Theor Chim Acta (1993) 85: 189–197

Theoretica Chimica Acta

Peptide models II. Intramolecular interactions and stable conformations of glycine, alanine, and valine peptide analogues*

Wladia Viviani[1], Jean-Louis Rivail[1], and Imre G. Csizmadia[2]

[1] Laboratoire de Chimie Théorique, U.A. 510 CNRS, Université de Nancy-I, BP 239, F-54506 Vandœuvre-les-Nancy, Cedex, France

[2] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 1A1, Canada

Received September 7, 1991/Accepted April 14, 1992

Summary. The stable conformations of N and C protected amino acids of the type: $HCONH–CHR–CONH_2$ of glycine, L-alanine and L-valine have been obtained by fully optimized *ab-initio* computations with a 3-21G basis set. An original procedure has been devised to extract the side-chain/backbone interaction energy and the backbone and side-chain distortion energies. This enables us to analyze the stabilization/destabilization effects introduced by the side-chain in terms of electrostatic, induction and steric hindrance contributions.

Key words: Peptide models – Amino acids – Conformations, stable – Proteins

1 Introduction

The study of the conformations of amino acid units entering the proteins is a key to the understanding of the secondary structure of proteins. Molecular Mechanics (MM) [1] has already made possible a first analysis at their degree of approximation although reference computational data are better obtained by *ab-initio* quantum chemical techniques [2, 3]. Nowadays the continuous improvements of high capacity computers and the refinement of the codes make possible the *ab-initio* computations of large molecular structures at a reasonable accuracy and with full optimization of geometrical parameters [4]. Recently such computations have been performed on three molecules, models of peptide units: N-formyl glycinamide and its two simplest homologues derived from alanine [5, 2, 6], and valine [7]. They revealed that some minimum energy conformers were missing from the nine legitimate structures predicted by multidimensional conformational analysis [2, 8].

The conformations of peptidic units are usually described by the two Ramachandran angles [9]: ϕ which characterizes the rotation around the C^{α}–N bond and ψ related to the rotation around the C^{α}–carbonyl bond. They are defined in Fig. 1.

In this paper, we denote the nine legitimate conformations by using the first five letters of the Greek alphabet, from α to ϵ, and two subscripts, L and D. The

* Dedicated to Dr. A. Pullman

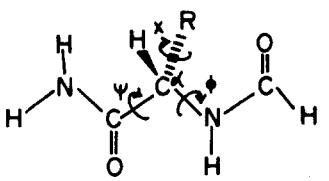
basis of this notation and its correspondence with Scheraga classification [10] are elaborated on [2]. The definition is summarized in Fig. 2. Five of these nine possible conformations have been recognized in proteins. These were the left and right handed α helices, the β pleated sheets and normal and inversed γ turns.

In the simplest case of N-formyl glycinamide, the ideal situation described before is already strongly perturbed. The angles corresponding to the stable conformation depart from the values indicated in Fig. 2, and the pair of ϵ conformations do not correspond to a minimum of the Potential Energy Surface. Nevertheless, the D and L conformations of a given backbone conformer (α, β, γ, δ, ϵ) still have the same energy, due to the achiral nature of the glycine. In the case of a substitution on the α carbon of the chain, which introduces a chiral center, this last regularity disappears and the distinction between D and L conformations is supported by energy differences.

The presence of a side-chain also modifies the list of stable conformers. For example, the replacement of the hydrogen atom by a methyl group when passing from glycine to alanine not only makes one of the α minima disappear but also makes one of the missing ϵ conformers reappear [2]. Things are becoming more complicated when the methyl group is replaced by an isopropyl group with three non-equivalent positions corresponding to three 120° rotations around the C^{α}–C^{β} bond characterized by a third angular variable χ (cf. Fig. 1). Besides, the size of the isopropyl substituent is expected to exhibit some phenomena related to steric hindrance.

In order to get a greater insight into the main contributions to the intramolecular interactions which stabilize or destabilize the conformations of the amino acid units in peptides and proteins, it then seemed useful to carefully analyze the results of *ab-initio* computations performed on the three simplest amino acid derivatives, in which the side-chain is a hydrophobic hydrocarbon radical, in order to look for some rationale which could be helpful in predicting the effects appearing in the other related amino acid units.

1



2

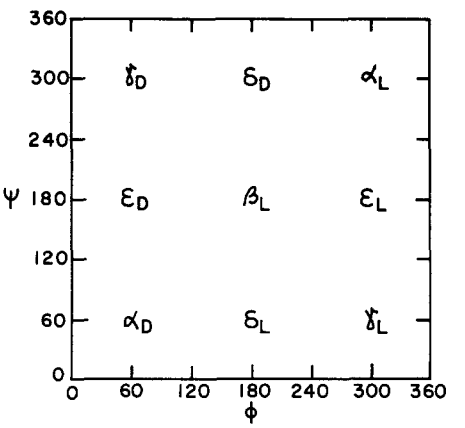


Fig. 1. Schematic molecular structure of N-formyl amino acid/amides with the specification of ϕ, ψ and χ torsional angles

Fig. 2. A schematic Ramachandran type map, $E = E(\phi, \psi)$, indicating the idealized locations of the expected 9 legitimate minima

Energy values can be used safely to compare the conformations of a given molecular species in which the nature and the number of the atoms are constant. In contrast, it is very difficult to compare two conformations of two different molecules since the energy variations due to the modifications of the number of atoms are several orders of magnitude larger than the small energy differences which occur between two conformations of a same molecule. It then became necessary to extract pertinent differential quantities from the global data of molecular energies. In particular, backbone/side-chain interactions, which have been recognized as playing an important part in the structure of peptides [13, 14], deserve our special attention. In this work we developed methodology which enables us to extract reliable quantities such as side-chain/backbone interaction or backbone distortion energies. This methodology is presented in the next section of this paper.

2 Method

Ab-initio SCF calculations have been carried out using GAUSSIAN-88 [11] and MONSTER-GAUSS [12] programs at the 3-21G basis set level. The geometry optimizations were stopped when the gradients were less than 5×10^{-4} mdyne/Å. Some molecular properties such as dipole moment, electric potential, electric field and Mulliken charges were calculated for selected conformations, whose stable structures were optimized along all the internal coordinates.

In order to extract useful energetic data from the substituent homologues of N-formyl glycinamide we proceed as follows.

1. replace the side-chain R along the C^{α}–C^{β} bond by a hydrogen atom and only optimize the bond length of this second C^{α}–H bond. One then obtains a distorted glycine conformer whose energy is denoted E_{DGL};

2. considering the C^{α}–R fragment, replace the C^{α}–C and C^{α}–N bonds by two C–H bonds and optimize the bond lengths, but maintain the bond angles at the values they had in the amino acid. The energy of the R–CH_3 molecule is denoted by E_{RMe};

3. starting from the distorted glycine molecule defined in 1, one builds up a methane molecule with the C^{α} by the same procedure as in 2. The energy is E_{HMe}.

Hence, the difference $E_{RMe}-E_{HMe}$ represents the energy increment due to the replacement of the H atom on the α carbon by the substituent R, in the absence of any interaction of R with the rest of the molecule, and the sum $E_{DGL} + (E_{RMe} - E_{HMe})$ would be the total energy of the amino acid derivative in the conformation of interest if the energy of interaction between the side-chain and the backbone were equal to zero. Therefore, the difference of this quantity with the actual energy E is expected to be a good evaluation of the energy of the interaction E_{BR} of the side-chain with the backbone in the geometry of the conformer.

$$E_{BR} = E - (E_{DGL} + E_{RMe} - E_{HMe}) \quad (1)$$

Similarly the difference E_D between E_{DGL} and the energy of the corresponding conformer of the fully relaxed glycine derivative evaluates the energy of distortion of the main chain when passing from the glycine derivative to the

corresponding conformer of the substituted amino acid derivative. Finally the distortion of the side-chain can be evaluated by means of the same procedure by comparing the energies of the relaxed hydrocarbon molecules to E_{HMe} and E_{RMe}.

3 Results and discussion

The main results are summarized in Tables 1–3 in which μ is the computed dipole moment of the conformer and μ_{DGL} is the computed dipole moment of the distorted glycine corresponding to the backbone geometry. They show a rather contrasted situation from one conformer to another, and differences between alanine and valine.

Table 1. Total energy and dipole moment for all backbone conformations of glycine derivative[a]

Conformation	E	μ
α	−373.641601	6.56
β	−373.647718	3.36
γ	−373.648707	3.54
δ	−373.643495	4.91
(ϵ)[b]	−373.638902	5.47

[a] Energy values in Hartrees, dipole moments in Debyes

[b] In order to evaluate the above quantities for the missing ϵ minimum, its geometry was assumed to be equivalent to the geometry of the optimized corresponding conformation of alanine

Table 2. Backbone distortion energy, backbone/side-chain interaction energy, dipole moment of the amino acid derivative and of the corresponding distorted glycine, for all backbone conformations of alanine[a]

Conformation	E_D	E_{BR}	μ	μ_{DGL}	θ_{HCO}	θ_{NCO}
α_D	+25.151	−10.849	6.56	7.36	113.98	110.33
(α_L)[b]	—	+33.230	6.98	6.76	90.34	52.83
β_L	+0.380	−5.585	3.23	3.50	62.52	115.37
γ_D	+1.349	−4.511	3.83	3.87	65.39	126.51
γ_L	+0.387	−6.180	3.28	3.44	127.05	47.92
δ_D	+3.375	−5.170	4.92	5.07	71.91	128.01
δ_L	+17.262	−7.656	3.97	4.31	114.53	95.55
ϵ_D	—	−3.734	5.12	5.48	45.28	136.00
(ϵ_L)[b]	—	−3.712	6.55	6.38	48.52	43.24

[a] Energy values in kcal/mol, dipole moments in Debyes and angles in degrees

[b] In order to evaluate the above quantities for missing expected minima, their geometry was assumed to be equivalent to the geometry of the corresponding optimized conformation obtained by interchanging the subscript D or L, i.e. by changing the sign of the value of the torsional angles: $\phi = -\phi$; $\psi = -\psi$

Table 3. Backbone distortion energy, backbone/side-chain interaction energy, dipole moment of the amino acid derivative and of the corresponding distorted glycine, for all backbone conformations of valine at the $\chi = 180°$ side-chain conformation[a]

Conformation	E_D	E_{BR}	μ	μ_{DGL}	θ_{HCO}	θ_{NCO}
α_D	+2.780	−6.401	6.51	7.05	123.26	115.36
$(\alpha_L)^b$	—	4.386	7.21	7.29	72.10	50.10
β_L	+2.714	−5.730	2.37	2.54	86.41	88.46
γ_D	+1.595	−4.835	3.61	3.75	61.10	126.32
γ_L	+1.023	−6.498	2.90	3.20	123.36	51.31
δ_D	+5.439	−5.528	6.08	6.30	65.93	121.65
$(\delta_L)^b$	—	199.332	7.51	6.67	125.39	171.88
ϵ_D	—	−3.960	5.93	6.56	70.11	145.23
$(\epsilon_L)^b$	—	32.373	7.76	7.06	24.59	40.24

[a] Energy values in kcal/mol, dipole moments in Debyes and angles in degrees

[b] In order to evaluate the above quantities for missing expected minma, their geometry was assumed to be equivalent to the geometry of the corresponding optimized conformation obtained by interchanging the subscript D or L, i.e. by changing the sign of the value of the torsional angles: $\phi = -\phi$; $\psi = -\psi$

3.1 Electrostatic interactions in the α and ϵ conformers

Brant et al. [15] have shown that dipole-dipole interactions between two adjacent peptidic units play an important role in the difference of stability between α helices and β strands. This can be easily seen in Table 1 where the global dipole moment can be analyzed as the vectorial sum of the moments of the two amidic groups. A large value of the global dipole moment indicates a conformation in which the angle between the two moments is small i.e. in which the electrostatic interaction is fairly repulsive. It then appears that this kind of interaction is the most important in the case of the α conformations.

The absence of a minimum corresponding to the α_L conformation of N-formyl-L-alanine amide has already been interpreted [2] as a consequence of the unfavourable interaction between the small dipole moment of the C^α–CH_3 group with these two large moments associated with the amidic groups. Conversely, this interaction is favourable in the case of the enantiomeric α_D conformation (see Fig. 3).

This interpretation can be further supported by various results. One first notices that the contribution of the backbone to the total dipole moment, which is approximated by the value of μ_{DGL}, is larger than the dipole moment of the α_D conformer of alanine, but smaller than the dipole moment computed for the hypothetic α_L conformation. If one takes the geometrical features into account, the difference, 0.2 *D*, is consistent with the estimated value for the contribution of a methyl group [16]. The same reasoning holds for the α conformers of valine.

If one assumes that the peptidic bond dipole moment is directed along the carbonyl bond, the relative orientation of the permanent dipole can also be clarified by the angles θ_{HCO} and θ_{NCO} between the RC^α vector and the CO vectors of the carbonyl bonds of the formly group at the amine end and of the acid end, respectively. The dipole-dipole electrostatic term is stabilizing when these angles are greater than 90° and destabilizing in the other case as suggested

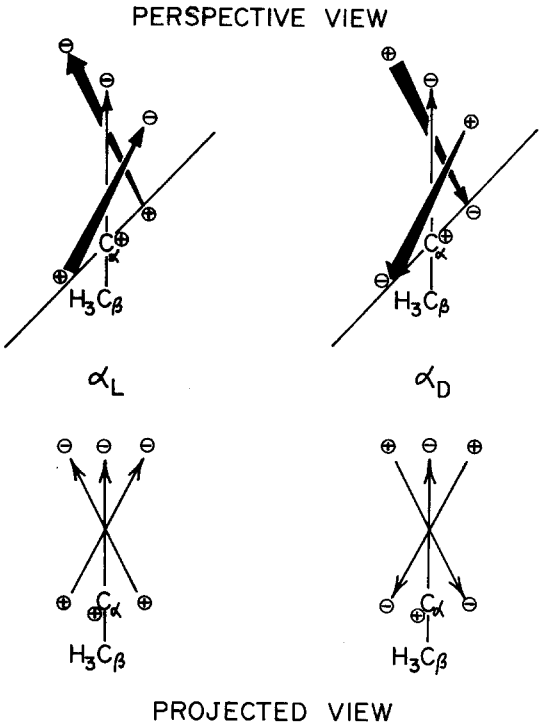


Fig. 3. Schematic vectorial illustration of the dipole-dipole interactions in the α_L and α_D conformers of the L-alanine derivative HCONHCH–(CH_3)CONH$_2$. The vectors point from the positive to the negative charge, accordint to chemical convention

by Fig. 1. The data given in Tables 2 and 3 show a favourable interaction for the α_D conformation and slightly unfavourable for the α_L.

The dipole-dipole interaction between the peptidic bonds is also expected to be an important contribution to the conformational energy. The variations of this contribution can be followed through the variations of the backbone dipole moment μ_{DGL}, large values of this quantity indicating a destabilizing relative orientation of the individual dipoles.

One is first incited to assume that this unfavourable interaction probably contributes to the absence of minimum corresponding to the ϵ geometry in the glycine derivative. In the case of alanine and valine, the ϵ_L conformations are found and the extra stability gained by the presence of the substituent can be explained as above.

3.2 Induction energy

One notices that in the alanine derivative, and valine derivatives, the side-chain/backbone interaction is stabilizing in all the conformations corresponding to a minimum of the PES. Although the electrostatic interactions still play a role, the magnitude of the interaction can hardly be explained by the electrostatic effect alone. Tables 2 and 3 clearly indicate that except for the α_D conformation the

orientation of the carbonyl groups is such that their contributions to the electrostatic interactions with the RC^{α} dipole nearly cancel. This finding strongly suggests that the induced dipole of the polarisable side-chain should play an important part leading to an induction contribution to the conformational energy [17] which is always stabilizing. Besides, this induced moment is expected to subtract to the total dipole moment, owing to the location of the side-chain between the two polar ends. The data of Tables 2 and 3 show that for all the stable conformers the dipole moment is smaller than μ_{DGL}, the expected contribution of the backbone.

An overview of the various effects can be reached by a direct comparison of the conformers of the three amino acid derivatives. This comparison is not possible with the *ab-initio* energies since the number of particles vary from one amino acid to another. Nevertheless, our analysis makes possible a scaling of the alanine and valine derivatives to glycine. For this purpose we added the *ab-initio* distortion energy of the backbone and of the side-chain, and the side-chain/backbone interaction energy for a given conformer, leading to scaled conformational energies represented on Table 4. These values clearly show the stabilization of all the stable conformations of alanine, which we interpret as mainly due to induction. The valine derivatives exhibit a more contrasted behaviour which can be interpreted by a competition between the stabilizing effects analyzed in the valine derivative and steric hindrance effects due to the size of the side-chain.

3.3 Steric hindrance in valine derivatives

The previous reasoning would lead us to expect a greater stabilization in the case of the valine compound compared with the corresponding configurations of the alanine derivative due to the larger polarizability of the isopropyl group. This is not observed, obviously because the size of the isopropyl group introduces a non-negligible repulsive contribution known as steric hindrance, even in the most stable conformations. Furthermore, no minimum can be found for δ_L conformer for $\chi = 180°$. Looking at the interatomic distances one notices that, as a consequence of the backbone folding direction in δ_L structures (Fig. 3a), most of the atoms of the side-chain in the $\chi = 180°$ conformation are found to be very close to the backbone atoms, so that the distance backbone/side-chain is shorter than 3 Å at many points (Fig. 4a,b).

Table 4. Scaled energies[a] for all stable conformations of glycine, L-alanine and L-valine derivatives. The reference level is the γ conformation of the glycine compound

	GLY	*ALA*	*VAL* ($\chi = 60°$)	*VAL* ($\chi = 180°$)	*VAL* ($\chi = 300°$)
α_D	4.459	0.304	3.279	1.339	2.440
α_L	4.459	—	—	—	—
β_L	0.621	−4.372	−3.429	−2.065	−3.608
γ_D	0.000	−3.101	−0.176	−3.210	0.199
γ_L	0.000	−5.622	−4.746	−5.413	−5.053
δ_D	3.271	1.680	2.526	3.603	5.475
δ_L	3.271	−1.820	−1.178	—	−2.117
ϵ_D	—	2.544	6.942	5.312	5.646

[a] in kcal/mol. For definition see text

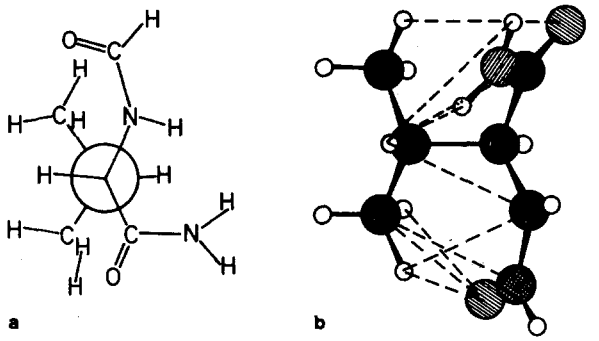


Fig. 4. Molecular structure for the expected δ_L conformation at $\chi_0 = 180°$ for N-for L-valine amide (the torsional angles ϕ and ψ are assumed to have the values of 60° and 180°, respectively). (a) Newman projection, (b) 3-D representation. Atom distances smaller than 3 Å are evidenced by *dotted lines*

3.4 Intramolecular hydrogen bonds

It is well known that intramolecular hydrogen bonds contribute to stabilization of the γ conformers independently of the nature of the side-chain [5, 2]. The variation of this contribution to the stability of the various γ conformers may be anticipated from the variations of the length of the hydrogen bond which occurs between a hydrogen atom bonded to the terminal nitrogen and the oxygen of the carboxylic group at the other end. The formation of this hydrogen bond between both ends of the backbone results in the formation of a cycle in which the optimum hydrogen bond length is modified by the angular tension.

If we assume glycine as having the natural conformation, that means in absence of side-chain influence, the internal torsion in the cycle leads to a relatively weak hydrogen bond with R(O–H) = 2.01 Å. When a side-chain is introduced, the geometry of the cycle is modified and the result is a variation of the length of this hydrogen bond, which is shortened in the case of the γ_D conformation and slightly lengthened in the γ_L's. Since the stabilization is always in favour of the γ_L conformations one is led to conclude that the variations of the hydrogen bond energy plays a negligible role in the stabilization and that the electrostatic and induction contributions can be considered in all cases as the main causes of stabilization arising from the C^α substitution by a hydrocarbon chain.

4 Conclusion

Ab-initio SCF calculations allow us to compute the energy both of a whole molecule and of selected fragment molecules in order to evaluate partial contributions to the total energy and consequently analyze properly the different factors intervening. Consequently, it becomes possible to highlight phenomena as dipole-dipole interactions, hydrogen bonds, induction effects.

On the basis of such an analysis one can then predict in a qualitative way the effects of the introduction of a given side-chain in an amino acid and even

estimate the influence of an eventual proximity of a portion of that side-chain to another part of the protein.

The method can be applied to any kind of natural or modified amino acid and is expected to contribute to a better understanding of some small effects which sometimes strongly influence the structure of a polypeptide or a protein.

Acknowledgements. The *ab-initio* computations were carried out at CIRCE (Orsay, FRANCE). The author thank IBM-France and CNRS for the generous allocation of computer time within the framework of the GS (Groupement Scientifique) "Modélisation Moléculaire". The continued financial support of the NSERC of CANADA is gratefully acknowledged.

References

1. Vásquez M, Nemethy G, Scheraga HA (1983) Macromolecules 16:1043
2. Perczel A, Angyan JG, Kajtar M, Viviani W, Rivail JL, Marcoccia JF, Csizmadia IG (1991) J Am Chem Soc 113:6256
3. Pullman B (1990) In: Rivail JL (ed) Modelling of molecular structures and properties. Elsevier, Amsterdam, p. 1
4. Schäfer L, Klimkowski VJ, Momany FA, Chuman H, VanAlsenoy C (1984) Biopolymers 23:2335
5. Head-Gordon T, Head-Gordon M, Frisch MJ, Brooks III C, Pople J (1989) Int J Quantum Chem, Quantum Biol Symp 16:311
6. Head-Gordon T, Head-Gordon M, Frisch MJ, Brooks III C, Pople J (1991) J Am Chem Soc 113:5989
7. Viviani W, Rivail JL, Perczel A, Csizmadia IG, submitted to J Am Chem Soc
8. Mezey PG (1987) Potential energy hypersurfaces. Elsevier, Amsterdam, p 227
9. Ramakrishnan C, Ramachandran GN (1965) Biophys J 5:909
10. Zimmerman SS, Pottle MS, Nemethy G, Scheraga HA (1977) Macromolecules 10:1
11. Frisch MJ, Head-Gordon M, Schlegel HB, Raghavachari K, Binkley JS, Gonzales C, Defrees DJ, Fox DJ, Whiteside RA, Seeger R, Melius CF, Baker J, Martin RL, Kahn LR, Steward JJP, Fluder EM, Topiol S, Pople JA, Gaussian Inc., Pittsburgh, PA
12. Peterson MR, Poirier RA, Chem Dept, University of Toronto, Canada
13. Gram TM, Matthews BW (1984) J Mol Biol 175:75
14. Scheraga HA (1985) Proc Natl Acad Sci USA 82:5585
15. Brant DA, Miller WG, Florey PJ (1967) J Mol Biol 23:47
16. Minkin VI, Osipov OA, Zhdanov YA (1970) Dipole moments in chemistry. Plenum Press, NY, p 79
17. Buckingham AD (1978) In: Pullman B (ed) Molecular interactions from diatomics to bipolymers. Wiley, NY, p 1