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/

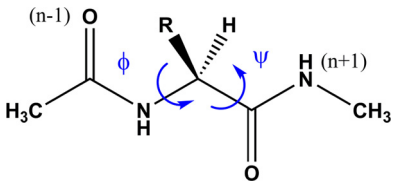

Fig. 1. Scheme of 2-acetamido-N-methylpropanamide structures worked. The φ and ψ angle rotations are pointed out with blue arrows. O (n-1) refers to the oxygen in the position $(n-1)$ where $(n-1)$ is the previous amino acid to the central amino acid (n). In the same way H $(n+1)$ refers to the hydrogen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