

An exhaustive survey of regular peptide conformations using a new metric for backbone handedness (h)

Ranjan V. Mannige^{1,2,*}

¹ Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, U.S.A.

² Present address: The Multiscale Institute, Redwood City, CA, U.S.A.

* ranjanmannige@gmail.com

ABSTRACT

The Ramachandran plot is important to structural biology as it describes a peptide backbone in the context of its dominant degrees of freedom – the backbone dihedral angles ϕ and ψ (Ramachandran *et al.*, 1963). Since its introduction, the Ramachandran plot has been a crucial tool to characterize protein backbone features. However, the conformation or twist of a backbone as a function of ϕ and ψ has not been completely described for both *cis* and *trans* backbones. Additionally, little intuitive understanding is available about a peptide's conformation simply from knowing the ϕ and ψ values of a peptide (e.g., is the regular peptide defined by $\phi = \psi = -100^\circ$ left-handed or right-handed?). This report provides a new metric for backbone handedness (h) based on interpreting a peptide backbone as a helix with axial or vertical displacement d and angular displacement θ , both of which are derived from a peptide backbone's internal coordinates, especially dihedral angles ϕ , ψ and ω . In particular, $h = \sin(\theta)d/|d|$ with range $[-1, 1]$ and negative (or positive) values indicating left(or right)-handedness. The metric h is used to characterize the handedness of every region of the Ramachandran plot for both *cis* ($\omega = 0^\circ$) and *trans* ($\omega = 180^\circ$) backbones, which provides the first exhaustive survey of twist handedness in Ramachandran (ϕ, ψ) space. These maps fill in the 'dead space' within the Ramachandran plot, which are regions that are not commonly accessed by globular (structured) proteins, but which may be accessible to intrinsically disordered proteins, short peptide fragments, and protein mimics such as peptoids. Finally, building on the work of Zacharias and Knapp (2013), this report presents a new plot based on d and θ that serves as a universal and intuitive alternative to the Ramachandran plot. The universality arises from the fact that the co-inhabitants of such a plot include every possible peptide backbone including *cis* and *trans* backbones. The intuitiveness arises from the fact that d and θ provide, at a glance, numerous aspects of the backbone including compactness, handedness, and planarity.

INTRODUCTION

The backbone of a protein (Fig. 1a) can twist and turn into numerous conformations (folds), in part due to the amino acid sequence that the protein displays. Understanding how a backbone twists is of great importance to the field of biochemistry, since understanding the structure of a protein goes a long way towards understanding how a protein functions (Alberts *et al.*, 2002; Berg *et al.*, 2010). While the conformation of a peptide backbone is dependent on a number of parameters (bond lengths, bond angles, and dihedral angles), Ramachandran *et al.* (1963) recognized that the twist of a peptide backbone can be described to a great degree by the dihedral angles ϕ and ψ (Fig. 1a).

Today, two-dimensional (ϕ, ψ) plots are called Ramachandran plots (or 'maps'), and are introduced in undergraduate biology textbooks as a guide for understanding a peptide backbone's general conformational state or 'twistedness' at a glance (Bragg *et al.*, 1950; Pauling and Corey, 1951b; Pauling *et al.*, 1951; Linderstrøm-Lang, 1952; Laskowski *et al.*, 1993; Chothia *et al.*, 1997; Hooft *et al.*, 1997; Cooper and Hausman, 2013; Alberts *et al.*, 2002; Laskowski, 2003; Ho *et al.*, 2003; Eisenberg, 2003; Berg *et al.*, 2010; Mannige, 2014). The Ramachandran plot is especially useful because (stable) proteins are hierarchical in structure (Linderstrøm-Lang, 1952): the final (tertiary) conformation of a *structured* protein is composed of discrete secondary structures – *regular* structures – that interact with each other and are strung together by loops that are less regular (Alberts *et al.*, 2002; Berg *et al.*, 2010). Each regular secondary

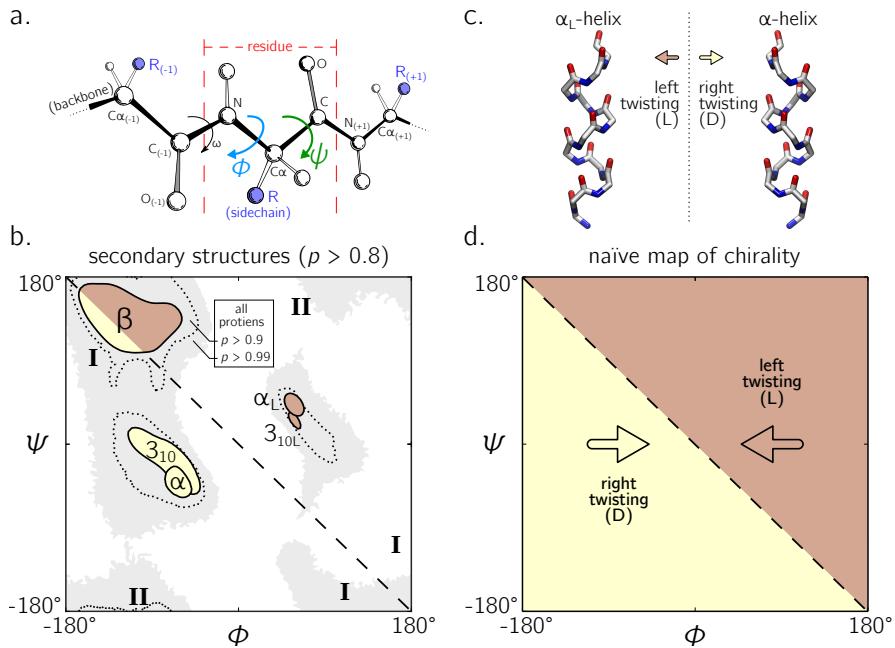


Figure 1. The backbone of a single residue (a) can be described by its dihedral angles ϕ and ψ (and in smaller part, ω , which is predominantly *trans* or $\sim 180^\circ$). For example, repeating a value for ϕ and ψ for every residue within a peptide – i.e., selecting a point on the Ramachandran plot (b) – results in regular secondary structures that then combine to form most globular proteins in the protein databank (Berman *et al.*, 2000). The twist of each perfectly regular secondary structure possesses a distinct chirality: left-handed or right-handed; examples are shown in (c). all molecular representations in this text are shown in ‘licorice’ form, with the colors red, blue and white representing the oxygen, nitrogen and carbon atoms. As evidenced in (c), the -ve diagonal within the Ramachandran plot (dashed line described by $\phi = \psi$) divides right-handed peptides from left-handed peptides, which leads to the naïve picture of handedness (d). Zacharias and Knapp (2013) showed that this picture is over simplistic, however an in-depth characterization of the backbone in all regions was not performed, and will be done here for both *cis* ($\omega = 0$) and *trans* backbones ($\omega = \pi$). Due to low count within the studied database (see Methods), the two left-handed helices in (b) are arbitrarily marked and have no statistical significance.

47 structure describes a backbone whose per-residue (ϕ, ψ) values are generally the same, and therefore
48 their ‘locations’ on the Ramachandran plot act as structural landmarks (Fig. 1b).

49 So far, our understanding of the Ramachandran plot has been limited mostly to ‘structured’ proteins
50 that display stable conformations (Berman *et al.*, 2000; Alberts *et al.*, 2002). These types of proteins
51 occupy only a limited region of the plot (dotted regions in Fig. 1b). The regular backbone conformations
52 in these regions are well understood; for example, known regular structures that are to the right of
53 the negatively sloping diagonal (dashed line in Fig. 1b; henceforth denoted as the ‘-ve diagonal’) are
54 left-handed in backbone twist, while those that are to the left of the diagonal are right-handed (left- and
55 right-handed regions are respectively shaded brown and gold¹). For example, the position of the idealized
56 left- and right-handed α -helices (Fig. 1c) – respectively denoted as α_L and α in Fig. 1b – are on opposite
57 sides of the -ve diagonal. The ‘naïve view’ of handedness, obtained from looking only at structured
58 proteins, would be the expectation that the -ve diagonal neatly separates the Ramachandran plot into
59 regions of left- and right-handedness (Fig. 1d).

60 However, structured proteins represent only a fraction of functional proteins. Indeed, up to 15%
61 of mammalian proteins are completely disordered – they natively display multiple, often extended,
62 conformations (Mannige, 2014) – and up to 50% of the mammalian proteins display large (> 30 residue)
63 stretches of disorder (Iakoucheva *et al.*, 2002; Ward *et al.*, 2004; Orosz and Ovádi, 2011). Interestingly,
64 when compared to structured backbones, structurally degenerate or disordered backbones occupy many

¹Go Hillies!

more regions within the Ramachandran plot (Beck *et al.*, 2008).

Additionally, a number of peptide mimics – especially peptoids (Sun and Zuckermann, 2013) – display novel secondary structures that occupy regions that are strictly disallowed by proteins due to steric clashes. For example, a ‘higher-order’ peptoid secondary structure – the Σ -strand (Mannige *et al.*, 2015; Robertson *et al.*, 2016) – samples regions of the Ramachandran plot (‘I’ in Fig. 1b) that are not permitted within natural proteins (this is because of the lack of backbone hydrogen bond donors in peptoids). Another secondary structure – the ‘ ω -strand’ (Gorske *et al.*, 2016) – samples similarly historically uncharted regions of the Ramachandran plot (‘II’ in Fig. 1b). Importantly, backbone twist handedness plays a crucial part in explaining these new motifs: as one goes along the backbones of these secondary structures, alternating residues occupy equal but opposite backbone twists [for this reason, the Σ -strand is relatively linear, albeit meandering; Mannige *et al.* (2015); Mannige *et al.* (2016)].

Despite these recent discoveries of natively disordered proteins and novel peptidomimetic structures, a complete understanding of backbone conformations that stray from the ‘structured’ regions on the Ramachandran plot is missing, which impedes our ability to identify and explore such conformations. Towards filling this gap in understanding, this report outlines a detailed study of how regular backbones twist in every region of the Ramachandran plot. In particular, this report develops and explores a new metric for handedness (h) based on modeling a regular backbone (described below) as a helix (Shimanouchi and Mizushima, 1955; Miyazawa, 1961; Zacharias and Knapp, 2013). The metric is used to exhaustively chart the handedness of regular backbones in every region of the Ramachandran plot. In doing so, this survey provides a new graphical format to explore new types of secondary structures being discovered (Mannige *et al.*, 2015; Gorske *et al.*, 2016), while dispelling the naïve view of handedness (Fig. 1d) by showing that the distribution of handedness as a function of ϕ and ψ is more complicated than the distribution allowed by the naïve view. The results also show that the Ramachandran plot whose ϕ and ψ values range between 0° and 360° is more intuitive and visually meaningful (compared to those that range between -180° and 180°), particularly for *cis* backbones. Finally, this work builds on a previous report (Zacharias and Knapp, 2013) and helps complete our understanding of the ways in which a peptide backbone twists, which is a basic component of structural biology.

METHODS

While angular units in this report switch between radians and degrees, their units in any particular situation may be inferred by the presence or absence of the degree symbol ($^\circ$). All methods and materials required to produce this manuscript are freely available at https://github.com/ranjanmannige/backbone_chirality.

Deriving measures for backbone handedness

Numerous metrics for molecular chirality and handedness have so far been discussed (Harris *et al.*, 1999). For example, metrics for chirality have been introduced that focus on vector orientations (Kwiecińska and Cieplak, 2005; Kabsch and Sander, 1983; Gruziel *et al.*, 2013), optical activity (Osipov *et al.*, 1995), and molecular shape (Ferrarini and Nordio, 1998). However, this report will focus on a simpler metric for chirality associated with an idealized helix within which all (regular) backbone atoms of one type sit (Shimanouchi and Mizushima, 1955; Miyazawa, 1961; Zacharias and Knapp, 2013). Here, a ‘regular’ backbone indicates that each tunable parameter within a unit or ‘residue’ – say a particular dihedral angle – remains the same for all residues. Below, regular backbones are modeled in context of helical parameters that, when combined, form an intuitive metric for backbone handedness.

Describing a regular backbone as a helix

Interest in how a backbone may be represented as a helix emerged shortly after the first secondary structures were introduced (Pauling *et al.*, 1951; Pauling and Corey, 1951b,a). In particular, Shimanouchi and Mizushima (1955) had derived a set of equations that fit a platonic helix to the atoms within a regular backbone. While the formalisms described by Shimanouchi and Mizushima (1955) [and later on by Miyazawa (1961), discussed below] apply to repeating linear polymers of arbitrary complexity, this report focuses specifically on how peptides may be modeled. Fig. 2 describes an arbitrary peptide backbone that may be represented either using internal coordinates (*i*) or helical coordinates (*ii*).

Internal coordinates are associated with stereochemical terms: bond lengths (v_{ij}) between adjacent atoms *i* and *j*, bond angles (σ_i) between the two backbone bonds adjacent to atom *i*, and dihedral or

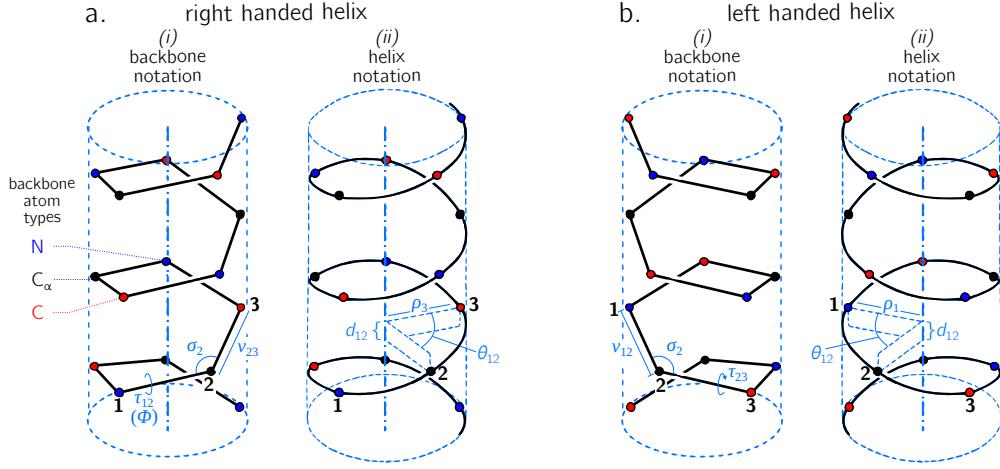


Figure 2. Internal coordinate (i) and helical coordinate (ii) representations of right-handed (a) and left-handed (b) regular backbones. Internal coordinates are a function of bond lengths (e.g., v_{12}), angles (σ_2), and dihedral angles (τ_{12}), while helical coordinates are a function of displacement along the helical axis (d_{12}), angular displacement in the plane perpendicular to the helical axis (θ_{12}) and shortest distance of an atom of type i from the helical axis (ρ_i). Representations are derived from Figs. 1 and 2 in Shimanouchi and Mizushima (1955).

torsion angles (τ_{ij}), which are associated with the bond $i - j$ and the atoms preceding and succeeding these atoms. Helical coordinates (Fig. 2(ii)) are described using measures of vertical displacement (d ; this is related to the pitch of a platonic helix), angular displacement (θ) and the radius of the helix (ρ_i) that hosts all backbone atoms of type i . Therefore, the single cylinder shown in Fig. 2 is too simplistic as there should be one distinct cylinder or radius per atom type.

Given that there are three atoms associated with a residue (Fig. 1a), $d = d_{n,\alpha} + d_{\alpha,c} + d_{c,n}$ and $\theta = \theta_{n,\alpha} + \theta_{\alpha,c} + \theta_{c,n}$. Here, $d_{i,j}$ and $\theta_{i,j}$ respectively refer to the vertical and angular displacement between adjacent atoms i and j . Subscripts ‘ n ’, ‘ α ’, and ‘ c ’ respectively refer to the backbone nitrogen, α -carbon and carbonyl carbon atoms (Fig. 1a). The notation used by Shimanouchi and Mizushima (1955) was in terms of matrices, which were then simplified by Miyazawa (1961) into trigonometric terms. In particular, Miyazawa (1961) noted that the total residue-residue vertical displacement (d) and angular displacement (θ) may be retrieved using the following two equations.

$$\begin{aligned} \cos\left(\frac{\theta}{2}\right) = & \cos\left(\frac{+\phi + \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right) \sin\left(\frac{\sigma_\alpha}{2}\right) \sin\left(\frac{\sigma_c}{2}\right) \\ & - \cos\left(\frac{+\phi - \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right) \cos\left(\frac{\sigma_\alpha}{2}\right) \cos\left(\frac{\sigma_c}{2}\right) \\ & - \cos\left(\frac{+\phi + \psi - \omega}{2}\right) \cos\left(\frac{\sigma_n}{2}\right) \sin\left(\frac{\sigma_\alpha}{2}\right) \cos\left(\frac{\sigma_c}{2}\right) \\ & - \cos\left(\frac{-\phi + \psi + \omega}{2}\right) \cos\left(\frac{\sigma_n}{2}\right) \cos\left(\frac{\sigma_\alpha}{2}\right) \sin\left(\frac{\sigma_c}{2}\right) \end{aligned} \quad (1)$$

$$\begin{aligned}
d \sin\left(\frac{\theta}{2}\right) = & (+v_{n,\alpha} + v_{\alpha,c} + v_{c,n}) \sin\left(\frac{+\phi + \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right) \sin\left(\frac{\sigma_\alpha}{2}\right) \sin\left(\frac{\sigma_c}{2}\right) \\
& - (+v_{n,\alpha} - v_{\alpha,c} + v_{c,n}) \sin\left(\frac{+\phi - \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right) \cos\left(\frac{\sigma_\alpha}{2}\right) \cos\left(\frac{\sigma_c}{2}\right) \\
& - (+v_{n,\alpha} + v_{\alpha,c} - v_{c,n}) \sin\left(\frac{+\phi + \psi - \omega}{2}\right) \cos\left(\frac{\sigma_n}{2}\right) \sin\left(\frac{\sigma_\alpha}{2}\right) \cos\left(\frac{\sigma_c}{2}\right) \\
& - (-v_{n,\alpha} + v_{\alpha,c} + v_{c,n}) \sin\left(\frac{-\phi + \psi + \omega}{2}\right) \cos\left(\frac{\sigma_n}{2}\right) \cos\left(\frac{\sigma_\alpha}{2}\right) \sin\left(\frac{\sigma_c}{2}\right)
\end{aligned} \quad (2)$$

129 The ranges for d and θ , respectively, are $[-\lambda, +\lambda]$ and $[0, 2\pi]$ (the positive limit λ is defined by allowed
130 values for the various internal coordinates). As above, subscripts ‘n’, ‘ α ’, and ‘c’ respectively refer to
131 the backbone nitrogen, α -carbon and carbonyl carbon atoms; v_{ij} refers to the distance between adjacent
132 atoms i and j ; σ_i refers to the angle between the adjacent bonds that have i as the common atom. The
133 dihedral angles ϕ , ψ , and ω represent the traditional symbols for backbone dihedral angles, which may
134 be otherwise denoted as $\tau_{n,\alpha}$, $\tau_{\alpha,c}$, and $\tau_{c,n(+1)}$, respectively.

Finally, for any type of atom (say α -carbons), the radius or distance from the helical axis ρ_α is defined by

$$\begin{aligned}
2\rho_\alpha^2 [1 - \cos(\theta)] + d^2 = & v_{\alpha,c}^2 + v_{c,n}^2 + v_{n,\alpha}^2 - 2v_{c,n} [v_{\alpha,c} \cos(\sigma_c) + v_{n,\alpha} \cos(\sigma_n)] \\
& + 2v_{\alpha,c} v_{n,\alpha} [\cos(\sigma_c) \cos(\sigma_n) - \sin(\sigma_c) \sin(\sigma_n) \cos(\tau_{c,n})]
\end{aligned} \quad (3)$$

Miyazawa (1961) noted that the right-hand side of Eqn. 3 is also the squared distance between adjacent atoms of the same type (which we denote as d_α^2 for α -carbons), which allows for a more simplified form

$$\rho_\alpha = \sqrt{\frac{d_\alpha^2 - d^2}{2 - 2\cos(\theta)}} \quad (4)$$

135 The distance between adjacent α -carbons (d_α) is $\sim 3.8\text{\AA}$ for *trans* peptides and $\sim 3.2\text{\AA}$ for *cis* peptides.
136 Other radii (ρ_c, ρ_n) can be obtained by cycling through (α, c, n) subscripts within Eqns. 3 and 4². Note
137 that all ρ_i 's are functions of θ, d (along with other internal coordinates), and so one may use two of
138 the three terms (d, θ, ρ_i) – to describe the helical state of a peptide. Since there is only one d and θ per
139 backbone (compared to three ρ_i 's, one per atom type), this report utilizes d and θ as the two descriptors
140 [discussions on this choice have also been made by Zacharias and Knapp (2013)].

Eqns. 1 and 2 may be substantially simplified (Miyazawa, 1961), given that backbone bond lengths and angles are much less ‘tunable’ when compared to dihedral angles (Ramachandran *et al.*, 1963; Imrota *et al.*, 2015a; Esposito *et al.*, 2013; Imrota *et al.*, 2015b). In particular, most backbone bond lengths and angles display one equilibrium value (Imrota *et al.*, 2015a; Esposito *et al.*, 2013; Imrota *et al.*, 2015b), while the backbone dihedral angles ϕ and ψ occupy a range of possible values and minima, e.g., regions in the Ramachandran plot that describe α -helices and β -sheets (Fig. 1b). With this in mind, Miyazawa (1961) set $\omega = \pi$ (*trans*) and substituted average (equilibrium) values for bond angles and lengths into Eqns. 1 and 2 to arrive at a simpler equation for *trans* backbones. Zacharias and Knapp (2013) published an updated version of this set of equations, which follows³.

$$\cos\left(\frac{\theta}{2}\right) = -0.8235 \sin\left(\frac{\phi + \psi}{2}\right) + 0.0222 \sin\left(\frac{\phi - \psi}{2}\right), \quad (5)$$

$$d \sin\left(\frac{\theta}{2}\right) = 2.9986 \cos\left(\frac{\phi + \psi}{2}\right) - 0.6575 \cos\left(\frac{\phi - \psi}{2}\right). \quad (6)$$

This equation is especially relevant to peptides as they occur predominantly in *trans* conformations ($\omega = \pi$). However, given the prevalence of *cis* backbones in peptide mimics such as peptoids (Mirijanian

² ρ_c and ρ_n are obtained by the following subscript conversions: ($\alpha \rightarrow c, c \rightarrow n, n \rightarrow \alpha$) and ($\alpha \rightarrow n, c \rightarrow \alpha, n \rightarrow c$).

³Values used here are also used by Zacharias and Knapp (2013). These values are:

$v_{n,\alpha} = 1.459\text{\AA}$, $v_{\alpha,c} = 1.525\text{\AA}$, $v_{c,n(+1)} = 1.336\text{\AA}$, $\sigma_\alpha = 111.0$, $\sigma_c = 117.2^\circ$, and $\sigma_n = 121.7^\circ$.

For reference, Miyazawa (1961) originally used the following values:

$v_{n,\alpha} = 1.470\text{\AA}$, $v_{\alpha,c} = 1.530\text{\AA}$, $v_{c,n(+1)} = 1.320\text{\AA}$, $\sigma_\alpha = 110.0$, $\sigma_c = 114.0$, and $\sigma_n = 123.0$.

et al., 2014; Gorske *et al.*, 2016), for completeness, the corresponding relationships for a *cis* ($\omega = 0$) backbone follows.

$$\cos\left(\frac{\theta}{2}\right) = 0.4052 \cos\left(\frac{\phi + \psi}{2}\right) - 0.4932 \cos\left(\frac{\phi - \psi}{2}\right), \quad (7)$$

$$d \sin\left(\frac{\theta}{2}\right) = 2.3093 \sin\left(\frac{\phi + \psi}{2}\right) + 0.0028 \sin\left(\frac{\phi - \psi}{2}\right). \quad (8)$$

Note that Eqns. 5 through 8 are simplifications of Eqns. 1 and 2, and are therefore prone to some limitations that are not present in Eqns. 1 and 2. For example, bond lengths (Impronta *et al.*, 2015a) and bond angles (Esposito *et al.*, 2013; Impronta *et al.*, 2015b) display some dependence on local backbone conformation. These subtle variations have great implications when dealing with a large number of residues, especially when considering bond angles. For example, when attempting to recreate a protein conformation from an original conformation's ϕ and ψ values (ignoring deviations in ω , bond angles, and lengths), the original and recreated conformations tend to deviate dramatically due to an accumulation of errors [by up to 22 Å in root mean squared deviation; Tien *et al.* (2013)]. However, when studying changes in conformationally regular and *local* stretches of peptides, such deviations are not likely to change relevant features such as handedness and extent of twistedness. If circumstances indicate that the backbone values for bond angles and ω of the peptide may be strained from their equilibrium values (e.g., due to bulky sidechains), only Eqns. 1 and 2 can be expected to faithfully (and perfectly) represent backbone features such as handedness of twist. However, the approximations of Eqns. 5 through 8 are sufficient for the purposes of this report, given that this report primarily discusses features within platonic regular backbones.

On the one-to-one correspondence between $[\phi, \psi, \omega]$ and $[d, \theta]$

Given a particular value of ω , every $[\phi, \psi]$ pair points to exactly one $[d, \theta]$. However, when using Eqns. 1 and 2, one value of ω can not be replaced with a *periodically equivalent* version of ω . For example, using $\omega = x + 2\pi$ instead of $\omega = x$ will maintain the magnitude of d and θ , but the signs will not remain conserved. This is because every summand in Eqns. 1 and 2 contains either a sine or cosine of $[\pm\phi \pm \psi \pm \omega]/2$. The issue arises because of the '2': even though the angle x is considered to be equivalent to the angle $x + 2\pi$, and even though $\cos(x + 2\pi)$ equals $\cos(x)$ (due to angle periodicity), $\cos([x \pm 2\pi]/2) = \cos(x/2 \pm \pi) = -\cos(x/2)$ (note the negative sign). Similarly, $\sin([x \pm 2\pi]/2) = -\sin(x/2)$. Therefore, even though the angles ω and $\omega + 2\pi$ may be considered to be equivalent angles, expressions such as $\cos([x - \omega + 2\pi]/2)$ and $\cos([x - \omega]/2)$ are only equal in magnitude and not in sign. I.e., one-to-one correspondence between $[\phi, \psi]$ and $[d, \theta]$ is only possible if we insist on specific values for ω s. For this reason, we propose to wrap the value of an amide backbone ω' between $[\Delta, \Delta + 360^\circ)$ using

$$\omega = (\omega' - \Delta)\%360 + \Delta, \quad (9)$$

where $\%$ represents the modulus function, and Δ describes the start of the range. Choosing $\Delta = -90^\circ$ would ensure that the distribution of both *cis* ($\omega = 0 \pm 20^\circ$) and *trans* ($\omega = 180 \pm 20^\circ$) will remain contiguous. Using this system, *cis* and *trans* backbones are respectively represented by $\omega = 0$ (and not 2π) and $\omega = \pi$ (not $-\pi$) for *trans* backbones. The rest of this report assumes these values of ω for *cis* and *trans* backbones.

These points lead to the conclusion that a *strict* one to one-to-one correspondence between (ϕ, ψ, ω) and (d, θ) does not exist, since multiple sets of the former may be backmapped from the latter (by reconfiguring Eqns. 1 and 2). Yet, a one-to-one correspondence may be ensured by discarding as solutions all but the one set of (ϕ, ψ, ω) , whose ω does not change after being wrapped by Eqn. 9.

Introducing an equation for backbone handedness

The helical parameters d and θ host a wealth of information, some of which is discussed in the Results section. For the purpose of developing an equation for backbone handedness, it is only important to recognize, as was done before (Zacharias and Knapp, 2013), that θ and d together are instrumental in describing backbone handedness.

The relationship between d , θ and handedness is shown in Fig. 3. While θ indicates the extent to which a regular backbone curves along a helical path, the handedness of a backbone is dependent on both

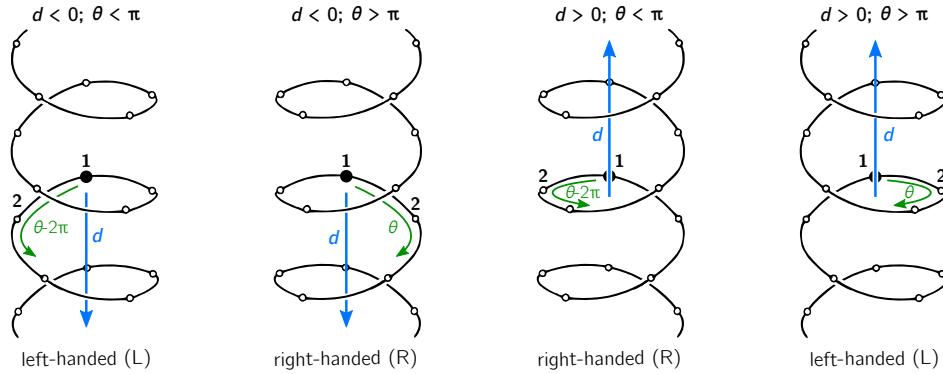


Figure 3. The handedness of a helix is a function of angular displacement perpendicular to the helical axis (θ ; green curved arrows) and linear displacement along the helical axis d (blue, vertical arrows). Note that left- and right-handed backbone twists are respectively associated with the L and D chiralities within the Fisher Projection system and S and R chiralities within the Cahn–Ingold–Prelog system (Cross and Klyne, 2013); however, as discussed in the Methods section, this report makes a distinction between helix handedness and molecular chirality.

173 θ and d . This is because the sign of d provides a frame of reference for interpreting θ . In particular, if d
 174 is positive, then $0 < \theta < \pi$ indicates right-handedness, while $\pi < \theta < 2\pi$ indicates a left-handed helix.
 175 However, if d is negative, then the manner in which the helix is ‘built’ reverses (Fig. 3), and $0 < \theta < \pi$
 176 indicates left-handedness, while $\pi < \theta < 2\pi$ indicates right-handedness.

These relationships allow for a simple metric for backbone handedness that depends on the sign of d and the value of θ :

$$h = \frac{d}{|d|} \sin(\theta). \quad (10)$$

177 Here, $h \in \{-1, \dots, 1\}$ is negative (or positive) when the overall twist of the backbone is left(or right)-
 178 handed. Note that $d/|d|$ is related to the traditional sign function $\text{sgn}(d)$, but deviates at $d = 0$, where
 179 the former term is undefined while the latter term is 0. Additionally, h will equal 0 if $d = 0$ or if $\theta = x\pi$
 180 (where x is an integer); for more on the meaning of d and θ in context of handedness and peptide geometry,
 181 please refer to the Results and Discussions section and Fig. 4 in particular.

182 Alternative measures of handedness

Two estimates for chirality, χ_1 and χ_2 , used to validate the new measure of handedness h (Eqn. 10), were previously used by Kwiecińska and Cieplak (2005) and Kabsch and Sander (1983), respectively. The equations are:

$$\chi_1 = \frac{1}{N} \sum_{i=2}^{N-2} \frac{(\mathbf{v}_{i-1} \times \mathbf{v}_i) \cdot \mathbf{v}_{i+1}}{\mathbf{v}_{i-1} \cdot \mathbf{v}_i \cdot \mathbf{v}_{i+1}}, \quad (11)$$

$$\chi_2 = \frac{1}{N} \sum_{i=2}^{N-2} \arctan2(\mathbf{v}_i \cdot \mathbf{v}_{i-1} \cdot \mathbf{v}_i \times \mathbf{v}_{i+1}, \mathbf{v}_{i-1} \times \mathbf{v}_i \cdot \mathbf{v}_i \times \mathbf{v}_{i+1}). \quad (12)$$

183 Here, N is the peptide length, i is the peptide residue number and the position of each α -carbon is
 184 \mathbf{N}_i , with vector $\mathbf{v}_k \equiv \mathbf{N}_{k+1} - \mathbf{N}_k$. The scalar component of the vector \mathbf{v}_i is denoted as v_i . Eqn. 11 has
 185 range $[-1, 1]$. Eqn. 12, also used by Gruziel *et al.* (2013), is the dihedral angle associated with the four
 186 contiguous α -carbons (one preceding and two succeeding the residue i), and ranges between $[-\pi, \pi]$
 187 radians. For both metrics, values deviating more from 0 are more chiral (or ‘twisted’ or ‘handed’), and
 188 left-handed twists are negative while right-handed twists are positive. Only α -carbon atom positions are
 189 used for the calculation.

Finally, a more backbone-agnostic metric of chirality has been introduced by Solymosi *et al.* (2002), which is replicated here purely for completeness:

$$\chi_3 = \frac{4!}{3N^4} \sum_{i,j,k,l \in N} \frac{((\mathbf{v}_{ij} \times \mathbf{v}_{kl}) \cdot \mathbf{v}_{il})(\mathbf{v}_{ij} \cdot \mathbf{v}_{jk})(\mathbf{v}_{jk} \cdot \mathbf{v}_{kl})}{(\mathbf{v}_{ij} \cdot \mathbf{v}_{jk} \cdot \mathbf{v}_{kl})^2 v_{il}}. \quad (13)$$

¹⁹⁰ χ_3 , of arbitrary range, is known as the chirality index G_{0S} in Solymosi *et al.* (2002) and Neal *et al.* (2003).
¹⁹¹ (i, j, k, l) are exhaustive permutations of $\{1, 2, \dots, N\}$. While unshown here, this metric qualitatively
¹⁹² matches the values of Eqn. 11 and 12.

233 Backbone structure generation

234 The metric h (Eqn. 10) is purely analytical and does not need structures to be computationally generated,
235 since Eqns. 5 through 8 that provide d and θ require only pairs of ϕ and ψ angles. However, if values
236 for bond angles, lengths and dihedral angles are expected to deviate greatly from equilibrium values, θ
237 and d can only be obtained from the more detailed Eqns. 1 and 2, whose parameters would likely be
238 obtained from a structure. On the other hand, as χ_1 (Eqn. 11) and χ_2 (Eqn. 12) work explicitly with atom
239 positions, these metrics explicitly need the generation of structures. In order to use these metrics, peptides
240 (poly-glycines) of arbitrary length were generated using the Python-based PeptideBuilder library (Tien
241 *et al.*, 2013). Analysis was performed using BioPython (Cock *et al.*, 2009) and Numerical Python (Van
242 Der Walt *et al.*, 2011). Ramachandran plots that describe chirality (e.g., Fig. 5a) were generated using a
243 grid spacing (in degrees) of $\phi, \psi \in \{-180, -178, \dots, 178, 180\}$.

244 Backbone chirality \neq backbone handedness

245 Finally, it is important to recognize the distinction between backbone (twist) handedness and backbone
246 (molecular) chirality. Naively, chirality is a simple concept: a molecular conformation is achiral if its
247 mirror image can be superimposed onto itself, otherwise that conformation is chiral (Gold *et al.*, 1997)
248 (alternatively, and less commonly, achiral molecules possess inversion centers). Despite this intuitive
249 definition, chirality has remained a confusing concept ever since its introduction (Bentley, 2010; Wallentin
250 *et al.*, 2009), which is possibly due to the fact that ‘context’ is very important when discussing chirality
251 (Mislow, 2002). For example, when looking at a peptide at the residue or ‘local’ level, every amino
252 acid (excepting glycine) is chiral due to the presence of a chiral α -carbon (its mirror image can not be
253 superimposed onto itself). Yet, at the macromolecular level, even an all-glycine (and therefore locally
254 achiral) peptide will display *conformations* that are not superimposable onto each other, and so such
255 conformations would be chiral. Alternatively, when considering handedness, if a backbone is completely
256 flat (say, a ring, where $d = 0$), handedness (h) will be undefined, and so one can not speak of handedness
257 of the twist. Yet, the backbone may still remain chiral; e.g., cisplatin and transplatin are planar molecules
258 that are nonetheless chiral opposites (Testa, 2013). It is for this reason that this report chooses to be
259 careful to not claim that Eqn. 10 is a metric for peptide/backbone chirality, but of peptide backbone *twist*
260 handedness. However, estimates for backbone chirality (e.g., Eqns. 11 and 12) may be used as surrogates
261 for twist chirality to validate h (Eqn. 10), as both are related but not the same.

262 Obtaining secondary structure statistics

263 Statistics about secondary structures – particularly α -helices, 3_{10} -helices and β -sheets – used in Figs. 1b
264 and 9 were identified using the DSSP algorithm (Kabsch and Sander, 1983), although the STRIDE
265 algorithm (Frishman and Argos, 1995) provides qualitatively identical distributions. This algorithm was
266 applied to a database of 13,760 three-dimensional protein conformations (one domain per conformation)
267 with lower than 40% sequence identity, obtained from the Structural Classification of Proteins or SCOPe
268 website [Release 2.06; Fox *et al.* (2014)]. This database is currently available as: <http://scop.berkeley.edu/downloads/pdbstyle/pdbstyle-sel-gs-bib-40-2.06.tgz>.

269 RESULTS AND DISCUSSION

270 Relevance of θ and d

271 When discussing peptide backbones, two possible definitions of backbone ‘flatness’ (or linearity) are
272 possible: flatness at a residue level and flatness at the atomic level. In the former, all atoms of the same *type*
273 are coplanar (examples of atom types are the backbone nitrogens, carbonyl carbons, α -carbons, or even

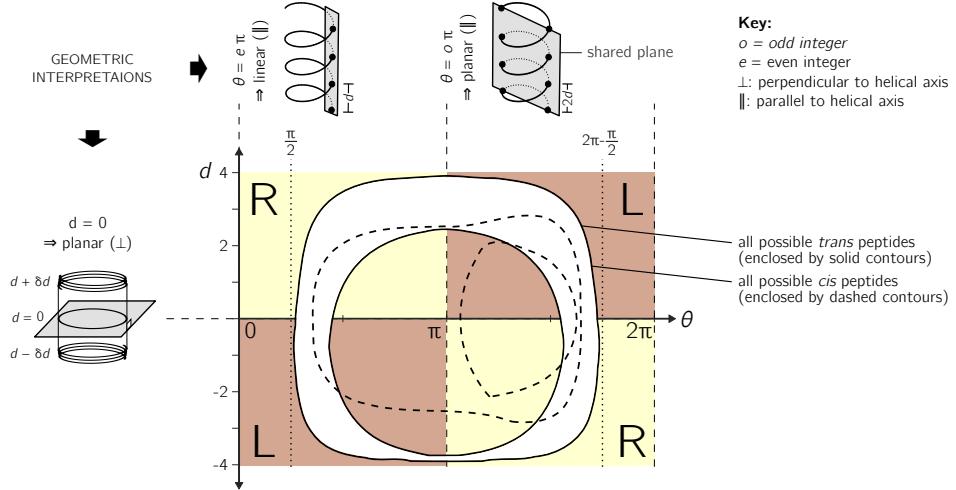


Figure 4. Further discussion on the meaning of d and θ . As shown in Fig. 3, axial separation d and angular separation θ between adjacent atoms of the same type combine to define handedness. The shaded regions within the graph show the distribution of handedness as a function of d and θ . The relevant boundaries – $\theta = x\pi$ (where x is a non-negative integer) and $d = 0$ – separate the map into four quadrants of left- and right-twisting backbones ('L' and 'R', respectively). Geometric interpretations of various boundaries, discussed in the text, are shown to the top and left of the graph. The toroid enclosed by two solid lines (and shaded white) represents all possible conformations for *trans* peptides ($\omega = 180 \pm 10^\circ$). Similarly, the region allowed for *cis* peptides are bound by the two dashed contours. As expected, *trans* and *cis* peptides behave very differently, further predicated the need to explore the two types of backbones as distinct.

235 sidechain β -carbons). In the latter, *all* atoms within the backbone are coplanar. For the discussions below,
 236 the former definition is chosen as the relevant scope for flatness (and ‘linearity’), as residue-by-residue
 237 behavior of the peptide is of primary relevance.

238 As described in Fig. 3, the helical parameters d and θ respectively refer to a **vertical** displacement
 239 along the helical axis and an angular displacement in a plane perpendicular to the helical axis. For
 240 example, $d = 0$ indicates a helix flattened along its helical axis (Fig. 4). This means that all regular
 241 peptides with $d = 0$ will be ring-like at some peptide length (shown in a following figure for a range of
 242 peptides). As expected from Eqn. 10, at $d = 0$, one can not tell how the helix was built, since coplanar
 243 peptides can not be described as either left- or right-twisting. Therefore, even though $d = 0$ indicates
 244 highly twisted peptides, these twists do not possess handedness. This shows up in the h metric because, at
 245 $d = 0$, $d/|d|$ is undefined.

246 Additionally Fig. 4 describes two important values for θ : $e\pi$ and $o\pi$, where e and o are even and odd
 247 non-negative integers. In particular, for any d , $\theta = e\pi$ indicates zero angular displacement along the axis,
 248 which puts all atoms of the same type on the same line parallel to the helical axis (Fig. 4). Similarly,
 249 $\theta = o\pi$ indicates that every alternate atom (of the same type) along the backbone will be linear, and every
 250 adjacent atom will be diametrically opposite to each other (Fig. 4); i.e., $\theta = o\pi$ indicates that all atoms of
 251 the same type will lie on a plane that is parallel to the helical axis. In short, $\theta = 0$ codes for backbones
 252 that are linear (optimally extended for a fixed d) and $\theta = \pi$ describe peptides that zig-zag along a plane
 253 perpendicular to the helical axis (for a fixed d). Finally, as evident in Fig. 4, $\theta = e\pi$ conformations are not
 254 available for peptide backbones; in these situations, $\theta = o\pi$ (e.g., π or 180°) will be the most extended
 255 type of backbone (for a fixed d). These ideas give us a sense, *a priori*, about the meaning of a backbone
 256 curve with particular d and θ . These ideas, along with the combination rules for θ and d (Fig. 3) that
 257 provide us with the handedness of a backbone (h ; Eqn. 10), provide a rich landscape onto which peptides
 258 can be discussed and explored (discussed further at the end of this section).

259 Finally, θ may serve as an important single-number metric for describing backbone configurations.
 260 Mannige *et al.* (2016) developed one such number – a Ramachandran number (\mathcal{R}) – that combines the
 261 values ϕ and ψ to produce a structurally meaningful number. This number depends on the fact that

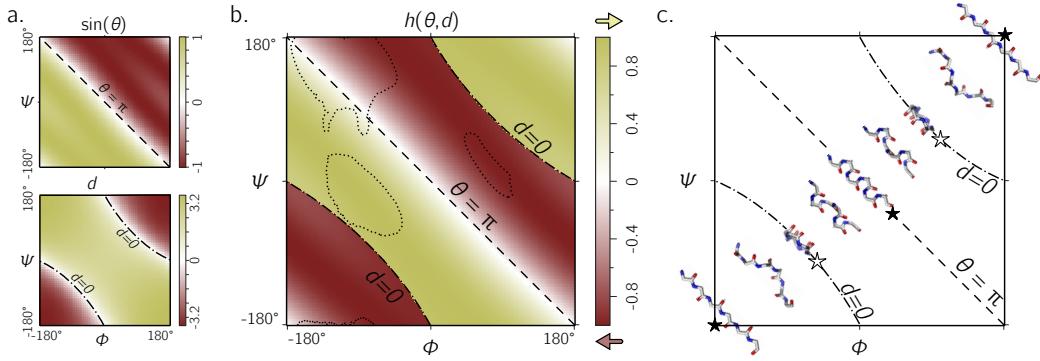


Figure 5. The chirality of an ordered *trans* peptide within the Ramachandran plot. Panel (a) displays the relationship between backbone parameters (ϕ, ψ) and the associated helix parameters of curvature $\sin(\theta)$ (top; Eqn. 1) and pitch d (bottom; Eqn. 2). The boundaries, $\theta = \pi$ (‘ $--$ ’) and $d = 0$ (‘ $- -$ ’), correspond to backbones that are equally flat, but which are respectively optimally extended and curved (see discussion in text). As shown in Fig. 3, the handedness of a helix is a function of these two variables (h ; Eqn. 10). Panel (b) is a map of backbone chirality (h) as a function of ϕ and ψ . This map shows that the naïve expectation of handedness in a Ramachandran plot (Fig. 1d) is inaccurate. Interestingly, our naïve expectations would be upheld if one were only to have sampled regions of the Ramachandran plot dominated by known proteins (a; regions enclosed by ‘ $....$ ’ indicate 90% occupancy). An example of the behavior of one ‘slice’ of (b) is shown in (c). Each snapshot represents a peptide backbone that is either in a distinct region of handedness or at a boundary.

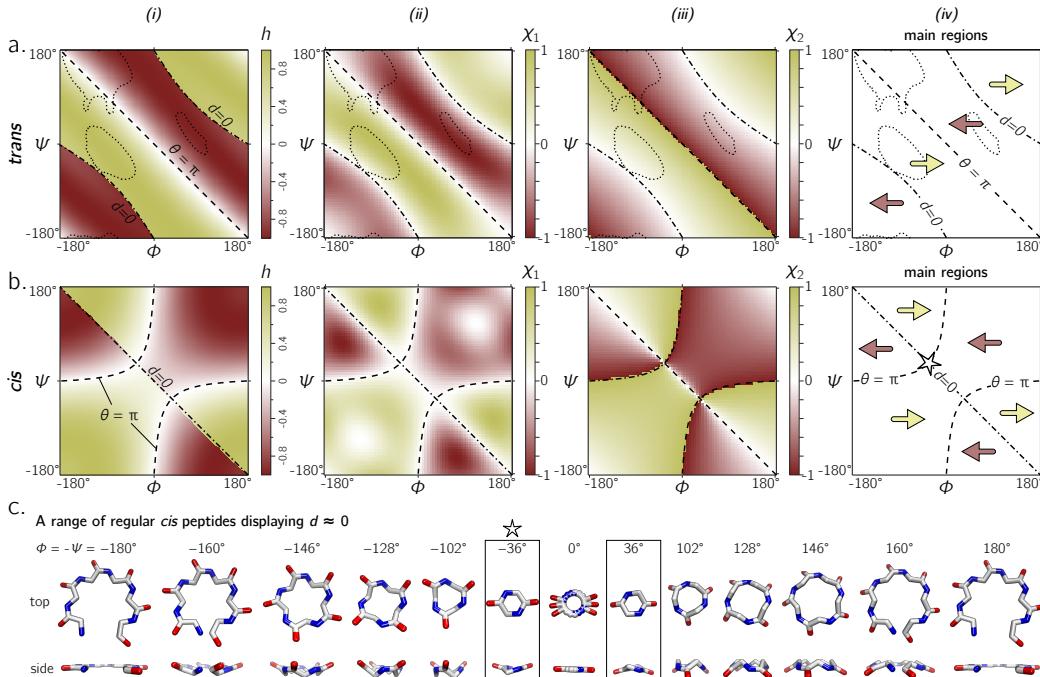


Figure 6. Panels (a) and (b) describe the handedness of backbone twists whose amide dihedral angles are *trans* ($\omega = \pi$) and *cis* ($\omega = 0$), respectively. Column (i) describes handedness calculated using h (Eqn. 10), which does not require structures to be computationally generated. Columns (ii) and (iii) respectively show vector-based estimates of backbone handedness – χ_1 (Eqn. 11) and χ_2 (Eqn. 12) – which are calculated from computationally generated peptides (see Methods). Regions of left- and right-handedness are identical for all measures (i–iii). This general map of how handedness is distributed within the Ramachandran plot is shown in (iv). Finally, Panel (c) displays a range of regular *cis* peptide backbones with $d \approx 0$. As explained by Fig. 4, $d = 0$ indicates a flatness of the peptide perpendicular to the helical axis, which results in peptides that close in on themselves at varying radii. Interestingly, a point in the Ramachandran plot exists exclusively for *cis* peptides, where $d = 0$ and $\theta = \pi$: $\phi = -\psi = \pm 36^\circ$ [\star in Panels(b)–(iv) and (c)].

structural features of the backbone (e.g., radius of gyration) vary least when one slices through the *trans* Ramachandran plot along negative-sloping lines that conserve $\phi + \psi$ (Ho *et al.*, 2003; Zacharias and Knapp, 2013; Mannige *et al.*, 2016). Interestingly, θ follows that trend too, which – in combination with the fact that regions of the Ramachandran plot are sparse (Mannige *et al.*, 2016) – means that θ and its derivatives (e.g., h) are *universal* Ramachandran numbers. The universality arises from the fact that *cis* Ramachandran plots do *not* conserve structure along lines that conserve $\phi + \psi$ (and so \mathcal{R} only works for *trans* backbones), yet any two backbones with nearly identical θ 's will also be conserved in structure (see, e.g., Fig. 5a, top). This feature of θ will be true irrespective of the nature of the amide dihedral angle ω (Eqn. 1).

Handedness of *trans* backbones

Fig. 5a describes the behavior of d and $\sin(\theta)$ as a function of ϕ and ψ (assuming an all-*trans* backbone; $\omega = \pi$ or 180°). Fig. 5b describes the behavior of handedness (h ; Eqn. 10) as a function of ϕ and ψ . This map is a complete description of the handedness of an all-*trans* (regular) peptide backbone. Fig. 5c describes some structures at various regions within the plot. As discussed above (Fig. 4), $d = 0$ (' $\star\star$ ') indicates that each residue is at the same ‘altitude’, i.e., the helix is perfectly flat and maximally curved (at that particular θ). Note that any path on the Ramachandran plot that transitions from negative to positive d will encounter an infinitesimal region in its path that is $d = 0$: it therefore appears as a sharp transition from one handedness to the other. When $\theta = \pi$, then the backbone is also flat (see ‘ \star ’ in Fig. 5c); however, atoms of the same type lie in a single plane that is perpendicular to the helical axis (Fig. 4). In short, within the Ramachandran plot, $d = 0$ (' $--$ ') and $h = \pi$ (' $--$ ') code for *flat* backbones that are respectively either optimally curved (at a given θ) or optimally extended (at a given d). A future report will discuss how these simple rules may be combined to make conjectures about novel secondary and tertiary structures.

Fig. 6a shows that the equation for h match other metrics for handedness, as interpreted by other metrics for chirality (Kwiecińska and Cieplak, 2005; Kabsch and Sander, 1983; Gruziel *et al.*, 2013). In particular, Fig. 6a displays the Ramachandran plot colored by h [(i); Eqn. 10] next to estimates calculated using χ_1 [(ii); Eqn. 11] and χ_2 [(iii); Eqn. 12]. Each panel describes identical regions of left- and right-handedness, which is shown as a cartoon in (iv). However, given that χ_1 and χ_2 are estimates of chirality and not backbone handedness, their exact values differ from the primary metric for handedness (h) provided here.

Handedness of *cis* backbones

In the same vein as Fig. 6a, Fig. 6b displays h , χ_1 and χ_2 as a function of ϕ and ψ for all-*cis* backbones. This appears to be the first complete description of chirality of an all-*cis* backbone ($\omega = 0$). Interestingly, the boundaries for $d = 0$ and $\theta = \pi$ switch in *cis* backbones, with the -ve diagonal and curved boundaries being caused by d and θ , respectively. Additionally, Fig. 6a reiterates the idea that *cis* peptides are quite different when compared to *trans* peptides: the regions and boundaries of left- and right-handedness within the Ramachandran plot differ for *cis* versus *trans*. Finally, points on the *cis* map ($\phi = \pm 36^\circ$, $\psi = \mp 36^\circ$) exist where $d = 0$ and $\theta = \pi$. An example of this, along with other $d = 0$ configurations, is shown in Fig. 6c for a six-residue peptide.

At first, this appears to be contradiction, because $d = 0$ indicates the most curved backbone at a fixed θ , and $\theta = \pi$ indicates the most linear backbone at a fixed d . Of course, this structure would only be possible for cyclic peptides with length two, given that any peptoid of length greater than two would result in many overlapping atoms. However, such a structure (one with $d = 0$ and $\theta = \pi$) is not possible in *trans* peptides, even in theory, because the boundaries associated with $d = 0$ and $\theta = \pi$ do not intersect (Fig. 6a); this is also evident in Fig. 4, where *trans* peptides are shown to not occupy regions of $[d, \theta] = [0, \pi]$, while *cis* peptides do.

How does the amide backbone (ω) change the chirality landscape?

So far, only discrete values for ω are discussed (either *cis* or *trans* peptides respectively corresponding to $\omega = 0$ and π). The picture that arises is that our naïve picture of chirality – that the -ve diagonal separates the right-twisting backbones from the left-twisting backbones (Fig. 1d) is wrong. In fact, the Ramachandran plots for both *cis* and *trans* backbones describe not two but four regions of distinct chirality (Fig. 6). However, deviations from $\omega = 0$ or π are evident in the Protein Databank; see, e.g., discussions by Imrota *et al.* (2011). This raises the question of how does varying ω change the landscape? Fig. 7

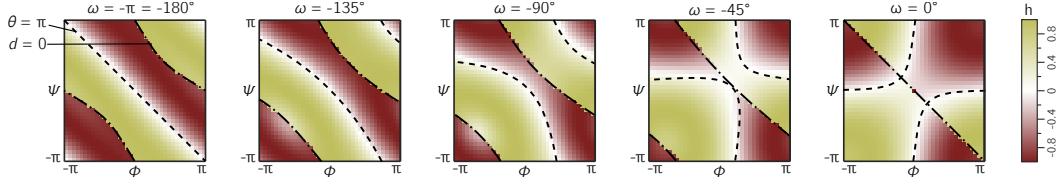


Figure 7. The landscape of backbone chirality as a function of amide dihedral angle ω . As ω is changed, the features of the landscape smoothly transform from the landscapes of $\omega = \pm\pi$ to $\omega = 0$. For all values of ω , it is evident that the naïve view of chirality (Fig. 1d) is wrong: at least four distinct regions of chirality [separated by boundaries $d = 0$ and $\sin(\theta) = \pi$] are evident in each scenario. Although only five snapshots (values of ω) are shown, all integer values of ω were tested, which corroborates the fact that the naïve view of backbone handedness (Fig. 1d) is universally incorrect.

describes some Ramachandran plots that show chirality in terms of varying values for ω . Fig. 7 shows that our naïve expectations of handedness (Fig. 1d) is too simple, irrespective of amide dihedral angle.

[-π,π] or [0,2π]: which frame of reference to use?

In structural biology, ϕ and ψ within the Ramachandran plot has been historically set to range between the values $[-\pi, \pi]$ radians [see, e.g., textbooks by Berg *et al.* (2010) and Alberts *et al.* (2002)]. However, Ramachandran *et al.* (1963) had originally used the range of $[0, 2\pi]$. Today, the range $[-\pi, \pi]$ is used predominantly by structural biologists (Laskowski *et al.*, 1993; Laskowski, 2003; Zacharias and Knapp, 2013), while some have turned to $[0, 2\pi]$ as the norm (Némethy *et al.*, 1966; Voelz *et al.*, 2011).

Given the periodicity of the Ramachandran plot, these differences do not make significant difference scientifically; however the value of the Ramachandran plot is undeniably relative to its boundaries: it is a map of important features of proteins relative to the various regions, quadrants, and diagonals in the map [see, e.g., discussions by Beck *et al.* (2008)]. The Ramachandran plot's value lies in being able to convey large amounts of information in easy to read logo- or picto-grams. For that reason, switching the map from one range to another means that two norms and two types of scientists will not be able to converse as seamlessly.

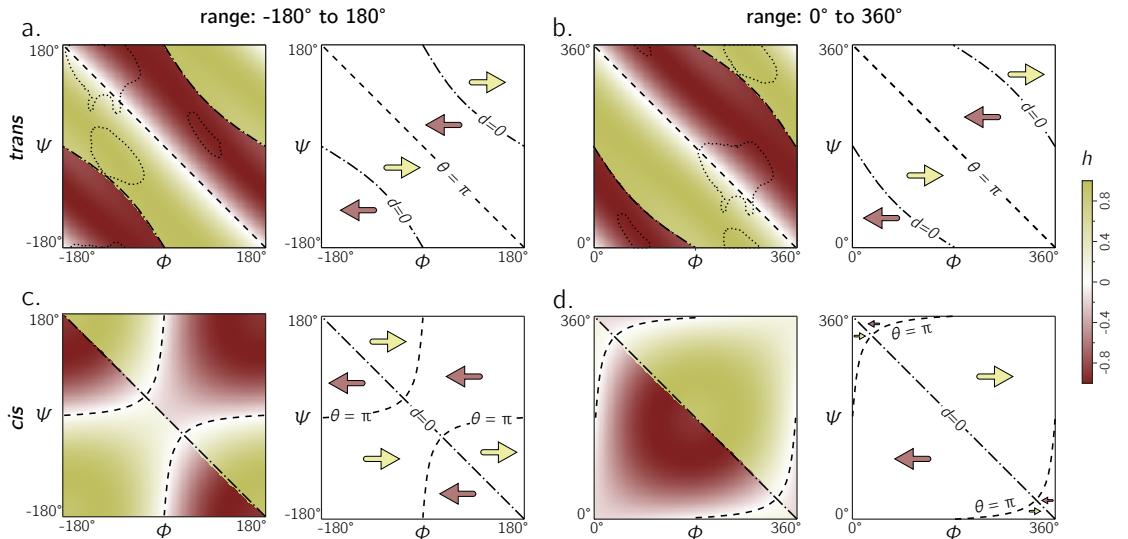


Figure 8. Each panel describes handedness data (right) and boundaries (left). Both frames of references $[-\pi, \dots, \pi]$ and $[0, \dots, 2\pi]$ yield similar trends for *trans* backbones (a,b); however, for *cis* backbones, the latter frame of reference (d) appears to more neatly apportion the handedness of the backbone rather than the traditional frame of reference (c). As in Figs. 5 and 6, ‘ $--$ ’ and ‘ $-.-$ ’ respectively correspond to boundaries defined by $\theta = \pi$ and $d = 0$. Also, regions bound by dotted contours indicate dominant regions within which proteins reside ($p = 0.9$).

Key: left-handed right-handed peptide occupancy (90%) regions within a Ramachandran plot (*trans*)

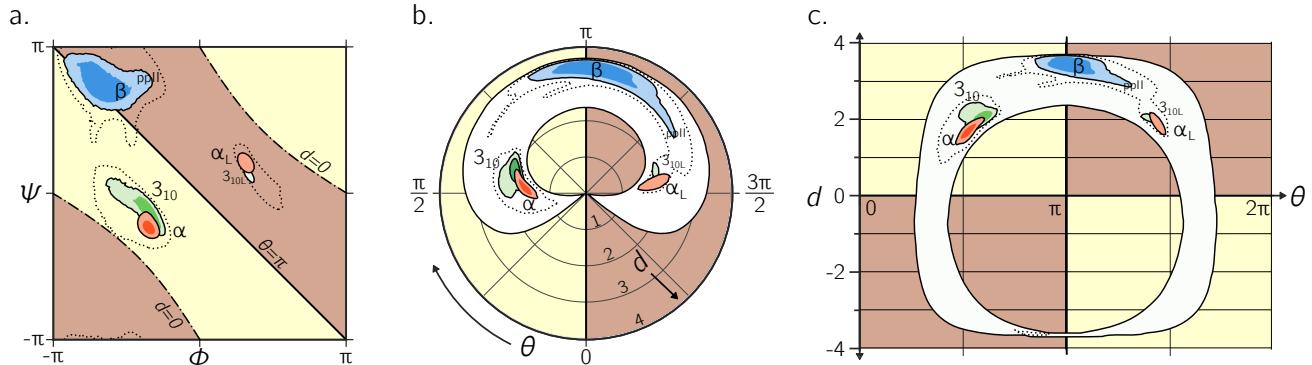


Figure 9. Alternative representations of the Ramachandran plot. While the Ramachandran plot is useful to map characteristics of secondary structures (a), it is not intuitive. For example, the relationship between the Ramachandran parameters (ϕ , ψ) and the handedness of a backbone is not obvious (see, e.g., the non-obvious distribution of left- and right-handed peptides as a function of ϕ and ψ). For this reason, Zacharias and Knapp (2013) introduced a graphical format involving the helical parameters d and θ in polar coordinate space (b), where the regions of left- and right-handedness are obvious [their format differs from (b) in that their θ increases in counter-clockwise fashion]. Panel (c), which is an extension of Fig. 4, introduces another graphical representation of backbone degrees of freedom based on (θ, d) , but in Cartesian space. While both (b) and (c) are equally useful in understanding regions available to a protein, the text discusses some benefits of (c) as a universal map for exploring new conformations and secondary structures. Excepting the left-handed helices (α_L -, 3_{10L} -helices), each secondary structure has two contours signifying $p = 0.5$ and 0.8 .

330 Therefore, the following question must arise: which range $[-\pi, \pi]$ or $[0, 2\pi]$ – is able to convey
 331 more information with the least amount of effort? Fig. 8 shows the behavior of a *trans* backbone (a,b) and
 332 *cis* backbone (c,d) in the two frames of reference. From (a) and (b) it is evident that general trends in
 333 the map for *trans* backbones remain the same in both frames of reference: the negative diagonal ($\sigma = \pi$)
 334 locally separates right-handed regions from left-handed regions, while the curved line ($d = 0$) – which
 335 also separates handedness – also appears to be in generally the same regions (albeit inverted in curvature).
 336 The *cis* backbones, however, look dramatically different in the two frames of reference: the range $[-\pi, \pi]$
 337 separates handedness in a more complicated manner (c), while, for the most part, the -ve diagonal appears
 338 to meaningfully separate handedness when the plot ranges from 0 to 2π (d). For this reason, purely when
 339 looking at handedness, and especially in the case of *cis* backbones, the Ramachandran plot that ranges
 340 between 0 and 2π appears to be more meaningful.

341 A universal alternative to the Ramachandran plot

342 While the Ramachandran plot is useful enough to earn a place in undergraduate-level biology textbooks,
 343 as discussed throughout this report, it is not easy to estimate features of a peptide backbone just from its
 344 (ϕ , ψ) angles (Fig. 9a). This prompted Zacharias and Knapp (2013) to introduce a new representation
 345 for backbone degrees of freedom in the form of a polar graph. In this polar representation, the θ is the
 346 angular coordinate (azimuth) and d is the radial coordinate. An example of one such representation is
 347 shown in Fig. 9b, with the direction of increasing θ reversed (compared to the cited report) to maintain
 348 relative positions of secondary structures within the Ramachandran plot (Fig. 9a). Zacharias and Knapp
 349 (2013) stated an additional reason for the introduction of the polar representation (Fig. 9b): θ , which is an
 350 angle and therefore periodic, can remain periodic as the angular coordinate in the graph.

351 However, the format proposed by Zacharias and Knapp (2013) (Fig. 9b) is incomplete for a few
 352 reasons: 1) $d < 0$ peptides (the bottom-left and top-right regions of Fig. 5a, bottom) will never be observed
 353 in this map since only structures with $d \geq 0$ are allowed; 2) all peptides with $d = 0$ (marked by ‘- -’
 354 in every preceding Ramachandran plot) will be compressed into one point at the center, even though
 355 Fig. 6c shows a range of legitimate $d = 0$ conformations; 3) while the graph is θ -periodic, the values for
 356 θ in peptides are constrained within one $[0, 2\pi]$ period (peptides range between $\theta = \pi/4$ and $2\pi - \pi/4$;
 357 vertical dotted lines in Fig. 4); i.e., periodicity in θ is not required for the faithful representation of
 358 peptides. Fortunately, even though this system is not universal (again, since $d < 0$ structures are not

359 accommodated), most conformations in globular proteins display positive d , and so the representation
360 presented by Zacharias and Knapp (2013) is a reasonable one for most proteins with known structure.

361 Interestingly, Fig. 9c – which arranges the parameters θ and d along Cartesian axes – serves as both a
362 universal *and* intuitive map for peptide backbone geometry. This is because: 1) as shown in Fig. 4, such
363 maps reveal a wealth of information about the peptide backbone, 2) both positive and negative values of d
364 are allowed (compared to Fig. 9b), due to the shift in the coordinate system from polar to Cartesian, and 3)
365 this format accommodates *every* type of peptide conformation: any peptide (or its mimic) has a place in
366 this map irrespective of whether the amide backbone is *cis* or *trans* or any other value; additionally, if the
367 backbone is distorted, such distortions can also be accounted for since d and θ includes such distortions
368 (Eqns. 1 and 2). This is impossible to do using a single Ramachandran plot without making sweeping
369 assumptions about backbone parameters that are not ϕ and ψ . The (θ, d) plot opens up the possibility for
370 a new, intuitive, and *universal* kind of graphical representation as a supplement to the Ramachandran plot.
371

372 A departure from perfect regularity

373 So far, this report has focused on *simple* regular backbone conformations: those that are formed from
374 the same ϕ and ψ angles repeated along the backbone. The text so far discusses the treatment of regular
375 backbones particularly because a simple correspondence exists between a regular backbone (described by
376 myriad internal coordinates) and a helix that is described simply by (d, θ, ρ_i) . This simple correspondence
377 allows for an intuitive understanding as to how to envision regular backbones as a function of d and θ .
378 There is a possibility that d and θ are useful even in isolation, when the unreasonable constraint of perfect
379 backbone regularity is lifted. Below we provide an example of such a departure from regularity, along
380 with a guide regarding how distinct but consecutive (d, θ) 's may be combined to great effect.

381 Some secondary structures are characterized by the regular combination of two or more sets of (ϕ, ψ)
382 pairs (Pauling and Corey, 1951b,a; Armen *et al.*, 2004; Daggett, 2006; Hayward and Milner-White,
383 2008; Mannige *et al.*, 2015, 2016). For example, the Σ -strand is constructed by alternating between
384 two backbone states $(\phi, \psi, \omega) = (-A, B, 180^\circ)$ and $(A, -B, 180^\circ)$, where $A \approx 120$ and $B \approx 90$ [Fig. 4h in
385 Mannige *et al.* (2015)]. It was found that the two states are equal in the extent to which the backbone
386 twists, but opposite in handedness, which allows for these secondary structures to remain linear, albeit
387 in a meandering way (Mannige *et al.*, 2015). Interestingly, from Eqn. 10, we easily find that both states
388 have handedness (h) equal in magnitude but opposite in sign/handedness (the two h 's are approximately
389 ± 0.5144). Similarly, the α -sheet proposed by Pauling and Corey (1951a) is constructed by alternating
390 between α and α_L backbone states, yet this motif is linear because each state describes equal but opposite
391 handedness $h = \pm 0.9979$. These points raise the idea that, even in the absence of perfect backbone
392 regularity, the values d and θ may be considered to be residue-specific properties that may be combined
393 to readily provide insights about higher order structures.

394 CONCLUSIONS

395 This report introduces a metric for backbone handedness (h) that is based on modeling the backbone as
396 a helix [Fig. 2; Miyazawa (1961)]. In particular, h , which is a combination of the helical parameters
397 θ (angular displacement) and d (vertical displacement), ranges from -1 and 1, and is negative (or
398 positive) when the backbone twist is left(or right)-handed (with larger values indicating greater extent of
399 twistedness). This metric (h) was used to characterize every regular backbone's twist, via Ramachandran
400 plots, for both *cis* and *trans* peptides. In doing so, this report dispels a naïve view of handedness (Fig. 1d),
401 which states that backbone handedness in the Ramachandran plot is separated by the negative-sloped (-ve)
402 diagonal. Interestingly, the reason for the naïve view makes sense when considering only *trans* peptides:
403 the -ve diagonal ('--' in Figs. 5a) separates *D* and *L* twists if we consider only the regions dominantly
404 occupied by structured proteins ('....' in Figs. 5a). Plotting the backbone handedness (h) in the common
405 frames of reference – $\phi, \psi \in [-\pi, \pi]$ and $[0, 2\pi]$ – indicates that the less commonly used frame $[0, 2\pi]$
406 may be more appropriate for interpreting *cis* backbones (Fig. 8).

407 The behavior of a backbone in *cis* and *trans* Ramachandran plots look dramatically different (Fig. 6),
408 and so scientists dealing with new structures having a combination of *cis* and *trans* backbones can not
409 use one Ramachandran plot to faithfully describe these structures. Interestingly, the parameters θ and d
410 combine all features (internal coordinates) of a contorting backbone, including the amide dihedral angle
411 ω , which means that (θ, d) can describe *any* peptide backbone, irrespective of ω . The Cartesian plot with

412 θ and d as the x- and y-axis, respectively, serves as a unique plot for *any* peptide backbone (Fig. 9), with
413 specific values and boundaries containing deep structural meaning (Fig. 4). These discussions, the author
414 hopes, clarify a number of concepts associated with the Ramachandran plot, while providing new insights
415 into how to interrogate the features of new protein and protein-like structures.

416 ACKNOWLEDGMENTS

417 RVM was *partially* supported by the Defense Threat Reduction Agency under contract no. IACRO-
418 B0845281. RVM thanks Alana Canfield Mannige for her support and input. This work was done *partially*
419 at the Molecular Foundry at Lawrence Berkeley National Laboratory (LBNL), supported by the Office
420 of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No.
421 DE-AC02-05CH11231.

422 REFERENCES

- 423 **Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P.** 2002. Molecular biology of the cell.
424 new york: Garland science; 2002. *Classic textbook now in its 5th Edition* .
- 425 **Armen RS, DeMarco ML, Alonso DO, Daggett V.** 2004. Pauling and corey's α -pleated sheet structure
426 may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proceedings of the National
427 Academy of Sciences of the United States of America* **101**(32):11622–11627.
- 428 **Beck DA, Alonso DO, Inoyama D, Daggett V.** 2008. The intrinsic conformational propensities of the
429 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the
430 National Academy of Sciences* **105**(34):12259–12264.
- 431 **Bentley R.** 2010. Chiral: a confusing etymology. *Chirality* **22**(1):1–2.
- 432 **Berg JM, Tymoczko JL, Stryer L.** 2010. *Biochemistry, International Edition*. WH Freeman & Co.,
433 New York, 7 edition.
- 434 **Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov IN, Bourne PE.**
435 2000. The protein data bank. *Nucleic acids research* **28**(1):235–242.
- 436 **Bragg L, Kendrew JC, Perutz MF.** 1950. Polypeptide chain configurations in crystalline proteins. *Pro-
437 ceedings of the Royal Society of London Series A Mathematical and Physical Sciences* **203**(1074):321–
438 357.
- 439 **Chothia C, Hubbard T, Hubbard S, Barns H, Murzin A.** 1997. Protein folds in the all-beta and
440 all-alpha classes. *Annu Rev Biophys Biomol Struct* **26**:597–627.
- 441 **Cock P, Antao T, Chang J, Chapman B, Cox C, Dalke A, Friedberg I, Hamelryck T, Kauff F,
442 Wilczynski B, de Hoon M.** 2009. Biopython: freely available python tools for computational molecular
443 biology and bioinformatics. *Bioinformatics* **25**(11):1422–1423.
- 444 **Cooper GM, Hausman RE.** 2013. *The Cell: A Molecular Approach*. Sinauer Associates, Inc., Sunderland,
445 MA, 6 edition.
- 446 **Cross L, Klyne W.** 2013. *Rules for the Nomenclature of Organic Chemistry: Section E: Stereochemistry
447 (Recommendations 1974)*. Elsevier.
- 448 **Daggett V.** 2006. α -sheet: the toxic conformer in amyloid diseases? *Accounts of chemical research*
449 **39**(9):594–602.
- 450 **Eisenberg D.** 2003. The discovery of the α -helix and β -sheet, the principal structural features of proteins.
451 *Proceedings of the National Academy of Sciences* **100**(20):11207–11210.
- 452 **Esposito L, Balasco N, De Simone A, Berisio R, Vitagliano L.** 2013. Interplay between peptide bond
453 geometrical parameters in nonglobular structural contexts. *BioMed research international* **2013**.
- 454 **Ferrarini A, Nordio PL.** 1998. On the assessment of molecular chirality. *Journal of the Chemical
455 Society, Perkin Transactions 2* (2):455–460.
- 456 **Fox NK, Brenner SE, Chandonia JM.** 2014. Scope: Structural classification of proteins—extended,
457 integrating scop and astral data and classification of new structures. *Nucleic Acids Res* **42**(Database
458 issue):D304–D309. doi:10.1093/nar/gkt1240.
- 459 **Frishman D, Argos P.** 1995. Knowledge-based protein secondary structure assignment. *Proteins:
460 Structure, Function, and Bioinformatics* **23**(4):566–579.
- 461 **Gold V, Loening K, McNaught A, Shemi P.** 1997. Iupac compendium of chemical terminology.
462 *Blackwell Science, Oxford* .

- 463 **Gorske BC, Mumford EM, Conry RR.** 2016. Tandem incorporation of enantiomeric residues engenders
464 discrete peptoid structures. *Organic letters*.
- 465 **Gruziel M, Dzwolak W, Szymczak P.** 2013. Chirality inversions in self-assembly of fibrillar superstruc-
466 tures: a computational study. *Soft Matter* **9**(33):8005–8013.
- 467 **Harris AB, Kamien RD, Lubensky TC.** 1999. Molecular chirality and chiral parameters. *Reviews of*
468 *Modern Physics* **71**(5):1745.
- 469 **Hayward S, Milner-White EJ.** 2008. The geometry of α -sheet: Implications for its possible function as
470 amyloid precursor in proteins. *Proteins: Structure, Function, and Bioinformatics* **71**(1):415–425.
- 471 **Ho BK, Thomas A, Brasseur R.** 2003. Revisiting the ramachandran plot: Hard-sphere repulsion,
472 electrostatics, and h-bonding in the α -helix. *Protein Science* **12**(11):2508–2522.
- 473 **Hooft RW, Sander C, Vriend G.** 1997. Objectively judging the quality of a protein structure from a
474 ramachandran plot. *Computer applications in the biosciences: CABIOS* **13**(4):425–430.
- 475 **Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK.** 2002. Intrinsic disorder in
476 cell-signaling and cancer-associated proteins. *Journal of molecular biology* **323**(3):573–584.
- 477 **Improta R, Vitagliano L, Esposito L.** 2011. Peptide bond distortions from planarity: new insights from
478 quantum mechanical calculations and peptide/protein crystal structures. *PLoS One* **6**(9):e24533.
- 479 **Improta R, Vitagliano L, Esposito L.** 2015a. Bond distances in polypeptide backbones depend on the
480 local conformation. *Acta Crystallographica Section D: Biological Crystallography* **71**(6):1272–1283.
- 481 **Improta R, Vitagliano L, Esposito L.** 2015b. The determinants of bond angle variability in pro-
482 tein/peptide backbones: A comprehensive statistical/quantum mechanics analysis. *Proteins: Structure,*
483 *Function, and Bioinformatics* **83**(11):1973–1986.
- 484 **Kabsch W, Sander C.** 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-
485 bonded and geometrical features. *Biopolymers* **22**(12):2577–2637. doi:10.1002/bip.360221211.
- 486 **Kwiecińska JI, Cieplak M.** 2005. Chirality and protein folding. *Journal of Physics: Condensed Matter*
487 **17**(18):S1565.
- 488 **Laskowski RA.** 2003. Structural quality assurance. *Structural Bioinformatics, Volume 44* pages 273–303.
- 489 **Laskowski RA, MacArthur MW, Moss DS, Thornton JM.** 1993. Procheck: a program to check the
490 stereochemical quality of protein structures. *Journal of applied crystallography* **26**(2):283–291.
- 491 **Linderstrøm-Lang KU.** 1952. *Lane Medical Lectures: proteins and enzymes*, volume 6. Stanford
492 University Press.
- 493 **Mannige RV.** 2014. Dynamic new world: Refining our view of protein structure, function and evolution.
494 *Proteomes* **2**(1):128–153.
- 495 **Mannige RV, Haxton TK, Proulx C, Robertson EJ, Battigelli A, Butterfoss GL, Zuckermann RN,**
496 **Whitelam S.** 2015. Peptoid nanosheets exhibit a new secondary structure motif. *Nature* **526**:415–420.
- 497 **Mannige RV, Kundu J, Whitelam S.** 2016. The Ramachandran number: an order parameter for protein
498 geometry. *PLoS One* **11**(8):e0160023.
- 499 **Mirijanian DT, Mannige RV, Zuckermann RN, Whitelam S.** 2014. Development and use of an
500 atomistic charmm-based forcefield for peptoid simulation. *J Comput Chem* **35**(5):360–370.
- 501 **Mislow K.** 2002. Stereochemical terminology and its discontents. *Chirality* **14**(2-3):126–134.
- 502 **Miyazawa T.** 1961. Molecular vibrations and structure of high polymers. ii. helical parameters of infinite
503 polymer chains as functions of bond lengths, bond angles, and internal rotation angles. *Journal of*
504 *Polymer Science* **55**(161):215–231.
- 505 **Neal MP, Solymosi M, Wilson MR, Earl DJ.** 2003. Helical twisting power and scaled chiral indices.
506 *The Journal of chemical physics* **119**(6):3567–3573.
- 507 **Némethy G, Leach S, Scheraga HA.** 1966. The influence of amino acid side chains on the free energy
508 of helix-coil transitions1. *The Journal of Physical Chemistry* **70**(4):998–1004.
- 509 **Orosz F, Ovádi J.** 2011. Proteins without 3d structure: definition, detection and beyond. *Bioinformatics*
510 **27**(11):1449–1454.
- 511 **Osipov M, Pickup B, Dunmur D.** 1995. A new twist to molecular chirality: intrinsic chirality indices.
512 *Molecular Physics* **84**(6):1193–1206.
- 513 **Pauling L, Corey RB.** 1951a. Configurations of polypeptide chains with favored orientations around
514 single bonds: two new pleated sheets. *Proceedings of the National Academy of Sciences of the United*
515 *States of America* **37**(11):729.
- 516 **Pauling L, Corey RB.** 1951b. The pleated sheet, a new layer configuration of polypeptide chains.
517 *Proceedings of the National Academy of Sciences of the United States of America* **37**(5):251.

- 518 **Pauling L, Corey RB, Branson HR.** 1951. The structure of proteins: two hydrogen-bonded helical
519 configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* **37**(4):205–
520 211.
- 521 **Ramachandran G, Ramakrishnan C, Sasisekharan V.** 1963. Stereochemistry of polypeptide chain
522 configurations. *Journal of molecular biology* **7**(1):95–99.
- 523 **Robertson EJ, Battigelli A, Proulx C, Mannige RV, Haxton TK, Yun L, Whitelam S, Zuckermann
524 RN.** 2016. Design, synthesis, assembly, and engineering of peptoid nanosheets. *Accounts of chemical
525 research* **49**(3):379–389.
- 526 **Shimanouchi T, Mizushima Si.** 1955. On the helical configuration of a polymer chain. *The Journal of
527 Chemical Physics* **23**(4):707–711.
- 528 **Solymosi M, Low RJ, Grayson M, Neal MP.** 2002. A generalized scaling of a chiral index for molecules.
529 *The Journal of chemical physics* **116**(22):9875–9881.
- 530 **Sun J, Zuckermann RN.** 2013. Peptoid polymers: a highly designable bioinspired material. *ACS Nano*
531 **7**(6):4715–4732. doi:10.1021/nn4015714.
- 532 **Testa B.** 2013. Organic stereochemistry. part 2. *Helvetica Chimica Acta* **96**(2):159–188.
- 533 **Tien MZ, Sydykova DK, Meyer AG, Wilke CO.** 2013. Peptidebuilder: A simple python library to
534 generate model peptides. *PeerJ* **1**:e80.
- 535 **Van Der Walt S, Colbert SC, Varoquaux G.** 2011. The numpy array: a structure for efficient numerical
536 computation. *Computing in Science & Engineering* **13**(2):22–30.
- 537 **Voelz VA, Dill KA, Chorny I.** 2011. Peptoid conformational free energy landscapes from implicit-solvent
538 molecular simulations in amber. *Peptide Science* **96**(5):639–650.
- 539 **Wallentin CJ, Orentas E, Wärnmark K, Wendt OF.** 2009. Chirality, a never-ending source of con-
540 fusion. *Zeitschrift für Kristallographie International journal for structural, physical, and chemical
541 aspects of crystalline materials* **224**(12):607–608.
- 542 **Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT.** 2004. Prediction and functional analysis of
543 native disorder in proteins from the three kingdoms of life. *J Mol Biol* **337**(3):635–645.
- 544 **Zacharias J, Knapp EW.** 2013. Geometry motivated alternative view on local protein backbone
545 structures. *Protein Science* **22**(11):1669–1674.