π interaction, operative between the C^{α}-H of Pro and π-electrons of $\tilde{A}$.[8,18] CH···π interaction is known to exert a strong influence on protein structure[30] and function.[31] Weak interactions like CH···π interactions and, very recently, Si-H···π interaction have been used to tune the band structures of graphene and silicene, respectively,[32] emphasizing the importance of understanding these interactions in greater detail.

An aromatic residue can influence the fate of a peptidyl−-prolyl bond not only when it precedes Pro but also when it follows Pro, or when present in its vicinity.[5,10,19,22,23] Although isolated studies have been reported, unlike the case of the ***ÃP*** motif, there is no systematic experimental study of how the presence of an aromatic residue succeeding a Pro residue (***PÃ***) could influence the *cis* content of a peptidyl−prolyl bond. For example, the ***PÃ*** motif was studied first by Kemmink and Creighton.[22,23] Comparing the populations of NMR-detected conformations of peptides containing the ***XPY*** and ***XPA*** motifs (in ***GAPYTGA***, ***GAPATGA***, and ***EAPY***), they showed that the *cis* content of the peptidyl−prolyl bond in ***XPY*** is enhanced compared to that in ***XPA***.[22,23] However, the study was not comprehensive because only three peptides were studied (***X*** = Ala, Pro, and Gly). In a recent study, we also showed that ***PPÃ*** and ***APW*** exhibit enhanced *cis* content compared to those of control peptides with ***A*** in place of ***Ã***/***W***.[10] In another study, performed on peptides with the ***XPÃ*** motif (***SXPYDV***), Dyson and co-workers[21] suggested that hydrophobic residues at position ***X*** are marginally better at promoting the formation of *cis*-peptidyl−prolyl bonds than hydrophilic ones. However, they did not compare their results with a control peptide (without the ***Ã*** residue). Therefore, although published results suggest that an aromatic residue following Pro enhances the *cis* content of the peptidyl−prolyl bond, to fully understand the ***XPÃ*** motif, a systematic study of peptides with the general formula ***XPÃ*** and ***XPA*** (control) is needed.

The goal of this study was to probe the effect of the presence of an aromatic residue succeeding a peptidyl−prolyl bond in a systematic manner. This led us to compare the *cis* content of the ***XP*** motif in two series of peptides, ***XPY*** and ***XPA***, by systematically varying ***X*** in each. Effectively, this allowed us to quantify the effect of C-capping the ***XP*** motif with an aromatic residue, Tyr. Compared to ***XP*** peptides C-capped by Ala, the Tyr C-capped peptides showed an enhanced *cis* population for almost all variations of ***X***, with Pro, Ala, and Gly exerting the strongest effect. The mechanism behind the enhancement is CH···π interaction and was confirmed from the proximity of the C^{α}-H proton of ***X*** and the ring atoms of ***Y*** or the *cis* conformers in ***GPY*** and ***APY***. This is consistent with similar results for ***PPY***, ***PPF***, and ***PPW***, obtained from NOE cross-peaks.[10] Direct evidence of CH···π interactions has been reported previously for a different system.[33] Although no such direct interaction was observed for other members, a linear correlation between the upfield chemical shifts of the C^{α}-H proton of ***X*** in the *cis* conformer and the free energy of the *trans* → *cis* transition is consistent with a CH···π interaction-stabilized *cis*-peptidyl−prolyl bond for almost all amino acids present at the ***X*** position.

Interestingly, the *trans* → *cis* transition was associated with a negative enthalpy change for all types of ***X*** in ***XPY*** (the corresponding entropy change being negative or unfavorable) and was linearly correlated with the corresponding free energy values. This also strongly indicates that the additional *cis* conformation stabilizing effect, when ***XP*** is C-capped by Tyr, is enthalpic in nature and consistent with the presence of CH···π interaction. The enthalpic gain (negative ΔH°) associated with *trans* → *cis* isomerization in ***XPY*** peptide series comes at an entropic price (negative ΔS°). The strong linear correlation between the ΔH° and ΔS° (correlation coefficient of 0.94) shows the importance of enthalpy−entropy compensation for this isomerization equilibrium.

To test if the ***XPA/XPY*** results are consistent with ***XPA/XPÃ*** peptides, we synthesized and studied three additional peptides, one with ***W*** as the aromatic group (***APW***) and two with ***F*** as the aromatic group (***GPF*** and ***NPF***). Like ***APY***, ***APW*** exhibits an enhanced *cis* population (21.6%) compared to that of ***APA*** (8%).[10] The enhanced *cis* content is correlated with a large upfield C^{α}-H chemical shift of Ala (4.39 and 3.30 ppm in *trans* and *cis* conformers, respectively). Similarly, ***GPF*** also exhibited an enhanced *cis* population (23.4%) compared to that of ***GPA*** (11.1%). The enhancement was accompanied by a large upfield C^{α}-H chemical shift of Gly (4.09 and 3.43/2.60 ppm in *trans* and *cis* conformers, respectively). The upfield chemical shifts (−1.09 ppm for ***APW*** and −0.66/−1.49 ppm for ***GPF***) are compatible with the average upfield shift observed in ***XPY*** peptide series (−0.64 ± 0.25 ppm). Similar to ***NPY***, ***NPF*** showed a reduced *cis* content (9.4%) compared to that of ***NPA*** (12.6%), and like ***NPY***, this could be attributed to the presence of a side chain−backbone H-bond, present in only the *trans* state [ROE cross-peak between the NH (terminal NH_2) and C^{β}-H of Asn].

Our results establish ***XPÃ*** to be a sequence motif that stabilizes the *cis*-peptidyl−prolyl bond, especially when ***X*** is Pro, Ala, or Gly. In this context, it should be pointed out that as a pre-*cis*Pro residue, Gly also exhibits other special characteristics. For example, while Li^{2+} is known to stabilize the *cis*Pro conformation of ***XP*** units in short peptides, Gly-Pro is inert to Li^{2+}.[34] Gly-*cis*Pro is over-represented in β-sheets, occurring as a united residue and with functional importance.[35] In ∼40% of cases, the β-(Gly-*cis*Pro) motif is followed by an aromatic residue.

It was interesting to note that some amino acids, like Asn, Asp, Ser, and Thr, when present at the ***X*** position of ***XPY***, despite exhibiting the CH···π interaction in the ***XpY*** conformation, did not induce stabilization of ***XpY*** (vs ***XpA***), as expected from the population of the $g^{(-)}$ state of Y in ***XpY***. Upon further investigation, it was found that the presence of unique side chain−backbone interactions, present in only the ***XPY*** state, counters the *cis*-stabilizing CH···π interactions of the ***XpY*** state in these peptides. Thus, although CH···π interactions do play an important role in stabilizing *cis*-peptidyl−prolyl bonds, the importance of other side chain−backbone interactions cannot be ignored, as it is the net effect that controls the *cis*−*trans* equilibrium of the peptidyl−prolyl bond.

We also explored the effect of placing an aromatic amino acid (Tyr) one residue from Pro in a peptide (***PPGY***) and compared its *cis* content with that of a control peptide without an aromatic side chain (***PPGA***). Compared to ***PPGA***, ***PPGY*** exhibited an enhanced *cis* content, and the Pro1 C^{α}-H proton resonance (NMR) in ***PPGY*** was upfield-shifted. Therefore, even when Tyr was noncontiguous to *cis*Pro, its presence did enhance the *cis* content of ***PP***, and the enhancement correlated with the proximity (only in the *cis* conformation) of ***P*** (N-terminal Pro of ***PP***) and the Tyr side chain. The ***PPGY/PPGA***

dx.doi.org/10.1021/bi4007918 | *Biochemistry* 2013, 52, 6348−6357

result is also noteworthy because a related sequence motif, ***NPXY***, is biologically important. The ***NPGY*** motif occurs seven times in a span of 40 residues in the region of residues 53−94 of chicken prion protein.[36] Two ***NPXY*** motifs in the intracellular domain of LRP1 regulate the function of the receptor. We investigated the *cis* populations of peptides ***NPGY*** and ***NPGA*** and found the *cis* population of ***NPGY*** (13.0%) to be higher than that of ***NPGA*** (9.6%). However, in terms of the upfield chemical shift of the *cis*Pro C^{α}-H proton resonance, the two peptides were not much different ($\Delta\Delta\delta \sim 0.02$ ppm). The slightly higher *cis* population of ***NPGY*** may be due to specific interactions of the Asn side chain. Results for these two isolated peptide pairs indicate the role of aromatic residues in influencing the *cis*−*trans* equilibrium of a peptidyl−prolyl bond even when it is not an immediate neighbor of Pro. However, without a detailed systematic study, general conclusions cannot be drawn about motifs that stabilize the *cis*-peptidyl−prolyl bonds that contain an aromatic residue noncontiguous to Pro.

The experimental data presented in this work are based on very short peptides, merely three residues long. A pertinent question is how relevant the data are to protein structures. Compelled by local interactions, the motifs will not only adopt a particular conformation in a folded protein but also be guided by global interactions. A comparison of the percent *cis* population of ***XPÑ*** fragments in proteins and the percent *cis* population of ***XPA*** peptides was not very insightful because the range of variation is limited in both. However, what is clear is that the percent *cis* population of ***XPA*** is always greater in peptides, especially for aromatic residues (Figure 6). The range of percent *cis* populations not only was much larger for ***XPÃ*** (in proteins) and ***XPY*** peptides but also exhibited a linear correlation (Figure 6). The linear fit runs almost parallel to the $Y = X$ axis, the former being vertically shifted up. Thus, what is true for three-residue peptides also holds for proteins: ***XpÃ*** fragments in proteins appear predominantly due to local interactions. The slightly higher percent *cis* population for ***XPY*** peptides, versus that in proteins, probably indicates the presence of a constant *cis* destabilizing global effect in proteins, irrespective of the nature of ***X***. Interestingly, *cis* populations of ***ÃPY*** or ***ÃPA*** peptides are consistently much higher than the corresponding *cis* populations of ***ÃPÃ*** or ***ÃPÑ*** motifs in proteins. It is possible that the pre-Pro ***Ã*** residue in ***ÃPA***, which in short peptides strongly interacts with ***P*** and stabilizes the *cis*Pro conformation,[18] becomes much less available in proteins because of the presence of alternate short- or long-range interactions. The same applies to ***ÃPÃ*** motifs. In addition, while both ***Ã*−*Ã*** and ***Ã*−*P*** interactions can occur alternately or simultaneously in short peptides, probably only one of the two is allowed in proteins.

In summary, this work establishes ***XPY***, and by analogy ***XPÃ***, as a sequence motif that stabilizes the *cis*-peptidyl−prolyl conformation mediated by CH···π interaction. Zondlo and co-workers have elegantly demonstrated how modulation of the π-electron density of an aromatic residue (acceptor) influences the peptidyl−prolyl *cis*−*trans* isomerization equilibrium in a linear fashion.[37,38] Because the π-electron rich acceptor is constant (Tyr) in the peptide series we have studied, in an analogous fashion, this work can be viewed as a reverse study, in which the influence of modulating the donor strength (acidity of C^{α}-H) on the peptidyl−prolyl *cis*−*trans* isomerization equilibrium is monitored, in the presence of a constant acceptor. We showed that the presence of an aromatic residue (Tyr) also modulates, although to a much lesser degree, the *cis* population of a noncontiguous peptidyl−prolyl group. Clearly, more sequence motifs, containing an aromatic residue and a noncontiguous peptidyl−prolyl group, need to be studied in a systematic way. The thermodynamic stability scale, presented in this work, for ***XpY*** (and ***XpA***), will be helpful in understanding conformational preferences of sequence motifs containing Pro.[39] We also showed that unique side chain−backbone H-bonds in the *trans* state of ***XPY*** can influence the enhancement (compared to ***XPA***) of the *cis* content of ***XPY***. Finally, the strong correlation between peptide NMR and protein structural database studies reaffirms that *cis*-peptidyl−prolyl bonds appear dominantly because of local interactions.

■ ASSOCIATED CONTENT

Supporting Information

Six tables (S1−S6) and one figure (S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VIIM, Kolkata 700054, India. Telephone: +91-33-2569-3215. Fax: +91-33-23553886. Email: gautamda@gmail.com or gautam@boseinst.ernet.in.

Funding

This work was supported by funds from the Department of Science and Technology (India) to G.B. (SR/SO/BB-11/2008) and the Council of Scientific and Industrial Research (India) to H.K.G.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Tanmoy Debnath and Barun Majumder for their help with peptide synthesis and NMR spectroscopy, respectively.

■ ABBREVIATIONS

X, generic amino acid; ***Ã***, aromatic residue; ***Ñ***, nonaromatic residue; ***Z***, non-proline residue; *p*, *cis*Pro; *P*, *trans*Pro.

■ REFERENCES

(1) Schulz, G. D., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.

(2) Creighton, T. E. (1993) *Proteins: Structure and Molecular Properties*, 2nd ed., W. H. Freeman and Co., New York.

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

(4) Choudhary, A., Gandla, D., Krow, G. R., and Raines, R. T. (2009) Nature of Amide Carbonyl−Carbonyl Interactions in Proteins. *J. Am. Chem. Soc. 131*, 7244−7246.

(5) Wathen, B., and Jia, Z. (2008) Local and Nonlocal Environments around Cis Peptides. *J. Proteome Res. 7*, 145−153.

(6) MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol. 218*, 397−412.

(7) 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.

(8) Pal, D., and Chakrabarti, P. (1999) Cis peptide bonds in proteins: Residues involved, their conformations, interactions and locations. *J. Mol. Biol. 294*, 271−288.

(9) 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.
(10) 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.
(11) Brauer, A. B., Domingo, G. J., Cooke, R. M., Matthews, S. J., and Leatherbarrow, R. J. (2002) A conserved cis peptide bond is necessary for the activity of Bowman-Birk inhibitor protein. *Biochemistry 41*, 10608−10615.
(12) Birolo, L., Malashkevich, V. N., Capitani, G., De Luca, F., Moretta, A., Jansonius, J. N., and Marino, G. (1999) Functional and structural analysis of cis-proline mutants of *Escherichia coli* aspartate aminotransferase. *Biochemistry 38*, 905−913.
(13) Mallis, R. J., Brazin, K. N., Fulton, D. B., and Andreotti, A. H. (2002) Structural characterization of a proline-driven conformational switch within the Itk SH2 domain. *Nat. Struct. Biol. 9*, 900−905.
(14) Jin, L., Stec, B., and Kantrowitz, E. R. (2000) A cis-proline to alanine mutant of *E. coli* aspartate transcarbamoylase: Kinetic studies and three-dimensional crystal structures. *Biochemistry 39*, 8058−8066.
(15) Lu, K. P., Finn, G., Lee, T. H., and Nicholson, L. K. (2007) Prolyl *cis-trans* isomerization as a molecular timer. *Nat. Chem. Biol. 3*, 619−629.
(16) Lorenzen, S., Peters, B., Goede, A., Preissner, R., and Frommel, C. (2005) Conservation of cis prolyl bonds in proteins during evolution. *Proteins 58*, 589−595.
(17) Andreotti, A. H. (2003) Native state proline isomerization: An intrinsic molecular switch. *Biochemistry 42*, 9515−9524.
(18) 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.
(19) Dasgupta, B., Chakrabarti, P., and Basu, G. (2007) Enhanced stability of *cis*Pro-Pro peptide bond in Pro-Pro-Phe sequence motif. *FEBS Lett. 581*, 4529−4532.
(20) Grathwohl, C., and Wuthrich, K. (1976) The X−Pro peptide bond as an NMR probe for conformational studies of flexible linear peptides. *Biopolymers 15*, 2025−2041.
(21) 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.
(22) Kemmink, J., and Creighton, T. E. (1993) Local conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor and their roles in folding. *J. Mol. Biol. 234*, 861−878.
(23) 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.
(24) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience Publications, New York.
(25) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235−242.
(26) Wang, G., and Dunbrack, R. L., Jr. (2005) PISCES: Recent improvements to a PDB sequence culling server. *Nucleic Acids Res. 33*, 94−98.
(27) Pachler, K. G. R. (1963) Nuclear magnetic resonance study of some α-amino acids-I: Coupling constants in alkaline and acidic medium. *Spectrochim. Acta 19*, 2085−2092.
(28) Stimson, E. R., Meinwald, Y. C., Montelione, G. T., and Scheraga, H. A. (1986) Conformational properties of *trans* Ac-Asn-Pro-Tyr-NHMe and *trans* Ac-Tyr-Pro-Asn-NHMe in dimethylsulfoxide and in water determined by multinuclear n.m.r. spectroscopy. *Int. J. Pept. Protein Res. 27*, 569−582.
(29) Eberhardt, E. S., Loh, S. N., and Raines, R. T. (1993) Thermodynamic origin of prolyl peptide bond isomers. *Tetrahedron Lett. 34*, 3055−3056.
(30) Brandl, M., Weiss, M. S., Jabs, A., Suhnel, J., and Hilgenfeld, R. (2001) C-H···π-interactions in proteins. *J. Mol. Biol. 307*, 357−377.
(31) Harigai, M., Kataoka, M., and Imamoto, Y. (2006) A single CH/π weak hydrogen bond governs stability and the photocycle of the photoactive yellow protein. *J. Am. Chem. Soc. 128*, 10646−10647.
(32) Li, Y., and Chen, Z. (2013) XH/π (X = C, Si) Interactions in Graphene and Silicene: Weak in Strength, Strong in Tuning Band Structures. *J. Phys. Chem. Lett. 4*, 269−275.
(33) 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.
(34) Kunz, C., Jahreis, G., Günther, R., Berger, S., Fischer, G., and Hofmann, H. J. (2012) Influence of lithium cations on prolyl peptide bonds. *J. Pept. Sci. 18*, 400−404.
(35) Das, M., and Basu, G. (2012) Glycine Rescue of β-Sheets from cis-Proline. *J. Am. Chem. Soc. 134*, 16536−16539.
(36) Wopfner, F., Weidenhöfer, G., Schneider, R., von Brunn, A., Gilch, S., Schwarz, T. F., Werner, T., and Schätzl, H. M. (1999) Analysis of 27 mammalian and 9 avian PrPs reveals high conservation of flexible regions of the prion protein. *J. Mol. Biol. 289*, 1163−1178.
(37) 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.
(38) Zondlo, N. J. (2013) Aromatic-proline interactions: Electronically tunable CH/π interactions. *Acc. Chem. Res. 46*, 1039−1049.
(39) Austin, E. W., 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.