Eur. Phys. J. D (2017) 71: 203
DOI: 10.1140/epjd/e2017-80187-5

THE EUROPEAN PHYSICAL JOURNAL D

Regular Article

Conformational landscape of isolated capped amino acids: on the nature of non-covalent interactions[⋆,⋆⋆]

Jorge González[1], Rodrigo Martínez[2], José A. Fernández[1], and Judith Millan[2,a]

[1] Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco-UPV/EHU, Barrio Sarriena s/n, 48940 Leioa, Spain
[2] Departamento de Química, Facultad de Ciencia y Tecnología, Universidad de La Rioja, Madre de Dios, 53, 26006 Logroño, Spain

Received 17 March 2017 / Received in final form 19 May 2017
Published online 3 August 2017 –

Abstract. The intramolecular interactions for isolated capped amino acids were investigated computationally by characterizing the conformers for selected amino acids with charged (arginine), polar (asparagine and glutamine), non-polar (alanine, valine and isoleucine), and aromatic (phenylalanine, tryptophan and tyrosine) side chains. The computational method applied combined a molecular mechanics conformational search (with an MMFFs forced field) followed by structural and vibrational density-functional calculations (M06-2X with a triple-ζ Pople's basis set). The intramolecular forces in each amino acid were analyzed with the Non-Covalent Interactions (NCI) analysis. The results for the 15 most stable conformers studied showed that the structure of isolated capped amino acids resembles those found in proteins. In particular, the two most stable conformers of the nine amino acids investigated exhibit γ_L and β_L conformations with 7- and 5-membered rings, respectively, as a result of the balance between non-covalent interactions (hydrogen bonds and van der Waals).

1 Introduction

Proteins are fundamental macromolecules for life, with a diversity of functions. Proteins act as channels through cellular walls, catalyzers, DNA benders, etc. The quaternary, tertiary, and secondary structures are key in the role that proteins play, being these macro-structures stabilized mainly by non-covalent interactions among the backbone and the lateral chain of their amino acids [1].

The wide range of functions and macro-structures found in proteins comes up from the combination of only 20 amino acids. Thus, the amino acid sequence of a protein (primary structure) determines its three-dimensional structure and therefore, its biological function [1–5]. The 20 amino acids can be classified by the nature and composition of their lateral chain: there are aliphatic or non-polar, acidic or basic, sulfur-containing, aromatic amino acids and so on. These general properties influence the conformational behavior of the isolated amino acids as it is shown in theoretical and experimental studies found in the literature [6–23].

Related to the nature of the lateral chain of the amino acids (non-polar, polar, charged and aromatic) there are several statistical studies about the propensity of contacts between the amino acids and the nucleotides in a protein-DNA complex. These contacts cover the interactions at the level of the phosphate, sugar and DNA-base separately or with the entire nucleotide [24,25]. In our previous studies, we developed a theoretical method which applies molecular and quantum mechanics to model a reductionist approach to calculate the propensities between amino acids and DNA bases, obtaining a good correlation between our results and statistical data [26].

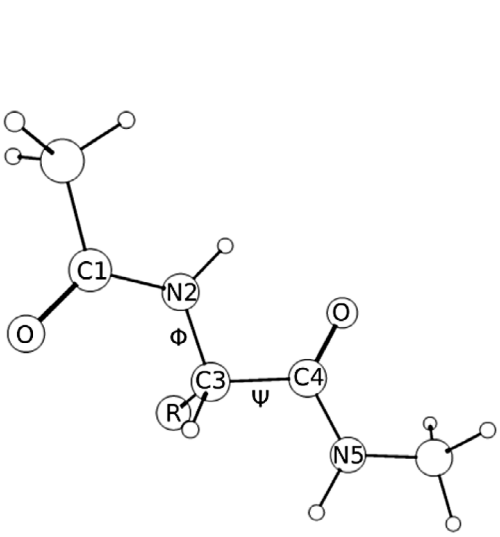
Here, we present a study on the conformational preferences of isolated capped amino acids using the same theoretical methodology to test whether the intramolecular interactions stabilize structures in gas phase that resemble incipient recognized secondary structures. It is not possible to correlate the results of the configuration obtained for the isolated amino acids to the secondary structures in proteins because of the size of the investigated systems [27–34]. However, as these amino acids are present in the structures found in proteins, the knowledge of the conformation adopted by the capped amino acids may be useful for the extrapolation of their properties to larger systems, as polypeptides or proteins, demonstrating that even in simple peptides and in absence of water, the biologically-relevant structures are already present.

To answer these question, we chose nine amino acids, according to the nature of their lateral chain: three with non-polar side chain (alanine, valine and isoleucine), three with polar or charged side chain (asparagine, glutamine

⋆ Contribution to the Topical Issue "Dynamics of Molecular Systems (MOLEC 2016) ", edited by Alberto Garcia-Vela, Luis Banares and Maria Luisa Senent.

⋆⋆ Supplementary material in the form of one pdf file available from the Journal web page at
https://doi.org/10.1140/epjd/e2017-80187-5

[a] e-mail: judith.millan@unirioja.es



R= -CH_3 (Alanine)

-$CH(CH_3)_2$ (Valine)

-$CH(CH_3)CH_2CH_3$ (Isoleucine)

-CH_2-C(=O)-NH_2 (Asparagine)

-CH_2CH_2-C(=O)-NH_2 (Glutamine)

-$CH_2CH_2CH_2$-NH-C(=NH_2^+)-NH_2 (Arginine)

-CH_2-C_6H_5 (Phenylalanine)

-CH_2-C_6H_4OH (Tyrosine)

-CH_2-C_8H_5N (Tryptophan)

Fig. 1. Structure and atom labeling for the studied amino acids (R denotes the side chain). The torsional angles ϕ and ψ are noted [36].

and arginine) and three with aromatic side chain (phenylalanine, tryptophan and tyrosine). Since the amino acids are inmersed in the protein, the terminal carboxylic and amino groups are not available for the interaction. Therefore, to have a system that resembles the real situation in the protein, the amino and acid groups were replaced by an amide group at both ends of the molecule. The atom labeling for the amino acids (with side chain R), the definition of the torsional dihedral angles ϕ and ψ, and the structure of the side chains for the investigated amino acids, may be found in Figure 1 [4,35,36].

The methodology combines a careful exploration of the conformational landscape of each isolated amino acid using molecular mechanics (MM), followed by quantum mechanical calculations using density-functional-theory (DFT) [26,37,38]. As we are especially interested in the interactions which stabilize the structures, we carried out an analysis of the electron density and its reduced density gradient for the optimized structures [39]. The non-covalent interactions (NCI) plots allowed us to characterize the non-covalent interactions in the amino acids and, particularly, to obtain a generalization of hydrogen bonds by the analysis of electron density at the bond critical points and on the surrounding regions, revealing also van der Waals interactions [40,41].

We present the conformational space for the nine capped amino acids investigated, characterizing the most relevant non-covalent interactions that govern the final configuration of the molecules, together with an analysis of their behavior along the temperature. We try to probe if these interactions are the responsible of the conformation that finally adopt the amino acids, looking for trends depending on the lateral chain, which can elucidate the orientation in larger biological systems.

2 Computational method

A first exploration of the conformational landscape using the Merck Molecular Mechanics force field (MMFFs) [42] was performed to locate the relevant structures for each individual capped amino acid within an energy window of 30 kJ mol^{-1}. The conformational map was obtained by combining Monte Carlo procedure and the "large-scales low-mode" method based in vibrational normal-mode analysis, as implemented in Macromodel [43]. The resulting conformations were optimized using the M06-2X DFT functional, which is very appropriate for systems where dispersion interactions are relevant [44,45]. A triple-ζ basis set with polarization and diffused orbitals (6-311++G(d,p)) [46–49] was used in the calculations. Our previous experiments proved that this theoretical level is reliable for systems with similar size in which non-covalent interactions are also present [26,37,38]. Furthermore, B3LYP-D with Grimme's D3 correction [50,51] and MP2 calculations [52] were also carried out for alanine in order to test the computational methodology for these systems. The geometry optimizations were completed with a normal-mode analysis, as implemented in Gaussian09 [53], to check that the structures were true minima and obtain the zero-point energy (ZPE) corrections. Thus, the relative energies presented in this work contain such correction.

After that, a Non-Covalent Interaction (NCI) analysis was done for the two most stable conformations previously

calculated for each amino acid. The NCI approach evaluates the intermolecular interactions, based on the behavior of the reduced density gradient, s, with respect to the electron density:

$$s = \frac{|\nabla\rho|}{c_F}\rho^{4/3}, \quad (1)$$

localizing the non-covalent interactions between the atoms, which appear as peaks in the $s(\rho)$ diagram and are then classified [39]. The limiting cases are strongly attractive hydrogen bonds, van der Waals and steric repulsions which are reflected with a coloring scheme (Blue, Green, Red, respectively, BGR code) [54]. The NCI plots were drawn with the NCIPLOT program and were visualized with Visual Molecular Dynamics (VMD)[1] [55].

3 Results

A significantly large amount of structures were found for each capped amino acid after applying the above-described methodology. Therefore, only the global minimum and the structure of the second most stable conformation of each amino acid will be shown. Representation of the 15 most stable structures for each amino acid, together with their energetics, can be found in Figures S1–S9 (Supplementary material, SM). Moreover, the 2D-NCI plots of the two most stable conformations found for each amino acid (Figs. S10–S12) and their structures in Cartesian coordinates (Figs. S13–S15) are also collected in SM.

The final structures are named as aXn, where X = A, V, I, N, Q, R, F, W and Y denotes the amino acid (alanine, valine, isoleucine, asparagine, glutamine, arginine, phenylalanine, tryptophan and tyrosine respectively) and $n = 1, 2, 3, \ldots$ denotes the relative stability of a particular structure, being $n = 1$ the most stable one.

3.1 Non-polar amino acids

Exploration of the conformational landscape of alanine, valine and isoleucine generated 23, 29, and 66 structures, respectively. The two most stable structures of each amino acid calculated at M06-2X/6-311++G(d,p) level are shown in Figure 2, left panels, together with the value of their relative energies, the distances of the most relevant interactions, and the NCI plots (right panels). The values of the distances and angles of the hydrogen bonds and the electron density of these interactions are collected in Tables 1 and 2.

The most stable structure of all three amino acids presents a C1=O$\cdots$HN5 7-membered ring in what is known as an γ_L conformation [4]. Despite that C1=O$\cdots$HN5 interaction is also present in γ_D conformations, γ_L ones are more stable due to an additional C1=O$\cdots$HC3 interaction which stabilizes the final structure and becomes the global minimum.

[1] http://www.ks.uiuc.edu/Research/vmd

Table 1. Distances (Å) and angles (°) of the most relevant intramolecular interactions observed for aA, aV and aI structures (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type		
		H bond	vdW	
		C1=O$\cdots$HN5	C1=O$\cdots$HN5	N2H$\cdots$O=C4
γ_L	aA1	2.086/141.4		
	aV1		2.174/136.6	
	aI1		2.153/138.6	
β_L	aA2			2.183/106.0
	aV2			2.185/105.7
	aI2			2.189/105.7

Table 2. Electron density (a.u.) of the most relevant interactions observed for aA, aV and aI calculated from the NCI grid with $0.1 \times 0.1 \times 0.1$ increments (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type		
		H bond	vdW	
		C1=O$\cdots$HN5	C1=O$\cdots$HN5	N2H$\cdots$O=C4
γ_L	aA1	0.0192		
	aV1		0.0164	
	aI1		0.0169	
β_L	aA2			0.0197
	aV2			0.0196
	aI2			0.0194

According to the NCI plots, the C1=O$\cdots$HN5 interaction is a hydrogen bond, as indicated by the dark blue disc in the NCI plot in Figure 2. However, the angle of the interaction (Tab. 1) indicates a better overlap between the orbitals of the interacting atoms for alanine than for valine and isoleucine, due to steric reasons [56]. In fact, the value of electron density confirms this assessment, which is slightly higher than in valine and isoleucine (Tab. 2): 0.0192, 0.0164 and 0.0169 a.u. respectively. Furthermore, additional van der Waals interactions are also found, such as the above-mentioned C1=O$\cdots$HC3 interaction, depicted with a bicolor blue/red isosurface due to the attractive contributions between the carbonyl and the hydrogen (blue) and the repulsive effect of the ring closure (red), together with some weak interactions with the hydrogens of the lateral chain depicted in green.

Thus, the final structures are a subtle balance of forces, where C1=O$\cdots$HN5 is the leading interaction which helps to close the 7-membered ring. As the size of the side chain increases from valine to isoleucine, the interactions with the peptidic bond increase, opening the ring, and reducing its relative contribution to the overall stability of the system (see C1=O$\cdots$HN5 distances in Tab. 1). This is reflected in the energy difference between the global minimum and the second most stable structure, which is an β_L conformation, which decreases from alanine to valine and isoleucine in parallel with the increase in the C1=O$\cdots$HN5 distance (5.19, 1.13 and 1.64 kJ mol^{-1}, respectively).

Regarding the second most stable type of structures, β_L conformation for the three amino acids, the most

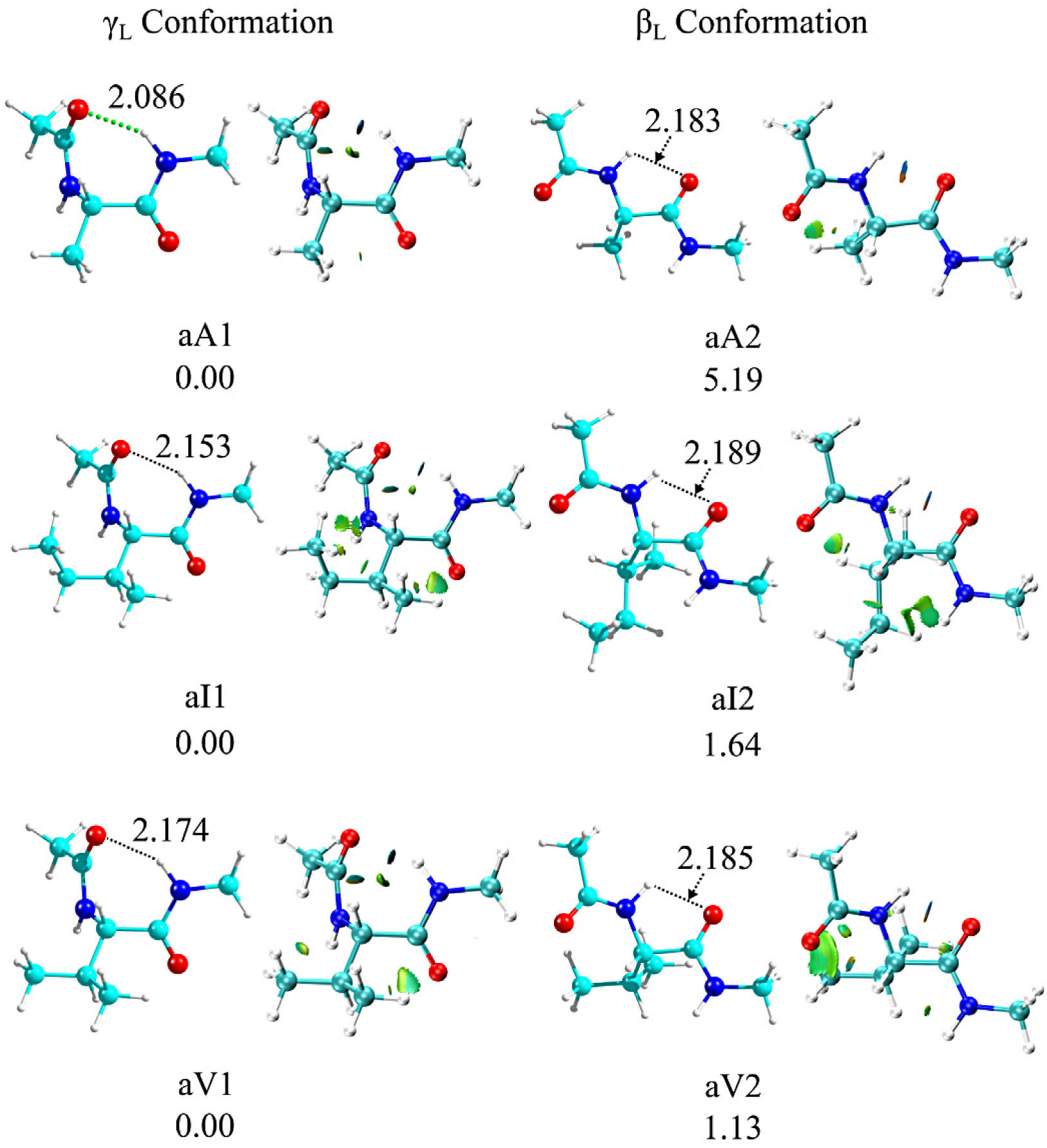


Fig. 2. The two most stable conformations, γ_L and β_L, for aA, aI and aV with distances (Å) of the hydrogen bonds and the most relevant van der Waals interactions and NCI plots with the RBG code (right panel). Relative energy values (kJ mol^{-1}) are also shown.

relevant intermolecular interaction N2H$\cdots$O=C4 forms an intramolecular 5-membered ring, which causes significantly stronger steric hindrances reflected by the red and blue bicolor isosurface in the NCI plots (Fig. 2). The blue zone is related to the intramolecular attractive interaction (N2H$\cdots$O=C4), while the red one is related to the steric effect due to the small number of members of the ring. There are also additional weak van der Waals interactions, mostly due to the contact between C1=O and N5H with the hydrogens of the lateral chain.

For alanine, a comparison between three computational methods (M06-2X, B3LYP-D with Grimme's D3 correction and MP2) was carried out in order to test our theoretical method. Figure 3 collects the structural parameters and the relative energy in kJ mol^{-1} of the two most stable conformers of aA found for these three methods. No relevant differences were found, what reinforces that M06-2X is a suitable method to investigate the capped amino acids as we found in previous results for systems in which non-covalent interactions are prevalent.

3.2 Polar amino acids

The size of the polar amino acids investigated, asparagine, glutamine and arginine, together with the presence of a carbonyl and an amino group at the end of their lateral chains, results in an increase in the number of structures found: 51, 46 and 44 for aN, aQ and aR, respectively. The two most stable conformers found for each amino acid

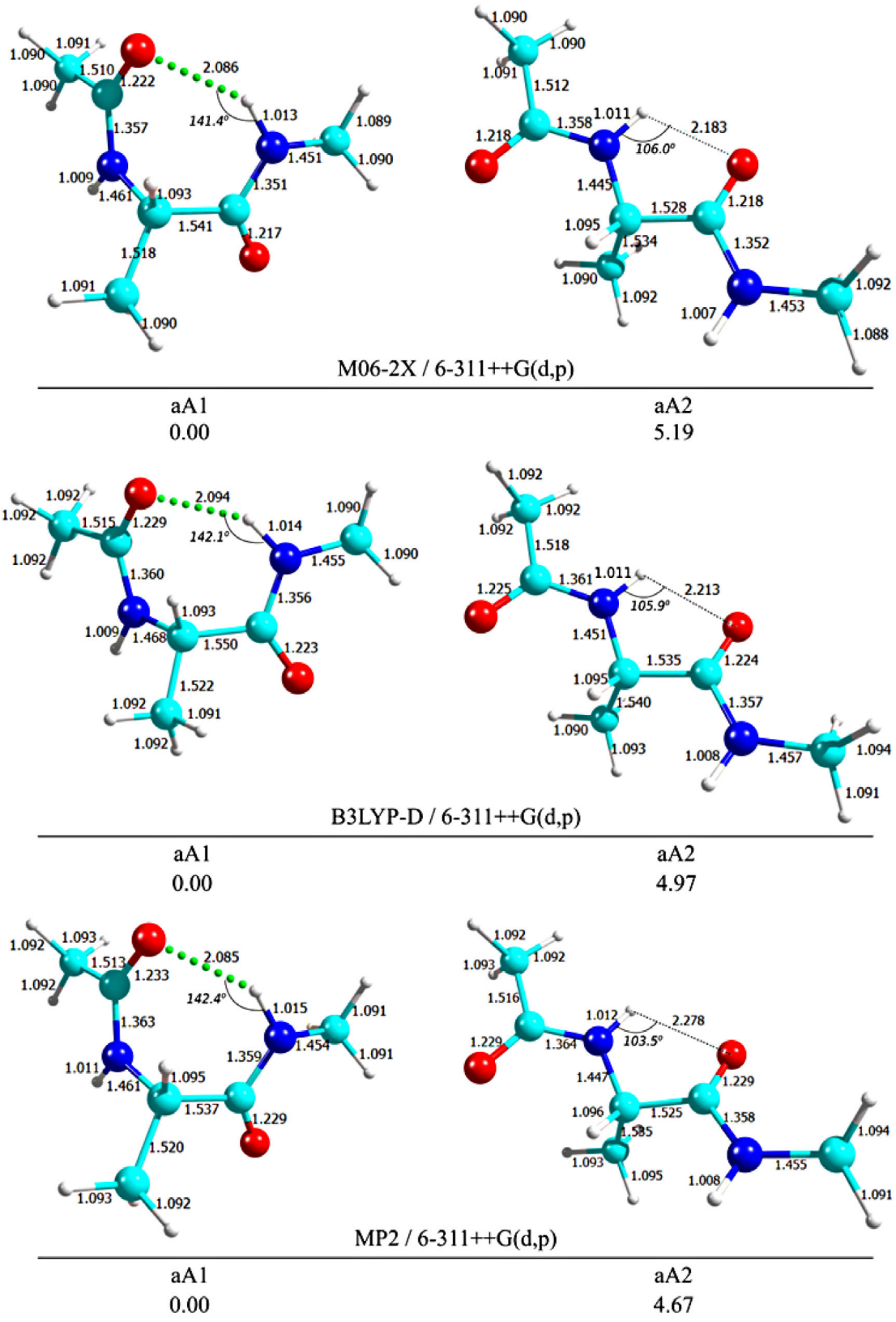


Fig. 3. The two most stable conformations for aA calculated at M06-2X/6-311++G(d,p), B3LYP-D/6-311++G(d,p) and MP2/6-311++G(d,p) levels with structural parameters (distances in Å, angle in °) depicted. The relative energy values (kJ mol^{-1}) are also shown.

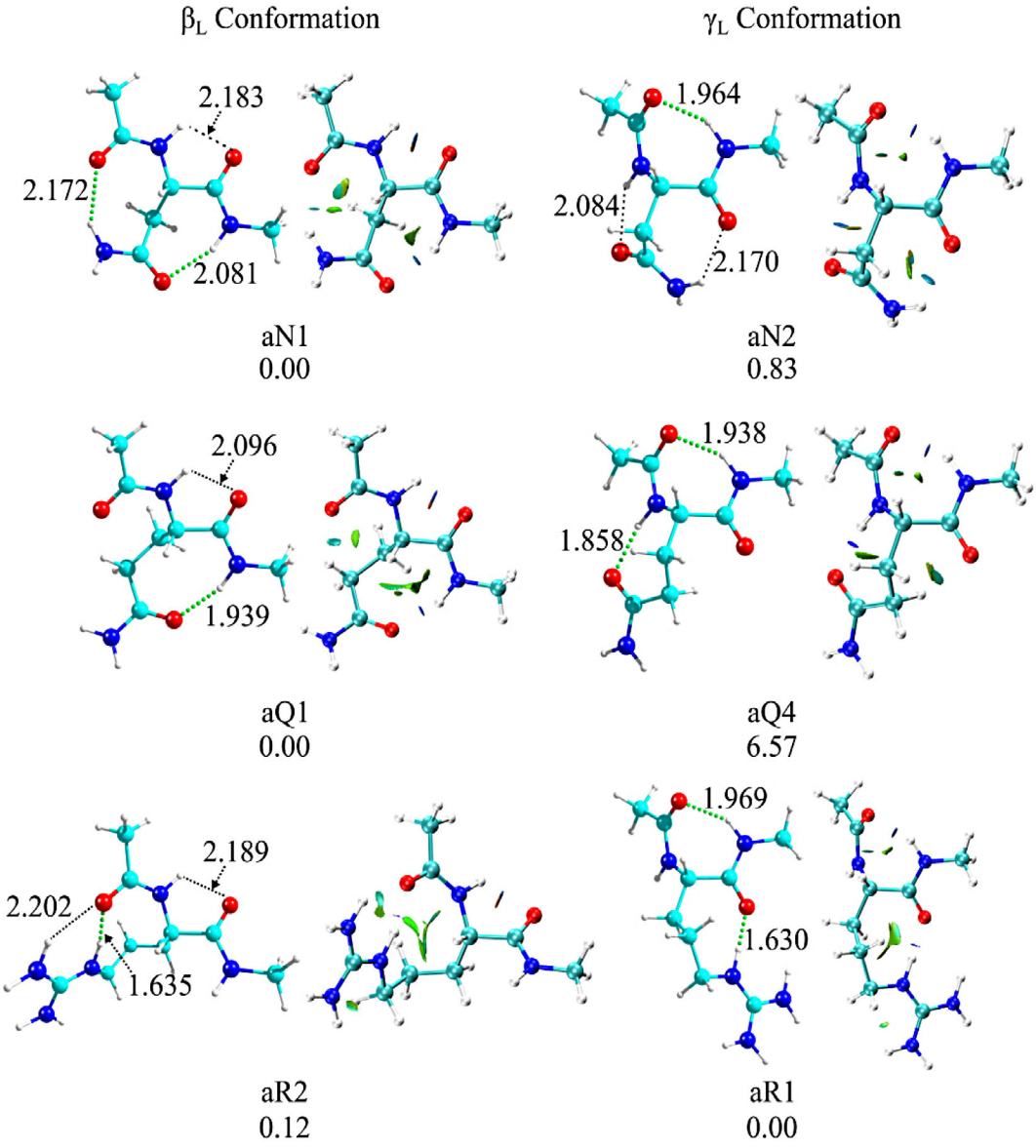


Fig. 4. The two most stable conformations, β_L and γ_L, for aN, aQ and aR with distances (Å) of the hydrogen bonds and the most relevant van der Waals interactions and NCI plots with the RBG code (right panel). Relative energy values (kJ mol^{-1}) are also shown.

are displayed in Figure 4, together with the relative energies, the distance of the most relevant interactions, and the NCI plots in the right panels. The values of the distances and angles for the hydrogen bonds and the values of the electron density of the interactions are collected in Tables 3 and 4.

Leaving aside the structural differences with the non-polar amino acids, the two most stable structures of the three polar amino acids are also γ_L and β_L. However, the relative energetic order is reversed for asparagine and glutamine, for which the β_L conformer is more stable. Furthermore, the three most stable conformers for glutamine are β_L conformations and only differ in the rotation of the lateral chain. For this amino acid, the first structure with conformation γ_L is aQ4, as displayed in Figure 4.

In addition to the N2H$\cdots$O=C4 van der Waals interaction already observed in non-polar amino acids in β_L conformations, the presence of heteroatoms in polar amino acids results in a rich collection of interactions, as it can be seen in Figure 4. The amide group of asparagine and glutamine side chains is prone to establish intramolecular interactions. So, the aN1 conformer exhibits two intramolecular hydrogen bonds with the carbonyl and the amine moieties of the lateral chain. On the other hand, despite that glutamine's side chain only differs from asparagine in a methyl group, the orientation it adopts is very different. Thus, the carbonyl moiety presents a single hydrogen bond with the amine group of the peptidic backbone, although with shorter distance and higher electron density than the equivalent in asparagine

Table 3. Distances (Å) and angles (°) of the relevant intramolecular interactions observed for aN, aQ and aR structures (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type								
		H bond					vdW			
		R:NH···O=C1	R:C=O···HN5	C1=O···HN5	R:C=O···NH2	R: H···O=C4	N2H···O=C4	R:NH···O=C1	R:NH···O=C4	R:C=O···NH2
β_L	aN1	2.173/145.9	2.081/153.3				2.154/107.0			
	aQ1		1.939/160.5				2.096/109.2			
	aR2	1.635/166.4					2.138/105.7	2.202/136.3		
γ_L	aN2			1.964/147.4					2.170/127.7	2.084/134.7
	aQ4			1.938/148.3	1.858/158.5					
	aR2			1.969/143.3		1.630/174.4				

Table 4. Electron density (a.u.) of the relevant interactions observed for aN, aQ and aR, calculated from the NCI grid with $0.1 \times 0.1 \times 0.1$ increments (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type								
		H bond					vdW			
		R:NH···O=C1	R:C=O···HN5	C1=O···HN5	R:C=O···HN2	R: H···O=C4	N2H···O=C4	R:NH···O=C1	R:NH···O=C4	R:C=O···HN2
β_L	aN1	0.0151	0.0188				0.0206			
	aQ1		0.0242				0.0226			
	aR2	0.0543					0.0214	0.0142		
γ_L	aN2			0.0264					0.0171	0.0196
	aQ4			0.0259	0.0310					
	aR2			0.0243		0.0532				

(Tabs. 3 and 4). So, it seems that the lateral chain in glutamine somehow reduces the number of hydrogen bonds, preferentially maximizing one of them. The side chain of arginine seems also suitable to establish intramolecular interactions. In fact, its NH group forms an extremely strong hydrogen bond with the C=O moiety of the peptidic chain, remarked by the high value of the electron density (0.0543 a.u., Tab. 4) and the short binding distance (1.635 Å, Tab. 3). Furthermore, an additional stabilizing interaction is observed between one of the NH_2 groups and the same carbonyl, depicted by a light blue zone in the NCI plot (Fig. 4), with an angle lower than 140°, due to the size and nature of the lateral chain [56]. Some weaker van der Waals interactions are also present in the structure.

The main interaction in the γ_L conformation of polar amino acids exhibits a reduced interaction distance when it is compared with the non-polar ones. Thus, the C1=O···HN5 distance varies from 2.086, 2.153 and 2.174 Å in aA1, aI1 and aV1 respectively (see Tab. 1) to 1.964, 1.938 and 1.969 Å for aN2, aQ4 and aR1 (see Tab. 3). The higher values of electron density in these interactions (0.0264, 0.0259 and 0.0243 a.u., respectively, see Tab. 4) endorse the stronger character of this hydrogen bond in polar amino acids. Moreover, additional interactions with the electronegative moieties of the side chain seem to stabilize the final structures. One strong hydrogen bond is observed for glutamine (C=O···NH2, 0.0310 a.u.) and especially in arginine (H···O=C4, 0.0532 a.u.), which remarks the relevance of the intramolecular side chain interactions for the stabilization of these structures. For asparagine, two weaker stabilizing interactions with the amide group of the side chain are also observed. Furthermore, weaker van der Waals interactions related to the ring closure or to electronegative groups interacting with aliphatic hydrogen, are also found.

3.3 Aromatic amino acids

Because the aromatic lateral chains are voluminous and lack the flexibility of the chains of previous systems, their conformational landscape is less complicated than for non-polar or polar systems. Consequently we obtained 29, 46 and 36 conformers for aF, aW and aY, respectively in the 30 kJ mol^{-1} window.

As in previous systems, the most stable structures are γ_L or β_L conformations, as it can be seen in Figure 5; the relevant parameters of the most important interactions were collected in Tables 5 and 6. The difference between γ_L and β_L conformations is mainly the relevance of the interaction between the peptidic bond and the aromatic ring. So, when the amino acid adopts an γ_L geometry, direct C=O···π and N-H···π interactions take place, as NCI plots displayed in Figure 5. However, β_L conformation allows the molecule to establish an N2-H···π interaction. Thus, the difference between both types of conformations seems very small – within the computation error – as indicated by the relative energy values and the interaction distances.

It is interesting to note that, despite the presence of a polar group in the lateral chain of aY, its behavior is different from that of previous polar amino acid. Thus, the rigidity of the lateral ring hampers the establishment

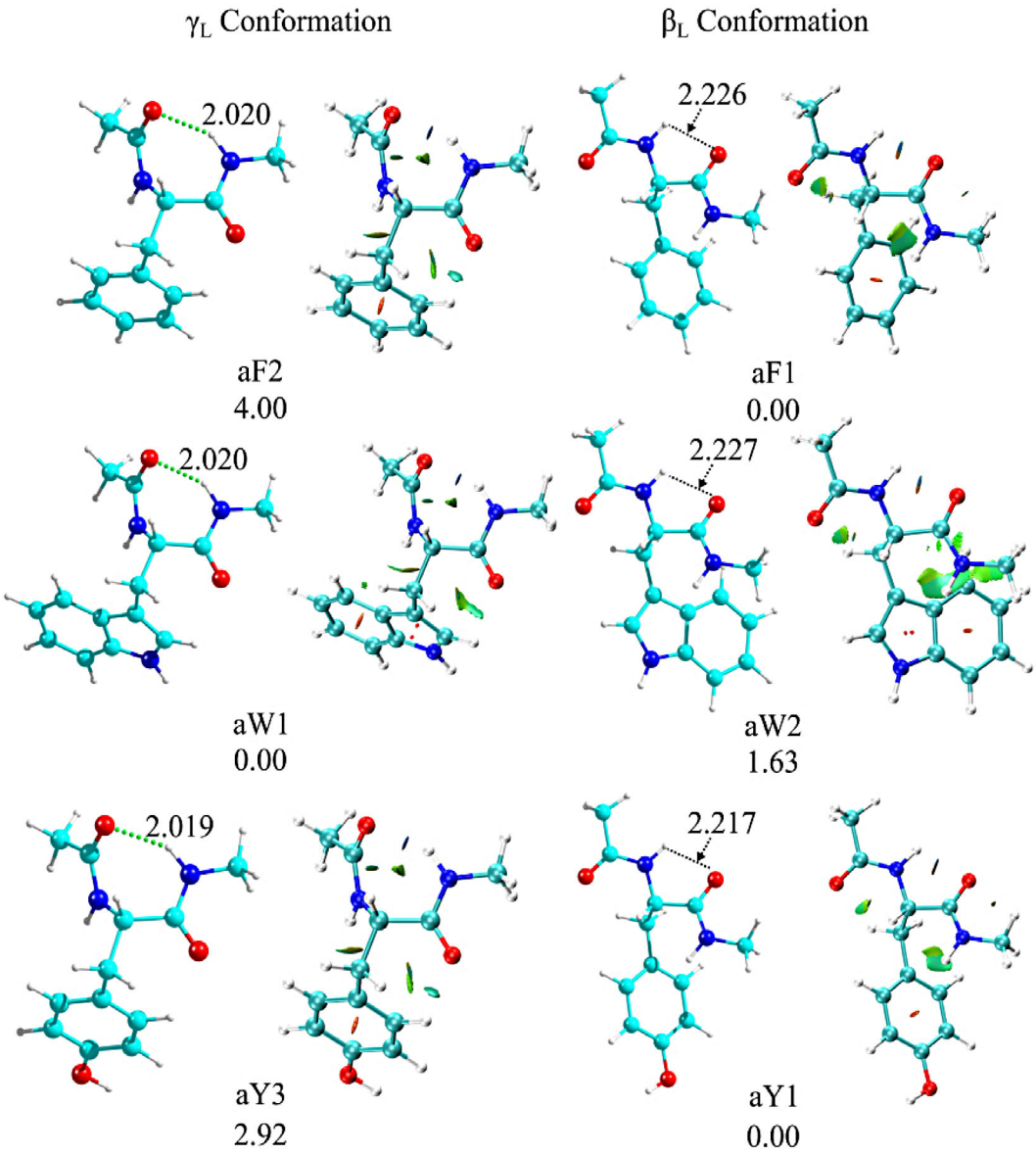


Fig. 5. The two most stable conformations, γ_L and β_L, for aF, aW and aY with distances (Å) of the hydrogen bonds and the most relevant van der Waals interactions and NCI plots with the RBG code (right panel). Relative energy values (kJ mol^{-1}) are also shown.

of intramolecular interactions with the peptide backbone, as it occurs for aF and aW.

4 Discussion

The conformational distribution and the number of structures that each capped amino acid may adopt in the stability window investigated are displayed in Figure 6. Although the calculation level of the study is relatively high, changes in the relative stability of some structures can still occur if the calculation level is further increased. However, the tendency of all the amino acids investigated to adopt 7-member ring conformations (γ_L or γ_D, represented as black squares and crosses, respectively) or 5-member conformations (β_L: red squares) is clear and governs the region nearby to the global minimum. Thus, the third different conformation in non-polar amino acids appears above 6 kJ mol^{-1}; for polar, above 10 kJ mol^{-1} and above 7 kJ mol^{-1} for aromatic. The only exception is aR, which presents several structures adopting α_D conformations. Despite the four most stable conformers are γ_L or β_L, the interactions with the lateral chain also favors this third type of structure (the first conformer of this family, aR5, is found at 2.17 kJ mol^{-1}). Our results for the most stable structures for alanine and valine are in agreement with spectroscopic IR/UV data that show a large propensity for both amino acids to adopt locally an γ_L conformation in short peptides followed by a second preference for the β_L conformation [57]. Indeed for alanine, FTMW

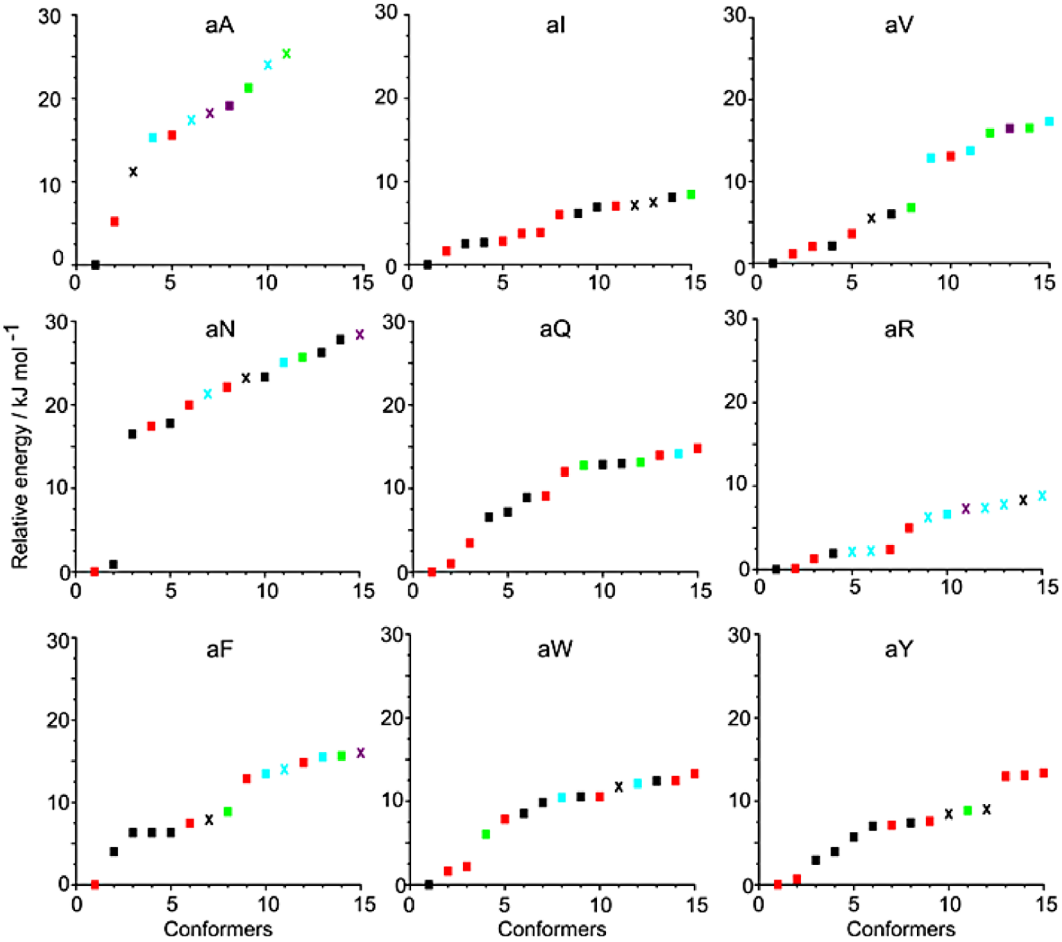


Fig. 6. Conformational distribution of the 15 most stable conformers of aA, aI, aV, aN, aQ, aR, aF, aW, and aY. Black: γ-conformation; red: β-conformation; blue: α-conformation; green: δ-conformation; purple: ε-conformation (□: L-disposition; x: D-disposition).

Table 5. Distances (Å) and angles (°) of the relevant intramolecular interactions observed for aF, aW and aY structures (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type		
		H bond	vdW	
		C1=O···HN5	N2H···O=C4	R:N2H···π
β_L	aF1		2.226/104.8	2.600/ –
	aW2		2.227/104.9	2.700/ –
	aY1		2.217/105.0	2.600/ –
γ_L	aF2	2.020/147.3		2.500/ –
	aW1	2.019/147.1		2.500/ –
	aY3	2.019/147.4		2.500/ –

Table 6. Electron density (a.u.) of the relevant interactions observed for aF, aW and aY, calculated from the NCI grid with $0.1 \times 0.1 \times 0.1$ increments (H bond: hydrogen bond, vdW: van der Waals).

		Interaction type	
		H bond	vdW
		C1=O···HN5	N2H···O=C4
β_L	aF1		0.0185
	aW2		0.0185
	aY1		0.0186
γ_L	aF2	0.0214	
	aW1	0.0215	
	aY3	0.0215	

experiments found that the most stable conformation for Ac-Ala-NHMe, is γ_L, which also confirm our conclusions for this amino acid [58]. However, for aromatic residues in short peptides, the IR/UV results show a preference to adopt β_L conformations to allow the establishment of a NH···π interaction, as we found for aF and aY [57]. Although for aW we obtain a reversal of both conformations, the difference between them is very small (1.63 kJ mol^{-1}) and could be considered within the computation error.

It is worthy to mention that other conformations, such as α_L or α_D (blue symbols), δ_L or δ_D (green symbols) and ε_L or ε_D (purple symbols) are also found, but less frequently. Therefore, we have investigated a representative member of those dispositions in order to analyze their structures. Hence, we found a weak NH···N interaction for α_D and δ_L with a distance too long to be taken as a 5 membered-ring. For the ε_D orientations, the N is not even

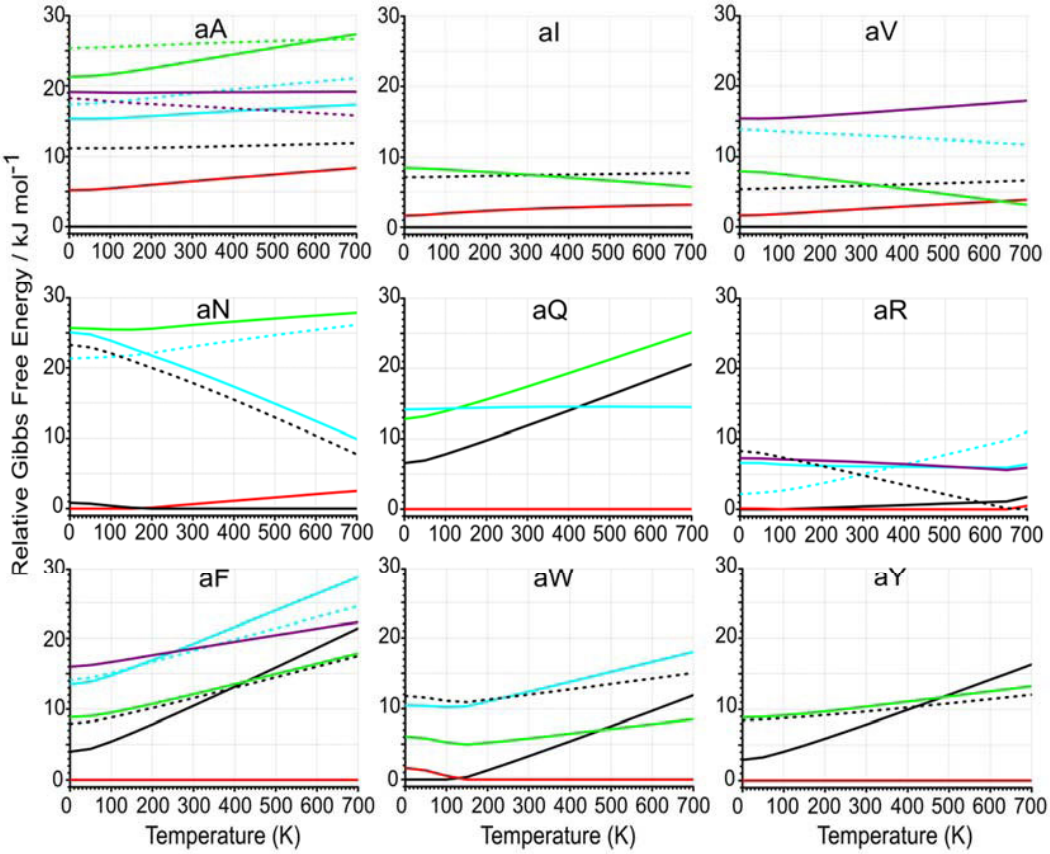


Fig. 7. Relative Gibbs free energy (kJ mol^{-1}) of aA, aI, aV, aN, aQ, aR, aF, aW, and aY calculated in the 0–700 K range. Black: γ-conformation; red: β-conformation; blue: α-conformation; green: δ-conformation; purple: ε-conformation (continous line: L-disposition; dot line: D-disposition).

pointing to the NH group, so it is not possible to form a ring. The rest of the non-covalent interactions found for α_D, δ_L and ε_D are not adequate to set cyclic structures. Thus, we conclude that α, δ and ε do not have interactions strong enough or with an adequate orientation to form ringed geometries. Moreover, this variety of α, δ or ε appears at high energy values (ranging from 6.05 kJ mol^{-1} for aW4 to 15.28 kJ mol^{-1} for aA4), remarking that the most stable conformers are stabilized with the formation of cyclic structures in γ and β. The only exception is aR, as commented previously.

In order to study the biological behavior of the amino acids, we calculated the variation of the Gibbs free energy from 0 to 700 K, the temperature at which organic matter decomposes[2] [59,60]. The relative ΔG values are presented in Figure 7. As it can be seen, γ_L and β_L dispositions are still the most favored structures for all the amino acids investigated between 0 and 700 K. Focusing in the temperature range of life (300–310 K), it is observed that γ_L conformation is the preferred one for non-polar amino acids, being β_L less stable as the temperature increases. For polar amino acids, aN and aR are clearly leaded by γ_L and β_L, which are separated by less than 2.5 kJ mol^{-1} in the whole interval of temperature. In fact, despite that β_L is the most stable orientation for aN, γ_L becomes more favorable between 150 and 700 K. On the other hand, for aR, γ_L and β_L conformers are isoenergetic. However, as the temperature increases, β_L conformation evolves into the most favored. Furthermore, aQ clearly exhibits a β_L preference, being even more separated from γ_L with the increase of the temperature. Finally, despite aromatic amino acids exhibited different global minimum conformations, when the temperature is taken into account, the preference of aF and aY for β_L conformations is clear in the whole range of temperature and above 150 K for aW.

[2] http://www.nist.gov/mml/csd/informatics_research/thermochemistry_script.cfm

5 Conclusions

We investigated the conformational landscape of nine capped amino acids (alanine, valine, isoleucine, asparagine, glutamine, arginine, phenylalanine, tryptophan and tyrosine) using a combination of computational methods which included molecular mechanics, DFT calculations and Non-Covalent Interaction analysis.

We found that non-covalent intramolecular interactions, hydrogen bonds and van der Waals dispersive forces, control the orientation of the 15 selected conformers for each amino acid, for which we perform a detailed

conformational map. However, they play a different role in the studied amino acids, according to the character of their lateral chain.

Thus, for non-polar amino acids (alanine, valine and isoleucine), the lateral chain stays away from the backbone for the two most stable conformers, whereas, for polar amino acids (asparagine, glutamine and arginine), the lateral chain plays a main role in the stability of the system. The same applies to the aromatic amino acids (phenylalanine, tryptophan and tyrosine), which exhibit interactions between the backbone and the aromatic ring. The analysis of the most relevant non-covalent intramolecular interactions found for the two most stable conformations of the nine amino acids showed that they are always cyclic structures, having a ring of 7 (γ_L or γ_D) or 5 (β_L) members. Both conformations are of similar stability, with differences in most cases within the computation error.

Fort the rest of structures investigated, γ or β conformations are still the preferred dispositions in all the range of temperatures investigated. Furthermore, α, δ or ε conformations do not present interactions that help to the establishment of cycles or rings and appear less frequently and at high values of energy.

So, we found the incipient dispositions that are recognized in larger systems (γ_L or β_L for the most stable conformers), as in polypeptides or protein structures. Therefore, the information about the conformations adopted by the capped amino acids will be useful for their application to bigger systems because even for these small molecules and in absence of water, the biological structures are also present.

Supplementary material

The 15 most stable structures of each amino are collected in Figures S1–S9. The 2D-NCI plots of the two most stable conformers found for each amino acid are collected in Figures S10–S12. The structures of the two most stable structures for each amino acid in Cartesian coordinates are displayed in Figures S13–S15.

The research leading to these results has received funding from the Spanish MINECO/FEDER EU (CTQ2015-68148-C2-1-P). J.G. thanks the UPV/EHU for a predoctoral fellowship. Computational resources from the SGI/IZO-SGIker network were used for this work.

Author contribution statement

The results collected in the paper are a part of Jorge González's Doctoral Thesis. He performed the calculations in collaboration with Dr. Rodrigo Martínez. Drs. Jose A. Fernández and J. Millan are the supervisors of Jorge Gonzalez's. Thesis and wrote the manuscript in collaboration with Jorge Gonzalez and Dr. Martínez.

References

1. D.L. Nelson, M.M. Cox, *Lehninger Principles of Biochemistry* (W.H. Freeman & Company, New York, 2008)
2. J.M. Berg, J.L. Tymoczko, L. Striyer, *Biochemistry* (W.H. Freeman and Company, New York, 2002)
3. J.S. Richardson, *The Anatomy and Taxonomy of Protein Structure* (Academic Press Inc., Durham, 2007)
4. G.A. Chass, M.A. Sahai, J.M.S. Law, S. Lovas, Ö. Farkas, A. Perczel, J.-L. Rivail, I.G. Csizmadia, Int. J. Quantum Chem. **90**, 93 (2002)
5. G.A. Petsko, D. Ringe, *Protein Structure and Function* (New Science Press, London, 2004)
6. A.G. Császár, J. Phys. Chem. **100**, 3541 (1996)
7. A.G. Császár, J. Mol. Struct. Theochem. **346**, 141 (1995)
8. M. Rassolian, G.A. Chass, D.H. Setiadi, I.G. Csizmadia, J. Mol. Struct. Theochem. **666**, 273 (2003)
9. A. Barth, Prog. Biophys. Mol. Biol. **74**, 141 (2000)
10. M.N. Barroso, E.S. Cerutti, A.M. Rodríguez, E.A. Jáuregui, O. Farkas, A. Perczel, R.D. Enriz, J. Mol. Struct. Theochem. **548**, 21 (2001)
11. B. Boeckx, G. Maes, Biophys. Chem. **165**, 62 (2012)
12. C. Cabezas, M. Varela, I. Pena, S. Mata, J.C. Lopez, J.L. Alonso, Chem. Commun. **48**, 5934 (2012)
13. F.C. Calaza, M.V. Rigo, A.N. Rinaldoni, M.F. Masman, J.C.P. Koo, A.M. Rodríguez, R.D. Enriz, J. Mol. Struct. Theochem. **634**, 201 (2003)
14. M. Chen, Z. Huang, Z. Lin, J. Mol. Struct. Theochem. **719**, 153 (2005)
15. M.A. da Silva, M. da Silva, A.F. Santos, M.V. Roux, C. Foces-Foces, R. Notario, R. Guzman-Mejia, E. Juaristi, J. Phys. Chem. B **114**, 16471 (2010)
16. C.M. Deane, F.H. Allen, R. Taylor, T.L. Blundell, Protein Eng. **12**, 1025 (1999)
17. P.D. Godfrey, S. Firth, L.D. Hatherley, R.D. Brown, A.P. Pierlot, J. Am. Chem. Soc. **115**, 9687 (1993)
18. S. Gronert, R.A.J. O'Hair, J. Am. Chem. Soc. **117**, 2071 (1995)
19. B. Hernández, F. Pflüger, M. Nsangou, M. Ghomi, J. Phys. Chem. B **113**, 3169 (2009)
20. B. Lakard, J. Mol. Struct. Theochem. **681**, 183 (2004)
21. A. Lesarri, R. Sánchez, E.J. Cocinero, J.C. López, J.L. Alonso, J. Am. Chem. Soc. **127**, 12952 (2005)
22. S. Dokmaisrijan, V.S. Lee, P. Nimmanpipug, J. Mol. Struct. Theochem. **953**, 28 (2010)
23. Y. Kawashima, T. Usami, N. Ohashi, R.D. Suenram, J.T. Hougen, E. Hirota, Acc. Chem. Res. **39**, 216 (2006).
24. R. Sathyapriya, K.V. Brinda, S. Vishveshwara, J. Chem. Inf. Model **46**, 123 (2006)
25. R. Sathyapriya, S. Vishveshwara, M.S. Vijayabaskar, PLoS Comput. Biol. **4**, e1000170 (2008)
26. J. González, I. Baños, I. León, J. Contreras-García, E.J. Cocinero, A. Lesarri, J.A. Fernández, J. Millan, J. Chem. Theory Comput. **12**, 523 (2016)
27. C.N. Pace, J.M. Scholtz, Biophys. J. **75**, 422 (1998)
28. M. Blaber, X.-J. Zhang, B.W. Matthews, Science **260**, 1637 (1993)
29. D.L. Minor, P.S. Kim, Nature **367**, 660 (1994)
30. S. Costantini, G. Colonna, A.M. Facchiano, Biochem. Biophys. Res. Commun. **342**, 441 (2006)
31. K. Fujiwara, H. Toda, M. Ikeguchi, BMC Struct. Biol. **12**, 1 (2012)

32. E.J. Milner-White, R. Poet, Trends Biochem. Sci. **12**, 189 (1987)
33. E.J. Milner-White, J. Mol. Biol. **216**, 385 (1990)
34. J.S. Richardson, D.C. Richardson, Science **240**, 1648 (1988)
35. I. Hudáky, R. Kiss, A. Perczel, J. Mol. Struct. Theochem. **675**, 177 (2004)
36. G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, J. Mol. Biol. **7**, 95 (1963)
37. I. León, J. Millan, F. Castaño, J.A. Fernández, Chem. Phys. Chem. **13**, 3819 (2012)
38. I. León, E.J. Cocinero, J. Millan, A.M. Rijs, I. Usabiaga, A. Lesarri, F. Castaño, J.A. Fernández, J. Chem. Phys. **137**, 074303 (2012)
39. R. Chaudret, B. de Courcy, J. Contreras-García, E. Gloaguen, A. Zehnacker-Rentier, M. Mons, J.-P. Piquemal, Phys. Chem. Chem. Phys. **16**, 9876 (2014)
40. B.J. Miller, J.R. Lane, H.G. Kjaergaard, Phys. Chem. Chem. Phys. **13**, 14183 (2011)
41. J.R. Lane, J. Contreras-Garcia, J.-P. Piquemal, B.J. Miller, H.G. Kjaergaard, J. Chem. Theory Comput. **9**, 3263 (2013)
42. T.A. Halgren, J. Comput. Chem. **20**, 730 (1999)
43. *MacroModel*, version 9.9 (Suite 2012) (Schrödinger, LLC, New York, 2012)
44. Y. Zhao, D.G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008)
45. Y. Zhao, D.G. Truhlar, Acc. Chem. Res. **41**, 157 (2008)
46. R. Ditchfield, W.J. Hehre, J.A. Pople, J. Chem. Phys. **54**, 724 (1971)
47. W.J. Hehre, R. Ditchfield, J.A. Pople, J. Chem. Phys. **56**, 2257 (1972)
48. M.M. Francl, W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. DeFrees, J.A. Pople, J. Chem. Phys. **77**, 3654 (1982)
49. M.J. Frisch, J.A. Pople, J.S. Binkley, J. Chem. Phys. **80**, 3265 (1984)
50. S. Grimme, J. Comput. Chem. **27**, 1789 (2006)
51. S. Grimme, J. Antony, S. Ehrlich, H. Krieg, J. Chem. Phys. **132**, 1 (2010)
52. M. Head-Gordon, J.A. Pople, M.J. Frisch, Chem. Phys. Lett. **153**, 503 (1988)
53. M. Frisch et al., *Gaussian 09, Revision A.02* (Gaussian Inc., Wallingford, CT, 2009)
54. J. Contreras-García, E.R. Jonhson, S. Keinan, R. Chaudret, J.-P. Piquemal, D.N. Beratan, W. Yang, J. Chem. Theory Comput. **7**, 625 (2011)
55. W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. **14**, 33 (1996)
56. G.A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press Inc., New York, 1997)
57. W. Chin, F. Piuzzi, I. Dimicoli, M. Mons, Phys. Chem. Chem. Phys. **8**, 1033 (2006)
58. R.J. Lavrich, D.F. Plusquellic, R.D. Suenram, G.T. Fraser, A.R.H. Walker, M.J. Tubergen, J. Chem. Phys. **118**, 5497 (2003)
59. I. Usabiaga, J. González, P.F. Arnáiz, I. León, E.J. Cocinero, J.A. Fernández, Phys. Chem. Chem. Phys. **18**, 12457 (2016)
60. J. González, I. Usabiaga, P.F. Arnáiz, I. León, R. Martínez, J. Millan, J.A. Fernández, Phys. Chem. Chem. Phys. **19**, 8826 (2017)