Journal of Molecular Graphics and Modelling 50 (2014) 125–133

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Computational study of protein secondary structure elements: Ramachandran plots revisited

Francisco Carrascoza*, Snezana Zaric, Radu Silaghi-Dumitrescu

Department of Chemistry, Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 11 Arany Janos Str, Cluj-Napoca RO-400028, Romania

ARTICLE INFO

Article history:
Accepted 2 April 2014
Available online 13 April 2014

Keywords:
Ramachandran
Secondary structure
Amino acid
DFT
Potential energy surface

ABSTRACT

Potential energy surface (PES) were built for nineteen amino acids using density functional theory (PW91 and DFT M062X/6-311**). Examining the energy as a function of the φ/ψ dihedral angles in the allowed regions of the Ramachandran plot, amino acid groups that share common patterns on their PES plots and global minima were identified. These patterns show partial correlation with their structural and pharmacophoric features. Differences between these computational results and the experimentally noted permitted conformations of each amino acid are rationalized on the basis of attractive intra- and intermolecular non-covalent interactions. The present data are focused on the intrinsic properties of an amino acid – an element which to our knowledge is typically ignored, as larger models are always used for the sake of similarity to real biological polypeptides.

1. Introduction

The secondary structure conformation of a protein can be expressed as a function of its backbone dihedrals expressed in (φ, ψ) pairs that can be represented in a Ramachandran type graphic for easier interpretation. These plots are typically split in forbidden and allowed regions [1]. Around 40% of all the amino acids in a structure are contained in just the 2% of the Ramachandran plot – the so-called "allowed areas" [2,3]. The non-allowed regions are those defined by a minimal contact distance between atoms of the neighbor amino acids ($n+1$) and ($n-1$), 'n' being the amino acid with the central alpha-carbon of reference [4]. Measuring the changes in energy by rotating the ψ and φ dihedral angles around the α-carbon may help to understand in a quantitative way the conformational preferences and allow predictions of three-dimensional structures [4–6]. Choices for a given type of secondary structure are dictated by interatomic interactions – both repulsive and attractive (primarily classical and non-classical hydrogen bonds) [7]. The understanding of the nature of the secondary structure has been approached statistically and theoretically.

Statistical approaches for secondary structure prediction are based on the probability of finding an amino acid in certain conformation; they use large protein X-ray diffraction databases. For instance, the Position-Specific Scoring Matrix (PSSM) implemented in a neural network, is based on similarity comparisons and predicted the tridimensional structure of a polypeptide with a success of 74% [8]. Similar success rates were reported by others, eventually going up to 80% [9–12]. It was also possible to establish with statistical approaches that each amino acid has certain preferences for secondary structure types, with structural elements proposed to explain such differences between amino acids [13]. Interestingly, rather than matching with the classical pharmacophoric classifications (non-polar, polar negative charged, positive charged, uncharged), these preferences appear to be dictated more specifically by structural factors such as steric hindrance and charge repulsion. Nevertheless, the prediction of the behavior of each amino acid in a real system becomes more complex due to the mixture of the factors within the neighbor environments [14,15].

On the theoretical side, the "Ramachandran plot" was a theoretical attempt reported in the early 60s [4] to explain the amino acid conformation by setting very logical rules of inter-atomic attraction and repulsion into a simple map, in times when crystallographic data were not developed. Latter *ab initio* calculations made their contribution with the results obtained at Hartree–Fock level and compiled in force fields, with AMBER [16] and CHARMM [17] as some of the most representative. However, these theoretical calculations were predicting a minimum in the 2.2_7 ribbon zone, whereas statistical methods showed that other regions were more populated. The problem with their force fields was linked to the implementation of the non-bonded interactions. David Baker tried to make improve by predicting correct phi, psi angles incorporating a predefined library of statistically probable

* Corresponding author. Tel.: +40 752524278.
E-mail address: franciscocarrascoza@gmail.com (F. Carrascoza).

http://dx.doi.org/10.1016/j.jmgm.2014.04.001
1093-3263/

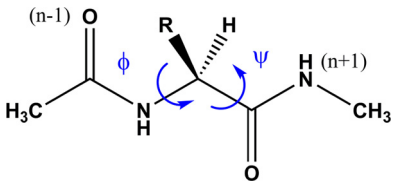

Fig. 1. Scheme of 2-acetamido-N-methylpropanamide structures worked. The φ and ψ angle rotations are pointed out with blue arrows. O (n-1) refers to the oxygen in the position $(n-1)$ where $(n-1)$ is the previous amino acid to the central amino acid (n). In the same way H $(n+1)$ refers to the hydrogen in position $(n+1)$ respect to the amino acid (n). *R* is used to point out the *R* chain of the amino acid.

local fragment conformations [18]; although this was successful in terms of correct prediction of turn type conformations, the approach does not offer a solution to the basic theory. AMBER and CHARMM force fields incorporated later a statistical correction term improving notably their results [19,20].

Stronger quantum mechanical calculations at DFT and MP2 level of theory have been developed more recently [21–23]. Tsai and co-workers reported that is possible to get a Ramachandran plot very close to statistical results for Gly_3 peptides by taking in account solvent and some near carboxy and amino interactions. Also Zhu and co-workers has shown good results on χ_1 and χ_2 dihedrals at MP2 level of theory [24].

Our research group has performed also previous DFT studies for α-helix and β-sheet polypeptide types [25] with an analysis of those hydrogen bond interactions and the solvent effect, as non-bonded interactions that affect the conformation of the amino acid and the secondary structure itself. Here we report potential energy surfaces (PES) scans at DFT, M062X and PW91 levels of theory, in order to evaluate exclusively bonded and near intra-molecular factors that affect the conformation of the backbone, for 19 of the most common amino acids. Similarities and differences are observed between these 19 PES's, as well between them and the canonical Ramachandran plots. These are rationalized on the basis of steric repulsion counterbalanced by weak pseudo-hydrogen bond attractive interactions, as well as on the basis of the availability for further intermolecular interactions. The PES's were also analyzed by a similarity score algorithm, which allowed us to cluster them and make comparisons with previous clusters of amino acids based on statistics.

Our results match previous theoretical studies at lower levels of theory. Although in contrast with statistical results they predict different conformations, the amino acids are still interestingly clustered in very similar groups to the statistically determined ones. This comparison gives a quantitative effect of that intrinsic force within the amino acid that drive its conformational preferences in general – preferences that are then expected to also be at work in peptide structures. These results may be of interest for those deriving new force fields and in protein structure predictions.

2. Materials and methods

Nineteen analog structures of (S) 2-acetamido-N-methylpropanamide were constructed, each them similar to each of the nineteen key amino acids, as indicated in Table 1 and Fig. 1; essentially, these are amino acids capped at each end with a peptide bond so as to define the φ and ψ angles. The peptide bonds are each capped with a methyl group, which thus models in a neutral way the carbon of the neighboring amino acids in a putative polypeptide incorporating the amino acid examined. Proline was not examined, because its dihedral angles are constrained due to its internal molecular structure.

Potential energy surfaces were built using density functional theory (DFT), M06-2x/6-311(d,p) [26] at vacuum as implemented in the Gaussian09 software package. Additionally similar PW91 calculations were performed for all models. For selected models, other levels of theory were also employed, as indicated in text. In order to obtain a complete plot of energy over the changes in the dihedral angles ψ and φ, a scan was performed rotating around these angles in twelve steps of 30° each, giving a total of 169 structures into a matrix of 13×13 molecules from −180°, until +180° for ψ and −180° to +180° for φ. Restrictions were applied over the φ and ψ angles for each compound and minimizations over all the structure were allowed.

The Asn case was studied with and without constrains, to highlight the effect of the R-chain into its backbone. Asp, Arg, Glu, His and Lys were computed with 0 charge.

Table 1
Structures employed and their similarity with corresponding amino acids.

Name	Amino acid
(S)-2-acetamido-N-methylpropanamide	Alanine
(S)-2-acetamido-5-(diaminomethylamino)-N-methylpentanamide	Arginine
(S)-2-acetamido-N1-methylsuccinamide	Asparagine
(S)-3-acetamido-4-(methylamino)-4-oxobutanoic acid	Aspartic acid
(S)-2-acetamido-3-mercapto-N-methylpropanamide	Cysteine
(S)-2-acetamido-N1-methylpentanediamide	Glutamine
(S)-4-acetamido-5-(methylamino)-5-oxopentanoic acid	Glutamic Acid
2-acetamido-N-methylacetamide	Glycine
(S)-2-acetamido-3-(2,3-dihydro-1H-imidazol-4-yl)-N-methylpropanamide	Histidine
(2S,3S)-2-acetamido-N,3-dimethylpentanamide	Isoleucine
(S)-2-acetamido-N,4-dimethylpentanamide	Leucine
(S)-2-acetamido-6-amino-N-methylhexanamide	Lysine
(S)-2-acetamido-N-methyl-4-(methylthio)butanamide	Methionine
(S)-2-acetamido-N-methyl-3-phenylpropanamide	Phenilalanine
(S)-2-acetamido-3-hydroxy-N-methylpropanamide	Serine
(2S,3R)-2-acetamido-3-hydroxy-N-methylbutanamide	Threonine
(S)-2-acetamido-3-(1H-indol-3-yl)-N-methylpropanamide	Tryptophan
(S)-2-acetamido-3-(4-hydroxyphenyl)-N-methylpropanamide	Tyrosine
(S)-2-acetamido-N,3-dimethylbutanamide	Valine

3. Results and discussion

3.1. DFT-derived Ramachandran plots

Fig. 2 shows DFT-derived Ramachandran plots for the compounds listed in Table 1 (cf. Fig. 1), which represent 19 of most common amino acids (proline not examined). Beyond some general similarities between the plots, one may note salient differences between them. Fig. 3 illustrates the classical descriptions of the preferred and allowed areas in a Ramachandran plot, superimposed over experimental data.

3.2. α-Helix vs 2.2_7 ribbon conformation

As illustrated in Fig. 2, most of the 19 amino acids display global minima localized at φ, $\psi = -90$, 60, which is the so called 2.2_7 ribbon conformation, or γ-turn (Fig. 4). This type of structure is located in the Ramachandran plot at an intermediate point between β-sheet and α-helix type. Our results are in agreement with those obtained by other authors with different levels of theory: HF/6-31G(d)//HF/6-31(d) vacuum and HF/6-311G(d,p) vacuum [28] AMBER force field [16] and CHARMM [17]. Iwaoka and co-workers pointed out in their work the solvent effect, which is later described at higher level of theory by Zhu and co-workers in 2012 [24] for χ_1 and χ_2 angles. The present work is, on the other hand, focused on the intrinsic differences between amino acids, before external factors (be they solvent or neighboring amino acids)

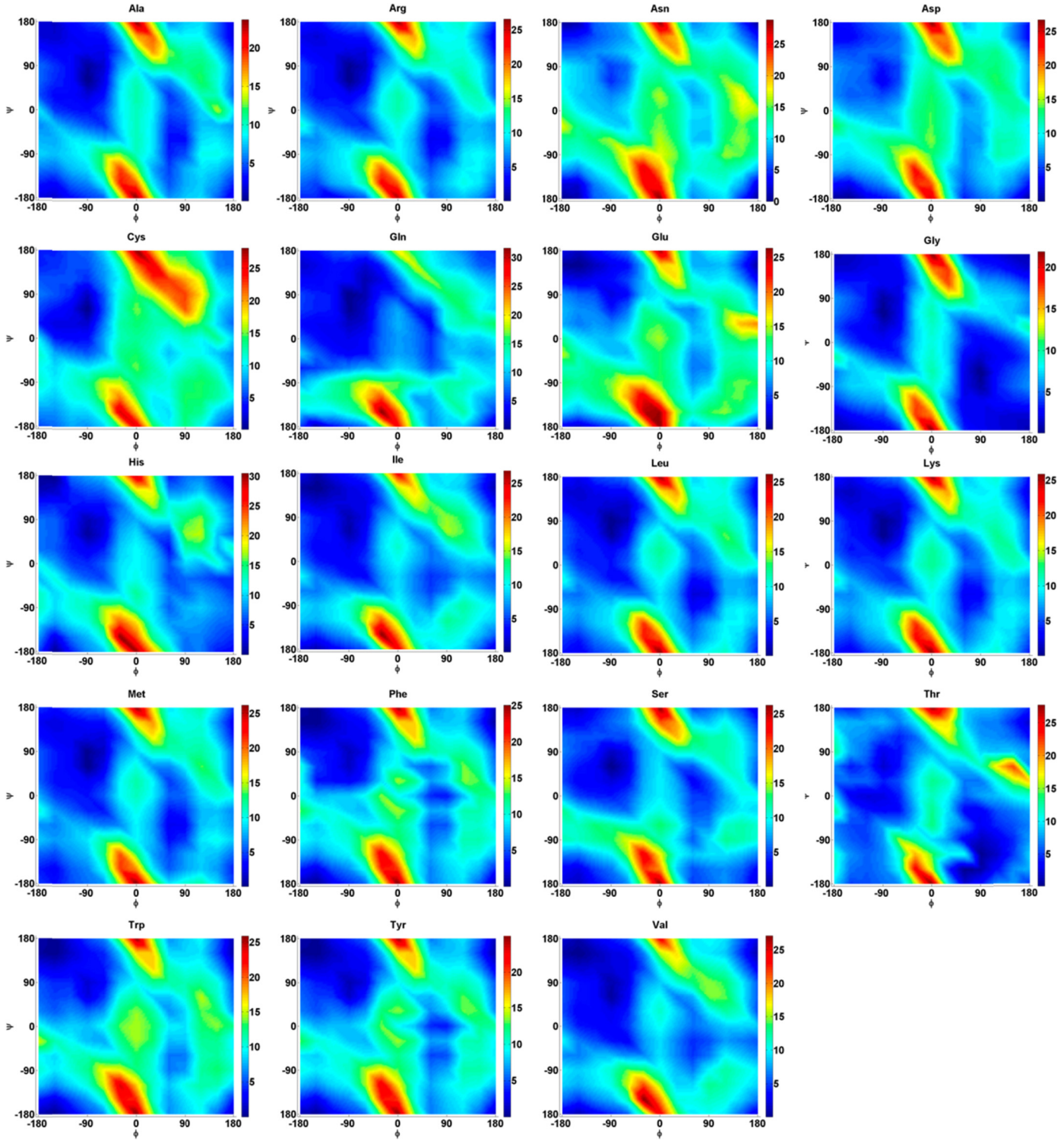

Fig. 2. PES scans for the 19 amino acid models examined in the present study; DFT M06-2x/6-311G(d,p).The energy axis (colored) is in kcal/mol, and in each individual plot it is referenced to the internal lowest-energy point.

are taken into account. In experimental data [29] most of the amino acids are found in α-helix or in β-sheet conformations, the 2.2_7 ribbon conformations in principle also efficiently avoids sterical clashed between the R chains. In fact, careful inspection of Fig. 3A indicates that the 2.2_7 ribbon area is also populated, and that many of the experimental points lie significantly outside the canonical α-helix and β-sheet areas.

3.3. *Standard conformations vs global minima; dependence on methodology*

Table 2 lists relative energies computed for data points corresponding to canonical φ, ψ pairs in alanine, not only as computed using DFT but also as derived from post-Hartree–Fock calculations. As discussed above, the global minimum corresponds to a

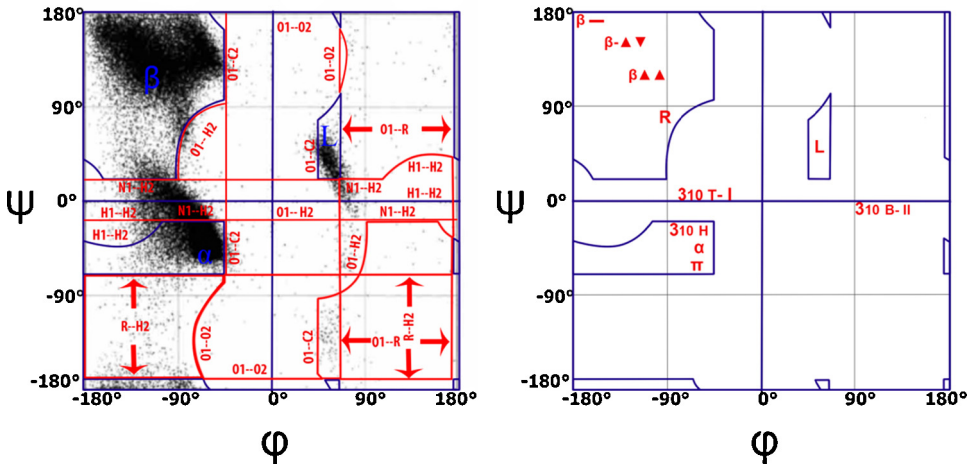

Fig. 3. *Left* – φ/ψ plot as defined by Ramachandran [4] with steric clashes reported by Ho and co-workers [27]. Collisions between groups leading to prohibited regions are marked in red lines, highlighting the atoms that collide in that conformation. As background is shown a φ/ψ plot reported by Ho and co-workers [27] that contains 18 of the common amino acids, but not Pro nor Gly. *Right* – Typical Ramachandran plot definition: β— = β-sheet planar; β▲▼ = β-sheet antiparallel; β▲▲ = β-sheet parallel; **R** = 2.2_7 ribbon; 3_{10}T-I = 3_{10}Turn type I; 3_{10} B-II = 3_{10} Bend type II; 3_{10} H = 3_{10} Helix; α = α-helix; **L** = Left-handed α-helix; **π** = π-helix.

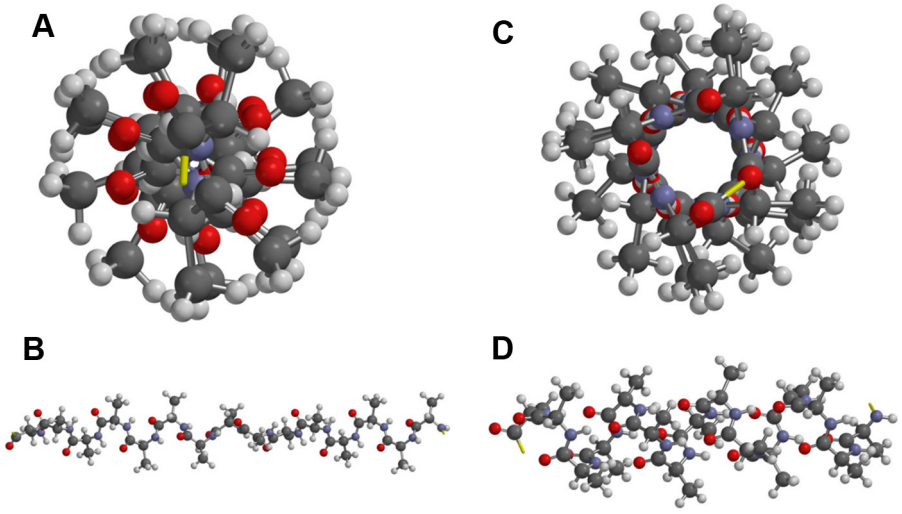

Fig. 4. 2.2_7 Ribbon and α-helix conformations for an hypothetical peptide $[Ala]_{10}$. (A) 2.2_7 ribbon frontal view, (B) 2.2_7 ribbon lateral view, (C) α-helix frontal view, and (D) α-helix lateral view.

2.2_7 ribbon. Planar β sheets follow suit closely, at 2–3 kcal/mol (depending on the method). Very close in energy are also the β turns (~4 kcal/mol, with all methods examined). The β pleated sheet and the π helix have very similar energies, at 5–6 kcal/mol, followed by the α-helix at 6–7 kcal/mol. By far the least advantageous are the left-hand helix and the 3_{10} helix (11–12 kcal/mol). Notably, these data, as obtained on our small models, reflect primarily the intrinsic sterical and electronic properties of the amino acid, and do not take into account its possibilities to engage in supra- or intra-molecular hydrogen bonds with neighboring amino acids; it would be such inter-amino acid interactions, that strongly stabilize the (cf. Table 2) otherwise very common α-helix and pleated β strands.

Table 2
Relative difference from the global minima φ, $\psi = (-90, 60)$. Values expressed in kcal/mol.

Conformation (φ, ψ)	M062x	MP2	CCSDT
β-Planar sheet (180°, 180°)	2.86	3.56	3.43
β-Planar sheet (−180°, 180°)	2.86	4.30	4.29
β-Turn (−60°, −30°)	4.13	3.76	3.80
β-Pleated sheet (−140°, 135°)	5.43	5.09	4.82
α-Helix (−58°, −47°)	7.11	6.02	6.19
Left-hand α helix (58°, 47°)	12.16	11.69	11.90
3_{10} Helix (−49°, −26°)	10.95	10.33	10.64
π-Helix −57°, 70°)	5.63	5.12	5.38

Small differences were observed when applying a different functional (PW91) as compared to the M062x, as illustrated in Fig. 5. Numerically, the differences are generally distinctly lower than 2 kcal/mol – i.e., not much different from those seen in Table 2. The two functionals apparently place different weights on hydrogen-bonding vs sterical clashes especially at the 0^0, 0^0 point. In this region the C=O ($n-1$) clashes with HN ($n+1$) and M062x expectedly reports this clash as a maximum (cf. Fig. 2); however, PW91 reports only a saddle point, relatively lower than its correspondent point in the M062x method – by up to −2.00 kcal/mol. PW91-derived Ramachandran plots are shown for all amino acids in Supporting Information, showing a general agreement with the M06-2x data from Fig. 2.

3.4. Asparagine: the side chain–back bone interactions

One may take Asn as a typical example of side chain–back bone interactions. By using the M062x functional and rotating the φ/ψ

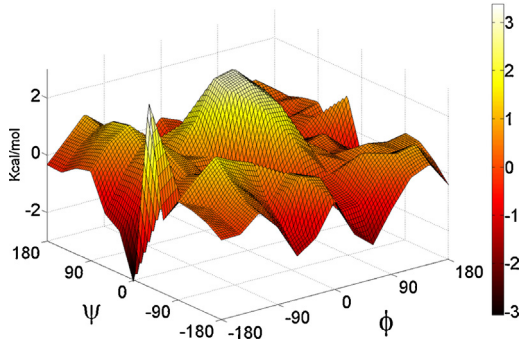

Fig. 5. Difference average between DFT M062x and PW91PW91 methods. A complete description of the graphics obtained by PW91 method can be seen in the Supplementary Material section.

angles we obtained 2 different plots (Fig. 6A and B), depending on the rotation of the χ_1 and χ_2 angles. In the case of Fig. 6E, the global minimum clearly points to a 2.2_7 ribbon zone (6A), because of 3 hydrogen bonds providing stability – while in the case of Fig. 6F only 2 hydrogen bonds are possible and the global minimum is located in the β-sheet planar region (6B). The same plot was obtained at PW91 level (6C). We took as reference for this study the case where only 2 hydrogen bonds are formed (Fig. 6F, $\chi_1 + 77^\circ$) because it provides less interaction side chain–back bone. Fig. 6D shows our own statistical study on the Gly-Asn-Gly occurrence, illustrating that the preferred statistical conformation seems to be the 3_{10}Turn type I region.

3.5. Comparisons between the 19 amino acids

In Fig. 2 one may visually recognize some groups. Group 1 would include Ala, Arg, Leu, Lys and Met. They match in the 2.2_7 ribbon and the Bend type II areas, as well they have similar secondary alternatives. Group 2, by simple visual inspection, may be composed by Asn, Glu, Ile, Tyr, Val and Ser, which have the planar β-sheet and 2.2_7 ribbon zones as first choice. Then Group 3 would contain only Gln and Thr, which have as only preference the 2.2_7 ribbon zone. The next group, His, Phe, and Trp, have a preference for the planar β-sheet region. The plots corresponding to Gly and Cys are too different to match any group.

In order to confirm these observations, a numerical method was used in order to compare the level of similarity between an amino acid and each one of the groups proposed, using a mathematical algorithm:

$$D_{Tx} = \frac{\sum_1^{169} D_x}{169}$$

where D_{Tx} is the total deviation of the relative energies for amino acid X, and D_x is defined as

$$D_x \left| E_x - \frac{\sum_1^{169} E_{Gj}}{n_{Gj}} \right|$$

where E_x is a matrix of energy values for the amino acid x and the term

$$\frac{\sum_1^{169} E_{Gj}}{n_{Gj}}$$

describes the average energy per point in a matrix (13×13 points) of energy of the Group J (Gj).

Thus it was possible to compare quantitatively the level of similarity between an amino acid respect each group formed. The results are presented in Table 3, Gly, His and Cys did not appear to fit in any group.

The amino acids of Table 3 were ordered with respect to the deviation from average of all the PES scan plot of all the amino acids calculated with the formula described above. The level of difference between the members of a group was confirmed creating an average between these members and obtaining its corresponding deviation. Thus, the members of a group should have the minimum deviation with respect to the average of the other groups, if they really belong to this group. For example: Ala was included to calculate the average of the Group 1. Then Alanine has a deviation from this average of 0.813633, which is the lowest deviation

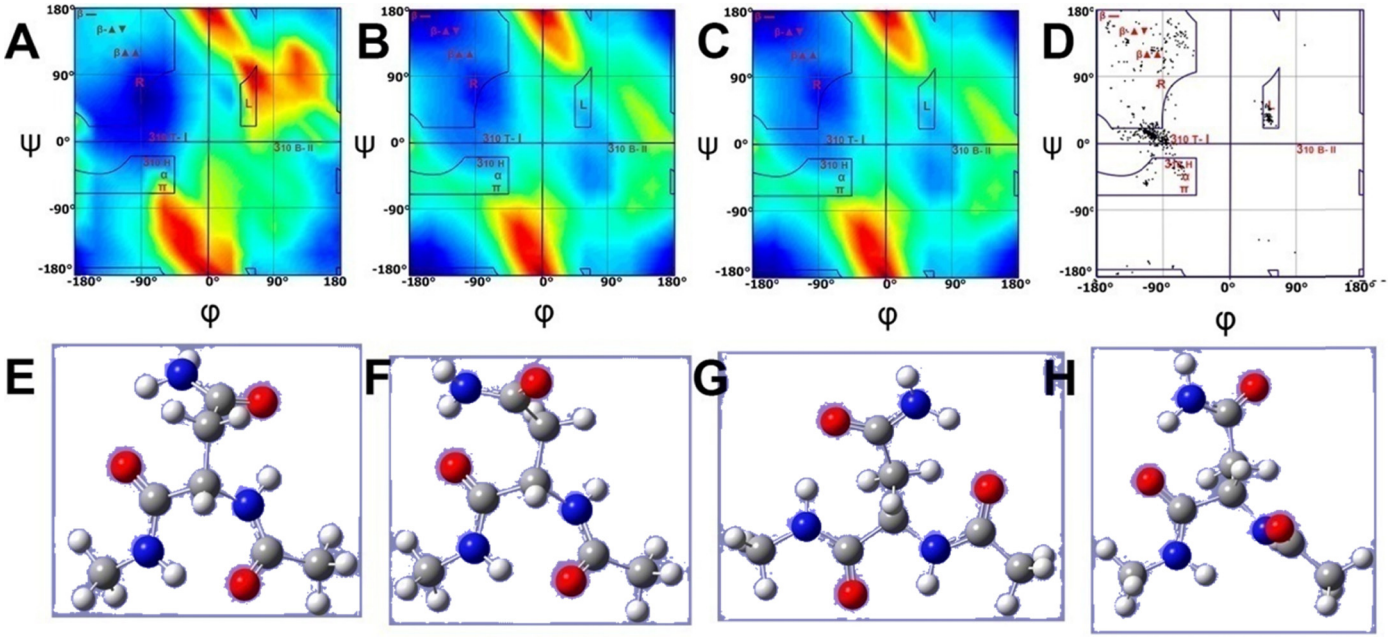

Fig. 6. Comparisons of the effect of side chain in Asn. (A) M062x PES plot $\chi_1 - 73^\circ$; (B) M062x PES plot; (C) PW91PW91 PES plot $\chi_1 + 77^\circ$; (D) statistical occurrence of Gly-Asn-Gly sequence in 6900 proteins with resolution ≤ 1.5 Å, extracted from Protein Data Bank; Asn conformations with (φ, ψ, χ_1) angles: (E) (−90, 60, −73); (F) (−90, 60, +77); (G) (−180,+180,+108); (H) (−90,+30,−70). All the angles reported in degrees.

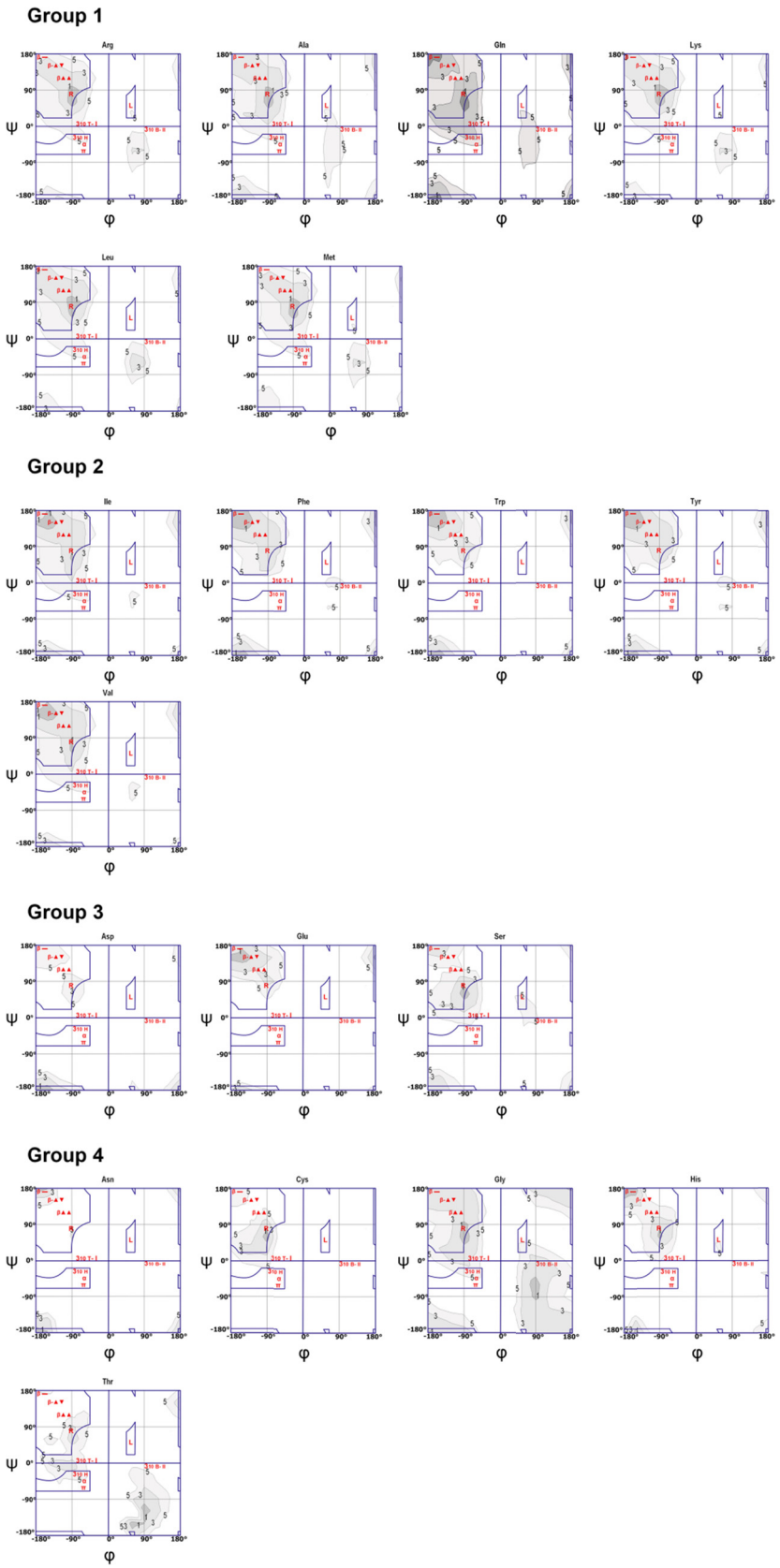

Fig. 7. M06-2x-derived Ramachandran plots, showing only areas as high as 5 kcal/mol (isocontour values of 1, 3 and 5 kcal/mol, respectively, are indicated in each plot, alongside classical assignments of the areas specific to canonical secondary structures.

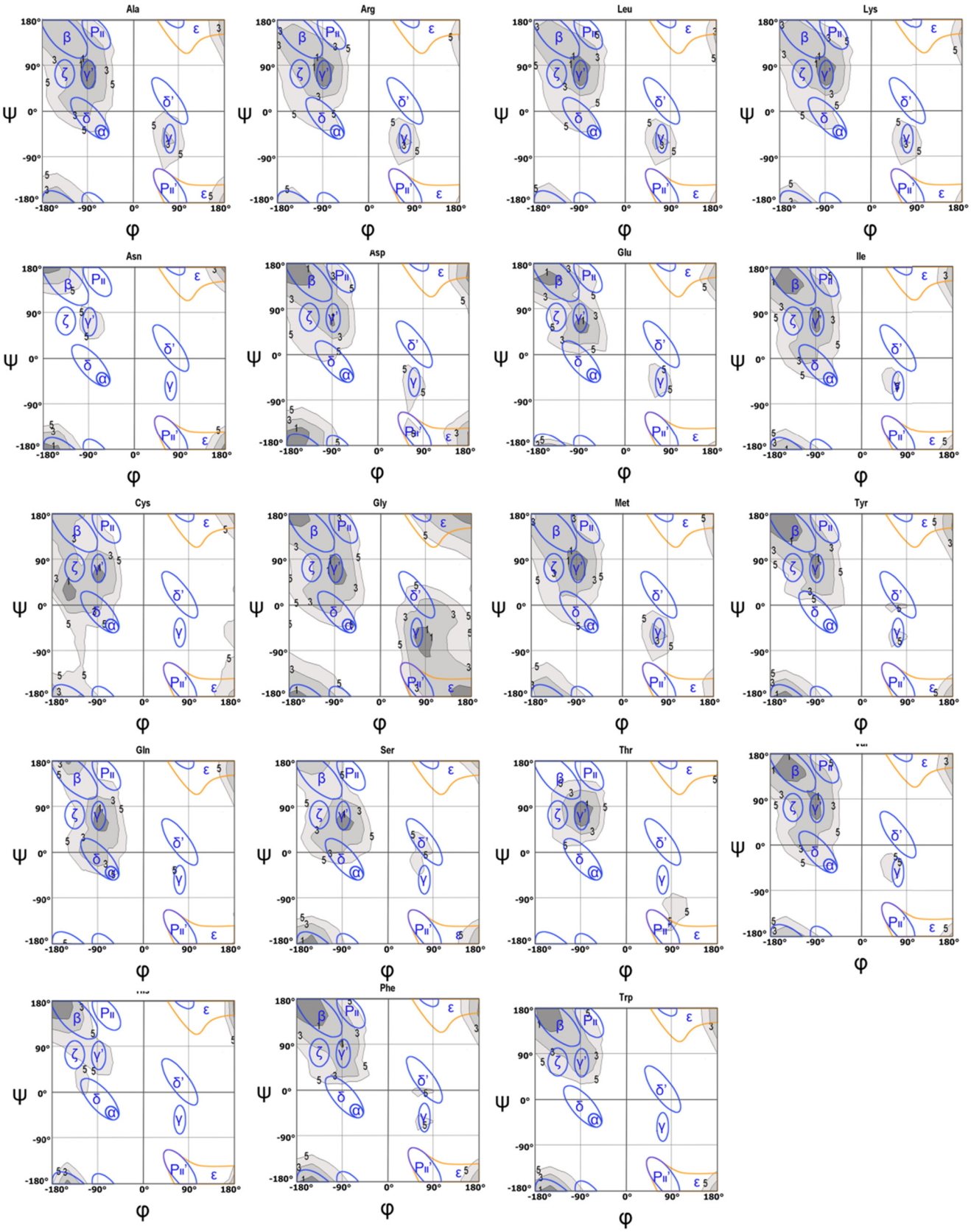

Fig. 8. DFT/M062x results analyzed with the new nomenclature proposed by Hollingsworth and Karplus [29]. α = α-helices; β = β-strands; PII = PII-spirals (pre-proline); γ = γ-turns; γ′ = γ′-turns; ε = ε region populated by glycine residues; δ = bridge region; populated by type I; type III and type II′ turns.

Table 3
Deviations calculated for the entire PES scan Plots. ATD means Average Total Deviation. Groups 1, 2 and 3 points out the amino acids that fits in each group, they are split by an horizontal line. The numbers highlighted in bold points out the members of a group used to calculate the average of this group.

Group	Amino Acid	ATD	Group 1	Group 2	Group 3
1	Gly	17.3444	2.420963	2.676374	3.506577
	Ala	15.7382	**0.813633**	1.293726	2.040494
	Gln	15.5868	1.100353	1.779178	2.26657
	Lys	15.1054	**0.415846**	1.267972	1.860956
	Leu	15.0796	**0.461985**	1.396356	1.94503
	Met	15.0648	**0.61618**	1.414849	2.130802
	Arg	14.9607	**0.523775**	1.305816	1.842973
	Ile	14.9506	**0.795223**	1.640883	1.984035
	Val	14.8783	**0.854747**	1.696393	2.028414
2	Thr	14.7826	2.466858	**2.129595**	2.785898
	Phe	14.7456	1.486954	**1.048019**	1.187413
	Tyr	14.7207	1.445923	**1.029253**	1.048664
	Ser	14.6695	1.700903	**1.177481**	1.545723
3	Trp	13.896	1.820893	1.404989	**0.813087**
	His	13.6361	2.281154	1.93197	1.91931
	Asp	13.5625	2.086454	1.530448	**0.853438**
	Glu	12.8696	2.623419	2.203506	**1.442626**
	Cys	11.9348	3.302523	3.26889	2.864504
	Asn	11.5901	3.943392	3.392175	**2.421025**

respect to the deviations for alanine in the Groups 2, 3 and 4. Then it is confirmed that Alanine is similar to the member of the Group 1. Gly case is opposite: it was taken into the average of Group 3 as it was considered so different from any other amino acids. Then, although Gly was part of this average, its deviation in Group 3 is the highest, meaning that it is totally different from the Group 3; nevertheless its minimum difference of deviation is within Group 1.

It is important to say that there is no linear correlation between the deviations of the members of a group and deviations of the other groups, because the deviation distance between two amino acids in a Cartesian plot is going to change depending on the point of reference in the multidimensional space.

We have previously reported a statistical analysis of amino acid preferences for various types of secondary structure, based on data from the Protein Data Bank [13]. A correlation was at that point noted between the secondary structure preferences and three structural features revolving around the nature of the β carbon: the presence of a heteroatom on Cβ or Cγ (turn and bend preferrers), the presence of an aromatic or branching point at Cβ (β strand preferrers), or the lack of any of these two features (α helix preferrers).

The statistical classification [13] gave His and Cys as outliers, in the sense that they seem to have no particular preference; indeed, Table 3 also points out His and Cys as the ones with biggest deviation respect to the four groups defined.

Also from statistics [13], Trp, Tyr, Phe, Ile, Val, and Thr were found to cluster together, in a group noted for its preference for β strands; Thr was a somewhat special case, as its preference for coils/turns/bends was also very strong. Indeed, the general shapes of the corresponding DFT-derived Ramachandran plots in Fig. 2 are clearly similar for Trp, Tyr, Phe. However, our numerical analysis points out Trp to be in a different group.

Table 4 further lists a more detailed account of the DFT-derived preferences, alongside with preferences previously derived from PDB-based statistics [13]. The two sets of data may be taken to generally agree. Thus, most of the β sheet preferrers of Malkov and co-workers [13] are found to feature preferred DFT-derived minima in the β region. Half of these six amino acids belong to Group 2 (cf Table 3). All the Malkov "helix preferrers" [13] were predicted by DFT to prefer 2.2_7-ribbon; the exception is Glu, for which the ribbon is a second, not a first option.

The fine detail of the energy minima defined in Fig. 3 (and implicitly Table 4) are subject to weak non-classical interactions

Table 4
Preferences of amino acids for various types of secondary structure, as derived from statistics (Malkov et al. [13], column 2) and DFT (present work; see also Fig. 7).

Amino acid	Group by similarity	Group by statistics	DFT favored minima	DFT-derived second choice
Ala	1	α-Helix	2.2_7-Ribbon	β-Sheet
Leu	1	α-Helix	2.2_7-Ribbon	β-Parallel, bend
Glu	3	α-Helix	β-Sheet antiparallel	2.2_7-Ribbon
Gln	1	α-Helix	2.2_7-Ribbon	β-Sheet antiparallel
Arg	1	α-Helix	2.2_7-Ribbon	β-Sheet, bend
Met	1	α-Helix	2.2_7-Ribbon	β-Sheet, bend
Lys	1	α-Helix	2.2_7-Ribbon	β-Sheet, bend
Val	1	Strand	β-Sheet antiparallel, 2.2_7-ribbon	β-Sheet parallel
Ile	1	Strand	β-Sheet antiparallel, 2.2_7-ribbon	β-Sheet parallel
Tyr	2	Strand	β-Sheet planar, antiparallel	β-Sheet parallel, 2.2_7-ribbon
Phe	2	Strand	β-Sheet planar, antiparallel	β-Sheet parallel, 2.2_7-ribbon
Thr	2	Strand	Bend type II	2.27-Ribbon, 3_{10}Turn I
Trp	3	Strand	β-Sheet-planar, antiparallel	β-Sheet-parallel, 2.2_7-ribbon
Gly	–	Other	2.2_7-Ribbon	β-Sheet, bend
Asn	3	Other	β-Sheet planar	2.2_7-Ribbon
Pro	N.A.	Other	N.A.	N.A.
Asp	3	Other	β-Sheet planar	2.2_7-Ribbon
Ser	2	Other	2.2_7-Ribbon	β-Sheet planar
Cys	–	No defined	2.2_7-Ribbon	3_{10}Turn I
His	–	No defined	2.2_7-Ribbon, β-sheet planar	β-Sheet antiparallel

such as the C=O—HC, all of which are expected to entail energy costs of only a few kcal/mol – if not less. When embedded in a medium where neighboring residues provide functional groups more adept at forming hydrogen bonds (i.e., with larger energies of stabilization), these groups will reorient themselves toward partners outside the current amino acid; all of this may well affect the secondary structure preference of the amino acid. In this respect, our data in Figs. 2 and 3 holds closer relevance to isolated small peptides than to real-life proteins – and indeed suggests that ribbons and planar β sheets are among the most likely choices of secondary structures for such peptides.

Amino acids in Fig. 8 were ordered by its similarity observed by eye-check, which partially agree with the numerical analysis (in Table 3). The new nomenclature defined by Hollingsworth and Karplus [29] also recognize PII-spirals, γ′-turns and $\varepsilon = \varepsilon$ region formally, which match better our results.

4. Conclusions

Three-dimensional Ramachandran plots were derived from DFT calculations on simple individual amino acid models, showing the extent to which certain areas would be prohibited energetically. Disagreements are found between these data and statistics-based classical Ramachandran plots; these are explained using supramolecular chemistry considerations: some regions which show local minima in the single-amino acid models are in fact shown to be non-profitable at supramolecular level in the sense that interactions with other polypeptide chains or with other amino acids in the same polypeptide chain are hampered. The present data are focused on the intrinsic properties of an amino acid – an element which to our knowledge is typically ignored, as larger models are always used for the sake of similarity to real biological polypeptides. Nevertheless, for small peptides, including, e.g., prebiotic chemistry, such small models hold information not readily available from full protein/polypeptides. Further work is in progress to quantify the effects of intramolecular near-neighboring amino acids, relative to the intra-amino acid ones.

Acknowledgement

Funding from the Romanian Ministry of Education and Research, (grant PN II 312/2008) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2014.04.001.

References

[1] C. Ramakrishnan, G.N. Ramachandran, Stereochemical criteria for polypeptide and protein chain conformations: II. Allowed conformations for a pair of peptide units, Biophys. J. 5 (1965) 909–933.
[2] S.C. Lovell, et al., Structure validation by Calpha geometry: phi, psi and Cbeta deviation, Proteins 50 (2003) 437–450.
[3] R.W. Hooft, C. Sander, G. Vriend, Objectively judging the quality of a protein structure from a Ramachandran plot, Comput. Appl. Biosci. 13 (1997) 425–430.
[4] G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, Stereochemistry of polypeptide chain configurations, J. Mol. Biol. 7 (1963) 95–99.
[5] G.J. Kleywegt, T.A. Jones, Databases in protein crystallography, Acta Crystallogr. D: Biol. Crystallogr. 54 (1998) 1119–1131.
[6] G.J. Kleywegt, T.A. Jones, Phi/psi-chology: Ramachandran revisited, Structure 4 (1996) 1395–1400.
[7] J. Wang, X. Zheng, Comparison of protein secondary structures based on backbone dihedral angles, J. Theor. Biol. 250 (2008) 382–387.
[8] M.J. Wood, J.D. Hirst, Protein secondary structure prediction with dihedral angles, Proteins 59 (2005) 476–481.
[9] W. Pirovano, J. Heringa, Protein secondary structure prediction, Methods Mol. Biol. 609 (2010) 327–348.
[10] J. Zhao, P.-M. Song, Q. Fang, J.-H. Luo, Protein secondary structure prediction using dynamic programming, Acta Biochim. Biophys. Sin. (Shanghai) 37 (2005) 167–172.
[11] R. Bondugula, A. Wallqvist, M.S. Lee, Can computationally designed protein sequences improve secondary structure prediction? Protein Eng. Des. Sel. 24 (2011) 455–461.
[12] H.-N. Lin, T.-Y. Sung, S.-Y. Ho, W.-L. Hsu, Improving protein secondary structure prediction based on short subsequences with local structure similarity, BMC Genomics 11 (Suppl. 4) (2010) S4.
[13] S.N. Malkov, M.V. Zivković, M. Beljanski, V. Hall, M.B.S.D. Zarić, A reexamination of the propensities of amino acids towards a particular secondary structure: classification of amino acids based on their chemical structure, J. Mol. Model. 14 (2008) 769–775.
[14] D. Ting, et al., Neighbor-dependent Ramachandran probability distributions of amino acids developed from a hierarchical Dirichlet process model, PLoS Comput. Biol. 6 (2010) e1000763.
[15] T.T. Wu, E.A. Kabat, An attempt to evaluate the influence of neighboring amino acids $(n-1)$ and $(n+1)$ on the backbone conformation of amino acid (n) in proteins: use in predicting the three-dimensional structure of the polypeptide backbone of other proteins, J. Mol. Biol. 75 (1973) 13–31.
[16] W.D. Cornell, et al., A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[17] A.D. MacKerell, et al., All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.
[18] C. Bystroff, D. Baker, Prediction of local structure in proteins using a library of sequence-structure motifs, J. Mol. Biol. 281 (1998) 565–577.
[19] V. Hornak, et al., Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (2006) 712–725.
[20] A.D. Mackerell, Empirical force fields for biological macromolecules: overview and issues, J. Comput. Chem. 25 (2004) 1584–1604.
[21] P.D. Renfrew, G.L. Butterfoss, B. Kuhlman, Using quantum mechanics to improve estimates of amino acid side chain rotamer energies, Proteins 71 (2008) 1637–1646.
[22] O. Carugo, K. Djinovic-Carugo, Half a century of Ramachandran plots, Acta Crystallogr. D: Biol. Crystallogr. 69 (2013) 1333–1341.
[23] M.I.-H. Tsai, Y. Xu, J.J. Dannenberg, Ramachandran revisited DFT energy surfaces of diastereomeric trialanine peptides in the gas phase and aqueous solution, J. Phys. Chem. B 113 (2009) 309–318.
[24] X. Zhu, P.E.M. Lopes, J. Shim, A.D. MacKerell, Intrinsic energy landscapes of amino acid side-chains, J. Chem. Inf. Model. 52 (2012) 1559–1572.
[25] A. Lupan, A.-Z. Kun, F. Carrascoza, R. Silaghi-Dumitrescu, Performance comparison of computational methods for modeling alpha-helical structures, J. Mol. Model. 19 (2013) 193–203.
[26] Y. Zhao, D.G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function, Theor. Chem. Acc. 120 (2007) 215–241.
[27] B.K. Ho, A. Thomas, R. Brasseur, Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and H-bonding in the alpha-helix, Protein Sci. 12 (2003) 2508–2522.
[28] M. Iwaoka, M. Okada, S. Tomoda, Solvent effects on the φ–ψ potential surfaces of glycine and alanine dipeptides studied by PCM and I-PCM methods, J. Mol. Struct. THEOCHEM 586 (2002) 111–124.
[29] S.A. Hollingsworth, P.A. Karplus, A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins, Biomol. Concepts 1 (2010) 271–283.