Prediction of Xaa-Pro peptide bond conformation from NMR chemical shifts

Yang Shen and Ad Bax

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, U.S.A.

Supplementary Material

Table S1. Amino acid occurrence for residues neighboring Pro residues with cis and trans Xaa-Pro peptide bonds, normalized to their average occurrence in the database

	cis(i-1)	***trans(i-1)***	***cis(i+1)***	***trans(i+1)***
A	0.93	0.93	1.26	1.06
C	0.77	1.19	1.18	0.83
D	0.59	1.09	0.66	1.08
E	1.02	0.76	0.58	1.37
F	1.50	0.92	1.62	0.95
G	1.46	0.74	0.95	1.05
H	0.80	1.22	1.26	0.90
I	0.48	1.15	0.68	0.86
K	1.11	1.07	0.75	1.01
L	0.65	1.10	0.93	0.93
M	0.72	1.04	0.78	0.93
N	1.29	1.19	1.02	0.91
P	1.27	0.89	1.32	0.89
Q	0.99	0.97	0.96	1.08
R	0.82	1.06	0.99	0.95
S	1.21	0.98	0.99	1.00
T	0.83	1.14	0.87	0.90
V	0.60	1.01	0.99	1.03
W	2.11	0.77	1.30	1.02
Y	2.08	0.96	1.75	0.95
c	0.99	1.19	1.76	0.70

Table S2. Average secondary chemical shifts and standard deviations for residues adjacent to Pro residues with cis or trans Xaa-Pro peptide bonds

X	$<\Delta\delta_{cis}^{X,i-1}>$	$\sigma_{cis}^{X,i-1}$	$<\Delta\delta_{trans}^{X,i-1}>$	$\sigma_{trans}^{X,i-1}$	$<\Delta\delta_{cis}^{X,i+1}>$	$\sigma_{cis}^{X,i+1}$	$<\Delta\delta_{trans}^{X,i+1}>$	$\sigma_{trans}^{X,i+1}$
^{15}N	-0.63	4.30	-0.19	4.07	0.67	4.97	-1.28	4.10
$^{1}H^{\alpha}$	-0.06	0.49	0.26	0.42	0.15	0.52	-0.03	0.44
$^{13}C'$	-1.38	1.47	-1.78	1.62	-0.48	1.38	0.08	1.74
$^{13}C^{\alpha}$	-1.84	1.71	-1.76	2.45	0.50	2.11	0.63	2.15
$^{13}C^{\beta}$	1.15	2.10	0.49	1.91	0.97	2.36	0.53	1.81
$^{1}H^{N}$	-0.19	0.82	-0.08	0.74	0.20	0.62	0.09	0.83

Table S3. Statistics of some Promega prediction outliers in Figure 2B

BMRB ID	Xray structure[a]	Resid[b]	isomer[c]	P_{cis}[d]	$\delta^{13}C^{\gamma}$ [ppm]	NMR structure[e]	NMR observation [f]
5496	2ZEQ	75	cis	0.00	27.6	1MG8	Strong Hα-Hδ NOEs; trans Xaa-Pro bond
6114	1T15	18	cis	0.13	26.9	1OQA	Strong Hα-Hδ NOEs; trans Xaa-Pro bond
15332	2O2V	40	trans	0.99	24.3	2JRH	Strong Hα-Hα NOEs; Cis Xaa-Pro bond
6111	1M8N	91	trans	0.99	24.8	1Z2F	Cis Xaa-Pro bond
15796	1FBV	41	trans	0.99	24.3	2K4D	Strong Hα-Hβ NOEs; Cis Xaa-Pro bond

[a] PDB code of the X-ray reference structure
[b] residue number of the center Proline residue
[c] type of the Xaa-Pro peptide bond observed in X-ray structure
[d] P_{cis} score calculated with $^{13}C^{\gamma}$ chemical shift included
[e] PDB code of the NMR reference structure to which the NMR chemical shifts correspond
[f] configuration of the Xaa-Pro peptide bond in the NMR structure or the isomer-specific sequential NOEs

Table S4. Promega predictions of Xaa-Pro peptide bonds for proteins not included in the database

Protein	*PDB/BMRB ID*[a]	*resID*[b]	*isomer*[c]	P_{cis}/P_{cis}^{norm} *(w/o* $\delta^{13}C^{\gamma}$*)*	P_{cis}/P_{cis}^{norm} *(with* $\delta^{13}C^{\gamma}$*)*
MEKK3 PB1	2JRH/15332	23	cis	**0.950/0.500**	0.999/0.986
		40	cis	**0.951/0.506**	1.000/0.999
	2PPH/15355	23	cis	**0.950/0.500**	0.998/0.972
		40	trans	**0.246/0.017**	0.000/0.000
caenopore-5	2JS9/15330	99	cis	**0.932/0.418**	0.997/0.945
	2JSA/15358	99	trans	**0.188/0.012**	No $^{13}C^{\gamma}$
PCuA	2K70/15897	14	cis	**0.998/0.964**	1.000/1.000
	2K6Z/15896	14	trans	**0.029/0.002**	0.000/0.000
DivIB	2ALJ/6395	108	cis	**0.998/0.959**	1.000/1.000
	1YR1/6395	108	trans	**0.002/0.000**	0.000/0.000

[a] PDB code of the reference structure and BMRB code of the chemical shift data
[b] residue number of the center Pro residue
[c] type of Xaa-Pro peptide bond observed in reference structure

Table S5. Promega Xaa-Pro dipeptide bond conformation prediction and CS-Rosetta structure prediction of protein XcR50

ResID[a]	P_{cis}/P_{cis}^{norm} ***(w/o*** $\delta^{13}C^{\gamma}$***)***	P_{cis}/P_{cis}^{norm} ***(w/*** $\delta^{13}C^{\gamma}$***)***	***Observed NOEs***[d]	***Xray isomer***[e]	***NMR isomer***[e]	***csRosetta isomer 1***[f]	***csRosetta isomer 2***[g]
49	1.000/0.993	1.000/0.998		cis	trans (20)	trans(10)	cis(10)
54	0.006/0.000	0.000/0.000	Hα-Hδ	trans	trans (20)	trans(10)	trans(10)
62[h]	0.601/0.078	0.601/0.078		cis	trans (20)	cis(9)	cis(10)
66	0.000/0.000	0.000/0.000	Hβ-Hδ	trans	trans (20)	trans(10)	trans(10)
76	0.273/0.019	0.000/0.000	Hβ-Hδ	cis	trans (20)	trans(10)	trans(10)

[a] residue number of the center Pro residue
[b] P_{cis} score calculated without $^{13}C^{\gamma}$ chemical shift
[c] P_{cis} score calculated with $^{13}C^{\gamma}$ chemical shift
[d] isomer-specified sequential NOEs
[e] type of the Xaa-Pro peptide bond in Xray (PDB ID: 1TTZ) or NMR (1XPV) structure; for NMR structure, the number of conformers (out of a total of 20) with the listed type of the Xaa-Pro peptide bond, is given in parentheses
[f] type of the Xaa-Pro peptide bond in the 10 lowest energy CS-Rosetta models obtained using a standard MFR module; the number of models, with the listed type of the Xaa-Pro peptide bond given in parenthesis
[g] type of the Xaa-Pro peptide bond in the 10 lowest energy CS-Rosetta model obtained using a MFR module that includes a Promega term
[h] no backbone and $^{13}C^{\beta}$ chemical shifts available for Pro62

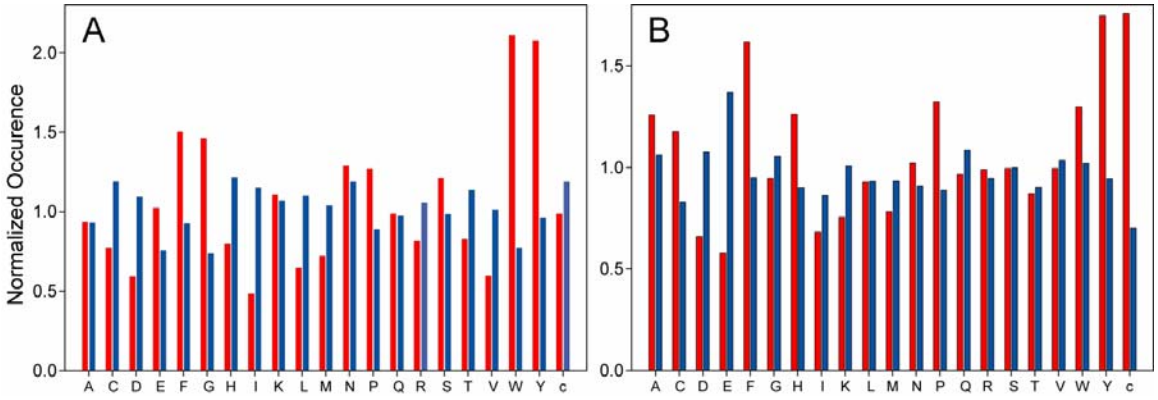

Figure S1. Amino acid occurrence of residues neighboring Pro. (A) Residues preceding and (B) following Pro, with cis (red) or trans (blue) Xaa-Pro peptide bond. The plotted normalized occurrence refers to the probability of finding a given residue type in that position and adjacent to a cis or trans Pro, relative to its probability to find it anywhere in the database (see Eq. 1). Lower case c refers to oxidized Cys.

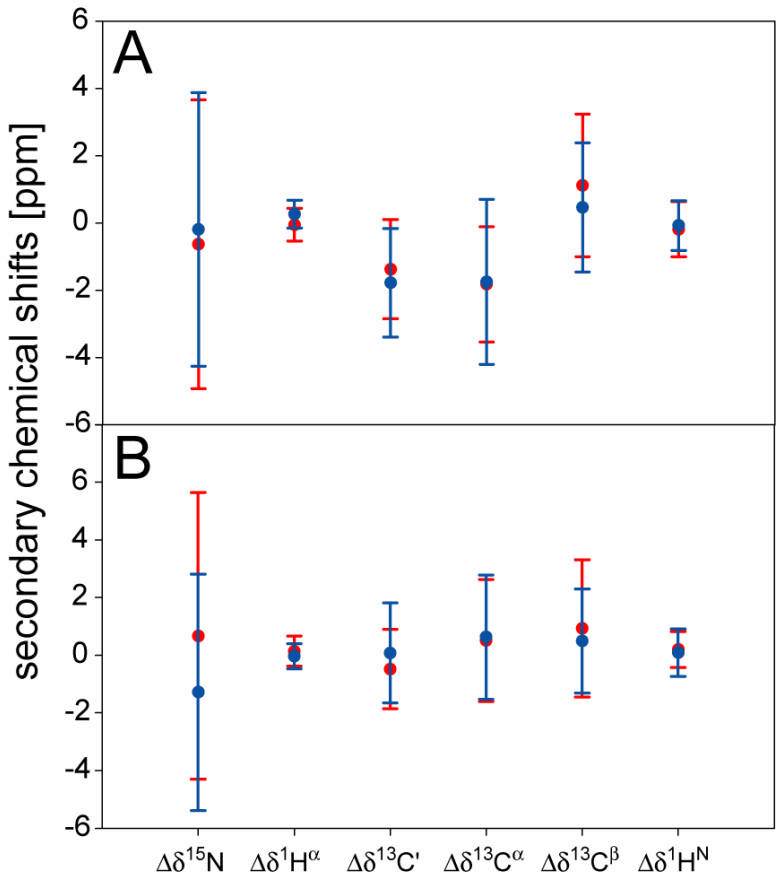

Figure S2. Chemical shift patterns of residues adjacent to Pro. The average secondary chemical shifts and their standard deviations are plotted for ^{15}N, $^{1}H^{\alpha}$, $^{13}C'$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{1}H^{N}$ of residues (A) preceding and (B) following Pro residues with cis (red) or trans (blue) Xaa-Pro peptide bonds.

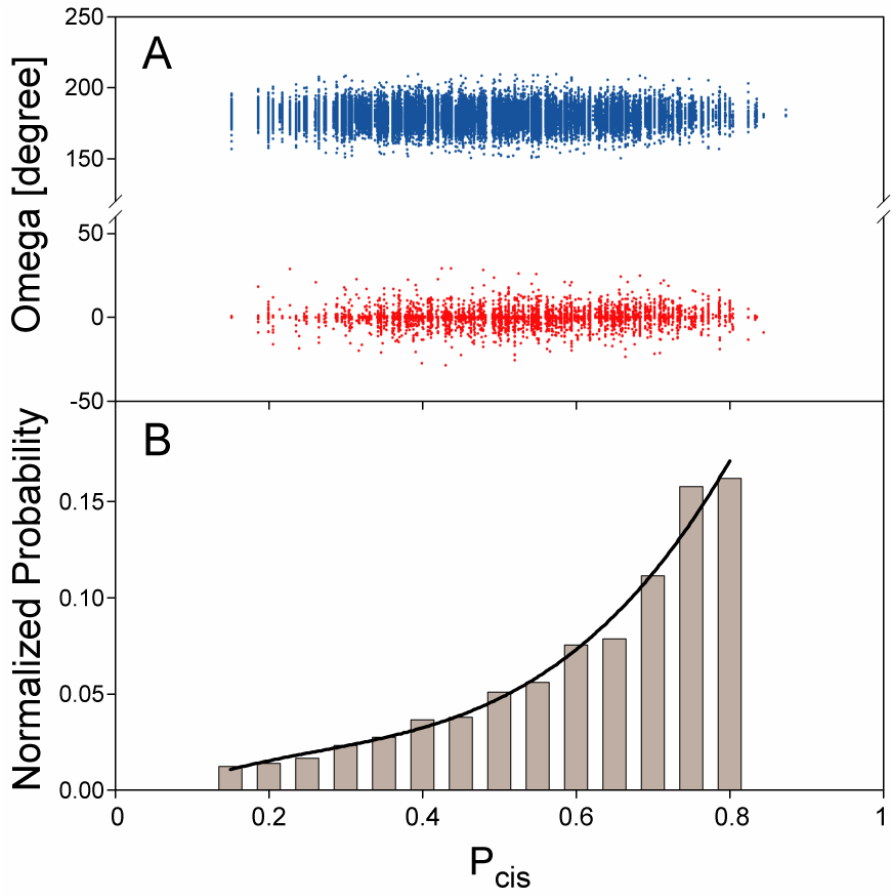

Figure S3. Distribution of Promega Xaa-Pro P_{cis} scores when only amino acid type information is used. (**A**) The peptide bond torsion angle ω is plotted against the probability score P_{cis}, calculated by using eq 5 with no chemical shifts included for either trans (blue) or cis (red) Pro residues (total of 109,396 Pro-centered tri-peptides) in the database. (**B**) Histogram of the fraction of database Pro residues with cis peptide bonds as a function of the P_{cis} score. The solid line represents P_{cis}^{norm}, calculated using eq 6.