Chirality and the Ramachandran plot

Ranjan V. Mannige[1,*] **and Other Names TBD**[1]

[1] **Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, 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 has been used since the 1960's to describe the nature in which the backbone twists into regular structures(Ramachandran *et al.*, 1963). While the regions within the Ramachandran plot that are well-populated by proteins is well known, new molecules are being designed today that are not bound to the traditional regions of the Ramachandran plot. This has sparked new interest in the basic behavior of a backbone within all regions of the Ramachandran plot (not just those available to a canonical protein). In these lines, this short paper describes: 1) a complete characterization of the way a backbone twists in all regions of the Ramachandran plot – this would serve as a reference point for understanding new types of peptide and peptidomimetic structures – and 2) a succinct set of Python scripts that show how these types of studies are accessible to undergraduate students with basic computational and biological training.

INTRODUCTION

The Ramachandran plot (Ramachandran *et al.*, 1963) is a two-dimensional map that describes the per-residue conformation of a peptide backbone (Berg *et al.*, 2010; Alberts *et al.*, 2002). The Ramachandran plot is plotted as a function of a peptide residue's dihedral angles ϕ and ψ (Fig. 1a). Each point (ϕ, ψ) represents a 'twist' of a peptide backbone in three-dimensional space. Any peptide built with uniform twist or backbone paramters (say, $\phi = X^\circ$ and $\psi = Y^\circ$) will result in a regular structure; some regular structures are thermodynamically stable 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). As such, Ramachandran plots have been useful for understanding a peptide backbone's general conformational state or 'twistedness' at a glance (Berg *et al.*, 2010; Alberts *et al.*, 2002; Subramanian, 2001; Laskowski *et al.*, 1993; Hooft *et al.*, 1997; Laskowski, 2003). While the idea of a curved peptide backbone appears to be the domain of a mathematical puzzle or diversion, in practice, the curve of the peptide backbone completely defines the general structure of a protein: proteins are peptides whose backbones occupy specific conformations[1]. This is important because, in the molecular world, the conformations available to a protein (or any molecule) plays a large part in defining the possible functionalities available to that molecule (Berg *et al.*, 2010; Alberts *et al.*, 2002).

So far, our understanding the Ramachandran plot been limited to the secondary structures and loops posed by proteins (Berman *et al.*, 2000). For example, structural biologists are aware that the negatively sloping diagonal (dashed line in Fig. 1b; henseforth denoted as the '-ve diagonal') demarcates a change in backbone chirality. For example, the position of the idealized left- and right-handed α-helices (Fig. 1c) – respectively denoted as α_L and α in Fig. 1b – are on opposite sides of the the -ve diagonal[2]. Additionally, the the β-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)[3]]. Indeed, all regular/ordered motifs within proteins that lie to the right (or top) of the -ve diagonal

[1] These conformations can be specific folds in globular proteins or larger structural ensembles in intrinsically disordered proteins; Mannige (2014).

[2] 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).

[3] Note that a different metric for handedness in β sheets exists today that addresses the handedness of a particular hydrogen

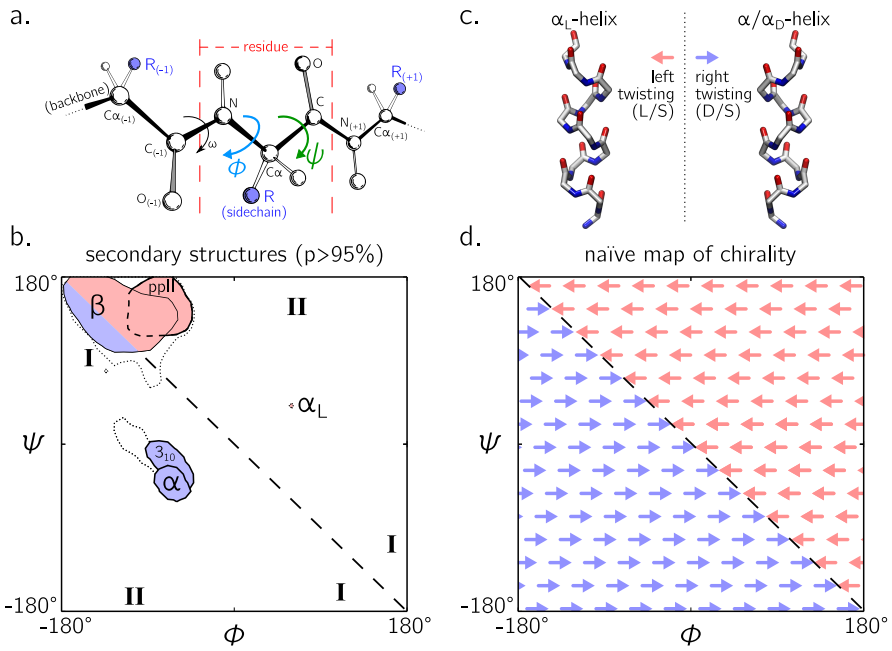

Figure 1. The general structures available to peptides and proteins are characterized the way in which the peptide backbone twists in three-dimensional space.

are left-handed in backbone twist; as a corolary, all ordered backbones to the left (or under) the -ve diagonal are right handed in nature (primarily, the α-, π- and 3_{10}-helix; Fig. 1b; π-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 was cognizant of this general trend in peptide secondary structure [see, e.g., his notes on what is now known as the α-sheet; Pauling *et al.* (1951); Pauling and Corey (1951b,a)], the picture in Fig. 1d has so far remained an untested hypothesis.

While chirality or 'twist' in regular protein secondary structures is well understood, peptide mimics – especially pep*toids* – 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 (Fig. 4). Certain peptides display secondary structures in regions of 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 Σ-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 'ω-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 σ 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 test the hypothesis in Fig. 1d. That is, we ask below: which way does the peptide/peptoid backbone twist at any point on the Ramachandran plot? This question, while simple, would complete our understanding of how twists combine to form structures.

bonding network and not the handedness of the participating peptide's backbone. See, e.g., Schulz *et al.* (1974) and Chothia *et al.* (1977).

METHODS

In order to satisfy historical notation (Berg *et al.*, 2010; Alberts *et al.*, 2002; Laskowski *et al.*, 1993; Laskowski, 2003), we assume ϕ and ψ range between -180° and 180°[4].

Backbone twist handedness. To explore the nature of backbone handedness, we use a metric that was previously used as a measure of helix twist handedness (Kwiecińska and Cieplak, 2005)[5]:

$$\chi_{\mathrm{b}} = \frac{1}{N}\sum_{i=2}^{N-2}\frac{(\boldsymbol{v}_{i-1}\times\boldsymbol{v}_i)\cdot\boldsymbol{v}_{i+1}}{v_{i-1}v_iv_{i+1}}. \tag{1}$$

Here, N is the peptide length, and the average peptide backbone handedness χ_{b} 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 α-carbon atom positions are used for the calculation. Each α-carbon belonging to residue $i \in \{1,2,\ldots N\}$ has position $\boldsymbol{N}_i$. Vector $\boldsymbol{v}_k \equiv \boldsymbol{N}_{k+1} - \boldsymbol{N}_k$. The scalar component of the vector $\boldsymbol{v}_i$ 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$3Å and $\sim$3.85Å for backbones whose amide dihedral angles respectively are *cis* ($\omega \approx 0^\circ$) and *trans* ($\omega \approx 180^\circ$). If we know that the backbone amide is either all-*cis* or all-*trans*, then the demnominator can be simplified to v_i^3.

While we use this metric as the primary measure of handedness, we also use the following two metrics for reference:

$$\chi_{\mathrm{b},2} = \frac{1}{\pi N}\sum_{i=2}^{N-2}\arctan2\left(v_i\boldsymbol{v}_{i-1}\cdot\boldsymbol{v}_i\times\boldsymbol{v}_{i+1}, \boldsymbol{v}_{i-1}\times\boldsymbol{v}_i\cdot\boldsymbol{v}_i\times\boldsymbol{v}_{i+1}\right). \tag{2}$$

$\chi_{\mathrm{b},2}$ is known as χ in Gruziel *et al.* (2013), absent the normalization by π used here to set the range from -1 to 1 in stead of $-\pi$ to π.

Generating regular peptides Peptides (poly-glycines; $N = 5$) were generated using the Python-based PeptideBuilder library (Matthew *et al.*, 2013). Analysis was performed using BioPython (Cock *et al.*, 2009) and Numerical Python (Dubois *et al.*, 1996). Ramachandran plots that describe chirality (e.g., Fig. 2a) were generated using a grid spacing of $\phi, \psi \in \{-180, -178, \ldots, 178, 180\}$.

Protein secondary structure statistics. α-helices, 3_{10}-helices and β-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.

RESULTS

As depicted in Fig. 1a, the backbone each residue i of a peptide possesses two main degrees of freedom: the dihedral angles ϕ_i, and ψ_i. An additional degree of freedom exists, ω_i, which in peptides predominantly occupies values of $\equiv 180^\circ$ (*trans*), and infrequently occupies a lower-probability (higher-energy) value of $\equiv 0^\circ$ (*cis*). Fig. 2 discusses the behavior of an all-*trans* backbone, while Fig. 3 describes the behavior of both all-*trans* and all-*cis* backbones.

Fig. 2a describes the twist or chirality (χ_{b}; Eqn. 1) of a regularly arranged backbone as a function of (ϕ, ψ) (additionally, an example of the behavior of one 'slice' of Fig. 2a is shown in Fig. 2b). Fig. 3(i) and Fig. 3(ii) extends on Fig. 2 and describes the handedness of both *trans* and *cis* backbones in using two metrics (χ_{b},$\chi_{\mathrm{b},2}$). In each graph, negative (red) and positive (blue) values of handedness respectively indicate left- and right-handed backbone twists. The color white indicates that all α-carbon atoms in the conformation are coplanar.

Figs. 2 and 3 assert the following: 1) the naïve view of chirality (Fig. 1d) is inaccurate for both the *trans* (Figs. 2a and 3*a(i,ii)*) and *cis* backbones (Fig. 3*b(i,ii)*). For the *trans* backbone, while the

[4]An angle ζ can be wrapped within the range $[-180, 180)$ using $\zeta' = -180 + (\zeta + 180)\%(360)$, where % represents the modulus function.

[5]With the exception of certain regions within the Ramachandran plot (e.g., $\phi = \psi = 180^\circ$), all regularly arranged backbones are helices. Indeed, the word 'helix' has a broad meaning that comes from the greek word ἕλικ ξ that means 'twisted, curved'; Liddell *et al.* (1894).

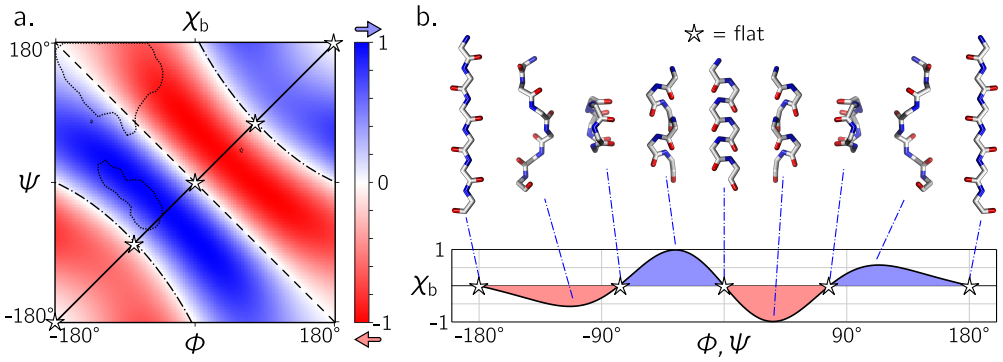

Figure 2. The chirality of an ordered peptide within the Ramachandran plot. Each point in (a) represents the chirality or handedness – χ_b (Eqn. 1) – of a peptide backbone with uniform (ϕ, ψ) diehdral angles. Panel (a) shows that the naïve expectation of handedness in a Ramachandran plot (Fig. 1d) is too simplistic. 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 (a) is shown in (b). The graph in Fig. 2b represents the handedness of peptides (y-axis) whose backbones are built by setting $\phi_i = \psi_i = A$ ($A \in [-180°, 180°]$). The graph is accompanied by snashots of peptides sampled at regions of the slice where χ_b is either 0 or at a local maximum or minimum. As expected, peptides whose carbon atoms are co-planar result in $\chi_b = 0$.

-ve diagonal ('– –') does separate right(D)- and left(L)-handed backbone twists, it does not *exclusively* partition regions of oppossite handedness: two additional borders exist that also serve as L-D demarcations ('– · –' in Fig. 2). For the *cis* backbone, the -ve diagonal does not even split the neighborhood into two parts: the -ve diagonal is adjacent to four regions of distinct handedness.

In each Ramachandran plot within Fig. 3, both metrics for handedness – χ_b *(i)* and $\chi_{b,2}$*(ii)* – describe four regions of handedness, which are distinctly shaded in *(iii)*.

Interestinglyl, the reason for the naïve view makes senses when considering only peptides: the -ve diagonal ('– –') separates D and L twists if we consider only the regions occupied by structured proteins

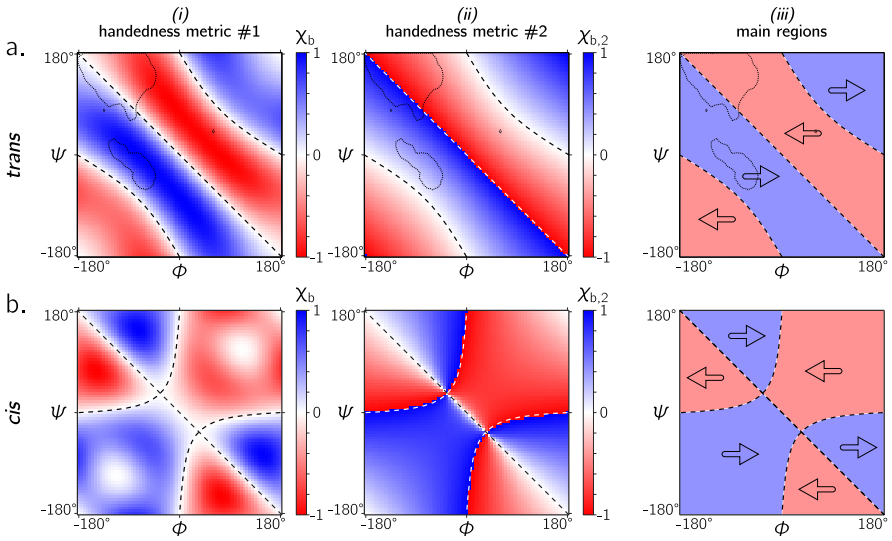

Figure 3. Panels (a) and (b) describe behavior of twists for backbones that are *trans* ($\omega = 180°$) and *cis* ($\omega = 0°$), respectively. Columns ***(i)*** and ***(ii)*** show, respectively, the behavior of the backbone as a measure of two metrics for twist handedness: χ_b (Eqn. 1) and $\chi_{b,2}$ (Eqn. 2) . Both plots show qualitatively identical apportionments of the Ramachandran plot into left and right handed twists. This general map is shown in column ***(iii)***.

(‘……’ in Figs. 2a and 3a).

Additionally, the -ve diagonal does not behave as mirror symmetry element[6]. In stead, the point $(0,0)$ serves as a two-fold point-group symmetry element with

of one point on the ramachandran plot along the

While this distribution of handedness within the Ramachandran plot deviates from the naïve view (Fig. 1d), the -ve diagonal nonetheless plays an important role: two points that are related by a reflection along the -ve diagonal also have their corresponding backbone structures related by mirror symmetry.

regions are more

In any frame of reference

While even Pauling was cognizant of this general trend in peptide secondary structure [see, e.g., his notes on what is now known as the α-sheet; Pauling *et al.* (1951); Pauling and Corey (1951b,a)], the picture in Fig. 1d has so far remained an untested hypothesis.

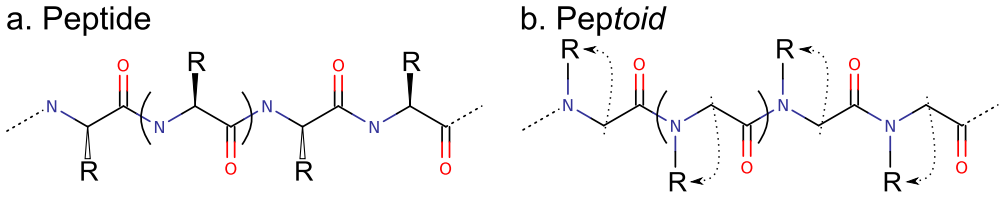

Figure 4. Peptoids and peptides are distinguished by the position of their sidechain (R) groups. They therefore behave differently. For example, the peptoid backbone is more flexible than the protein one, and peptoid backbones lack canonical hydrogen bond donors, which are crucial to the formation of traditional secondary structures in proteins such as α-helices β-sheet (Berg *et al.*, 2010).

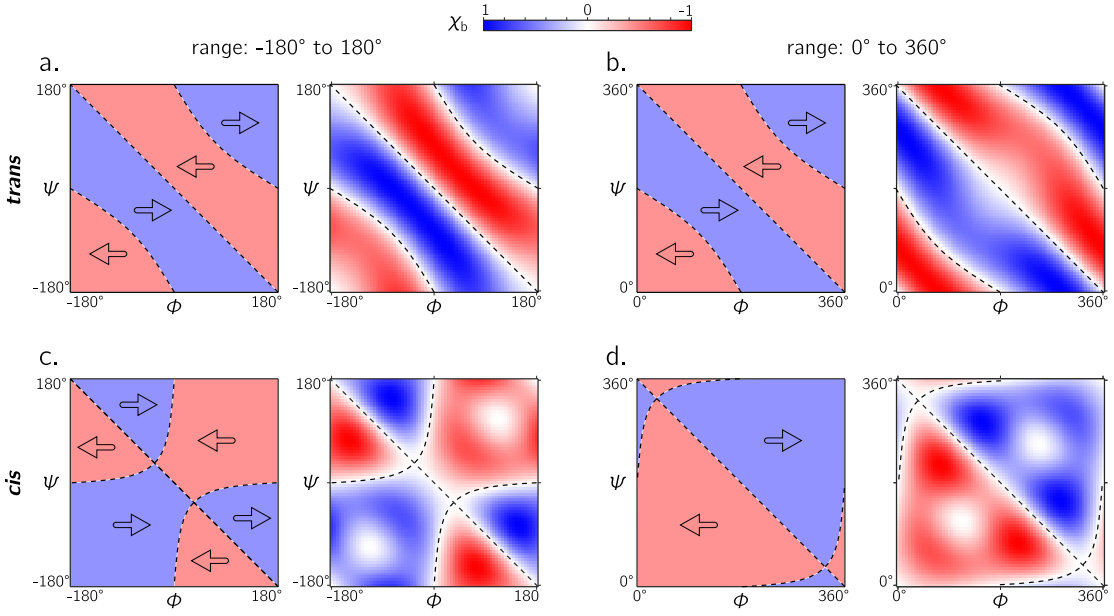

Figure 5. Each panel describes handedness data (right) and boundaries (left). While both frames of references $[-180°,\ldots,180°]$ (a,c) and $[0°,\ldots,360°]$ (b,d) yield similar trends for *trans* backbones (a,b); however, for *cis* backbones, the latter frame of reference (d) appears to more neatly apportion the behavior of backbone than the traditional frame of reference (c).

Pauling and Corey (1951b); Pauling *et al.* (1951) (Mannige *et al.*, 2016) ppI helix (Adzhubei and Sternberg, 1993)

Backbones of natural amino acids with ‘L’ chirality dominantly occupy the “left side” of the plot ($\phi < 0$; Fig. **??**a, top), a consequence of the chirality of the backbone α-carbon (Berg *et al.*, 2010; Cintas, 2002) and steric hindrance between the backbone carbonyl and sidechain atoms (Branden *et al.* (1999)).

[6]I.e., points on the Ramachandran plot that are related by a relection across the -ve diagonal do not code for backbones of opposing symmetry.

Interestingly, switching the C_α's chirality for every residue results in structures that exist on the "other side" of the plot, which a majority of its density at $\phi > 0$; (Zawadzke and Berg (1993); Hung *et al.* (1998)).

ACKNOWLEDGMENTS

RVM was supported by the Defense Threat Reduction Agency under contract no. IACRO-B0845281. RVM thanks Alana Canfield Mannige for her critique. 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

Adzhubei AA, Sternberg MJ. **1993**. Left-handed polyproline ii helices commonly occur in globular proteins. *Journal of molecular biology* **229(2)**:472–493.

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.

Branden CI, *et al.* **1999**. *Introduction to protein structure*. Garland Science.

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.

Cintas P. **2002**. Chirality of living systems: a helping hand from crystals and oligopeptides. *Angewandte Chemie International Edition* **41(7)**:1139–1145.

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 (α-peptoids) and derivatives. *Molecules* **15(8)**:5282–5335.

Dubois PF, Hinsen K, Hugunin J. **1996**. Numerical python. *Computers in Physics* **10(3)**.

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.

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.

Hung LW, Kohmura M, Ariyoshi Y, Kim SH. **1998**. Structure of an enantiomeric protein, d-monellin at 1.8 å resolution. *Acta Crystallographica Section D: Biological Crystallography* **54(4)**:494–500.

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.

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.

Pauling L, Corey RB. **1951**a. 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. **1951**b. 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 β-amino acids: toward a rational design of sheet-forming β-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.

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.

Zawadzke LE, Berg JM. **1993**. The structure of a centrosymmetric protein crystal. *Proteins: Structure, Function, and Bioinformatics* **16(3)**:301–305.

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.