International Scientific Organization
<http://iscientific.org/>
 Chemistry International
www.bosaljournals.com/chemint/

Theoretical study of the potential energy surface of N-formyl-L-cysteine-amide by using a genetic algorithm multi niche crowding

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, Morocco

*Corresponding author's E. mail: malika.bourjila@gmail.com

ARTICLE INFO

Article type:

Research article

Article history:

Received May 2018

Accepted November 2018

April 2019 Issue

Keywords:

Conformational space

Cysteine

AM1

Genetic algorithm

ABSTRACT

The generation of molecular structures constituting the conformational space of trans-N-For-L-Cys-NH₂ was accomplished using the genetic algorithm MNC coupled to the semi-empirical AM1 method, AM1/GA-MNC. The structural and energy analysis of the obtained conformational space $E=E(\varphi, \psi)$ locates 7 regions or minima γ_L , γ_D , β_L , δ_D , δ_L , ϵ_D and ϵ_L . The combination of these folds to structuring modes adopted by the side chain CH₂-SH has allowed us to identify 27 stable geometric structures. The regions corresponding to helical folds α_D and α_L are not favorable for the system that is the subject of current study.

© 2019 International Scientific Organization: All rights reserved.

Capsule Summary: Molecular structure of trans-N-For-L-Cys-NH₂ was studied using the genetic algorithm MNC coupled to the semi-empirical AM1 method, AM1/GA-MNC and combination of the folds to structuring modes adopted by the side chain CH₂-SH revealed the possibility of 27 stable geometric structures of N-For-L-Cys-NH₂.

Cite This Article As: M. Bourjila, A. El Gridani, R. Tijar, B. El Merbouh and R. D. El Bouzaidi. Theoretical study of the potential energy surface of N-formyl-L-cysteine-amide by using a genetic algorithm multi niche crowding. Chemistry International 5(2) (2019) 132-142. <https://doi.org/10.5281/zenodo.1482898>

INTRODUCTION

For several years, many theorists and experiments focus their efforts to analyze and understand the topology of the potential energy surfaces of polypeptides and proteins. The Ramachandran Map (Ramachandran et al., 1963) is a theoretical strategy that appeared in the 60 (Ramachandran, 1968) to explain through a simple map of interactions (interatomic repulsion and attraction) the conformations adopted by amino acids in a time when the crystallographic data are limited. It is a graphical representation associated with the behavior of the amino acid residues included in a peptide chain as a function of the φ and ψ angles describing

the structure mode of the main chain or backbone peptide. The conformational analysis of potential energy surfaces of diamides models such as R'CONH-CHR-CNOR'H is being developed. Thus, the conformational properties of the polypeptides and proteins can be deduced from those of the blocked or protected units.

In recent years, several theoretical approaches (Baldoni et al., 1999; Head-Gordan et al., 1989; 1991; McAllister et al., 1998; Perczel et al., 1991; Rodriguez et al., 1998;) have been used to study two-dimensional potential energy surface $E(\varphi, \psi)$ of CH₃COHN-CHR-CONHCH₃ or HCONH-CHR-CONH₂. The multi-dimensional conformational analysis process (MDCA) states (Perczel et al., 1991) the existence of nine minimum, designated by Greek letters

attached to L or D to denote the chirality of the conformer. These codes are noted γ_L , γ_D , β_L , δ_L , α_D , α_L , ϵ_D , ϵ_L and δ_D (Fig. 1,a). For amino acid residues that have complex side chains, analyzing their conformational space impose conformational study of two sub-spaces; one that describes the mode of backbone rearrangement $E(\varphi, \psi)$ and the two-dimensional space $E(\chi_1, \chi_2)$ that defines the structure mode of the side chain and which is governed by χ_1 et χ_2 angles (torsion angles around C α -C β and C β -X bonds respectively). Each twist angle characterizing the side chain should have 3 positions (g+, a, g-). Thus, the potential energy surface $E=E(\chi_1, \chi_2)$ will be defined in the ideal case by 9 conformations Figure (1,b).

Serine and cysteine are amino acids which have the simplest side chains. Thus, for the trans-For-L-Ser-NH₂ (Farkas et al., 1995; Perczel et al., 1995; 1996a; 1996b; 1998; 2003), over 81 conformations were defined by the MDCA. This number was reduced to 44 and 36 structures if methods RHF/3-21G and RHF/6-31++G(d,p) respectively are used. In this case, the most favored structure (global minimum) is the $\gamma_L(g+ g+)$ followed by C5 conformation [$\beta_L(a g+)$] which is 2.10 kcal/mol relative to $\gamma_L(g+ g+)$ (Perczel et al., 2003).

The tetra-dimensional Ramachandran surface of Cysteine is defined by four variables $E = E(\varphi, \psi, \chi_1, \chi_2)$. Given its polar character, the functional group SH may be donor as it can be proton acceptor. In fact, several types of intramolecular interactions are going to be at the heart of discussions such as backbone-backbone interaction and/or side-chain backbone.

Ab initio and DFT Studies (Bombasaro et al., 2005; Zamora et al., 2001) have been devoted to the study of the conformational space of N-acetyl-L-Cystéine-N-methylamide, while for N-Formyl-L-Cys-amide and according to our knowledge, the data are still limited (Zamora et al., 2002).

The conformation adopted by the peptide backbone defined by the HCO-NH-CH α -CONH₂ fragment depends on three angles φ , ψ and ω . The peptide angle may be around 0° or 180° according to Cis or Trans form of the peptide. 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; Ramachandran., 1968). Indeed, we are interested in studying the trans form of N-For-L-Cys-NH₂ by fixing OCNC angle near 0°.

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, Ala₂CysNH₂, Ala₃CysNH₂, Ala₄CysNH₂, CysAlaNH₂ et CysAla₂NH₂. 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. The objective of this work is the structural and energy analysis of minimum constituting the conformational space $E=E(\varphi, \psi)$ of Trans-N-For-L-Cys-NH₂ in isolated state.

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 et al., 2014). It's a stochastic techniques based on the simulation of a biological system. These are based on the mechanisms of natural selection and genetic recombination. 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.

RESULTS AND DISCUSSION

The exploitation and exploration of the conformational space of the blocked diamide of cysteine HCO-NH-L-CH(CH₂-SH)CO-NH₂ using the genetic algorithm, based on the MNC method, coupled with the semi-empirical AM1 method, allows to graphically represent the two-dimensional sub-space $E = E(\varphi, \psi)$ as shown in Figure 3. The representation is done in two-dimensional space [-180°, 180°] according to the convention of the IUPAC-IUB.

An initial examination of the obtained conformational surface (Fig. 3) shows that the 500 conformations of the final population is grouped in 7 niches or permitted regions. Each of these regions corresponds to a type of folding or minimum (γ_L , γ_D , α_L) based on the values of φ and ψ angles that describe the structure mode adopted by the main chain or backbone.

The first region (γ_L) is a region where a large number of conformations is grouped as a niche, and it is favored with 29% of 500 conformations constituting the final population. In this region, the backbone of N-For-L-Cys-NH₂ diamide adopts a structure γ -turn by forming a C_{7^{eq}} conformation. The second region occupies 14% of 500 conformations constituting the population, it is denoted (γ_D). In this case, the side chain-CH₂-SH, denoted C_{7^{ax}}, is in axial position relative to the backbone. The third area denoted (β_L) occupies 38% in the two edges of the left side of the obtained Ramachandran surface. In this region the peptide backbone adopts an extended structure denoted C5. The fourth and the fifth region denoted ϵ_L and ϵ_D occupy 8% and 4.8% respectively. The last regions denoted δ_D and δ_L occupy 5.2% and 1% respectively. The backbones of the two conformations are stabilized mainly by hydrogen bonds involving the two amide functions O₆C₅N₄H₁₂ and O₇C₂N₁H₉. Each identified minimum is associated with the rearrangement mode adopted by the groups of the side chain -CH₂-SH since each torsional mode (χ_1 and χ_2) is expected to have three minima (g+, a, g-).

The γ_L backbone conformation

The backbone $\gamma_L(\varphi \approx -82.8^\circ, \psi \approx 64.4^\circ)$ is stabilized by two hydrogen bonds such backbone/backbone involving the two amide functions O₆C₅N₄H₁₂ and O₇C₂N₁H₉.

Fig. 1: (a,b) : Idealized PES topology for a single amino acid residue PCO-NH-CHR-CO-NHQ (MDCA). The central square, marked by broken lines, specifies the cut of the PES accepted by IUPAC convention: $-180^\circ \leq \phi \leq +180^\circ$ and $-180^\circ \leq \psi \leq +180^\circ$ (Perczel et al., 1991).

Fig. 2: Numbering atoms of N-Formyl-L-Cys-amide

The oxygen (O_6) of carbonyl bonded to the amide nitrogen N_4 interacts with one of the hydrogens carried by the heteroatom N_1 . Thus, a hydrogen bond $O_6 \dots H_9$ with a medium length of 2.18 Å was formed. On the other hand, nitrogen N_4 comes into interaction of weak character (2.94 Å) with the same hydrogen $N_4 \dots H_9$ (Fig. 4).

By analyzing the different conformations γ_L (Fig. 4), three hydrogen bonds such side-chain backbone are conceivable depending on the values of χ_1 and χ_2 angles. By combining the type of rearrangement adopted by the side chain $-\text{CH}_2\text{-SH}$ in the peptide backbone γ_L , 5 conformations were identified namely $\gamma_L[\text{g}+\text{g}+]$, $\gamma_L[\text{g}-\text{g}-]$, $\gamma_L[\text{a}-\text{g}-]$, $\gamma_L[\text{g}-\text{g}+]$ and $\gamma_L[\text{g}-\text{a}]$ (Table 1). Geometrically, in addition to the two

deterministic hydrogen bonds ($O_6 \dots H_9$, $N_4 \dots H_9$) of backbone γ_L , the sulfur can be an atomic center donor and / or proton acceptor.

In terms of energy, as shown in Table 1 where are represented energy differences between different structures γ_L , we find that the $\gamma_L[\text{g}+\text{g}+]$ is the most favored by representing the global minimum from the AM1 of N-For-L-Cys-NH₂ in the isolated state. The structures $\gamma_L[\text{g}-\text{g}-]$, $\gamma_L[\text{a}-\text{g}-]$, $\gamma_L[\text{g}-\text{g}+]$ and $\gamma_L[\text{g}-\text{a}]$ are less stable and are 1.6, 1.9, 2.3 and 2.5 kcal/mol respectively, compared to the global minimum. This order of stability can be justified in discussing the hydrogen bonds involving the functional groups of the side chain $-\text{CH}_2\text{-SH}$ in each structure (Fig. 4).

Fig. 3: Distribution of conformations constituting the conformational space of Trans-N-For-L-Cys-NH₂ according to the angles ϕ and ψ of the backbone, AM1/GA-MNC

Fig. 4: γ_L minima of N-For-L-Cys-NH₂ obtained by AM1/GA-MNC, hydrogen bonds length are in Å

Fig. 5: γ_D minima of N-For-L-Cys-NH₂ calculated by using AM1/GA-MNC, hydrogen bonds length are in Å

Fig. 6: β_L minima obtained for N-For-L-Cys-NH₂ by using AM1/GA-MNC, hydrogen bonds length are in Å

Fig. 7: ϵ_L minima calculated for N-For-L-Cys-NH₂ by using AM1/GA-MNC, hydrogen bonds length is in Å

Fig. 8: ϵ_D minima calculated for the N-For-L-Cys-NH₂ using AM1/GA-MNC, hydrogen bonds length are in Å

The comparison of mean values of angles φ and ψ of backbone γ_L found in this study with their counterparts from other theoretical calculation methods and for other diamides (Table 2) shows that the used research technique (AM1/GA-MNC) well locates γ_L conformation at values of

angles that are in perfect agreement with those found either by the same method AM1 or those from the ab initio. Moreover, we can say that the increase in the size of the side chain -R linked to the alpha carbon does not lead to a geometric modification of the backbone γ_L .

Fig. 9: δ_D conformations obtained for the N-For-L-Cys-NH₂ using AM1/GA-MNC, hydrogen bonds length are in Å

Fig. 10: The $\delta_L[g+g-]$ conformation calculated for N-For-L-Cys-NH₂, AM1/GA-MNC, hydrogen bonds length are in Å

The γ_D backbone conformation

The second allowed region characterizing the conformational space of N-For-L-Cys-NH₂ is denoted γ_D ($\varphi \approx 73.8$, $\psi \approx -61.5$) (C_{7ax}). Structurally, the peptide backbone of γ_D conformation is stabilized by the same bonds of backbone-backbone type as those of γ_L ($N_4 \dots H_9$ et $O_6 \dots H_9$). The corresponding lengths are of the order of 2.15

and 2.95 Å, respectively. The backbone γ_D wherein the side chain is in the axial position is less stable than γ_L where the $-CH_2-SH$ is in the equatorial position relative to the peptide backbone. By analyzing the space $E_{\gamma_D}(\chi_1, \chi_2)$, 6 conformations were located, $\gamma_D[g-g+]$, $\gamma_D[g-g-]$, $\gamma_D[g-a]$, $\gamma_D[ag-]$, $\gamma_D[ag+]$ and $\gamma_D[g+g-]$ (Fig. 5).

The energies barriers noticed between γ_D structures result in small geometric changes. In a previous paper (Bourjila et al., 2017), We carried out a conformational study on the preferred minima γ_D (C_{7ax}) of trans N-For-L-Cys-NH₂ in isolated state by AM1/GA-MNC calculation. All conformations located were subjected to geometry optimization at the HF/3-21G and the HF/6-31G++ (d, p) levels of theory. The three levels of theory provide the same results.

The β_L backbone conformation

The third area populated by the diamide N-For-L-Cys-NH₂ is located in both sides of the left side of the obtained Ramachandran surface ($\varphi \approx -150.1$ and $\psi \approx 177.2$) (figure 3). In this region denoted β_L , were gathered conformations which have an expanded geometry C_5 where the peptide backbone arranges in a manner to create interaction between hydrogen amide carried by the nitrogen N4 and oxygen of carbonyl bonded to the alpha carbon ($O_7 \dots H_{12}$). The length of the hydrogen bond thus formed (deterministic folding β_L) is around 2.41 ± 5 Å. By analyzing the three possible positions of functional groups SH of the side chain according to the values of the corresponding angles χ_1 and χ_2 , 4 conformations β_L may be adopted. In Table 3, we have grouped the energies gaps of β_L structures relative to the global minimum $\gamma_L[g+g+]$ and the values of the corresponding angles φ , ψ , χ_1 and χ_2 . The β_L structures are less stable than the global minimum $\gamma_L[g+g+]$. The energy gap increases from one conformation to another Table 3. This can be explained by analyzing the intramolecular interactions stabilizing each structure of Figure 6. As regards the conformation $\beta_L[a+g+]$, the functional group SH is arranged in a way to make its acceptor atomic center S in interaction of strong character (2.44 Å) with the hydrogen carried by the nitrogen N₁ ($S \dots H_9$). The conformation $\beta_L[g+a]$, no binding is provided by the side chain $-CH_2-SH$ which reduces its stability of 6.7 kcal/mol relative to the global minimum.

The ϵ_L backbone conformation

The ϵ_L conformation results from a rotation around the $-C\alpha-CO-$ bond of β_L backbone. Thus, the obtained folding defines the fourth area occupied by the diamide N-For-L-Cys-NH₂ in an isolated state.

By combining the rearrangement mode of the side chain to ϵ_L folding, 3 conformations were located namely $\epsilon_L[g+g+]$, $\epsilon_L[a+g-]$ and $\epsilon_L[a+a]$. Their corresponding geometric parameters are grouped in the Table 4.

Table 1: Relatives energies in (kcal/mol) and angles values of γ_L minima calculated for N-For-L-Cys-NH₂. AM1/GA-MNC

Backbone [sidechain]	$\Delta\Delta H_f$	$\varphi(^{\circ})$	$\psi(^{\circ})$	$\chi_1(^{\circ})$	$\chi_2(^{\circ})$
$\gamma_L[g+g+]$	0.00	-82.4	62.8	44.8	73.2
$\gamma_L[g-g-]$	1.6	-82.7	68.0	-45.9	-63.7
$\gamma_L[a g-]$	1.9	-82.7	65.5	-160.0	-76.8
$\gamma_L[g-g+]$	2.3	-80.8	61.9	-72.8	63.4
$\gamma_L[g-a]$	2.5	-85.4	63.8	-57.9	-177.8

Table 2: Comparison of angles values (φ , ψ) of γ_L backbone (obtained in this work) with those of N-For-L-Gly-NH₂, N-For-L-Ala-NH₂ and N-For-L-Ser-NH₂ ^{a,b}

Systems	$\varphi(^{\circ})$	$\psi(^{\circ})$
N-For-L-Gly-NH ₂ (HF/3-21G)	-83.9	67.8
N-For-L-Ala-NH ₂ (AM1)	-82.2	65.7
N-For-L-Ala-NH ₂ (HF/3-21G)	-84.5	67.3
N-For-L-Ser-NH ₂ (HF/3-21G)	-82.2	67.4
N-For-L-Cys-NH ₂ (AM1/GA-MNC)	-82.8	64.4

^a Rodriguez et al. (1998), ^b Perczel et al. (1996)

Table 3: Relatives energies and angles values characterizing the minimum β_L obtained from AM1/GA-MNC calculation

Backbone [sidechain]	$\Delta\Delta H_f(\text{kcal/mol})$	$\varphi(^{\circ})$	$\psi(^{\circ})$	$\chi_1(^{\circ})$	$\chi_2(^{\circ})$
$\beta_L[a g+]$	2.6	-150.5	167.8	-156.6	73.0
$\beta_L[g- g+]$	3.8	-151.0	157.6	-111.3	57.1
$\beta_L[g+g-]$	5.1	-134.7	170.2	67.3	-60.1
$\beta_L[g+a]$	6.7	-157.5	167.2	64.7	-160.9

The $\epsilon_L[g-g+]$, $\epsilon_L[a g-]$ and $\epsilon_L[a a]$ structures are 3.5, 4.2 and 4.3 kcal/mol, respectively relative to minimum global. Geometrically, in addition to the hydrogen bonding (O₇...H₁₂) common among the three structures, a second interaction is generated by the SH group as shown in Figure 7.

Regarding conformations $\epsilon_L[a g-]$ and $\epsilon_L[a a]$, the same geometrical structure is adopted. However, in $\epsilon_L[a a]$ conformation, the acidic hydrogen of the thiol function is rotated around the S-CH₂ bond. The corresponding dihedral angle χ_2 varies from -98.3 ° to -140.5 °.

The ϵ_D backbone conformation

Symmetrically with ϵ_L conformation, the minimum denoted ϵ_D appears in the right side of the obtained Ramachandran surface. The ϵ_D backbone adopts an extended structure favored by the formation of hydrogen bonds involving hydrogen amide H₁₂ and the carbonyl oxygen bonded to the alpha carbon (O₇...H₁₂).

According to the values of χ_1 and χ_2 angles that define the type of rearrangement adopted by the side chain -CH₂-SH, 4 conformations have been associated with

folding ϵ_D namely $\epsilon_D[a g+]$, $\epsilon_D[a a]$, $\epsilon_D[g-g+]$ and $\epsilon_D[g-g-]$ (Table 5 is shown energy gaps of structures with respect to the global minimum $\gamma_L[g+g+]$ and the values of corresponding dihedral angles.

The structures of the Table 5 are classed in ascending order of their formation energies. The energy deviation from the global minimum increase of 5.6 kcal/mol for $\epsilon_D[a g+]$ to 9.7 kcal/mol for $\epsilon_D[g-g-]$. This order of stability may be due to effect (steric or electronic) which can be generated by the functional groups of the side chain. Highlighted minimum ϵ_D for the N-For-L-Cys-NH₂ are grouped in Figure 8.

Structural analysis of the different conformations of Figure 8 shows that at most both types of intramolecular interactions are possible. In addition to hydrogen bonding (O₇...H₁₂) characterizing the ϵ_D backbone, the first two conformations $\epsilon_D[a g+]$ and $\epsilon_D[a a]$ are stabilized by a second hydrogen bond (average length of 2.36 Å) which involves the acceptor atomic center S of the thiol group and one of the hydrogens carried by N₁ (S...H₉). These two structures adopt similar geometries. However, the energy barrier of 0.4 kcal/mol is likely due to the redirection of the acidic hydrogen in $\epsilon_D[a a]$.

Table 4: Energies deviations from the global minimum and angles characterizing the minimum ϵ_L of N-For-L-Cys-NH₂, AM1/GA-MNC

Backbone [sidechain]	$\Delta\Delta H_f$ (kcal/mol)	$\varphi(^{\circ})$	$\psi(^{\circ})$	$\chi_1(^{\circ})$	$\chi_2(^{\circ})$
$\epsilon_L[g-g+]$	3.5	-102.5	143.7	-72.8	60.7
$\epsilon_L[a g-]$	4.2	-120.5	155.0	-165.5	-98.3
$\epsilon_L[a a]$	4.3	-122.0	159.2	-157.0	-140.5

Table 5: Energies gaps and angles characterizing the minima ϵ_D of N-For-L-Cys-NH₂, obtained from AM1/GA-MNC calculation

Backbone [sidechain]	$\Delta\Delta H_f$ (kcal/mol)	$\varphi(^{\circ})$	$\psi(^{\circ})$	$\chi_1(^{\circ})$	$\chi_2(^{\circ})$
$\epsilon_D[a g+]$	5.6	93.5	-167.9	-155.2	65.8
$\epsilon_D[a a]$	6.0	91.2	-170.5	-149.3	-174.7
$\epsilon_D[g-g+]$	7.0	99.9	161.2	-71.8	83.0
$\epsilon_D[g-g-]$	9.7	77.6	170.0	-63.4	-58.8

Table 6: Energies deviations from the global minimum $\gamma_L[g+g+]$ and the values of the angles φ , ψ , χ_1 and χ_2 of the 4 minimum δ_D of N-For-L-Cys-NH₂ system. AM1/GA-MNC

Backbone [side chain]	$\Delta\Delta H_f$ cal/mol	$\varphi(^{\circ})$	$\psi(^{\circ})$	$\chi_1(^{\circ})$	$\chi_2(^{\circ})$
$\delta_D[g+g-]$	3.8	-130.2	-60.3	51.9	-76.9
$\delta_D[g+a]$	4.8	-127.8	-57.3	49.2	156.6
$\delta_D[g-g+]$	5.9	-131.8	-65.4	-81.5	25.6
$\delta_D[ag+]$	6.2	-133.3	-67.8	173.3	58.2

The δ_D backbone conformation

The δ_D region constitutes the sixth populated area of the Ramachandran surface $E = E(\varphi, \psi)$ of N-For-L-Cys-NH₂ system. In all δ_D structures, the amide hydrogen of the peptide bond H₁₂N₄C₅O₆ is attracted by the acceptor atomic center N₁ of the second amide function. Thus, a hydrogen bond N₁...H₁₂ of backbone-backbone type and with an average length 2.78 Å was established. By combining the geometric shape adopted by the -CH₂-SH fragment to fold δ_D 4 conformations are at the heart of discussions as shown in Figure 9.

Structurally, the two structures $\delta_D[g+g-]$ and $\gamma_L[g+g+]$ undergo completely different geometrical rearrangements. Thus, the formation of C7 ring and the equatorial position of the side chain -CH₂-SH make the conformation $\gamma_L[g+g+]$ less hindered and favor the formation of strong hydrogen bonds by two thiol.

In Table 6 are shown the energies deviations from the global minimum $\gamma_L[g+g+]$ and the values of the angles φ , ψ , χ_1 et χ_2 of the 4 minimum δ_D of N-For-L-Cys-NH₂ system.

In the structure $\delta_D[g+g-]$, the position of the -CH₂-SH fragment relative to δ_D backbone leads to the interaction of the thiol with hydrogen H₉ (S...H₉) (2.43 Å) with the lone pair of the nitrogen amide N₄ (N₄...H₁₇) (2.96 Å, which increase the effect of the steric hindrance, and consequently the decrease in stability of 3.8 kcal / mol.

From a conformation to another, the energy barrier ΔH_f (δ_D)- ΔH_f ($\gamma_L[g+g+]$) increases until it reaches 6.2 kcal / mol.

The δ_L backbone conformation

The least populated region (1%) of Ramachandran surface of N-For-L-Cys-NH₂ is denoted δ_L ($\varphi = -114.7$, $\psi = 16.7$). As for the δ_D conformation, the backbone δ_L is stabilized by hydrogen bonding (N₄...H₉) between the two amide functions. The corresponding length is of the order of 2.57 Å. The side chain -CH₂-SH rearranges in a manner to create interaction of strong character (2.59 Å) between the sulfur and the amide hydrogen (S...H₁₂). The values of χ_1 and χ_2 angles are of about 37.8 and -91.7°, respectively. Energetically the $\delta_L[g+g-]$ conformation ($\Delta H_f = -76.2$ kcal/mol) is 5.1 kcal / mol relative to the global minimum $\gamma_L[g+g+]$. This can be justified by comparing the number of interactions stabilizing each minima (Fig. 10).

The structure $\gamma_L[g+g+]$ is the most stable found along the surface conformation of N-For-L-Cys-NH₂. According to our knowledge, a conformational and electronic study of energetically favored conformations of N-For-L-Cys-NH₂ diamide at HF/3-21G and B3LYP/6-31G(d,p) (Zamora et al., 2002), shows that the most stable structure (global minimum) of this system adopts the structure $\gamma_L[g+g+]$. In this work we found the same result with our technical computing AM1/GA-MNC. Results

revealed that this techniques is useful for the generation of molecular structures constituting the conformational space of protein and could possibly be extended for the determination of molecular structures of other related molecules (Bora et al., 2019; Custódio et al., 2014; Daeyaert et al., 2005; Fu et al., 2018; Kaur and Kadam, 2018; Lei et al., 2018; Lucasius and Kateman, 1994; Luke, 1996; Song et al., 2018; Tian et al., 2018; Vitela and Castaños, 2012).

CONCLUSIONS

The conformational analysis of trans-N-For-L-Cys-NH₂ using genetic algorithm MNC coupled to the semi-empirical calculation method AM1 was studied and 7 regions or folds were located based on the mode of rearrangement of backbone, namely $\gamma_L, \gamma_D, \beta_L, \delta_D, \delta_L, \epsilon_D$ and ϵ_L and 27 have been identified if the rearrangement mode is associated of the side chain to these minima. The minimum corresponding to helical folds α_D and α_L are not favorable for this system.

REFERENCES

- Baldoni, H.A., Rodriguez, A.M., 1999. Peptide models XXIV: An ab initio study on N-formyl-L-prolinamide with trans peptide bond. The existence or non-existence of (3L) and (3L) conformations. *Journal of Molecular Structure (Theochem)* 465, 79–91.
- Bombasaro, J. A., Zamora, M. A., Baldoni, H. A., Enriz, R. D., 2005. An Exhaustive Conformational Analysis of N-Acetyl-L-cysteine-N-methylamide. Identification of the Complete Set of Interconversion Pathways on the ab Initio and DFT Potential Energy Hypersurface. *Journal of the American Chemical Society* A 109, 874-884.
- Bora, T.C., Mariani, V.C., Coelho, L.d.S., 2019. Multi-objective optimization of the environmental-economic dispatch with reinforcement learning based on non-dominated sorting genetic algorithm. *Applied Thermal Engineering* 146, 688-700.
- Bourjila, M., El Merbouh, B., Tijar, R., EL Guerdaoui, A., Drissi El Bouzaidi, R., El Gridani, A., EL Mouhtadi, M., 2016. Polyalanine gas phase acidities determination and conformational space analysis by genetic algorithm assessment. *Chemistry International* 2(3), 145-167.
- Bourjila, M., Tijar, R., El Merbouh, B., El Guerdaoui, A., A., Drissi El Bouzaidi, R., El Gridani, A., 2017. Conformational Properties of N-Formyl-L-Cys-NH₂ in Its γ_D -Backbone Conformations as Computed by Semi-Empirical and *Ab Initio* Methods. *Journal of Chemical and Pharmaceutical Research* 9(2), 176-180.
- Custódio, F.L., Barbosa, H.J.C., Dardenne, L.E., 2014. A multiple minima genetic algorithm for protein structure prediction. *Applied Soft Computing* 15, 88-99.
- Daeyaert, F., de Jonge, M., Heeres, J., Koymans, L., Lewi, P., van den Broeck, W., Vinkers, M., 2005. Pareto optimal flexible alignment of molecules using a non-dominated sorting genetic algorithm. *Chemometrics and Intelligent Laboratory Systems* 77, 232-237.
- El Merbouh, B., Bourjila, M., Tijar, R., El Bouzaidi, R. D., El Gridani, A., EL Mouhtadi, M., 2014. Conformational space analysis of neutral and protonated glycine using a genetic algorithm for multi-modal search. *Journal of Theoretical and Computational Chemistry* 13(8), 1450067.
- Farkas, O., Perczel, A., Marcoccia, J., Hollosi, M., Csizmadia, I.G., 1995. Peptide models XIX: Side-chain conformational energy surface $E = f(\chi_1, \chi_2)$ and amide I vibrational frequencies of N-formyl-L-phenylalaninamide (For-Phe-NH₂) in its γ_L or γ^{inv} or C_7^{eq} backbone conformation. *Journal of Molecular Structure* 331, 27-36.
- Fu, X., Lei, L., Yang, G., Li, B., 2018. Multi-objective shape optimization of autonomous underwater glider based on fast elitist non-dominated sorting genetic algorithm. *Ocean Engineering* 157, 339-349.
- Head-Gordan, T., Head-Gordan, M., Frisch, M.J., Brooks III, C., Pople, J., 1989. A theoretical study of alanine dipeptide and analogs. *International Journal of Quantum Chemistry* 16, 311-322.
- Head-Gordon, T., Head-Gordon, M., Frisch, M.J., Brooks III, C., Pople, J., 1991. Theoretical study of blocked glycine and alanine peptide analogs. *Journal of the American Chemical Society* 113, 5989-5997.
- Kaur, M., Kadam, S., 2018. A novel multi-objective bacteria foraging optimization algorithm (MOBFOA) for multi-objective scheduling. *Applied Soft Computing* 66, 183-195.
- Lei, X., Zhang, J., Wang, H., Wang, M., Khu, S.-T., Li, Z., Tan, Q., 2018. Deriving mixed reservoir operating rules for flood control based on weighted non-dominated sorting genetic algorithm II. *Journal of Hydrology* 564, 967-983.
- Lucasius, C.B., Kateman, G., 1994. Understanding and using genetic algorithms Part 2. Representation, configuration and hybridization. *Chemometrics and Intelligent Laboratory Systems* 25, 99-145.
- Luke, B.T., 1996. An Overview of Genetic Methods, in: Devillers, J. (Ed.), *Genetic Algorithms in Molecular Modeling*. Academic Press, London, pp. 35-66.
- McAllister, M.A., Perczel, A., Csizsar, P., Viviani, W., Rivail, J.L., and Csizmadia, I. G., 1993. Peptide models 4. Topological features of molecular mechanics and ab initio 2D-ramachandran maps: Conformational data for for-Gly-NH₂, for-L-Ala-NH₂, Ac-L-Ala-NHMe and for-L-Val-NH₂. *Journal of Molecular Structure* 288, 161-179.

- Perczel, A., Csizmadia, I.G., 1995. Searching for the simplest structural units to describe the three-dimensional structure of proteins. *Journal International Reviews in Physical Chemistry* 14, 127-136.
- Perczel, A., Angyan, J.G., Kajtar, M., Viviani, W., Rivail, J.L., Marcoccia, J.F., Csizmadia, I.G., 1991. Peptide Models. 1. Topology of Selected Peptide Conformational Potential Energy Surfaces (Glycine and Alanine Derivatives). *Journal of the American Chemical Society* 113, 6256-6265.
- Perczel, A., Farkas, O., Csizmadia, I. G., 1996. *Journal of Computational Chemistry* 17, 821-834.
- Perczel, A., Farkas, O., Csizmadia, I. G., 1996. Peptide Models 18. Hydroxymethyl Side-Chain Induced Backbone Conformational Shifts of L-Serine Amide. All ab Initio Conformers of For-L-Ser-NH₂. *Journal of the American Chemical Society* 118, 7809-7817.
- Perczel, A., Farkas, O.J., akli, I., Csizmadia, I.G., 1998. Peptide models XXI. Side-chain/backbone conformational interconversions in HCO-l-Ser-NH₂. Tracing relaxation paths by ab initio modeling. An exploratory study. *Journal of Molecular Structure* 455, 315-338.
- Perczel, A., Fuzery, A.K., Csaszar, A. G., 2003. Toward Direct Determination of Conformations of Protein Building Units from Multidimensional NMR Experiments. V. NMR Chemical Shielding Analysis of N-formyl-serinamide, a Model for Polar Side-Chain Containing Peptides. *Journal of Computational Chemistry* 24, 1157-1171.
- Ramachandran, G.N., 1968b. Conformation of polypeptides and proteins. *Biopolymers* 6, 1255-1262.
- Ramachandran, G.N., Ramakrishnan, C and Sasisekharan, V., 1963a. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* 7, 95-99.
- Rodriguez, A.M., Baldoni, H.A., Suvire, F., Vazquez, R. N., Zamarbide, G., Enriz, R.D., Farkas, O., Perczel, A., McAllister, M.A., Torday, L.L., Papp, J.G., Csizmadia, I. G., 1998. Characteristics of Ramachandran maps of l-alanine diamides as computed by various molecular mechanics, semiempirical and ab initio MO methods. A search for primary standard of peptide conformational stability. *Journal of Molecular Structure (Theochem)* 455, 275-301
- Song, S., Gao, S., Chen, X., Jia, D., Qian, X., Todo, Y., 2018. AIMOES: Archive information assisted multi-objective evolutionary strategy for ab initio protein structure prediction. *Knowledge-Based Systems* 146, 58-72.
- Stewart, J.J.P., 1989. Optimization of parameters for semiempirical methods I. *Journal of Computational Chemistry* 10, 209-220.
- Tian, Z., Yan, Y., Hong, Y., Guo, F., Ye, J., Li, J., 2018. Improved genetic algorithm for optimization design of a three-dimensional braided composite joint. *Composite Structures* 206, 668-680.
- Vitela, J.E., Castaños, O., 2012. A sequential niching memetic algorithm for continuous multimodal function optimization. *Applied Mathematics and Computation* 218, 8242-8259.
- Zamora, M.A., Baldoni, H.A., Bombasaro, J.A., Mak, M.L., Perczel, A., Farkas, O., Enriz, R.D., 2001. An exploratory ab initio study of the full conformational space of N-acetyl-l-cysteine-N-methylamide. *Journal of Molecular Structure* 540, 271-283.
- Zamora, M. A., Baldoni, H.A., Rodriguez, A.M., Enriz, R.D., Sosa, C.P., Perczel, A., Kucsman, A., Farkas, O., Deretey, E., Vank, J.C., Csizmadia, I.G., 2002. Peptide model XXVIII: An exploratory ab initio and density functional study on the side-chain- backbone interaction in N-acetyl-L-cysteine-N-methylamide and N-formyl-L-cysteinamide in their γ_L -backbone conformations. *Journal of Chemistry* 80, 832-844.

Visit us at: <http://bosajournals.com/chemint/>
Submissions are accepted at: editorci@bosajournals.com