Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Quantitative evaluation of positive ϕ angle propensity in flexible regions of proteins from three-bond *J* couplings

Jung Ho Lee, Jinfa Ying and Ad Bax*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[*] To whom correspondence should be addressed

Electronic Supplementary Information

Table S1. $^{3}J_{\mathrm{C'H}\alpha}$, $^{3}J_{\mathrm{HNH}\alpha}$, $^{3}J_{\mathrm{C'C'}}$, $^{3}J_{\mathrm{HNC'}}$, and $^{2}J_{\mathrm{C'H}\alpha}$ coupling values in N-terminally acetylated α-synuclein.

	$^{3}J_{\mathrm{C'H}\alpha}$/Hz [a]	$^{3}J_{\mathrm{HNH}\alpha}$/Hz [b]	$^{3}J_{\mathrm{C'C'}}$/Hz [c]	$^{3}J_{\mathrm{HNC'}}$/Hz [a]	$^{2}J_{\mathrm{C'H}\alpha}$/Hz [a]
M1	2.36		0.72	0.99	-4.66
D2	2.30		0.74	0.98	-4.54
V3	2.12	7.25	0.76	0.91	-4.89
F4	2.13	7.13	0.78	0.96	-4.59
M5	2.58	7.13	0.81	0.98	-4.80
K6	2.16	6.11	0.74	1.02	-4.63
G7	3.92		0.66	0.68	-5.00
L8	2.07	7.01	0.73	0.66	-4.50
S9	2.07	6.53	0.83	0.88	-4.37
K10	2.31		0.89	1.11	-4.55
A11	1.76	5.30	0.67	1.21	-3.79
K12	2.18	6.64	0.81	0.90	-4.25
E13	2.12	6.18	0.74	0.92	-4.21
G14			0.72		
V15	2.26	7.37	0.83	0.72	-4.04
V16	2.18	7.63	0.92	0.72	-3.99
A17	1.88	5.45	0.71	1.32	-3.85
A18	1.88	5.22	0.76	1.24	-4.01
A19	1.85	5.28	0.73	1.19	-4.17
E20	1.92	6.13	0.67	0.82	-4.16
K21	2.16	6.53	0.80	0.84	-4.52
T22	2.22	6.99	0.83	0.82	-4.72
K23	2.22	6.65	0.84	0.89	-4.39
Q24	2.08	6.69	0.86	0.75	-4.23
G25			0.72		
V26	2.24	7.12	0.88	0.98	-4.29
A27	1.89	5.58	0.73	1.06	-4.16
E28	1.92	6.08	0.70	1.04	-4.22
A29	2.01		0.73	1.17	-4.23
A30	2.02		0.71	1.26	-4.27
G31	3.69		0.75	0.54	-4.96
K32	2.26	7.01	0.86	0.64	-4.47
T33	2.33	7.22	0.93	0.70	-4.49
K34	2.31		0.91	0.95	-4.36
E35		6.30		0.80	-4.23
G36			0.72		
V37	2.32	7.50	0.86	0.59	-4.15

L38	2.25	7.04	0.73[d]	0.58	-4.20
Y39	2.40	7.47	0.98	0.80	-3.89
V40	2.62	8.01	1.16	0.83	-4.53
G41	3.84		0.82	1.01	-4.66
S42	2.34	6.49	0.87	0.93	-4.19
K43	2.34	7.00	0.85	0.82	-4.55
T44	2.40	7.26	0.87	0.78	-4.54
K45	2.33	6.65	0.86	0.91	-4.44
E46		6.37	0.73	0.93	-4.24
G47			0.74		
V48	2.27	7.58	0.87	0.68	-3.99
V49	2.50		0.98	0.50	-3.98
H50	2.77	7.59	1.09	1.00	-4.41
G51			0.84		
V52	2.41		0.99	0.82	-4.11
A53	2.07	5.79	0.75	1.12	-3.77
T54	2.33	7.56	0.93	0.70	-4.49
V55	2.40		0.99	0.92	-4.17
A56	2.04	5.61	0.74	1.03	-3.84
E57	1.89	6.15	0.72	1.00	-3.96
K58	2.15	6.57	0.79	0.95	-4.23
T59	2.32	7.01	0.81	0.88	-4.39
K60	2.15		0.83	1.03	-4.44
E61	2.04	6.17	0.72	0.80	-3.98
Q62	2.21	7.01	0.83	0.83	-4.14
V63	2.29	7.70	0.93	0.82	-4.11
T64	2.58	7.87	1.01	0.97	-4.74
N65	3.02	7.51	0.93	0.88	-4.62
V66	2.43	7.34	0.91	0.95	-4.53
G67			0.76	0.78	
G68	4.03		0.78	0.74	-4.84
A69	1.98	5.83	0.74	1.11	-3.78
V70	2.31	7.53	0.87	0.68	-3.72
V71	2.43	8.15	1.12	0.61	-3.95
T72	2.50	7.70	1.02	1.00	-4.80
G73	3.90		0.78	0.94	-4.85
V74	2.34		0.90	0.75	-4.17
T75	2.38	7.62	0.95	0.80	-4.56
A76	2.05	5.87	0.81	1.12	-3.88
V77	2.23	7.45	0.93	0.80	-4.01
A78	2.01	5.62	0.71	1.30	-3.75

Q79	2.09	6.79	0.82	0.93	-4.02
K80	2.21	6.71	0.86	0.99	-4.10
T81	2.33	7.52	0.93	0.66	-4.17
V82	2.36	7.65	0.98	0.83	-4.18
E83	2.06	6.19	0.78	0.97	-3.96
G84	3.78		0.79	0.65	-4.93
A85	2.12	5.64	0.71	1.40	-4.43
G86	3.89		0.72	0.65	-4.99
S87	2.24	6.82	0.90	0.94	-4.61
I88	2.32	7.46	0.89	0.80	-4.39
A89	1.89	5.57	0.73	1.30	-3.97
A90	1.85	5.62		0.98	-4.05
A91	1.99	5.77	0.71	1.04	-4.11
T92	2.39	7.50	0.99	0.79	-5.09
G93	3.98		0.76	0.40	-4.81
F94	2.21	6.92	0.94	0.89	-4.02
V95	2.51	8.04	1.22	0.74	-3.85
K96	2.03	6.51	0.71	0.97	-3.68
K97			0.92		
D98	2.45	6.72	0.72[d]	0.82	-4.53
Q99	2.42	7.23	0.89	0.88	-4.78
L100	2.26	6.80	0.71	0.86	-4.66
G101	3.99		0.72	0.69	-4.60
K102	2.24	6.83	0.91	1.10	-4.39
N103	3.01	7.20	0.84	0.91	-4.84
E104	2.30	6.79	0.82	0.89	-4.59
E105	2.12	6.39	0.74	0.90	-4.33
G106	3.82		0.79	0.65	-5.03
A107	1.89	5.92		1.12	-3.09
Q109	2.35	6.94	0.89	0.77	-4.29
E110	2.21	6.52	0.82	0.98	-4.20
G111	3.64		0.75	0.73	-4.47
I112	2.39	7.75	0.89	0.69	-4.10
L113	2.35	7.16	0.79	0.87	-4.15
E114	2.31	6.72	0.84	1.00	-4.34
D115	2.82	6.85	0.83	0.84	-4.23
M116	2.45	7.17		0.72	-3.47
V118	2.39	7.68	0.91	0.63	-4.05
D119		6.79		0.97	
D121	2.49	7.14		0.42	-4.91
N122	2.69	7.29		0.72	-5.00

E123	2.24	6.55	0.71	0.95	-4.59
A124	2.18	6.23	0.81	0.89	-4.20
Y125	2.28	7.09	1.05	0.96	-4.26
E126	2.45	7.29	1.05	0.75	-3.97
M127	2.02	6.62		1.17	-3.34
S129	2.15	6.65	0.93	0.91	-4.16
E130	2.32	6.86	0.88	0.88	-4.53
E131	2.09	6.41		0.98	-4.29
G132	3.94		0.79	0.82	-4.73
Y133	2.29	6.77	0.96	1.00	-4.17
Q134	2.50	7.55	1.16	1.04	-4.53
D135	2.61	6.97		1.03	-4.62
Y136	2.23	7.17	1.11	1.04	-4.06
E137	2.47	7.63		0.94	-3.58
E139	2.19	6.71	0.74	0.84	-4.02
A140	2.07	6.65		1.43	-5.06

[a] $^{3}J_{\mathrm{C'H\alpha}}$, $^{2}J_{\mathrm{C'H\alpha}}$, and $^{2}J_{\mathrm{HNC'}}$ couplings are measured for a ^{13}C/^{15}N-enriched α-synuclein sample in 20 mM sodium phosphate, 50 mM NaCl, and 5% D_2O at pH 6, 288 K, and 600 MHz ^{1}H frequency.

[b] $^{3}J_{\mathrm{HNH\alpha}}$ couplings are from [1].

[c] $^{3}J_{\mathrm{C'C'}}$ couplings are measured under the same condition as above, except for 20 mM NaCl and a ^{1}H frequency of 900 MHz. Correlation between the $^{3}J_{\mathrm{C'C'}}$ couplings in this table and Lee et al. [2] is shown in ESI Fig. S1.

[d] $^{3}J_{\mathrm{C'C'}}$ coupling from Lee et al.[2].

Table S2. ${}^{3}J_{\mathrm{C'H\alpha}}$ and ${}^{2}J_{\mathrm{C'H\alpha}}$ coupling values in GB3. [a]

	${}^{3}J_{\mathrm{C'H\alpha}}$/Hz	${}^{2}J_{\mathrm{C'H\alpha}}$/Hz
M1		-3.45
Q2	2.61	-3.53
Y3	2.81	-4.35
K4	3.06	-4.31
L5	3.23	-3.15
V6	2.83	-3.58
I7	3.11	-3.46
N8	3.49	-4.46
K10	1.56	-5.41
T11	3.1	-6.73
L12	2.39	-3.48
K13	2.84	-3.98
E15	2.57	
T16		-5.96
T17	3.14	-5.32
T18	2.67	-4.92
K19	2.83	-4.07
A20	2.53	-4.85
V21	2.33	-5.53
D22	2.18	-5.61
A23	1.73	-4.97
E24	1.32	-4.78
T25		-5.78
A26	0.94	-5.42
K28	1.14	-5.16
A29	1.05	
F30		-6.52
K31	1.2	-5.71
Q32	1.2	-5.01
Y33	0.96	-4.55
A34	1.06	-5.18
N35	1.15	-5.27
D36	1.35	-5.63
N37	2.68	-6.45
V39	2.63	-3.66
D40	2.33	-4.17
V42	2.38	-3.09
W43	2.86	-3.5

T44	3.21	
Y45		-4.39
D46	3.2	-3.67
D47	1.68	
A48	1.44	-5.61
T49	2.81	-6.3
K50	6.7	-6.46
T51	3.01	
F52		-4.18
T53	3.01	
V54		-4.46
T55	3.19	-4.33
E56	2.34	

[a] $^{3}J_{\mathrm{C'H}\alpha}$ and $^{2}J_{\mathrm{C'H}\alpha}$ couplings are measured for a uniformly ^{13}C/^{15}N-enriched 1.2 mM GB3 sample in 50 mM sodium phosphate, 50 mM NaCl, and 5% D_2O at pH 6.5, 288 K, and 600 MHz ^{1}H frequency.

Table S3. $^{3}J_{\mathrm{C'H\alpha}}$ and $^{2}J_{\mathrm{C'H\alpha}}$ coupling values in ubiquitin. [a]

	$^{3}J_{\mathrm{C'H\alpha}}$/Hz	$^{2}J_{\mathrm{C'H\alpha}}$/Hz
M1		-4.14
Q2	2.76	-4.02
I3	3.19	-5.65
F4	3.16	
V5		-4.05
K6	2.71	-3.69
T7	2.5	
L8	1.68	-5.4
T9	2.26	
K11	1.69	-3.21
T12	3.14	-4.63
I13	3	-4.08
T14	2.97	-4.18
L15	3.14	-4.62
E16	2.63	-3.58
V17	2.76	
E18	3.07	
P19		-4.82
S20	2.06	-5.35
D21	1.11	-3.39
T22	1.91	-4.64
I23	1.43	
E24		-5.22
N25	1.23	
V26		-5.49
K27	1.1	-5.57
A28	1.14	
K29	1.65	-5.48
I30	1.14	-5.64
Q31	0.91	-4.86
D32	1.2	
K33		-5.82
E34	2.9	
I36	2.23	
P38		-4.74
D39	1.67	
Q40	2.69	-6.01
Q41	1.98	-3.58

R42	2.92	
L43		-4.55
I44	3.01	-4.11
F45	2.87	-4.05
A46	6.89	
K48	2.54	-3.87
Q49	1.83	-3.21
L50	2.53	-3.45
E51	2.57	
D52	1.2	
R54	2.13	-4.35
T55	3.08	
L56	1.04	
S57	1.35	
D58	1.7	-5.88
Y59	2.56	-5.63
N60	6.85	-5.9
I61	2.13	-3.01
Q62	2.87	
K63	1.1	-2.98
E64	6.99	-7.05
S65	1.31	-3.62
T66	3.29	-4.48
L67	2.91	
H68	3.45	-4.28
L69	2.64	-3.62
V70	3.04	-4.39
L71	2.36	-4.19
R72	2.1	-3.73
L73	2.29	-4.06
R74	2.12	-4.07

[a] $^{3}J_{C'H\alpha}$ and $^{2}J_{C'H\alpha}$ couplings are measured for a uniformly $^{13}C/^{15}N$-enriched 2.8 mM ubiquitin sample in 20 mM imidazole and 7% D2O at pH 6.0, 288 K, and 600 MHz ^{1}H frequency.

Table S4. Parameters calculated from the newly generated coil database [3] for $\phi > 0$ conformations.

	N(ϕ<0)[a]	N(ϕ>0) [b]	$<^3J_{\mathrm{C'H\alpha}}(\phi>0)>$ [c]	$<^3J_{\mathrm{HNH\alpha}}(\phi>0)>$	$<^3J_{\mathrm{C'C'}}(\phi>0)>$
ARG	8764	354	6.53	6.92	0.62
LYS	10657	347	6.53	6.90	0.64
ASP	11034	743	6.79	7.20	0.56
GLU	11836	352	6.37	6.74	0.74
SER	13273	386	6.17	6.51	0.77
THR	12905	110	5.93	6.10	0.81
ASN	7492	1168	6.90	7.35	0.55
GLN	6320	219	6.48	6.88	0.68
ALA	11346	314	6.23	6.61	0.77
VAL	9802	72	5.89	6.14	0.88
ILE	6939	36	6.15	6.44	0.83
LEU	10625	274	6.53	6.90	0.63
MET	2208	75	6.78	7.15	0.53
PHE	5873	273	6.86	7.30	0.54
TYR	5442	177	6.69	7.12	0.60
TRP	1984	54	6.55	7.06	0.66
HIS	4573	258	6.58	7.03	0.65
CYS	1944	89	6.84	7.29	0.53
GLY	6667	10199	5.19	5.05	0.87
PRO	25206	5	5.39	5.03	0.99

[a] Total number of $\phi < 0$ residues for each amino acid type in the coil database.

[b] Total number of $\phi > 0$ residues for each amino acid type in the coil database.

[c] The rigid limit Karplus equation, $^3J_{\mathrm{C'H\alpha}} = 4.17 \times \cos^2(\phi - 60°) + 2.00 \times \cos(\phi - 60°) + 1.02$, was used to calculate $^3J_{\mathrm{C'H\alpha}}$ couplings for all $\phi > 0$ residues for each amino acid type and averaged afterwards. The procedure was used to calculate the next two columns using rigid limit Karplus equations.[2].

Table S5. Positive ϕ population (P^{+}) calculated by the MERA webserver,[3] by the interpolation method using ${}^{3}J_{\mathrm{HNH}\alpha}$ and ${}^{3}J_{\mathrm{C'H}\alpha}$, and by the iterative method using ${}^{3}J_{\mathrm{HNH}\alpha}$, ${}^{3}J_{\mathrm{C'C'}}$, and ${}^{3}J_{\mathrm{C'H}\alpha}$.[a]

	P^{+} (MERA) / %	P^{+} (interpolation method) / %	P^{+} (iterative method) / %
D2	4		
V3	0	-5	-4
F4	4	-3	-3
M5	4	7	7
K6	2	3	3
L8	3	-4	-3
S9	2	-1	-1
K10	4		
A11	2	-2	-2
K12	2	1	1
E13	2	2	2
V15	0	-2	-1
V16	0	-6	-5
A17	2	0	0
A18	2	1	0
A19	2	0	-1
E20	1	-2	-2
K21	2	1	1
T22	0	0	0
K23	3	2	1
Q24	5	-2	-2
V26	0	-1	-1
A27	2	0	-1
E28	2	-2	-2
A29	3		
A30	3		
K32	3	1	1
T33	0	1	1
K34	4		
E35	1		
V37	0	-1	-1
L38	2	0	1
Y39	5	1	1
V40	0	4	3
S42	4	6	5
K43	4	3	3

T44	0	3	3
K45	3	4	4
E46	1		
V48	0	-3	-3
V49	0		
H50	8	9	9
V52	0		
A53	2	3	2
T54	0	-1	-1
V55	0		
A56	2	3	3
E57	1	-3	-3
K58	2	1	1
T59	0	2	2
K60	2		
E61	2	0	0
Q62	2	-1	0
V63	0	-3	-3
T64	0	4	3
N65	13	15	14
V66	0	3	3
A69	2	1	0
V70	0	-2	-1
V71	0	-3	-3
T72	0	3	2
V74	0		
T75	0	0	0
A76	3	2	1
V77	0	-3	-3
A78	3	2	2
Q79	2	-2	-2
K80	1	1	1
T81	0	-1	-1
V82	0	-1	-1
E83	2	1	0
A85	3	5	5
S87	3	1	1
I88	0	-1	-1
A89	2	0	-1
A90	2	-1	
A91	2	1	1

T92	0	1	1
F94	3	0	0
V95	0	0	0
K96	4	-2	-1
K97	5		
D98	8	6	7
Q99	3	3	3
L100	2	2	2
K102	4	1	1
N103	15	16	16
E104	3	3	3
E105	4	1	1
A107	2	-2	
Q109	2	3	3
E110	2	2	2
I112	0	-1	-1
L113	2	2	2
E114	3	4	3
D115	6	14	14
M116	6	4	
V118	0	0	0
D119	7		
D121	5	5	
N122	12	8	
E123	3	3	3
A124	3	3	3
Y125	2	1	0
E126	2	4	3
M127	4	-3	
S129	2	0	-1
E130	2	3	3
E131	2	0	
Y133	3	3	2
Q134	2	3	3
D135	7	8	
Y136	3	-1	-1
E137	2	2	
E139	4	1	1
A140	1	-2	

[a] P^+ errors for both iterative and interpolation methods are dominated by the intrinsic uncertainty in the $^3J_{\mathrm{C'H\alpha}}$ Karplus curve, estimated at 0.25 Hz based on the fits to ubiquitin and GB3 ϕ angles, yielding P^+ uncertainties of *ca* 5%. The true error is likely to be somewhat smaller, since factors other than ϕ backbone torsion angle would largely average to zero in IDPs and IDRs.

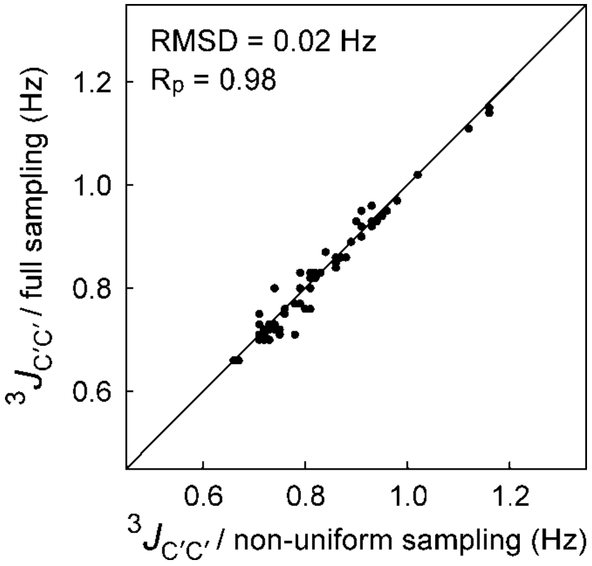

Fig. S1 Correlation between ${}^3J_{C'C'}$ coupling values derived from HN(COCO)NH with a full sampling scheme[2] and HN(COCO)NH with a 2.5% non-uniform sampling scheme.

1. J. Roche, J. Ying, A. S. Maltsev and A. Bax, *ChemBioChem*, 2013, **14**, 1754-1761.
2. J. H. Lee, F. Li, A. Grishaev and A. Bax, *J. Am. Chem. Soc.*, 2015, **137**, 1432-1435.
3. A. B. Mantsyzov, Y. Shen, J. H. Lee, G. Hummer and A. Bax, *J. Biomol. NMR*, 2015, **in press**.