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

Ranjan V. Mannige

Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, United States
Multiscale Institute, Redwood City, CA, United States

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, Ramakrishnan & Sasisekharan, 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 $\varphi = \psi = -100°$ 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 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 equals $\sin(\theta)d/|d|$, 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* ($\omega = 0°$) and *trans* ($\omega = 180°$) 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 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 & 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
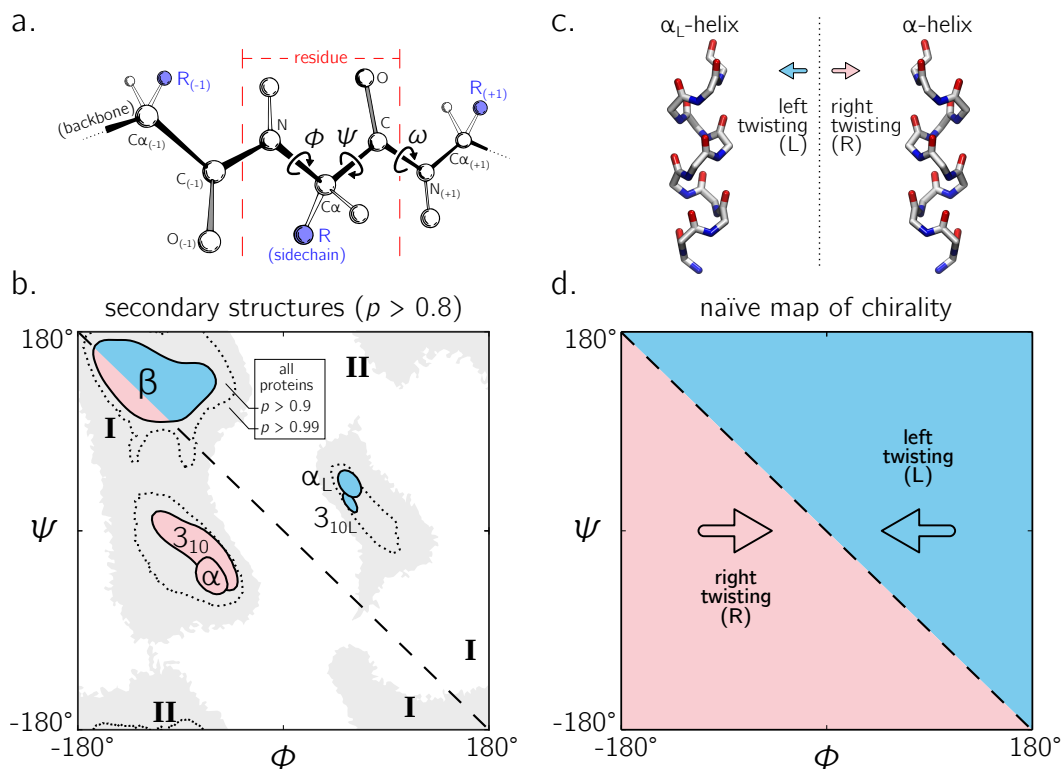
Figure 1 **The backbone of a single residue (A) can be described by its dihedral angles ϕ and ψ (and in smaller part, ω, which is predominantly *trans* or $\sim180°$).** The Ramachandran plot is important because a number of regular conformations important to biology—secondary structures—are located at specific regions of the plot (B). For the most part, regular peptide backbones twist in either a left-handed or right-handed fashion; examples are shown in (C). As evidenced in (B), 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 & 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 ($\omega = \pi$). (A) is modified from *Mannige, Kundu & Whitelam (2016)*. Due to low incidence within the studied database (see Methods), the two left-handed helices in (B) are arbitrarily marked and have no statistical significance. All molecular representations in this text are shown in 'licorice' form, with the colors red, blue and white representing oxygen, nitrogen and carbon atoms.

the structure of a protein goes a long way towards understanding how a protein functions (*Alberts et al., 2002*; *Berg, Tymoczko & Stryer, 2010*). While the conformation of a peptide backbone is dependent on a number of parameters (bond lengths, bond angles, and dihedral angles), *Ramachandran, Ramakrishnan & Sasisekharan (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, Kendrew & Perutz, 1950*; *Pauling & Corey, 1951b*; *Pauling, Corey & Branson, 1951*; *Linderstrøm-Lang, 1952*; *Laskowski et al., 1993*; *Chothia et al., 1997*; *Hooft, Sander & Vriend, 1997*;

Cooper & Hausman, 2013; *Alberts et al., 2002*; *Laskowski, 2003*; *Ho, Thomas & Brasseur, 2003*; *Eisenberg, 2003*; *Berg, Tymoczko & Stryer, 2010*; *Mannige, Kundu & Whitelam, 2016*). 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 which are strung together by loops that are less regular (*Alberts et al., 2002*; *Berg, Tymoczko & Stryer, 2010*). Each regular peptide structure describes a backbone whose per-residue (ϕ, ψ) values are generally the same, and therefore their 'locations' on the Ramachandran plot act as structural landmarks (Fig. 1B).

So far, our understanding of the Ramachandran plot has been limited mostly to structured proteins that display stable conformations (*Berman et al., 2000*; *Alberts et al., 2002*). These types of proteins occupy only a limited region of the plot (dotted regions in Fig. 1B). The regular backbone conformations in these regions are well understood. For example, known regular structures that are to the right of the negatively sloping diagonal (dashed line in Fig. 1B; henceforth denoted as the ' −ve diagonal') are left-handed in backbone twist, while those that are to the left of the diagonal are right-handed (left- and right-handed regions are respectively shaded blue and red). For example, the position of the idealized left- and right-handed α-helices (Fig. 1C)—respectively denoted as α_L and α in Fig. 1B—are on opposite sides of the −ve diagonal. The 'naïve view' of handedness, obtained from looking only at structured proteins, would be the expectation that the −ve diagonal neatly separates the Ramachandran plot into regions of left- and right-handedness (Fig. 1B).

However, structured proteins represent only a fraction of functional proteins. Indeed, up to 15% of mammalian proteins are completely disordered—they natively display multiple, often extended, conformations—and up to 50% of the mammalian proteins display large (>30 residue) stretches of disorder (*Iakoucheva et al., 2002*; *Ward et al., 2004*; *Orosz & Ovádi, 2011*; *Mannige, 2014*). Interestingly, when compared to structured backbones, structurally degenerate or disordered backbones occupy many more regions within the Ramachandran plot (*Beck et al., 2008*).

Additionally, a number of peptide mimics—especially peptoids (*Sun & Zuckermann, 2013*)—have been found to 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*)—is believed to sample regions of the Ramachandran plot ('**I**' in Fig. 1B) that are not permitted within natural proteins (this is because peptoid backbones lack hydrogen-bond donors). Another peptoid secondary structure—the 'ω-strand' (*Gorske, Mumford & Conry, 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 display backbone twists that are equal in magnitude but opposite in handedness (for this reason, the Σ-strand is relatively linear, albeit meandering; *Mannige et al., 2015*; *Mannige, Kundu & Whitelam, 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 for both *cis* and *trans* peptides. 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 & Mizushima, 1955*; *Miyazawa, 1961*; *Zacharias & Knapp, 2013*). The metric is used to exhaustively chart the handedness of regular backbones. In doing so, this survey provides a new graphical format to explore new types of secondary structures being discovered (*Mannige et al., 2015*; *Gorske, Mumford & Conry, 2016*). Also, this survey dispels 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. Finally, 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. This work builds on a previous report (*Zacharias & 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 (°). 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, Kamien & Lubensky, 1999*). For example, metrics for chirality have been introduced that focus on vector orientations (*Kwiecińska & Cieplak, 2005*; *Kabsch & Sander, 1983*; *Gruziel, Dzwolak & Szymczak, 2013*), optical activity (*Osipov, Pickup & Dunmur, 1995*), and molecular shape (*Ferrarini & Nordio, 1998*). However, this report will focus on a simpler metric for handedness associated with an idealized helix within which all (regular) backbone atoms of one type sit (Fig. 2; *Shimanouchi & Mizushima (1955)*; *Miyazawa (1961)*; *Zacharias & 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, Corey & Branson, 1951*; *Pauling & Corey, 1951b*; *Pauling & Corey, 1951a*). In particular, *Shimanouchi & Mizushima (1955)*
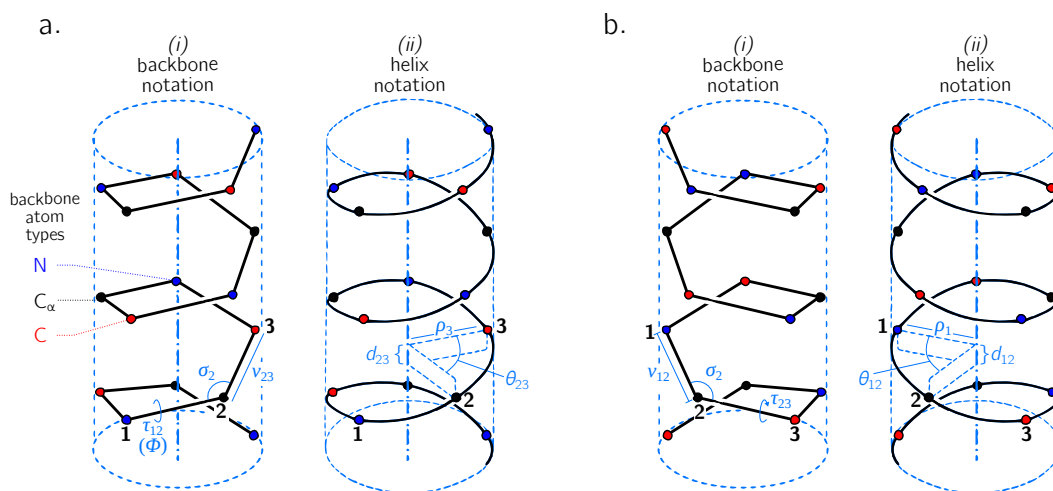
Figure 2 **Internal coordinate *(i)* and helical coordinate *(ii)* representations of right-handed (A) and left-handed (B) regular backbones.** Internal coordinates are a function of bond lengths (e.g., v_{23}), angles (σ_2), and dihedral angles (τ_{12}), while helical coordinates are a function of displacement 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 to the helical axis (ρ_i). Representations are derived from Figs. 1 and 2 in *Shimanouchi & 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 & 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. Figure 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 bonds adjacent to atom i, and dihedral or torsion angles (τ_{ij}), which involve atoms associated with the bond $i-j$ and the two adjacent atoms. Helical coordinates (Fig. 2 *(ii)*) are described using measures of axial displacement between two successive atoms of the same type (d; this is related to the pitch of a platonic helix), angular displacement between two successive atoms of the same type (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 backbone 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 axial 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 & 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 axial displacement (d) and angular displacement (θ) may be retrieved using the following two equations.

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

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

The ranges for d and θ, respectively, are $[-\lambda, +\lambda]$ and $[0, 2\pi)$ (the positive limit λ is defined by allowed values for the various internal coordinates). As above, subscripts 'n', 'α', and 'c' respectively refer to the backbone nitrogen, α-carbon and carbonyl carbon atom types. The dihedral angles ϕ, ψ, and ω represent the traditional symbols for backbone dihedral angles, which may 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

$$2\rho_\alpha^2[1-\cos(\theta)]+d^2 = v_{\alpha,c}^2+v_{c,n}^2+v_{n,\alpha}^2-2v_{c,n}\left[v_{\alpha,c}\cos(\sigma_c)+v_{n,\alpha}\cos(\sigma_n)\right]$$
$$+2v_{\alpha,c}v_{n,\alpha}\left[\cos(\sigma_c)\cos(\sigma_n)-\sin(\sigma_c)\sin(\sigma_n)\cos(\tau_{c,n})\right] \tag{3}$$

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

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

The distance between adjacent α-carbons (d_α) is ~ 3.8 Å for *trans* peptides and ~ 3 Å for *cis* peptides. Other radii (ρ_c, ρ_n) can be obtained by cycling through (α, c, n) subscripts within Eqs. (3) and (4).[1] Note that all ρ_i's are functions of θ and d (along with other internal coordinates), and so one may use two of the three terms in (d, θ, ρ_i)— to describe the helical state of a peptide. Since there is only one d and θ per backbone (compared to three ρ_i's, one per atom type), this report utilizes d and θ as the two descriptors (other discussions on this choice have also been made by *Zacharias & Knapp (2013)*).

[1] ρ_c and ρ_n are obtained by the following subscript conversions: ($\alpha \to c, c \to n, n \to \alpha$) and ($\alpha \to n, c \to \alpha, n \to c$).

Eqs. (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, Ramakrishnan & Sasisekharan, 1963*; *Improta, Vitagliano & Esposito, 2015a*; *Esposito et al., 2013*; *Improta, Vitagliano & Esposito, 2015b*). In particular, most backbone bond lengths and angles display one equilibrium value (*Improta, Vitagliano & Esposito, 2015a*; *Esposito et al., 2013*; *Improta, Vitagliano & Esposito, 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 Eqs. (1) and (2) to arrive at a simpler equation for *trans* backbones. *Zacharias & Knapp (2013)* published an updated version of this set of equations, which follows.[2]

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

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

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 et al., 2014*; *Gorske, Mumford & Conry, 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 Eq. (5) through Eq. (8) are simplifications of Eqs. (1) and (2), and are therefore prone to some limitations that are not present in Eq. (1) and Eq. (2). For example, bond lengths (*Improta, Vitagliano & Esposito, 2015a*) and bond angles (*Esposito et al., 2013*; *Improta, Vitagliano & Esposito, 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 may be strained from their equilibrium values (e.g., due to bulky sidechains), only Eqs. (1) and (2) can be expected to faithfully (and perfectly) represent backbone features such as handedness of twist. However, the approximations of Eq. (5) through Eq. (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 (ϕ, ψ, ω) and (d, θ)

Given a particular *value* of ω, every (ϕ, ψ) pair points to exactly one (d, θ). However, when using Eqs. (1) and (2), one value of ω can not be replaced with a *periodically equivalent* version of ω (the same can be said for ϕ and ψ). 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 Eqs. (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., a one-to-one correspondence between (ϕ, ψ) and (d, θ) is only possible if one insists on specific values for ωs. For this reason, this report proposes to wrap the value of an amide backbone ω' between $[\Delta, \Delta + 360°)$ using

$$\omega = (\omega' - \Delta)\%360 + \Delta, \tag{9}$$

where % represents the modulus function, and Δ describes the start of the range $[\Delta, \Delta + 2\pi)$. Choosing $\Delta = -90°$ would ensure that the distribution of both *cis* ($\omega = 0 \pm 5°$) and *trans* ($\omega = 180 \pm 5°$) 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 Eqs. (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 ϕ and ψ lie within a preset range –e.g., $[0, 2\pi)$ or $[-\pi, \pi)$ –and whose ω does not change after being wrapped by Eq. (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 & Knapp, 2013*), that θ and d together are instrumental in describing backbone handedness.

The relationship between handedness and (d, θ) 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 θ and d. This is because the sign of d provides a frame of reference for interpreting θ. In particular, if d is negative, then $0 < \theta < \pi$ indicates left-handedness (Fig. 3A), while $\pi < \theta < 2\pi$ indicates a right-handed helix (Fig. 3B). However, if d is positive, then the manner in which the helix is 'built' reverses, and $0 < \theta < \pi$ indicates right-handedness (Fig. 3C), while $\pi < \theta < 2\pi$ indicates left-handedness (Fig. 3D).

a. $d < 0; \theta < \pi$ left-handed (L)

b. $d < 0; \theta > \pi$ right-handed (R)

c. $d > 0; \theta < \pi$ right-handed (R)
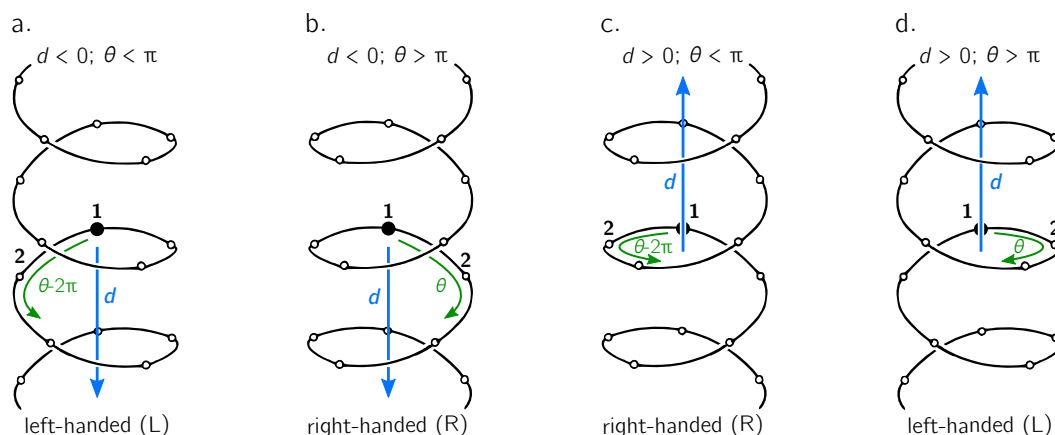
d. $d > 0; \theta > \pi$ left-handed (L)

Figure 3 **The handedness of a helix is a function of angular displacement θ perpendicular to the helical axis (green curved arrows) and linear displacement d along the helical axis (blue, vertical arrows).** Note that left-handed (L) and right-handed (R) 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 & Klyne, 2013*); however, as discussed in the Methods section, this report makes a distinction between helix handedness and molecular chirality.

Given these relationships, this paper proposes a new metric for backbone handedness that depends on the sign of d and the value of θ:

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

The range of h is $[-1, 1]$, with negative (or positive) values indicating that the overall twist of the backbone is left(or right)-handed. Also, $|h|$ is proportional to the extent to which the backbone is twisted. Note that $d/|d|$ is *related* to the traditional sign function $\text{sgn}(d)$, but deviates at $d = 0$, where the former term is undefined while the latter term is 0. Additionally, h will equal 0 if $d = 0$ or if $\theta = x\pi$ (where x is an integer); for more on the meaning of d and θ in context of handedness and peptide geometry, please refer to the 'Results and Discussions' and Fig. 4 in particular.

Alternative measures of handedness

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

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

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

Here, N is the peptide length, i is the peptide residue number and the position of each α-carbon is $\mathbf{N}_i$, 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. Eq. (11) has range $[-1, 1]$. Eq. (12), also used by *Gruziel, Dzwolak &*

Szymczak (2013), is the dihedral angle associated with the four contiguous α-carbons (one preceding and two succeeding the residue i), and ranges between $[-\pi, \pi]$ radians. For both metrics, values deviating more from 0 are more chiral (or 'twisted' or 'handed'), and left-handed twists are negative while right-handed twists are positive. Only α-carbon atom positions are 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:

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

χ_3, of arbitrary range, is known as the chirality index G_{0S} in *Solymosi et al. (2002)* and *Neal et al. (2003)*. (i, j, k, l) are exhaustive permutations of $\{1, 2, \dots, N\}$. This metric qualitatively matches the values of Eqs. (11) and (12) and, while not shown, the relationship between (ϕ, ψ) and χ_3 is available in the online GitHub repository.

Backbone structure generation

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

Obtaining secondary structure statistics

Statistics about secondary structures—particularly α-helices, 3_{10}-helices and β-sheets— were identified using the DSSP algorithm (*Kabsch & Sander, 1983*), although the STRIDE algorithm (*Frishman & Argos, 1995*) provides qualitatively identical distributions. The DSSP algorithm was applied to a database of 13,760 three-dimensional protein conformations (one domain per conformation) with lower than 40% sequence identity, obtained from the Structural Classification of Proteins or SCOPe website (Release 2.06; *Fox, Brenner & Chandonia (2014)*). This database is currently available as: http://scop.berkeley.edu/downloads/pdbstyle/pdbstyle-sel-gs-bib-40-2.06.tgz.

Backbone chirality $\neq$ backbone handedness

Finally, it is important to recognize the distinction between backbone (twist) handedness and backbone (molecular) chirality. Naïvely, chirality is a simple concept: a molecular conformation is achiral if its mirror image can be superimposed onto itself, otherwise that conformation is chiral (*Gold et al., 1997*) (alternatively, and less commonly, achiral

molecules possess inversion centers). Despite this intuitive definition, chirality has remained a confusing concept ever since its introduction (*Bentley, 2010*; *Wallentin et al., 2009*), which is possibly due to the fact that 'context' is very important when discussing chirality (*Mislow, 2002*). For example, when looking at a peptide at the residue or 'local' level, every amino acid (excepting glycine) is chiral due to the presence of a chiral α-carbon (its mirror image can not be superimposed onto itself). Yet, at the macromolecular level, even an all-glycine (and therefore locally achiral) peptide will display *conformations* that are not superimposable onto each other, and so such conformations would be chiral. Alternatively, when considering handedness, if a backbone is completely flat (say, a ring, where $d = 0$), handedness (h) will be undefined, and so one can not speak of handedness of the twist. Yet, the backbone may still remain chiral; e.g., cisplatin and transplatin are planar molecules that are nonetheless chiral opposites (*Testa, 2013*). It is for this reason that this report chooses to be careful to not claim that Eq. (10) is a metric for peptide/backbone chirality, but of peptide backbone *twist* handedness. However, estimates for backbone chirality (e.g., Eqs. (11) and (12)) may be used as surrogates for twist chirality to validate h (Eq. (10)), as both are related but not the same.

RESULTS AND DISCUSSION

Relevance of θ and d

When discussing peptide backbones, two possible definitions of backbone 'flatness' (or linearity) are possible: flatness at a residue level and flatness at the atomic level. In the former, all atoms of the same *type* are coplanar (examples of atom types are the backbone nitrogens, carbonyl carbons, α-carbons, or even sidechain β-carbons). In the latter definition of flatness, *all* atoms within the backbone are coplanar. For the discussions below, since the residue-by-residue behavior of the peptide is of primary relevance, the former definition is chosen as the relevant scope for flatness.

As described in Fig. 3, the helical parameters d and θ respectively refer to an axial displacement along the helical axis and an angular displacement in a plane perpendicular to the helical axis. For example, $d = 0$ indicates a helix flattened along its helical axis (Fig. 4, Scenario 1). This means that all regular peptides with $d = 0$ will be ring-like at some peptide length (shown in a following figure for a range of peptides). As expected from Eq. (10), at $d = 0$, one can not tell how the helix was built, since coplanar peptides can not be described as either left- or right-twisting. Therefore, even though $d = 0$ indicates highly twisted peptides, these twists do not possess handedness. This shows up in the h metric because, at $d = 0$, $|d|^{-1}$ is undefined.

Additionally Fig. 4 describes two important values for θ: $e\pi$ (Scenario 2) and $o\pi$ (Scenario 3), where e and o are even and odd integers. In particular, for any d, $\theta = e\pi$ indicates zero angular displacement along the axis, which puts all atoms of the same type on the same line parallel to the helical axis (Fig. 4, Scenario 2). Similarly, $\theta = o\pi$ indicates that every alternate atom (of the same type) along the backbone will be linear, and every adjacent atom will be diametrically opposite to each other (Fig. 4, Scenario 3); i.e., $\theta = o\pi$ indicates that all atoms of the same type will lie on a plane that is parallel to the helical axis. In short,
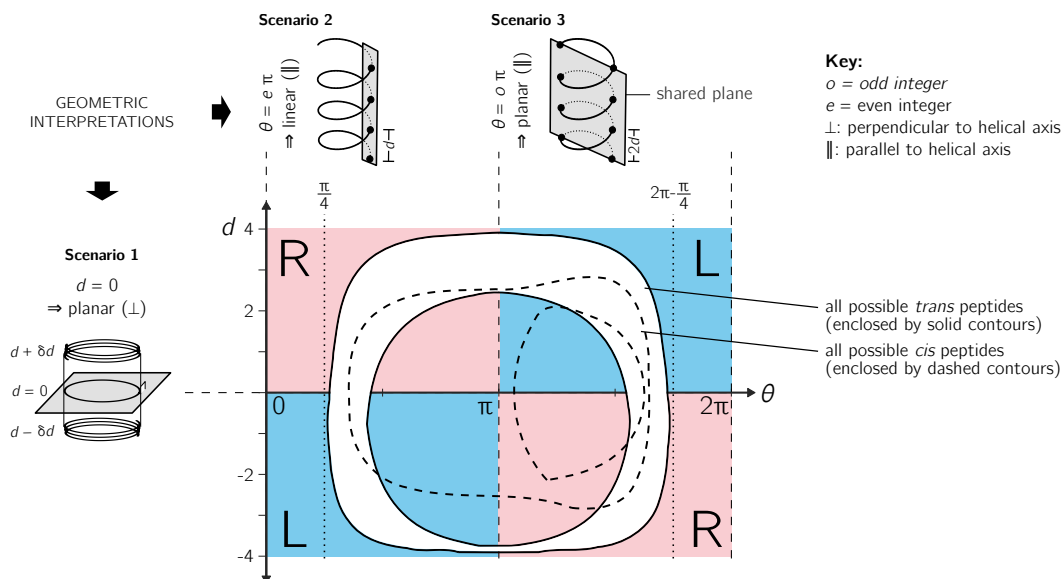
Figure 4 Further discussion on the meaning of *d* and *θ*. As shown in Fig. 3, axial separation *d* and angular separation *θ* between adjacent atoms of the same type combine to define handedness. The blue (dark) and red (light) shaded quadrants within the graph show the distribution of handedness as a function of *d* and *θ*. The relevant boundaries—*θ* = *x*π (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 as three scenarios. The toroid enclosed by two solid lines (and shaded white) represents all possible conformations for *trans* peptides ($\omega = 180 \pm 5°$). Similarly, the region allowed for *cis* peptides ($\omega = 0 \pm 5°$) are bound by the two dashed contours.

$\theta = 0$ codes for backbones that are linear (optimally extended for a fixed *d*) and $\theta = \pi$ describe peptides that zig-zag along a plane perpendicular to the helical axis. Finally, as is evident in Fig. 4, $\theta = e\pi$ conformations are not available to peptide backbones. Therefore, $\theta = o\pi$ (e.g., π or 180°) will be the most extended type of backbone (for a fixed *d*). These relationships show how, *a priori*, the curve of a backbone with particular (d, θ) may be interpreted.

Finally, θ may serve as an important single-number metric for describing backbone configurations. *Mannige, Kundu & Whitelam (2016)* developed one such number—a Ramachandran number ($\mathcal{R}$)—that is a structurally meaningful combination of ϕ and ψ. This number depends on the fact that 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, Thomas & Brasseur, 2003*; *Zacharias & Knapp, 2013*; *Mannige, Kundu & Whitelam, 2016*). Interestingly, θ follows that trend too, which—in combination with the fact that regions of the Ramachandran plot are sparse (*Mannige, Kundu & Whitelam, 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
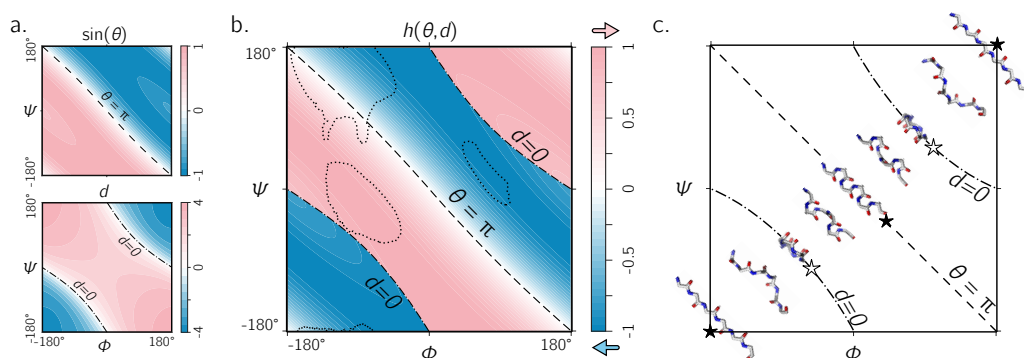
Figure 5 The handedness of an ordered *trans* peptide within the Ramachandran plot. (A) displays the relationship between backbone parameters (ϕ, ψ) and the associated helix parameters of curvature $\sin(\theta)$ (top; Eq. (1)) and axial displacement d (bottom; Eq. (2)). As shown in Fig. 3, the handedness of a helix is a function of these two variables (h; Eq. (10)). (B) is a map of backbone handedness (h) as a function of ϕ and ψ. 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). (B) 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.

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 (Eq. (1)).

Handedness of *trans* backbones

Figure 5A describes the behavior of $\sin(\theta)$ and d as a function of ϕ and ψ (assuming an all-*trans* backbone; $\omega = \pi$ or 180°). Figure 5B describes the behavior of backbone handedness (h; Eq. (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, $d = 0$ ('☆') 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 where $d = 0$ and so h is undefined there. This, along with the recognition that $d = 0$ indicates highly curved backbones, means that such transitions would be concomitant with a sharp change in handedness. When $\theta = \pi$, then the backbone is also flat (see '★' 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.

Figure 6A shows that the equation for h match other metrics for handedness, as interpreted by other metrics for chirality (*Kwiecińska & Cieplak, 2005*; *Kabsch & Sander, 1983*; *Gruziel, Dzwolak & Szymczak, 2013*). In particular, Fig. 6A displays the Ramachandran plot colored by h ((*i*); Eq. (10)) next to estimates calculated using χ_1
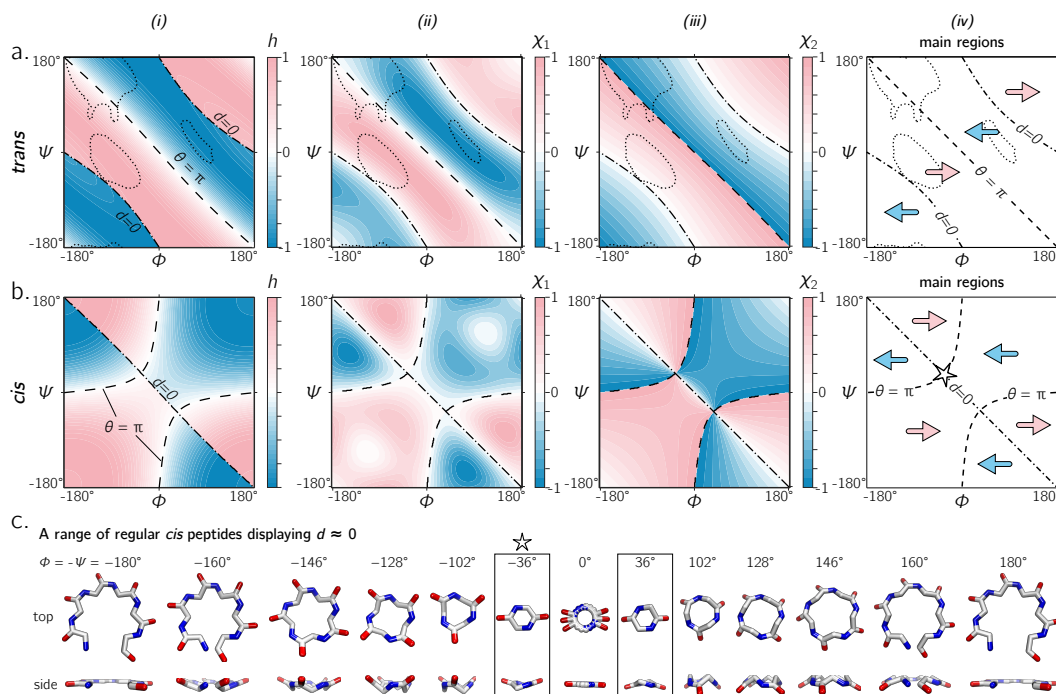
Figure 6 (A) and (B) describe the handedness of backbone twists whose amide dihedral angles are *trans* ($\omega = \pi$) and *cis* ($\omega = 0$), respectively. Column *(i)* describes handedness (h; Eq. (10)), which does not require structures to be computationally generated. Columns *(ii)* and *(iii)* respectively show vector-based estimates of backbone chirality— χ_1 (Eq. (11)) and χ_2 (Eq. (12))—which are calculated from computationally generated peptides (see 'Methods'). Regions of left- and right-handedness are identical for all measures *(i–iii)*. A cartoon representation of distinct regions of handedness is shown in *(iv)*. Finally, (C) displays a range of regular *cis* peptide backbones with $d \approx 0$. As explained in Fig. 4, $d = 0$ indicates a flat backbone that lies perpendicular to the helical axis, which results in ring-like peptides. Interestingly, a point in the Ramachandran plot exists exclusively for *cis* peptides, where $d = 0$ and $\theta = \pi$: $\phi = -\psi = \pm 36°$ ('☆' in (B)-*(iv)* and (C)).

((ii); Eq. (11)) and χ_2 *((iii)*; Eq. (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 for 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* regular 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 ($\phi = \pm 36°$, $\psi = \mp 36°$) 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 glance, this appears to be contradiction, because $d = 0$ indicates the most
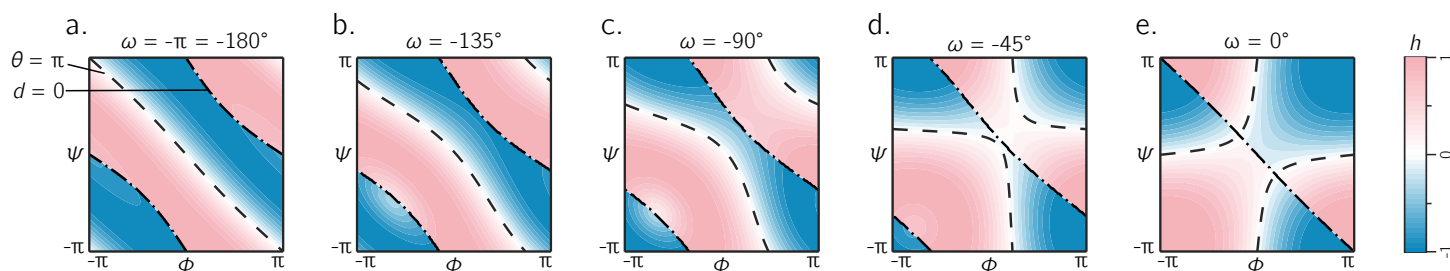
Figure 7 **The landscape of backbone chirality as a function of amide dihedral angle ω.** As ω is changed, the features of the landscape smoothly transform from the landscapes of $\omega = \pm\pi$ to $\omega = 0$. For all values of ω, it is evident that the naïve view of chirality (Fig. 1D) is wrong: at least four distinct regions of chirality (separated by boundaries $d = 0$ and $\theta = \pi$) 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.

curved backbone at a fixed θ, and $\theta = \pi$ indicates the most linear backbone at a fixed d; however, it is purely due to the nature of the *cis* backbone that this indeed is possible. 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 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.

The exhaustive survey of regular *cis* ($\omega = 0$) and *trans* ($\omega = \pi$) peptides (Fig. 6) proves that the naïve picture of chirality—that the $-$ve diagonal separates the right-twisting backbones from the left-twisting backbones (Fig. 1D)—is wrong. However, deviations from $\omega = 0$ or π are evident in the Protein Databank; see, e.g., discussions by *Improta, Vitagliano & Esposito (2011)*. This raises the question: how does varying ω through non-traditional values change the handedness landscape? Figure 7 describes Ramachandran plots that show handedness in terms of varying ω, which shows that this complicated separation of handedness in *cis* and *trans* backbones also holds for other values of ω. Therefore, the naïve expectation of handedness (Fig. 1D) is too simple, irrespective of amide dihedral angle.

$[-\pi, \pi)$ or $[0, 2\pi)$: 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, Tymoczko & Stryer (2010)* and *Alberts et al. (2002)*). However, *Ramachandran, Ramakrishnan & Sasisekharan (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 & Knapp, 2013*), while some have turned to $[0, 2\pi)$ as the norm (*Némethy, Leach & Scheraga, 1966; Voelz, Dill & Chorny, 2011*).

Given the periodicity of the Ramachandran plot, the two frames of reference are scientifically identical; however the value of the Ramachandran plot lies in its utility as a *map*: 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
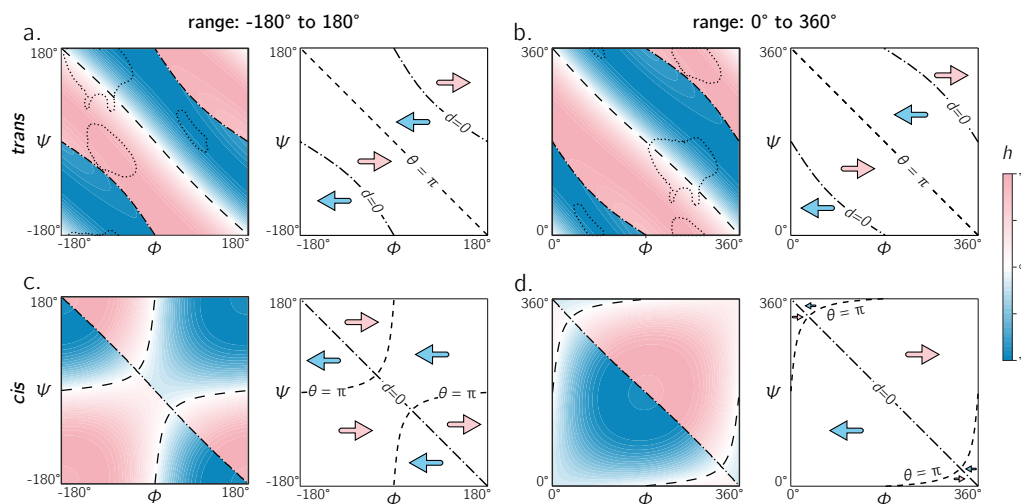
Figure 8 **The two frames of reference (or ranges) for the Ramachandran plot for *trans* and *cis* backbones.** Both ranges $[-\pi, \pi)$ and $[0, 2\pi)$ yield similar trends for *trans* backbones (A, B); however, for *cis* backbones, the latter frame of reference (D) appears to more neatly apportion 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 ($p = 0.9$).

pictograms. For that reason, switching the map from one range to another means that the two types of scientists—each used to a distinct range—will not be able to converse as seamlessly.

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

A universal alternative to the Ramachandran plot

While the Ramachandran plot is useful enough to earn a place in undergraduate-level biology textbooks (*Berg, Tymoczko & Stryer, 2010*; *Alberts et al., 2002*), as discussed throughout this report, it is not easy to estimate features of a peptide backbone just from its (ϕ, ψ) angles (Fig. 9A). This prompted *Zacharias & Knapp (2013)* to introduce a new representation for backbone degrees of freedom in the form of a polar graph. In this polar representation, the θ is the angular coordinate (azimuth) and d is the radial coordinate.
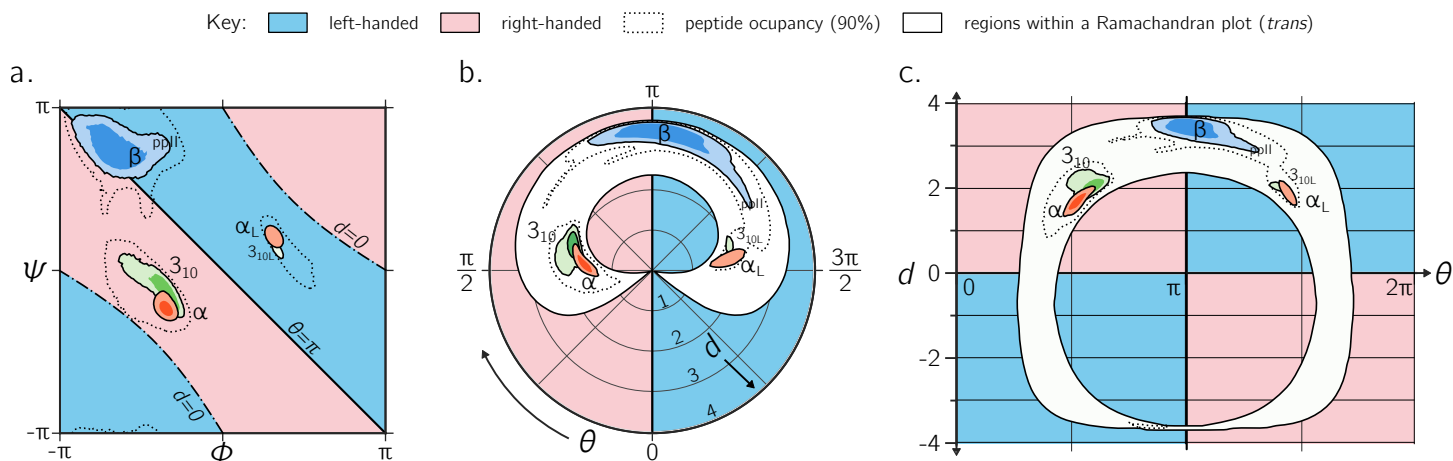
Figure 9 Alternative representations of the Ramachandran plot. While the Ramachandran plot is useful to map characteristics of secondary structures (A), it is not intuitive. For example, the relationship between the Ramachandran parameters (ϕ, ψ) and the handedness of a backbone is not obvious (see, e.g., the non-obvious distribution of left- and right-handed peptides as a function of ϕ and ψ). For this reason, *Zacharias & 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). (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; see 'Methods'), each secondary structure has two contours signifying $p = 0.5$ and 0.8.

An example of one such representation is shown in Fig. 9B, with the direction of increasing θ reversed (compared to the cited report) to maintain relative positions of secondary structures within the Ramachandran plot (Fig. 9A). *Zacharias & Knapp (2013)* stated an additional reason for the introduction of the polar representation (Fig. 9B): θ, which is an angle and therefore periodic, can remain periodic as the angular coordinate in the graph.

However, the format proposed by *Zacharias & Knapp (2013)* (Fig. 9B) is incomplete for a few reasons: (1) $d < 0$ peptides (the bottom-left and top-right regions of Fig. 5A, bottom) will never be observed in this map since only structures with $d \geq 0$ are allowed; (2) all peptides with $d = 0$ (marked by '—·—' in every preceding Ramachandran plot) will be compressed into one point at the center, even though Fig. 6C shows a range of legitimate $d = 0$ conformations; (3) while the graph is θ-periodic, the values for θ in peptides are constrained within one $[0, 2\pi]$ period (peptides range between $\theta = \pi/4$ and $2\pi - \pi/4$; vertical dotted lines in Fig. 4); i.e., periodicity in θ is not required for the faithful representation of peptides. Fortunately, even though this system is not universal (again, since $d < 0$ structures are not accommodated), most conformations in globular proteins display positive d, and so the representation presented by *Zacharias & Knapp (2013)* is a reasonable one for most proteins with known structure.

Interestingly, Fig. 9A—which arranges the parameters θ and d along Cartesian axes— serves as both a universal *and* intuitive map for peptide backbone geometry. This is because: (1) as shown in Fig. 4, such maps reveal a wealth of information about the peptide backbone, (2) both positive and negative values of d are allowed (compared to Fig. 9B), due to the shift in the coordinate system from polar to Cartesian, and (3) this format

accommodates *every* type of peptide conformation: any peptide (or its mimic) has a place in this map irrespective of whether the amide backbone is *cis* or *trans* or any other value; additionally, if the backbone is distorted, such distortions can also be accounted for since d and θ account for such distortions (Eqs. (1) and (2)). This is impossible to do using a single Ramachandran plot without making sweeping assumptions about backbone parameters that are not ϕ and ψ. The (θ, d) plot opens up the possibility for a new, intuitive, and *universal* kind of graphical representation as a supplement to the Ramachandran plot.

A departure from perfect regularity

So far, this report has focused on regular or *simple* backbone conformations, i.e., those that are formed from the same ϕ and ψ angles repeated along the backbone. This is particularly because a simple and visually intuitive correspondence exists (Figs. 3 and 4) between a regular backbone (described by myriad internal coordinates) and a helix that is described simply by (d, θ). However, there is a possibility that d and θ are useful even in isolation, when the unreasonable constraint of perfect backbone regularity is lifted. An example of such a departure from regularity follows.

Some secondary structures are characterized by the regular combination of two or more sets of $[\phi, \psi]$ (*Pauling & Corey, 1951b*; *Pauling & Corey, 1951a*; *Armen et al., 2004*; *Daggett, 2006*; *Hayward & Milner-White, 2008*; *Mannige et al., 2015*; *Mannige, Kundu & Whitelam, 2016*). For example, the Σ-strand is constructed by alternating between two backbone states $(\phi, \psi, \omega) = (-A, B, 180°)$ and $(-B, A, 180°)$, where $A \approx 120$ and $B \approx 90$ (Fig. 4 in *Mannige et al. (2015)*). It was found that the two states are similar in the extent to which the backbone twists, but opposite in handedness, which allows for these secondary structures to remain linear, albeit in a meandering way (*Mannige et al., 2015*). Equation (10) also describes these two states as opposite in handedness and similar in twist extent: the h for the two states are -0.34 and 0.51, respectively (the difference in magnitude is within the range of the standard deviation in h [0.391] for the β-sheet). Similarly, the α-sheet proposed by *Pauling & Corey (1951a)* is constructed by alternating between $\alpha_{(D)}$ and α_L backbone states, yet this motif is linear because each state describes equal but opposite handedness $h = \pm 0.41$. These points raise the possibility that, even in the absence of perfect backbone regularity, the values d, θ, and h may be considered to be residue-specific properties that may be combined to readily provide insights about higher order structures.

CONCLUSIONS

This report introduces a metric for backbone handedness (h) that is based on modeling the backbone as a helix (Fig. 2; *Miyazawa (1961)*). In particular, h, which is a combination of the helical parameters θ (angular displacement) and d (axial displacement), ranges from -1 to 1, and is negative (or positive) when the backbone twist is left(or right)-handed (with larger $|h|$ indicating greater extent of twistedness). This metric (h) was used to characterize every regular backbone's twist within the Ramachandran plot, for both *cis* and *trans* peptides. In doing so, this report dispels a naïve view of handedness (Fig. 1D), which states that backbone handedness in the Ramachandran plot is separated by the negative-sloped ($-$ve) diagonal. Interestingly, the reason for the naïve view makes

senses when considering only *trans* peptides: the $-$ve diagonal ('$--$' in Fig. 5A) separates right-handed and left-handed twists *if* one considers only the regions dominantly occupied by structured proteins ('dotted line' in Fig. 5A). Plotting the backbone handedness (h) in the two common frames of reference— $\phi, \psi \in [-\pi, \pi)$ and $[0, 2\pi)$—indicates that the less commonly used frame $[0, 2\pi)$ may be more appropriate for interpreting *cis* backbones (Fig. 8).

The behavior of a backbone in *cis* and *trans* Ramachandran plots look dramatically different (Fig. 6), and so scientists dealing with new structures that have a combination of *cis* and *trans* backbones can not use one Ramachandran plot to faithfully describe these structures. Interestingly, the parameters θ and d combine all features (internal coordinates) of a contorting backbone, including the amide dihedral angle ω, which means that (θ, d) can describe *any* peptide backbone, irrespective of ω. Therefore, the Cartesian plot with θ and d as the x- and y-axis, respectively, serves as a unique plot for *any* peptide backbone (Fig. 9), with specific values and boundaries containing deep structural meaning (Fig. 4). These discussions, the author hopes, clarify a number of concepts associated with the Ramachandran plot, while providing new insights into how to interrogate the features of new protein and protein-like structures.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Grant Disclosures

Competing Interests

The author declares that there are no competing interests.

Author Contributions

- Ranjan V. Mannige conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

Data Availability

The following information was supplied regarding data availability:

The author has uploaded a publicly accessible GitHub repository containing Python scripts that recreate all relevant data depicted within the manuscript. The repository is available here:

https://github.com/ranjanmannige/backbone_chirality.

This repository may also be downloaded as a compressed zip file at the following URL: https://github.com/ranjanmannige/backbone_chirality/archive/master.zip.

REFERENCES

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. *Molecular biology of the cell.* 5th Edition. New York: Garland Science.

Armen RS, DeMarco ML, Alonso DO, Daggett V. 2004. Pauling and Corey's α-pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proceedings of the National Academy of Sciences of the United States of America* **101(32)**:11622–11627 DOI 10.1073/pnas.0401781101.

Beck DA, Alonso DO, Inoyama D, Daggett V. 2008. The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences of the United States of America* **105(34)**:12259–12264 DOI 10.1073/pnas.0706527105.

Bentley R. 2010. Chiral: a confusing etymology. *Chirality* **22(1)**:1–2 DOI 10.1002/chir.20699.

Berg JM, Tymoczko JL, Stryer L. 2010. *Biochemistry, international edition.* 7th edition. New York: WH Freeman & Co.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov IN, Bourne PE. 2000. The protein data bank. *Nucleic Acids Research* **28(1)**:235–242 DOI 10.1093/nar/28.1.235.

Bragg L, Kendrew JC, Perutz MF. 1950. Polypeptide chain configurations in crystalline proteins. *Proceedings of the Royal Society of London Series A Mathematical and Physical Sciences* **203(1074)**:321–357 DOI 10.1098/rspa.1950.0142.

Chothia C, Hubbard T, Brenner S, Barns H, Murzin A. 1997. Protein folds in the all-beta and all-alpha classes. *Annual Review of Biophysics and Biomolecular Structure* **26**:597–627 DOI 10.1146/annurev.biophys.26.1.597.

Cock P, Antao T, Chang J, Chapman B, Cox C, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, De Hoon M. 2009. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25(11)**:1422–1423 DOI 10.1093/bioinformatics/btp163.

Cooper GM, Hausman RE. 2013. *The cell: a molecular approach*. 6th edition. Sunderland: Sinauer Associates, Inc.

Cross L, Klyne W. 2013. *Rules for the nomenclature of Organic Chemistry: Section E: stereochemistry (Recommendations 1974)*. Oxford: Pergamon.

Daggett V. 2006. α-sheet: the toxic conformer in amyloid diseases? *Accounts of Chemical Research* **39(9)**:594–602 DOI 10.1021/ar0500719.

Eisenberg D. 2003. The discovery of the α-helix and β-sheet, the principal structural features of proteins. *Proceedings of the National Academy of Sciences of the United States of America* **100(20)**:11207–11210 DOI 10.1073/pnas.2034522100.

Engh RA, Huber R. 1991. Accurate bond and angle parameters for x-ray protein structure refinement. *Acta Crystallographica Section A: Foundations of Crystallography* **47(4)**:392–400 DOI 10.1107/S0108767391001071.

Engh R, Huber R. 2006. Structure quality and target parameters. In: *International tables for crystallography volume F: crystallography of biological macromolecules*. Dordrecht: Springer, 382–392.

Esposito L, Balasco N, De Simone A, Berisio R, Vitagliano L. 2013. Interplay between peptide bond geometrical parameters in nonglobular structural contexts. *BioMed Research International* **2013**:Article 326914.

Ferrarini A, Nordio PL. 1998. On the assessment of molecular chirality. *Journal of the Chemical Society, Perkin Transactions* **2(2)**:455–460.

Fox NK, Brenner SE, Chandonia JM. 2014. Scope: structural classification of proteins–extended, integrating scop and astral data and classification of new structures. *Nucleic Acids Research* **42(Database Issue)**:D304–D309.

Frishman D, Argos P. 1995. Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function, and Bioinformatics* **23(4)**:566–579 DOI 10.1002/prot.340230412.

Gold V, Loening K, McNaught A, Shemi P. 1997. *IUPAC compendium of chemical terminology*. Oxford: Blackwell Science.

Gorske BC, Mumford EM, Conry RR. 2016. Tandem incorporation of enantiomeric residues engenders discrete peptoid structures. *Organic Letters* **18(11)**:2780–2783.

Gruziel M, Dzwolak W, Szymczak P. 2013. Chirality inversions in self-assembly of fibrillar superstructures: a computational study. *Soft Matter* **9(33)**:8005–8013 DOI 10.1039/c3sm27961h.

Harris AB, Kamien RD, Lubensky TC. 1999. Molecular chirality and chiral parameters. *Reviews of Modern Physics* **71(5)**:1745–1757 DOI 10.1103/RevModPhys.71.1745.

Hayward S, Milner-White EJ. 2008. The geometry of α-sheet: implications for its possible function as amyloid precursor in proteins. *Proteins: Structure, Function, and Bioinformatics* **71(1)**:415–425 DOI 10.1002/prot.21717.

Ho BK, Thomas A, Brasseur R. 2003. Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and h-bonding in the α-helix. *Protein Science* **12(11)**:2508–2522.

Hooft RW, Sander C, Vriend G. 1997. Objectively judging the quality of a protein structure from a Ramachandran plot. *Computer Applications in the Biosciences* **13(4)**:425–430 DOI 10.1093/bioinformatics/13.4.425.

Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK. 2002. Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of Molecular Biology* **323(3)**:573–584 DOI 10.1016/S0022-2836(02)00969-5.

Improta R, Vitagliano L, Esposito L. 2011. Peptide bond distortions from planarity: new insights from quantum mechanical calculations and peptide/protein crystal structures. *PLOS ONE* **6(9)**:e24533 DOI 10.1371/journal.pone.0024533.

Improta R, Vitagliano L, Esposito L. 2015a. Bond distances in polypeptide backbones depend on the local conformation. *Acta Crystallographica Section D: Biological Crystallography* **71(6)**:1272–1283 DOI 10.1107/S1399004715005507.

Improta R, Vitagliano L, Esposito L. 2015b. The determinants of bond angle variability in protein/peptide backbones: a comprehensive statistical/quantum mechanics analysis. *Proteins: Structure, Function, and Bioinformatics* **83(11)**:1973–1986 DOI 10.1002/prot.24876.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22(12)**:2577–2637 DOI 10.1002/bip.360221211.

Kwiecińska JI, Cieplak M. 2005. Chirality and protein folding. *Journal of Physics: Condensed Matter* **17(18)**:S1565.

Laskowski RA. 2003. Structural quality assurance. *Methods of Biochemical Analysis* **44**:273–303.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. Procheck: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26(2)**:283–291 DOI 10.1107/S0021889892009944.

Linderstrøm-Lang KU. 1952. *Lane medical lectures: proteins and enzymes*. Vol. 6. Stanford: Stanford University Press.

Mannige RV. 2014. Dynamic new world: refining our view of protein structure, function and evolution. *Proteomes* **2(1)**:128–153 DOI 10.3390/proteomes2010128.

Mannige RV, Haxton TK, Proulx C, Robertson EJ, Battigelli A, Butterfoss GL, Zuckermann RN, Whitelam S. 2015. Peptoid nanosheets exhibit a new secondary structure motif. *Nature* **526**:415–420 DOI 10.1038/nature15363.

Mannige RV, Kundu J, Whitelam S. 2016. The Ramachandran number: an order parameter for protein geometry. *PLOS ONE* **11(8)**:e0160023 DOI 10.1371/journal.pone.0160023.

Mirijanian DT, Mannige RV, Zuckermann RN, Whitelam S. 2014. Development and use of an atomistic charmm-based forcefield for peptoid simulation. *Journal of Computational Chemistry* **35(5)**:360–370 DOI 10.1002/jcc.23478.

Mislow K. 2002. Stereochemical terminology and its discontents. *Chirality* **14(2–3)**:126–134 DOI 10.1002/chir.10069.

Miyazawa T. 1961. 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* **55(161)**:215–231 DOI 10.1002/pol.1961.1205516121.

Neal MP, Solymosi M, Wilson MR, Earl DJ. 2003. Helical twisting power and scaled chiral indices. *Journal of Chemical Physics* **119(6)**:3567–3573 DOI 10.1063/1.1590307.

Némethy G, Leach S, Scheraga HA. 1966. The influence of amino acid side chains on the free energy of helix-coil transitions1. *Journal of Physical Chemistry* **70(4)**:998–1004 DOI 10.1021/j100876a008.

Orosz F, Ovádi J. 2011. Proteins without 3d structure: definition, detection and beyond. *Bioinformatics* **27(11)**:1449–1454 DOI 10.1093/bioinformatics/btr175.

Osipov M, Pickup B, Dunmur D. 1995. A new twist to molecular chirality: intrinsic chirality indices. *Molecular Physics* **84(6)**:1193–1206 DOI 10.1080/00268979500100831.

Pauling L, Corey RB. 1951a. Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets. *Proceedings of the National Academy of Sciences of the United States of America* **37(11)**:729–740 DOI 10.1073/pnas.37.11.729.

Pauling L, Corey RB. 1951b. The pleated sheet, a new layer configuration of polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America* **37(5)**:251–256 DOI 10.1073/pnas.37.5.251.

Pauling L, Corey RB, Branson HR. 1951. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America* **37(4)**:205–211 DOI 10.1073/pnas.37.4.205.

Ramachandran G, Ramakrishnan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **7(1)**:95–99 DOI 10.1016/S0022-2836(63)80023-6.

Robertson EJ, Battigelli A, Proulx C, Mannige RV, Haxton TK, Yun L, Whitelam S, Zuckermann RN. 2016. Design, synthesis, assembly, and engineering of peptoid nanosheets. *Accounts of Chemical Research* **49(3)**:379–389 DOI 10.1021/acs.accounts.5b00439.

Shimanouchi T, Mizushima SI. 1955. On the helical configuration of a polymer chain. *Journal of Chemical Physics* **23(4)**:707–711 DOI 10.1063/1.1742083.

Solymosi M, Low RJ, Grayson M, Neal MP. 2002. A generalized scaling of a chiral index for molecules. *Journal of Chemical Physics* **116(22)**:9875–9881 DOI 10.1063/1.1476321.

Sun J, Zuckermann RN. 2013. Peptoid polymers: a highly designable bioinspired material. *ACS Nano* **7(6)**:4715–4732 DOI 10.1021/nn4015714.

Testa B. 2013. Organic stereochemistry. Part 2. *Helvetica Chimica Acta* **96(2)**:159–188 DOI 10.1002/hlca.201200470.

Tien MZ, Sydykova DK, Meyer AG, Wilke CO. 2013. Peptidebuilder: a simple python library to generate model peptides. *PeerJ* **1**:e80 DOI 10.7717/peerj.80.

Van Der Walt S, Colbert SC, Varoquaux G. 2011. The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering* **13(2)**:22–30.

Voelz VA, Dill KA, Chorny I. 2011. Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in amber. *Peptide Science* **96**(5):639–650 DOI 10.1002/bip.21575.

Wallentin CJ, Orentas E, Wärnmark K, Wendt OF. 2009. Chirality, a never-ending source of confusion. *Zeitschrift für Kristallographie International Journal for Structural, Physical, and Chemical Aspects of Crystalline Materials* **224**(12):607–608.

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology* **337**(3):635–645 DOI 10.1016/j.jmb.2004.02.002.

Zacharias J, Knapp EW. 2013. Geometry motivated alternative view on local protein backbone structures. *Protein Science* **22**(11):1669–1674 DOI 10.1002/pro.2364.