Supplementary Information to:

50 Years of Lifson-Roig Models: Application to Molecular Simulation Data

Andreas Vitalis[1]* and Amedeo Caflisch[1]

[1]*Department of Biochemistry*

University of Zurich

Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

*: To whom correspondence should be addressed:

Andreas Vitalis: Tel: +41446355597, E-mail: a.vitalis@bioc.uzh.ch

Illustration of the standard Lifson-Roig model for a tetrapeptide

The partition function of the simplest Lifson-Roig model for a tetrapeptide is the sum of possible statistical weights for each realizable sequence. The latter are listed in Table S1:

State	Weight	N_s	N_h	N_l	State	Weight	N_s	N_h	N_l
cccc	$u_{33}u_{33}u_{33}u_{33}$	0	0	0	hcch	$v_{32}u_{23}u_{33}v_{32}$	0	0	2
ccch	$v_{32}u_{23}u_{33}u_{33}$	0	0	1	hchc	$v_{32}u_{23}v_{32}u_{23}$	0	0	2
cchc	$u_{33}v_{32}u_{23}u_{33}$	0	0	1	hhcc	$v_{31}v_{12}u_{23}u_{33}$	1	0	0
chcc	$u_{33}u_{33}v_{32}u_{23}$	0	0	1	chhh	$u_{33}v_{31}wv_{12}$	1	1	0
hccc	$u_{33}u_{33}u_{33}v_{32}$	0	0	1	hchh	$v_{32}u_{23}v_{31}v_{12}$	1	0	1
cchh	$u_{33}u_{33}v_{31}v_{12}$	1	0	0	hhch	$v_{31}v_{12}u_{23}v_{32}$	1	0	1
chch	$u_{33}v_{32}u_{23}v_{32}$	0	0	2	hhhc	$v_{31}wv_{12}u_{23}$	1	1	0
chhc	$u_{33}v_{31}v_{12}u_{23}$	1	0	0	hhhh	$v_{31}wwv_{12}$	1	2	0

Table S1: A tetrapeptide has 2^4=16 unique conformational states (indices for distinguishing residues have been dropped for clarity, and subscripts indicate matrix elements as in equation S1). The values for N_s, N_h and N_l defined in the main text are listed for each state along with its statistical weight that is computed as the product of the weights for individual residues. Obviously, N_l is not or negatively correlated to the other two quantities. Positive coupling between N_s and N_h is expected, and becomes even larger if two-residue segments are discarded.

The partition function is simply the sum of all the weights. The transfer matrix approach is a representation of Z as matrix products, and allows – in its original form – to only compute properties that saliently emerge from sequences of length 3:

$$Z=\begin{pmatrix}0\\1\\1\end{pmatrix}^{T}\cdot\begin{pmatrix}{}^{1}w & {}^{1}v_{12} & 0\\ 0 & 0 & {}^{1}u_{23}\\ {}^{1}v_{31} & {}^{1}v_{32} & {}^{1}u_{33}\end{pmatrix}\cdot\begin{pmatrix}{}^{2}w & {}^{2}v_{12} & 0\\ 0 & 0 & {}^{2}u_{23}\\ {}^{2}v_{31} & {}^{2}v_{32} & {}^{2}u_{33}\end{pmatrix}\cdot\begin{pmatrix}{}^{3}w & {}^{3}v_{12} & 0\\ 0 & 0 & {}^{3}u_{23}\\ {}^{3}v_{31} & {}^{3}v_{32} & {}^{3}u_{33}\end{pmatrix}\cdot\begin{pmatrix}{}^{4}w & {}^{4}v_{12} & 0\\ 0 & 0 & {}^{4}u_{23}\\ {}^{4}v_{31} & {}^{4}v_{32} & {}^{4}u_{33}\end{pmatrix}\cdot\begin{pmatrix}0\\1\\1\end{pmatrix} \qquad \text{(S1)}$$

From the resultant weights in Table S1, equations 4 and 5 in the main text become obvious. Such

partial derivatives of Z with respect to specific elements in the matrices corresponds to isolating different contexts in those sequences of length 3. For instance, elements v_{12} will always be followed by a *u*-weight, and will be preceded by either *w* or v_{31}. It therefore corresponds to an end of a stretch of at least two residues in helical conformation ending at the third residue. Similarly, the context around elements v_{32} is always that of two residues in coil conformation (u_{23} or u_{33}). If we drop the residue identity labels, equation 5 in the main text becomes explicitly:

$$\frac{\delta \ln Z}{\delta \ln v_{32}} = \frac{v_{32}}{Z} \cdot \left(2u_{23}v_{31}v_{12} + 3u_{33}^2 u_{23} + u_{33}^3 + 2u_{23}^2 v_{32} + 4u_{23}u_{33}v_{32} \right) = Z^{-1} \sum\nolimits_k W_k \cdot N_1(k) = \langle N_1 \rangle \quad \text{(S2)}$$

In equation S2, W_k is the weight of state *k* as found in Table S1 for the 16 states.

Supplementary Methods

Description of Computational Model:

The ABSINTH continuum solvation model (see Vitalis and Pappu, *J. Comput. Chem.* 2009, **30**, 673-699, and references therein) describes the process of aqueous solution by formally decomposing large solute molecules into building blocks that correspond to small molecules for which transfer free energies from the gas phase into aqueous solution have been measured. Using a volume-based approach, all atomic solvent accessibilities are computed from the positions of all explicitly represented atoms in the vicinity. The resultant solvation states are used to create a weighted average over all the atoms comprising a specific solvation group. From this, a group solvent accessibility (normalized to an interval between 0 and 1) can be computed, and is multiplied with the reference transfer free energy for the underlying model compound to obtain the contribution this particular group makes to the total direct mean-field interactions (DMFI) between solute and water. The DMFI encapsulates polar and nonpolar terms and utilizes explicitly quantities that can be measured experimentally. Dielectric screening is handled in a way similar to generalized Born-type (GB) models of solvation using again volume-based measures to derive the effective dielectric acting between two (partial) charges. The main difference between GB models and ABSINTH is the different decomposition. GB models capture polar components of the DMFI as well as dielectric screening within the same framework, but require a separate model to describe any nonpolar contributions. ABSINTH captures all components of the DMFI within the same framework, but requires an additional model for dielectric screening.

The force field paradigm used in conjunction with the ABSINTH model differs from standard protocols in that 1) dihedral angle potentials are considered only to maintain planarity at bonds that are electronically restricted (like the peptide bond); and 2) short-range electrostatic interactions are pruned in an effort to maintain only those interactions that correspond to pairwise terms between charge groups

(usually net neutral units) that are separated by enough bonds such that nonbonded energy functions apply. Parameters are taken from OPLS-AA/L (Kaminski *et al*., *J. Phys. Chem. B*, 2001, **105**, 6474-6487 → partial charges and bonded parameters), were specifically reparameterized (Lennard-Jones radii and dispersion parameters), or are available directly from experiments (group reference free energies of solvation).

Monte Carlo simulations:

The FS-peptide was simulated for $50x10^6$ elementary steps, the first $20x10^6$ of which were discarded as equilibration. The employed moveset was taken similar to prior work (Vitalis and Caflisch, *J. Mol. Biol.* 2010, **403**, 148-165). At very low temperatures sampling is difficult, and we therefore only obtained data for the replica exchange schedule at higher temperatures. The underlying range of temperatures (260-440K) is identical to prior MC simulations on this system (Vitalis and Pappu, *J. Comput. Chem.* 2009, **30**, 673-699). All other settings were chosen identically to those for the molecular dynamics runs. Results from the Monte Carlo simulations are shown in Figure S1 only. While it may have been an obvious choice to utilize MC results as the "rigidified" model throughout, we were specifically interested in being able to systematically alter the applied constraints. We also wanted to sidestep possible concerns regarding comparisons between sets of data relying on fundamentally different methodologies. Therefore, the MC data serve exclusively to demonstrate that mass-metric tensor artifacts are quantitatively much smaller than the differences imposed by the constraints *per se*.

Derivation of Properties of the Simplified Model

The simplified model as proposed in equations 7 and 8 in the main text has the partition function:

$$Q=1+v\cdot N_r+v^2\cdot(N_r-1)+\sum_{i=2}^{N_r-1} v^2 w^{i-1}\cdot(N_r-i) \tag{S3}$$

For convenience, we work in reduced units given that all processes are unimolecular (normalization by the activity of the all-coil state). The average number of hydrogen bonds is obtained as:

$$\langle N_h\rangle=\frac{\partial \ln Q}{\partial \ln w}=\frac{w}{Q}\cdot\frac{\partial Q}{\partial w}=\frac{w}{Q}\sum_{i=2}^{N_r-1} v^2 w^{i-2}\cdot(i-1)\cdot(N_r-i) \tag{S4}$$

This is equivalent to the result in equation 8 in the main text. The generalized equilibrium constant Θ_m is defined as follows:

$$\Theta_m=\frac{\langle N_h\rangle}{N_r-2-\langle N_h\rangle}=\frac{f_h}{1-f_h} \quad \text{with} \quad f_h=\frac{\langle N_h\rangle}{N_r-2} \tag{S5}$$

The one derivative we will inevitably need to characterize the dependence of Θ_m on w is the following:

$$\begin{aligned}
&\frac{\partial\langle N_h\rangle}{\partial w}=\frac{\partial}{\partial w}[\frac{1}{Q}\sum_{i=2}^{N_r-1} v^2 w^{i-1}\cdot(i-1)\cdot(N_r-i)] \\
&\frac{\partial\langle N_h\rangle}{\partial w}=\frac{1}{Q}\cdot[\sum_{i=2}^{N_r-1} v^2 w^{i-2}\cdot(i-1)^2\cdot(N_r-i)] \\
&-\frac{1}{Q^2}\cdot[\sum_{i=2}^{N_r-1} v^2 w^{i-1}\cdot(i-1)\cdot(N_r-i)]\cdot[\sum_{i=2}^{N_r-1} v^2 w^{i-2}\cdot(i-1)\cdot(N_r-i)] \\
&\frac{\partial\langle N_h\rangle}{\partial w}=\frac{1}{wQ}\cdot[\sum_{i=2}^{N_r-1} v^2 w^{i-1}\cdot(i-1)^2\cdot(N_r-i)]-\frac{1}{w}\langle N_h\rangle^2
\end{aligned} \tag{S6}$$

Given equation S4, it should be apparent that – excepting the w in the denominator - the first term on the right-hand side in the last line of equation S6 corresponds to the ensemble average of the square of the number of hydrogen bonds, $\langle N_h^2\rangle$. Hence:

$$\frac{\partial \langle N_h \rangle}{\partial w} = \frac{1}{w} \cdot [\langle N_h^2 \rangle - \langle N_h \rangle^2] \tag{S7}$$

The direct derivative and thereby the logarithmic derivative are directly proportional to the second central moment of the distribution of the number of hydrogen bonds. The logarithmic derivative of Θ_m with w is as follows:

$$\begin{aligned} \frac{\partial \ln \Theta_m}{\partial \ln w} &= \frac{w}{\Theta_m} \cdot \frac{\partial \Theta_m}{\partial w} = \frac{w}{\Theta_m} \cdot \frac{\partial f_h}{\partial w} \cdot (1 - f_h)^{-2} = \frac{w}{\Theta_m} \cdot \frac{\partial \langle N_h \rangle}{\partial w} \cdot \frac{(1 - f_h)^{-2}}{N_r - 2} \\ \frac{\partial \ln \Theta_m}{\partial \ln w} &= \frac{\langle N_h^2 \rangle - \langle N_h \rangle^2}{\langle N_h \rangle \cdot (1 - f_h)} \end{aligned} \tag{S8}$$

Unfortunately, it does not appear trivial to obtain analytical expressions for maxima in the function in equation S8. We address this numerically in Figure S5 where we show maximal values for this function using values of v and N_r that cover a regime that corresponds to what has been tested and proposed *in vitro*. Due to the use of the simplified model, this should be an upper bound estimate as can be gleaned from Figure 5 in the main text. The most realistic values for v (0.05 and 0.15) yield maximum values for the derivative in equation S8 that for short chain lengths are indeed close to (N_r-2)/2 as stated in the main text.

Details on the Partition Function Implied by Equation 9 in the Main Text

Equation 9 constructs a new model by considering a weighted hybrid of two alternative LR models as follows:

$$Z_{total} = Z_{N_r=7}^{3f_3} Z_{N_r=21}^{(1-f_3)} \quad \text{(S9)}$$

Here, $Z_{Nr=21}$ is the standard partition function for a peptide of 21 residues. It obeys the standard LR rules, *i.e.*, the statistical weights are assigned as demonstrated above (see Table S1). Similarly, $Z_{Nr=7}$ is the standard LR partition function for a peptide of 7 residues. Individually, it again obeys the rules outlined above. In the limiting case of f_3 approaching unity, the total partition function of the modified model is that of three independent helical peptides that are each 7 residues long. This implies both a quenching of states that are representable in $Z_{Nr=21}$, and an extension to new states. The quenched ones are those possessing helical segments longer than 7 residues (in the LR partition function, terminal residues can never be assigned state *w*). The maximally helical sequence possible is "$(vwwwwwv)^3$", which can be thought of as "*vwwwwwv*|*vwwwwwv*|*vwwwwwv*" when mapped back onto the 21 residues of the actual peptide (vertical lines indicating boundaries between 7-residue segments). This particular sequence is actually an example of an extension of the original partition function, as it is not realizable in $Z_{Nr=21}$. This is because the two consecutive *v*-residues would automatically be turned into state *w*, *i.e.*, $Z_{Nr=21}$ only allows helix interruptions if *u*-residues are in between. The reason that the first term on the right-hand side of equation S9 is beneficial to the fitting (see Figures 7 in the main text and S6) lies in the fact that it allows states with both high helicity and larger values for the number of helical segments to be represented without requiring very large values of *v*. The sequence above would be counted as $N_h = 15$ and $N_s = 3$. Conversely, in the coil phase, the two limiting partition functions in equation S9 (f_3 approaching either zero or unity) become indistinguishable, *i.e.*, sequences like

"*uvwwwwvuuuuuuuuuuuuuuu*" are invariant whether they are represented in the framework of $Z_{Nr=21}$ or $Z_{Nr=7}^3$. To see this, consider explicitly the sum entering $Z_{Nr=7}$.:

$$Z_{N_r=7}^3=\left[u^7+7u^6v+42u^5v^2+\ldots+2uv^2w^4+v^2w^5\right]^3 \tag{S10}$$

The first term in equation S10 is the all-coil state, the second and third terms are those with just one or two residues in helical conformation (no hydrogen bond formed). Terms with 1, 2, or 3 hydrogen bonds are then omitted due to the complicated combinatorial factors involved. The second last term corresponds to the exactly two ways of having four hydrogen bonds (the single *u*-residue has to be terminal), and the last term is the aforementioned all-helical conformation with 5 hydrogen bonds. It is important to note that the exponentiation by 3 generates all the possible cross-terms, *i.e.*, "*uvwwwwv*|*uuuuuuu*|*uuuuuuu*" is representable within the scope of $Z_{Nr=7}^3$. Deriving properties from the hybrid partition function is simple:

$$\begin{aligned}\frac{\partial \ln\left(Z_{N_r=7}^{3f_3}Z_{N_r=21}\right)}{\partial \ln w}&=\frac{\partial}{\partial \ln w}\left[3f_3\ln Z_{N_r=7}+\left(1-f_3\right)\ln Z_{N_r=21}\right]\\ \frac{\partial \ln\left(Z_{N_r=7}^{3f_3}Z_{N_r=21}\right)}{\partial \ln w}&=3f_3\frac{\partial \ln Z_{N_r=7}}{\partial \ln w}+\left(1-f_3\right)\frac{\partial \ln Z_{N_r=21}}{\partial \ln w}=\langle N_h\rangle\end{aligned} \tag{S11}$$

Equation S11 is identical to equation 9 in the main text. It shows that due to the logarithmic derivatives, ensemble averages can simply be derived using properties of the underlying standard LR partition functions for the two different peptide lengths. Analogous equations hold for $\langle N_s\rangle$ and $\langle N_l\rangle$.

Supplementary Figure Captions

Figure S1: Comparison to data from Monte Carlo (MC) Simulations. This figure is identical to Figure 1 in the main text with the two exceptions that data from MC simulations are added, and that Panel B lacks the uncorrected DSSP-derived values for $\langle N_s \rangle$. The data demonstrate that differences between the MC data and those in the presence of backbone bond angle constraints are negligible for $\langle N_s \rangle$ and $\langle N_h \rangle$ except at the four lowest temperatures. Small, systematic deviations are seen for $\langle N_l \rangle$. The comparison demonstrates that the majority of differences between flexible and rigidified cases indeed stems from constraints themselves, and not from artefactual sources, *i.e.*, mass-metric tensor artifacts (see Methods in the main text). The small deviations are just as likely a result of the additional constraints present.

Figure S2: Pair correlation functions for sodium$^+$ *vs*. chloride$^-$ ions (green colors), the peptide arginine$^+$ sidechain carbon atom of the guanidino group *vs*. sodium$^+$ ions (red colors), and for the peptide arginine$^+$ sidechain carbon atom of the guanidino group *vs*. chloride$^-$ ions (blue colors). Panel **A** shows data for the case with flexible backbone bond angles for a few selected temperatures as indicated, and Panel **B** does the same for the case with rigidified backbone bond angles (data are taken from low temperature replica-exchange simulations only). Peptide-ion pair correlation functions do not converge exactly to 1.0 because of the boundary condition (spherical droplet with soft wall), and – more importantly – the finite volume of the peptide. Sodium$^+$ is preferentially excluded from the vicinity of the arginine$^+$ sidechains, while chloride$^-$ ions are preferentially found around the peptide, but show very little direct binding (small peak at ~5Å). Ion-ion pair correlation functions reveal a weak direct interaction peak (~4Å), followed by a desolvation barrier, and a solvent-separated peak (~7.5Å). Direct contact formation is expected to be much stronger in explicit solvent (see for example Chen, A. A. and Pappu, R. V. *J. Phys. Chem. B*, 2007, **111**, 6469-6478) meaning that possible processes that are

explicitly mediated by direct ion binding would not be well described. Data are overall only very weakly dependent on temperature.

Figure S3: Quality of fits of LR theory to helical content data as a function of temperature for the FS-peptide when using a temperature-independent value for the nucleation parameter ($v = 0.127$). Data shown as solid lines in Panels A-C are identical to those provided in Figure 1, and the same color code applies. Fitted values according to LR theory are shown as symbols only (using equations 4 and 5 in the main text). In contrast to Figure 3, all three readouts were fit to simultaneously ($\langle N_h \rangle$, $\langle N_s \rangle$, and $\langle N_l \rangle$), and only w was allowed to vary during the fitting procedure. Panel B uses the same legend as Panel A. Panel D shows the underlying values for w as a function of temperature. This plot is analogous to Figure 4B in the main text, and again uses the same color code.

Figure S4: Quality of fits of LR theory to helical content data as a function of temperature for the FS-peptide when including data for $\langle N_l \rangle$ in the fitting. This figure is exactly analogous to Figures 3 (Panels A-C) and 4 (Panels D-E) in the main text with the exception that data for $\langle N_l \rangle$ were included in the fitting. Panel B uses the same legend as Panel A, and Panel E uses the same legend as Panel D.

Figure S5: Analysis of maximal slopes in double logarithmic plots of the generalized equilibrium constant, Θ_m, as a function of the propagation parameter w. Values are plotted as a function of peptide length for five different values of the nucleation parameter, v. The maximal slopes were obtained by taking numerical derivatives of data as shown in Figure 5 in the main text. This implies that w-values ranging from 0.355 to 3.345 were used that utilize a discretization interval of 0.01. As can be seen from Figure 5, the underlying function is generally smooth and exhibits a well-defined region of maximal slopes for values of Θ_m in the vicinity of 10^{-1}. The dotted and dashed black lines are lines with slopes of N_r-2 and (N_r-2)/2, respectively. It is obvious that the case of N_r-2 is never reached even if the nucleation

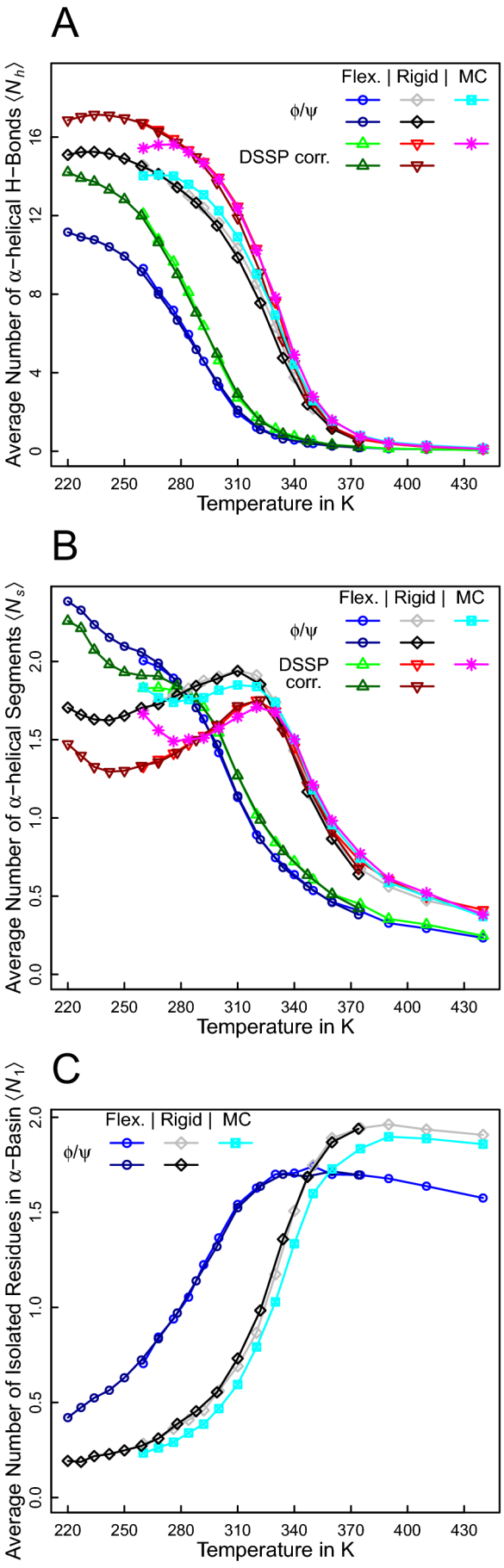
parameter is vanishingly small.

Figure S6: Quality of fits of a modified LR model to helical content data as a function of temperature for the FS-peptide . This figure is identical to Figure 3 in the main text with the exception that data for $\langle N_1 \rangle$ were included in the fitting, and that equation 9 and analogs for $\langle N_s \rangle$ and $\langle N_1 \rangle$ were used instead of equations 4-5.

Figure S7: Considering alternative models for fitting temperature-dependent helix-coil equilibrium data. As an example, we can construct a partition function allowing for two helices of variable length (compare to equation S9): $Z_{total} = Z_{N_r=N}^{f_2} Z_{N_r=N_r-N}^{f_2} Z_{N_r=21}^{(1-f_2)}$

To avoid overfitting such a model, some values have to be provided by independent means. Here, we choose to set the nucleation parameter directly to the values from dipeptide data shown in Figure 7 of the main text. The three fitted quantities are then the subsegment length N, the fractional occupancy f_2 (both in Panel **A**), and the propagation parameter w (Panel **B**). As can be seen, the fitting quality is unsatisfactory when compared to the data in Figures 7 and S6, in particular for flexible backbone bond angles (Panels **C**-**E**). As with the model in equation 9, values for f_2 (and similarly N) become meaningless in the coil regime. This is because peptide annotation strings with few and short helical segments are equally well-representable in both limiting cases of the partition function (*i.e.*, $f_2 \to 0$ and $f_2 \to 1$), and – by extension – also by intermediate values of f_2. As a peculiarity of the particular model explored here, $N \to 0$ will also recover the original LR partition function regardless of the value of f_2. Results are shown for ϕ/ψ statistics, but similar results regarding fitted quantities and fit quality are obtained with DSSP-based data.

Figure S1


A
Flex. | Rigid | MC
ϕ/ψ
DSSP corr.
Average Number of α−helical H−Bonds ⟨N_h⟩
0
4
8
12
16
220
250
280
310
340
370
400
430
Temperature in K
B
Flex. | Rigid | MC
ϕ/ψ
DSSP corr.
Average Number of α−helical Segments ⟨N_s⟩
0.0
0.5
1.0
1.5
2.0
220
250
280
310
340
370
400
430
Temperature in K
C
Flex. | Rigid | MC
ϕ/ψ
Average Number of Isolated Residues in α−Basin ⟨N_1⟩
0.0
0.5
1.0
1.5
2.0
220
250
280
310
340
370
400
430
Temperature in K

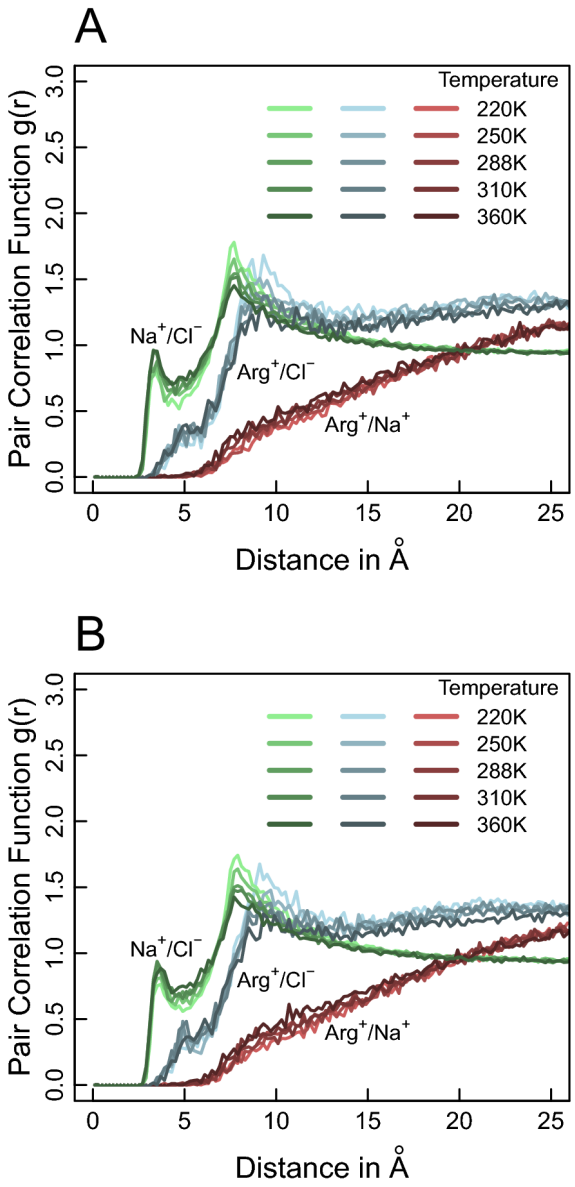

A
Temperature
220K
250K
288K
310K
360K
Na+/Cl−
Arg+/Cl−
Arg+/Na+
Pair Correlation Function g(r)
Distance in Å
B
Temperature
220K
250K
288K
310K
360K
Na+/Cl−
Arg+/Cl−
Arg+/Na+
Pair Correlation Function g(r)
Distance in Å

Figure S2

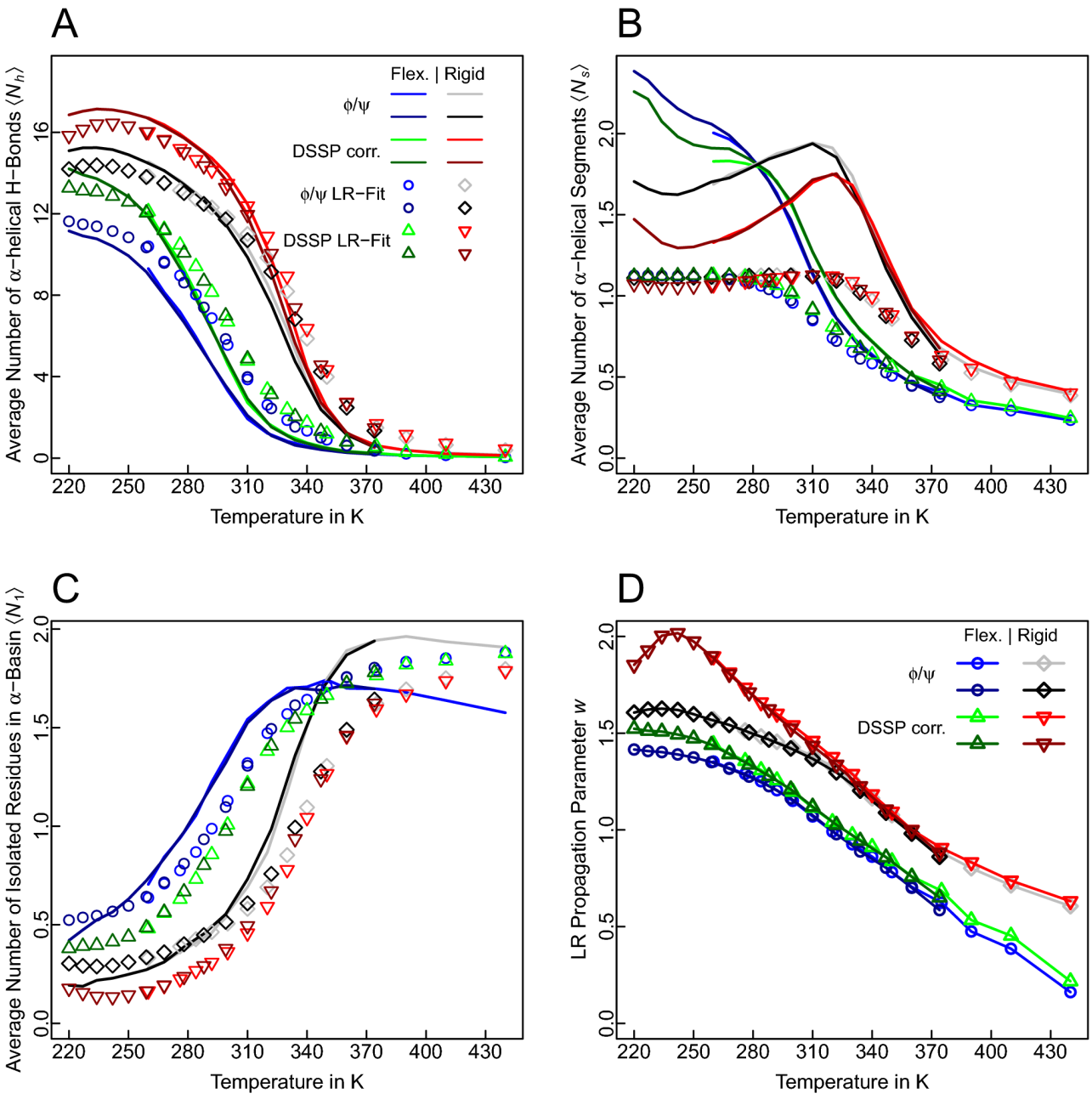

A
Flex. | Rigid
ϕ/ψ
DSSP corr.
ϕ/ψ LR−Fit
DSSP LR−Fit
Average Number of α−helical H−Bonds ⟨N_h⟩
Temperature in K
B
Average Number of α−helical Segments ⟨N_s⟩
Temperature in K
C
Average Number of Isolated Residues in α−Basin ⟨N_1⟩
Temperature in K
D
Flex. | Rigid
ϕ/ψ
DSSP corr.
LR Propagation Parameter w
Temperature in K

Figure S3

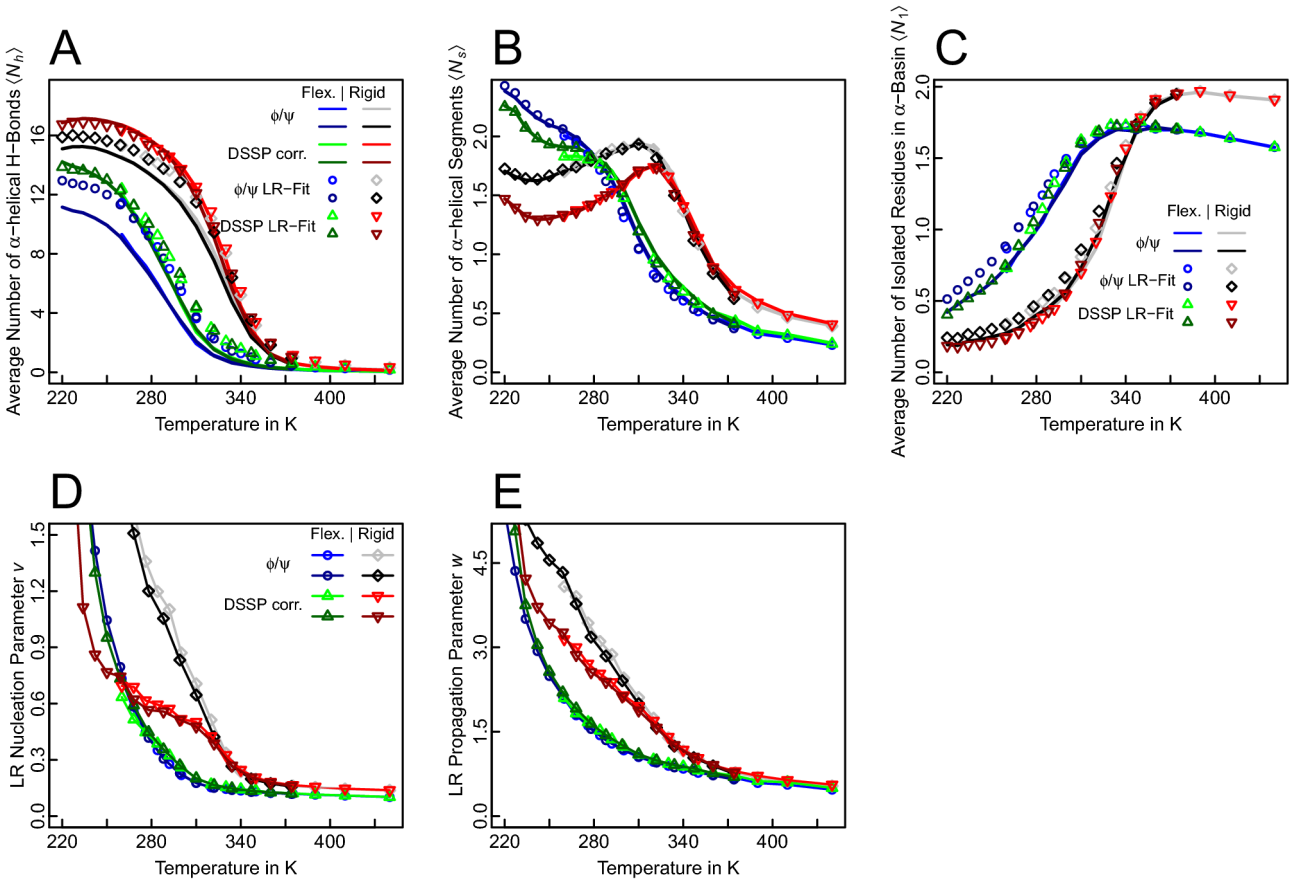

A
Average Number of α−helical H−Bonds ⟨N_h⟩
Flex. | Rigid
ϕ/ψ
DSSP corr.
ϕ/ψ LR−Fit
DSSP LR−Fit
Temperature in K
B
Average Number of α−helical Segments ⟨N_s⟩
Temperature in K
C
Average Number of Isolated Residues in α−Basin ⟨N_1⟩
Flex. | Rigid
ϕ/ψ
ϕ/ψ LR−Fit
DSSP LR−Fit
Temperature in K
D
LR Nucleation Parameter v
Flex. | Rigid
ϕ/ψ
DSSP corr.
Temperature in K
E
LR Propagation Parameter w
Temperature in K

Figure S4

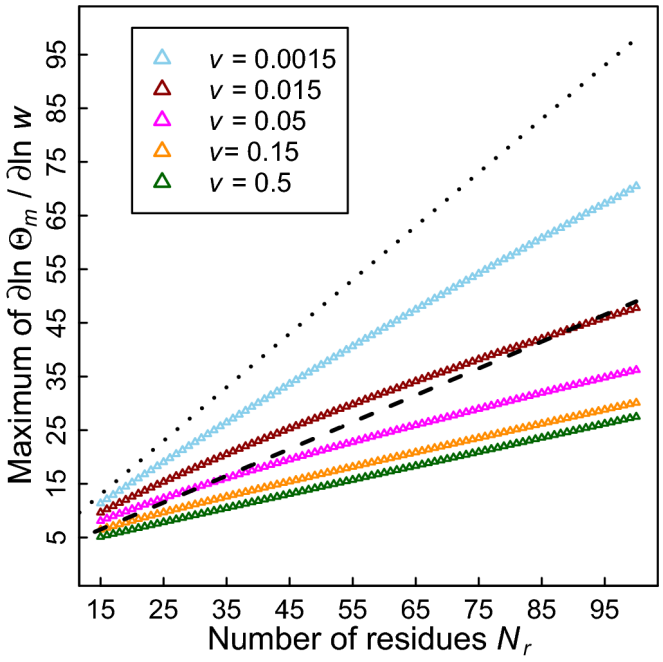
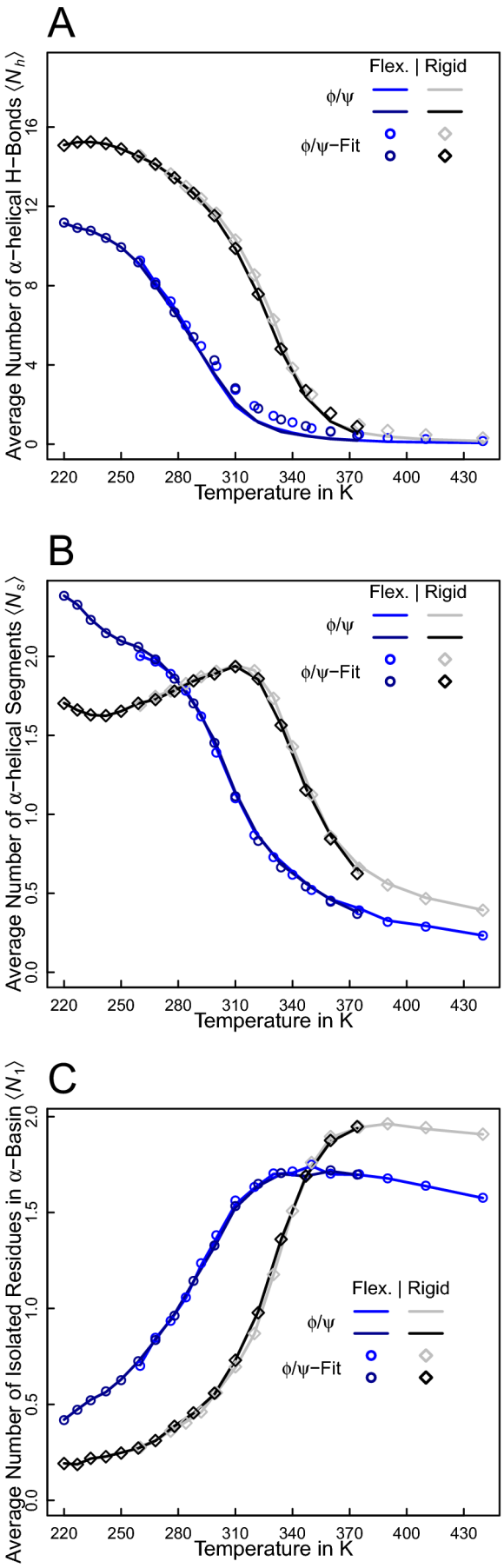

v = 0.0015
v = 0.015
v = 0.05
v= 0.15
v = 0.5
Maximum of ∂ln Θm / ∂ln w
Number of residues Nr
5
15
25
35
45
55
65
75
85
95

Figure S5

Figure S6


A
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit
Average Number of α−helical H−Bonds ⟨N_h⟩
0
4
8
12
16
220
250
280
310
340
370
400
430
Temperature in K
B
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit
Average Number of α−helical Segments ⟨N_s⟩
0.0
0.5
1.0
1.5
2.0
220
250
280
310
340
370
400
430
Temperature in K
C
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit
Average Number of Isolated Residues in α−Basin ⟨N_1⟩
0.0
0.5
1.0
1.5
2.0
220
250
280
310
340
370
400
430
Temperature in K

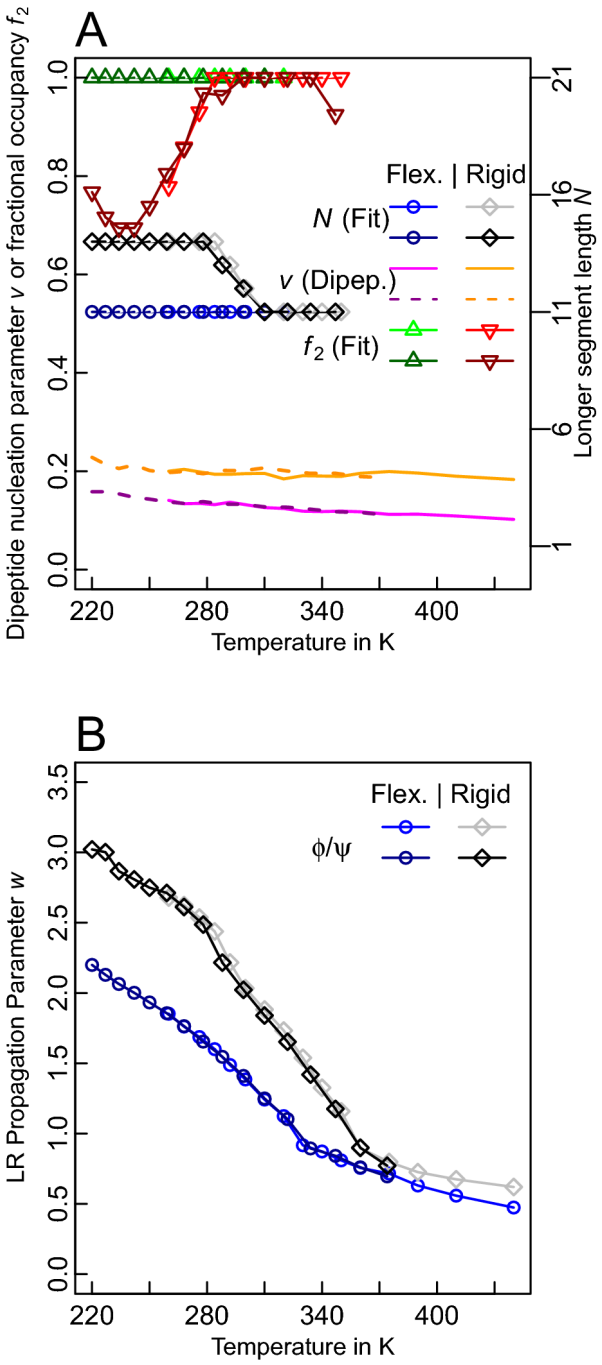

A
Dipeptide nucleation parameter v or fractional occupancy f_2
Longer segment length N
Temperature in K
Flex. | Rigid
N (Fit)
v (Dipep.)
f_2 (Fit)
B
LR Propagation Parameter w
Temperature in K
Flex. | Rigid
ϕ/ψ
C
Average Number of α−helical H−Bonds $\langle N_h \rangle$
Temperature in K
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit
D
Average Number of α−helical Segments $\langle N_s \rangle$
Temperature in K
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit
E
Average Number of Isolated Residues in α−Basin $\langle N_1 \rangle$
Temperature in K
Flex. | Rigid
ϕ/ψ
ϕ/ψ−Fit

Figure S7