Biophysical Journal, Volume *110*

Supplemental Information

Unconventional N-H…N Hydrogen Bonds Involving Proline Backbone Nitrogen in Protein Structures

R. N. V. Krishna Deepak and Ramasubbu Sankararamakrishnan

SUPPORTING INFORMATION

Unconventional N-H…N Hydrogen Bonds Involving Proline Backbone Nitrogen in Protein Structures

R. N. V. Krishna Deepak and Ramasubbu Sankararamakrishnan*

Department of Biological Sciences and Bioengineering, Indian Institute of Technology Kanpur, Kanpur – 208016, India

Figure S1

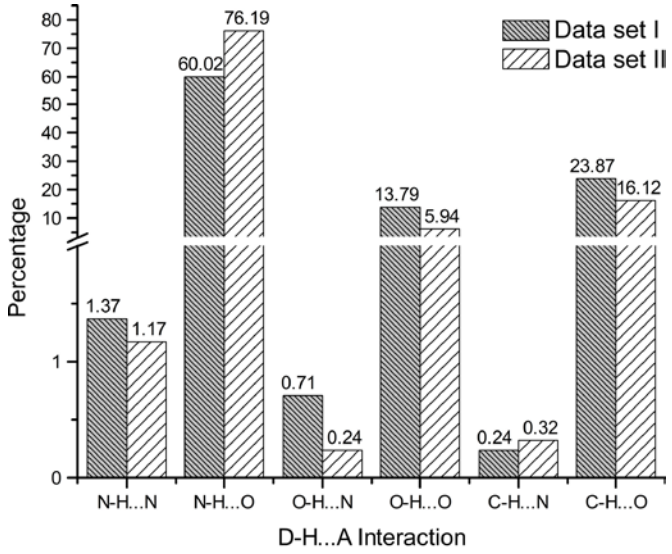

Figure S1: Histogram showing the percentages of different types of hydrogen bonds from Datasets I and II. Here the donor atom can be nitrogen, oxygen or carbon and acceptor atom is either nitrogen or oxygen atom. The numbers shown above each bar indicate the percentage of that particular hydrogen bond in Dataset I or II. We identified 114 and 14347 examples from Dataset-I and Dataset-II respectively in which the geometric criteria for N-H…N type of hydrogen bond were satisfied.

Figure S2

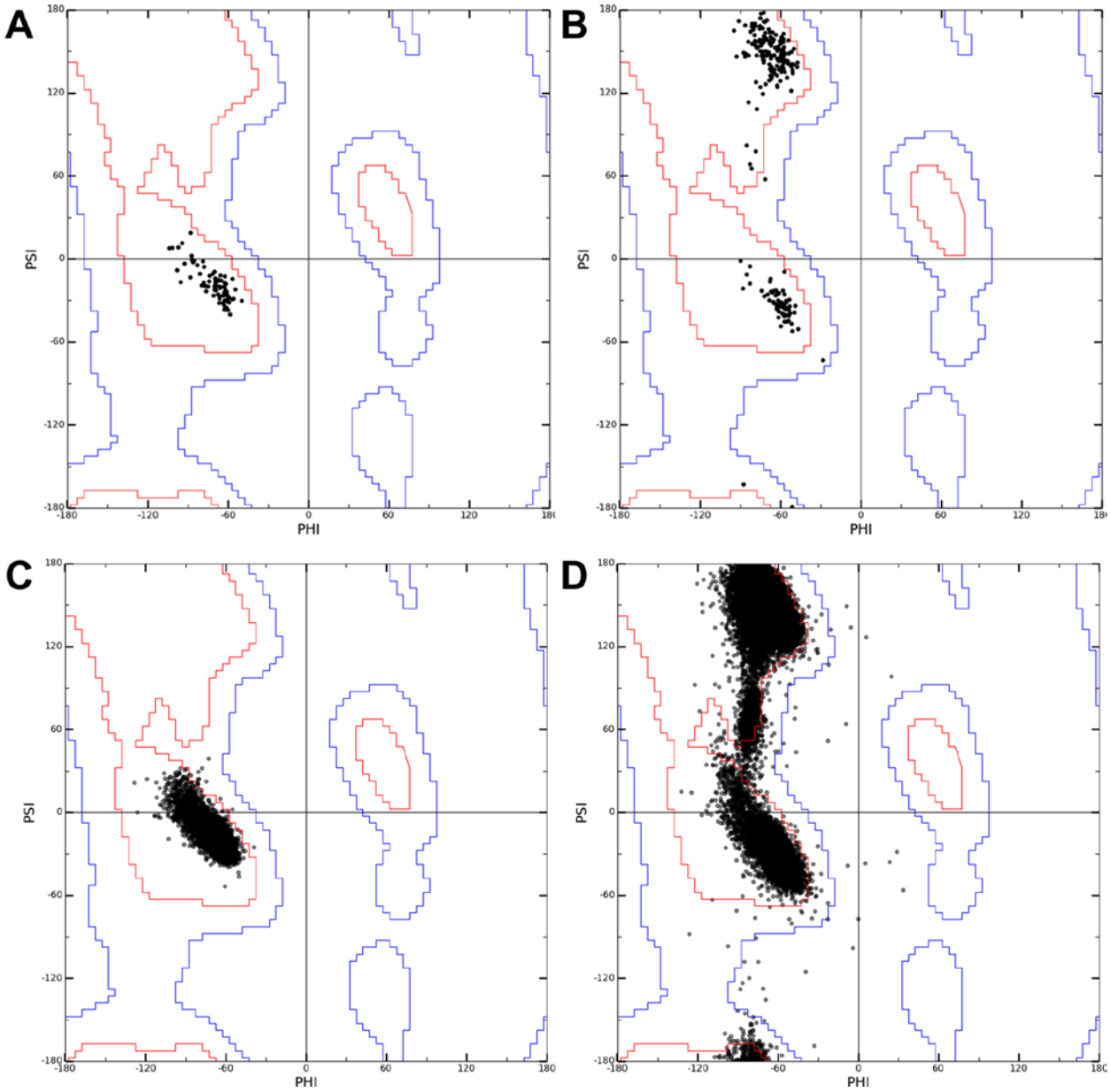

Figure S2: Ramachandran map showing the (φ, ψ) angles of HB-forming proline residues (A and C) and all other proline residues (B and D) from Dataset-I (A and B) and Dataset-II (C and D).

Figure S3

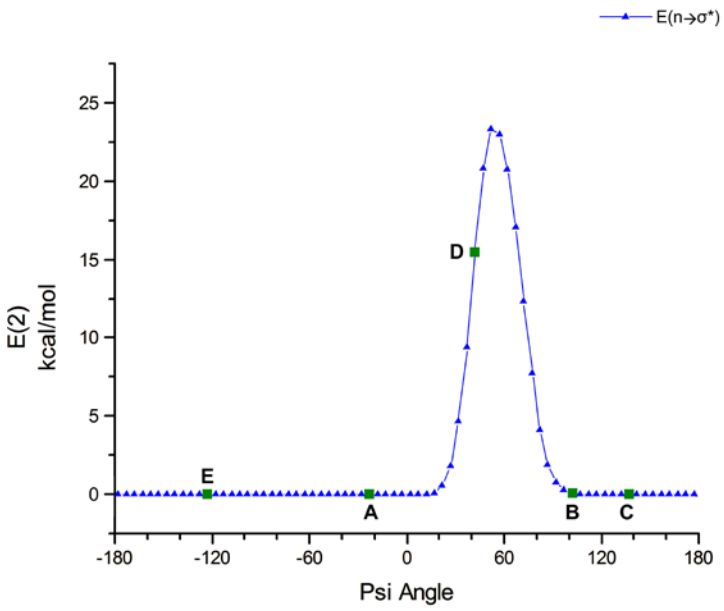

Figure S3: The second order perturbation energy profile for Ni+1-Hi+1…Oi-1 hydrogen bond calculated for all the 72 conformations of Ac-Pro-NMe. The labels A to E represent the different conformations analyzed in detail. NBO calculations were performed as described in the Methods section in the main text.

Figure S4

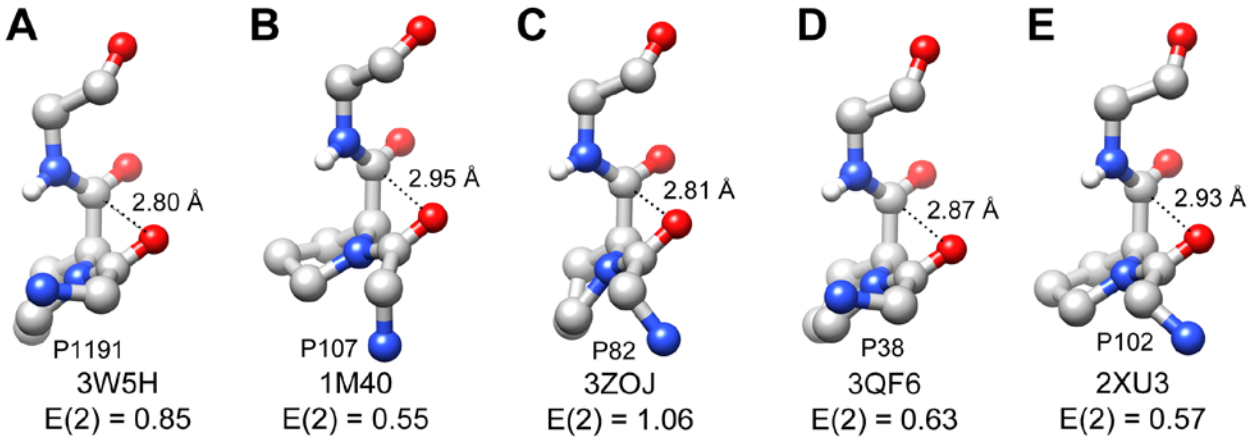

Figure S4: Structures of Ace-Pro-NMe with conformations of proline residues involved in n → π* interactions adopted from different protein structures for which NBO calculations were performed. Hydrogen atoms are optimized as described in the Methods section. The unique PDB Ids of the proteins which possess the n → π* interactions along with the second order perturbation energies are given for each case. The distances between the O_{i-1} of residue preceding proline and C_i of proline are also provided.

Figure S5

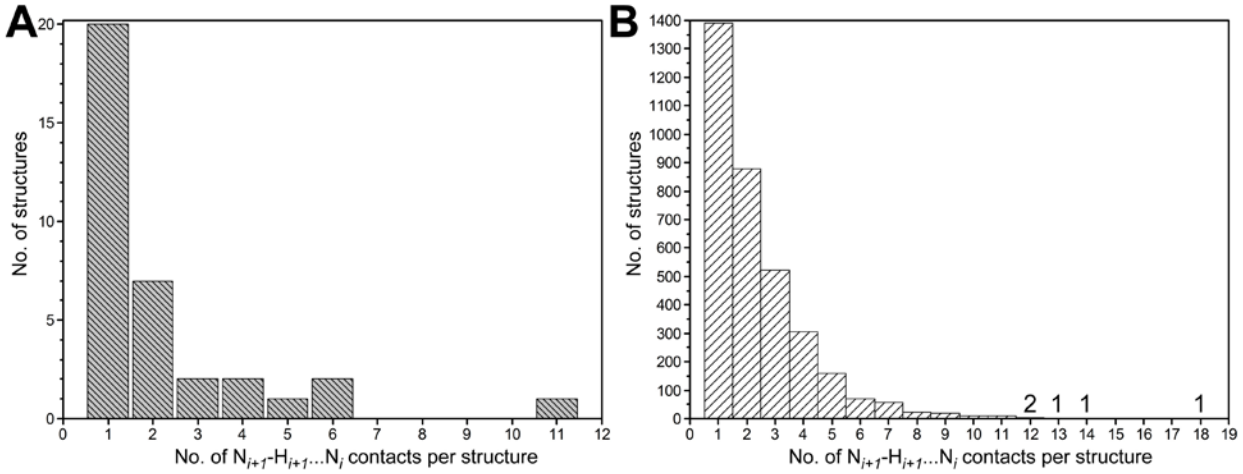

Figure S5: Histograms showing the frequency of occurrence of proline residues with N_{i+1}-H_{i+1}…N_i hydrogen bonds in individual proteins for (A) Dataset-I and (B) Dataset-II. In Dataset-II, the number of proteins having higher number of HB-forming proline residues is explicitly mentioned in the plot.

Table S1: Details of residues participating in N-H…N hydrogen bonds in protein structures[a]

S.No:	DONOR	ACCEPTOR	Dataset-I (%)	Dataset-II (%)
1	Main-chain -NH	Main-chain N (PRO)	66.7	57.0
2	Main-chain -NH	Side-chain N (HIS)	16.7	19.9
3	Side-chain -NH	Side-chain N (HIS)	7.9	9.7
4	All other N-H…N hydrogen bonds		8.7	13.4

[a]This data provides information about the number of N-H…N contacts in the two datasets of protein structures that satisfy the geometric criteria for forming the hydrogen bonds. See the main text for details.

Table S2: Average (Standard Deviation) of geometrical parameters[a] that define n → π* interactions in HB-forming proline and all other proline residues from Dataset-II

Residue	n → π* interaction[b]	$d(O_{i-1}\ldots C_i)$ (Å)	$\theta(O_{i-1}\ldots C_i{=}O_i)$ (°)
HB-forming prolines	Present (2408)	2.95 (0.09)	114.35 (3.60)
	Absent (5197)	3.12 (0.16)	126.76 (6.19)
	Cis prolines (501)	4.49 (0.07)	160.58 (5.43)
All other prolines	Present (19456)	2.93 (0.11)	107.11 (5.29)
	Absent (24826)	3.02 (0.18)	100.26 (19.40)
	Cis prolines (2215)	4.37 (0.13)	95.74 (33.56)

[a]n → π* interactions were identified if $d(O_{i-1}\ldots C_i) \leq 3.22$ Å and $99° \leq \theta(O_{i-1}\ldots C_i{=}O_i) \leq 119°$. The index *i* refers proline residue

[b]The numbers given in the brackets are the number of proline residues under the particular category

Table S3: NPA (Natural Population Analysis) charges on atoms participating in n → σ* and/or n → π* interactions in different conformers

Conformation[a]	N_i	N_{i+1}	H_{i+1}	O_{i-1}	C_i	O_i
A	-0.48273	-0.62771	+0.40851	-0.60770	+0.67181	-0.63640
B	-0.44942	-059404	+0.42294	-0.65579	+0.65624	-0.65376
C	-0.44735	-0.60960	+0.40664	-0.63671	+0.66899	-0.64715
D	-0.44404	-0.62832	+0.42941	-0.63109	+0.65217	-0.65734
E	-0.47064	-0.62214	+0.39315	-0.57581	+0.64548	-0.59607

[a]These conformations correspond to those labeled in the energy profile shown in Figure 3 in the main text.

Table S4: Hyperconjugation between donor and acceptor NBOs and second order perturbation energies for Ac-Pro-NMe in which the single point energies were calculated corresponding to proline conformations from different protein structures

PDB ID[a]	Proline residue[b]	$d(H_{i+1}\ldots N_i)$ (Å)[c]	φ (°)[d]	ψ (°)[d]	E(n → σ*) (kcal/mol)[e,g]	E(n → π*) (kcal/mol)[f,g]
4CVI	Pro-94	2.32 (2.45)	-69.4	-22.9	0.75 (0.11)	0.56 (0.67)
1GKT	Pro-61	2.23 (2.33)	-69.4	-22.5	0.53 (0.11)	0.39 (0.56)
3W5H	Pro-1191	2.41 (2.49)	-58.2	-40.4	0.30 (0.07)	0.85 (1.00)
3W5H	Pro-1249	2.32 (2.43)	-64.4	-31.5	0.52 (0.10)	0.40 (0.54)
1M40	Pro-107	2.34 (2.43)	-67.0	-26.3	0.44 (0.12)	0.55 (0.86)
2R24	Pro-112	2.32 (2.45)	-73.9	-16.4	0.63 (0.08)	0.31 (0.42)
1VYR	Pro-247	2.36 (2.41)	-78.7	-1.7	0.33 (ND)	0.14 (ND)
1VYR	Pro-236	2.33 (2.39)	-92.9	-3.1	0.54 (ND)	0.03 (ND)
2B97	Pro-56	2.31 (2.39)	-103.9	+8.0	0.84 (ND)	< 0.01 (ND)
3ZOJ	Pro-82	2.62 (2.77)	-53.4	-44.7	0.13 (<0.01)	1.06 (1.31)
3QF6	Pro-38	2.41 (2.58)	-61.5	-34.5	0.38 (0.02)	0.63 (0.80)
2XU3	Pro-102	2.40 (2.56)	-60.7	-25.2	0.22 (0.01)	0.57 (0.83)
4NSV	Pro-213	2.38 (2.52)	-82.1	-5.7	0.36 (ND)	0.09 (ND)
4F1V	Pro-1288	2.48 (2.61)	-87.5	-22.5	0.29 (ND)	0.14 (ND)
1C57	Pro-202	2.12 (2.16)	-77.8	-17.4	1.36 (0.90)	0.11 (0.12)
3KKX	Pro-83	2.26 (2.26)	-86.9	+2.5	0.39 (0.38)	0.02 (0.02)
3Q3L	Pro-53	2.18 (2.29)	-85.9	-3.0	1.00 (0.37)	0.10 (0.08)
1VYR	Pro-17	2.28 (2.35)	-81.2	-10.5	0.67 (ND)	0.04 (ND)

[a]The unique PDB IDs of selected protein structures belonging to Dataset-I are given.
[b]The different conformations of Ac-Pro-NMe correspond to these proline residues were considered for single point calculations.
[c]The distances between the atoms H_{i+1} and N_i are provided after the hydrogen positions are optimized. The values given in the brackets correspond to the distances calculated from experimental structures
[d]The backbone (φ, ψ) values of proline residues under consideration are given.
[e]Second order perturbation energy corresponding to N_{i+1}-$H_{i+1}\ldots N_i$ interaction is given. Values in brackets are second order perturbation energies before optimizing the hydrogen positions.
[f]Second order perturbation energy corresponding to $O_{i-1}\ldots C_i{=}O_i$ interaction is given. Values in brackets are second order perturbation energies before optimizing the hydrogen positions.
[g]"ND" refers "Not Determined". The program terminated without calculating the energies.

Table S5: Secondary structure preferences for proline residues with N-H…N hydrogen bonds

Dataset	Proline residues	Helix	Strand	Loop
I	With N-H..N hydrogen bond	26.7%	1.3%	72%
	With bifurcated hydrogen bonds[a]	33.3%	2.8%	63.9%
	All other proline residues	20.5%	10.0%	69.5%
II	With N-H…N hydrogen bonds	24.6%	0.2%	75.1%
	With bifurcated hydrogen bonds[a]	29.4%	0.3%	70.3%
	All other proline residues	20.6%	8.2%	71.2%

[a]Prolines with the bifurcated hydrogen bonds are the subset of HB-forming proline residues. About 50% of proline residues involved in N_{i+1}-H_{i+1}…N_i hydrogen bonds also participate in additional hydrogen bond interactions and in majority of them N_{i+1}-H_{i+1} form hydrogen bonds with the carbonyl oxygen of *i-2*/*i-1* residue.