Supplementary material

Geometry motivated alternative view on local protein backbone structures

Jan Zacharias and Ernst Walter Knapp
Institute of Chemistry and Biochemistry, Freie Universität Berlin, D-14195 Berlin, Germany
Running title: alternative to Ramachandran plot

The formulas (1) and (2) from the main text as found by Miyazawa[1] before insertion of the peptide plane parameters, assuming trans conformation for the peptide plane ($\omega = 180^0$) read

$$\begin{aligned} \cos\left(\frac{\vartheta}{2}\right) &= -(k_1 + k_3)\sin\left(\frac{\varphi+\psi}{2}\right) + (k_2 - k_4)\sin\left(\frac{\varphi-\psi}{2}\right) \\ d\sin\left(\frac{\vartheta}{2}\right) &= (m_1 k_1 + m_3 k_3)\cos\left(\frac{\varphi+\psi}{2}\right) - (m_2 k_2 + m_4 k_4)\cos\left(\frac{\varphi-\psi}{2}\right) \end{aligned} \qquad \text{(S1)}$$

For a definition of the helix parameters (d, ϑ, r) see Fig. S1.
The parameters m_i, k_i are functions of the bond angles τ_i and the bond lengths d_i. The latter are given by

$\tau_1 = \angle(N, C_\alpha, C) = 111.0°, \quad \tau_2 = \angle(C_\alpha, C, N) = 117.2°, \quad \tau_3 = \angle(C, N, C_\alpha) = 121.7°$

$d_1 = d(C_\alpha, C) = 1.525\text{Å}, \quad d_2 = d(C, N) = 1.336\text{Å}, \quad d_3 = d(N, C_\alpha) = 1.459\text{Å}$

according to [2, 3]. And the parameters m_i, k_i are defined as

$m_1 = +d_1 + d_2 + d_3 = 4.320\text{Å}, \quad m_2 = -d_1 + d_2 + d_3 = 1.270\text{Å}$

$m_3 = +d_1 - d_2 + d_3 = 1.648\text{Å}, \quad m_4 = +d_1 + d_2 - d_3 = 1.402\text{Å}$

$$k_1 = \sin\left(\frac{\tau_1}{2}\right)\sin\left(\frac{\tau_2}{2}\right)\sin\left(\frac{\tau_3}{2}\right) = 0.614, \quad k_2 = \cos\left(\frac{\tau_1}{2}\right)\cos\left(\frac{\tau_2}{2}\right)\sin\left(\frac{\tau_3}{2}\right) = 0.258$$

$$k_3 = \sin\left(\frac{\tau_1}{2}\right)\cos\left(\frac{\tau_2}{2}\right)\cos\left(\frac{\tau_3}{2}\right) = 0.209, \quad k_4 = \cos\left(\frac{\tau_1}{2}\right)\sin\left(\frac{\tau_2}{2}\right)\cos\left(\frac{\tau_3}{2}\right) = 0.235$$

All together this leads to the relations

$$\begin{aligned} \cos\left(\frac{\vartheta}{2}\right) &= -0.8235\sin\left(\frac{\varphi+\psi}{2}\right) + 0.0222\sin\left(\frac{\varphi-\psi}{2}\right) \\ d\sin\left(\frac{\vartheta}{2}\right) &= 2.999\cos\left(\frac{\varphi+\psi}{2}\right) - 0.657\cos\left(\frac{\varphi-\psi}{2}\right) \end{aligned} \qquad \text{(S2)}$$

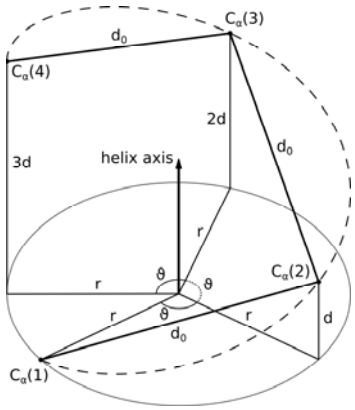

Figure S1: Definition of the three helix parameters (d, ϑ, r), d_0 is the distance between covalently bound C_α atoms.

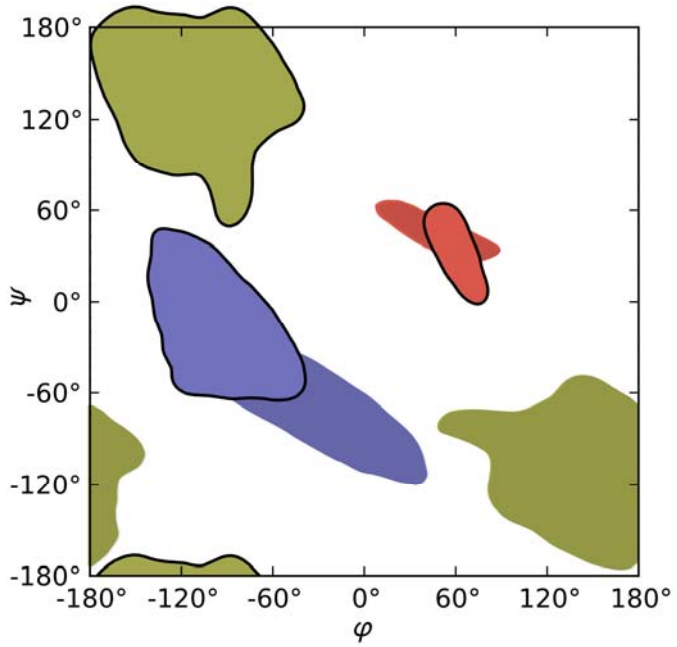

Figure S2: Ramachandran plot (R-plot) obtained by using the back transform from (d, ϑ)-space to (φ, ψ)-space for the easily accessible areas (blue: right handed helices, red: left handed helices, green: strands). Due to the twofold degeneracy that for each (d, ϑ) value there exist two different (φ, ψ) values, the easily accessible areas of the (d, ϑ)-space appear twice in the R-plot. The relevant easily accessible areas in the R-plot are surrounded by back lines. For the strand area there is no overlap, i.e. only one of the two (φ, ψ) values belongs to the easily accessible strand area. For the helix areas (left and right) there is an overlap. However, the (φ, ψ) values in the overlap regime are very similar. Hence, the formal loss in information going from (φ, ψ)-space to , ϑ)-space is not critical. Note that there is an approximate mirror symmetry between the pairs of (φ, ψ) areas obtained by the back transform, which is due to the dominance of the (φ + ψ)/2 angle in the equations (S2) defining the transform.

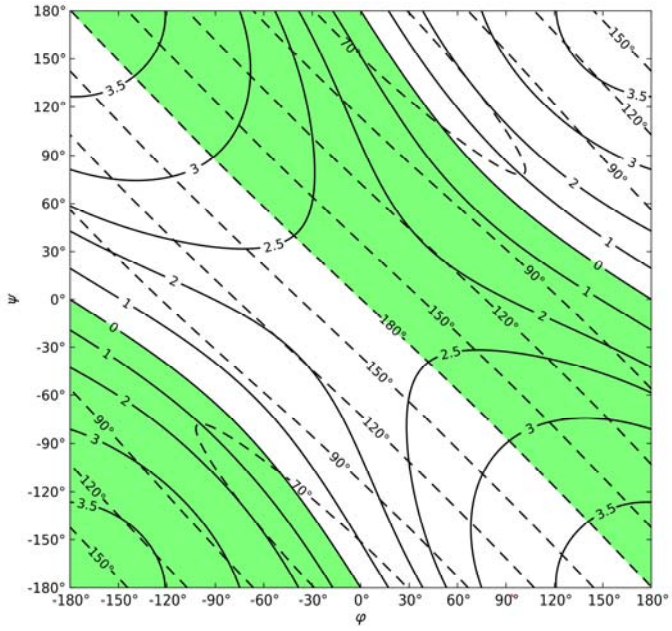

Figure S3: Classical R-plot with contour lines of constant d (solid) and constant ϑ (dashed). Left-handed regions are marked green, right-handed regions are marked white. The borders between left- and right-handed areas are defined by the contour lines with $d = 0$ and $\vartheta = 180^0$. The values of d are given in [Å]. Note that in the left-handed areas either the rise d or the angle ϑ will be considered negative.

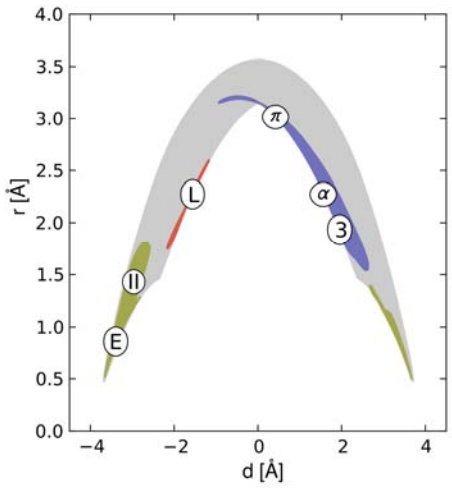

Figure S4: (*d*, *r*)-plot. Negative values of the parameter *d* are considered to discriminate the handedness of the polypeptide backbone structure. Left-handed conformations are on the left- and right-handed on the right-side of the (*d*, *r*)-plot. The easily accessible domains are marked in color (green: strand; blue: helix; red: left-handed helical conformations).

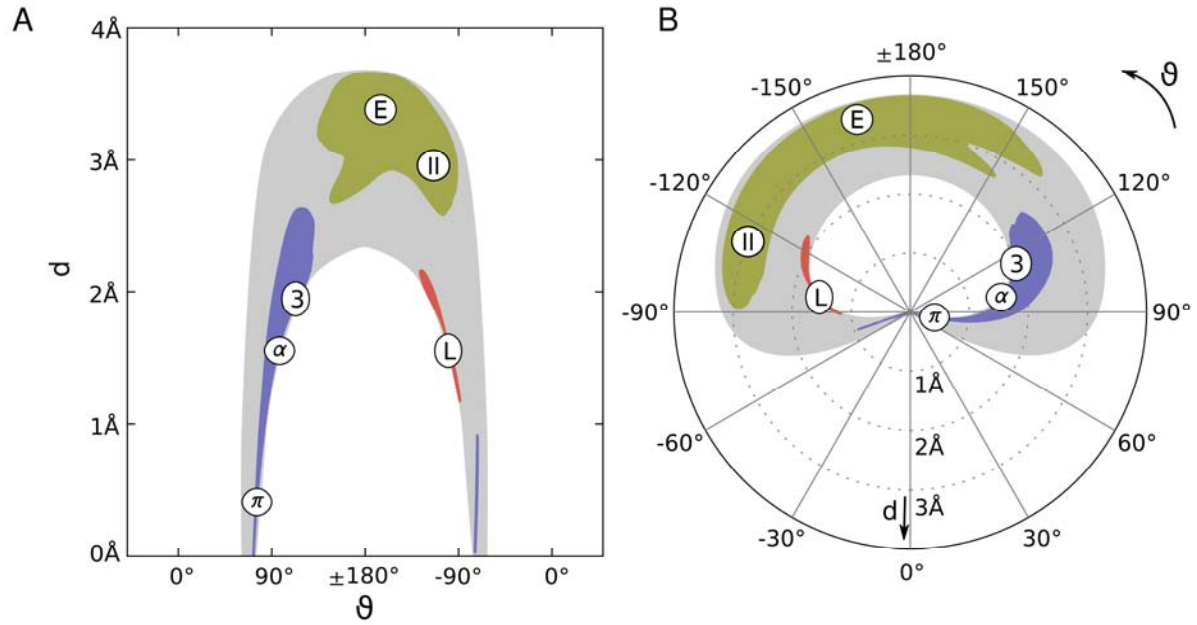

Figure S5: (*d*, *ϑ*)-plots. A: Cartesian (*d*, *ϑ*)-plot; B: polar (*d*, *ϑ*)-plot. Gray shaded area: formally accessible region, i.e. the whole (*φ*, *ψ*)-space; green: strand; blue: helix; red: left-handed helical conformations.

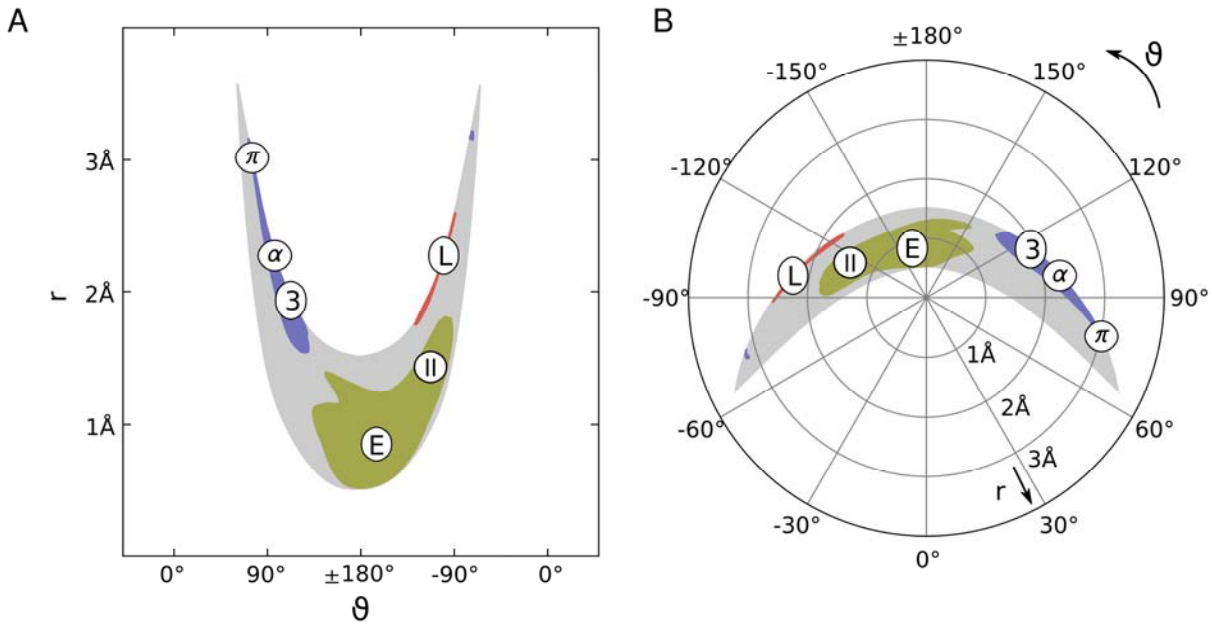

Figure S6: The (r, ϑ)-plot as possible alternative to the R-plot. A: Cartesian representation; B: polar plot. Gray shaded area: formally accessible region, i.e. the whole (φ, ψ)-space; green: strand; blue: helix; red: left-handed helical conformations.

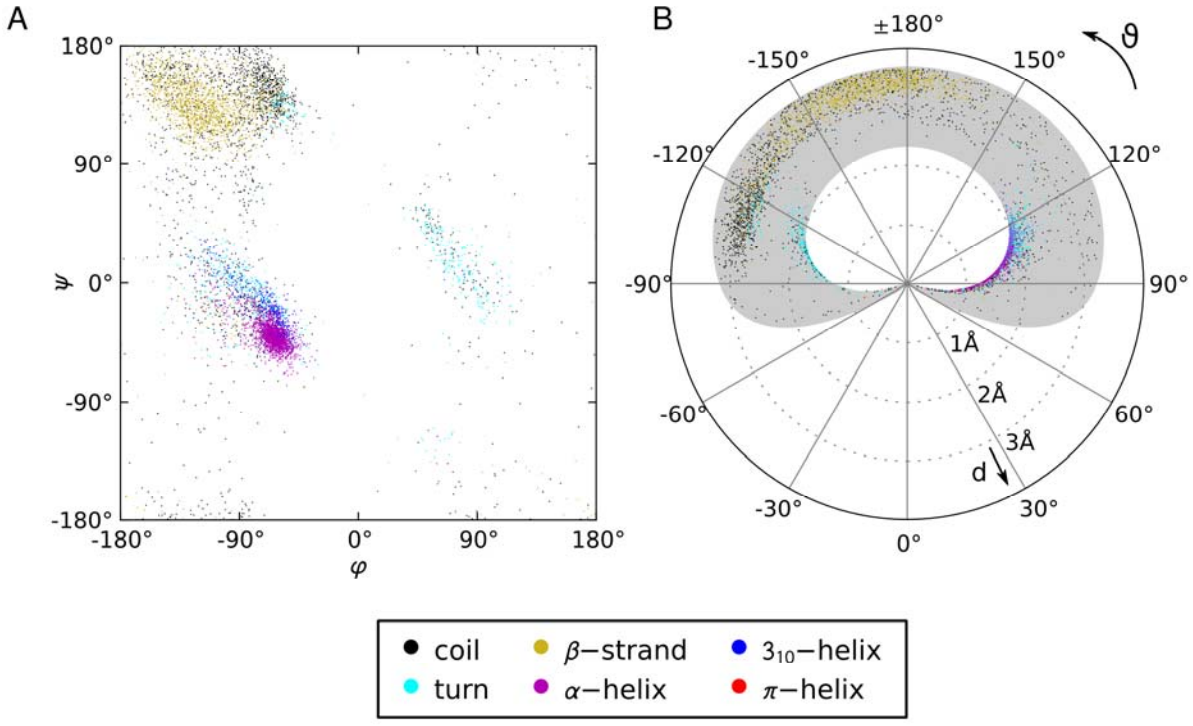

Figure S7: Scatter plot for 50 representative proteins taken from Astral40 database[4] A: Ramachandran-Plot, B: (d, ϑ)-plot; in gray the formally accessible region.

The 50 selected proteins for Figure S7:
d1coza_, d1a04a1, d1hd2a_, d2g82a2, d2heyr3, d1mdoa_, d1adta1, d1ijqa2, d1nekb2, d1pjqa1, d1mg7a1, d1qfja2, d1r9wa_, d1aqta1, d2ghpa3, d1ok7a1, d1xb4a1, d2i9ua1, d2r8oa2, d1ifqa_, d1nd9a_, d2ds5a1, d1gu2a_, d2imha1, d1vcta2, d1r6la2, d1xp4a2, d1st9a_, d2rlda1, d1ppjc2, d2a9sa1, d2bz2a1, d1ebfa2, d1ooya1, d1o89a2, d1ufoa_, d1vlya1, d1q4ua_, d2isba1, d1a8da1, d1pn2a2, d1piwa1, d1ca1a1, d1a6bb_, d1ykga1, d1o59a1, d1sgma1, d2np3a1, d2h3ja1, d1tx2a_.

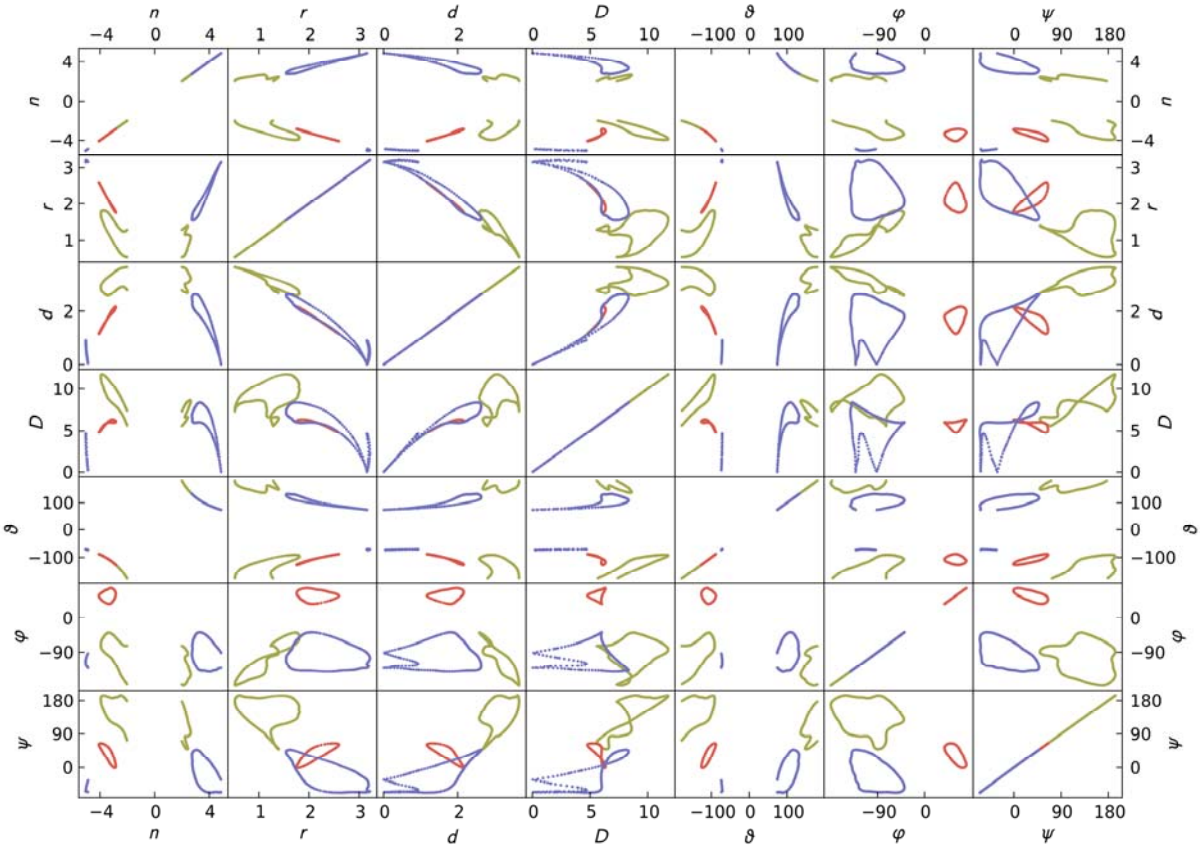

Figure S8: Matrix showing two-dimensional plots of the possible combinations for the variables (*n, r, d, D*, ϑ, φ, ψ), where $D = d\,n$; $n = 360^0 / \vartheta$. Shown are the contours of the easily accessible regions (blue: helical, green: mostly beta-strand, red: left-handed helical). Note that in this plot the handedness is only included in variables *n* and ϑ. Thus, for example, the (*d, r*) plot shows some overlaps that can be avoided as shown in Fig S3.

References:

1. Miyazawa, T., *Molecular vibrations and structure of high polymers. II. Helical parameters of infinite polymer chains as functions of bond lengths, bond angles, and internal rotation angles.* Journal of Polymer Science, 1961. **55**: p. 215-231.
2. Engh, R.A. and R. Huber, *Accurate bond and angle parameters for X-ray protein structure refinement.* Acta Crystallographica Section A, 1991. **47**: p. 392-400.
3. Engh, R.A. and R. Huber, *Structure quality and target parameters*, in *International Tables for Crystallography Volume F: Crystallography of biological macromolecules SE - 42*, M.G. Rossmann and E. Arnold, Editors. 2001, Springer Netherlands. p. 382-392.
4. Chandonia, J.-M., et al., *The ASTRAL Compendium in 2004.* Nucleic acids research, 2004. **32**: p. D189-92.