Do the Neighboring Residues in a Polypeptide Affect the Electron Distribution of an Amino Acid Significantly? A Quantitative Study Using the Quantum Theory of Atoms in Molecules (QTAIM)

Luis Lorenzo, María J. González Moa, Marcos Mandado, and Ricardo A. Mosquera*

Departamento de Química Física, Facultade de Química, Universidade de Vigo, Lagoas-Marcosende s/n 36310-Vigo, Galicia, Spain

Received May 8, 2006

Geometries, as well as bond and atomic properties obtained with the atoms-in-molecules theory applied on B3LYP/6-31++G**//B3LYP/6-31G** charge densities, of the N-formyl amides of the nine tripeptides obtained by combining glycine, alanine, and serine around a central glycine residue were analyzed to check how the properties of the central residue are modified by other amino acids bonded to it. All of the molecules were optimized from an α-helix conformation that was also displayed by the optimized structure. Significant variations of the geometry (especially remarkable for dihedral angles) and atomic properties of the central glycine residue are observed when it is attached to a serine residue whose side chain is involved in a hydrogen bond.

INTRODUCTION

Natural proteins can be considered as lineal and aperiodical polymers whose primary elements (monomers) come from the group made by the 21 L-α-amino acids (Aa's).[1] The peptidic unity ($-C_i^{\alpha}-CO-NH-C_{i+1}^{\alpha}-$) responsible for the union between monomers and the *i* residues of the amino acids ($-NH-HRC_i^{\alpha}-CO-$) are the elements that describe the primary structures of the proteins.[2] This primary structure contains a great number of atoms that is normally beyond the limit that can be handled with quantum chemical calculations nowadays. Thus, the computational study of the conformational properties of polypeptides has been possible because of the use of molecular mechanics force fields that are specifically parametrized.[3−6] Such parametrizations assume the transferability of the properties for each of the residues and peptidic units, keeping common properties in different environments and even in different polypeptides.

Although several theories, such as the Stone's distributed multipole analysis,[7,8] have been used to study the assumed transferability in polypeptides,[9,10] the quantum theory of atoms in molecules (QTAIM)[11−13] provides a tool to analyze the transferability that is only based upon first physical principles.[14,15] Therefore, some works have been carried out with this method to check the reliability of the Aa approximate transferability. Thus, Popelier and Bader analyzed with QTAIM the electron density of one specific amino acid sequence (N-formyltriglycine amide) in different conformations, finding no important variations of the atomic properties along the conformational change.[16] In another QTAIM study, Chang and Bader considered HF/4-31G**//HF/4-31G charge densities for the conformers with main dihedral angles of 180° for N-formyl amides of alanine (Ala) and glycine (Gly); dipeptides Ala−Gly, Gly−Ala, and Gly−Gly; and tripeptide Gly−Gly−Gly, concluding that the geometry and atomic and bond properties are nearly transferable among these systems.[17] Recently, Matta and Bader reported an exhaustive investigation on the geometry and atomic and bond properties of the genetically encoded amino acids in a series of papers.[18−20] Part I of this series reviews QTAIM (a reader unfamiliar with this theory is referred to this paper[18] or to refs 11−13) and establishes the transferability of the geometric, atomic, and bond properties among the diverse conformers of leucine and the short-range nature of the perturbations induced by the formation of its zwitterionic forms. A high degree of transferability among the genetically encoded amino acids was obtained for the geometry of the main chain in part II.[19] A similar transferability of the side chains is found for common functional groups of two or more amino acids.[19] The atomic and bond properties of amino acids presented in part III indicate that the main-chain group (MCG) $NH_2-HC^{\alpha}-COOH$ can be regarded as a common group. The comparison of the results obtained for Gly−Aa−Gly tripeptides with those computed for the isolated Aa is employed to check the reliability of the MCG transferability.[20]

The influence of neighboring Aa's on the central residue has only being studied focusing on geometry aspects[21,22] but not using electron densities. Thus, the main goal of this paper is to analyze if the QTAIM properties of the Aa residues and peptidic unities are independent of the Aa's that surround them. To this end, a series of nine tripeptides where Gly is the central residue bonded to other Aa's (glycine, alanine, and serine) was considered. We have also upgraded the computational level to include electron correlation.

On the other hand, it has been shown that the employment of atomic normalized Shannon entropies of the atomic electronic distribution computed in the position space within the framework of the QTAIM[23] is a useful tool to define nearly transferable atoms. Besides, the analysis of this

* Corresponding author. Fax: +34986812321. E-mail: mosquera@uvigo.es.

10.1021/ci600184t CCC: $33.50
Published on Web 08/04/2006

quantity has led to the modification of previously established conclusions on the approximate transferability of *n*-alkanes.[24]

Because of the size of this kind of system, tripeptides made from the union of small amino acids (glycine, G; alanine, A; serine, S) were taken as models. The amino acids located at the ends of the chain were bonded (through peptidic bonds) to hydrogen atoms, to yield the corresponding N-formyl-tripeptide amide, which allows the model to approach the situation of the amino acids in a protein.

COMPUTATIONAL DETAILS

The charge densities used here were calculated with the program Gaussian 03[25] at the B3LYP/6-31++G** level over B3LYP/6-31G** completely optimized geometries. Atomic, bond, and ring properties were determined using the program AIMPAC.[26] This program was modified to allow the calculation of the unitary Shannon entropy in the space of positions for the electronic distribution in each atom, $Sh(\Omega)$.

According to the QTAIM, the molecular electron density is partitioned into disjointed atomic fragments, Ω. Each of them contains a charge density maximum (or attractor) that is usually placed in the proximity of the nucleus. The atomic fragments are delimited by surfaces defined by a zero-flux condition for the gradient of the charge density and, with some exceptions, by an isocontour where the electron density vanishes (i.e., 10^{-5} au). The zero-flux surfaces contain certain critical points of the charge density, located between every two bonded nuclei and characterized by having two negative eigenvalues for the Hessian matrix of the charge density, that are known as bond critical points (BCPs). The properties of these points characterize the corresponding chemical bonds. When a set of bonds forms a ring, a charge density critical point with only one negative eigenvalue of the Hessian matrix appears. This is called a ring critical point (RCP).

Values of the following atomic properties for the central glycine residue atoms were analyzed in this study: $Sh(\Omega)$; electronic population, $N(\Omega)$; energy, $E(\Omega)$; dipole moment, $\mu(\Omega)$; and the volume delimited by the zero-flux surfaces and a contour with a density charge of 0.001 au, $v(\Omega)$. We have also considered the values obtained for several properties at the BCP of the central residue and its bonds to the N-terminal and C-terminal residues: the charge density, ρ_c; its laplacian, $\nabla^2\rho_c$; and the energy density, H_c. Bond lengths, R; the distances between the nuclei and the BCP, r_c; and the bond ellipticity, ϵ, were also studied.

The systems included in this work are named by the sequence of the constitutive amino acids (expressed by their one-letter symbols), starting from that one bonded to the N-terminal extreme in the central glycine and finishing on that one attached to the C-terminal extreme. In some cases, the individual residues are named using the initial followed by the number that identifies its position in the sequence.

In all cases, the resulting conformer of the tripeptide optimized starting from an α-helix without restrictions was considered. The effect of the conformation over the properties of the central residue in all of the studied systems is expected to be similar to that one observed in triglycine.[16] To describe the disposition of the conformer obtained, the Ramachandran angles ϕ_i ($C^{i-1}-N^i-C_\alpha{}^i-C^i$) and ψ_i ($N^i-C_\alpha{}^i-C^i-N^{i+1}$) that describe each residue i, together with other dihedral angles that connect two residues, ω_{i-i+1}($C_\alpha{}^i-C^i-N^{i+1}-C_\alpha{}^{i+1}$), of the main chain (Figure 1) were used. Other dihedral angles, χ_{ij}, are used to indicate the disposition of the side chain in the alanine ($\chi_i = H-C-C_\alpha{}^i-C^i$) and serine ($\chi_{i1} = O-C-C_\alpha{}^i-C^i$), and $\chi_{i2} = (H-O-C-C_\alpha{}^i)$.

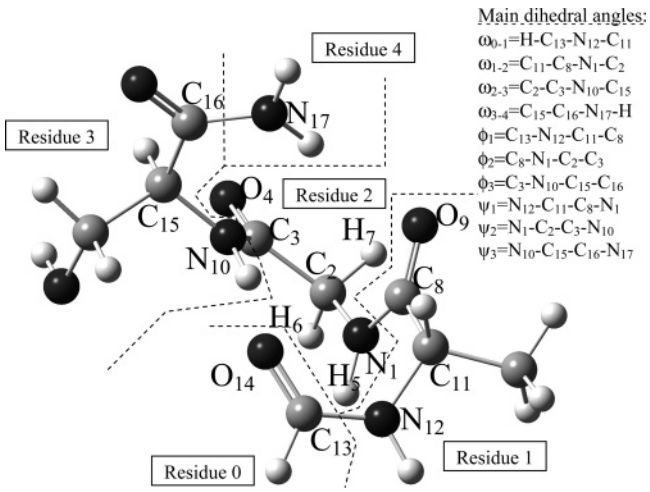

Figure 1. Main dihedral angles and atom numbering used in this study for the main chain of N-formil amides of the nine tripeptides here considered (AGS in this case) for all of the atoms in the central glycine residue G_2 and their neighboring atoms in the carbonyl group of residue 1 (alanine A_1 in this case) and the amidic nitrogen of residue 3 (serine S_3 in this case). Dashed lines separate the five residues in these molecules.

RESULTS

Structures Analyzed and Energy Additivity. The conformers obtained for the nine tripeptides from the complete optimization of the initial α-helix conformation have characteristic Ramachandran angles for an α-helix structure (Table 1).

The total electronic energies obtained at the B3LYP/6-31++G**//B3LYP/6-31G** level for the nine N-formil-tripeptide amides considered can be reproduced by fitting eq 1. In this equation, n_G, n_A, and n_S represent, respectively, the number of glycine, alanine, and serine residues in the compound of total energy E. $E(HCONNH_2)$, $E(G)$, $E(A)$, and $E(S)$ are the fitting parameters and represent the average values for the energy of formamide (actually, for the addition of terminal groups HCO− and $-NH_2$), glycine, alanine, and serine residues. In this case, the values obtained were −169.8464, −247.3690, −208.0490, and −322.5845 au, respectively.

$$E = E(HCONH_2) + n_G E(G) + n_A E(A) + n_S E(S) \qquad (1)$$

The calculated B3LYP/6-31++G**//B3LYP/6-31G** energies are well-fitted by the previous equation as it is demonstrated by several statistical parameters. Thus, the r^2 coefficient for eq 1 differs from the unity in 1.0×10^{-10}, and the probability that this excellent lineal adjustment is due to random coincidence is estimated by parameter F as 7.8×10^{-16}. In addition, the differences between computed energies and those obtained by the fitting equation do not reach in any case 1 kcal mol^{-1} (Table 1). All of these data point to the suitability of the energetic additivity models for this series of compounds.

Table 1. Disposition of the Main Dihedral Angles (degrees) after B3LYP/6-31G** Optimization, Total Electronic Energy (au), *E*, after the B3LYP/6-31++G** Single-Point Calculation, and Molecular Virial Ratio, γ[a]

	AGA	AGG	AGS	GGA	GGG	GGS	SGA	SGG	SGS
$-E$	872.6332	833.3140	947.8493	833.3126	793.9934	908.5276	947.8493	908.5300	1023.0640
$10^3(\gamma-2)$	9.329	9.304	9.237	9.302	9.275	9.223	9.233	9.205	9.158
ΔE	0.5	−1.6	−1.1	2.1	0.0	3.4	−1.1	−2.4	1.1
ω_{0-1}	−168.2	−168.0	−168.7	−169.1	−168.9	−169.4	−164.6	−164.7	−165.0
ϕ_1	−68.0	−67.9	−67.0	−66.8	−66.8	−65.6	−59.5	−59.8	−59.4
ψ_1	−21.6	−20.2	−24.2	−22.5	−21.1	−24.9	−33.3	−31.6	−34.0
ω_{1-2}	175.8	175.4	176.5	176.0	175.5	175.4	−178.7	−179.5	−179.7
ϕ_2	−70.3	−68.4	−71.9	−70.4	−68.5	−76.4	−69.8	−67.8	−74.9
ψ_2	−6.1	−8.9	−1.3	−5.9	−8.8	9.2	−6.9	−9.8	7.9
ω_{2-3}	173.2	171.2	176.0	173.3	171.6	175.5	173.8	171.8	173.4
ϕ_3	−104.7	−101.5	−117.0	−105.0	−102.4	−140.8	−105.1	−102.7	−139.6
ψ_3	11.7	10.6	13.8	11.8	11.3	26.2	11.5	10.6	26.5
ω_{3-4}	176.6	176.6	177.1	176.7	176.6	178.7	175.5	175.2	177.0
χ_1/χ_{11}	172.3	172.1	172.7				−175.4	−175.6	−175.5
χ_{12}							−177.9	−178.0	−177.1
χ_3/χ_{31}	168.9		175.7	168.8		−177.0	169.1		−177.5
χ_{32}			−62.2			−40.0			−38.1

[a] ΔE (kJ mol^{-1}) is the difference between *E* and the energy obtained by fitting eq 1. Subscript 1 refers to the amino acid bonded to the N-terminal ending of the central glycine, and subscript 3 refers to the one bonded to the C-terminal ending. The numbers 0 and 4 refer, respectively, to the HCO and NH_2 groups bonded to the tripeptidic N-terminal and C-terminal endings.

Table 2. Central Glycine Residue B3LYP/6-31G** Geometry[a]

	GGG	AGA	AGG	AGS	GGA	GGS	SGA	SGG	SGS
C2−C3	1.5406	1.0		−1.0	0.6	−3.4			−3.8
C2−H6	1.0925					−0.8			−0.9
C2−H7	1.0925			0.5		1.0			1.3
C3−N10	1.3604	−1.9		−3.3	−1.8	−4.6	−3.0	−1.1	−6.8
C3−O4	1.2241	0.7		1.3	0.7	5.2	1.1		5.9
C8−N1	1.3560			1.6	0.5	3.5		−0.7	2.4
C8−O9	1.2273	1.3	1.7	0.5		−2.0	1.9	2.3	0.8
N1−C2	1.4541					−1.6	0.9	0.8	−0.6
N1−H5	1.0088								0.7
C2−C3−N10	117.00	−0.18		−0.56	−0.15	−0.19	−0.17		−0.16
C2−C3−O4	118.45	−0.11		0.31	−0.13	0.42	−0.21	−0.13	0.30
C8−N1−C2	121.20	0.08		0.34		0.60	0.35	0.23	0.90
C8−N1−H5	118.43	−0.23	−0.17	−0.25	−0.07	−0.31	−0.11	−0.05	−0.30
H5−N1−C2	119.82		0.12	−0.15		−0.61	0.06	0.12	−0.47
H6−C2−C3	108.25	−0.07		−0.19	−0.10	−0.38	−0.24	−0.11	−0.31
H7−C2−C3	105.53		−0.07	0.26	0.06	0.46		−0.12	0.17
H7−C2−H6	107.71			0.05		0.09	−0.12	−0.08	
N10−C3−O4	124.19	0.38	0.64	0.57	0.63	0.36	0.71	0.53	0.10
N1−C2−C3	116.61	0.11		−0.07	0.16		0.57	0.41	0.44
N1−C2−H6	108.80	0.13		0.27	0.11	0.61	−0.20	−0.28	0.30
N1−C2−H7	109.61	−0.25		−0.34	−0.25	−0.82	−0.08	0.12	−0.63
O9−C8−N1	123.52	−0.12	−0.20			0.32	−0.41	−0.50	−0.17
C8−N1−C2−C3	−68.5	−1.7		−3.4	−1.9	−7.9	−1.3	0.8	−6.5
C8−N1−C2−H6	168.8	−1.8		−3.3	−2.0	−7.9	−1.2	0.8	−6.6
C8−N1−C2−H7	51.2	−1.7		−3.3	−1.8	−7.8	−0.9	1.0	−6.4
H5−N1−C2−C3	102.8	−2.6	−0.6	−3.8	−2.2	−10.0	1.6	3.6	−5.3
H5−N1−C2−H6	−19.1	−2.7	−0.7	−3.7	−2.3	−10.0	1.6	3.6	−5.5
H5−N1−C2−H7	−137.4	−2.7	−0.7	−3.8	−2.3	−10.1	1.9	3.8	−5.3
H6−C2−C3−N10	114.2	2.8		7.7	3.1	18.5	1.9	−1.2	17.2
H6−C2−C3−O4	−66.3	2.4		5.9	2.7	15.0	1.7	−1.0	13.6
H7−C2−C3−N10	−130.7	2.8		7.8	3.1	18.7	1.6	−1.4	17.1
H7−C2−C3−O4	48.8	2.4		6.0	2.7	15.1	1.5	−1.2	13.5
N1−C2−C3−N10	−8.8	2.6		7.5	2.9	18.0	1.9	−1.1	16.7
N1−C2−C3−O4	170.8	2.1		5.6	2.4	14.4	1.7	−0.9	13.0
O9−C8−N1−C2	−5.2						3.4	2.9	2.5
O9−C8−N1−H5	−176.6	0.7		0.8	0.5	1.5	0.6		1.4

[a] Absolute values for GGG and those relative to them for the rest of the tripeptides. Distances in Å and angles in degrees. Differences in length multiplied by 10^3. Differences inferior to 5×10^{-4} Å (bonds), 0.05° (bond angles), and 0.5° (dihedral angles) are not shown.

As transferability implies relatively common geometries, it is necessary to evaluate the constancy in the geometry of the central glycine residue, G_2 (Table 2). In fact, little variations can be observed in the geometry of G along the different tripeptides here studied, with regard to that displayed in GGG. These variations are concentrated in the C_3-N_{10}, $C_3=O_4$, and C_8-N_1 bonds for the tripeptides containing a serine residue. The most affected bond angle is $N_{10}-C_3=O_4$, which experiences sensitive variations (but always smaller than 1°) when a glycine residue is replaced by serine

Table 3. Main Characteristics (B3LYP/6-31++G**//B3LYP/6-31G**) of the Hydrogen Bonds Responsible for the α-Helix Structure[a]

bond		AGA	AGG	AGS	GGA	GGG	GGS	SGA	SGG	SGS
$N_{10}-H\cdots O_{14}$	R/Å	2.138	2.135	2.111	2.148	2.143	2.087	2.049	2.042	2.009
	θ	166.3	167.3	165.5	166.4	167.5	166.0	167.6	168.6	167.7
	$10^2\rho_c$	1.59	1.60	1.69	1.55	1.57	1.77	1.91	1.93	2.08
	$10^2\epsilon$	4.9	4.7	5.1	5.0	4.8	5.4	5.1	5.0	5.3
	$10^2\nabla^2\rho_c$	4.95	4.95	5.21	4.86	4.88	5.45	5.98	6.05	6.49
	10^3H_c	0.11	0.10	0.07	0.15	0.14	0.07	0.12	0.13	0.14
$N_{17}-H\cdots O_9$	R/Å	2.240	2.204	2.356	2.248	2.224	2.615	2.210	2.173	2.460
	θ	161.3	162.8	163.4	161.7	162.8	163.4	160.9	161.8	162.1
	$10^2\rho_c$	1.27	1.36	0.98	1.24	1.30	0.56	1.34	1.45	0.78
	$10^2\epsilon$	6.0	6.2	5.5	6.1	6.2	14.9	6.5	6.7	7.3
	$10^2\nabla^2\rho_c$	4.07	4.32	3.32	4.01	4.16	2.20	4.31	4.59	2.83
	10^3H_c	0.30	0.21	0.55	0.33	0.28	0.91	0.26	0.19	0.77

[a] ρ_c, $\nabla^2\rho_c$, and H_c in au.

or alanine. Again, the replacement by serine gives higher geometrical modifications in the set of bond angles, whereas the dihedral angles experience more significant modifications (close to 20°) in serine-containing tripeptides.

In all cases, the modifications obtained after the B3LYP optimization were higher than the ones observed after HF/6-31G* optimization, carried out during the first steps of this work. At this calculation level, the highest difference in the bond distances occurs in $C_3{=}O_4$ in the SGS system and only reaches 2.4×10^{-3} Å. For bond angles, the higher difference, 0.7°, corresponds to the $N_1-C_2-C_3$ angle of the SGA system. These variations are on the order of magnitude found by Bader and Chang from a HF/4-31G study of flat dipeptides.[17] For the dihedral angles, the systems containing C-terminal serine (S_3) present variations around −20° in the ϕ_{+1} angle, while the ones that contain N-terminal serine show variations around −13° in the ψ_{-1} angle. The GGS system also presents other noticeable alterations in the torsion angles of the central residue, G_2: around 13° for ψ_2 and $N-C_\alpha-C{=}O$ and −11° for $H-N-C_\alpha-C$ and ϕ_2. The AGS system presents the same trend, but with smaller variations (9 and −7°, respectively). ω angles, which indicate the planarity of the peptidic bond, are smaller than 5° (reaching this value for ω_{2-3} in SGA, though) and are what indicate the planarity of these bonds.

In conclusion, the geometrical analysis shows that, except for the case of the main dihedral angles of tripeptides GGS and SGS, the variations in the valence geometry are irrelevant in the central glycine residue.

Intramolecular Hydrogen Bonds. The α-helix structure is characterized by certain values of the Ramachandran angles but also by the presence of hydrogen bonds (HBs) between the amidic H of residue $i + 1$ and the carbonylic oxygen of residue $i - 1$, which generates 10-member cycles. The topological analysis of the charge density with the QTAIM lets us locate the BCPs due to these HBs and the RCPs of the corresponding cyclic structures. Thus, in all of the considered tripeptides, we found two HBs, one between the carbonylic oxygen of the formyl group on the N-terminal side and the amidic hydrogen of residue 3 ($O_{14}\cdots H-N_{10}$ bond in Figure 1 and bonds of type I in Figure 2). The other HB is established between the carbonylic oxygen of residue 1 and one of the amidic hydrogens bonded on the C-terminal side ($O_9\cdots H-N_{17}$ bond in Figure 1 and bonds of type II in Figure 2).

All of the HBs present a bond path and have been characterized through the corresponding BCP (Table 3). The analysis of this table shows that the geometrical characteristics of the hydrogen bonds are kept in all cases,[27] with distances smaller than 2.5 Å (if we exclude GGS) and donor−H-acceptor angles (θ) close to 180°. Also, HBs of type I are slightly stronger, with bond distances, R, shorter than those of the second type (II) and higher BCP charge densities. For both types of bonds, a linear dependence between ρ_c and R^{-1} can be observed (Figure 2). In the first case, the θ angles are also closer to linearity. On the other hand, in both classes of bonds, the values for ϵ, $\nabla^2\rho_c$, and H_c agree with what can be expected for a HB: low ellipticities, low and positive values for $\nabla^2\rho_c$, and almost zero values for H_c. In addition, the rest of the criteria described for HBs by Koch and Popelier[28] and by Bader and Carroll[29] are met (data not shown). Finally, the HBs of type II are weakened in the GGS and SGS tripeptides, as the longer bond distances and lower ρ_c indicate. This weakening also gives higher ellipticity, less positive values for $\nabla^2\rho_c$, and a reinforcement of the positive character of H_c.

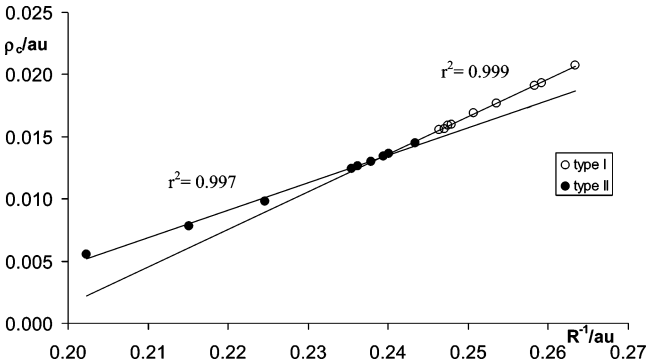

Figure 2. Representation of ρ_c vs $1/R$ for the hydrogen bonds that are part of the α-helix structure. Type I refers to $O_{14}\cdots H-N_{10}$ and type II to $O_9\cdots H-N_{17}$.

Besides these HBs, we have found other bond paths between two nonbonded atoms in the Lewis structure (Table 4). Some of them correspond to HBs between the oxygen of the carbonylic group and the OH of the serine side chain (GGS and SGS). These bonds were not detected when the OH orientation of the side chain was not the adequate (AGS, SGA, and SGG). This new HB between residues 2 and 3 of GGS and SGS can be interpreted as the factor that debilitates the $N_{17}-H\cdots O_9$ bond. Thus, this region of the molecule adopts a compromise geometry that allows the formation of both HBs, but none of them in optimal conditions. This interpretation also agrees with the fact that the higher modifications observed in the geometry of the G_2 residue

Table 4. Main Characteristics (B3LYP/6-31++G**//B3LYP/6-31G**) of Other Bonds and Interactions that Do Not Correspond to the Lewis Structure of the Studied Tripeptides[a]

	type	interaction	R/Å	ΔR/Å	θ	$10^2\rho_c$	$10^2\epsilon$	$10^2\nabla^2\rho_c$	10^3H_c
AGA	NH···N	N12···H5−N1	2.355	0.232	105.2	1.55	1.4	6.36	2.28
AGA	CH···O	C−H···O14	2.969	0.033	132.0	0.32	0.6	1.28	0.72
AGG	NH···N	N12···H5−N1	2.352	0.225	105.2	1.56	1.3	6.36	2.28
AGS	NH···N	N12···H5−N1	2.361	0.266	105.0	1.55	1.7	6.39	2.35
AGS	N···N	N1···N10	2.730	0.315		1.51	14.1	6.78	2.86
GGA	N···N	N12···N1	2.819	0.362		1.46	5.5	6.32	2.53
GGA	CH···O	C−H···O14	2.988	0.034	132.0	0.31	0.6	1.23	0.71
GGG	N···N	N12···N1	2.818	0.378		1.47	4.1	6.28	2.49
GGS	N···N	N12···N1	2.818	0.374		1.47	4.6	6.34	2.52
GGS	OH···O	O−H···O4	2.371	0.063	130.8	1.14	0.1	4.00	0.68
SGA	CH···O	C−H···O14	2.984	0.038	129.9	0.32	0.8	1.27	0.72
SGS	OH···O	O−H···O4	2.354	0.054	132.6	1.17	0.1	4.03	0.58

[a] ΔR = length of the bond path − internuclear distance (R). ρ_c, $\nabla^2\rho_c$, and H_c in au.

Table 5. Electronic Population (au) of the Different Atoms of Residue G_2 in the Tripeptides of This Study[a]

	GGG	AGA	AGG	AGS	GGA	GGS	SGA	SGG	SGS	$10^4L_m(\Omega)$
N_1	8.2417	0.7	1.1	−0.8	−0.1	−3.5	−0.9	−0.6	−5.3	3.7
C_2	5.5975	0.9	0.3	−1.0	1.5	−5.3	0.1	−1.2	−5.4	*b*
C_3	4.4821	−1.1	−1.4	−5.3	−1.3	4.2	−5.9	−3.6	−1.3	18.5
O_4	9.2184	2.0	0.4	7.1	1.7	7.4	2.2	0.5	8.6	0.8
H_5	0.5698	0.0	−1.0	−1.0	0.6	−0.2	−4.0	−4.9	−5.4	1.3
H_6	0.9689	0.9	0.5	−2.2	0.2	−5.9	0.4	0.2	−6.0	0.3
H_7	0.9368	0.8	0.6	1.0	0.0	−2.8	3.9	3.5	0.7	0.2
C_8	4.4765	13.8	13.8	13.8	0.0	1.5	6.8	6.8	9.6	*c*
O_9	9.2213	2.8	3.4	3.3	−0.5	−2.6	−2.0	−1.4	−2.2	1.1
N_{10}	8.2519	0.3	0.7	−3.1	−0.5	−2.5	8.3	7.8	6.0	4.8

[a] Absolute values for GGG and those relative to them (multiplied by 10^3) for the rest of the compounds. L_m(C) represents the maximum absolute value for $L(\Omega)$ in each case. [b] $N^0(\Omega)$ values obtained using independent lineal adjustments for each atom of the type $N(\Omega) = N^0(\Omega) + a\,L(\Omega)$. [c] $N^0(\Omega)$ values obtained as in b but considering the following groups of compounds: (GGG, AGA, and AGG), (AGS and GGG), (GGS), (SGA and SGG), and (SGS).

correspond to the elongation of the $C_3{=}O_4$ bond in the tripeptides GGS (5.2×10^{-3} Å) and SGS (5.9×10^{-3} Å).

Weak HBs between the oxygen of one carbonyl group and a C−H bond of the side chain of alanine are also observed. This bond occurs on O_{14}, when the alanine residue acts as residue 3 in the chain, forming a 12-member cycle. If we consider the participation of O_{14} in every case and also in a HB with $H-N_{10}$ stabilizing the α-helix structure, it can be concluded that O_{14} has the structure of a bifurcated hydrogen-bond acceptor.[30]

The rest of the bond paths localized correspond to different interactions. For instance, bond paths between nitrogen atoms of contiguous residues were located. Even though the presence of similar interactions (different from HBs or bonds present in the Lewis structure) in other compounds originated different interpretations in the bibliography,[31−34] the clear positive character of H_c, in this case, indicates that these interactions occur among complete shells. Even the increase of the positive value of H_c, when the intramolecular HBs of the α-helix become weaker, points to a consideration that these very positive values associate the N···N bond paths with repulsive interactions. In agreement with this idea, the N−H···N bond paths found in different tripeptides with geometrical (θ angles closer to a right angle than to a lineal one) and topological parameters improper of HBs can also be associated with repulsive interactions.

The values of the properties at the BCP shown in Table 4 let us distinguish four different kinds of interactions: O−H···O and C−H···O HBs and N···N and N−H···N repulsions. Thus, it can be noticed that, even though the interatomic distances of the N−H···N interactions are similar to the ones in O−H···O, the first ones present longer bond paths. These differences in the length of the bond path and the internuclear distance are over 0.3 Å for N···N interactions, 0.2 Å for N−H···N interactions, 0.02−0.03 Å for C−H···O interactions, and slightly over 0.05 Å for O−H···O HBs. On the other hand, the interaction angles, θ, that we propose as repulsive are much smaller (around 105°) than the weak HB angles (over 130° in O−H···O and C−H···O, which are still far from linearity though).

The ρ_c is higher than the one for the HBs in the α-helix for the repulsive interactions, but it is smaller for the weak HBs (between 1.1×10^{-2} and 1.2×10^{-2} for O−H···O and around 10^{-3} for C−H···O). The ellipticities change significantly in these bonds: over 10^{-2} in N···N, around that value in N−H···N, smaller than 10^{-2} in C−H···O, and smaller than 10^{-3} in O−H···O. Finally, both H_c and $\nabla^2\rho_c$ are clearly positive in the repulsive interactions and almost zero in the weak bonds shown in Table 4.

The formation of hydrogen bonds with the side chains of the residues is the factor that originates the only significant geometrical variations observed. So, the next objective is to study the influence of such HBs over the electronic distribution of the central glycine residue, G_2.

Electronic Distribution in the Central Residue. The B3LYP/6-31++G** charge densities were analyzed using the diverse atomic properties calculated with the QTAIM. In this analysis, it was necessary to pay attention to the effect that the error in the integration and determination of the interatomic surfaces in the molecule, measured by $L(\Omega)$, has

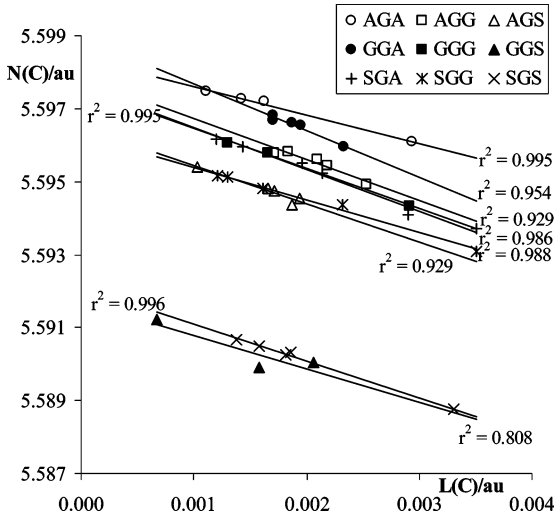

Figure 3. $N(C)$ vs $L(C)$ plots for diverse integrations of the α carbon, C_2, at the central glycine residue.

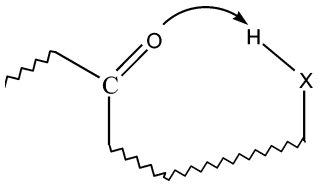

Figure 4. Traditional representation of the hydrogen-bond formation.

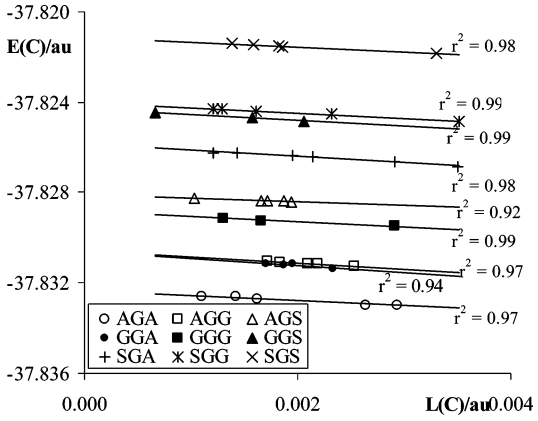

Figure 5. Representation of $E(C)$ vs $L(C)$ for several integrations of C_2. Relative values of the intercept with regard to that of GGG (−37.828 23 au) are (in kJ mol^{-1}) −9.2, −4.5, 1.5, −4.6, 11.9, 7.8, 12.7, and 20.3 for, respectively, AGA, AGG, AGS, GGA, GGS, SGA, SGG, and SGS.

over the values computed for the properties of the molecule. Specifically, it was very difficult to reduce the absolute value of $L(\Omega)$ below 10^{-3} au for the carbons in the central residue and several atoms in its vicinity, even after performing multiple integrations for each atom with different integration parameters and considering the possibility of more than one intersection of the integration rays with the zero-flux surface. It was investigated whether, as it was communicated by Aiken and Popelier,[35] an increase in the number of Gaussian quadrature points used in the integration, or in the group of gradient trajectories originated by the proper vectors associated with the negative proper values in the BCPs that are used to delimit the interatomic surfaces, does not guarantee a reduction of the absolute value of $L(\Omega)$. Moreover, it was found that $L(\Omega)$ behaves erratically with the rigor of the integration process.

In these conditions, the individual $N(\Omega)$ values can lead to false conclusions over the constancy in the charge distribution properties of the tripeptides. To avoid this problem, we took advantage of the lineal relations between $N(\Omega)$ and $L(\Omega)$, previously proposed by our group[36] and confirmed for diverse homologous series.[24,37−41] Figure 3 shows the application of these relations to the α carbon of the central residue (C_2), one of the atoms that has presented more problems during the integration. It can be observed that there are two correlation lines with a significantly lower intercept. They correspond to the tripeptides where the serine side chain is involved in an O−H···O_4 HB (SGS and GGS).

Table 5 contains the electronic populations obtained for the different atoms of residue G_2 and for the neighbor atoms (C_8, O_9, and N_{10}). It is estimated that these populations are affected by an error of 10^{-3} au, except in the case of C_3, where it can reach 3×10^{-3} au. These populations were obtained through a lineal fitting for C_2 and C_8 and, in the rest of the cases, from the integrated $N(\Omega)$ value obtained with the lowest $|L(\Omega)|$. Table 5 also shows the highest absolute values of $L(\Omega)$ in each series of atoms.

Inside the G_2 residue, the largest $N(\Omega)$ variations correspond to O_4 in the two compounds where it participates in

Table 6. Absolute Values (au) of $\mu(\Omega)$ for the Atoms in Residue G_2 in GGG and Those Relative (multiplied by 10^3) to Them for the Rest of the Tripeptides[a]

	GGG	AGA	AGG	AGS	GGA	GGS	SGA	SGG	SGS
N_1	0.1826					−5.1			−5.8
C_2	0.4615								
C_3	0.7196					12.6	−6.5	−5.2	8.4
O_4	0.4323			−9.5		−21.8	−6.4		−24.2
H_5	0.1758								
H_6	0.1312								
H_7	0.1302								

[a] Relative values inferior to $\pm 5.0 \times 10^{-3}$ au are not shown.

Table 7. Variations of the Shannon Entropies for the Special Electronic Distribution (in au) of C_2, $\Delta Sh(\Omega)$, in the Different Tripeptides with Indication of the Maxima and Minima Values of $L(\Omega)$ (in au)

	AGA	AGG	AGS	GGA	GGG	GGS	SGA	SGG	SGS
$10^3\Delta Sh(C_2)$	2.5	1.4	1.2	1.4	2.1	1.8	3.3	2.2	2.3
$10^3L_1(C_2)$	1.1	1.7	1.0	1.7	1.3	0.7	1.2	1.2	1.4
$10^3L_2(C_2)$	2.9	2.5	1.9	2.3	2.9	2.1	3.5	3.5	3.3

Table 8. Variations Observed in $Sh(\Omega)$ of the Atoms from the Central Residue and Neighbors in the Series of Tripeptides[a]

	$Sh_2(\Omega)$	$Sh_1(\Omega)$	$\Delta Sh(\Omega)$	$10^4L_2(\Omega)$	$10^4L_1(\Omega)$
N_1	2.827	2.825	0.003	17.1	6.7
C_2	2.322	2.317	0.006	18.5	−2.2
C_3	1.732	1.726	0.006	12.3	−1.1
O_4	2.784	2.781	0.003	1.3	−0.3
H_5	2.347	2.338	0.009	0.3	−0.02
H_6	2.907	2.893	0.014	0.2	0.1
H_7	2.852	2.841	0.011	3.7	0.4
C_8	1.727	1.719	0.008	4.8	−2.3
O_9	2.778	2.767	0.011	0.8	0.1
N_{10}	2.831	2.816	0.015	1.1	0.1

[a] Maxima, $Sh_2(\Omega)$, and minima, $Sh_1(\Omega)$, values and the amplitude of the interval, $\Delta Sh(\Omega)$, are indicated together with the corresponding maxima and minima values of $L(\Omega)$ in au

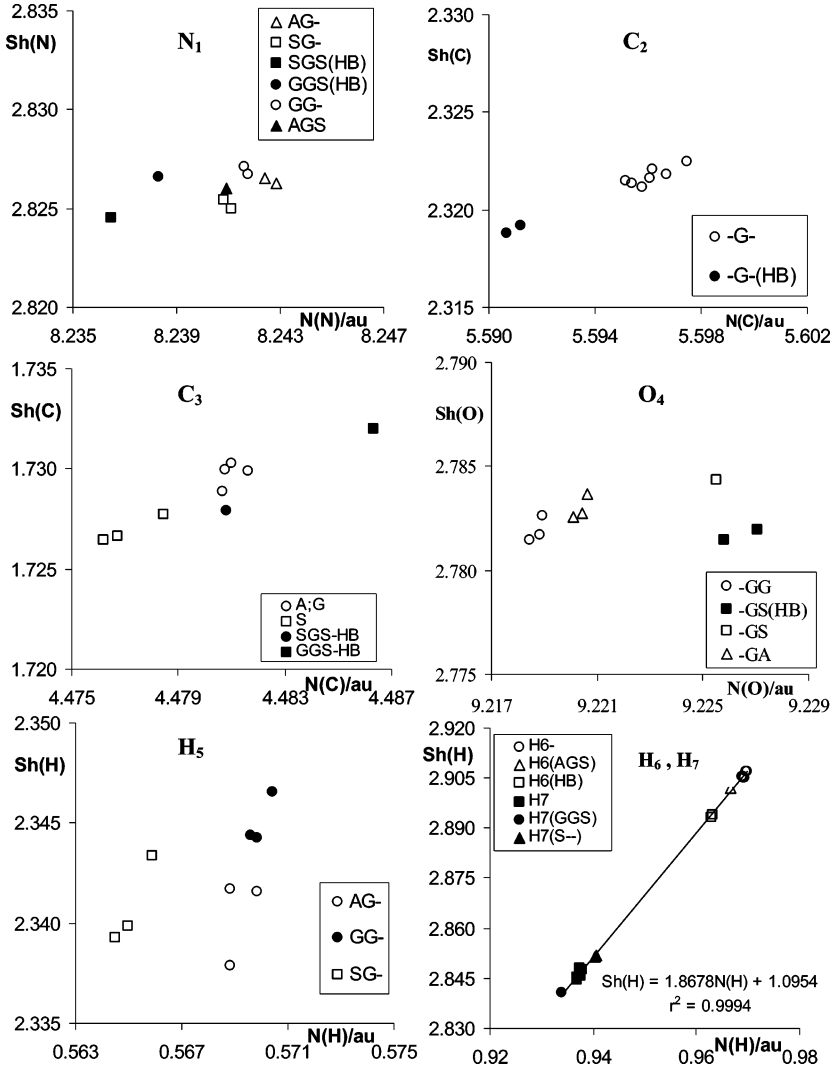

Figure 6. *Sh*(Ω) vs *N*(Ω) plots for the atoms in residue G_2 in the series of tripeptides.

an O−H···O_4 bond with the side chain of serine (GGS and SGS). We can also observe that the formation of this HB gives rise to important variations in the electron population of N_1, C_2, and H_6. It should be noticed that, as previously found for other systems with QTAIM,[42,43] the electronic population of the HB acceptor, $N(O_4)$, increases in the compounds where this HB exists. This fact contradicts the traditional interpretation of the HBs as a partial donation of a hydrogen from an X−H bond to the acceptor atom, O_4, that shares (at least partially) one of the electron pairs with this hydrogen (Figure 4). In agreement with the QTAIM results, the formation of the HBs comes together with a reduction in the electronic population of atoms C_2 and H_6. H_6 is closer to the hydroxyl placed on the same side of the plane of the peptidic bond between residues 2 and 3.

The union of G_2 to a serine residue (as residue 1 or 3) also originates variations that are higher than the maximum error that can be attributed to the data in Table 5, even when there is no HB with the side chain. The variation experienced by C_3 and O_4 in AGS and C_3 and H_5 in SGA and SGG is worth mentioning.

The effect of $L(\Omega)$ over $E(\Omega)$ was also analyzed in the atom C_2. An excellent lineal correlation between both values was found (Figure 5). In this case, it also has to be considered that atomic energies are computed by correcting atomic electron kinetic energies, $K(\Omega)$, with the molecular virial ratio, γ. When virial ratios span over a wide range, this procedure can lead to undesirable artifacts, as pointed out by Mandado et al.[44] Although these problems can be avoided using self-consistent virial scaling optimizations,[45] the variations displayed by γ in this work are small enough (Table 1) to guarantee that they affect the relative atomic energies by less than 34, 25, 17, and 0.3 kJ mol^{-1} for O, N, C, and H, respectively.

Thus, $E(C_2)$ for SGS is the only intercept in Figure 5 that can be considered significantly different from the corresponding value in GGG. The other system with O−H···O_4 HB (GGS) displays an intercept that is close to the significance limit.

Looking at the $E(\Omega)$ values presented by the remaining atoms of G_2 and those directly bonded to them (data not shown), we observe that some atoms (like N_1 and O_9) display variations that are below the significance limit in all of the tripeptides here considered. On the contrary, the highest variations (certainly over the significance limit) correspond to O_4 in GGS and SGS (52 and 63 kJ mol^{-1}, respectively). This reinforces the specific nature of part of the G_2 residue in these compounds that includes a HB with the S_3 side chain.

Table 9. Main QTAIM Properties of the G_2 Residues[a]

	GGG	AGA	AGG	AGS	GGA	GGS	SGA	SGG	SGS	δ_m
$N(G_2)$	30.015	4.1	0.5	−2.3	2.5	−6.1	−4.2	−6.1	−14.0	0.047
$Sh(G_2)$	17.756	0.2	−7.2	−6.5	2.1	−17.2	1.8	−2.6	−22.2	0.125
$E(G_2)$	−208.0490	−30.7	−27.1	28.9	−8.6	79.2	47.4	55.1	115.4	0.021
$v(G_2)$	438.64	0.09	−0.58	0.54	−0.04	−3.99	−0.26	−0.98	−4.90	1.117

[a] Absolute values for GGG in au (except *Sh*) and those relative to GGG for the rest of the compounds in au and multiplied by 10^3 (*N*), multiplied by 10^3 (*Sh*), in kJ mol^{-1} (*E*), and in au (*v*). δ_m represents the maximum relative variation of each property in the series of compounds expressed in %.

Important variations are also observed for $E(N_{10})$. Thus, N_{10} is stabilized by 75 and 103 kJ mol^{-1} in GGS and SGS, respectively, with regard to GGG.

The dominant character of the hydrogen-bond formation with the side chain and the inductive effects of the serine residue on the charge distribution are observed also in Table 6. This table only shows $\mu(\Omega)$ variations higher than 5×10^{-3} au in the atoms of residue G_2. The most significant variations with regard to GGG are displayed by $\mu(O_4)$ in the compounds with O−H···O_4 HB, and as expected according to Koch and Popelier criteria,[28] they are reductions. Also, the only significant variations of $v(\Omega)$ correspond to size reductions of O_4 in the GGS and SGS tripeptides (−3.4 and −3.5 au, respectively) that should accompany the formation of HBs according to Koch and Popelier criteria.[28]

To use the $Sh(\Omega)$ values, it was also necessary to analyze the effect of $L(\Omega)$ on this quantity. To estimate this effect, the calculation of the integrated properties was repeated several times using different integration parameters, for each atom. Once more, because of the difficulty in reducing $|L(\Omega)|$ for C_2, this is the atom where the study was centered. Thus, Table 7 shows the maxima variations of $Sh(C_2)$ for the same molecule, when $L(\Omega)$ values were included between $L_1(\Omega)$ and $L_2(\Omega)$ (shown in the same table). According to these data, it can be concluded that the effect of $L(\Omega)$ on $Sh(\Omega)$ in the series of atoms and compounds considered is not higher than 3.5×10^{-3}. It can also be noticed that the variation of $Sh(C_2)$ is never higher than $L_2(\Omega)$. For that reason, a first approximation to estimate the error in the $Sh(\Omega)$ computed for the charge distribution expressed in au is its corresponding $L(\Omega)$ value.

Table 8 shows the variation intervals of $Sh(\Omega)$ for each one of the 10 central atoms along the series of tripeptides. The variations observed are higher (if we do not consider the carbon atoms) by 1 order of magnitude than the maximum value of $L(\Omega)$, indicating that the variations displayed by $Sh(\Omega)$ are not due to integration errors. Therefore, the variation of $Sh(\Omega)$ can be employed to analyze the modification experienced by the charge distribution. This analysis is carried out using the $Sh(\Omega)$ versus $N(\Omega)$ representations previously introduced for studying the transferability of *n*-alkanes.[41] Thus, Figure 6 shows these representations for the atoms in residue G_2 in the nine tripeptides considered, and Figure 7 shows this information for the neighboring atoms.

From the analysis of Figure 6, we can conclude that the α carbon, C_2, only suffers important modifications in the two compounds containing a HB with the S_3 side chain [points −G−(HB)]. The formation of such bonds produces a reduction of both $Sh(C_2)$ and $N(C_2)$. This means that, in SGG and SGS, the electron population of C_2 becomes lower and more concentrated in certain regions of the space. The charge distribution for C_2 does not seem to be altered significantly by the fact that the residue attached to the G_2 N extreme is serine, alanine, or glycine. However, for the rest of the atoms, the alterations seem to come from the HB formation, and also from the charge transference from the other residues. A similar trend, with no significant reduction of $Sh(\Omega)$, though, is observed for N_1, where the properties of the atom are only significantly different for SGS and GGS. Properties of C_3 and O_4 are mainly affected by union to a serine a group and O−H···O_4 HB formation. It has to be indicated that the biggest variations of $Sh(\Omega)$ in G_2 are for atoms H_6 and H_7. For these atoms, a good correlation

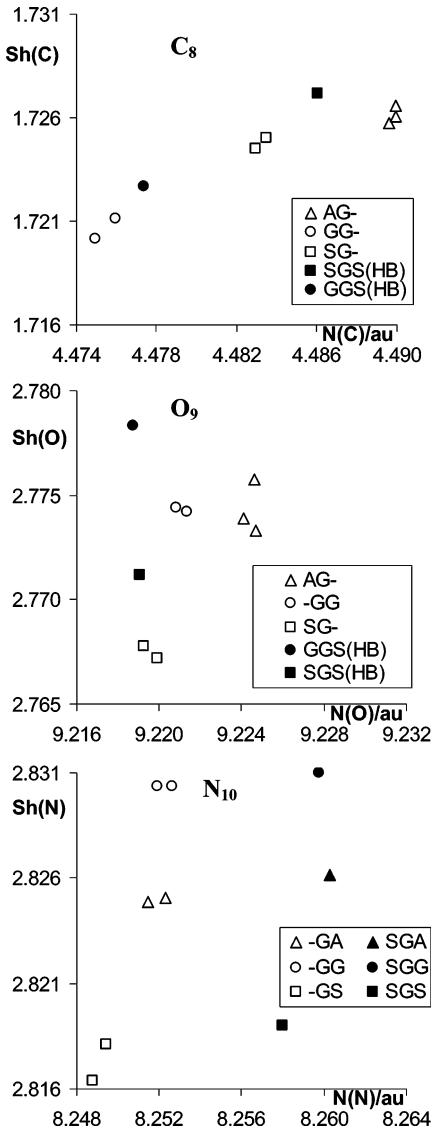

Figure 7. $Sh(\Omega)$ vs $N(\Omega)$ plots for neighboring atoms of G_2.

between *Sh*(H) and *N*(H) is found, indicating that the origin of the *Sh* variation is a charge transference and not a polarization. The opposite situation occurs in H_5 when AG− and GG− compounds are compared. In this case, the substitution by alanine in residue 1 produces a diminution of *Sh*(Ω), keeping practically the same charge and illustrating an atomic polarization without charge transference.

In Figure 7, we observe that, in the vicinity of G_2 on the N-terminal extreme, for C_8 and O_9, a specific character of each residue in position 1 can be observed, with variations inferior to 10^{-3} au, when the side chain does not participate in HBs. This fact can be taken as an indication of the good transferability of residue 1 properties, when the only HBs in the molecule are those due to the α-helix structure. However, the atom N_{10}, which belongs to residue 3, suffers a significant electron population increase when there is a serine in position 1.

Another important aspect is the study of the global properties of the G_2 residue. Table 9 shows that the residue is not neutral and even in the tripeptide GGG presents a negative charge (15×10^{-3} au). This charge is modified depending on the other residues. In general, S residues withdraw electronic charge from G_2, while A residues act as donors. The energy of the central residue also changes notably, varying in this series of compounds by over 146 kJ mol^{-1}, which is almost 50 times greater than the highest discrepancy proportionated by the fitting equation, eq 1. Thus, it can be concluded that the utility of the group contribution models is not based on the charge distribution constancy, but as in other series of compounds,[46−49] group additivity is reached through energy compensation.[49] Moreover, the energies of certain atoms experience significant variations (more than 100 kJ mol^{-1} for N_{10}) within each residue along the molecules here considered. The property that experiences the highest variation is the volume of the group (over 1%). However, the precision obtained in the volume integration is small and does not present a neat dependency with *L*(Ω) to allow the estimation of the error. For that reason, *Sh*(Ω), the following quantity in relative variation and obtained with better precision, becomes an interesting property to analyze the variation of the electronic charge distributions.

CONCLUSIONS

In this study, we have considered the α-helix structures of the N-formyl amides of nine tripeptides containing a central glycine residue bonded to alanine, glycine, or serine. The study of their B3LYP/6-31G** geometries and their B3LYP/6-31G**//B3LYP/6-31++G** molecular energies and electron densities has led us to the following conclusions:

1. The total molecular energies of these compounds can be fitted within 1 kcal mol^{-1} by a contribution model with unique energies for alanine, glycine, serine, and formamide.

2. Little variations can be observed in the geometry of G_2 along the different tripeptides here studied, with regard to that displayed in GGG. The exception is the main dihedral angles of tripeptides GGS and SGS.

3. Several bond paths are observed between atoms that are not bonded according to the Lewis structure. They correspond to the HBs involved in the α-helix structure and also to O−H···O HBs between the side chain of S_3 and the carbonyl group of G_2 in GGS and SGS, C−H···O weak HBs established between a carbonyl group and the side chain of an arginine residue, and diverse N···N and N−H···N repulsive interactions.

4. The O−H···O HBs observed in GGS and SGS can be considered the origin of the geometry distortion of central glycine observed in these compounds.

5. Several integrated atomic properties, *N*(Ω), *E*(Ω), $\mu(\Omega)$, $\nu(\Omega)$, and *Sh*(Ω), of G_2 computed in GGS and SGS display significantly different values from those presented in the remaining compounds here studied.

6. The atomic populations of O_4 in GGS and SGS are larger than those in the rest of the compounds here studied. This does not agree with the traditional picture of O−H···O═C hydrogen bonding.

7. The excellent group additivity observed for molecular energies is not based on the charge distribution constancy, but as previously shown for other series of compounds,[46−49] it is reached through energy compensation.[49]

Finally, answering the question from the title of this work, we can say that along the series of tripeptides the central glycine residue has different properties in the presence of a serine residue whose side chain forms a HB with the carbonyl group of G_2. However, even if these changes could be enough to explain specific properties of certain Aa's in a protein, they are not big enough to make necessary the employment of specific parameters in the molecular mechanics force fields usually used to study the geometry of these systems. In contrast, the effects due to neighboring Aa's are not important in the remaining cases here studied.

REFERENCES AND NOTES

(1) Tapia, O. Dinámica Molecular de Proteinas y Reacciones Enzimáticas. In *Química Teórica*; Fraga, S., Ed.; CSIC: Madrid, Spain, 1989; Vol. 2, pp 155−175.
(2) Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure*; Springer-Verlag: New York, 1979.
(3) Prigogine, I.; Rice, S. A. In *Computational Methods for Protein Folding. Advances in Chemical Physics*; Friesner, R. A., Ed.; J. Wiley & Sons: New York, 2001.
(4) Stocker, U.; van Gunsteren, W. F. Molecular dynamics simulation of hen egg white lysozyme: A test of the GROMOS96 force field against nuclear magnetic resonance data. *Proteins* **2000**, *40*, 145−153.
(5) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Freguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(6) Lazaridis, T.; Karplus, M. Effective energy functions for protein structure prediction. *Curr. Opin. Struct. Biol.* **2000**, *10*, 139−145.
(7) Stone, A. J. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **1981**, *83*, 233−239.
(8) Stone, A. J.; Alderton, M. Distributed multipole analysis − methods and applications. *Mol. Phys.* **1985**, *56*, 1047−1064.
(9) Faerman, C. H.; Price, S. L. A transferable distributed multipole model for the electrostatic interactions of peptides and amides. *J. Am. Chem. Soc.* **1990**, *112*, 4915−4926.
(10) Price, S. L.; Stone, A. J. Electrostatic models for polypeptides: Can we assume transferability? *J. Chem. Soc., Faraday Trans.* **1992**, *88*, 1755−1763.
(11) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory*; Clarendon Press: Oxford, U. K., 1990.
(12) Bader, R. F. W. Principle of Stationary Action and the Definition of a Proper Open System. *Phys. Chem. B* **1994**, *49*, 13348−13356.
(13) Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893−928.
(14) Bader, R. F. W.; Popelier, P. L. A.; Chang, C. Similarity and complementarity in chemistry. *THEOCHEM* **1992**, *255*, 145−171.
(15) Bader, R. F. W.; Popelier, P. L. A.; Keith, T. A. Theoretical definition of a functional group and the molecular orbital paradigm. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 620−631.

(16) Popelier, P. L. A.; Bader, R. F. W. Effect of Twisting a Polypeptide on Its Geometry and Electron Distribution. *J. Phys. Chem.* **1994**, *98*, 4473−4481.
(17) Chang, C.; Bader, R. F. W. Theoretical construction of a polypeptide. *J. Phys. Chem.* **1992**, *96*, 1654−1662.
(18) Matta, C. F.; Bader, R. F. W. An atoms-in-molecules study of the genetically-encoded amino acids: I. Effects of conformation and of tautomerization on geometric, atomic, and bond properties. *Proteins* **2000**, *40*, 310−329.
(19) Matta, C. F.; Bader, R. F. W. Atoms-in-molecules study of the genetically encoded amino acids. II. Computational study of molecular geometries. *Proteins* **2000**, *48*, 519−538.
(20) Matta, C. F.; Bader, R. F. W. Atoms-in-molecules study of the genetically encoded amino acids. III. Bond and atomic properties and their correlations with experiment including mutation-induced changes in protein stability and genetic coding. *Proteins* **2003**, *52*, 360−399.
(21) Schäfer, L.; Cao, M.; Ramek, M.; Teppen, B. J.; Newton, S. Q.; Siam, K. Conformational geometry functions: Additivity and cooperative effects. *J. Mol. Struct.* **1997**, *413*−*414*, 1997.
(22) Ramek, M.; Yu, C.; Schäfer, L. Ab initio conformational analysis of the model tripeptide *N*-formyl-L-alanyl-L-alanine amide. *Can. J. Chem.* **1998**, *76*, 566−575.
(23) Hô, M.; Simth, V., Jr.; Weaver, D. F.; Gatti, C.; Sagar, R. P.; Esquivel, R. O. Shannon information entropies of molecules and functional groups in the self-consistent reaction field. *J. Chem. Phys.* **2000**, *112*, 7572−7580.
(24) Lorenzo, L.; Mosquera, R. A. Approximate transferability in alkanes revisited. *Chem. Phys. Lett.* **2002**, *356*, 305−312.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C02; Gaussian Inc: Wallingford, CT, 2004.
(26) Bader, R. F. W. *AIMPAC: A Suite of Programs for the AIM Theory*; Mc Master University: Hamilton, Ontario, Canada. Contact bader@mcmail.cis.mcmaster.ca.
(27) Steiner, T. The hydrogen bond in the solid state. *Angew. Chem., Int. Ed.* **2001**, *41*, 48−76.
(28) Koch, U.; Popelier, P. L. A. Characterization of C−H−O hydrogen bonds on the basis of the charge density. *J. Phys. Chem.* **1995**, *99*, 9747−9754.
(29) Carroll, M. T.; Bader, R. F. W. An analysis of the hydrogen-bond in base-Hf complexes using the theory of atoms in molecules. *Mol. Phys.* **1988**, *65*, 695−722.
(30) Rozas, I.; Alkorta, I.; Elguero, J. Bifurcated hydrogen bonds: Three-centered interactions. *J. Phys. Chem. A* **1998**, *102*, 9925−9932.
(31) Cioslowski, J.; Edgington, L.; Stefanov, B. B. Steric overcrowding in perhalogenated cyclohexanes, dodecahedranes, and [60]fulleranes. *J. Am. Chem. Soc.* **1995**, *117*, 10381−10384.
(32) Bader, R. F. W. A bond path: A universal indicator of bonded interactions. *J. Phys. Chem. A* **1998**, *102*, 7314−7323.
(33) Tsirelson, V.; Abramov, Y.; Zavodnyk, V.; Stash, A.; Belokoneva, E.; Stahn, J.; Pietsch, U. Critical points in a crystal and procrystal. *Struct. Chem.* **1998**, *9*, 249−254.
(34) Pendás, A. M.; Costales, A.; Luaña, V. Ions in crystals: The topology of the electron density in ionic materials. 1. Fundamentals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1997**, *55*, 4275−4284.
(35) Aiken, F. M.; Popelier, P. L. A. Atomic properties of selected biomolecules. Part 1. The interpretation of atomic integration errors. *Can. J. Chem.* **2000**, *78*, 415−426.
(36) Graña, A. M.; Mosquera, R. A. The transferability of the carbonyl group in aldehydes and ketones. *J. Chem. Phys.* **1999**, *110*, 6606−6616.
(37) Vila, A.; Carballo, E.; Mosquera, R. A. AIM study on the transferability of the oxygen atom in linear ethers. *Can. J. Chem.* **2000**, *78*, 1535−1543.
(38) López, J. L.; Mandado, M.; Graña, A. M.; Mosquera, R. A. Approximate transferability in alkanenitriles. *Int. J. Quantum Chem.* **2002**, *86*, 190−198.
(39) Mandado, M.; Graña, A. M.; Mosquera, R. A. Appoximate transferability in alkanols. *THEOCHEM* **2002**, *584*, 221−234.
(40) Quiñónez, P. B.; Vila, A.; Graña, A. M.; Mosquera, R. A. AIM study on the influence of fluorine atoms on the alkyl chain. *Chem. Phys.* **2003**, *287*, 227−236.
(41) Vila, A.; Mosquera, R. A. An electron density analysis of the proximity effect in linear alkyl diethers. *Chem. Phys. Lett.* **2001**, *345*, 445−452.
(42) Vila, A.; Mosquera, R. A.; Hermida-Ramon, J. M. AIM characterization of hydrogen bonds in dimers of methoxymethane. *THEOCHEM* **2001**, *541*, 149−158.
(43) Vila, A.; Graña, A. M.; Mosquera, R. A. Electron density characterisation of intermolecular interactions in the formaldehyde dimer and trimer. *Chem. Phys.* **2002**, *281*, 11−22.
(44) Mandado, M.; Vila, A.; Graña, A. M.; Mosquera, R. A.; Cioslowski, J. Transferability of energies of atoms in organic molecules. *Chem. Phys. Lett.* **2003**, *371*, 739−743.
(45) Cortés-Guzmán, F.; Bader, R. F. W. Transferability of group energies and satisfaction of the virial theorem. *Chem. Phys. Lett.* **2003**, *379*, 183−192.
(46) Graña, A. M.; Mosquera, R. A. Transferability in aldehydes and ketones. II. Alkyl chains. *J. Chem. Phys.* **2000**, *113*, 1492−1500.
(47) Vila, A.; Mosquera, R. A. Transferability in alkyl monoethers. II. Methyl and methylene fragments. *J. Chem. Phys.* **2001**, *115*, 1264−1273.
(48) Mandado, M.; Graña, A. M.; Mosquera, R. A. On the effects of electron correlation and conformational changes on the distortion of the charge distribution in alkyl chains. *Chem. Phys. Lett.* **2002**, *355*, 529−537.
(49) Bader, R. F. W.; Bayles, D. Properties of atoms in molecules: Group additivity. *J. Phys. Chem. A* **2000**, *104*, 5579−5589.

CI600184T