

# On the other hand: chirality and the peptide backbone

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

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

\* rvmannige@lbl.gov

## ABSTRACT

Proteins are a class of biomolecules that display diverse conformations, which is the reason for their diverse functionality. These conformations are afforded in large part due to a protein's backbone, whose twists and contortions allow for a protein to fold into particular conformations. The Ramachandran plot describes a backbone as a function of its dominant degrees of freedom – the backbone dihedral angles  $\phi$  and  $\psi$  (Ramachandran *et al.*, 1963). Since its introduction in the 1960's Ramachandran plots have been a crucial tool to characterize protein backbone features. However, so far, the exact nature of the twist of a backbone, as a function of  $\phi$  and  $\psi$  has not been completely described for both *cis* and *trans* backbones. In these lines, this short paper provides 1) a simple metric ( $\chi$ ) for backbone chirality, 2) rules of thumb as to how to interpret chirality within a Ramachandran plot, and 3) a complete characterization of the way a backbone twists in all regions of the Ramachandran plot for both *cis* and *trans* backbones; this fills in the 'dead space' within the Ramachandran plot, which are regions that are not commonly accessed by proteins. Given the recent discussion on intrinsic disorder, whose locations on the Ramachandran plot are well characterized, and new peptide mimics that are not bound by traditional peptide rules, these results aim to serve as a guide for understanding new types of peptide and peptidomimetic structures. Finally, we provide Python scripts that replicate the results within the paper as a Github repository.

## INTRODUCTION

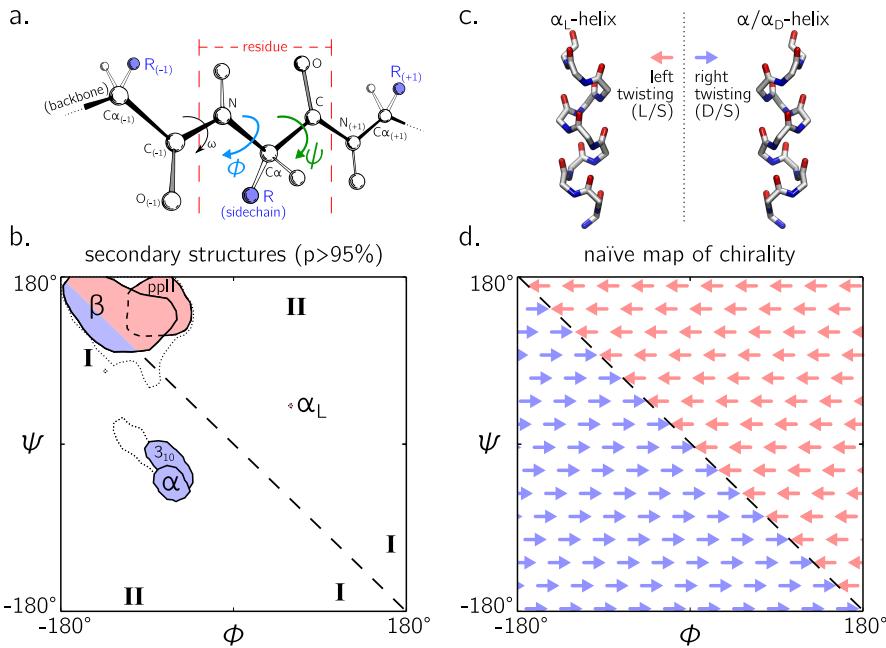
Proteins are a class of biomolecules that are unparalleled in their functionality (Berg *et al.*, 2010). Proteins are able to display diverse functions because of the diverse conformations that they are able to attain. When attempting to characterize a protein's conformation, it is useful to recognize that proteins are linear polymers: beads on a chain, where beads are amino acids, otherwise called residues (Fig. 1a). These polymers may be distinguished into two regions: the backbone (dark bonds in Fig. 1a) and residue-specific sidechains that 'decorate' the backbone (the 'R's in Fig. 1a). Therefore, a protein's conformation is highly dependent on the manner in which a backbone curves in space. For any protein, its collection or ensemble of conformations<sup>1</sup> define possible molecular functionality. In short, understanding how a backbone behaves is of great importance to the field of biochemistry, since understanding the structure of a protein goes a long way towards understanding how proteins function.

Ramachandran *et al.* (1963) recognized that the twist of the backbone is dependent on the dihedral angles  $\phi$  and  $\psi$  (Fig. 1a)<sup>2</sup>. Any peptide built with uniform twist or backbone parameters (say,  $\phi = X^\circ$  and  $\psi = Y^\circ$ ) will result in a regular structure; some regular structures are energetically stable (or metastable) and are called secondary structures (Fig. 1b). These secondary structures pack together with the help of loops, which are also twisted (Berg *et al.*, 2010). Today, two-dimensional ( $\phi, \psi$ ) plots are called Ramachandran plots, and have been used as a guide for understanding a peptide backbone's general conformational state or 'twistedness' at a glance (Ho *et al.*, 2003; Berg *et al.*, 2010; Alberts *et al.*, 2002; Subramanian, 2001; Laskowski *et al.*, 1993; Hooft *et al.*, 1997; Laskowski, 2003).

So far, our understanding of the Ramachandran plot has been limited to the secondary structures and loops displayed by proteins (Berman *et al.*, 2000; Alberts *et al.*, 2002). Structural biologists are aware that the negatively sloping diagonal (dashed line in Fig. 1b; henceforth denoted as the '-ve diagonal')

<sup>1</sup>These conformations can be specific folds in globular proteins or larger structural ensembles in intrinsically disordered proteins; Mannige (2014).

<sup>2</sup>Another degree of freedom – the amide dihedral angle ( $\omega$ ) – that is predominantly *trans* in peptides (i.e.,  $\omega = 180^\circ$ ), but which can also occupy *cis* conformations ( $\omega = 0^\circ$ ). The Ramachandran plot requires a presumption that the backbone is either *cis* or *trans*.



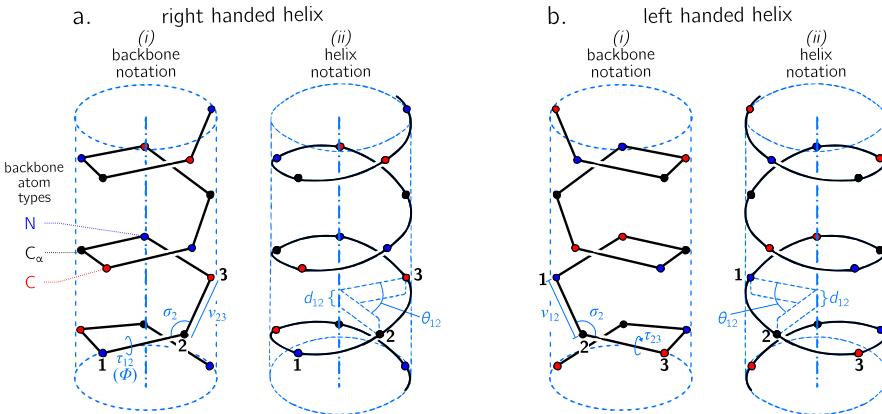
**Figure 1.** A peptide is composed of a linear chain of residues (a). Each peptide is structurally distinguished into backbone atoms (connected by dark shaded bonds in a) and sidechain motifs that stick out of each residue's  $\alpha$ -carbon ( $R$ ). The structures available to a peptide are mainly due to the dihedral angles  $\phi$  and  $\psi$  (and in smaller part,  $\omega$ ). For example, repeating a value for  $\phi$  and  $\psi$  for every residue within a peptide – i.e., selecting a point on the Ramachandran plot (b) – results in regular (secondary) structures that then combine to form most globular proteins in the protein databank (Berman *et al.*, 2000). The twist of each perfectly regular secondary structure possesses a distinct chirality: left-handed, or right-handed (examples shown in c); rarely, peptides may be non-chiral or flat (e.g., when  $\phi = \psi = \pi$ ). As evidenced in (c) the -ve diagonal within the Ramachandran plot (dashed line) 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 ( $\omega = \pi$ ).

demarcates a change in backbone chirality. For example, the position of the idealized left- and right-handed  $\alpha$ -helices (Fig. 1c) – respectively denoted as  $\alpha_L$  and  $\alpha$  in Fig. 1b – are on opposite sides of the the -ve diagonal<sup>3</sup>. Additionally, the the  $\beta$ -strand exists predominantly on the right of the -ve diagonal, and their backbones are predominantly left-handed [see, e.g., discussions within Quiocho *et al.* (1977) and Shaw and Muirhead (1977)<sup>4</sup>]. Indeed, all regular/ordered motifs within proteins that lie to the right (or top) of the -ve diagonal are left-handed in backbone twist; as a corollary, all ordered backbones to the left (or under) the -ve diagonal are right handed in nature (primarily, the  $\alpha$ -,  $\pi$ - and  $3_{10}$ -helix; Fig. 1b;  $\pi$ -helices are not shown due to low frequency in the protein databank). Naïvely, these observations lead to the hypothesis that the -ve diagonal of the Ramachandran plot separates the left handed backbones from the right handed backbones (Fig. 1d). While even Pauling and colleagues were cognizant of this general trend in peptide secondary structure [see, e.g., notes on what is now known as the  $\alpha$ -sheet; Pauling *et al.* (1951); Pauling and Corey (1951b,a)], the picture in Fig. 1d is an often assumed one by protein structural biologists, even though its incorrectness is evident in at least one earlier report (Zacharias and Knapp, 2013).

While ‘twist’ chirality in regular protein secondary structures is well understood, peptide mimics

<sup>3</sup>Note that left- and right-handed backbone twists are respectively associated with the L and D chiralities within the Fisher Projection system and S and R chiralities within the Cahn–Ingold–Prelog system (Cross and Klyne, 2013).

<sup>4</sup>Note that a different metric for handedness in  $\beta$  sheets exists today that addresses the handedness of a particular hydrogen bonding network and not the handedness of the participating peptide's backbone. See, e.g., Schulz *et al.* (1974) and Chothia *et al.* (1977).



**Figure 2.** Internal coordinate (i) and helical coordinate (ii) representations of right-handed (a) and left-handed (b) regular backbones. Internal coordinates are a function of bond lengths (e.g.,  $v_{12}$ ), angles ( $\sigma_2$ ), 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 from the helical axis ( $\rho$ ; not shown). Representations are adopted from Fig. 1 in (Shimanouchi and Mizushima, 1955).

– especially peptoids (Sun and Zuckermann, 2013) – display new secondary structures that fall out of the regions regularly occupied by proteins. Peptoids are distinguished from peptides by the position of each sidechain on the backbone [sidechains are attached to the backbone nitrogen atom rather than the  $\alpha$ -carbon; Sun and Zuckermann (2013)]. Certain peptoids display secondary structures in regions of the Ramachandran plot that are not well characterized (Sun and Zuckermann, 2013; Goodman *et al.*, 2007; Culf and Ouellette, 2010; Beke *et al.*, 2006; Pohl *et al.*, 2012; Zuckermann and Kodadek, 2009; Sun and Zuckermann, 2013). For example, a ‘higher-order’ peptoid secondary structure – the  $\Sigma$ -strand (Mannige *et al.*, 2015; Robertson *et al.*, 2016) – samples regions of the Ramachandran plot (‘I’ in Fig. 1d) that are not permitted within natural proteins (this is because of the lack of backbone hydrogen bond donors in peptoids). Another secondary structure – the ‘ $\omega$ -strand’ (Gorske *et al.*, 2016) – samples similarly ‘historically uncharted’ regions of the Ramachandran plot (‘II’ in Fig. 1d). Importantly, handedness plays a crucial part in explaining these new motifs: as one goes along the backbones of these secondary structures, alternating residues occupy equal but opposite backbone twists [for this reason, the  $\sigma$  sheet is relatively linear, albeit meandering; Mannige *et al.* (2015); Mannige *et al.* (2016)].

In light of these new secondary structures, and in anticipation of the discovery of additional (higher-order) secondary structures, we chose to perform a detailed study into how regular backbones twist in each region of the Ramachandran plot. This question builds on a previous report that discussed handedness in all-*trans*-backbones (Zacharias and Knapp, 2013), and would complete our understanding of a very basic component of biochemistry (handedness and twist of a backbone). An additional aim is to completely dispel the naïve view in Fig. 1d.

## 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; Gruziel *et al.*, 2013), optical activity (Osipov *et al.*, 1995), and molecular shape (Ferrarini and Nordio, 1998). However, we will focus on a simpler metric for chirality associated with an idealized helix within which all atoms of the backbone sit (Shimanouchi and Mizushima, 1955; Miyazawa, 1961; Zacharias and Knapp, 2013). We do this because any *regular* backbone can be associated with a specific helix<sup>5</sup>, and, as we will discuss below, helices have handedness baked into the equations. We will first introduce the use of helical coordinates (and how peptides may be treated in such contexts), after which we will show how some parameters within the helical coordinate system (namely  $d$  and  $\theta$ ) may be combined

<sup>5</sup>Indeed, the word ‘helix’ has a broader meaning that comes from the Greek word  $\xi\lambdaι\xi$  that means ‘twisted, curved’; Liddell *et al.* (1894).

to describe a physically unambiguous meaning of handedness of a backbone's twist.

### 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). Shimanouchi and Mizushima (1955) had derived a set of equations that fit a platonic helix to the atoms within a *regularly arranged* backbone. 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. Fig. 2 describes an arbitrary peptide backbone that may be represented either using internal coordinates (*i*) or helical coordinates (*ii*).

Internal coordinates are associated with (stereo)chemical terms: bond lengths ( $v_{ij}$ ) between adjacent atoms  $i$  and  $j$ , bond angles ( $\sigma_i$ ) between the two backbone bonds adjacent to atom  $i$ , and dihedral or torsion angles ( $\tau_{ij}$ ), which are associated with the bond  $i - j$  and the atoms preceding and succeeding these atoms. Helical coordinates (Fig. 2(*ii*)) are described using measures of vertical displacement ( $d$ ; this is related to the pitch of a platonic helix), angular displacement ( $\theta$ ) and the radius of the helix ( $\rho$ ; not shown in Fig. 2)<sup>6</sup>. Given that there are three atoms associated with a residue (Fig. 1a),  $d$  and  $\theta$ , which are per-residue values, are described in terms of distances ( $d_{hk}$ ) and angles ( $\theta_{hk}$ ) between adjacent atoms, where  $(h,k) \in \{(N_i, C_{\alpha,i}), (C_{\alpha,i}, C_i), (N_i, C_{i+1})\}$ . Here, N, C<sub>α</sub>, and C are the backbone nitrogen, α-carbon, and carbonyl carbon atoms (Fig. 1a), and the index  $i$  (and  $i + 1$ ) indicates residue position.

The notation used by Shimanouchi and Mizushima (1955) were in terms of matrices, which were then simplified by Miyazawa (1961) into trigonometric terms<sup>7</sup>. In particular, Miyazawa (1961) noted that the total residue-residue vertical displacement ( $d$ ) and angular displacement ( $\theta$ ) may be retrieved using the following two equations.

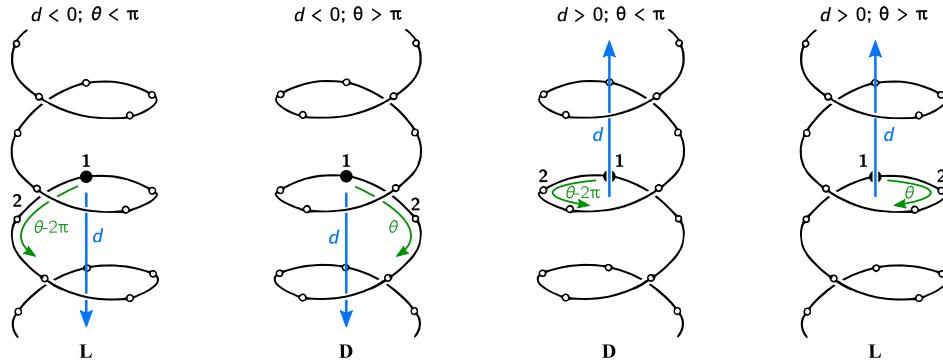
$$\begin{aligned} \cos\left(\frac{\theta}{2}\right) = & \cos\left(\frac{+\phi + \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right) \sin\left(\frac{\sigma_\alpha}{2}\right) \sin\left(\frac{\sigma_c}{2}\right) \\ & - \cos\left(\frac{+\phi - \psi + \omega}{2}\right) \sin\left(\frac{\sigma_n}{2}\right)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

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

Here, subscripts ‘n’, ‘α’, and ‘c’ respectively refer to the backbone nitrogen, α-carbon and carbonyl carbon atoms;  $v_{ij}$  refers to the distance between adjacent atoms  $i$  and  $j$ ;  $\sigma_i$  refers to the angle between the adjacent bonds that have  $i$  as the common atom. Finally,  $\phi$ ,  $\psi$ , and  $\omega$  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. Given that backbone bond lengths and angles are generally less tunable when compared to dihedral angles,

<sup>6</sup>Miyazawa (1961) also reported a helix parameter related to  $\theta$  and  $d$  – the radius ( $\rho$ ) of the helix (Eqn. 38 in that report) – which will not be discussed here, as its meaning with regards to handedness is redundant.

<sup>7</sup>Interestingly, the formalisms described by Miyazawa (1961) and Shimanouchi and Mizushima (1955) apply to repeating linear polymers of arbitrary complexity; therefore, these formalisms can be applied to even novel protein mimics that do not conserve the number of atoms within the backbone (Kandakov and Ginzburg, 2013), such as β-peptides.



**Figure 3.** The handedness of a helix is a function of angular displacement perpendicular to the helical axis ( $\theta$ ; green curved arrows) and linear displacement along the helical axis  $d$  (blue, vertical arrows). Here, the left- and right-handed twists are respectively referred to as **L** and **D** (shown at the bottom), as per the Fisher Projection system (Cross and Klyne, 2013).

we can keep them fixed. Using revised values for the bond angles and lengths<sup>8</sup>, we arrive at equations for  $\theta$  and  $d$  that are purely a function of the backbone dihedral angles. In particular, Miyazawa (1961) set  $\omega = \pi$  and used preset values for bond angles and lengths to arrive at a simpler equation for *trans* backbones. An updated version of this equation was reported by Zacharias and Knapp (2013) follows.

$$\cos\left(\frac{\theta}{2}\right) = -0.8235 \sin\left(\frac{\phi + \psi}{2}\right) + 0.0222 \sin\left(\frac{\phi - \psi}{2}\right), \quad (3)$$

$$d \sin\left(\frac{\theta}{2}\right) = 2.9986 \cos\left(\frac{\phi + \psi}{2}\right) - 0.6575 \cos\left(\frac{\phi - \psi}{2}\right). \quad (4)$$

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 *et al.*, 2016), for completeness, we report the corresponding relationships for a *cis* backbone ( $\omega = 0$ ) below.

$$\cos\left(\frac{\theta}{2}\right) = 0.4052 \cos\left(\frac{\phi + \psi}{2}\right) - 0.4932 \cos\left(\frac{\phi - \psi}{2}\right), \quad (5)$$

$$d \sin\left(\frac{\theta}{2}\right) = 2.3093 \sin\left(\frac{\phi + \psi}{2}\right) + 0.0028 \sin\left(\frac{\phi - \psi}{2}\right). \quad (6)$$

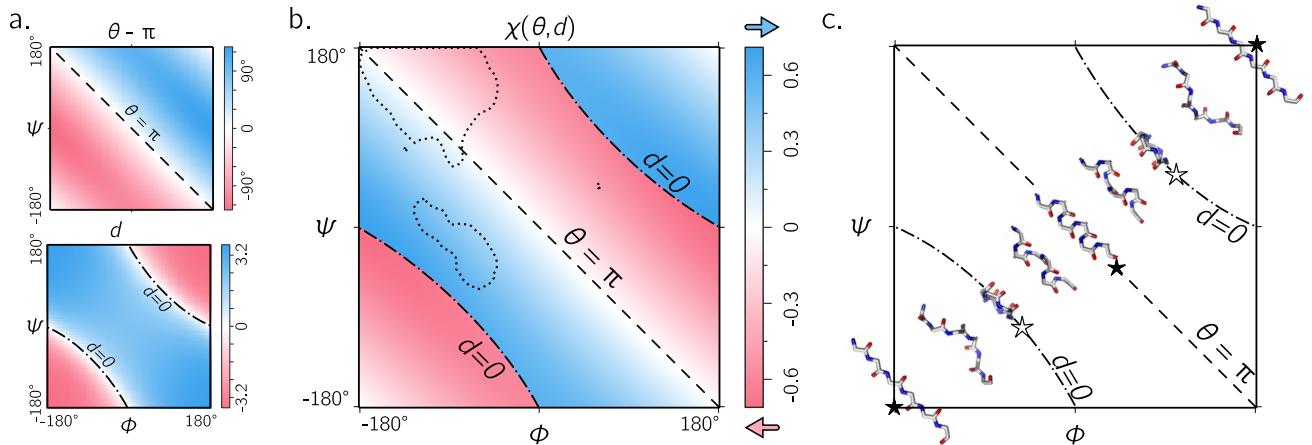
### An equation for backbone handedness.

The helical parameters  $d$  and  $\theta$  host a wealth of information. For example,  $d$  is related to the radius of gyration,  $d = 0$  indicates all atoms of a particular type (e.g., all  $\alpha$ -carbons) are in the same plane, and *asd*. Zacharias and Knapp (2013) observed that  $\theta$  and  $d$  are also instrumental in describing backbone handedness. We describe this relationship first by words and pictures, and then in equation form.

$\theta$ , with range  $[0, 2\pi]$ , gives us an idea of how curved or ‘tightly wound’ the backbone twists. The backbone is least curved when  $\theta$  is  $\pi$  (i.e., when  $\sin(\theta) = 0$ ) and is the most tightly wound when  $\theta$  assumes the value of 90 and 270 (i.e., when  $\sin(\theta) = \pm 1$ ). However,  $\theta$  alone can not distinguish between left- and right-handed backbones. This is because, regarding handedness,  $\theta$  must be interpreted in the context of helix pitch that is proportional to  $d$ . In particular, if  $d$  is positive, then  $\theta < 180$  (or  $\sin(\theta) > 0$ ) indicates right handedness, while  $\theta > 180$  (or  $\sin(\theta) < 0$ ) indicates a left-handed helix. However, if  $d$  is negative, then the conditions reverse, and  $\sin(\theta) > 0$  indicates left-handedness, while  $\sin(\theta) < 0$  indicates right-handedness. This relationship between  $d$ ,  $\theta$  and handedness is reiterated in Fig. 3.

Given that the range of  $\arccos(x)$  is  $[0, \pi]$ ,  $\theta$ ’s range is  $[0, 2\pi]$  (see Eqn. 1). Therefore, we provide the following, simple, metric for backbone handedness that depends on the sign of  $d$  and sign and magnitude

<sup>8</sup>Values used here are also used by Zacharias and Knapp (2013). These values are:  
 $v_{n,\alpha} = 1.459\text{\AA}$ ,  $v_{\alpha,c} = 1.525\text{\AA}$ ,  $v_{c,n(+1)} = 1.336\text{\AA}$ ,  $\sigma_\alpha = 111.0$ ,  $\sigma_c = 117.2^\circ$ , and  $\sigma_n = 121.7^\circ$ .  
 For reference, Miyazawa (1961) originally used the following values:  
 $v_{n,\alpha} = 1.470\text{\AA}$ ,  $v_{\alpha,c} = 1.530\text{\AA}$ ,  $v_{c,n(+1)} = 1.320\text{\AA}$ ,  $\sigma_\alpha = 110.0$ ,  $\sigma_c = 114.0$ , and  $\sigma_n = 123.0$ .



**Figure 4.** The chirality 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  $\theta - \pi$  (top; Eqn. 1) and pitch  $d$  (bottom; Eqn. 2). The boundaries,  $\theta = \pi$  ( $n \in \mathbb{Z}$ ; ‘–’–) and  $d = 0$  (‘–·–’), correspond to backbones that are equally flat, but which are respectively optimally extended and curved (see discussion in text). As shown in Fig. 3, the handedness of a helix is a function of these two variables ( $\chi$ ; Eqn. 7). Panel (b) is a map of backbone chirality ( $\chi$ ) as a function of  $\phi$  and  $\psi$ . This map shows that the naïve expectation of handedness in a Ramachandran plot (Fig. 1d) is inaccurate. Interestingly, our naïve expectations (Fig. 1d) would be upheld if we were only to have sampled regions of the Ramachandran plot dominated by known proteins (a; regions enclosed by ‘....’). 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.

of  $\theta$ :

$$\chi = \frac{d}{|d|} \frac{\pi - \theta}{\pi} \quad (7)$$

Here,  $\chi$  is negative (or positive) when the overall twist of the backbone is left (or right) handed<sup>9</sup>.

## DISCUSSION

### Relevance of $\theta$ and $d$ .

The helical parameters  $d$  and  $\theta$  respectively refer to a displacement along the helical axis and an angular displacement in a plane perpendicular to the helical axis (Fig. 3). Fig. 4a describes the behavior of  $d$  and  $\theta - \pi$  as a function of  $\phi$  and  $\psi$  (assuming an all-*trans* backbone;  $\omega = 0$ ). As discussed previously, both  $d$  and  $\theta$  are structurally important:  $d = 0$  codes for backbones that are both flat and optimally extended (for a fixed *theta*), while  $\theta = \pi$  ( $n \in \mathbb{Z}$ ) codes for backbones that are both flat and optimally collapsed or curved (for a fixed  $d$ ); these ideas are further discussed below in context of *trans* backbones.

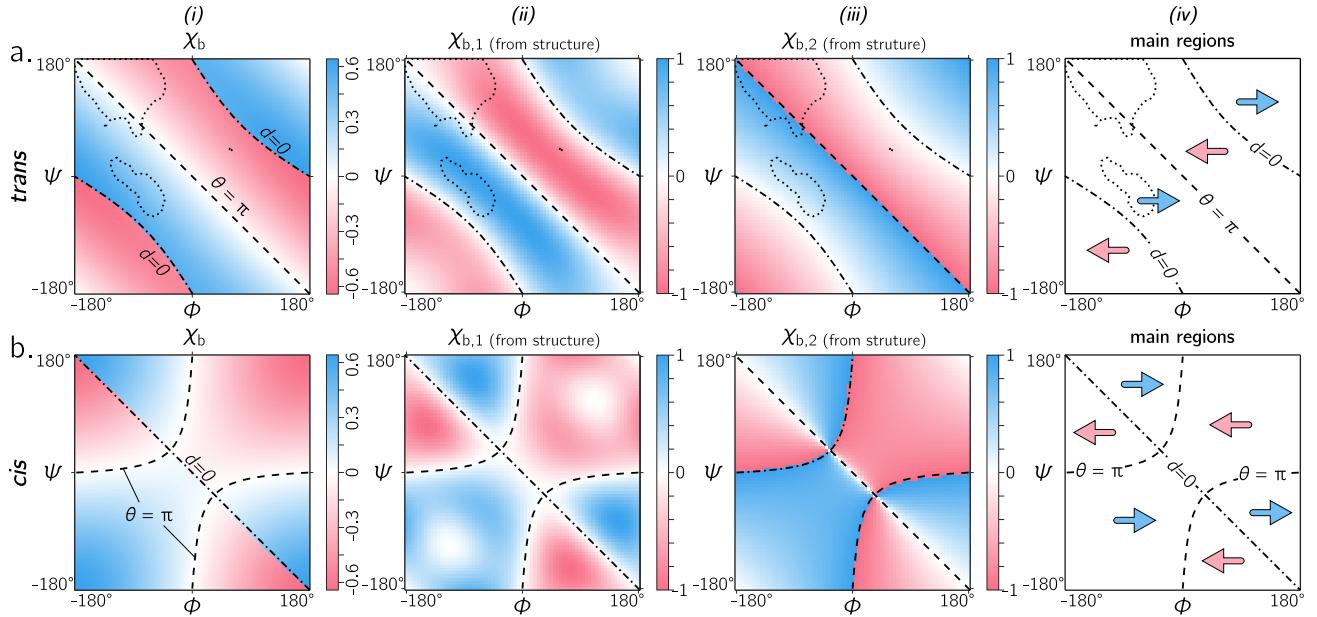
Two possible definitions of backbone ‘flatness’ are possible: flatness at a residue level and flatness at the atomic level. In the former, all atoms of the same *type* are coplanar (examples of atom types are backbone nitrogen, carbonyl carbon,  $\alpha$ -carbon, or even sidechain  $\beta$ -carbon). In the latter, *all* atoms within the backbone are coplanar. We choose the former definition as the relevant definition of flatness, as we are only concerned about the residue-by-residue behavior of the peptide.

Given basic combination guidelines (Fig. 3),  $\theta$  and  $d$  combine into  $\chi$  (Eqn. 7) to define the handedness of a backbone. In absence of handedness, when  $\chi = 0$ , we have an achiral or flat backbone.

### Handedness within *trans* backbones.

Fig. 4b is a complete description of the chirality of an all-*trans* peptide backbone. It describes the behavior of backbone handedness as a measure of  $\phi$  and  $\psi$ , while Fig. 4c describes some structures at relevant regions within the Ramachandran plot. When  $d = 0$  (‘★’), then each residue is at the same ‘altitude’, i.e., the helix is perfectly flat and curved. Note that any path on the Ramachandran plot that transitions

<sup>9</sup>A similar but not identical outcome could be obtained if we use the equation  $\chi' = \frac{d}{|d|} \sin(\theta)$ .



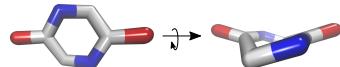
**Figure 5.** Panels (a) and (b) describe the handedness of backbones twists whose amide dihedral angles are *trans* ( $\omega = \pi$ ) and *cis* ( $\omega = 0$ ), respectively. Column (i) describes handedness calculated using  $\chi$  (Eqn. 7), which does not require structures to be computationally generated. Columns (ii) and (iii) respectively show vector-based estimates of backbone handedness –  $\chi_1$  (Eqn. 8) and  $\chi_2$  (Eqn. 9) – which are calculated from computationally generated peptides (see Methods). Regions of left- and right-handedness are identical for all measures of handedness (i–iii). This general map of how handedness is distributed within the Ramachandran plot is shown in (iv). As in Fig. 4, ‘–’ and ‘–·–’ respectively correspond to  $\theta = \pi$  and  $d = 0$ .

from negative to positive  $d$  will encounter an infinitesimal region in its path that is  $d = 0$ : it therefore does not appear white as would be expected for a region where  $\chi = 0$ , but as a sharp transition from one handedness to the other. When  $\theta = \pi$  (where  $n \in \mathbb{Z}$ ), then the backbone is also flat (see ‘★’ in Fig. 4c); however, these backbones are also linear (i.e., each atom type is *colinear*). In short: Within the Ramachandran plot,  $d = 0$  (‘–·–’) and  $\chi = \pi$  (‘–’–) code for *flat* backbones that are respectively either optimally collapsed (at a given  $\theta$ ) or optimally extended (at a given  $d$ ). We later show how these simple rules may be combined to make conjectures about novel secondary and tertiary structures.

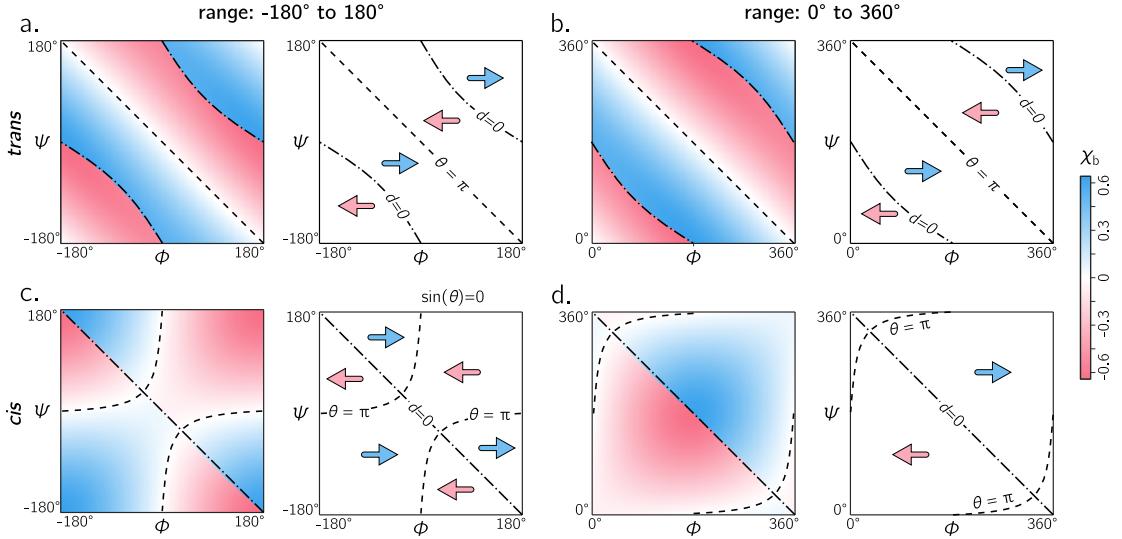
Fig. 5a shows that the equation for  $\chi$  match other metrics for handedness and chirality (Kwiecińska and Cieplak, 2005; Gruziel *et al.*, 2013). In particular, Fig. 5a displays the Ramachandran plot colored by  $\chi$  (i) next to estimates calculated using  $\chi_1$  ((ii); Eqn. 8) and  $\chi_2$  ((iii); Eqn. 9). 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, their exact values differ from the primary metric for handedness ( $\chi$ ) provided here.

### Handedness within *cis* backbones.

In the same vein as Fig. 5a, Fig. 5b displays  $\chi$ ,  $\chi_1$  and  $\chi_2$  as a function of  $\phi$  and  $\psi$  for all-*cis* backbones. To our recollection, this is 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. 5a reiterates the idea that *cis* peptides are a complete different animal, compared to *trans* peptides: the regions of left- and right-handedness are differ. Additionally, points on the *cis* map ( $\phi = \pm 36^\circ$ ,  $\psi = \mp 36^\circ$ ) exist where  $d = 0$  and  $\theta = \pi$ . An example of a six-residue peptide is shown:



This is almost a contradiction, because  $d = 0$  indicates the most curved backbone at a fixed  $\theta$ , and  $\theta = \pi$  indicates the most linear backbone at a fixed  $d$ . Of course, this structure would never be possible, since many atoms overlap. However, such a structure (one with  $d = 0$  and  $\theta = \pi$ ) is not possible in *trans*



**Figure 6.** Each panel describes handedness data (right) and boundaries (left). Both frames of references  $[-\pi, \dots, \pi]$  and  $[0, \dots, 2\pi]$  yield similar trends for *trans* backbones (a,b); however, for *cis* backbones, the latter frame of reference (d) appears to more neatly apportion the handedness of the backbone rather than the traditional frame of reference (c). As in Figs. 4 and 5, ‘–’ and ‘–·–’ respectively correspond to  $\theta = \pi$  and  $d = 0$ .

peptides, even in theory, because the boundaries associated with  $d = 0$  and  $\theta = \pi$  will never intersect (Fig. 5a).

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

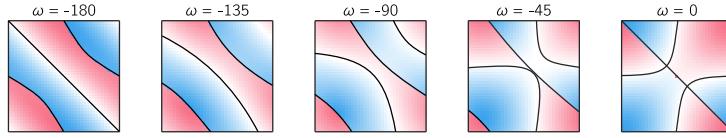
In structural biology,  $\phi$  and  $\psi$  within the Ramachandran plot has been historically set to range between the values  $[-\pi, \pi]$  radians [see, e.g., textbooks by Berg *et al.* (2010) and Alberts *et al.* (2002)]. However, Ramachandran *et al.* (1963) had originally used the range of  $[0,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

Given the periodicity of the Ramachandran plot, these differences do not make any difference scientifically; however the value of the Ramachandran plot is undeniably relative to its boundaries: it is a map of important features of proteins and their relatives in context to the various regions, quadrants, and diagonals in the map. The ramachandran plot’s value lies in being able to convey large amounts of information in easy to read logo- or picto-grams. For that reason, switching the map from one range to another means that two norms and two types of scientists will not be able to converse as seamlessly. Therefore, the following question must arise: which range is able to convey more information with the least amount of work? Fig. 6 shows the behavior of a *trans* backbone (a,b) *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 ( $\sigma = \pi$ ) locally separates right 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 diagonal appears meaningfully separates handedness when the plot ranges from  $0$  to  $2\pi$  (d), while the range  $[-\pi, \pi]$  separates handedness in a more complicated manner (c). For this reason, purely when looking at handedness, and especially in the case of *cis* backbones, the range  $0$  to  $2\pi$  appears to be more meaningful.

## CONCLUSIONS

1. The handedness of both *cis* and *trans* backbones are shown as a function of  $\phi$  and  $\psi$ .
2. Our naïve expectations of handedness is too simple in *trans* backbones and completely wrong in *cis* backbones: Each Ramachandran plot contains at least four regions of left- (red) and right-handed

(blue) regions (not two), irrespective of the state of the amide dihedral angle. While this was tested for all integer values, the following are some examples ( $\phi, \psi \in [-\pi, \dots, \pi]$ ):



3. Interestingly, the reason for the naïve view makes sense when considering only *trans* peptides: the -ve diagonal ('--' in Figs. 4a and 5a) separates *D* and *L* twists if we consider only the regions dominantly occupied by structured proteins ('....' in Figs. 4a and 5a).
4.  $d = 0$  codes for backbones that are flat and optimally extended (for a fixed *theta*).
5.  $\theta = \pi$  ( $n \in \mathbb{Z}$ ) codes for backbones that are flat and optimally collapsed or curved (for a fixed *d*).
6. Ramachandran plots have traditionally been shown in the range  $[-\pi, \pi]$ , however, Fig. 6 indicates that the alternative range  $[0, 2\pi]$  may be more meaningful when considering basic structural properties, especially when considering the *cis* backbone.

## METHODS

All methods and materials required to produce this manuscript are freely available at [https://github.com/ranjanmannige/backbone\\_chirality.git](https://github.com/ranjanmannige/backbone_chirality.git). The manuscript itself may be recreated by running ‘python compileme.py’ within the subdirectory manuscript/ (this command uses PDFLatex; however any other flavor of latex may be used instead).

### Other measures of handedness – calculating handedness from generated structures.

In addition to Eqn. 7, two estimates for handedness were used to validate Eqn. 7 ( $\chi$ ), which are shown for *trans* and *cis* backbones in Fig. 5.

The first metric, previously used as a measure of helix twist handedness (Kwiecińska and Cieplak, 2005), is the following.

$$\chi_1 = \frac{1}{N} \sum_{i=2}^{N-2} \frac{(\mathbf{v}_{i-1} \times \mathbf{v}_i) \cdot \mathbf{v}_{i+1}}{v_{i-1} v_i v_{i+1}}. \quad (8)$$

Here,  $N$  is the peptide length, and the average peptide backbone handedness  $\chi$  has range  $[-1, 1]$ . Values deviating from 0 are more chiral (or ‘twisted’ or ‘handed’), and left handed twists are negative while right handed twists are positive. Only  $\alpha$ -carbon atom positions are used for the calculation. Each  $\alpha$ -carbon belonging to residue  $i \in \{1, 2, \dots, N\}$  has position  $\mathbf{N}_i$ . Vector  $\mathbf{v}_k \equiv \mathbf{N}_{k+1} - \mathbf{N}_k$ . The scalar component of the vector  $\mathbf{v}_i$  is denoted as  $v_i$ . Each scalar in the denominator within the summand (e.g.  $v_i$ ) indicates the distance between neighboring alpha carbons, which is  $\sim 3\text{\AA}$  and  $\sim 3.85\text{\AA}$  for backbones whose amide dihedral angles respectively are *cis* ( $\omega \approx 0$ ) and *trans* ( $\omega \approx \pi$ ). If we know that the backbone amide is either all-*cis* or all-*trans*, then the denominator can be simplified to  $v_i^3$ .

The second measure for handedness, used by Gruziel *et al.* (2013), is

$$\chi_2 = \frac{1}{\pi N} \sum_{i=2}^{N-2} \arctan2(\mathbf{v}_i \mathbf{v}_{i-1} \cdot \mathbf{v}_i \times \mathbf{v}_{i+1}, \mathbf{v}_{i-1} \times \mathbf{v}_i \cdot \mathbf{v}_i \times \mathbf{v}_{i+1}). \quad (9)$$

Here,  $\chi_2$  is known as  $\chi$  in Gruziel *et al.* (2013), absent the normalization by  $\pi$  used here to set the range from  $-1$  to  $1$  in stead of  $-\pi$  to  $\pi$ .

### Generating regular peptides.

While Eqn. 7 is purely analytical and does not need structures to be computationally generated, Eqns. 8 and 9 do. In order to use these equations, peptides (poly-glycines) of arbitrary length were generated using the Python-based PeptideBuilder library (Matthew *et al.*, 2013)<sup>10</sup>. Analysis was performed using BioPython (Cock *et al.*, 2009) and Numerical Python (Van Der Walt *et al.*, 2011). Ramachandran plots that describe chirality (e.g., Fig. 4a) were generated using a grid spacing of  $\phi, \psi \in \{-180, -178, \dots, 178, 180\}$ .

<sup>10</sup>As we are dealing with regular peptides, peptide length must only be greater than 4 residues.

## Protein secondary structure statistics.

Secondary structure statistics were obtained from [Mannige et al. \(2016\)](#). Given that this is previously reported, we have not provided code for this procedure within the Github repository. As per [Mannige et al. \(2016\)](#),  $\alpha$ -helices,  $3_{10}$ -helices and  $\beta$ -sheets were identified using the DSSP algorithm ([Zhao et al., 2005; Kabsch and Sander, 1983; Joosten et al., 2011](#)) and sourced from protein structures within the 40% non-redundant database provided by the Structural Classification of Proteins or SCOP (Release 2.03; [Fox et al. \(2014\)](#)). The polyproline II helix statistics were obtained from segments within 16,535 proteins annotated by PolyprOnline ([Chebrek et al., 2014](#)) to contain three or more residues of the secondary structure.

## ACKNOWLEDGMENTS

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

## REFERENCES

- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002.** Molecular biology of the cell. new york: Garland science; 2002. *Classic textbook now in its 5th Edition* .
- Beke T, Somlai C, Perczel A. 2006.** Toward a rational design of beta-peptide structures. *J Comput Chem* **27**(1):20–38. doi:10.1002/jcc.20299.
- Berg JM, Tymoczko JL, Stryer L. 2010.** *Biochemistry, International Edition*. WH Freeman & Co., New York, 7 edition.
- 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.
- Chebrek R, Leonard S, de Brevern AG, Gelly JC. 2014.** Polypronline: polyproline helix ii and secondary structure assignment database. *Database* **2014**:bau102.
- Chothia C, Levitt M, Richardson D. 1977.** Structure of proteins: packing of alpha-helices and pleated sheets. *Proceedings of the National Academy of Sciences* **74**(10):4130–4134.
- 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.
- Cross L, Klyne W. 2013.** *Rules for the Nomenclature of Organic Chemistry: Section E: Stereochemistry (Recommendations 1974)*. Elsevier.
- Culf AS, Ouellette RJ. 2010.** Solid-phase synthesis of n-substituted glycine oligomers ( $\alpha$ -peptoids) and derivatives. *Molecules* **15**(8):5282–5335.
- 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 Res* **42**(Database issue):D304–D309. doi:10.1093/nar/gkt1240.
- Goodman CM, Choi S, Shandler S, DeGrado WF. 2007.** Foldamers as versatile frameworks for the design and evolution of function. *Nature chemical biology* **3**(5):252–262.
- Gorske BC, Mumford EM, Conry RR. 2016.** Tandem incorporation of enantiomeric residues engenders discrete peptoid structures. *Organic letters* .
- Gruziel M, Dzwolak W, Szymczak P. 2013.** Chirality inversions in self-assembly of fibrillar superstructures: a computational study. *Soft Matter* **9**(33):8005–8013.
- Harris AB, Kamien RD, Lubensky TC. 1999.** Molecular chirality and chiral parameters. *Reviews of Modern Physics* **71**(5):1745.
- Ho BK, Thomas A, Brasseur R. 2003.** Revisiting the ramachandran plot: Hard-sphere repulsion, electrostatics, and h-bonding in the  $\alpha$ -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: CABIOS* **13**(4):425–430.

- Joosten RP, te Beek TAH, Krieger E, Hekkelman ML, Hooft RWW, Schneider R, Sander C, Vriend G.** 2011. A series of pdb related databases for everyday needs. *Nucleic Acids Res* **39**(Database issue):D411–D419. doi:10.1093/nar/gkq1105.
- 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.
- Kandakov A, Ginzburg B.** 2013. Construction of myazawa's equations for the case of an arbitrary number of junction atoms. *Journal of Macromolecular Science, Part B* **52**(8):1056–1063.
- Kwiecińska JI, Cieplak M.** 2005. Chirality and protein folding. *Journal of Physics: Condensed Matter* **17**(18):S1565.
- Laskowski RA.** 2003. Structural quality assurance. *Structural Bioinformatics, Volume 44* pages 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.
- Liddell HG, Scott R, Drisler H.** 1894. *A greek-english lexicon*. Harper & brothers.
- Mannige RV.** 2014. Dynamic new world: Refining our view of protein structure, function and evolution. *Proteomes* **2**(1):128–153.
- 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.
- Mannige RV, Kundu J, Whitelam S.** 2016. The Ramachandran number: an order parameter for protein geometry. *PLoS One* **11**(8):e0160023.
- Matthew Z, Sydykova D, Meyer A, Wilke C.** 2013. Peptidebuilder: A simple python library to generate model peptides. *PeerJ* **1**:e80.
- Mirjanian DT, Mannige RV, Zuckermann RN, Whitelam S.** 2014. Development and use of an atomistic charmm-based forcefield for peptoid simulation. *J Comput Chem* **35**(5):360–370.
- 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.
- Némethy G, Leach S, Scheraga HA.** 1966. The influence of amino acid side chains on the free energy of helix-coil transitions1. *The Journal of Physical Chemistry* **70**(4):998–1004.
- Osipov M, Pickup B, Dunmur D.** 1995. A new twist to molecular chirality: intrinsic chirality indices. *Molecular Physics* **84**(6):1193–1206.
- 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.
- 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.
- 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* **37**(4):205–211.
- Pohl G, Beke-Somfai T, Csizmadia IG, Perczel A.** 2012. Exploiting diverse stereochemistry of  $\beta$ -amino acids: toward a rational design of sheet-forming  $\beta$ -peptide systems. *Amino Acids* **43**(2):735–749. doi:10.1007/s00726-011-1124-7.
- Quiocho FA, Gilliland GL, Phillips G.** 1977. The 2.8-a resolution structure of the l-arabinose-binding protein from escherichia coli. polypeptide chain folding, domain similarity, and probable location of sugar-binding site. *Journal of Biological Chemistry* **252**(14):5142–5149.
- Ramachandran G, Ramakrishnan C, Sasisekharan V.** 1963. Stereochemistry of polypeptide chain configurations. *Journal of molecular biology* **7**(1):95–99.
- 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.
- Schulz G, Elzinga M, Marx F, Schirmer R.** 1974. Three-dimensional structure of adenyl kinase. *Nature* **250**:120–123.
- Shaw PJ, Muirhead H.** 1977. Crystallographic structure analysis of glucose 6-phosphate isomerase at 3· 5 Å resolution. *Journal of molecular biology* **109**(3):475–485.
- Shimanouchi T, Mizushima Si.** 1955. On the helical configuration of a polymer chain. *The Journal of Chemical Physics* **23**(4):707–711.

- Subramanian E.** 2001. Gn ramachandran. *Nature Structural & Molecular Biology* **8**(6):489–491.
- Sun J, Zuckermann RN.** 2013. Peptoid polymers: a highly designable bioinspired material. *ACS Nano* **7**(6):4715–4732. doi:10.1021/nn4015714.
- 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.
- Zacharias J, Knapp EW.** 2013. Geometry motivated alternative view on local protein backbone structures. *Protein Science* **22**(11):1669–1674.
- Zhao Y, Schultz NE, Truhlar D.** 2005. Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *The Journal of chemical physics* **123**(16):161103.
- Zuckermann RN, Kodadek T.** 2009. Peptoids as potential therapeutics. *Curr Opin Mol Ther* **11**(3):299–307.