

# 1 An exhaustive survey of regular peptide 2 conformations using a new metric for 3 backbone handedness ( $h$ )

4 Ranjan V. Mannige<sup>1,2,\*</sup>

5 <sup>1</sup> Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, U.S.A.

6 <sup>2</sup> Present address: Multiscale Institute, Redwood City, CA, U.S.A.

7 \* ranjanmannige@gmail.com

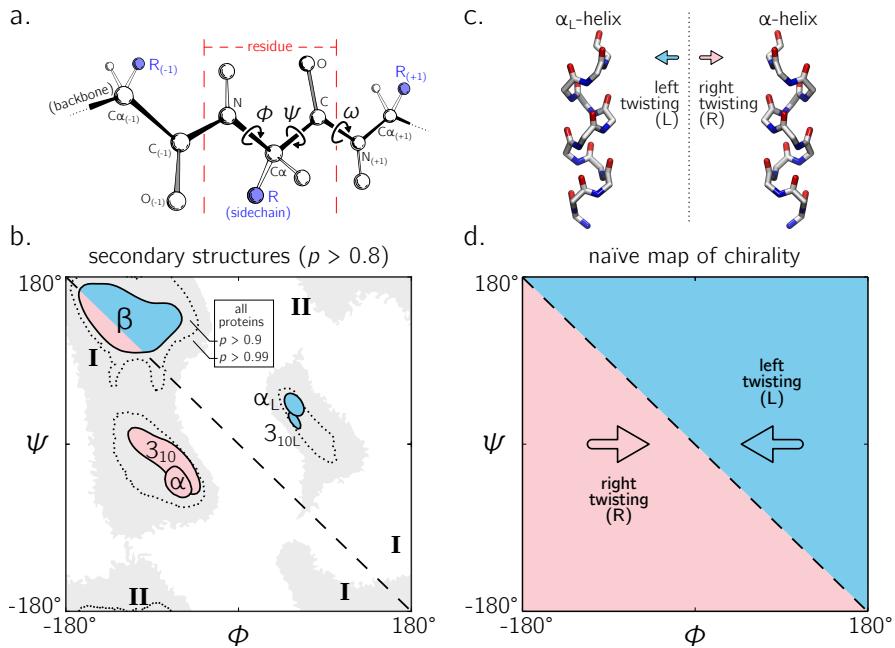
## 8 ABSTRACT

9 The Ramachandran plot is important to structural biology as it describes a peptide backbone in the  
10 context of its dominant degrees of freedom – the backbone dihedral angles  $\phi$  and  $\psi$  (Ramachandran  
11 *et al.*, 1963). Since its introduction, the Ramachandran plot has been a crucial tool to characterize protein  
12 backbone features. However, the conformation or twist of a backbone as a function of  $\phi$  and  $\psi$  has not  
13 been completely described for both *cis* and *trans* backbones. Additionally, little intuitive understanding  
14 is available about a peptide's conformation simply from knowing the  $\phi$  and  $\psi$  values of a peptide (e.g.,  
15 is the regular peptide defined by  $\phi = \psi = -100^\circ$  left-handed or right-handed?). This report provides  
16 a new metric for backbone handedness ( $h$ ) based on interpreting a peptide backbone as a helix with  
17 axial displacement  $d$  and angular displacement  $\theta$ , both of which are derived from a peptide backbone's  
18 internal coordinates, especially dihedral angles  $\phi$ ,  $\psi$  and  $\omega$ . In particular,  $h$  equals  $\sin(\theta)d/|d|$ , with range  
19  $[-1, 1]$  and negative (or positive) values indicating left(or right)-handedness. The metric  $h$  is used to  
20 characterize the handedness of every region of the Ramachandran plot for both *cis* ( $\omega = 0^\circ$ ) and *trans*  
21 ( $\omega = 180^\circ$ ) backbones, which provides the first exhaustive survey of twist handedness in Ramachandran  
22 ( $\phi, \psi$ ) space. These maps fill in the 'dead space' within the Ramachandran plot, which are regions  
23 that are not commonly accessed by structured proteins, but which may be accessible to intrinsically  
24 disordered proteins, short peptide fragments, and protein mimics such as peptoids. Finally, building on  
25 the work of Zacharias and Knapp (2013), this report presents a new plot based on  $d$  and  $\theta$  that serves  
26 as a universal and intuitive alternative to the Ramachandran plot. The universality arises from the fact  
27 that the co-inhabitants of such a plot include every possible peptide backbone including *cis* and *trans*  
28 backbones. The intuitiveness arises from the fact that  $d$  and  $\theta$  provide, at a glance, numerous aspects of  
29 the backbone including compactness, handedness, and planarity.

## 30 INTRODUCTION

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

38 Today, two-dimensional ( $\phi, \psi$ ) plots are called Ramachandran plots (or 'maps'), and are introduced in  
39 undergraduate biology textbooks as a guide for understanding a peptide backbone's general conformational  
40 state or 'twistedness' at a glance (Bragg *et al.*, 1950; Pauling and Corey, 1951b; Pauling *et al.*, 1951;  
41 Linderstrøm-Lang, 1952; Laskowski *et al.*, 1993; Chothia *et al.*, 1997; Hooft *et al.*, 1997; Cooper and  
42 Hausman, 2013; Alberts *et al.*, 2002; Laskowski, 2003; Ho *et al.*, 2003; Eisenberg, 2003; Berg *et al.*,  
43 2010; Mannige *et al.*, 2016). The Ramachandran plot is especially useful because (stable) proteins are  
44 hierarchical in structure (Linderstrøm-Lang, 1952): the final (tertiary) conformation of a *structured* protein  
45 is composed of discrete secondary structures – *regular* structures – that interact with each other and  
46 which are strung together by loops that are less regular (Alberts *et al.*, 2002; Berg *et al.*, 2010). Each



**Figure 1.** The backbone of a single residue (a) can be described by its dihedral angles  $\phi$  and  $\psi$  (and in smaller part,  $\omega$ , which is predominantly *trans* or  $\sim 180^\circ$ ). 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 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 ( $\omega = \pi$ ). Panel (a) is modified from Mannige *et al.* (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.

47 regular peptide structure describes a backbone whose per-residue ( $\phi, \psi$ ) values are generally the same,  
48 and therefore 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 that  
50 display stable conformations (Berman *et al.*, 2000; Alberts *et al.*, 2002). These types of proteins occupy  
51 only a limited region of the plot (dotted regions in Fig. 1b). The regular backbone conformations in these  
52 regions are well understood. For example, known regular structures that are to the right of the negatively  
53 sloping diagonal (dashed line in Fig. 1b; henceforth denoted as the ‘-ve diagonal’) are left-handed in  
54 backbone twist, while those that are to the left of the diagonal are right-handed (left- and right-handed  
55 regions are respectively shaded brown and gold<sup>1</sup>). For example, the position of the idealized left- and  
56 right-handed  $\alpha$ -helices (Fig. 1c) – respectively denoted as  $\alpha_L$  and  $\alpha$  in Fig. 1b – are on opposite sides  
57 of the -ve diagonal. The ‘naïve view’ of handedness, obtained from looking only at structured proteins,  
58 would be the expectation that the -ve diagonal neatly separates the Ramachandran plot into regions of left-  
59 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 – and up to 50% of the mammalian proteins display large ( $> 30$  residue) stretches of  
63 disorder (Iakoucheva *et al.*, 2002; Ward *et al.*, 2004; Orosz and Ovádi, 2011; Mannige, 2014). Interestingly,  
64 when compared to structured backbones, structurally degenerate or disordered backbones occupy many

<sup>1</sup>Go Hillies!

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

Additionally, a number of peptide mimics – especially peptoids (Sun and 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  $\Sigma$ -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 ‘ $\omega$ -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 display backbone twists that are equal in magnitude but opposite in handedness [for this reason, the  $\Sigma$ -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 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 and Mizushima, 1955; Miyazawa, 1961; Zacharias and 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 *et al.*, 2016). Also, this survey dispels the naïve view of handedness (Fig. 1d) by showing that the distribution of handedness as a function of  $\phi$  and  $\psi$  is more complicated than the distribution allowed by the naïve view. Finally, the results also show that the Ramachandran plot whose  $\phi$  and  $\psi$  values range between  $0^\circ$  and  $360^\circ$  is more intuitive and visually meaningful (compared to those that range between  $-180^\circ$  and  $180^\circ$ ), particularly for *cis* backbones. 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

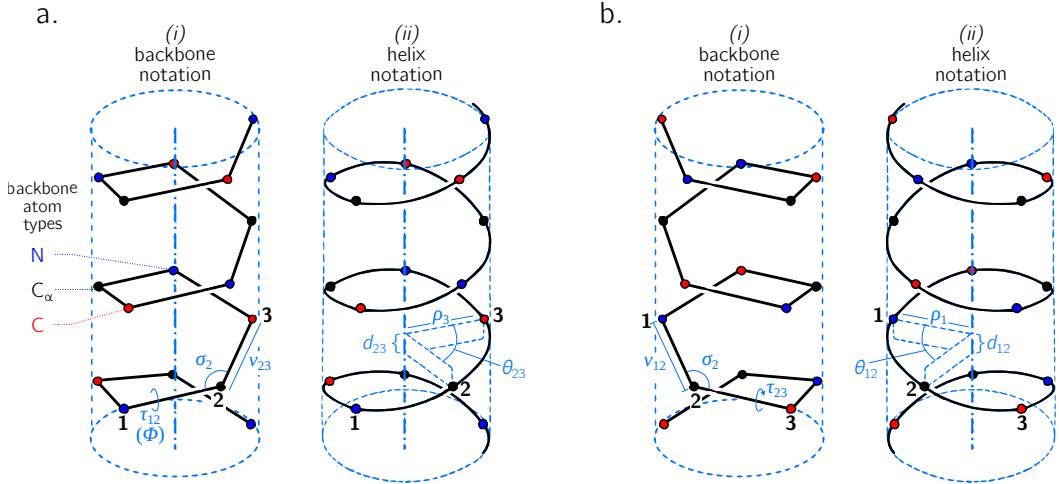
While angular units in this report switch between radians and degrees, their units in any particular situation may be inferred by the presence or absence of the degree symbol ( $^\circ$ ). All methods and materials required to produce this manuscript are freely available at [https://github.com/ranjanmannige/backbone\\_chirality](https://github.com/ranjanmannige/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 [Fig. 2; 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*).



**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 ( $\sigma_2$ ), and dihedral angles ( $\tau_{12}$ ), while helical coordinates are a function of displacement along the helical axis ( $d_{12}$ ), angular displacement in the plane perpendicular to the helical axis ( $\theta_{12}$ ) and shortest distance of an atom of type  $i$  to the helical axis ( $\rho_i$ ). Representations are derived from Figs. 1 and 2 in Shimanouchi and Mizushima (1955).

Internal coordinates are associated with stereochemical terms: bond lengths ( $v_{ij}$ ) between adjacent atoms  $i$  and  $j$ , bond angles ( $\sigma_i$ ) between the two bonds adjacent to atom  $i$ , and dihedral or torsion angles ( $\tau_{ij}$ ), 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 ( $\theta$ ), and the radius of the helix ( $\rho_i$ ) 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 ‘ $n$ ’, ‘ $\alpha$ ’, and ‘ $c$ ’ respectively refer to the backbone nitrogen,  $\alpha$ -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 axial displacement ( $d$ ) and angular displacement ( $\theta$ ) may be retrieved using the following two equations.

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

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

131 The ranges for  $d$  and  $\theta$ , respectively, are  $[-\lambda, +\lambda]$  and  $[0, 2\pi]$  (the positive limit  $\lambda$  is defined by allowed  
132 values for the various internal coordinates). As above, subscripts ‘n’, ‘ $\alpha$ ’, and ‘c’ respectively refer to the  
133 backbone nitrogen,  $\alpha$ -carbon and carbonyl carbon atom types. The dihedral angles  $\phi$ ,  $\psi$ , and  $\omega$  represent  
134 the traditional symbols for backbone dihedral angles, which may be otherwise denoted as  $\tau_{n,\alpha}$ ,  $\tau_{\alpha,c}$ , and  
135  $\tau_{c,n(+1)}$ , respectively.

Finally, for any type of atom (say  $\alpha$ -carbons), the radius or distance from the helical axis  $\rho_\alpha$  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 (denoted here as  $d_\alpha^2$  for  $\alpha$ -carbons), which allows for a more simplified form

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

136 The distance between adjacent  $\alpha$ -carbons ( $d_\alpha$ ) is  $\sim 3.8\text{\AA}$  for *trans* peptides and  $\sim 3\text{\AA}$  for *cis* peptides.  
137 Other radii ( $\rho_c, \rho_n$ ) can be obtained by cycling through ( $\alpha, c, n$ ) subscripts within Eqns. 3 and 4<sup>2</sup>. Note  
138 that all  $\rho_i$ 's are functions of  $\theta$  and  $d$  (along with other internal coordinates), and so one may use two of  
139 the three terms in  $(d, \theta, \rho_i)$  to describe the helical state of a peptide. Since there is only one  $d$  and  $\theta$  per  
140 backbone (compared to three  $\rho_i$ 's, one per atom type), this report utilizes  $d$  and  $\theta$  as the two descriptors  
141 [other 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; Improta *et al.*, 2015a; Esposito *et al.*, 2013; Improta *et al.*, 2015b). In particular, most backbone bond lengths and angles display one equilibrium value (Improta *et al.*, 2015a; Esposito *et al.*, 2013; Improta *et al.*, 2015b), while the backbone dihedral angles  $\phi$  and  $\psi$  occupy a range of possible values and minima, e.g., regions in the Ramachandran plot that describe  $\alpha$ -helices and  $\beta$ -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 follows<sup>3</sup>.

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

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

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

<sup>2</sup> $\rho_c$  and  $\rho_n$  are obtained by the following subscript conversions: ( $\alpha \rightarrow c, c \rightarrow n, n \rightarrow \alpha$ ) and ( $\alpha \rightarrow n, c \rightarrow \alpha, n \rightarrow c$ ).

<sup>3</sup>The values used by Zacharias and Knapp (2013), taken from Engh and Huber (1991, 2006), are:

$v_{n,\alpha} = 1.459\text{\AA}$ ,  $v_{\alpha,c} = 1.525\text{\AA}$ ,  $v_{c,n(+1)} = 1.336\text{\AA}$ ,  $\sigma_\alpha = 111.0^\circ$ ,  $\sigma_c = 117.2^\circ$ , 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  $\phi$  and  $\psi$  values (ignoring deviations in  $\omega$ , 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  $\omega$  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  $\omega$ , every  $(\phi, \psi)$  pair points to exactly one  $(d, \theta)$ . However, when using Eqns. 1 and 2, one value of  $\omega$  can not be replaced with a *periodically equivalent* version of  $\omega$  (the same can be said for  $\phi$  and  $\psi$ ). For example, using  $\omega = x + 2\pi$  instead of  $\omega = x$  will maintain the magnitude of  $d$  and  $\theta$ , 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  $\omega$  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  $(\phi, \psi)$  and  $(d, \theta)$  is only possible if one insists on specific values for  $\omega$ s. For this reason, this report proposes to wrap the value of an amide backbone  $\omega'$  between  $[\Delta, \Delta + 360^\circ]$  using

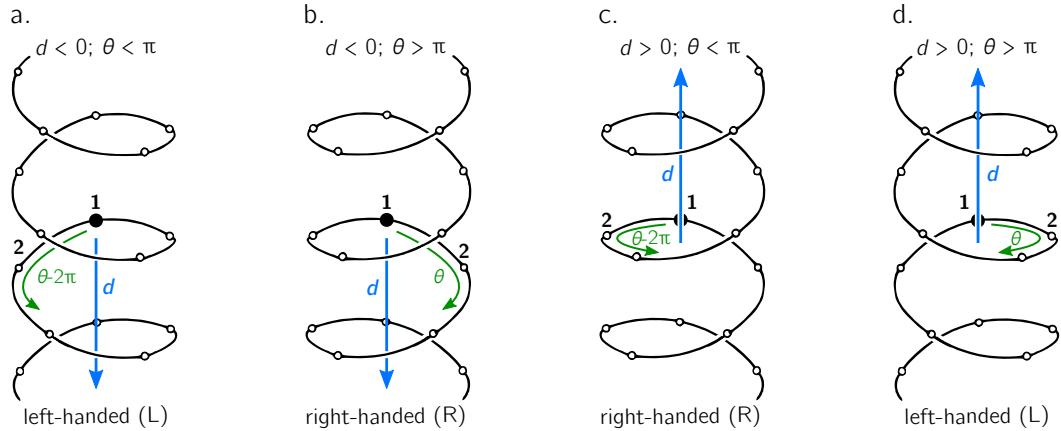
$$\omega = (\omega' - \Delta)\%360 + \Delta, \quad (9)$$

where  $\%$  represents the modulus function, and  $\Delta$  describes the start of the range  $[\Delta, \Delta + 2\pi]$ . Choosing  $\Delta = -90^\circ$  would ensure that the distribution of both *cis* ( $\omega = 0 \pm 5^\circ$ ) and *trans* ( $\omega = 180 \pm 5^\circ$ ) will remain contiguous. Using this system, *cis* and *trans* backbones are respectively represented by  $\omega = 0$  (and not  $2\pi$ ) and  $\omega = \pi$  (not  $-\pi$ ) for *trans* backbones. The rest of this report assumes these values of  $\omega$  for *cis* and *trans* backbones.

These points lead to the conclusion that a *strict* one to one-to-one correspondence between  $(\phi, \psi, \omega)$  and  $(d, \theta)$  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  $(\phi, \psi, \omega)$ , whose  $\phi$  and  $\psi$  lie within a preset range – e.g.,  $[0, 2\pi)$  or  $[-\pi, \pi)$  – and whose  $\omega$  does not change after being wrapped by Eqn. 9.

### Introducing an equation for backbone handedness

The helical parameters  $d$  and  $\theta$  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  $\theta$  and  $d$  together are instrumental in describing backbone handedness.



**Figure 3.** The handedness of a helix is a function of angular displacement  $\theta$  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 and Klyne, 2013); however, as discussed in the Methods section, this report makes a distinction between helix handedness and molecular chirality.

173 The relationship between handedness and  $(d, \theta)$  is shown in Fig. 3. While  $\theta$  indicates the extent to  
 174 which a regular backbone curves along a helical path, the handedness of a backbone is dependent on both  
 175  $\theta$  and  $d$ . This is because the sign of  $d$  provides a frame of reference for interpreting  $\theta$ . In particular, if  $d$  is  
 176 negative, then  $0 < \theta < \pi$  indicates left-handedness (Fig. 3a), while  $\pi < \theta < 2\pi$  indicates a right-handed  
 177 helix (Fig. 3b). However, if  $d$  is positive, then the manner in which the helix is ‘built’ reverses, and  
 178  $0 < \theta < \pi$  indicates right-handedness (Fig. 3c), while  $\pi < \theta < 2\pi$  indicates left-handedness (Fig. 3d).

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

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

179 The range of  $h$  is  $[-1, 1]$ , with negative (or positive) values indicating that the overall twist of the backbone  
 180 is left(or right)-handed. Also,  $|h|$  is proportional to the extent to which the backbone is twisted. Note  
 181 that  $d/|d|$  is related to the traditional sign function  $\text{sgn}(d)$ , but deviates at  $d = 0$ , where the former term  
 182 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  
 183 integer); for more on the meaning of  $d$  and  $\theta$  in context of handedness and peptide geometry, please refer  
 184 to the Results and Discussions section and Fig. 4 in particular.

### 185 Alternative measures of handedness

Two estimates for chirality,  $\chi_1$  and  $\chi_2$ , 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:

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

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

186 Here,  $N$  is the peptide length,  $i$  is the peptide residue number and the position of each  $\alpha$ -carbon is  
 187  $N_i$ , with vector  $\mathbf{v}_k \equiv N_{k+1} - N_k$ . The scalar component of the vector  $\mathbf{v}_i$  is denoted as  $v_i$ . Eqn. 11 has  
 188 range  $[-1, 1]$ . Eqn. 12, also used by Gruziel *et al.* (2013), is the dihedral angle associated with the four  
 189 contiguous  $\alpha$ -carbons (one preceding and two succeeding the residue  $i$ ), and ranges between  $[-\pi, \pi]$

190 radians. For both metrics, values deviating more from 0 are more chiral (or ‘twisted’ or ‘handed’), and  
191 left-handed twists are negative while right-handed twists are positive. Only  $\alpha$ -carbon atom positions are  
192 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:

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

193  $\chi_3$ , of arbitrary range, is known as the chirality index  $G_{0S}$  in Solymosi *et al.* (2002) and Neal *et al.* (2003).  
194 ( $i, j, k, l$ ) are exhaustive permutations of  $\{1, 2, \dots, N\}$ . This metric qualitatively matches the values of  
195 Eqns. 11 and 12, and, while not shown, the relationship between  $(\phi, \psi)$  and  $\chi_3$  is available in the online  
196 GitHub repository.

### 197 Backbone structure generation

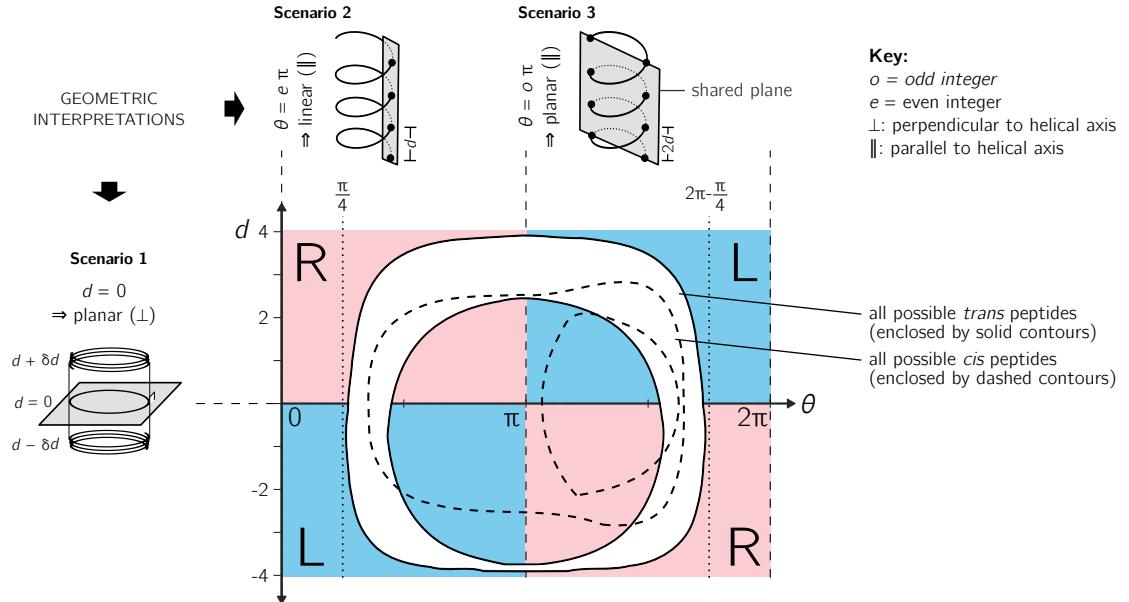
198 The metric  $h$  (Eqn. 10) is purely analytical and does not need structures to be computationally generated,  
199 since Eqns. 5 through 8 that provide  $d$  and  $\theta$  require only pairs of  $\phi$  and  $\psi$  angles. However, if values  
200 for bond angles, lengths and dihedral angles are expected to deviate greatly from equilibrium values,  $\theta$   
201 and  $d$  can only be obtained from the more detailed Eqns. 1 and 2, whose parameters would likely be  
202 obtained from a structure. On the other hand, as  $\chi_1$  (Eqn. 11) and  $\chi_2$  (Eqn. 12) work explicitly with atom  
203 positions, these metrics explicitly need the generation of structures. In order to use these metrics, peptides  
204 (poly-glycines) of arbitrary length were generated using the Python-based PeptideBuilder library (Tien  
205 *et al.*, 2013). Analysis was performed using BioPython (Cock *et al.*, 2009) and Numerical Python (Van  
206 Der Walt *et al.*, 2011). Ramachandran plots that describe chirality (e.g., Fig. 5a) were generated using a  
207 grid spacing (in degrees) of  $\phi, \psi \in \{-180, -178, \dots, 178, 180\}$ .

### 208 Obtaining secondary structure statistics

209 Statistics about secondary structures – particularly  $\alpha$ -helices,  $3_{10}$ -helices and  $\beta$ -sheets – were identified  
210 using the DSSP algorithm (Kabsch and Sander, 1983), although the STRIDE algorithm (Frishman and  
211 Argos, 1995) provides qualitatively identical distributions. The DSSP algorithm was applied to a database  
212 of 13,760 three-dimensional protein conformations (one domain per conformation) with lower than 40%  
213 sequence identity, obtained from the Structural Classification of Proteins or SCOPe website [Release  
214 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>.

### 216 Backbone chirality $\neq$ backbone handedness

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



**Figure 4.** Further discussion on the meaning of  $d$  and  $\theta$ . As shown in Fig. 3, axial separation  $d$  and angular separation  $\theta$  between adjacent atoms of the same type combine to define handedness. The brown (dark) and gold (light) shaded quadrants within the graph show the distribution of handedness as a function of  $d$  and  $\theta$ . 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 as three scenarios. The toroid enclosed by two solid lines (and shaded white) represents all possible conformations for *trans* peptides ( $\omega = 180 \pm 5^\circ$ ). Similarly, the region allowed for *cis* peptides ( $\omega = 0 \pm 5^\circ$ ) are bound by the two dashed contours.

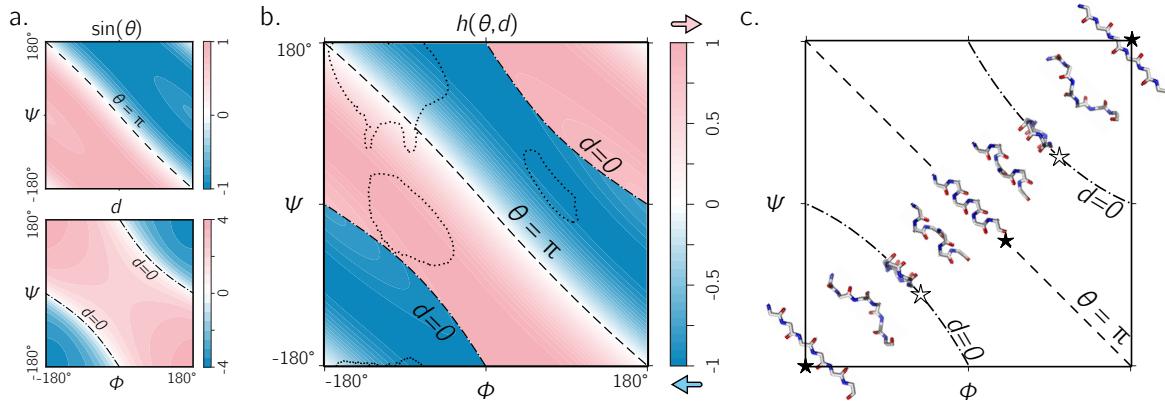
## RESULTS AND DISCUSSION

### Relevance of $\theta$ 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,  $\alpha$ -carbons, or even sidechain  $\beta$ -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  $\theta$  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 Eqn. 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  $\theta$ :  $e\pi$  (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,  $\theta = 0$  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 (for a fixed  $d$ ). Finally, as is evident



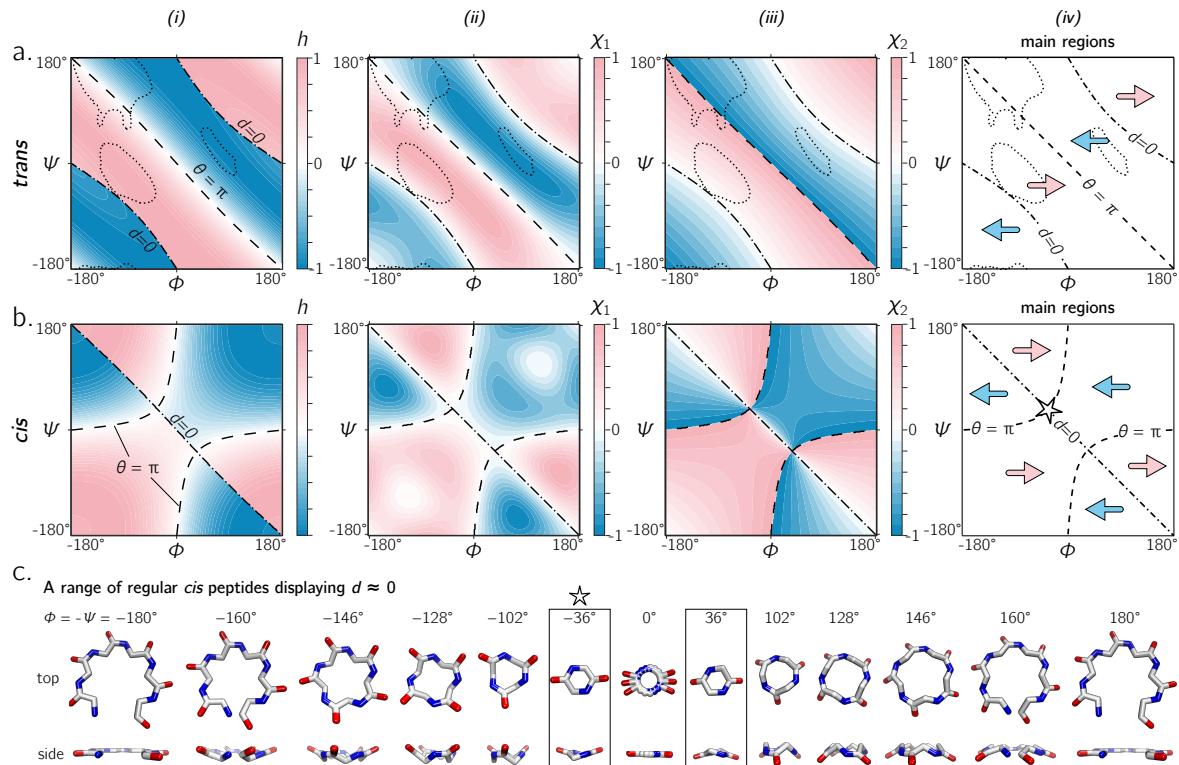
**Figure 5.** The handedness of an ordered *trans* peptide within the Ramachandran plot. Panel (a) displays the relationship between backbone parameters ( $\phi, \psi$ ) and the associated helix parameters of curvature  $\sin(\theta)$  (top; Eqn. 1) and axial displacement  $d$  (bottom; Eqn. 2). As shown 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  $\phi$  and  $\psi$ . 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). Panel (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.

in Fig. 4,  $\theta = e\pi$  conformations are not available to peptide backbones. Therefore,  $\theta = o\pi$  (e.g.,  $\pi$  or  $180^\circ$ ) 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 ( $d, \theta$ ) may be interpreted.

Finally,  $\theta$  may serve as an important single-number metric for describing backbone configurations. Mannige *et al.* (2016) developed one such number – a Ramachandran number ( $\mathcal{R}$ ) – that is a structurally meaningful combination of  $\phi$  and  $\psi$ . 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 *et al.*, 2003; Zacharias and Knapp, 2013; Mannige *et al.*, 2016). Interestingly,  $\theta$  follows that trend too, which – in combination with the fact that regions of the Ramachandran plot are sparse (Mannige *et al.*, 2016) – means that  $\theta$  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  $\theta$ 's will also be conserved in structure (see, e.g., Fig. 5a, top). This feature of  $\theta$  will be true irrespective of the nature of the amide dihedral angle  $\omega$  (Eqn. 1).

## Handedness of *trans* backbones

Fig. 5a describes the behavior of  $\sin(\theta)$  and  $d$  as a function of  $\phi$  and  $\psi$  (assuming an all-*trans* backbone;  $\omega = \pi$  or  $180^\circ$ ). Fig. 5b describes the behavior of backbone handedness ( $h$ ; Eqn. 10) as a function of  $\phi$  and  $\psi$ . 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,  $d = 0$  (‘ $\star$ ’) indicates that each residue is at the same ‘altitude’, i.e., the helix is perfectly flat and maximally curved (at that particular  $\theta$ ). 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. 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 ‘ $\star$ ’ 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  $\theta$ ) 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.



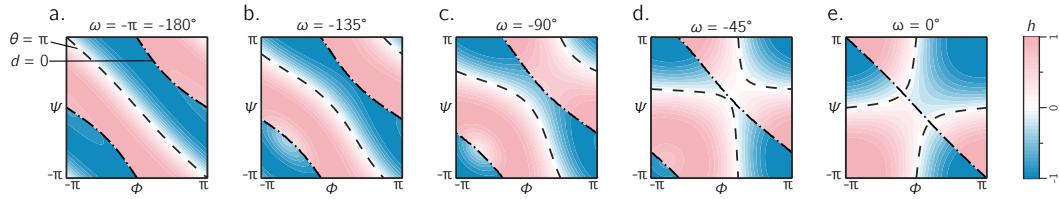
**Figure 6.** Panels (a) and (b) describe the handedness of backbone twists whose amide dihedral angles are *trans* ( $\omega = \pi$ ) and *cis* ( $\omega = 0$ ), respectively. Column (i) describes handedness ( $h$ ; Eqn. 10), which does not require structures to be computationally generated. Columns (ii) and (iii) respectively show vector-based estimates of backbone handedness –  $\chi_1$  (Eqn. 11) and  $\chi_2$  (Eqn. 12) – which are calculated from computationally generated peptides (see Methods). Regions of left- and right-handedness are identical for all measures (i–iii). A cartoon representation of distinct regions of handedness is shown in (iv). Finally, Panel (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  $\theta = \pi$ :  $\phi = -\psi = \pm 36^\circ$  [ $\star$  in Panels(b)–(iv) and (c)].

Fig. 6a shows that the equation for  $h$  match other metrics for handedness, as interpreted by other metrics for chirality (Kwicinski 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  $\chi_1$  [(ii); Eqn. 11] and  $\chi_2$  [(iii); Eqn. 12]. Each panel describes identical regions of left- and right-handedness, which is shown as a cartoon in (iv). However, given that  $\chi_1$  and  $\chi_2$  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  $h$ ,  $\chi_1$  and  $\chi_2$  as a function of  $\phi$  and  $\psi$  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  $\theta$ , 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 glance, this appears to be contradiction, because  $d = 0$  indicates the most curved backbone at a fixed  $\theta$ , 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



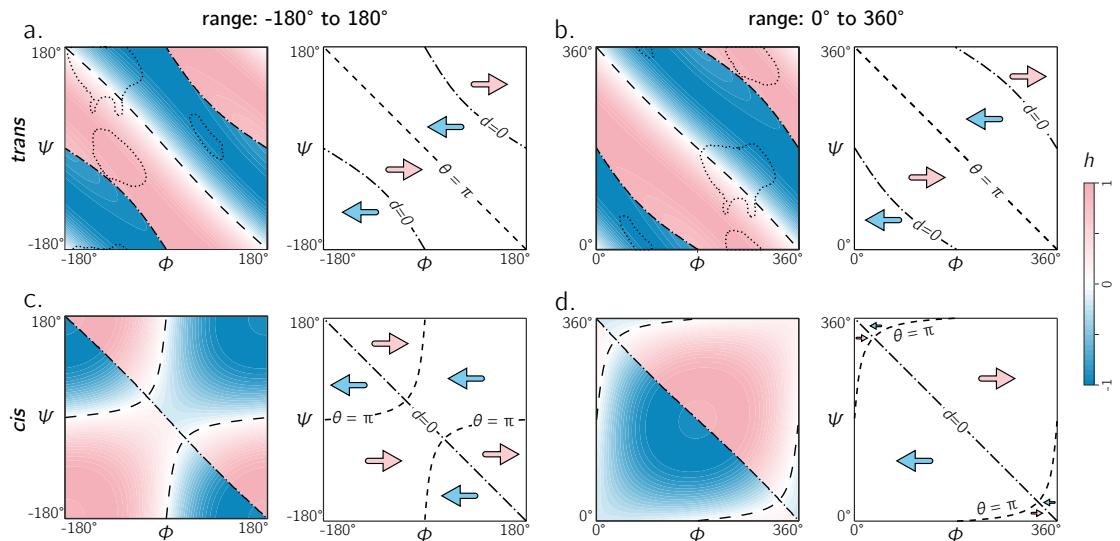
**Figure 7.** The landscape of backbone chirality as a function of amide dihedral angle  $\omega$ . As  $\omega$  is changed, the features of the landscape smoothly transform from the landscapes of  $\omega = \pm\pi$  to  $\omega = 0$ . For all values of  $\omega$ , it is evident that the naïve view 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  $\omega$ ) are shown, all integer values of  $\omega$  were tested, which corroborates the fact that the naïve view of backbone handedness (Fig. 1d) is universally incorrect.

306 with length two, given that any peptoid of length greater than two would result in overlapping atoms.  
 307 However, such a structure (one with  $d = 0$  and  $\theta = \pi$ ) is not possible in *trans* peptides, even in theory,  
 308 because the boundaries associated with  $d = 0$  and  $\theta = \pi$  do not intersect (Fig. 6a); this is also evident in  
 309 Fig. 4, where *trans* peptides are shown to not occupy regions of  $(d, \theta) = (0, \pi)$ , while *cis* peptides do.

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

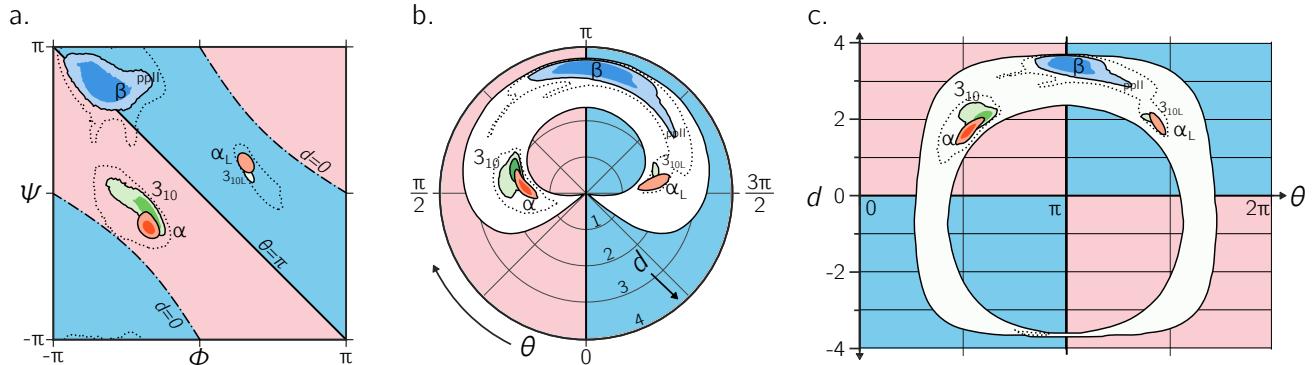
### 318 [-π, π] or [0, 2π]: which frame of reference to use?

319 In structural biology,  $\phi$  and  $\psi$  within the Ramachandran plot has been historically set to range between  
 320 the values  $[-\pi, \pi]$  radians [see, e.g., textbooks by Berg *et al.* (2010) and Alberts *et al.* (2002)]. However,



**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 ( $p = 0.9$ ).

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 ( $\phi$ ,  $\psi$ ) 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  $\phi$  and  $\psi$ ). For this reason, Zacharias and Knapp (2013) introduced a graphical format involving the helical parameters  $d$  and  $\theta$  in polar coordinate space (b), where the regions of left- and right-handedness are obvious [their format differs from (b) in that their  $\theta$  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  $(\theta, d)$ , 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 ( $\alpha_L$ -,  $3_{10L}$ -helices; see Methods), each secondary structure has two contours signifying  $p = 0.5$  and 0.8.

321 Ramachandran *et al.* (1963) had originally used the range of  $[0, 2\pi]$ . Today, the range  $[-\pi, \pi]$  is used  
 322 predominantly by structural biologists (Laskowski *et al.*, 1993; Laskowski, 2003; Zacharias and Knapp,  
 323 2013), while some have turned to  $[0, 2\pi]$  as the norm (Némethy *et al.*, 1966; Voelz *et al.*, 2011).

324 Given the periodicity of the Ramachandran plot, the two frames of reference are scientifically identical;  
 325 however the value of the Ramachandran plot lies in its utility as a *map*: it is a map of important features of  
 326 proteins relative to the various regions, quadrants, and diagonals in the map [see, e.g., discussions by Beck  
 327 *et al.* (2008)]. The Ramachandran plot's value lies in being able to convey large amounts of information  
 328 in easy to read pictograms. For that reason, switching the map from one range to another means that the  
 329 two types of scientists – each used to a distinct range – will not be able to converse as seamlessly.

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 handedness of a *trans* backbone (a,b)  
 332 and *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 ( $\theta = \pi$ )  
 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  $2\pi$  (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  $2\pi$  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 (Berg *et al.*, 2010; Alberts *et al.*, 2002), as discussed throughout this report, it is not easy to estimate  
 344 features of a peptide backbone just from its  $(\phi, \psi)$  angles (Fig. 9a). This prompted Zacharias and  
 345 Knapp (2013) to introduce a new representation for backbone degrees of freedom in the form of a  
 346 polar graph. In this polar representation, the  $\theta$  is the angular coordinate (azimuth) and  $d$  is the radial  
 347 coordinate. An example of one such representation is shown in Fig. 9b, with the direction of increasing  $\theta$   
 348 reversed (compared to the cited report) to maintain relative positions of secondary structures within the  
 349 Ramachandran plot (Fig. 9a). Zacharias and Knapp (2013) stated an additional reason for the introduction

350 of the polar representation (Fig. 9b):  $\theta$ , which is an angle and therefore periodic, can remain periodic as  
351 the angular coordinate in the graph.

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

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

### 373 A departure from perfect regularity

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

380 Some secondary structures are characterized by the regular combination of two or more sets of  
381  $[\phi, \psi]$  (Pauling and Corey, 1951b,a; Armen *et al.*, 2004; Daggett, 2006; Hayward and Milner-White,  
382 2008; Mannige *et al.*, 2015, 2016). For example, the  $\Sigma$ -strand is constructed by alternating between  
383 two backbone states  $(\phi, \psi, \omega) = (-A, B, 180^\circ)$  and  $(-B, A, 180^\circ)$ , where  $A \approx 120$  and  $B \approx 90$  [Fig. 4h in  
384 Mannige *et al.* (2015)]. It was found that the two states are similar in the extent to which the backbone  
385 twists, but opposite in handedness, which allows for these secondary structures to remain linear, albeit in a  
386 meandering way (Mannige *et al.*, 2015). Eqn. 10 also describes these two states as opposite in handedness  
387 and similar in twist extent: the  $h$  for the two states are  $-0.34$  and  $0.51$ , respectively (the difference in  
388 magnitude is within the range of the standard deviation in  $h$  [ $0.391$  for the  $\beta$ -sheet]). Similarly, the  $\alpha$ -sheet  
389 proposed by Pauling and Corey (1951a) is constructed by alternating between  $\alpha_{(D)}$  and  $\alpha_L$  backbone  
390 states, yet this motif is linear because each state describes equal but opposite handedness  $h = \pm 0.41$ .  
391 These points raise the possibility that, even in the absence of perfect backbone regularity, the values  $d$ ,  
392  $\theta$ , and  $h$  may be considered to be residue-specific properties that may be combined to readily provide  
393 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  $\theta$   
397 (angular displacement) and  $d$  (axial displacement), ranges from  $-1$  and  $1$ , and is negative (or positive)  
398 when the backbone twist is left(or right)-handed (with larger  $|h|$  indicating greater extent of twistedness).  
399 This metric ( $h$ ) was used to characterize every regular backbone’s twist within the Ramachandran plot,  
400 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  
402 (-ve) diagonal. Interestingly, the reason for the naïve view makes sense when considering only *trans*

403 peptides: the -ve diagonal ('--' in Figs. 5a) separates *D* and *L* twists if one considers only the regions  
404 dominantly occupied by structured proteins ('....' in Figs. 5a). Plotting the backbone handedness (*h*) in the  
405 two common frames of reference –  $\phi, \psi \in [-\pi, \pi]$  and  $[0, 2\pi]$  – indicates that the less commonly used  
406 frame  $[0, 2\pi]$  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 that have a combination of *cis* and *trans* backbones can  
409 not use one Ramachandran plot to faithfully describe these structures. Interestingly, the parameters  $\theta$   
410 and  $d$  combine all features (internal coordinates) of a contorting backbone, including the amide dihedral  
411 angle  $\omega$ , which means that  $(\theta, d)$  can describe *any* peptide backbone, irrespective of  $\omega$ . Therefore, the  
412 Cartesian plot with  $\theta$  and  $d$  as the x- and y-axis, respectively, serves as a unique plot for *any* peptide  
413 backbone (Fig. 9), with specific values and boundaries containing deep structural meaning (Fig. 4). These  
414 discussions, the author hopes, clarify a number of concepts associated with the Ramachandran plot, while  
415 providing new insights into how to interrogate the features of new protein and protein-like structures.

## 416 ACKNOWLEDGMENTS

417 The author thanks Alana Canfield Mannige, Ronald D. Hills Jr, the editor, and the reviewer for their  
418 constructive input.

## 419 REFERENCES

- 420 **Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002.** Molecular biology of the cell.  
421 new york: Garland science; 2002. *Classic textbook now in its 5th Edition* .
- 422 **Armen RS, DeMarco ML, Alonso DO, Daggett V. 2004.** Pauling and corey's  $\alpha$ -pleated sheet structure  
423 may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proceedings of the National  
424 Academy of Sciences of the United States of America* **101**(32):11622–11627.
- 425 **Beck DA, Alonso DO, Inoyama D, Daggett V. 2008.** The intrinsic conformational propensities of the  
426 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the  
427 National Academy of Sciences* **105**(34):12259–12264.
- 428 **Bentley R. 2010.** Chiral: a confusing etymology. *Chirality* **22**(1):1–2.
- 429 **Berg JM, Tymoczko JL, Stryer L. 2010.** *Biochemistry, International Edition*. WH Freeman & Co.,  
430 New York, 7 edition.
- 431 **Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov IN, Bourne PE.**  
432 2000. The protein data bank. *Nucleic Acids Research* **28**(1):235–242.
- 433 **Bragg L, Kendrew JC, Perutz MF. 1950.** Polypeptide chain configurations in crystalline proteins. *Pro-  
434 ceedings of the Royal Society of London Series A Mathematical and Physical Sciences* **203**(1074):321–  
435 357.
- 436 **Chothia C, Hubbard T, Brenner S, Barns H, Murzin A. 1997.** Protein folds in the all-beta and  
437 all-alpha classes. *Annual Review of Biophysics and Biomolecular Structure* **26**:597–627.
- 438 **Cock P, Antao T, Chang J, Chapman B, Cox C, Dalke A, Friedberg I, Hamelryck T, Kauff F,  
439 Wilczynski B, de Hoon M. 2009.** Biopython: freely available python tools for computational molecular  
440 biology and bioinformatics. *Bioinformatics* **25**(11):1422–1423.
- 441 **Cooper GM, Hausman RE. 2013.** *The Cell: A Molecular Approach*. Sinauer Associates, Inc., Sunder-  
442 land, MA, 6 edition.
- 443 **Cross L, Klyne W. 2013.** *Rules for the Nomenclature of Organic Chemistry: Section E: Stereochemistry  
(Recommendations 1974)*. Elsevier.
- 444 **Daggett V. 2006.**  $\alpha$ -sheet: the toxic conformer in amyloid diseases? *Accounts of Chemical Research*  
445 **39**(9):594–602.
- 446 **Eisenberg D. 2003.** The discovery of the  $\alpha$ -helix and  $\beta$ -sheet, the principal structural features of proteins.  
447 *Proceedings of the National Academy of Sciences* **100**(20):11207–11210.
- 448 **Engh RA, Huber R. 1991.** Accurate bond and angle parameters for x-ray protein structure refinement.  
449 *Acta Crystallographica Section A: Foundations of Crystallography* **47**(4):392–400.
- 450 **Engh R, Huber R. 2006.** Structure quality and target parameters. In *International Tables for Crystallog-  
451 raphy Volume F: Crystallography of biological macromolecules*, pages 382–392. Springer.
- 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**.

- 455 **Ferrarini A, Nordio PL. 1998.** On the assessment of molecular chirality. *Journal of the Chemical  
456 Society, Perkin Transactions 2* (2):455–460.
- 457 **Fox NK, Brenner SE, Chandonia JM. 2014.** Scope: Structural classification of proteins—extended, inte-  
458 grating scop and astral data and classification of new structures. *Nucleic Acids Research* **42**(Database  
459 issue):D304–D309.
- 460 **Frishman D, Argos P. 1995.** Knowledge-based protein secondary structure assignment. *Proteins:  
461 Structure, Function, and Bioinformatics* **23**(4):566–579.
- 462 **Gold V, Loening K, McNaught A, Shemi P. 1997.** IUPAC compendium of chemical terminology.  
463 *Blackwell Science, Oxford*.
- 464 **Gorske BC, Mumford EM, Conry RR. 2016.** Tandem incorporation of enantiomeric residues engenders  
465 discrete peptoid structures. *Organic Letters*.
- 466 **Gruziel M, Dzwolak W, Szymczak P. 2013.** Chirality inversions in self-assembly of fibrillar superstruc-  
467 tures: a computational study. *Soft Matter* **9**(33):8005–8013.
- 468 **Harris AB, Kamien RD, Lubensky TC. 1999.** Molecular chirality and chiral parameters. *Reviews of  
469 Modern Physics* **71**(5):1745.
- 470 **Hayward S, Milner-White EJ. 2008.** The geometry of  $\alpha$ -sheet: Implications for its possible function as  
471 amyloid precursor in proteins. *Proteins: Structure, Function, and Bioinformatics* **71**(1):415–425.
- 472 **Ho BK, Thomas A, Brasseur R. 2003.** Revisiting the ramachandran plot: Hard-sphere repulsion,  
473 electrostatics, and h-bonding in the  $\alpha$ -helix. *Protein Science* **12**(11):2508–2522.
- 474 **Hooft RW, Sander C, Vriend G. 1997.** Objectively judging the quality of a protein structure from a  
475 ramachandran plot. *Computer Applications in the Biosciences: CABIOS* **13**(4):425–430.
- 476 **Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK. 2002.** Intrinsic disorder in  
477 cell-signaling and cancer-associated proteins. *Journal of Molecular Biology* **323**(3):573–584.
- 478 **Impróta R, Vitagliano L, Esposito L. 2011.** Peptide bond distortions from planarity: new insights from  
479 quantum mechanical calculations and peptide/protein crystal structures. *PLoS One* **6**(9):e24533.
- 480 **Impróta R, Vitagliano L, Esposito L. 2015a.** Bond distances in polypeptide backbones depend on the  
481 local conformation. *Acta Crystallographica Section D: Biological Crystallography* **71**(6):1272–1283.
- 482 **Impróta R, Vitagliano L, Esposito L. 2015b.** The determinants of bond angle variability in pro-  
483 tein/peptide backbones: A comprehensive statistical/quantum mechanics analysis. *Proteins: Structure,  
484 Function, and Bioinformatics* **83**(11):1973–1986.
- 485 **Kabsch W, Sander C. 1983.** Dictionary of protein secondary structure: pattern recognition of hydrogen-  
486 bonded and geometrical features. *Biopolymers* **22**(12):2577–2637.
- 487 **Kwiecińska JI, Cieplak M. 2005.** Chirality and protein folding. *Journal of Physics: Condensed Matter*  
488 **17**(18):S1565.
- 489 **Laskowski RA. 2003.** Structural quality assurance. *Structural Bioinformatics, Volume 44* pages 273–303.
- 490 **Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993.** Procheck: a program to check the  
491 stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**(2):283–291.
- 492 **Linderström-Lang KU. 1952.** *Lane Medical Lectures: proteins and enzymes*, volume 6. Stanford  
493 University Press.
- 494 **Mannige RV. 2014.** Dynamic new world: Refining our view of protein structure, function and evolution.  
495 *Proteomes* **2**(1):128–153.
- 496 **Mannige RV, Haxton TK, Proulx C, Robertson EJ, Battigelli A, Butterfoss GL, Zuckermann RN,  
497 Whitelam S. 2015.** Peptoid nanosheets exhibit a new secondary structure motif. *Nature* **526**:415–420.
- 498 **Mannige RV, Kundu J, Whitelam S. 2016.** The Ramachandran number: an order parameter for protein  
499 geometry. *PLoS One* **11**(8):e0160023.
- 500 **Mirijanian DT, Mannige RV, Zuckermann RN, Whitelam S. 2014.** Development and use of an  
501 atomistic charmm-based forcefield for peptoid simulation. *Journal of Computational Chemistry*  
502 **35**(5):360–370.
- 503 **Mislow K. 2002.** Stereochemical terminology and its discontents. *Chirality* **14**(2-3):126–134.
- 504 **Miyazawa T. 1961.** Molecular vibrations and structure of high polymers. ii. helical parameters of infinite  
505 polymer chains as functions of bond lengths, bond angles, and internal rotation angles. *Journal of  
506 Polymer Science* **55**(161):215–231.
- 507 **Neal MP, Solymosi M, Wilson MR, Earl DJ. 2003.** Helical twisting power and scaled chiral indices.  
508 *Journal of Chemical physics* **119**(6):3567–3573.
- 509 **Némethy G, Leach S, Scheraga HA. 1966.** The influence of amino acid side chains on the free energy

- 510 of helix-coil transitions1. *Journal of Physical Chemistry* **70**(4):998–1004.
- 511 **Orosz F, Ovádi J. 2011.** Proteins without 3d structure: definition, detection and beyond. *Bioinformatics*  
512 **27**(11):1449–1454.
- 513 **Osipov M, Pickup B, Dunmur D. 1995.** A new twist to molecular chirality: intrinsic chirality indices.  
514 *Molecular Physics* **84**(6):1193–1206.
- 515 **Pauling L, Corey RB. 1951a.** Configurations of polypeptide chains with favored orientations around  
516 single bonds: two new pleated sheets. *Proceedings of the National Academy of Sciences of the United*  
517 *States of America* **37**(11):729.
- 518 **Pauling L, Corey RB. 1951b.** The pleated sheet, a new layer configuration of polypeptide chains.  
519 *Proceedings of the National Academy of Sciences of the United States of America* **37**(5):251.
- 520 **Pauling L, Corey RB, Branson HR. 1951.** The structure of proteins: two hydrogen-bonded helical  
521 configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* **37**(4):205–  
522 211.
- 523 **Ramachandran G, Ramakrishnan C, Sasisekharan V. 1963.** Stereochemistry of polypeptide chain  
524 configurations. *Journal of Molecular Biology* **7**(1):95–99.
- 525 **Robertson EJ, Battigelli A, Proulx C, Mannige RV, Haxton TK, Yun L, Whitelam S, Zuckermann  
526 RN. 2016.** Design, synthesis, assembly, and engineering of peptoid nanosheets. *Accounts of Chemical*  
527 *Research* **49**(3):379–389.
- 528 **Shimanouchi T, Mizushima Si. 1955.** On the helical configuration of a polymer chain. *Journal of  
529 Chemical Physics* **23**(4):707–711.
- 530 **Solymosi M, Low RJ, Grayson M, Neal MP. 2002.** A generalized scaling of a chiral index for molecules.  
531 *Journal of Chemical Physics* **116**(22):9875–9881.
- 532 **Sun J, Zuckermann RN. 2013.** Peptoid polymers: a highly designable bioinspired material. *ACS Nano*  
533 **7**(6):4715–4732.
- 534 **Testa B. 2013.** Organic stereochemistry. part 2. *Helvetica Chimica Acta* **96**(2):159–188.
- 535 **Tien MZ, Sydykova DK, Meyer AG, Wilke CO. 2013.** Peptidebuilder: A simple python library to  
536 generate model peptides. *PeerJ* **1**:80.
- 537 **Van Der Walt S, Colbert SC, Varoquaux G. 2011.** The numpy array: a structure for efficient numerical  
538 computation. *Computing in Science & Engineering* **13**(2):22–30.
- 539 **Voelz VA, Dill KA, Chorny I. 2011.** Peptoid conformational free energy landscapes from implicit-solvent  
540 molecular simulations in amber. *Peptide Science* **96**(5):639–650.
- 541 **Wallentin CJ, Orentas E, Wärnmark K, Wendt OF. 2009.** Chirality, a never-ending source of con-  
542 fusion. *Zeitschrift für Kristallographie International Journal for Structural, Physical, and Chemical*  
543 *aspects of Crystalline Materials* **224**(12):607–608.
- 544 **Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004.** Prediction and functional analysis of  
545 native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology* **337**(3):635–  
546 645.
- 547 **Zacharias J, Knapp EW. 2013.** Geometry motivated alternative view on local protein backbone  
548 structures. *Protein Science* **22**(11):1669–1674.