Biophysical Journal, Volume 114

Supplemental Information

Anticooperative Nearest-Neighbor Interactions between Residues in

Unfolded Peptides and Proteins

Reinhard Schweitzer-Stenner and Siobhan E. Toal

1. Using Experimental J-coupling Constants and Vibrational Band Profiles to Extract Conformational Propensities

In this study, we explore whether nearest-neighbor interactions (NNIs) between amino acid residues depend on the conformation of a given neighbor or solely physico-chemical properties of the neighbor. To this end, we utilize conformational propensities of x- and y- residues in GxyG peptides previously determined via a global fitting of a set of experimentally obtained J-coupling constants as well as amide I' band profiles. This fitting procedure is extensive and has been described in full by us previously (1). Briefly, we utilize a proven excitonic coupling algorithm to mathematically describe experimental amide I' profiles obtained from IR, Raman, and VCD studies, in conjunction with fitting restraints provided by six independent J-coupling constants extracted from hetero-nuclear 2D NMR experiments. Each coupling constant depends differently on a given residue's backbone angle according to empirically derived Karplus equations. Four of the coupling constants: 3J(H^α,C'), 3J(H^N,C'), 3J(H^N,Cα), 3J(H^N,C^β) depend solely on the angle ϕ , while the two remaining ones, $1J(NCa)$ and $2J(NCa)$ depend on ψ . Moreover, since the set of Karplus relationships for each J-coupling constant are out of phase from each other (2), these six restraints in combination with independent vibrational experiments have been shown to be sufficient for quantifying conformational distributions of amino acid residues ((1, 3, 4)). For instance, in a seminal study by Graf et al.(2) it was shown that the combined use of all NMR J-coupling constants alone along with distributions derived from constrained all atom MD simulations was sufficient to accurately obtain the conformational ensemble of various alanine based peptides. In addition, work by the Schwalbe group has shown that various J coupling constants along with SAXS data can be used to determine residue level structure in IDPs (5). Previous and extensive work by our group has also demonstrated the use coupling constants in combination with vibrational spectroscopy to examine conformational ensembles of amino acid residues (1, 3, 4, 6).

2. Calculation of interaction Gibbs energies.

We start with equations (3) in the main manuscript and solve eq. (3a) for $e^{(\Delta G_{P2\beta,1}+\delta G_{\beta(1)\beta(2)})/RT}$:

$$
e^{(\Delta G_{pPI\beta,1} + \delta G_{pPI\beta})/RT} = R_1 \cdot \left[e^{(\Delta G_{pPI\beta,2} + \delta G_{pPI\beta})/RT} + e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2} + \delta G_{\beta\beta})/RT} - e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2} + \delta G_{\beta\beta})/RT} \right]
$$
(S1)

Now, we insert eq.(S1) into eq. (3b) and solve it for $e^{(\Delta G_{P2\beta,2}+\delta G_{P2(1)\beta(2)})/RT}$:

$$
e^{(\Delta G_{pPI\beta,2}+\delta G_{pPI\beta})/RT} = R_2 \left\{ 1 + R_1 \left[1 + e^{(\Delta G_{pPI\beta,2}+\delta G_{pPI\beta})/RT} + e^{ \Delta G_{pPI\beta,2}/RT} \right] - e^{(\Delta G_{pPI\beta,1}+\Delta G_{pPI\beta,2})/RT} - e^{(\Delta G_{pPI\beta,1}+\Delta G_{pPI\beta,2})/RT} \right\}
$$

$$
- e^{(\Delta G_{pPI\beta,1}+\Delta G_{pPI\beta,2}+\delta G_{p0})/RT} - e^{(\Delta G_{pPI\beta,1}+\Delta G_{pPI\beta,2})/RT}
$$

$$
(\mathsf{S2})
$$

which could be rearranged into:

$$
e^{(\Delta G_{pPIB,2} + \delta G_{pPIB})/RT} = \frac{R_2 (1 + R_1) - (R_2 + 1) e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,2} + \delta G_{\beta\beta})/RT} - R_2 e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,2})/RT} - e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,2})/RT}}
$$
\n
$$
(1 - R_1 R_2)
$$

$$
(S3)
$$

from where eq.(4a) can be easily obtained. Eq. (4b) can be more easily obtained after solving eq. (3b) for $e^{(\Delta G_{pPIl\beta,2}+\delta G_{\beta pPIl})/RT}$

3. Error propagation

Calculating the error of $\delta G_{_{pPH\beta}}$ and $\delta G_{_{\beta pPH}}$ is straightforward but somewhat tedious. In view of the complexity of the considered function, it make sense to write eqs. (4a) and (4b) as:

$$
\delta G_{pPIB} = RT \cdot \ln f_1 \tag{S4a}
$$

$$
\delta G_{\beta pPII} = RT \cdot \ln f_2 \tag{S4b}
$$

so that the elements of the Jacobi matrix can be written as;

$$
\frac{\partial \delta G_{pPH\beta}}{\partial R_i} = RT \cdot \frac{\partial f_1/\partial R_i}{f_1}
$$
 (S5a)

$$
\frac{\partial \delta G_{\beta pPII}}{\partial R_i} = RT \cdot \frac{\partial f_2 / \partial R_i}{f_2}
$$
 (S5b)

where

$$
f_1 = \frac{R_2 (1 + R_1) - (R_2 + 1)e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2} + \delta G_{\beta\beta})/RT} - R_2 e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2} + \Delta G_{pPI
$$

$$
f_2 = R_1 \left(1 + e^{(\Delta G_{pPI\beta,2} + \delta G_{pPI\beta})/RT} + e^{\Delta G_{pPI\beta,2}} \right) - e^{(\Delta G_{pPI\beta,1} + \Delta G_2)/RT} - e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2})/RT} + \Delta G_{p_2\rho,1}
$$
(S6b)

The first partial derivatives of f_1 with regard to R_1 and R_2 reads as:

$$
\frac{\partial f_1}{\partial R_1} = \frac{R_2 (1 - R_1 R_2) + R_2^2 (1 + R_1) + R_2 (1 + R_2) e^{(\Delta G_{pPII\beta,1} + \Delta G_{pPII\beta,2} + \delta G_{\beta\beta})/RT}}{(1 - R_1 R_2)^2}
$$
(S7a)

$$
\frac{\partial f_1}{\partial R_2} = \frac{R_1 - \left[1 - R_1(2R_2 + 1)e^{(\Delta G_{P2\beta,1} + \Delta G_{P2\beta,2} + \delta G_{\beta\beta})/RT} - (1 + R_2)e^{(\Delta G_{P2\beta,1} + \Delta G_{P2\beta,2})/RT}\right]}{(1 - R_1 R_2)^2}
$$
(S7b)

Correspondingly, one obtains for the derivatives of *f2*:

$$
\frac{\partial f_2}{\partial R_2} = 1 + e^{(\Delta G_{pPI\beta,1} + \delta G_{\beta(1)pPI})/RT} + e^{\Delta G_{pPI\mu,1}/RT}
$$
(S8)

In addition we use eqs. (S7a) and (S7b) to utilize eq. (S5a), which can further be used in combination with:

$$
\frac{\partial f_2}{\partial \delta G_{pPIB}} = \frac{R_1}{RT} e^{(\Delta G_{pPIB,1} + \delta G_{pPIB})/RT}
$$
(S9)

Eqs. (S7)-(S9) can be used to calculate the statistical error of the interactions parameters in dependence of the estimated statistical errors of R_1 and R_2 .

In order to estimate the errors of R_1 and R_2 we very conservatively assumed a statistical error of 0.05 for the propensity values of both, pPII and β-strand.(1). We estimated relative statistical errors between 0.03 (for high pPII fractions) and 0.07 (for comparable pPII and βfractions), based on the uncertainties of the respective mole fractions .

4. Partition sum and mole fractions

The partition sum for GxyG peptides can be written as:

$$
Z = 1 + e^{(\Delta G_{pPIB,2} + \delta G_{pPIB})/RT} + e^{\Delta G_{pPIB,1}/RT} + e^{(\Delta G_{pPIB,1} + \delta G_{pPIB})/RT} + e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,2} + \delta G_{\beta\beta})/RT}
$$
\n
$$
+ e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,2})/RT} + e^{(\Delta G_{pPIB,2} + \Delta G_{pPIB,2})/RT} + e^{(\Delta G_{pPIB,1} + \Delta G_{pPIB,1} + \Delta G_{pPIB,2})/RT}
$$
\n(S10)

The mole fractions of the conformations pPII-pPII, pPII-β,β-pPII, etc can be calculated by using;

$$
\chi_{pPIpPI} = \frac{1}{Z} \tag{S11a}
$$

$$
\chi_{pPIB} = \frac{e^{(\Delta G_{pPIB,2} + \delta G_{pPIB})/RT}}{Z}
$$
(S11b)

$$
\chi_{\beta pPII} = \frac{e^{(\Delta G_{pPI\beta,1} + \delta G_{\beta pPI})/RT}}{Z}
$$
 (S11c)

$$
\chi_{\beta\beta} = \frac{e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\beta,2})/RT}}{Z}
$$
 (S11d)

$$
\chi_{\text{pPIII}} = \frac{e^{\Delta G^*_{\text{pPIII},2}/RT}}{Z} \tag{S11e}
$$

$$
\chi_{\text{tpPI}} = \frac{e^{\Delta G^*_{\text{pPIh},1}/RT}}{Z}
$$
 (S11f)

$$
\chi_{\beta t} = \frac{e^{(\Delta G_{pPI\beta,1} + \Delta G_{pPI\mu,2})/RT}}{Z}
$$
(S11g)

$$
\chi_{t\beta} = \frac{e^{(\Delta G_{pPH\beta 21} + \Delta G^*_{pPHt,1})/RT}}{Z}
$$
 (S11h)

$$
\chi_{tt} = \frac{e^{(\Delta G^*_{pPIII1} + \Delta G^*_{pPIII2})/RT}}{Z}
$$
 (S11i)

5. Calculation of mole fraction to calculate the temperature dependence of ∆ε²¹⁸ **of the**

MAX3 peptide

The total mole fractions of residues adopting pPII, β and turn-like structures in polypeptides can be calculated as follows:

$$
\chi_{k} = \frac{\sum_{\{k\}} \left(n_{k} q_{N} \prod_{i=0}^{N-1} P_{i+1i}^{(k^{*})} \right)}{N \cdot Z}
$$
(S12)

where represents one of three transfer matrices that are written as follows:

$$
P_{i+1i}^{pPII} = \begin{pmatrix} 1 & 0 & 0 \\ \exp(\delta G_{\beta(i)pPII(i+1)}/RT) & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}
$$
 (S13a)

$$
P_{i+1i}^{\beta} = \begin{pmatrix} 0 & \exp\left(\left(\Delta G_{pPI\beta,i} - \delta G_{pPI(i)\beta(i+1)}\right)/RT\right) & 0 \\ 0 & \exp\left(\Delta G_{pPI\beta,i}/RT\right) & 0 \\ 0 & \exp\left(\Delta G_{pPI\beta,i}/RT\right) & 0 \end{pmatrix}
$$
(S13b)

$$
P_{i+1i}^{t} = \begin{pmatrix} 0 & 0 & \exp(\Delta G_{pPIIt,i}/RT) \\ 0 & 0 & \exp(\Delta G_{pPIIIt,i}/RT) \\ 0 & 0 & \exp(\Delta G_{pPIIIt,i}/RT) \end{pmatrix}
$$
(S13c)

The index *k* in eq. (12) labels the residue conformation ($k=pPII$, β *and t*) while $\{k\}$ represents a sequence of *N* residues. Hence, the choice for $P_{i+1,i}^{(k)}$ depends on the specific conformation of the sequence *k'* at the position *j*. The statistical weight of each conformation is multiplied with the number n_k of residues that adopt the conformation k in the sequence. Note the interaction parameter $\delta G_{\beta\beta}$ has been omitted in the above formalism.

Table S1: List of interaction parameters and their estimated statistical errors in kJ/mol

a: GALG

b: GSAG

c: GDAG[1](#page-7-0)

d: GAV[G2](#page-7-1)

e: GDLG

 1δ 1δ G_{ββ}=-5.0 kJ/mol for T=298 K

[2](#page-7-3) δGββ=2.0 kJ/mol for T=298 K

f: GSLG[3](#page-8-0)

g: GLLG

Contract Contract

a se a construír a construí

h: GVLG

i: GDKG

j: GKLG

T[K]	$\delta G_{\rm pPIIB}$	$\Delta \delta G_{pPIIB}$	$\delta G_{\beta p$ PII	$\Delta \delta G_{\beta p$ pii
298	14	26	13	26
353	6.6	0.4	3.8	0.8

[³](#page-8-1)δG_{ββ}=-3.0 kJ/mol for 298 K

k: GKVG

l: GDVG

m: GSVG

Table S2: List of ³J(H^{NHα}) parameters used for the simulation of the coupling constants of E3 and D7 of $A\beta(1-9)$

Figure S1: Thermodynamic parameters *ΔGp2^β* (upper panel), *ΔHp2*^β (black, lower panel) and *ΔSp2^β* (grey, lower panel) of indicated amino acid residues in the absence and presence of nonglycine neighbors as derived from the temperature dependence of $3J(HNH^{\alpha})$ by Toal et al.(1) Each figure shows the parameters for a distinct residue and different neighbors. It should be noted that originally the corresponding plots in the paper of Toal et al.(1) contained some minor errors which were corrected for this figure. A corrigendum of the paper with the above figure is now available together online with the paper.

Figure S2: Graphic representation of the difference between the Gibbs energy gaps between pPII and β-strand of a guest residue in GxG and in respective GxyG host/guest peptides at room temperature and at 353 K calculated using earlier reported and values.(1, 3, 4, 6, 7)

Figure S3: Correlation plot relating the β-strand fraction χ_{β} of mole fractions of x- and y-residues to the β-strand fraction of the respective neighbor χβN for a reduced set of earlier investigated GxyG peptides at T=298 K (upper panel) and T=353 K (lower panel).

Figure S4: Correlation diagram relating the mole fraction of x- and y-residues of the complete set of earlier investigated GxyG peptides at T=293K (left) and T=353 K. Upper panel: polyproline II versus β-strand fraction of neighbor, lower panel: β-strand versus β-strand of neighbor. Solid lines result from the regression analysis described in the text. The obtained Pearson coefficients are 0.41 and 0.28 for T=293 K and 0.08 and 0,04 for T=353 K. These values imply that no significant correlation could be inferred from the complete data set, as indicated in the main text.

Figure S5: δG_{P2β} (black) and δG_{βP2} (red) values of GKVG at the indicated temperature calculated as a function of $\delta G_{\beta\beta}$ as described in the main text.

Figure S6a: Mole fractions of indicated conformations of the central residue dimer of GxyG (abscissa) calculated with the NNI-model described in the text ("explicit") compared with mole fractions of indicated conformations of the central residue dimer of GxyG (abscissa) calculated using the thermodynamic parameters reported by Toal et al.(1) ("non-explicit") This calculation does not assume any conformation specific interactions.

Figure S6b: Population of different conformations of selected GxyG peptides. The selection was made based on the statistical error of the respective interaction parameters.

Figure S7a. Correlation plot of calculated pPII fractions of residue x and β-strand fractions of the y-neighbor based on the NNI theory described in the main manuscript. Parameters: $\Delta G_{P2\beta,1}$ = 1.0*kJ* / mol; $\Delta G_{P2t,1}$ = $\Delta G_{P2t,2}$ 3.0*kJ* / mol , $\Delta G_{P2\beta,2}$ was varied between -4.0 and 4.0 kJ/mol. $\delta G_{_{P2(1)\beta(2)}}$ = 2.0*kJ / mol*, $\delta G_{_{\beta(1)P2(2)}}$ = 2.0,4.0,6.0 and 8.0 kJ/mol for the plots in the upper left, upper right, lower left and lower right, respectively.

Figure S7b: Correlation plot of calculated pPII fractions of residue x and β-strand fractions of the y-neighbor based on the NNI theory described in the main manuscript. Parameters: $\Delta G_{P2\beta,1}$ = 1.0*kJ* / mol; $\Delta G_{P2t,1}$ = $\Delta G_{P2t,2}$ 3.0*kJ* / mol , $\Delta G_{P2\beta,2}$ was varied between -4.0 and 4.0 kJ/mol. $\delta G_{_{P2(1)}\beta(2)}$ was stochastically varied between 0 and 5 kJ/mol and $\delta G_{_{\beta(1)P2(2)}}$ between 0 and 3 kJ/mol.

Figure S7c: Correlation plot of calculated pPII fractions of residue x and β-strand fractions of the y-neighbor based on the NNI theory described in the main manuscript. Parameters: $\Delta G_{P2\beta,1} = \Delta G_{P2\beta,2} = 1.0 kJ/mol; \Delta G_{P2t,1} = \Delta G_{P2t,2} 3.0 kJ/mol; \quad \delta G_{pPII(1)\beta(2)}$ and $\delta G_{\beta(1)pPII(2)}$ was stochastically varied between 0 and 8 kJ/mol, $\Delta G_{_{P}{P}I\!I\!t,1}$ and $\Delta G_{_{P}{P}I\!I\!t,2}$ between -1 and 2 kJ/mol. The calculation was carried out for a limited set of 30 data points. The linear correlation fits (solid lines) yielded 0.98 for pPII and 0.77 for βstrand. $\delta G_{\beta(1)\beta(2)}$ between -3 and 8 kJ/mol,

Figure S8: Mole fractions of the total pPII (solid), total β-strand (dotted), total turn-like content (long dashed) as well as of pPII-β (short dashed) and, pPII-β-pPII-β (dashed dot) segments of MAX3 plotted as function of temperature.

References

1. Toal, S. E., N. Kubatova, C. Richter, V. Linhard, H. Schwalbe, and R. Schweitzer-Stenner. 2015. Randomizing the Unfolded State of Peptides (and Proteins) by Nearest Neighbor Interactions between Unlike Residues. Chem. Eur. J. 21:5173-5192.

2. Graf, J., P. H. Nguyen, G. Stock, and H. Schwalbe. 2007. Structure and dynamics of the homologous series of alanine peptides: a joint molecular dynamics/NMR study. J. Am. Chem. Soc. 129:1179-1189.

3. Hagarman, A., T. J. Measey, D. Mathieu, H. Schwalbe, and R. Schweitzer-Stenner. 2010. Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I' band profiles and NMR scalar coupling constants. J. Am. Chem. Soc. 132:540-551.

4. Hagarman, A., D. Mathieu, S. Toal, T. J. Measey, H. Schwalbe, and R. Schweitzer-Stenner. 2011. Amino acids with hydrogen-bonding side chains have an intrinsic tendency to sample various turn conformations in aqueous solution. Chemistry 17:6789-6797.

5. Schwalbe, M., V. Ozenne, S. Bibow, M. Jaremko, L. Jaremko, M. Gajda, M. R. Jensen, J. Biernat, S. Becker, E. Mandelkow, M. Zweckstetter, and M. Blackledge. 2014. Predictive Atomic Resolution Description of Intrinsically Disordered hTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. Structure 22:1-12.

6. Rybka, K., S. E. Toal, D. J. Verbaro, D. Mathieu, H. Schwalbe, and R. Schweitzer-Stenner. 2013. Disorder and order in unfolded and disordered peptides and proteins: a view derived from tripeptide conformational analysis. II. Tripeptides with short side chains populating asx and beta-type like turn conformations. Proteins 81:968-983.

7. Schweitzer-Stenner, R., A. Hagarman, S. Toal, D. Mathieu, and H. Schwalbe. 2013. Disorder and order in unfolded and disordered peptides and proteins: a view derived from tripeptide conformational analysis. I. Tripeptides with long and predominantly hydrophobic side chains. Proteins 81:955-967.