PROTEINS, accepted (2007)

A chirality index for investigating protein secondary structures and their time evolution

Adriana Pietropaolo, Luca Muccioli, Roberto Berardi and Claudio Zannoni *

Abstract

We propose a methodology for the description of the secondary structure of proteins, based on assigning a chirality parameter to short aminoacid sequences according to their arrangement in space at a certain time. We validated the method on ideal and crystalline structures, showing that it can assign secondary structures and that this assignment is robust with respect to random conformational perturbations. From the values of the index and its pattern along a sequence it is possible to recognize many structural motifs of a protein, and in particular poly-L-proline II left-handed helices, often not

*E-mail: Claudio.Zannoni@unibo.it Dipartimento di Chimica Fisica ed Inorganica and INSTM, Università di Bologna, V.le Risorgimento, 4, I-40136 Bologna, Italy.

detected by secondary structure assignment algorithms.

Assigning an instantaneous chirality index to the fragments also allows the dynamics to be studied. With this purpose, molecular dynamics simulations were carried out in water for selected Hemoglobin (110 ns) and Immunoglobulin antigen fragments (50 ns), showing the capability of the chiral index in identifying the stable secondary structure elements, as well as in following their time evolution and conformational changes during the trajectory.

1 Introduction

The analysis and assignment of the secondary structure of proteins is a central problem in biophysics and algorithms designed for this purpose are indispensable tools not only for the assignment and classification of newly derived native protein structures but also for all the computational techniques that aim at structural predictions on the basis of the primary sequences, multiple sequence alignment and related statistical studies of local properties like solvent accessibility and native contacts. The first key contribution in the field was probably that of Ramachandran [1,2], correlating the native distribution of the -N–C_α- and -C_α–C- dihedral angles (ϕ, ψ) of constituent aminoacids in a given sequence to the protein secondary structure. The Ramachandran map for a sequence of aminoacids is indeed very helpful in the individua-

tion of the existent secondary elements, but it may fail in the case of high conformational flexibility, which leads to non standard backbone angles, in particular for peptides [3, 4]. To date, the most commonly used and authoritative secondary structure determination program is the "Dictionary of Protein Secondary Structures" (DSSP) [5]. The DSSP relies on an algorithm based on hydrogen bond patterns involving the C=O and N-H backbone atoms, neglecting the ϕ and ψ dihedral angles and classifying qualitatively the structure in eight classes. Despite its effectiveness, this choice does not help in the detection of small deviations of the backbone dihedral angles from the ideal structure, which may be important for the biological function of a protein. Furthermore, the DSSP analysis is known to be error–prone in the exact detection of the edges of a given motif [6]. An improvement which tries to address some of these limitations, such as the absence of description of thermal fluctuations present in experimental structures, is DSSPcont [7], which performs a continuous assignment of secondary structure by calculating weighted averages with different hydrogen bond thresholds.

More effectively, STRIDE [8] considers, in classifying secondary structures, both hydrogen bond patterns and backbone dihedrals. Many other variants and different criteria, like alpha carbon distances and angles, have been proposed over the years (see e.g. [9–12]), all achieving a high global agreement among them and with the PDB classification (higher than 80%).

Notwithstanding this success, further improvement would be important in the case of non-standard conformations strongly departing from ideal backbones (the "twilight zone" [9]), like the ones obtained by NMR experiments, and in particular polyproline II structures [13, 14].

Despite the intrinsic chiral nature of aminoacids and of many motifs, to our knowledge quantitative measurements of chirality have never been proposed as criteria in the field of protein structure analysis, with the notable exception of a variant of Gō–like folding models [15]. In the present study we wish to fill this gap, suggesting the use of a local chirality index that varies continuously as the conformation changes and that aims to provide a quantitative answer to the question "how chiral is a molecule?" [16]. Such an index must be invariant under similarity transformation, change sign upon reflection and be null for symmetric objects [17]. In particular chiral indices which are derived from the disposition in space of the atoms of a given molecule [18–21] have proved to be useful in the attempt of relating molecular structure with macroscopic properties, such as helical pitch [22], helical twisting power [23, 24] and facial diastereoselectivity [25]. In the following, we adopt the scaled chiral index of Solymosi et al. [21] for the analysis of the conformation of ideal backbones and real proteins, showing that local symmetry measurements can actually give reliable information of protein secondary structure.

2 Chirality calculation on ideal structures

A simple indicator of the conformational chirality of a molecule can be written down as a pseudoscalar combination of three molecule fixed vectors, analogously to the calculation of a dihedral angle. The idea of calculating molecular chirality from atomic coordinates is akin to a generalization of simple models of optical activity, in which a third rank tensor based on dipolar interaction products appears. In that case (see ref. [18] for details) the vectors are related to electronic transitions, but defining the tensor as a purely geometrical entity and reducing it to second-rank on the basis of symmetry arguments, Osipov et al. [18] identified in its trace a pseudoscalar quantity useful for determining molecular handedness. A scaled version of this index was subsequently introduced to facilitate the comparison between molecules of different size [21], leading to the following expression:

$$G = \frac{4!}{3N^4} \sum_{\substack{\text{all permutations of} \\ i,j,k,l=1\ldots N}} w_i w_j w_k w_l \frac{[(\mathbf{r}_{ij} \times \mathbf{r}_{kl}) \cdot \mathbf{r}_{il}](\mathbf{r}_{ij} \cdot \mathbf{r}_{jk})(\mathbf{r}_{jk} \cdot \mathbf{r}_{kl})}{(r_{ij} r_{jk} r_{kl})^n \; r_{il}^m}, \qquad (1)$$

where i, j, k, l are four of the N atoms belonging to the molecule, $\mathbf{r}_{ab}$ are interatomic distance vectors, w_i, w_j, w_k, w_l are suitably chosen weights for each atom, and n and m are arbitrary integers. This index is commonly employed in a dilatation–invariant form with $n = 2$ and $m = 1$, [21, 23–25], while the weights are set to unity (dimensionless form) or to atomic masses

(recalling the Cahn–Ingold–Prelog rules).

To apply this index to the analysis of protein secondary structures, some adjustments are necessary. First, since the structural motifs represent a local property of a small group of aminoacids, it is not very meaningful to consider in the calculation the chirality between all possible sets, getting a single value for the whole protein as in equation 1. Thus, we decided to focus only on backbone atoms (N, C_α and C) and to calculate the chiral index for sequences of connected atoms of length N_a. Secondly, we introduced in eq. 2 a cutoff radius to avoid the computation of unnecessary long-range terms, that give a negligible contribution to the overall chirality.

$$G^{a,N_a} = \frac{4!}{3N_a^4} \sum_{\substack{\text{all permutations} \\ \text{of } i,j,k,l}} \begin{cases} \frac{[(\mathbf{r}_{ij} \times \mathbf{r}_{kl}) \cdot \mathbf{r}_{il}](\mathbf{r}_{ij} \cdot \mathbf{r}_{jk})(\mathbf{r}_{jk} \cdot \mathbf{r}_{kl})}{(r_{ij} r_{jk} r_{kl})^2 r_{il}} & \begin{array}{l}\text{if } r_{ij}, r_{kl}, r_{il}, r_{jk} < r_c, \text{ and} \\ a \le i,j,k,l \le N_a + a - 1\end{array} \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

Considering the secondary structure a local geometry–dependent property of a small number of connected aminoacids, we studied the variation of the average $G^{N_a} = 1/(3N_R - N_a + 1) \sum_{a=1}^{3N_R - N_a + 1} G^{a,N_a}$, as function of the number of backbone atoms N_a, choosing the values of this parameter and of the cutoff distance r_c that maximize the local sensitivity for ideal backbones composed of 40 residues. For this purpose, the cutoff distance was increased until the stability of G^{N_a} values was achieved, as we obtained for r_c greater

than 10 Å. In practice we have chosen a cutoff of 12 Å, that is appropriate for an extension of the analysis to side chain atoms and should comprise all the possible aminoacid native contacts [26]. The value of N_a which allows the best differentiation of the secondary structures is 15, corresponding to five consecutive residues.

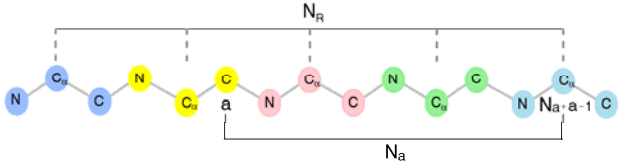


Figure 1: A backbone composed of 15 atoms (N_R=5 residues). For N_a=9, the sequence of atoms contributing to the calculation of G^{a,N_a} is indicated, starting from atom $a = 6$.

In building ideal secondary structures, we took into account binary or quaternary periodicity on the backbone angles (ϕ,ψ): ($-67°$, $-41°$) for α helix [27], ($-49°$, $-26°$) for 3_{10} helix [28], ($-67°$, $-59°$) for π helix [27], ($-60°$, $-30°$, $-90°$, $0°$) for type I β turns [29], ($-75°$, $147°$) for PPII helix [30] and ($-130°$, $130°$) for sheets regions [31], while the ω angles were kept fixed to the *trans* value of $180°$. Although Type I β turn conformation is not periodic in proteins, involving generally only 4 consecutive residues, we considered it periodic for ease of comparison with the other motifs.

In Figure 2 we report the behavior of the G index for ideal structures along the backbone, calculated with $N_a = 15$ and $r_c = 12$ Å. Different patterns are clearly distinguishable: in particular, the right handed α helix, type I β turn and 3_{10} helix possess negative chiral index values, which exhibit the correct periodicity when moving along the backbone. Furthermore, the left handed helix of poly-L-proline II shows a positive sign of chiral index, in accord with its opposite handedness with respect to the other helices. The β sheets structure, having a flat shape and symmetric ϕ and ψ dihedrals, shows a chiral index close to zero, as well as the π helix, which shows negative values approaching zero (ϕ and ψ respectively $-67°$, $-59°$). In summary we see that the various important motifs can be all assigned and differentiated on the basis of their intrinsic chirality.

3 Chirality of crystalline protein structures

After this preliminary study, we analyzed a set of seven real protein structures, collected from the Protein DataBank, containing the most important structural motifs. We chose the chain A of hemoglobin (pdb code 2MHB), as a globin representative α protein, and again for helix structures, we analyzed the avian prion globular domain (pdb code 1U3M), which contains three α helices [32] and ubiquitin (pdb code 1D3Z), with one α helix. Concerning

turn and sheet regions, we studied respectively the chain A of immunoglobulin antigen (pdb code 1REI), previously included in the DSSP data set [5], and serine protease, a turn rich protein (pdb code 1HPJ). Model peptide systems for 3_{10} helix (pdb code 1LB0) and poly-L-proline (pdb code 1JMQ 51-60) were also taken from the protein databank.

The helix of hemoglobin (5-18) and the helices 2 and 3 of avian prion protein (Figure 3), show the G pattern typical of ideal α helices, while for avian prion protein helix 1, G values reveal imperfections in the helix backbone, as also suggested by secondary structure prediction algorithms [33]. In fact, the index shows irregularities in the first few residues, assuming the values typical of an α helix only after the sixth (Figure 3). The abundance of α helices can also be visually noticed looking at the G values along the backbone of hemoglobin (Figure 4 [a]), with the motifs helix-turn-helix and a high positive peak due to the presence of residues with ϕ and ψ values typical of poly-L-proline II, in the region after residue 90. Like hemoglobin, also for avian prion globular domain (Figure 4 [b]), it is easy to distinguish the different secondary structures along the backbone, like the three helices followed by turns, the positive peaks around residues 140 and 175 and after residue 200, due to at least one residue adopting poly-L-proline conformation. The zero G values suggest the presence of β sheets, quite evident in the plateau region centered at residue 169. Ubiquitin (Figure 4 [c]) has only one

helix, and in fact only one region with negative periodic fluctuations of G is present, while at least four β-sheets can be identified; serine protease (Figure 4 [d]), possesses a high number of turn regions, detectable from the sudden alternation of negative and positive peaks, which are instead only negative for 3_{10} helices (cf Figure 3), being constituted by at least three residues.

Concerning the β sheet-containing peptides (Figure 5 [a-b]), the analysis appears to be more difficult, because these structures occur in proteins with parallel or antiparallel regions formed by groups of residues far away in the protein sequence. Consequently further investigations, like hydrogen bonds screening, should be carried out in these cases to match the sequences. However, the plateaus at zero values of G generally help in identifying such structures.

As previously said, the chiral index is very sensitive to poly-L-proline dihedrals: a positive peak underlines in fact that at least one aminoacid with PPII structural motif is present in a given protein region. Concerning the G of the model poly-L-proline peptide (fragment 51-60 of 1JMQ) reported in Figure 5 [b], a good overlap between the PPII ideal structure and the PPII model peptide results in the 3-5 region. After residue 5 the G values of 1JMQ drop as they take into account residues 7 and 8 which are not in PPII conformation. A full detection of PPII structure using DSSP-like algorithms is hampered because prolines do not form hydrogen bonds and

although adopted also by other aminoacids, its extended conformation (9.3 $\AA$ pitch) does not allow a well definite hydrogen bond pattern; therefore PPII regions are usually misclassified as loop or coils [14]. The sensitivity of G to PPII chirality seems important for a better identification of this class of structures. For a visual comparison with DSSP classification, we also report in Figures 4 and 5 the DSSP sequence assignment: we see that in all cases the qualitative agreement between the two indices is good, confirming the ability of the G index for discriminating secondary structures of real proteins.

To summarize the relation between chirality values and secondary structure, in Figure 6 we report the cumulative G distributions among all the structures analyzed in this section. The histogram shows clearly four maxima, corresponding to Type I β turn/3_{10} helix, α helix, β sheets and PPII respectively, which all present G values close to the ones of the ideal structures (blue dots in Figure 6), and reveals the approximate content of these motifs in the data set. Even if necessarily limited by our particular choice of proteins and peptides, this finding suggests the possible use of such distributions for a quick similarity check between two protein structures or data sets.

4 Stability of the chirality index

In the last section we have investigated the behavior of the chiral index for single structures of ideal and real proteins. Here, instead, we want to test the performance of the method in response to random thermal fluctuations and to conformational changes. This is important as in analyzing real structures, particularly in solution, fluctuations are unavoidable. To mimic this condition, we have studied the effects of a gaussian noise in the backbone angles values, altering the secondary structure periodicity. To this end, we have randomly built 2000 configurations for each type of secondary structure, extracting their ϕ and ψ from a gaussian distribution centered on the ideal ϕ and ψ values (see section 2) and from these, we have evaluated the overall average standard deviation of the chiral index as a function of the gaussian half-height amplitude η (Table I). We report in Table I such standard deviations with the α helix and poly-L-proline II structures showing the highest ones, thus the chirality of these two structures is more sensitive to backbone variations with respect to others. In general, the index does not seem overly sensitive to a random perturbation of the dihedral angles and thus appears to be sufficiently robust to follow the fluctuations of the protein structure during a computer simulation, without being disrupted by thermal noise.

As a final inspection of the applicability of the chirality index analysis we

wish to test secondary structure assignment in the more realistic situation of a protein in water, where the geometry fluctuations or possibly, the conformation changes, are also caused by the interaction with the solvent at certain thermodynamic conditions. Thus we have performed two rather long molecular dynamics runs [34]: a 110 ns simulation of a fragment of hemoglobin, and a 50 ns simulation of a fragment of immunoglobulin antigen, in which a sheet-turn-sheet motif is present. Both simulations were run in water using ORAC 4.0 code [35] and the Amber94 force field (FF) [36]. Cubic boxes containing the protein chain and 484 water molecules for hemoglobin, and 1359 for the immunoglobulin fragment, were used with periodic boundaries and isothermal-isobaric conditions [37] (P=1 atm, T=300 K). Temperature was controlled using a Nosé-Hoover thermostat [38, 39] and the SPC model [40] was used for water. An r-RESPA multiple time-step algorithm with a potential subdivision specifically tuned for proteins [41] was used for integrating the equations of motion, using an overall time step equal to 10 fs.

As we can see from Figure 7 [a], the negative periodic pattern of the chiral index is retained during the simulation of hemoglobin, reflecting the fact that its helical structure is not disrupted. Indeed the time behavior during the simulations of G for selected residues (Figure 8 [a,d]), shows that the chiral index is stable in an ensemble of configurations fluctuating around the same secondary structures, i.e. that folding/unfolding does not happen. More

interestingly, in the case that major conformational changes occur, as for the 1REI immunoglobulin antigen 1-30 fragment (Figure 7 [b]), the chiral index gives precious indications about the different conformational states that the fragment explores, detectable from the different values adopted by G during the time evolution (Figure 9 [a,d]). The comparison with the instantaneous DSSP classification in Figures 8 and 9 confirms the qualitative agreement between the two indices and the greater capabiliy of G in quantifying even small structural changes in time.

4.1 Chirality index dynamics and folding

Having established the link between chirality index and motif of a certain fragment, it is important to make full use of the fact that, differently from DSSP, the chiral index is a continuous dynamical quantity that can be employed to assess average structural changes during the simulation rather than just visually examine them along an individual trajectory. To this end we introduce a time correlation function between the chiral index of two fragment a, b, expressed as follows:

$$\chi^{a,b}(t) = \langle G^a(0) G^b(t) \rangle, \tag{3}$$

and a normalized version:

$$\chi_N^{a,b}(t) = \frac{\langle G^a(0) G^b(t) \rangle}{\langle G^a(0) G^b(0) \rangle}. \tag{4}$$

The normalized correlation has the advantage of bringing all the various fragment correlation in the same range, facilitating the comparison of time evolutions: values that remain close to one and slowly decaying indicate strong correlation, while functions reaching rapidly zero are proof of fast, uncorrelated motions. However the initial value is of course important, as it allows to distinguish the type of secondary structure. We used these equations for the calculation of auto–correlation functions, namely with $a = b$ in equations 3 and 4, reported in Figures 10 and 11 for hemoglobin and immunoglobulin respectively. Cross–correlation functions $\chi^{a,b}(t)$ and $\chi_1^{a,b}(t)$, were also calculated for selected residues, and their time behavior is shown in Figures 12 [a], [b] and 13 [a], [b]. In particular, examining the helix of hemoglobin we find that both the auto and cross–correlations functions have high values in the internal core of the helix structure while in the N-terminal domain the memory of the initial configuration is rapidly lost (Figures 10 [a], [b]) and 12 [a], [b]. This can be noticed from the asymptotic trend towards 1 of the functions $\chi_N^{22,22}$, $\chi_N^{37,37}$ and $\chi_N^{22,37}$, centered on residues 10 and 15 respectively, which correspond to the internal core (Figure 10 [a], [b] and 12 [a], [b]). More interestingly, a transition between α and 3_{10} helix is also observed in residue 5, corresponding to $\chi_N^{7,7}$. This is shown both by the decrease of the auto–correlation functions (Figures 10 [a], [b]) and by the variation of the index during the time, which exhibits evidently the transition approximately

after 40 ns for G^7 (Figure 8 [c]). The N-terminal region alternatively is unstructured, or assumes turn conformations, as seen from the negative peaks in the G value reported in Figures 8 [a],[b]; this is also shown by overlapping the structures obtained from MD simulations (Figure 10 [c]).

Concerning the 1REI immunoglobulin fragment, high flexible and unstructured regions are present. A multiple transition between coil–sheets–coil–turn–3_{10} and rarely α helix, occurs in residue 6, understandable from the variation of G index during the time (Figure 9 [a]) and from the auto–correlation functions in Figure 11 [a],[b], where three minima and one shoulder for the sheets–coil transition at 20 ns, are present. The other sheet region, centered at residue 23, becomes unstructured ((Figure 9 [c]), consequently the functions show a fluctuating behaviour, as underlined from Figure 11 [a] and more evidently in Figure 11 [b]. Residue 10 (Figure 9[b]) shows values of G typical of turn, coil and interestingly, of polyproline II at 35 ns, whose presence was confirmed with a check of backbone dihedral values, while residue 26 is in a less flexible region of the peptide (cf Figure 9 [d]). This is also confirmed by the auto correlation functions reported in Figure 11 [a],[b] showing both a flat shape. The cross–correlation functions of immunoglobulin fragment reported in Figure 13 [a],[b] show uncorrelated regions, thus pointing to high dynamical states.

Even if the time scale of the simulations performed does not allow a

complete exploration of the conformational space of these long peptides, and only a few exchanges between the most probable structures are sampled, the functions introduced here seem to be able to effectively quantify the time correlation between the different structures and between different regions of a given protein.

5 Conclusions

We have introduced a geometrical chirality index G (see eq.2) that can be easily calculated from the instantaneous conformation of a certain protein fragment [43].

This index assumes well defined values for the typical secondary structure elements and, differentiating from other methods, is particularly effective in detecting polyproline II motifs. We have shown that the index is robust towards random perturbations of the structures and that is stable for long molecular dynamics trajectories that conserve the motif. On the other hand, following the evolution of fragments chirality in time and its correlation offers a direct possibility of monitoring protein conformational changes and we have shown this analyzing 110 ns and 50 ns–long runs for selected hemoglobin and immunoglobulin segments.

We believe that the index proposed here can be a powerful tool in complementing existing structure assignment algorithms, in following folding and misfolding processes for proteins in solution [44, 45], and in particular in capturing the early stages of these extremely important processes.

6 Acknowledgments

We wish to thank Prof. E. Rizzarelli for useful discussions. We also thank University of Bologna, INSTM and FIRB project RBNE03PX83 for financial support.

References

[1] Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol 1963;7:95–99.

[2] Schlick T. Molecular Modeling and Simulation. New York: Springer; 2002. 656 p.

[3] Zagrovic B, Lipfert J, Sorin EJ, Millett IS, van Gunsteren WF, Doniach S, Pande VS. Unusual compactness of a polyproline type II structure. Proc Natl Acad Sci USA 2005;102:11698–11703.

[4] Pietropaolo A, Raiola L, Muccioli L, Tiberio G, Zannoni C, Fattorusso R, Isernia C, La Mendola D, Pappalardo G, Rizzarelli E. An NMR and

molecular dynamics investigation of the avian prion hexarepeat conformational features in solution. 2007.

[5] Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen–bonded and geometrical features. Biopolymers 1983;22:2577–2637.

[6] Mizuguchi K, GōN. Comparison of spatial arrangements of secondary structural elements in proteins. Protein Eng 1995;8:353–362.

[7] Anderson CAF, Palmer AG, Brunak S, Rost B. Continuum secondary structure captures protein flexibility. Structure 2002;10:175 184.

[8] Frishman D, Argos P. Knowledge-based protein secondary structure assignment. Proteins 1995;23:566–579.

[9] Martin J, Letellier G, Marin A, Taly JF, de Brevern AG, Gibrat JF. Protein secondary structure assignment revisited: a detailed analysis of different assignment methods. BMC Struct Biol 2005;5:17.

[10] Dupuis F, Sadoc JF, Mornon JP. Protein secondary structure assignment through Voronoi tesselation. Proteins 2004;55:519–528.

[11] Richards FM, Kundrot CM. Identification of structural motifs from protein coodinate data: secondary strucure and first level supersecondary structure. Proteins 1988;3:71–84.

[12] Majumdar I, Sri Krishna S, Grishin NV. PALSSE: A program to delineate linear secondary structural elements from protein structures. BMC Bioinformatics 2005;6:202.

[13] Stapley BJ, Creamer TR. A survey of left-handed polyproline II helices. Protein Sci 1999;8:587–595.

[14] Cubellis MV, Caillez F, Blundell TL, Lovell SC. Properties of polyproline II, a secondary structure element implicated in protein-protein interactions. Proteins 2005;58:880–892.

[15] Kwiecińska JI, Cieplak M. Chirality and protein folding. J Phys Condens Matter 2005;17:1565–1580.

[16] Ruch E. Chiral derivatives of achiral molecules: standard classes and the problem of a right–left classification. Angew Chem Int Ed Engl 1977; 16:65–72.

[17] Buda AB, Mislow K. A Hausdorff chirality measure. J Am Chem Soc 1992;114:6006–6012.

[18] Osipov MA, Pickup BT, Dunmur DA. A new twist to molecular chirality: intrinsic chirality indices. Mol Phys 1995;84:1193–1206.

[19] Ferrarini A, Nordio PL. On the assessment of molecular chirality. J Chem Soc, Perkin Trans 2 1998;2:455–460.

[20] Harris AB, Kamien RD, Lubensky TC. Molecular chirality and chiral parameters. Rev Mod Phys 1999;71:1745.

[21] Solymosi M, Low RJ, Grayson M, Neal MP. A generalized scaling of a chiral index for molecules. J Chem Phys 2002;116:9875–9881.

[22] Todd SM, Ferrarini A, Moro GJ. Molecular modelling of chiral nematics. Phys Chem Chem Phys 2001;3:5535–5541.

[23] Neal MP, Solymosi M, Wilson MR, Earl DJ. Helical twisting power and scaled chiral indices. J Chem Phys 2003;119:3567–3573.

[24] Earl DJ, Wilson MR. Predictions of molecular chirality and helical twisting powers: A theoretical study. J Chem Phys 2003;119:10280–10288.

[25] Berardi R, Cainelli G, Galletti P, Giacomini D, Gualandi A, Muccioli L, Zannoni C. Can the pi-facial selectivity of solvation be predicted by atomistic simulation? J Am Chem Soc 2005;127:10699–10706.

[26] Kinjo AR, Horimoto K, Nishikawa K. Predicting absolute contact numbers of native protein structure from amino acid sequence. Proteins 2005; 58:158–165.

[27] Pauling L, Corey RB, Branson HR. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci USA 1951;37:205–211.

[28] Marti DN, Schaller J, Llinás M. Solution structure and dynamics of the plasminogen kringle 2-AMCHA complex: 3(1)-helix in homologous domains. Biochemistry 1999;38:15741–15755.

[29] Chou KC. Prediction of tight turns and their types in proteins. Anal Biochem 2000;286:1–16.

[30] Adzhubei AA, Stemnberg M, JE. Left-handed polyproline-II helices commonly occur in globular-proteins. J Mol Biol 1993;229:472–493.

[31] Branden C, Tooze J. Introduction to Protein Structure. New York, NY: Garland Publishing; 1999. 410 p.

[32] Calzolai L, Lysek DA, Perez DR, Guntert P, Wüthrich K. Prion protein nmr structures of chickens, turtles, and frogs. Proc Natl Acad Sci USA 2005;102:651–655.

[33] Bazan JF, Fletterick RJ, McKinley MP, Prusiner SB. Predicted secondary structure and membrane topology of the scrapie prion protein. Protein Eng 1987;1:125–135.

[34] Adcock SA, McCammon JA. Molecular dynamics: Survey of methods for simulating the activity of proteins. Chem Rev 2006;106:1589–1615.

[35] Procacci P, Paci E, Darden T, Marchi M. Orac: A molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. J Comput Chem 1997;18:1848–1862.

[36] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J Am Chem Soc 1995;117:5179–5197.

[37] Allen MP, Tildesley DJ. Computer Simulation of Liquids. Oxford University Press, Oxford; 1989.

[38] Nosé S. A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 1984;81:511–519.

[39] Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Phys RevA 1985;31:1695–1697.

[40] Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Intermolecular Forces. Reidel, Dordrecht: B. Pullman, editor; 1981.

[41] Procacci P, Marchi M. Taming the Ewald sum in molecular dynamics simulations of solvated proteins via a multiple time step algorithm. J Chem Phys 1996;104:3003.

[42] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Molec Graphics 1996;14:33–38.

[43] A code is available upon request to the authors and will be put on the public domain.

[44] Daggett V. Protein folding-simulation. Chem Rev 2006;106:1898–1916.

[45] Wales DJ. Energy Landscapes. With Applications to Clusters, Biomolecules and Glasses. Cambridge: Cambridge U. P; 2003. 692 p.

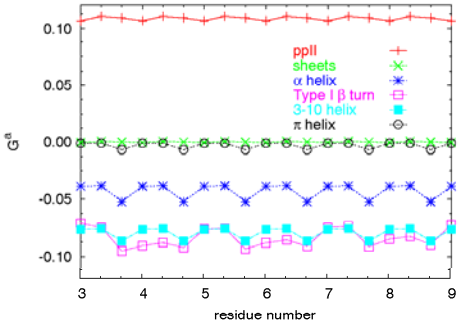


Figure 2: Chiral index, G along the backbone for different secondary ideal structures. The cutoff used is r_C=12 Å and we considered N_a=15 atoms.

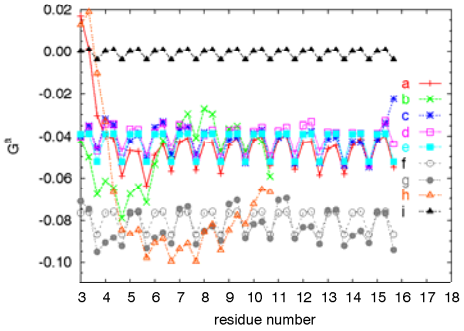


Figure 3: Chiral index, G, along the backbone, for α helices belonging to different proteins: [a]: hemoglobin 1-18 helix, [b]: ChPrP helix1, [c]: ChPrP helix2, [d]: ChPrP helix3, [e]: ideal α helix, [f]: ideal 3_{10} helix, [g]: ideal Type I β turn, [h]: 1LB0 3_{10} helix model peptide, [i]: polyalanine π helix. Type I β turn, 3_{10} and π helices are shown as comparison for ChPrP helix 1, which shows imperfections in the N-terminal region.

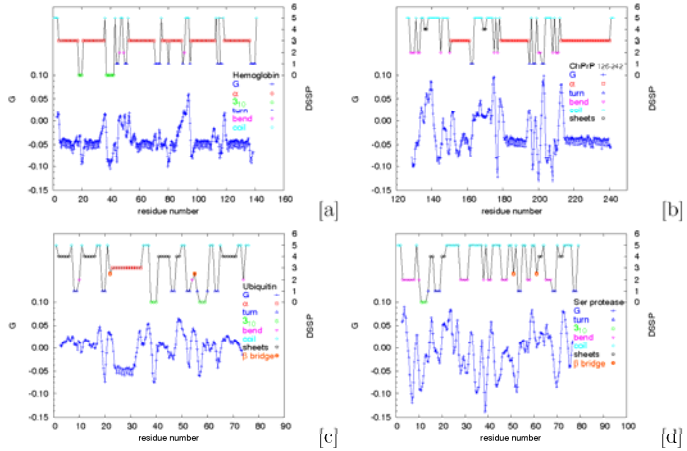
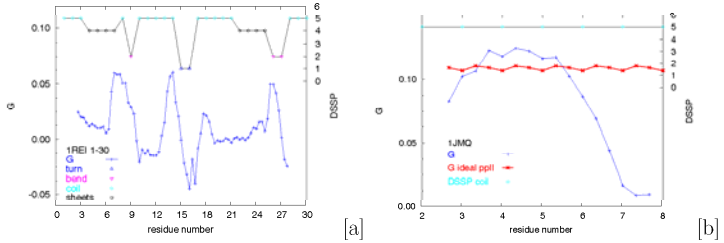


Figure 4: Chiral index, G, along the backbone for different crystalline proteins. Typical secondary structures, with the negative periodicity concerning the α helices, and with the G typical values for the other secondary structures (cf figure 2) are easily identified. The DSSP assignment is also plotted as the numeric code: 3_{10}=0, turn=1, bend=2, bridge=2.5, α=3, sheets=4, coil=5.



[a] [b]

Figure 5: Chiral index behavior for model peptides: [a] Immunoglobulin antigen 1-30 1REI (β sheet 4-7 and 22-25); [b] 1JMQ (poly-L-proline II between residues 3-6). The DSSP assignment is plotted according to a number code which mimics the variation of G (3_{10}=0, turn=1, bend=2, bridge=2.5, α=3, sheets=4, coil=5).

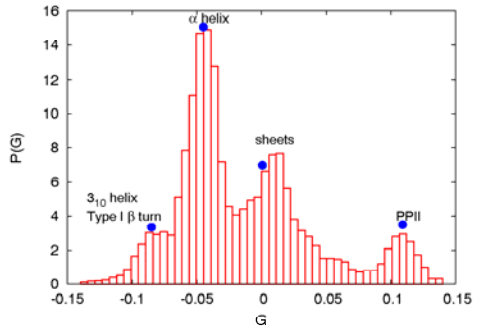
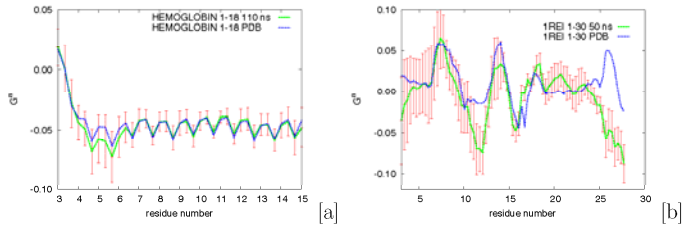


Figure 6: G distributions among the proteins and peptides analyzed in this work. The typical G values of the ideal structures are shown with blue dots.



[a] [b]

Figure 7: Standard deviations of G among hemoglobin 1-18 [a] and 1REI 1-30 [b] configurations. It is worth to note the persistence of the chiral index inside the average configurations for hemoglobin, while in the 1REI immunoglobulin antigen 1-30 fragment the crystal structure is not retained during the simulation. As comparison the G from PDB and from the trajectories is shown.

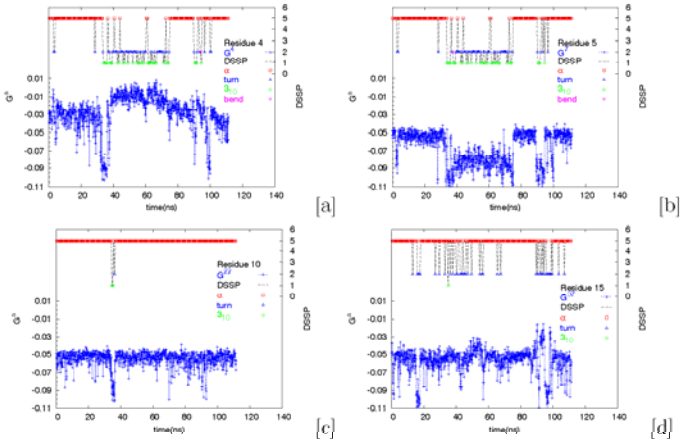


Figure 8: Time dependence of G for the fragments 4, 5, 10, 15 of hemoglobin 1-18 helix. The conversion from α helix to 3_{10} helix is underlined from the lowering of the G index (residue 5, G^7 [b]); the coil-turn transition is evident from the conversion to negative peaks (residue 4, G^4)[a]; the rigid core could be noticed from the constant G values [c],[d]. The DSSP assignment is plotted according to a number code which mimics the variation of G (3_{10}=0, turn=1, bend=2, bridge=2.5, α=3, sheets=4, coil=5).

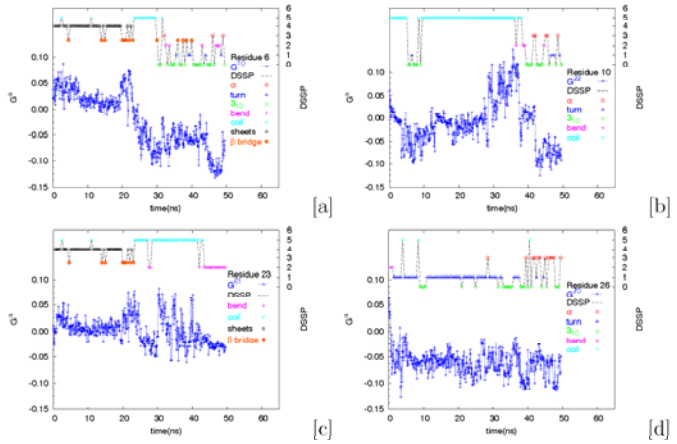


Figure 9: Time dependence of G for the fragments 6, 10, 23, 26 of immunoglobulin antigen 1-30 fragment. The conversion from sheets to coil, turn and 3_{10} helix is underlined from the lowering of the G index and few typical values of α helix are also detected at around 31 ns (residue 6, G^{10} [a]); PPII values can be individuated by the high positive peak present at 35 ns (residue 10, G^{22})[b]; the sheets conformation (residue 23, G^{61} [c]) can be distinguished from the values approaching zero and the transition to unordered states can be detected from the oscillations to positive and negative values near zero. The less flexible core could be noticed from the constant G values (residue 26, G^{70}) [d]). The DSSP assignment is plotted according to a number code which mimics the variation of G (3_{10}=0, turn=1, bend=2, bridge=2.5, α=3, sheets=4, coil=5).

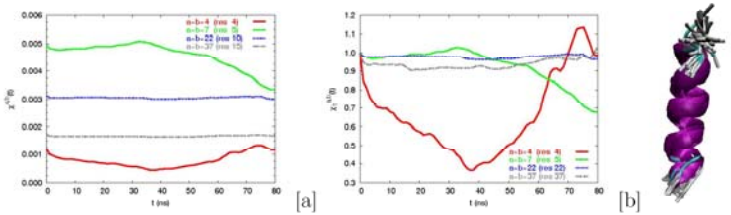


Figure 10: [a] Auto–correlation functions of the chiral index for residues 4, 5, 10, 15 of hemoglobin 1-18. [b] Auto–correlation functions scaled with the square of $G^{a}(0)$. $\chi^{4,4}$ (residue 4) shows the unordered N-terminal region; $\chi^{7,7}$ (residue 5) shows clearly the transition between α helix and 3_{10} helix; $\chi^{22,22}$, $\chi^{37,37}$ (residue 10 and 22 respectively) underline a rigid core in α helix for the fragment 10-18. [c]: VMD visualization [42] of the 1-18 hemoglobin helix, which underlines a flexible N-terminal region (turns are shown in cyan and coils in gray) and a rigid core structure adopting α helix, shown in violet.

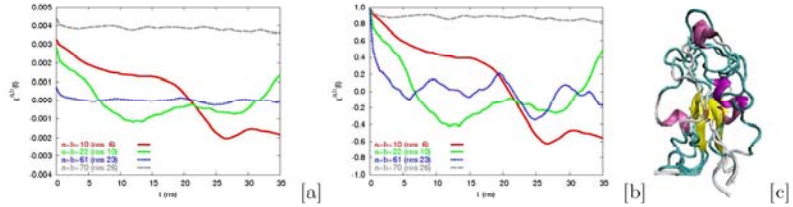
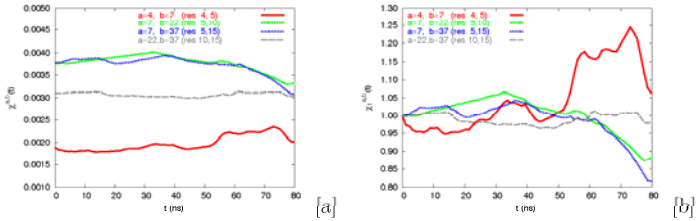


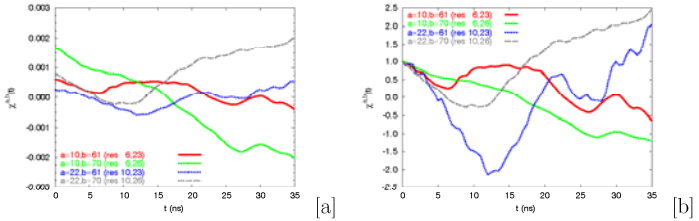
Figure 11: [a] Auto–correlation functions of the chiral index for residues 6, 10, 23, 26 of immunoglobulin antigen 1-30 peptide. [b] Auto–correlation functions scaled with the square of $G^{a}(0)$. $\chi^{10,10}$ (residue 6) shows two minima for the coil–sheets, coil–turn transitions and one shoulder for the sheets–coil transition; $\chi^{22,22}$ (residue 10) shows clearly the two transitions between coil and 3_{10} helix; $\chi^{61,61}$ (residue 23) underlines a less correlation in the trajectories, due to unordered dihedrals, $\chi^{70,70}$ (residue 26), underlines a less flexible core. [c]: VMD visualization [42] of the 1-30 immunoglobulin antigen fragment, which underlines a flexible structure. For residue 6 it can be noticed the transition from sheets to 3_{10} helix; (turns are shown in cyan, coils in gray, sheets in yellow, α helix in violet and 3_{10} in pink).



[a]

[b]

Figure 12: [a] Cross–correlation functions of the chiral index for residues 4, 5, 10, 15 of hemoglobin 1-18. [b] Cross–correlation functions scaled with the square of $G^{a}(0)$. The cross correlation clearly shows the presence of the rigid core for residue 10-15, namely the internal core of the helix and a less correlation between the central core and the N-terminal region.



[a]

[b]

Figure 13: [a] Cross–correlation functions of the chiral index for residues 6, 10, 23, 26 of immunoglobulin antigen 1-30 fragment. [b] Cross–correlation functions scaled with the square of $G^{a}(0)$. The cross correlations reveal a strong loss of correlation in the secondary structures adopted by the immunoglobulin antigen 1-30 fragment, thus showing a high flexibility explored by the 1REI peptide.