Minimalist model for the dynamics of helical polypeptides: a statistic-based parameterization

Giulia Lia Beatrice Spampinato[1], Giuseppe Maccari[2], Valentina Tozzini[1]*

[1] NEST, Istituto Nanoscienze – CNR and Scuola Normale Superiore, Piazza San Silvestro 12-56127 Pisa, Italy; [2] Center for Nanotechnology and Innovation @NEST, Istituto Italiano di Tecnologia, Piazza San Silvestro 12-56127 Pisa, Italy

	3_{10}			α		π	
	(PERUTZ, 1951)	(Pauling et al., 1951)	(Crisma et al., 2006)	(Pauling et al., 1951)	(Crisma et al., 2006)	(Low and Baybutt, 1952)	(Fodje and Al-Karadaghi, 2002)
Φ (deg)	-49	-74	-57	-58	-63	-57	-76
Ψ (deg)	-26	-4	-30	-47	-42	-70	-41
n	3.0	3.0	3.24	3.6	3.63	4.4	4.4
d (Å)	1.93	2.0	1.94	1.5	1.56	1.14	1.2
p (Å)	5.8	6.0	6.29	5.4	5.67	5.02	5.28
(i,i+m), m=	3	3	3	4	4	5	5

Table S1. Topology and structural parameters of the different kind of helices. Backbone conformational parameters from the three different types of helices in all-atom representation. Φ and Ψ are the backbone dihedral angles, n is the number of residues per helical turn, d is the axial translation per residue and p is the pitch or axial translation for helical turn. The row (i,i+m) identifies how many residues separate two hydrogen bonded amino acids.

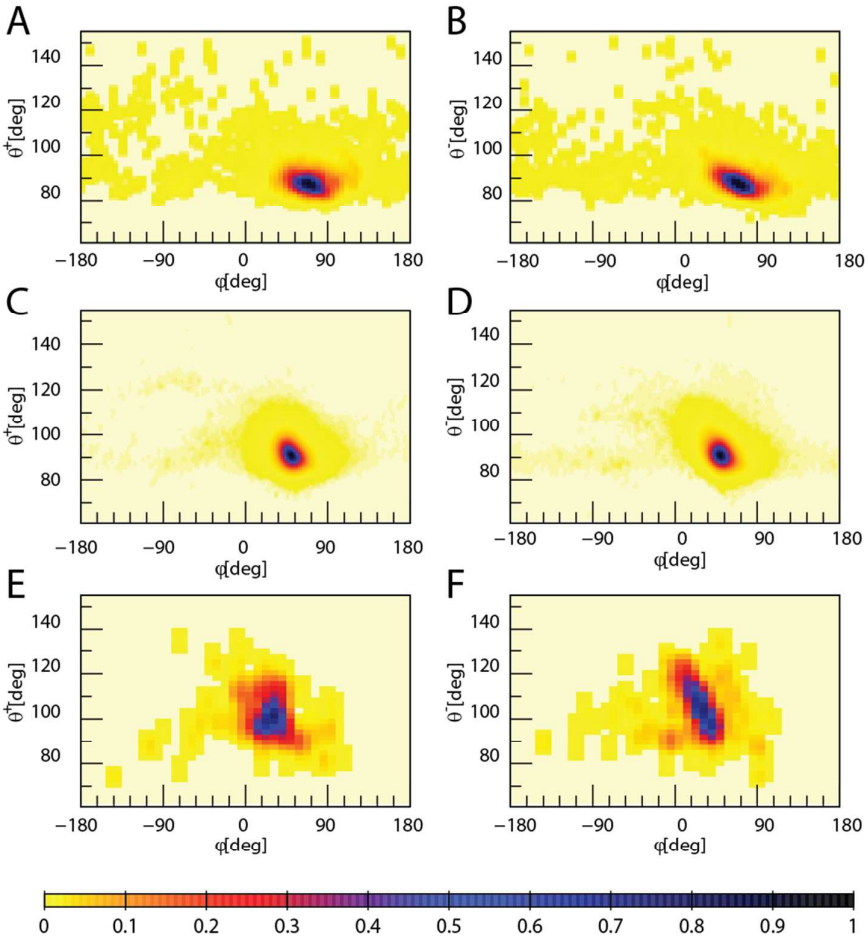

Figure S1. (θ+, φ) and (θ-, φ) correlation plots. A, B) 3_{10}-Helix; C,D) α-helix; E,F) π-helix. In this figure, the (θ^+,φ) and (θ^-,φ) plots are compared, evaluated from the final structural datasets**.** The comparison shows no statistically relevant difference, indicating that helical structures display substantial symmetry upon inversion of helix directionality sense. Thus, in the model and in the comparison with experiment the directionality is neglected and the (θ,φ) correlation plot, namely the average of (θ^+, φ) and (θ^-, φ) is used as reference for comparison between simulations and experiments.

Helix type	3_{10}-helix				α-helix				π-helix	
Experimental method	Xray		NMR		Xray		NMR		Xray	NMR
Sec Struct Identification	DSSP	PDB	DSSP	PDB	DSSP	PDB	DSSP	PDB	DSSP	DSSP
θ(deg)	89.7	89.7	88.11	88.1	91	91.35	90.8	90.8	100	100.33
φ(deg)	68	63	64	60	50.4	50.4	50.4	50.4	28.8	36
$r_{1\text{-}3}$(Å)	5.33	5.42	5.33	5.33	5.42	5.42	5.42	5.42	5.78	5.78
$r_{1\text{-}4}$ (Å)	5.75	5.6	5.61	5.33	5.15	5.15	5.15	5.15	5.96	5.6
$r_{1\text{-}5}$(Å)	8.2	7.76	8.12	8	6.14	6.14	6.14	6.05	4.88	4.52
$r_{1\text{-}6}$(Å)	10.2	10.2	9.92	9.92	8.66	8.66	8.57	8.66	6.32	

Table S2. Helical statistical parameters. Distribution parameters for the three different types of helices for the CG structures, obtained using Octave (Eaton et al., 2009). The values reported are the maximum peaks of the distributions. The first two rows contain the experimental method and the algorithm for the structure identification.

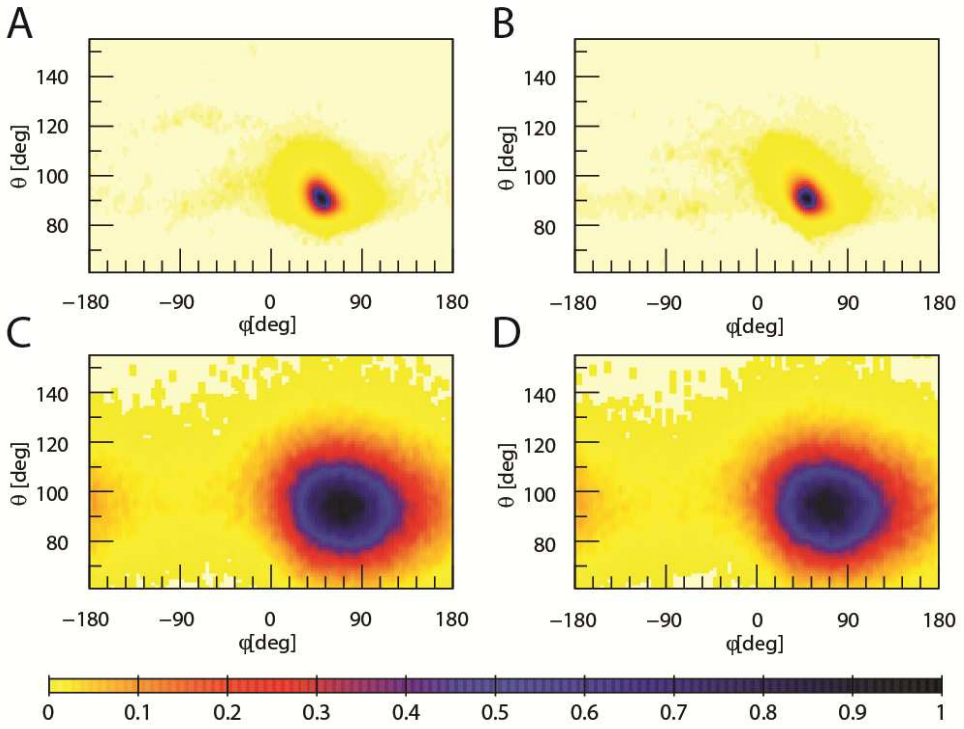

Figure S2. Comparison of (θ^{+-}, φ) experimental and simulation correlations plots of a 28mer without hydrogen bonds. (A,C) (θ^{-},φ) correlation map; (B,D) (θ^{+},φ) correlation map; A, B) experimental, from the α-helix dataset; C,D) simulation, from the equilibrium dynamics run without the hydrogen bonding terms. The absence of helical-stabilizing hydrogen bonding produces an anomalously large spreading and incorrect shape of the occupied area.

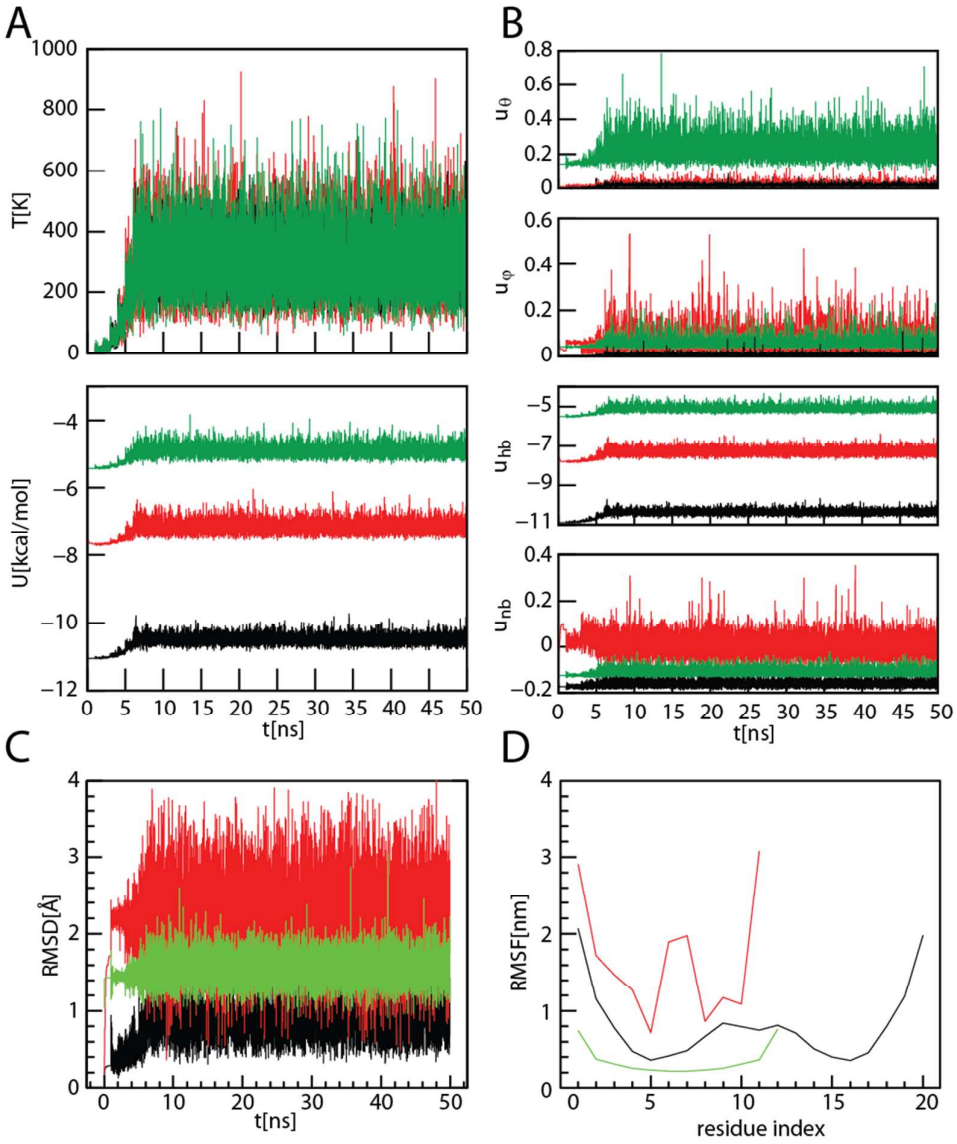

Figure S3. Equilibrium dynamics. Analysis of thermalization and production run of the three helices. Color codes: black=α-helix, red= 3_{10}-helix, green= π-helix. A) Temperature (top) and internal energy per residue (bottom); B) Separate FF terms energy contributions. From top to bottom: u_θ, u_φ, u_{hb}, u_{nb} (per residue);.C) Root mean squared deviation from the starting configuration; D) Simulation RMSF, evaluated in the production part (>10ns).

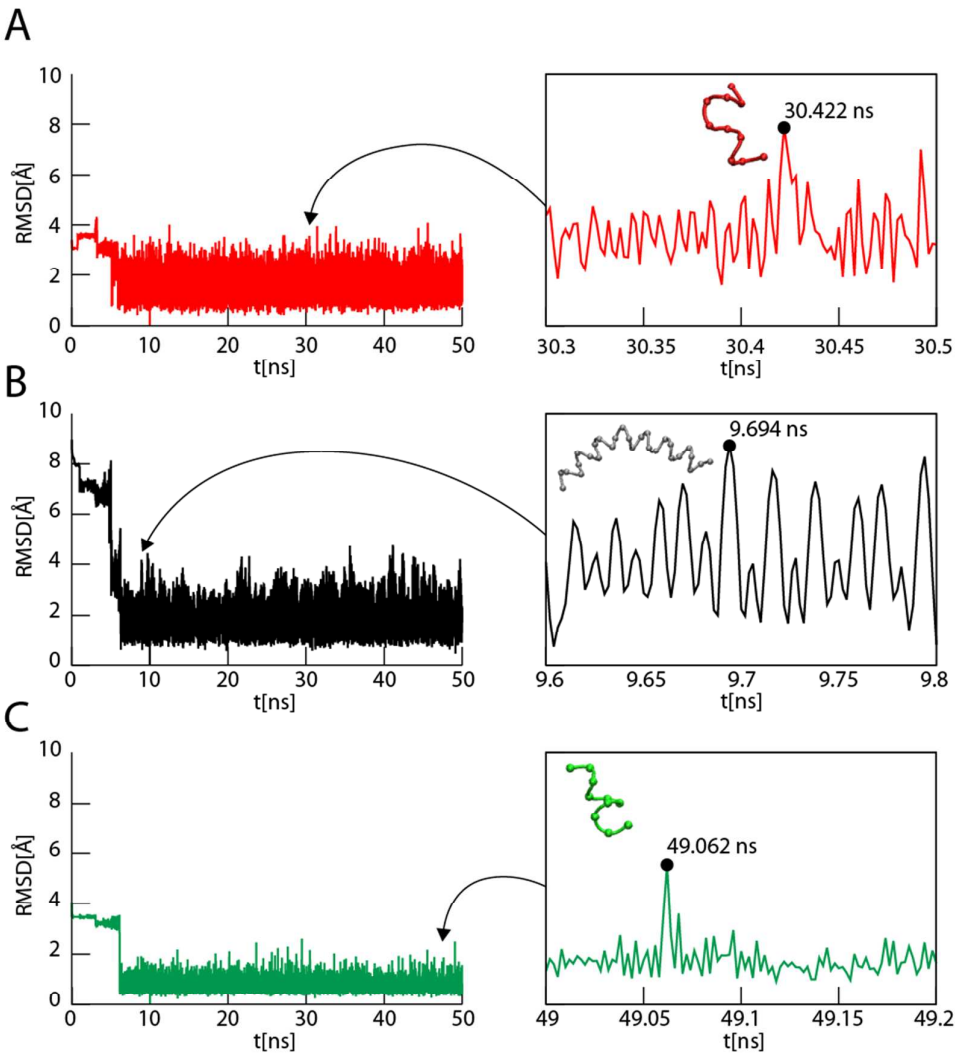

Figure S4. Unfolding events in folding dynamics. For each helical structure, the left panel shows the overall folding dynamic RMSD, while the right panel displays a sample event of unfolding during the simulation. A) 3_{10} helix; B) α-helix; C) π-helix. For α-helix, we show a kink forming event, rather than a real unfolding, which can be considered a precursor to the breaking of a long helix in two shorter one.

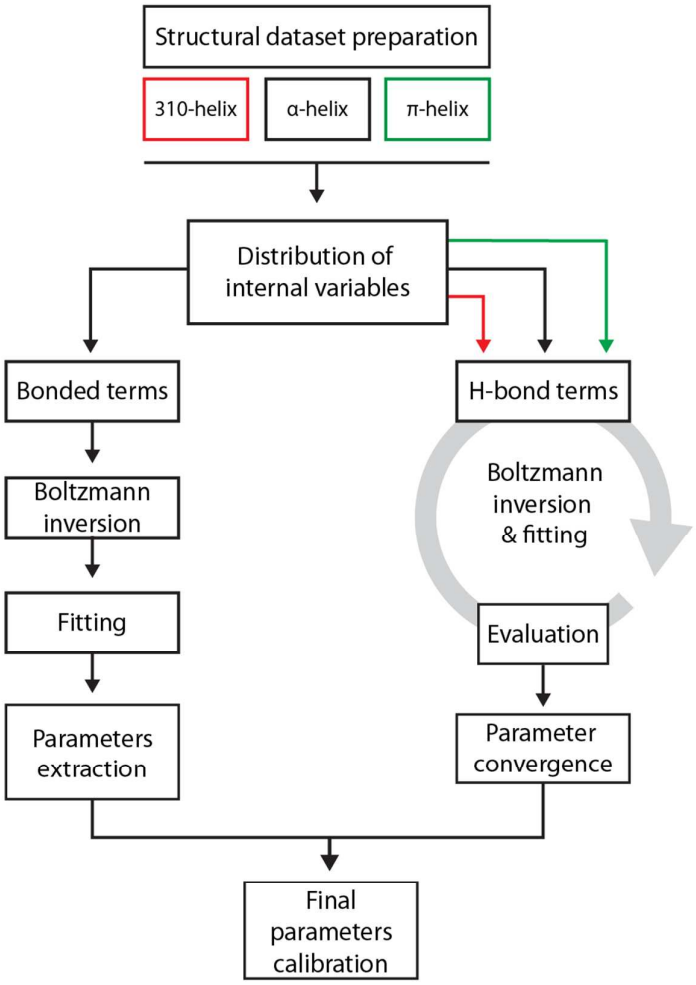

Figure S5. Parameterization flowchart. For each helical type, a structural dataset was compiled and internal variables distributions were calculated with SecStAnT. For the bonded terms (θ,φ), the parameters initial guesses were obtained fitting the direct Boltzmann inverted of the corresponding distribution with its functional form (see main text). For the h-bond terms, the first guesses were obtained with the same procedure, however separately for the three helices, including only the terms pertaining to each helix. The force constants are then separately optimized following an iterative BI procedure, comparing at each step all the internal variable distributions and correlations obtained by MD simulations.

Dataset preparation query parameters
Below, the query parameters for each dataset creation are listed. For each dataset, the original query is represented as given by the RCSB server.

- *310-helix XRAY*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=3-10 helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is X-RAY and Representative Structures at 30% Sequence Identity
- *310-Helix NMR*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=3-10 helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is SOLUTION NMR and Representative Structures at 30% Sequence Identity
- *α-helix XRAY*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=alpha helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is X-RAY and Representative Structures at 30% Sequence Identity
- *α-helix NMR*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=alpha helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is SOLUTION NMR and Representative Structures at 30% Sequence Identity
- *π-helix XRAY*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=pi helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is X-RAY and Representative Structures at 30% Sequence Identity
- *π-helix NMR*
  Chain Type: there is a Protein chain and SecondaryStructureLengthQuery: ss.method.value=DSSP helix.subtype.value=pi helix helix.length.comparator=between beta_strand.length.comparator=between and Experimental Method is SOLUTION NMR and Representative Structures at 30% Sequence Identity
- *Random Coil NMR*
  Chain Type: there is a Protein chain and Secondary structure has: 20 or less percent of elements are Alpha Helical and 20 or less percent of elements are Beta Sheet and Experimental Method is SOLUTION NMR and Representative Structures at 30% Sequence Identity

SUPPLEMENTAL REFERENCES

Crisma, M., Formaggio, F., Moretto, A., and Toniolo, C. (2006). Peptide helices based on alpha-amino acids. Biopolymers *84*, 3–12.

Eaton, J.W., Bateman, D., and Hauberg, S. (2009). {GNU Octave} version 3.0.1 manual: a high-level interactive language for numerical computations (CreateSpace Independent Publishing Platform).

Fodje, M.N., and Al-Karadaghi, S. (2002). Occurrence, conformational features and amino acid propensities for the -helix. Protein Eng. Des. Sel. *15*, 353–358.

Low BW, Baybutt RB. The Pi Helix--A Hydrogen Bonded Configuration of the Polypeptide Chain. J Am Chem Soc. 1952;*74*:5806–5807.

Pauling, L., Corey, R.B., and Branson, H.R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. Proc. Natl. Acad. Sci. *37*, 205–211.

PERUTZ, M.F. (1951). New X-Ray Evidence on the Configuration of Polypeptide Chains: Polypeptide Chains in Poly-γ-benzyl-L-glutamate, Keratin and Hæmoglobin. Nature *167*, 1053–1054.