Theoretical investigation of Ramachandran plot of N-formyl-L-alanine-amide

Malika Bourjila*, Abderrahmane El Gridani, Rachida Tijar, Brahim El Merbouh and Rachid Drissi El Bouzaidi

Département de chimie, faculté des Sciences, B.P. 8106, Université Ibn Zohr, 80000, Agadir, Maroc
*Corresponding author’s E. mail: malika.bourjila@gmail.com

ABSTRACT

The full conformational space of N-formyl-L-alanine-amide was explored by the semi-empirical method AM1 coupled to the Multi Niche Crowding (MNC) genetic algorithm implemented in a package of programs developed in our laboratory. The structural and energy analysis of the resulting conformational space E(Φ,ψ) exhibits 5 regions or minima γL, γD, εL, δD and αD. The technique provides better detection of local and global minima within a reasonable time.

INTRODUCTION

The conformation adopted by the peptide backbone such as CHO-NH-(R)CaH-CO-NH₂ depends on three angles, Φ, ψ and α. The peptide angle α may be around 0 ° or 180 ° according to cis or trans form of the peptide bond. In addition, it was recognized that the presence of the cis form of the peptide bond is rare because of the unfavorable contact between the amino acid residues (Pauling, 1960 and Ramachandran, 1968). In a previous paper (Bourjila et al., 2016), we were interested in the study of equilibrium structures and acidity in the gas phase of peptides based on Cysteine and Alanine including AlaCysNH₂, Ala:CysNH₂, Ala-CysNH₂, Ala+CysNH₂, CysAlaNH₂ and CysAla+NH₂. Therefore, we concluded that (i) these systems mainly adopt helical forms in their neutral forms, as they are rearranged in globular form in their deprotonated forms (ii) polyalanines acidity increases with increasing the length of the peptide chain (iii) the residue of cysteine is more acidic when it is placed on N-terminal side of the peptide chain. In the present paper, we wish to report the exploration of the conformational space of trans form of N-For-L-Ala-NH₂ (CHO-NH–(CH₂)CH–CONH₂) (Figure 1); define its allowed regions in the Ramachandran map, its global minimum, local minima and their energy and structural properties.

MATERIAL AND METHODS

The molecular potential energy surface PES was explored automatically by the multi-niche crowding genetic algorithm (MNC GA) programmed in our laboratory (El Merbouh, 2014). The genetic algorithm based on the Multi-Niche Crowding (MNC) method is used with the semi-empirical methods AM1. AM1/GA-MNC It's a stochastic technique based on the simulation of a biological system (Cedeneo et al, 1995) and founded on the mechanisms of natural selection and
genetic recombination. The individuals are represented by the conformations, the genes by the dihedral angles; the heat of formation (objective function) was calculated using the semi-empirical method AM1 and through four steps, selection (crowing selection), crossover, mutation and replacement (worst among most similar replacement WAMS) (Fig. 2).

The global minimum and all local minima are determined after several generations. This algorithm is implemented in a package of program interfaced with MOPAC (Stewart, 1989) (version 6.0) in order to evaluate the quality of the individual to insert into the population in each iteration. Table 1 regroups the different control parameters of the algorithm used in this study.

Table 1: Geometrical parameters of genetic algorithm program

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>500</td>
</tr>
<tr>
<td>Crowding selection size (Cs)</td>
<td>25</td>
</tr>
<tr>
<td>Crowding Factor size (Cf)</td>
<td>3</td>
</tr>
<tr>
<td>Intervalcrossoverparameter (ε)</td>
<td>10</td>
</tr>
<tr>
<td>Probability ofcrossover (Pc)</td>
<td>1</td>
</tr>
<tr>
<td>Probability of mutation (Pm)</td>
<td>0.06</td>
</tr>
<tr>
<td>Number of generation</td>
<td>100</td>
</tr>
</tbody>
</table>

genetic recombination. The individuals are represented by the conformations, the genes by the dihedral angles; the heat of formation (objective function) was calculated using the semi-empirical method AM1 and through four steps, selection (crowing selection), crossover, mutation and replacement (worst among most similar replacement WAMS) (Fig. 2).

The global minimum and all local minima are determined after several generations. This algorithm is implemented in a package of program interfaced with MOPAC (Stewart, 1989) (version 6.0) in order to evaluate the quality of the individual to insert into the population in each iteration. Table 1 regroups the different control parameters of the algorithm used in this study.

Table 1: Geometrical parameters of genetic algorithm program

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>500</td>
</tr>
<tr>
<td>Crowding selection size (Cs)</td>
<td>25</td>
</tr>
<tr>
<td>Crowding Factor size (Cf)</td>
<td>3</td>
</tr>
<tr>
<td>Intervalcrossoverparameter (ε)</td>
<td>10</td>
</tr>
<tr>
<td>Probability ofcrossover (Pc)</td>
<td>1</td>
</tr>
<tr>
<td>Probability of mutation (Pm)</td>
<td>0.06</td>
</tr>
<tr>
<td>Number of generation</td>
<td>100</td>
</tr>
</tbody>
</table>


RESULTS AND DISCUSSION

The exploitation of the conformations constituting the population of the last generation of the genetic algorithm used allows to draw the Ramachandran map E=E(Φ, ψ) represented in figure 3. In accordance with the values of the pairs of angles Φ and ψ, this representation is divided into allowed and not allowed regions.

On the Ramachandran map, the first region is denoted γl. It is favored with 34.4% of the 500 conformations constituting the final population. In this case, the values angles Φ and ψ vary on the following range -85° ≤ Φ ≤ -80° and 55° ≤ ψ ≤ 74°. The region-2 for which the value of the angle Φ is between 70 and 80°, the value of angle ψ is between -70 and -50° corresponds to the γD folding and occupies 26.4% of the final population. The most populated region-3 is favored to 37%, it corresponds to the extended structure εL. The values of Φ and ψ varying on the following range: -120° ≤ Φ ≤ -110° and 130° ≤ ψ ≤ 150°. The two less populated regions; region-4 (1.8%) and region-5 (0.4%) correspond to the backbone conformation δD and αD respectively. Their corresponding values of Φ and ψ are between -120° and -110°, -60° and -50° respectively for δD, 80° and 10°, 70° and 72° respectively for αD. The regions corresponding to the backbone conformation βL, δL, αL and ED are not favorable for the system that is the subject of this work.

The five identified regions of the Ramachandran map (Fig. 3) correspond to the five conformations that characterize N-For-L-Ala-NH₂ in the isolated state. Figure 4 and Table 2 are represented the minima highlighted for N-For-L-Ala-NH₂ relative energies (AM1) (kcal/mol), torsion angles (°) and bond length (Å).

A first examination of the structures obtained showed that one hydrogen bond backbone / backbone type may occur in the various conformers of N-For-L-Ala-NH₂; it involves the two amide functions O₅ and N₁₅ H₁₂ and O₇ and N₁₇ H₁₃ (Fig. 4).

Energetically and as shown in Table 2, the minimum γL is found the most stable due to its low energy ΔH = -87.0 kcal/mol. Thus, it is the global minimum of diamide N-For-L-Ala-NH₂ in the isolated state. The minima γD, εL, δL, αL and ED represent local minima classified by ascending order of their energies.

The global minimum γL (C=O)

The diamide N-For-L-Ala-NH₂ adopts the backbone conformation γL as an equilibrium structure. It is stabilized by an intramolecular interaction, typical in the seven-member ring of the γL conformer, the carbonyl oxygen of the N-terminal side interacts with the amide hydrogen on the C-terminal side to form the hydrogen bond O₆...H₈ (2.20 Å).
The local minimum $\gamma_D$ ($C_{\text{ax}}$)

The second backbone conformation whose $\Phi = 75.1^\circ$ and $\psi = -61.8^\circ$ is denoted $\gamma_D$. In this case, the backbone peptide is stabilized by the backbone / backbone C=O···HN-type of hydrogen bond, typical in the 7-membered ring of the $\gamma$ conformer. The corresponding length is of the order of 2.15 Å. The axial position of the side chain (-CH$_3$) decreases the stability of this minimum by 0.5 kcal/mol compared to the global minimum. This may be explained by the steric effect of the methyl group with the 7-membered ring, this effect is not present in the $\gamma_L$ structure.

The local minimum $\epsilon_L$

The extended structure $\epsilon_L$ is characterized by $\Phi = -111.4^\circ$ and $\psi = 145.1^\circ$. It is 1.5 kcal/mol relative to the global minimum $\gamma_L$ (Table 2). Geometrically, an intramolecular interaction involving the carbonyl oxygen of the C-terminal side and the amide hydrogen of the N-terminal side (O$_7$…H$_{11}$), contributes

![Fig. 3: Distribution of conformations constituting the conformational space of N-For-L-Ala-NH$_2$ according to the angles $\Phi$ and $\psi$ of the backbone](image)

![Fig. 4: Minima demonstrated for N-For-L-Ala-NH$_2$ calculated at AM1/GA-MNC](image)
to the stability of this minimum. The obtained hydrogen bond (2.58 Å) leads to the formation of a 5-membered ring C5.

**The local minima δ₀ and α₀**

The least populated regions of the Ramachandran map E(Φ, ψ) are denoted δ₀ (Φ = -120.2° et ψ = -57.1°) and α₀ (Φ = 12.9° et ψ = 63.3°). Geometrically, one hydrogen bond was established in the both cases. It involves the acceptor atomic centers N_h and N_a as well as their amide hydrogens. Their low stability compared with the global minimum (ΔH_f (δ₀) - ΔH_f (γ_L) = 3.2 kcal/mol, ΔH_f (α₀) - ΔH_f (γ_L) = 6.0 kcal/mol), is explained by the low character of the intramolecular interaction highlighted in these two structures, the corresponding bond length ≈ 2.82 Å.

Several groups have carried out the study of the Ramachandran map of N-for-Ala-NH2. In Table 3 are represented the number and type of minima highlighted for N-for-Ala-NH2 by using different methods and basis of theoretical calculation.

From the analysis of Table 3, it is clear that the number and the type of minima highlighted depend on the precision of the method of calculation used. Thus, using the semi-empirical methods, seven and four minima are located by using PM3 and AM1 respectively; four of them (γ_L, γ_D, δ₀ and ε_D) are common between these two methods. At the ab initio approach, we note that the number of minima detected decreases with the increase of the basis of calculation at both HF and B3LYP levels of theory.

**CONCLUSIONS**

Conformational analysis of trans-N-For-L-Ala-NH₂ using the MNC genetic algorithm coupled to the semi-empirical method AM1 was performed and 5 regions or folds were located based on the mode of rearrangement of the backbone, namely γ_L, γ_D, ε_L, δ₀, and α₀.

The γ_L structure is found as the most stable conformation along the conformational space of N-For-L-Ala-NH₂ at our technical computing, and is the same as the one detected by all methods of theoretical calculation cited in Table 3, except for the semi-empirical method PM3.

A comparison of our results from the AM1/GA-MNC computation with those of the literature using the same method of calculation AM1 shows the capacity of the algorithm to locate all allowed minima; the Bibliographic result from the AM1 is devoid of the minimum α₀.

The technique (AM1/GA-MNC) provides better detection of local and global minima within a reasonable range of stability.
time; it will be used to calculate the molecular PES of peptides of larger size.

REFERENCES


Ramachandran, G.N., 1968. Stereochemical criteria for polypeptides and proteins. IV. Standard dimensions for the cis-peptide unit and conformation of cis-polypeptid, Biopolymers 6, 1255-1262
