Theoretical studies of molecular conformation.
II: Application of the SIBFA procedure to molecules containing carbonyl and carboxylate oxygens and amide nitrogens

Nohad Gresh, Alberte Pullman, and Pierre Claverie

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, associé au C.N.R.S., 13 rue Pierre et Marie Curie, f-75005 Paris, France

The SIBFA procedure (Sum of Interactions Between Fragments computed *Ab initio*, Ref [1]) is extended to the study of the conformational behavior of representative molecules containing amide nitrogens and carbonyl and carboxylate oxygens. The molecules studied are C- and N-ethylammonium formamide, C- and N-ethanol-formamide, ethylammonium formate and ethanolformate. The cases investigated include interactions of the types which occur between functional groups in proteins or ionophores. The accuracy of the procedure, assessed by comparing the results to those of corresponding *ab initio* SCF computations, is very satisfactory. An application of the procedure to study the conformation of the glycyl and alanyl dipeptides as a function of the backbone torsional angles ϕ and ψ is presented.

Key words: SIBFA method—conformational analysis—peptide groups— carbonyl oxygens—carboxylate oxygens

1. Introduction

Conformational studies of organic and biological molecules constitute an important field of applications of theoretical chemistry. In a previous paper [1] we have presented a new methodology for the rapid computation of the variations in the conformational energies of large molecules, compatible with the simultaneous computation of intermolecular interaction energies. This procedure, named SIBFA, (Sum of Interactions Between Fragments computed *Ab initio*)

divides the macro-molecule investigated into fragments separated by single bonds, and computes the variation in the *intra*molecular interaction energy of the macromolecule upon rotation about these bonds as a sum of interaction energies between the fragments. Distinctive features of the procedure are the computation of the electrostatic and polarization components of the energy using a multipolar expansion of the *ab initio* SCF wave functions of the individual fragments, and the computation of the repulsion energy as a sum of bond and lone pair interactions. The treatment of the interactions involving the junctions between two successive fragments was described in [1].

The accuracy of the SIBFA procedure in reproducing the conformational behaviour of a series of test molecules involving ether-like oxygens and pyridine-like nitrogens was assessed by confronting the SIBFA results to those of corresponding *ab initio* SCF computations, providing a very encouraging agreement. In the present study, we extend the procedure to include molecules bearing amide nitrogen and carbonyl or carboxylate oxygens. The molecules investigated are: C and N-ethylammonium formamide, C and N-ethanol formamide, ethylammonium formate and ethanolformate.

These model studies are expected to cover a wide range of interactions susceptible of occurring between functional groups of proteins or ionophores.

2. Methodology

Let us recall that the variation in the intramolecular energy is computed as a sum of five components:

$$E = E_{MTP} + E_{pol} + E_{rep} + E_{dl} + E_{tor}$$

in which E_{MTP} is the electrostatic component, computed as a sum of multipole-multipole interactions between the interacting fragments, E_{pol} is a polarization contribution computed accordingly using the same multipolar expansions, E_{rep} is a sum of bond–bond, bond–lone pair, and lone pair–lone pair interactions, E_{dl} is a dispersion-like term containing the charge-transfer component when present [2], and E_{tor} is a transferable torsional energy contribution, calibrated on model systems.

The fragments utilized to build the investigated test molecules are formamide, formate, monomethylammonium, methane and water, with the appropriate bond lengths and valence angles which they have in the test molecules. Rigid rotation was assumed throughout this study. The *ab initio* SCF conformational computations as well as the *ab initio* SCF computations on the fragments, necessary to obtain the multipoles were performed using the same pseudopotential procedure and basis set as in paper 1.

The present extension requires the appropriate definition of lone-pairs for the carbonyl and carboxylate oxygens, and for the amide-type nitrogen, as well as the selection of the corresponding van der Waals radii.

In accordance with the general methodology described in [1], we attribute to each carbonyl and carboxylate oxygen, X_O, two lone-pairs defined by two fictitious atoms, L_O and $L_{O'}$, located in the plane of the carboxylate or of formamide and at 0.30 Å from X_O, corresponding to the barycenters of two equivalent sp^2 oxygen hybrids:

$$\chi_O = C_s 2_{s_0} + C_p 2p_{\sigma_0}$$

with the standard coefficients $C_s = 0.577$, $C_p = 0.816$, and a Slater exponent of 2.24 [3].

The π lone pair of the nitrogen atom of the formamide moiety is represented by two fictitious atoms, L_N, $L_{N'}$ one on each side of the formamide plane, with the segments $X_N L_N$ and $X_N L_{N'}$ perpendicular to C–N. The distances X_N–L_N and $X_N L_{N'}$ are computed as 0.52 Å using the barycenter of each of the lobes p_z of the orbital with $\zeta_N = 1.92$ [3].

Accordingly, on account of the half occupancy of this π lone pair and of its partial delocalization, in the computation of the repulsion contribution involving it and a given atom P or lone pair of another fragment, the term:

$$M_{PL}^2 = \frac{1}{v_p}\left(1 - \frac{q_p}{N_p^{\mathrm{val}}}\right) \quad \text{is divided by two.}$$

The determination of the necessary effective radii for the carbonyl and carboxylate oxygens, and for the amide nitrogens, as well as of their associated fictitious atoms, was done in a manner consistent with that adopted in [1] for ether-like oxygens and pyridine-like nitrogens, so as to retrieve the values of the corresponding *ab initio* pseudopotential energy differences between representative conformations (*gauche, trans* and *cis*) in: ethylammonium formate for anionic oxygens, C-ethylammonium formamide for carbonyl oxygens, and N-ethanol formamide (with a terminal *cis* hydroxyl) for amide nitrogens.

The first two molecules were selected to ensure large amplitudes of the variations of intramolecular energy due to the terminal ammonium end, hence a large sensitivity to the choice of the W's. For the N-substituted formamide, we have selected for the calibration of W_N and W_{LN} a *cis* terminal hydroxyl group rather than an ammonium end, since the short-range repulsive energy involving the N pi lone-pair and the OH bond have a larger amplitude of variation than the corresponding repulsive energy with either of the three NH bonds (which are staggered with respect to the methylene and farther from the amide nitrogen). Furthermore, this repulsion is not "obliterated" by too large an attraction with the nearby carbonyl oxygen, as is the case with the terminal ammonium.

This led us to the values $W_O = 1.235$ Å and 1.26 Å for carbonyl and carboxylate oxygens respectively, with the corresponding values of W_L equal to 0.925 Å and 0.945 Å respectively, this choice being consistent with the values of W_O and W_L adopted in [1] for O_{sp^3} oxygens (1.285 Å and 0.970 Å respectively), and consistent with the global standard values.

The value of the effective radius W_N adopted for the amide nitrogen is 1.20 Å, along the direction of the π lone pair and 1.75 Å along the directions of the N–C or the N–H bonds, and the effective radius of the fictitious atoms representing each of the half-occupied π lobe orbitals is 0.35 Å.

3. Results and discussion

3.1. Ethylammonium formate and ethanolformate

In this investigation of the two model compounds with a carboxylate moiety, one of the $C_1O^{\delta-}$ bonds, C_1O_1, is held in a cis conformation with respect to the C_2C_3 bond, and the evolution of the conformational energy is studied as a function of the torsions along the C_2C_3 bond. This preselected conformation of the carboxylate enables to "maximize" the amplitude of variation of its interactions, either attractive or repulsive, with the terminal ammonium or hydroxyl groups.

3.1.1. Ethylammonium formate. The calibration of the effective radii of carboxylate oxygens, and of their associated fictitious atoms, was performed on this model molecule. The evolution of the conformational energy as a function of the torsional angle τ is represented in Fig. 1.

The intrinsically preferred conformation of ethylammonium formate is *gauche*. It corresponds to a six-membered ring, stabilized by a hydrogen-bond between O_1 and one NH proton ($d_{O_1-H} = 1.70$ Å). The value of the torsional angle obtained at the SCF level is reproduced to within 5° by the SIBFA procedure. The bond $O_1\cdots H$ has a strongly ionic character, as evidenced by the steep increase of energy
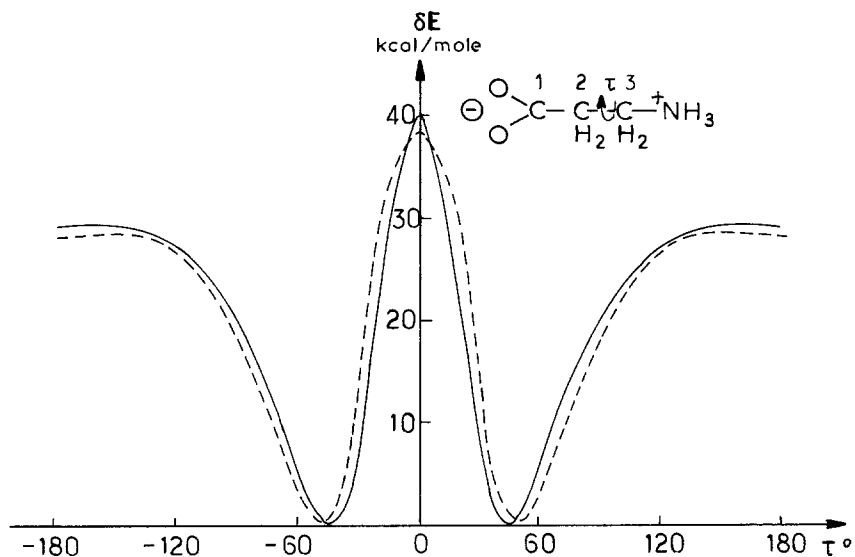

Fig. 1. Ethylammonium formate. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond CC–CN. Energy of the most stable conformer taken as energy zero

upon increasing $|\tau|$, and thereby pushing O_1 and H further apart. The "strength" of the ionic bond can be recognized in the magnitude of the energy difference between the *gauche* minimum and the *trans* conformation, in which the carboxylate and the ammonium end are the farthest apart. This energy difference amounts to 29.2 kcal (*ab initio*) or 28.1 kcal (SIBFA). These values are quite sizeable. They are, nevertheless, considerably smaller than those computed for the *inter*molecular interaction of a carboxylate with a monomethylammonium ion in their most favorable mutual positions, namely −125.6 kcal [4]. The latter value is, however, computed with respect to the infinite separation of the two entities, whereas in an *intra*molecular computation, the attraction of the oppositely charged ends remains felt in the *trans* conformation, while the interaction in the *gauche* conformation cannot be optimized fully on account of the constraints imposed by the molecular structure.

The evolution of the conformational energy is rather shallow in the range 140°–180° (and in the symmetric region) a feature reproduced by the SIBFA procedure.
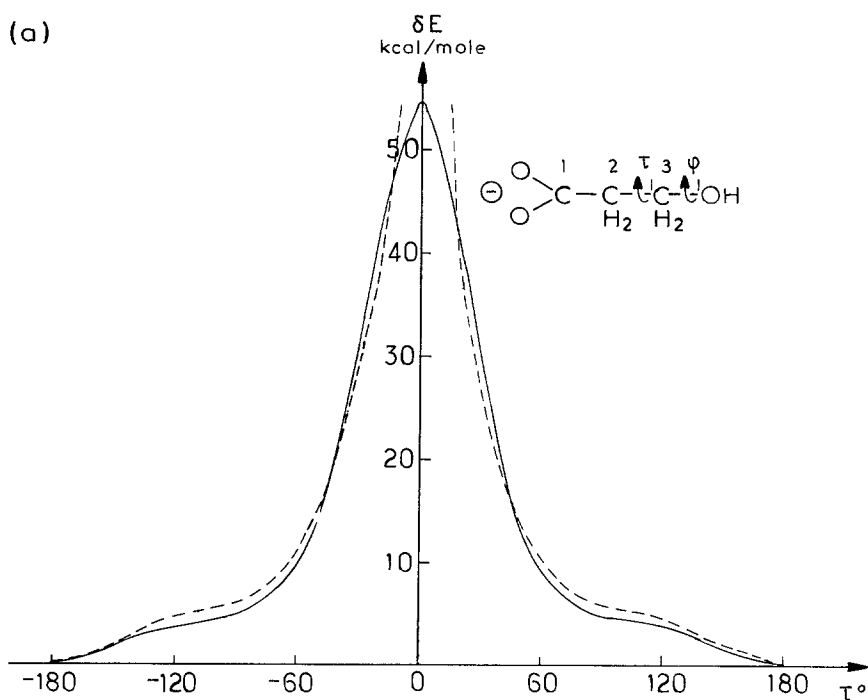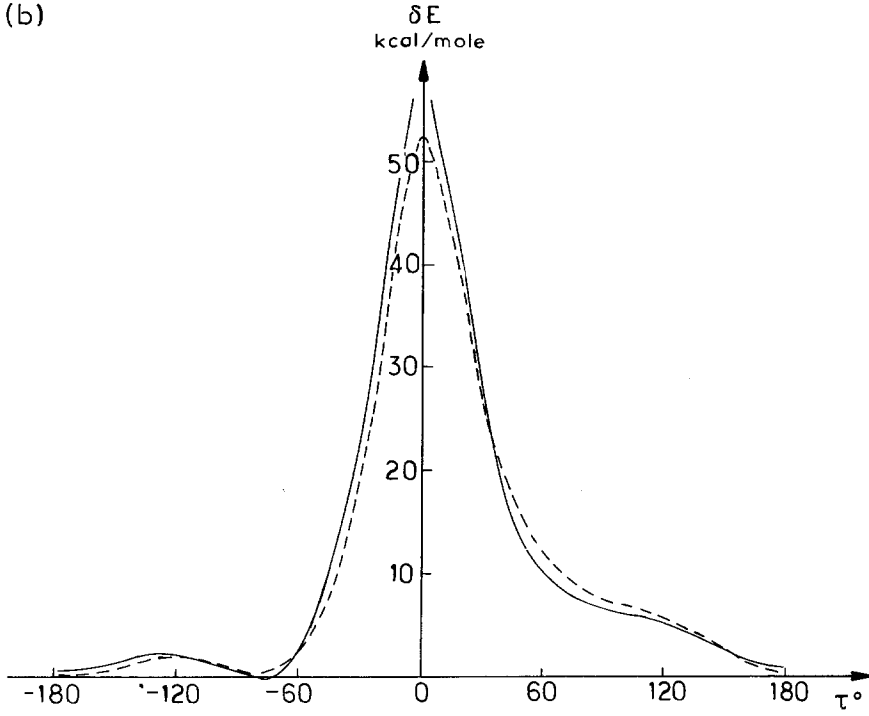

Fig. 2. Ethanolformate. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond CC–CO. Energy of the most stable conformer taken as energy zero. The torsion angle CC–OH (ϕ) is prefixed in four different values.
a) $\phi = 180°$
b) $\phi = 120°$
c) $\phi = 60°$
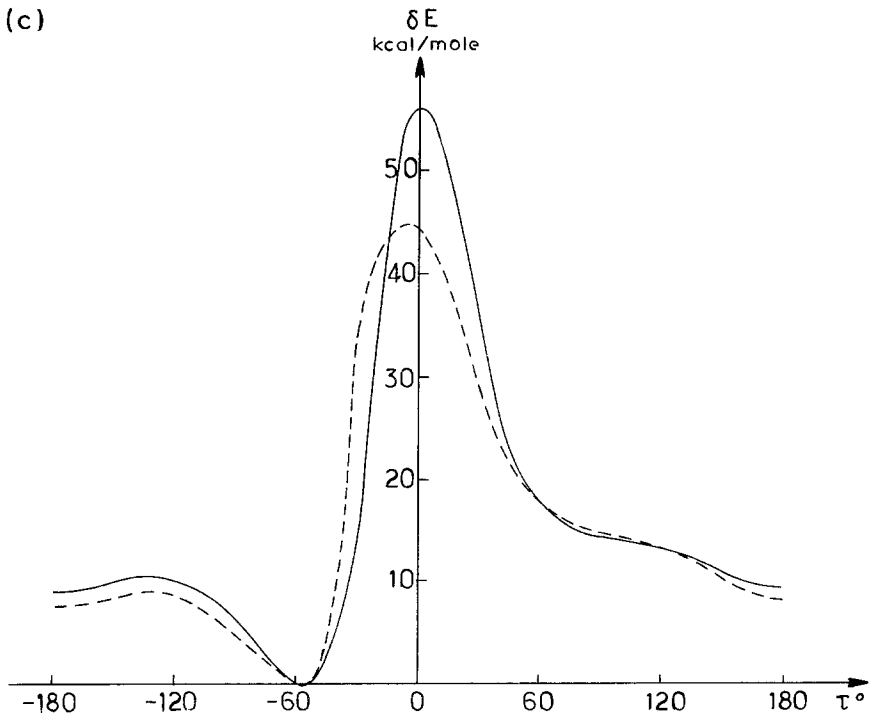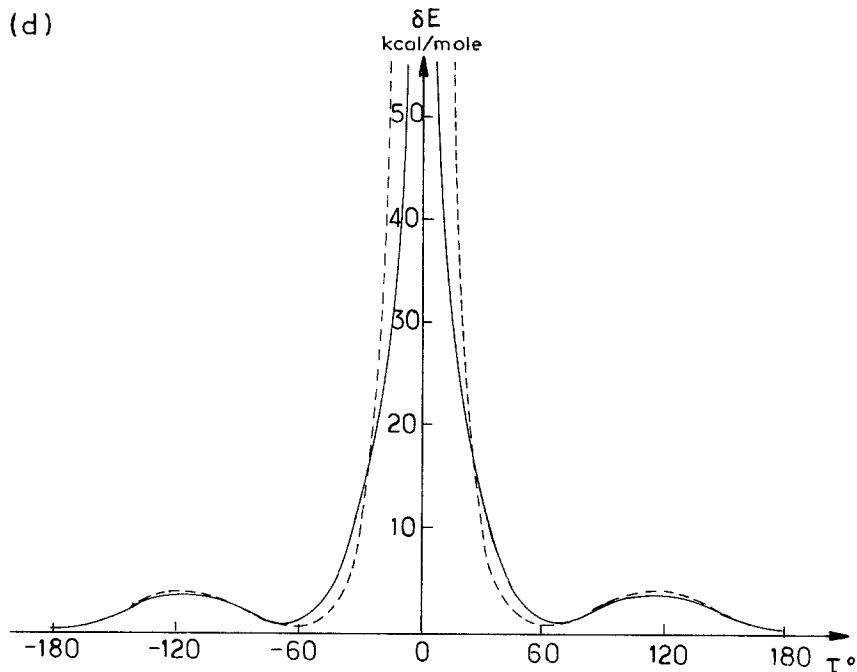d) $\phi = 0°$

(b)

(c)

Fig. 2 (*cont.*)

(d)

Fig. 2 (cont.)

A steep increase of the conformational energy occurs for $0° < |\tau| < 30°$, owing to the highly repulsive steric interactions between one NH bond and the $C-O_1$ bond. Thus, for $\tau = 30°$, the distance $O_1 \cdots H$ is 1.4 Å, appreciably shorter than the equilibrium distances encountered in either inter or intra-molecular interactions. Even though, in the present case, SIBFA reproduces rather well the ab initio results, a high accuracy of such a procedure is evidently not to be expected for such a range of distances. These are, at any rate, unlikely to be encountered in conformations having any energetical significance.

3.1.2. Ethanol formate. The variation of the conformational energy of ethanol formate was studied for rotations along the C_2-C_3 bond, the terminal hydroxyl group being held in either of the four conformations defined by the values of the torsion angle C_2C_3-OH (ϕ): 180°, 120°, 60° and 0°. (See Figs. 2a–d).

$\phi = 180°$ (Fig 2a).
The minimum energy conformation is the trans conformation. The conformational energy increases steadily from $|\tau| = 180°$ to $|\tau| = 80°$, with an inflexion point occuring for $|\tau| = 120°$, a feature accounted for by SIBFA ($\delta_{SCF} = 3.6$ kcal/mole; $\delta_{SIBFA} = 4.5$ kcal/mole).

The conformational energy increases steeply from $|\tau| = 80°$ to $|\tau| = 0°$, on account of the electrostatic repulsions involving O_1 and the hydroxyl oxygen, the energy peak being reached for $\tau = 0°$ when the distance $O \cdots O_1$ is 1.95 Å.

$\phi = 120°$ (Fig 2b).
Two nearly degenerate minima are found in the SCF and the SIBFA conformational curves. One minimum corresponds to the *trans* conformation ($\tau = 180°$). The other occurs for $\tau = -80°$. With respect to the $\phi = 180°$ situation, in which the latter minimum was absent, we observe that the electrostatic repulsion between O_1 and the hydroxyl oxygen is compensated by a favorable (albeit still remote) interaction between O_1 and the hydroxyl hydrogen. The *gauche* conformer is slightly favored by the SCF computations, whereas the *trans* conformer is slightly favored by the SIBFA computations but the energy separation between the two minima is so small (0.4 kcal/mole) that we do not consider the difference serious. A local maximum is found for $\tau = -120°$, for which the two C_3H bonds and the C_3-O bond eclipse C_1C_2 and the two C_2H bonds ($\delta_{SCF} = 2.0$ kcal/mole; $\delta_{SIBFA} = 1.8$ kcal/mole). An inflexion point is found by both procedures for $\tau = 100°$, and a steep increase occurs in the range $-60° < \tau < 60°$, on account of the electrostatic repulsions between the hydroxyl oxygen and O_1, the global maximum occurring for $\tau = 0°$ ($\delta_{SCF} = 57.3$ kcal/mole; $\delta_{SIBFA} = 52.3$ kcal/mole).

$\phi = 60°$ (Fig. 2c).
The intrinsically preferred conformation is *gauche* ($\tau = -60°$), and corresponds to the formation of a six-membered ring stabilized by a hydrogen bond between the hydroxyl hydrogen and $O_1(d_{O_1-H} = 1.98 Å)$. The *trans* conformation is a local energy minimum. The *trans-gauche* energy difference is high ($\delta_{SCF} = 8.9$ kcal/mole; $\delta_{SIBFA} = 7.7$ kcal/mole) translating the strength of the hydrogen bond formed. A local maximum is found for $\tau = -140°$, and an inflexion point appears for $\tau = 80°$. The global maximum appears for $\tau = 0°$.

Discrepancies appear between the *ab initio* and SIBFA curves for $|\tau| < 40°$. For $-40° < \tau < -20°$, the SIBFA curves are higher in energy than their *ab initio* conterpart, owing to very close steric contacts between the hydroxyl hydrogen and O_1. Thus, for $\tau = -40°$, $d_{O_1-H} = 1.53 Å$ and for $\tau = -20°$, $d_{O_1-H} = 1.46 Å$. As discussed above, the occurrence of such conformations is highly unlikely. For $0° < \tau < 20°$, the SIBFA curve on the other hand is distinctly lower than the *ab initio* curve. The points of the curve, are, however, more than 35 kcal/mole above the energy minimum.

$\phi = 0°$ (Fig. 2d).
The global energy minimum corresponds to the *trans* conformation. The *gauche* conformation ($\tau = \pm60°$) corresponds to a local minimum ($d_{O_1-H} = 2.37 Å$) ($\delta_{SCF} = 1.4$ kcal/mole; $\delta_{SIBFA} = 0.8$ kcal/mole). A local maximum is found for $\tau = \pm120°$. A steep increase occurs for $|\tau| < 40°$, on account of the increase of the steric repulsive interactions between these two atoms. These repulsions culminate for $\tau = 0°$, at the global maximum, as the distance between O_1 and H becomes equal to 1.0 Å.

3.2. C-ethylammonium formamide and C-ethanol formamide

In this investigation of the conformational behavior of the model compounds bearing on amide carbonyl, the carbonyl bond is held in a *cis* conformation with
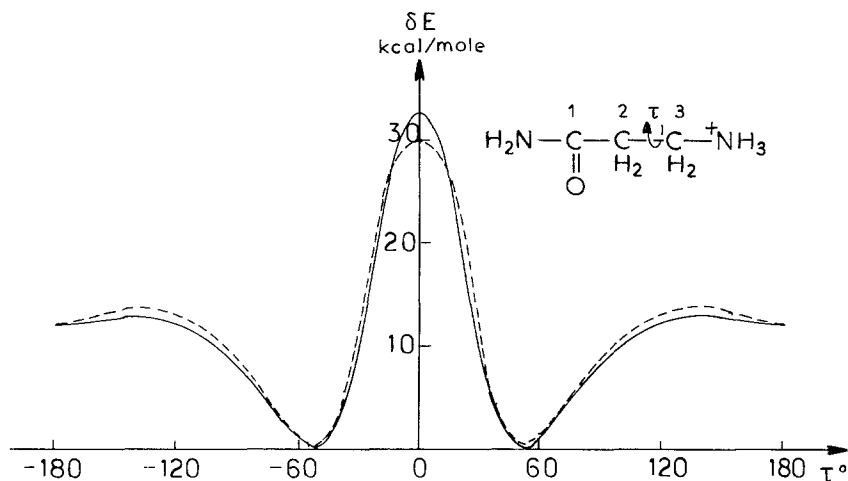
Fig. 3. C-ethylammonium formamide. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond CC–CN. Energy of the most stable conformer taken as energy zero

respect to the C_2C_3 bond and conformational variations are generated along the C_2C_3 bond. This conformation of the carbonyl group permits to "maximize" the amplitude of variations of the interactions, either attractive or repulsive, which involve the carbonyl and the terminal ammonium or hydroxyl end. For C-ethanol formamide, the terminal hydroxyl group is held successively in one of the four conformations defined by the values of the angle C_2C_3–OH (ϕ) = 180°, 120°, 60° and 0°.

3.2.1. C-ethylammonium formamide (Fig. 3). As mentioned above, the calibration of the effective radii of the carbonyl oxygens and of their associated fictitious atoms, was performed on this molecule.

The intrinsically preferred conformation is *gauche* ($\tau = \pm 50°$ in both *ab initio* and SIBFA), and is stabilized by a close approach to one ammonium hydrogen to the carbonyl oxygen ($d_{O-H} = 1.77$ Å), leading to the formation of a six-membered ring. The *trans* conformation is a local minimum, separated from the global minimum by a large energy difference of 12 kcal/mole in both procedures (compare with the corresponding energy difference of 29 kcal/mole found for ethylammonium formate). A local energy maximum exists for $\tau = 140°$ ($\delta_{SCF} = 12.8$ kcal/mole; $\delta_{SIBFA} = 13.8$ kcal/mole), whereas the global maximum is found for $\tau = 0°$, destabilized by strong steric repulsions between the carbonyl and two NH bonds ($d_{H_1}\cdots O = 1.72$ Å; $d_{H_2}\cdots O = 1.72$ Å; $d_{O-N} = 2.02$ Å).

3.2.2. C-ethanol formamide.

$\phi = 180°$ *(Fig. 4a).*
The global minimum is the *trans* conformation, clearly due to the electrostatic repulsions between the carbonyl and the hydroxyl oxygens. A local minimum
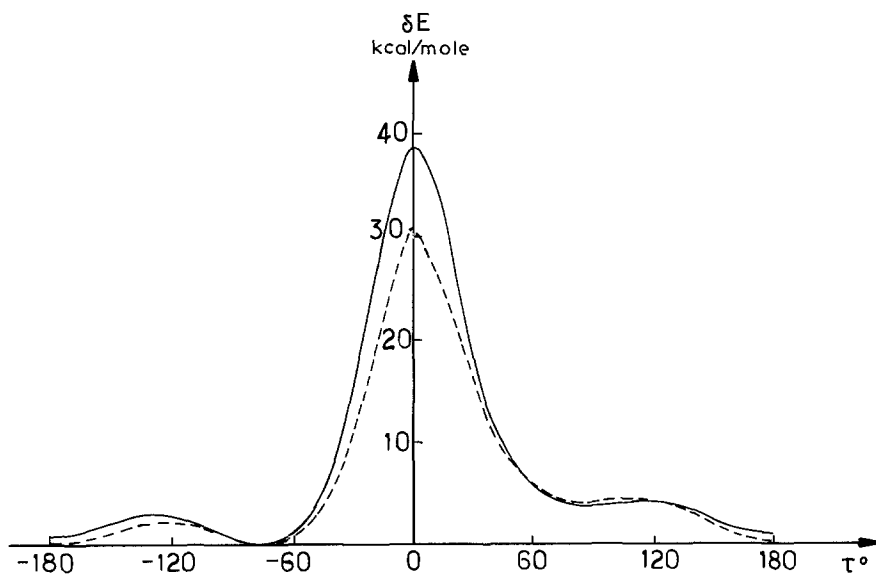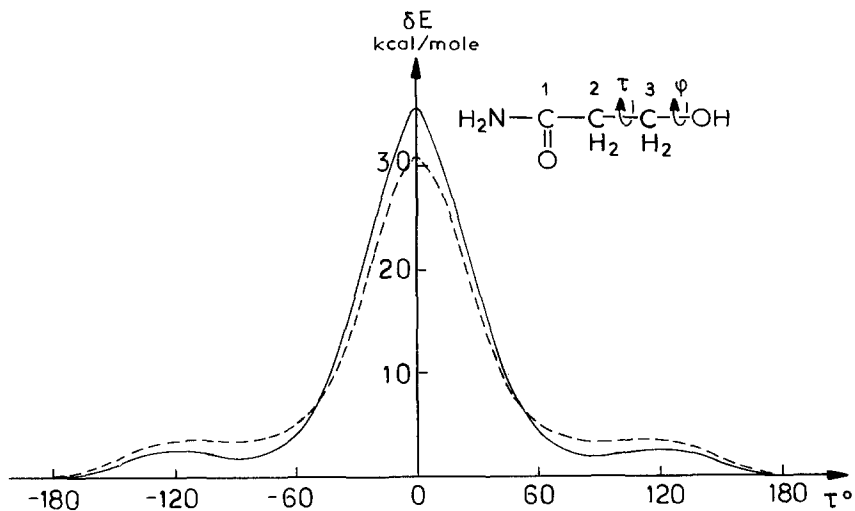
Fig. 4. C-ethanol formamide. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond CC–CO. Energy of the most stable conformer taken as energy zero. The torsion angle CC–OH (ϕ) is prefixed in four different values.

a) $\phi = 180°$

b) $\phi = 120°$

c) $\phi = 60°$

d) $\phi = 0°$

(c)

(d)
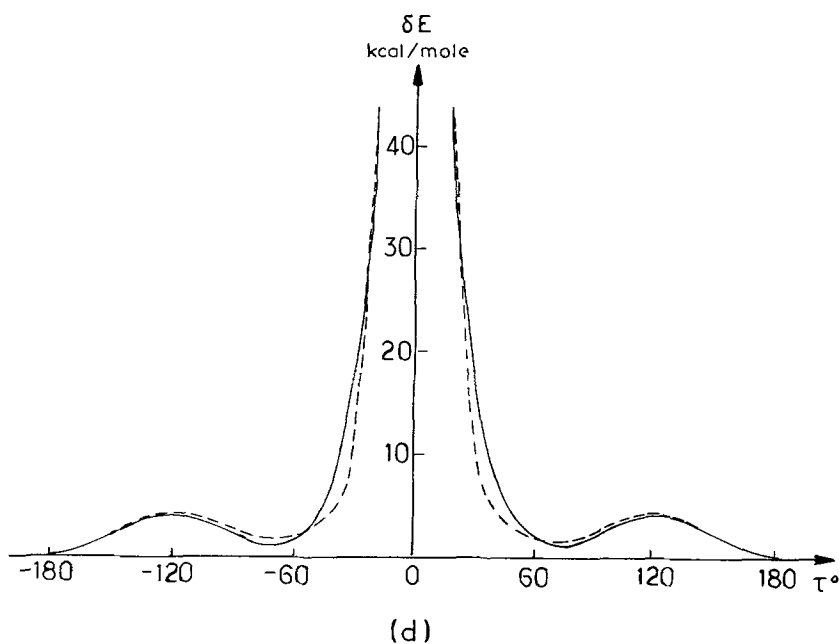
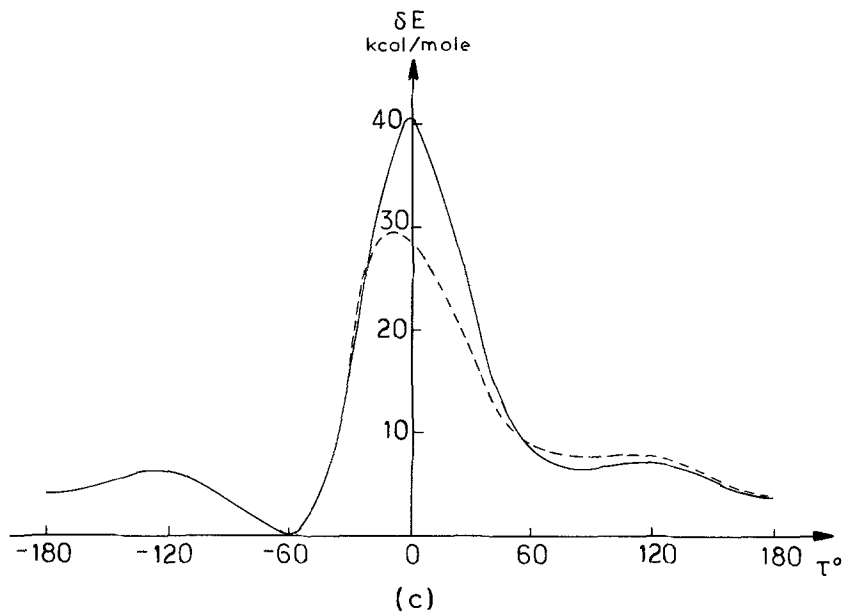**Fig. 4** (cont.)

occurs for $|\tau| = 80°$ ($\delta_{SCF} \approx 1.9$ kcal/mole; $\delta_{SIBFA} = 3.3$ kcal/mole), the energy curves being shallow in the range of $|\tau|$ values comprised between 80° and 140°, presenting a local maximum for $|\tau| = 120°$ ($\delta_{SCF} = 2.6$ kcal/mole; $\delta_{SIBFA} = 3.6$ kcal/mole). The global energy maximum occurs at $\tau = 0°$

$\phi = 120°$ *(Fig. 4b).*
Two energy minima are observed, one *gauche* ($\tau = -70°$) and one *trans* conformation. The electrostatic repulsions between the carbonyl and hydroxyl oxygens appear compensated by a favourable, albeit remote, attraction involving the carbonyl oxygen and the hydroxyl hydrogen (compare with the results obtained with C-ethanol formate for $\phi = 120°$). Both the SIBFA and the SCF calculations favour the gauche conformation (δ gauche-trans (SCF) = 1 kcal/mole, $\delta_{SIBFA} =$ 0.1 kcal/mole).

The global energy maximum is found for $\tau = 0°$, and two local maxima occur for $\tau = 120°$ ($\delta_{SCF} = 4.1$ kcal/mole, $\delta_{SIBFA} = 4.2$ kcal/mole), and $\tau = -120°$ ($\delta_{SCF} =$ 2.9 kcal/mole; $\delta_{SIBFA} = 2.4$ kcal/mole), the curves in the range $60° < \tau < 140°$ being very shallow.

$\phi = 60°$ *(Fig. 4c).*
The preferred conformation is *gauche* ($\tau = -60°$), stabilized by a hydrogen bond between the carbonyl oxygen and the hydroxyl hydrogen ($d_{O-H} = 2.00$ Å). The *trans* conformation corresponds to a local minimum. The energy separation with respect to the gauche conformer, which translates the strength of the hydrogen bond in the presence of the constraints imposed by the molecular structure, amounts to 4.0 (SCF) and 3.9 (SIBFA) kcal/mole respectively.

Two local maxima are found for $\tau = 110°$ ($\delta_{SCF} = 7.3$ kcal/mole; $\delta_{SIBFA} =$ 8.1 kcal/mole), and for $\tau = -120°$ ($\delta_{SCF} = 6.3$ kcal/mole; $\delta_{SIBFA} = 6.1$ kcal/mole), the curves in the range $60° < \tau < 140°$ being shallow. The global energy maximum occurs for $\tau = -10°$. The discrepancies between the two procedures in the interval $0° < |\tau| < 20°$ correspond, as emphasized above, to conformations which are highly unfavorable on account of steric repulsions between the carbonyl oxygen and the hydroxyl group.

$\phi = 0°$ *(Fig. 4d).*
The global minimum corresponds to the *trans* conformation. The *gauche* conformation ($|\tau| = 70°$) corresponds to a local minimum ($\delta_{SCF} = 0.8$ kcal/mole; $\delta_{SIBFA} =$ 1.2 kcal/mole). The attractive interactions between the carbonyl and the hydroxyl hydrogen are counteracted by repulsive steric interactions (compare with a similar situation occurring for $\phi = 0°$ with C-ethanol formate). A local maximum is found for $\tau = 120°$ ($\delta_{SCF} = 4.1$ kcal/mole; $\delta_{SIBFA} = 4.3$ kcal/mole), with a disappearance of the plateau region observed in the previous curves of C ethanol-formamide in the zone comprised between 80° and 140°.

A very steep increase of the conformational energy occurs for $|\tau| < 40°$, due to very close contacts between the carbonyl oxygen and the hydroxyl hydrogen (for $|\tau| = 20°$, $d_{O-H} = 1.32$ Å).

3.3. N *ethylammonium-formamide* and N *ethanol-formamide*

In this investigation of the intrinsic preferences of N-substituted amides, the dihedral angle CNC_2C_3 was held at 90°, and rotations were performed about the
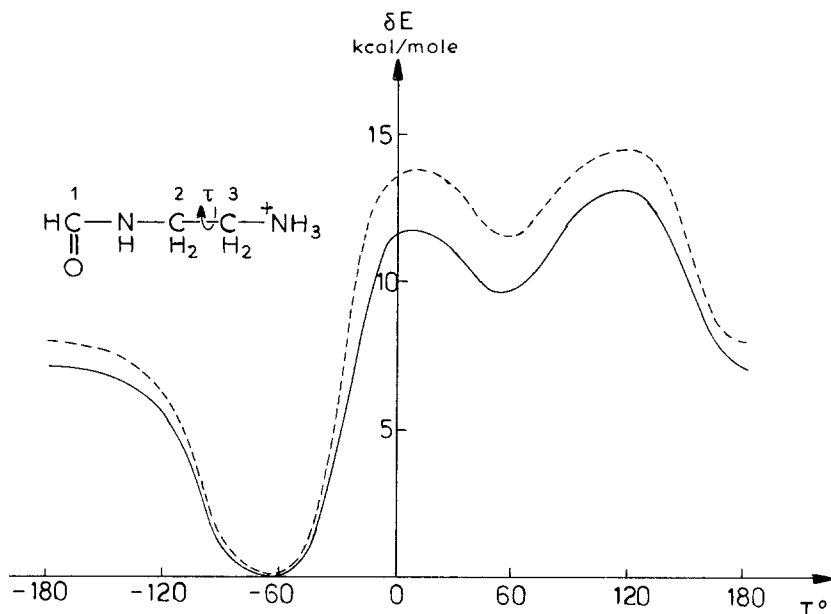
Fig. 5. N-ethylammonium formamide. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond NC-CN. Energy of the most stable conformer taken as energy zero

C_2-C_3 bond. The constraint on CNC_2C_3 was adopted so as to "force" the terminal ammonium or hydroxyl end to interact with the amide group over its plane, and in particular over the region of the nitrogen π lone pair in view of assessing the ability of the SIBFA procedure to reproduce the interactions, either attractive or repulsive, involving this lone pair, as well as the contribution arising from the carbonyl oxygen in the explored regions.

3.3.1. N ethylammonium formamide (Fig. 5). The intrinsically preferred conformation is *gauche* ($\tau = -60°$), stabilized by a hydrogen bond between one ammonium hydrogen and the carbonyl oxygen, from above the amide plane ($d_{O-H} = 2.50$ Å), leading to the formation of a seven-membered ring. A local energy minimum is found for $\tau = 60°$, but its energy separation from the global minimum is high ($\delta_{SCF} = 9.7$ kcal/mole; $\delta_{SIBFA} = 11.5$ kcal/mole). A rapid increase of the intramolecular energy occurs from $\tau = -60°$ to $\tau = 0°$, as the terminal ammonium moves over the C-N and NC_1 bonds, and above the nitrogen lone-pair. A high energy local maximum occurs close to $\tau = 0°$ ($\delta_{SCF} = 11.6$ kcal/mole; $\delta_{SIBFA} = 13.6$ kcal/mole), whereas the global maximum is located at $\tau = 120°$ ($\delta_{SCF} = 13.3$ kcal/mole; $\delta_{SIBFA} = 14.6$ kcal/mole).

3.2.2. N ethanol formamide. We have investigated the conformational behaviour of N ethanol formamide for three conformations of the terminal hydroxyl group defined by the values of the torsional angle C_2C_3-OH (ϕ): 180°, 0°, −60°.

(a)

(b)

(c)
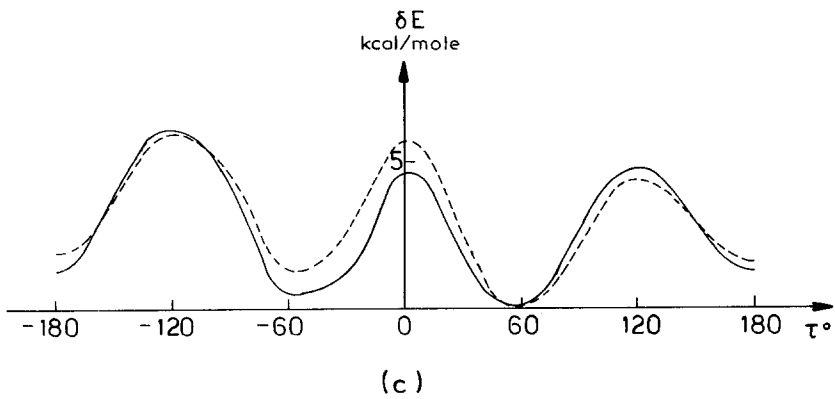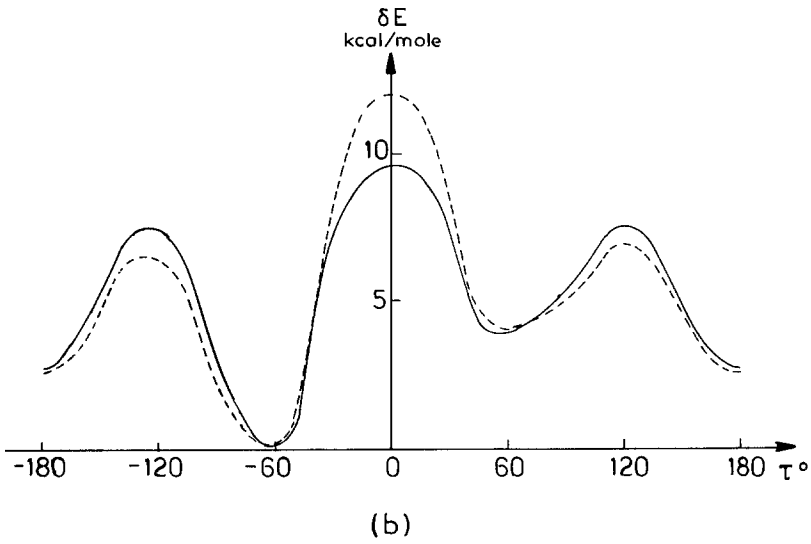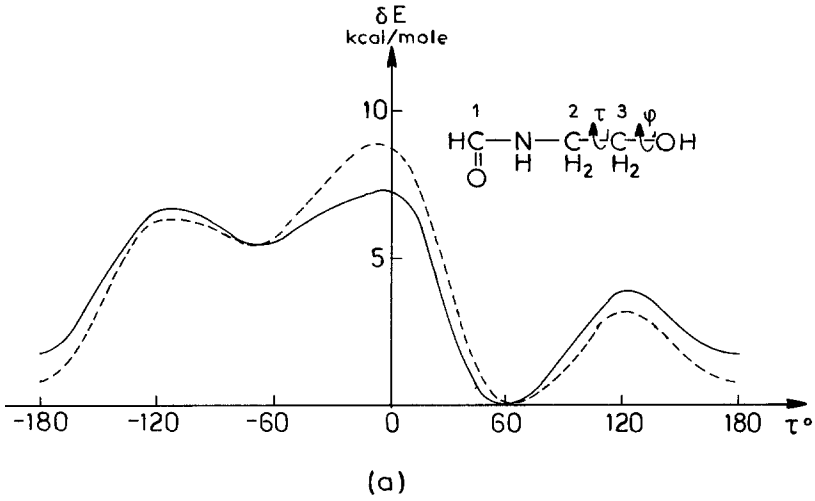
$\phi = 180°$ *(Fig. 6a)*.

The global energy minimum is *gauche* ($\tau = 60°$), stabilized by an attractive, albeit remote, interaction between the hydroxyl oxygen and the NH hydrogen ($d_{O-H} = 2.87$ Å). The *trans* conformation is a local minimum, energetically close to the first one ($\delta_{SCF} = 1.7$ kcal/mole; $\delta_{SIBFA} = 0.7$ kcal/mole). Another high-energy local minimum is found for $\tau = -60°$, but with an energy difference with respect to the global minimum appreciably larger ($\delta_{SCF} = \delta_{SIBFA} = 5.5$ kcal/mole). The global maximum is found close to $\tau = 0°$. Two local minima occur, one for $\tau = 120°$ ($\delta_{SCF} = 3.8$ kcal/mole; $\delta_{SIBFA} = 3.1$ kcal/mole), and the other for $\tau = -120°$ ($\delta_{SCF} = 6.6$ kcal/mole; $\delta_{SIBFA} = 6.2$ kcal/mole). The right side local maximum is appreciably lower in energy than the left side one, which is destabilized by repulsive electrostatic interactions between the carbonyl and hydroxyl oxygens.

$\phi = 0°$ *(Fig. 6b)*.

The calibration of the effective radii of the amide nitrogens and their associated fictitious atoms was performed on this molecule. The intrinsically preferred conformation is *gauche* ($\tau = -60°$), stabilized by a hydrogen bond between the hydroxyl hydrogen and the carbonyl oxygen ($d_{O-H} = 2.03$ Å). The *trans* conformation is a local minimum ($\delta_{SCF} = 2.7$ kcal/mole; $\delta_{SIBFA} = 2.6$ kcal/mole). Another minimum occurs at $\tau = 60°$ ($\delta_{SCF} = 3.9$ kcal/mole; $\delta_{SIBFA} = 4.1$ kcal/mole). The global maximum occurs at $\tau = 0°$ ($\delta_{SCF} = 9.8$ kcal/mole; $\delta_{SIBFA} = 12$ kcal/mole), and two local maxima appear at $\tau = 120°$ and $\tau = -120°$, the left side maximum being only slightly more stable than the right side one.

$\phi = -60°$ *(Fig. 6c)*.

Three energetically close minima appear. The most stable one, which corresponds to a *gauche* conformation ($\tau = 60°$), is stabilized by a "residual" attraction between the hydroxyl oxygen and the NH hydrogen (compare with the situation occuring for $\phi = 180°$). The second minimum, also a *gauche* conformation ($\tau = -60°$; $\delta_{SCF} = 0.5$ kcal/mole; $\delta_{SIBFA} = 1.3$ kcal/mole), is stabilized by a residual attraction between the hydroxyl hydrogen and the carbonyl oxygen (compare with the situation occurring with $\phi = 0°$). the third minimum corresponds to the *trans* conformation ($\delta_{SCF} = 1.2$ kcal/mole; $\delta_{SIBFA} = 1.5$ kcal/mole). The global energy maximum now occurs for $\tau = -120°$ ($\delta_{SCF} = 6.3$ kcal/mole; $\delta_{SIBFA} = 6.2$ kcal/mole), and two local maxima are found for $\tau = 0°$ and $\tau = 120°$.

4. Conformational study of the glycine and alanine dipeptides

An analysis of the conformational preferences expressed at the level of the dipeptide building-blocks is of fundamental importance for the understanding

---

Fig. 6. N-ethanol formamide. Evolutions of the *ab initio* (*full lines*) and the SIBFA (*dotted lines*) energy curves for torsions along the bond NC–CO. Energy of the most stable conformer taken as energy zero. The torsion angle NC–CO (ϕ) is prefixed in three different values.
a) $\phi = 180°$
b) $\phi = 0°$
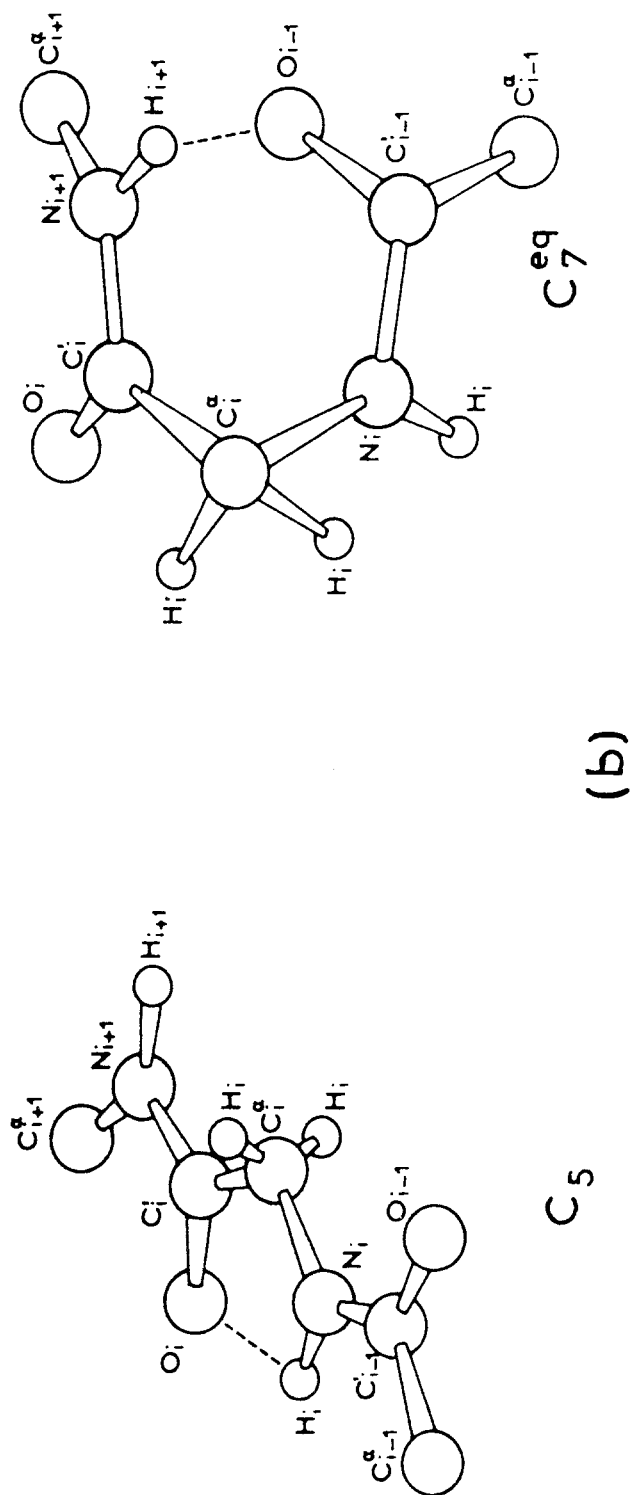c) $\phi = -60°$

(a)
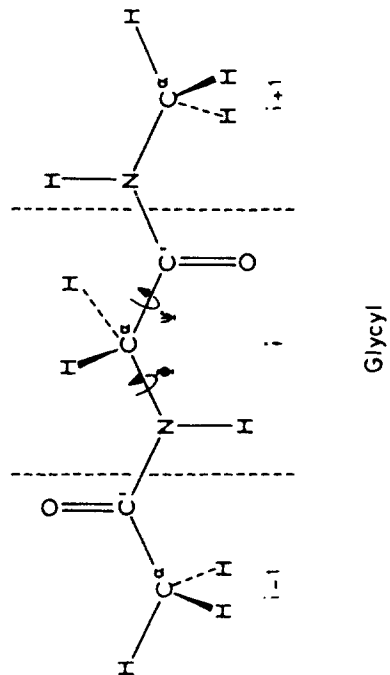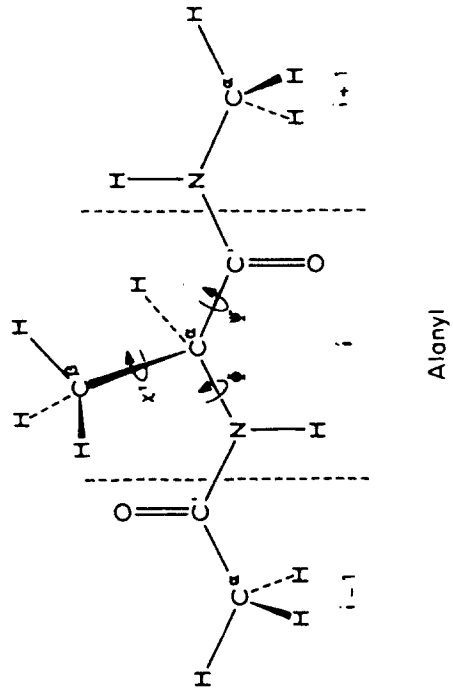
Fig. 7. The glycyl and alanyl dipeptides. (**a**) Molecular structure. (**b**) Representation of the C_5 and C_7^{eq} conformations

of the conformational properties of oligo- or polypeptides. Numerous investigations have been devoted to study the conformations of the glycine and alanine dipeptides as a function of their backbone torsional angles ϕ and ψ (see Fig. 7), using both spectroscopic techniques in apolar solvents [5–8] and theoretical computations [9–18]. Experimental studies [5–8], as well as *ab initio* SCF computations [9–15] or *refined* semi-empirical [16] or empirical [17–18] procedures indicate the prevalence of two essential conformations, each stabilized by an intramolecular hydrogen-bond: a five-membered ring denoted C_5, and a seven-membered ring denoted C_7^{eq} (see Fig. 7b). Furthermore, whereas the energy difference between the two conformers is small, the relative stabilization of the C_7^{eq} conformation with respect to the C_5 conformation is increased in the alanyl dipeptide as compared to the glycyl dipeptide.

The numerical experience gathered so far by previous theoretical computations will enable us to assess the possibility to utilize the SIBFA procedure in conformational studies of dipeptides. The model dipeptides N-acetyl-N'-methyl-glycylamide and N-acetyl-N' methylalanylamide are built out of the fragments formamide and methane with standard internal geometries [16].

The three-fold torsional energy contributions

$$E_{tor} = \tfrac{1}{2} V_0 (1 + \cos 3\chi)$$

for torsions along the bonds C–N and C–C_α were calibrated so as to reproduce the barrier height for torsion of the methyl group in the two compounds C-methylacetamide and N-methylacetamide respectively, as computed with our usual SCF pseudopotential procedure. In these two compounds, the barrier is the highest for conformations eclipsing the NH bond and one C–H methyl. (Compare with the results obtained in references [19–20]). The energy differences with respect to the staggered conformation are 1.2 kcal/mole and 0.50 kcal/mole respectively. The corresponding values of V_0 in the SIBFA procedure are 0.3 and 0.2 kcal/mole, respectively, the other terms in equation (1) providing the remainder of the barrier [1]. Thus, for torsions ϕ along the C–N bond, $\chi = 180 - \phi$ and for torsions ψ along the C–C bond, $\chi = \psi$.

We found furthermore that an additional term is required in order to account satisfactorily for the complex interplay of the different interactions between two amide moieties separated by an α carbon. This contribution acts in such a way as to favor conformations in which the two amide planes are coplanar, and to disfavor maximally conformations in which the two planes are perpendicular, with a magnitude larger in the glycyl than in the alanyl dipeptide, translating the possiblity of hyperconjugation in the former. We have adopted *tentatively* for this term a simple expression:

$$V_\pi = V_\pi^0 \left(1 - \left| \frac{2\gamma}{\pi} - 1 \right| \right)$$

in which γ denotes the angle between the normals to the two formamide planes and chose V_π^0 for each dipeptide so as to reproduce the values of the pseudopoten-

Table 1. The three representative conformers of the glycyl and alanyl dipeptides. Values of the *ab initio* SCF conformational energy differences with respect to the most stable conformer, and values of the SIBFA energy differences between the conformers. Energies in kcal/mole

	ϕ	ψ	δ_{SCF}	δ_{SIBFA}		ϕ	ψ	δ_{SCF}	δ_{SIBFA}
		Glycyl dipeptide					Alanyl dipeptide		
C_5	180	180	0.0	0.0	C_5	180	180	1.0	1.2
$C_{5'}$	−160	160	0.8	1.1	$C_{5'}$	−160	160	0.4	0.7
C_7	−80	80	1.4	1.1	C_7	−80	80	0.0	0.0

tial *ab initio* SCF energy differences computed between the three representative conformations: C_7^{eq}, C_5 and $C_{5'}$ (see, e.g. Refs. [13–19]), defined by the following set of angles (resp.): $\{\phi, \psi\} = \{-80, 80\}$, $\{-180, 180\}$ and $\{-160, 160\}$. Table 1 indicates the *ab initio* values and their SIBFA counterparts obtained with $V_\pi^0 = 2.3$ and 4.5 kcal/mole, respectively, for the alanyl and glycyl dipeptide. The conformational maps of the glycyl and alanyl dipeptides obtained with the final parameters are represented in Figs. 8a and 8b respectively.
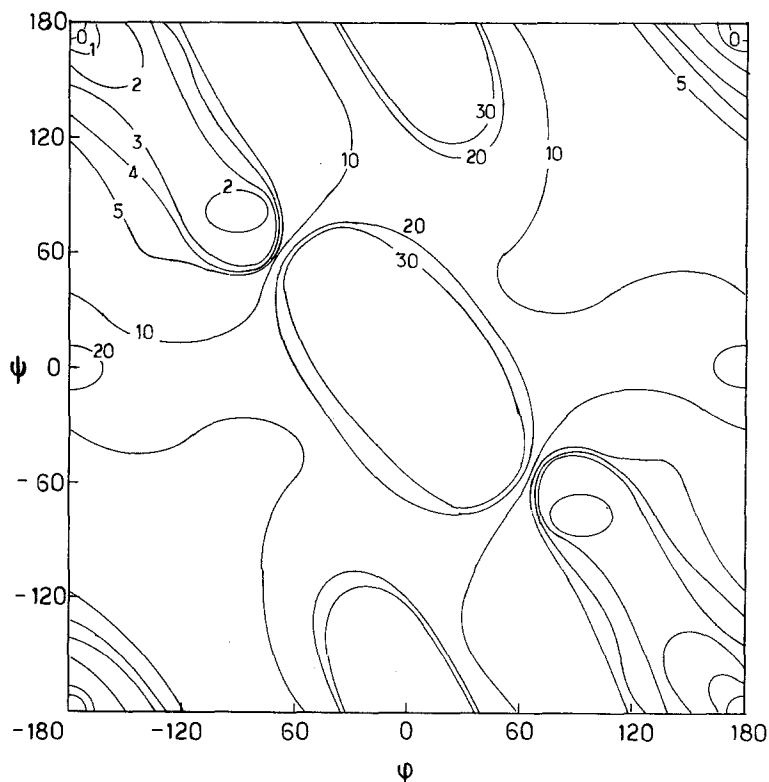
(a)

Fig. 8. Conformational maps of the glycyl (**a**) and the alanyl (**b**) dipeptides
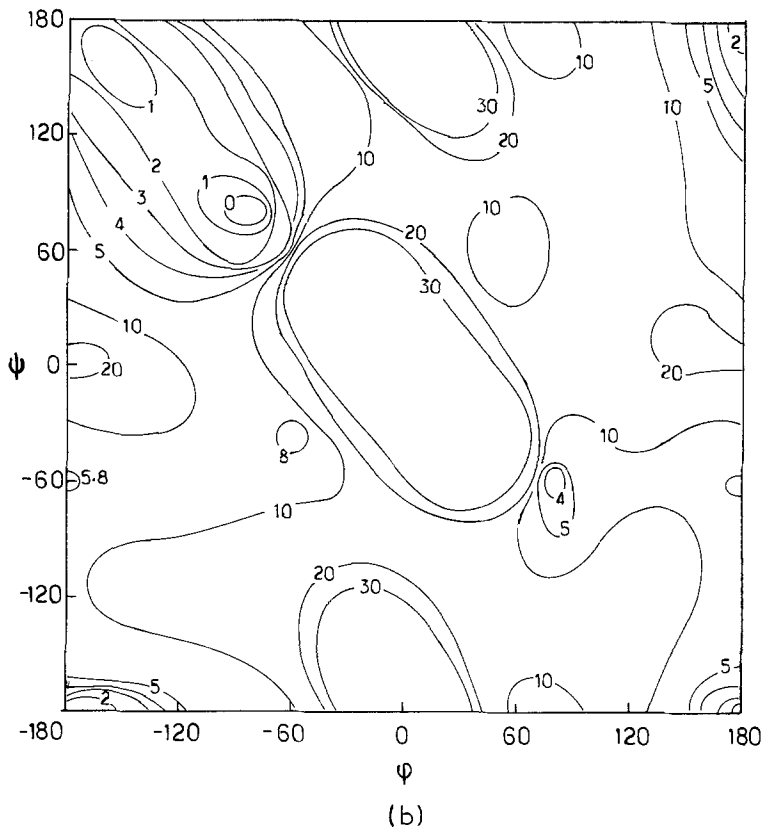
Fig. 8 *(cont.)*

In the glycyl dipeptide, the fully extended conformation corresponds to the global energy minimum, with the C_7 conformation as a local energy minimum at 1 kcal/mole above the first (Table 1). A 5 kcal/mole energy contour encircles the C_5-C_7 region. The α-helix is located at the limits of a contour encompassing points that are between 6 and 10 kcal/mole from the global minimum (for $\phi = \psi = -60°$, $\delta E = 10$ kcal/mole from the global energy minimum).

The conformational map is very similar to the ones drawn on the basis of *ab initio* SCF computations [11, 12, 13], with the reservation, however, that numerical differences may still appear between the different calculations due to the use of different basis sets or different internal geometries (as in [13]).

In the alanyl dipeptide, the C_7^{eq} conformation ($\phi = -80°$, $\psi = 80°$) is the conformation corresponding to the global energy minimum. A well-defined local energy minimum is now found for a *distorted* C_5 conformation ($\phi = -160°$, $\psi = 160°$). This effect of the methyl side chain on the distortion of the C_5 conformation was formerly evidenced by experimentation [5–8] as well as *ab initio* SCF computations [9, 12, 13, 15]. Other well-defined local minima are found for the C_7^{ax}

conformation ($\phi = 80°$, $\psi = -60°$, $\delta E = 3.7$ kcal/mole), for the α' conformation ($\phi = 180°$, $\psi = -60°$, $\delta E = 5.8$ kcal/mole), close to the α_R helix ($\phi = -60°$, $\psi = -80°$; $\delta E = 8.0$ kcal/mole), and for the α_L helix ($\phi = 60°$, $\psi = 60°$, $\delta E = 9.8$ kcal/mole).

The existence of a wide contour of points within 5 kcal/mole from the global minimum can be noted in the upper left corner of figure 8b and in the bottom right corner, encircling the $C_5-C_7^{eq}$ region. The $\beta_{\parallel}$ conformation ($\phi = -120°$, $\psi = 40°$) is located at the extreme limit of the contour ($\delta E = 4.7$ kcal/mole). The existence of a cluster of points within 10 kcal/mole from the global minimum can be finally noted for $60° < \phi < 80°$ and $160° < \psi < 180°$.

The map resembles closely those obtained for the alanyl dipeptide by *ab initio* SCF computations [12, 13] (more particularly to the map computed with the $3s2p$ and 4-31G basis sets). Again, however, it must be emphasized that numerical energy differences with these calculations may arise at given points of the conformational map owing to the use of different basis sets and/or different internal geometries.

5. Conclusions

The results of this study, in conjunction with those of Ref. [1], show that it is possible to reproduce very satisfactorily, in the framework of the SIBFA procedure, all the particular details of *ab initio* SCF conformational energy computations. The sole serious numerical discrepancies appear for conformations characterized by very close contacts between terminal ends, for instance typically for oxygen–hydrogen distances shorter than 1.5 Å. It is clear that such conformations are of little importance since their conformational energy is very high and that such intermolecular distances are not likely to occur anyway.

The interactions investigated in this work and in the preceding study [1] cover a wide range of interactions expected to occur between functional or liganding groups in proteins and ionophores, and we are presently engaged in applications of the procedure to study the interplay of intra- and intermolecular interactions. A recent study of this kind using the SIBFA procedure, and dealing with the gramicidin A transmembrane channel was recently performed in our laboratory [21].

References

1. Gresh, N., Claverie, P., Pullman, A.: Theoret. Chim. Acta (Biol.) **66**, 1 (1984)
2. Gresh, N., Claverie, P., Pullman, A.: Int. J. Quantum Chem., Symp. **13**, 243 (1979)
3. Clementi, E., Raimondi, D.: J. Chem. Phys. **38**, 2686 (1963)

4. Gresh, N., Pullman, B.: Biochim. Biophys. Acta **625**, 356 (1980)

5. Avignon, M., Huong, P., Lascombe, J., Marraud, M., Neel, J.: Biopolymers **8**, 69 (1969)

6. Marraud, M., Neel, J., Avignon, M., Huong, P.: J. Chim. Phys. **67**, 959 (1970)

7. Cung, M., Marraud, M., Neel, J.: in Conformation of biological molecules and polymers, the 5th Jerusalem Symposium in Quantum Chemistry and Biochemistry, Bergmann, E., Pullman, B. Eds. New York: Academic Press 1973

8. Burgess, A., Scheraga, H.: Biopolymers **12**, 2177 (1973)

9. Pullman, A., Berthod, H.: Compt. Rend. Acad. Sci. Paris, **277**, 2077 (1973)

10. Shipman, L., Christoffersen, R.: J. Am. Chem. Soc. USA **95**, 1408 (1973)

11. Robson, B., Hillier, I., Guest, M.: J. Chem. Soc. (Faraday Trans. II) **74**, 1311 (1978)

12. Hillier, I., Robson, B.: J. Theor. Biol. **76**, 83 (1979)

13. Peters, D., Peters, J.: J. Mol. Struct. **85**, 107 (1981)

14. Schäfer, L., Van Alsenoy, C., Scarsdale, J.: J. Chem. Phys. **76**, 1439 (1982)

15. Scarsdale, J. Van Alsenoy, C. Klimkowski, V., Schäfer, L., Momany, F.: J. Am. Chem. Soc. USA **105**, 3438 (1983)

16. Pullman, B., Pullman, A.: Adv. Protein Chem. **28**, 347 (1974), and references therein

17. Vasquez, M., Nemethy, G., Scheraga, H.: Macromolecules **16**, 1043 (1983)

18. Weiner, S., Kollman, P., Case, D., Chandra Singh, V., Ghio, C., Alagona, G., Profeta, Jr., S., Weiner, P.: J. Am. Chem. Soc. USA **106**, 765 (1984)

19. Perricaudet, M., Pullman, A.: Int. J. Peptide and Protein Res. **5**, 99 (1973)

20. Hagler, A., Leiserowitz, L., Tuval, M.: J. Am. Chem. Soc. USA **98**, 4600 (1976)

21. Etchebest, C., Pullman, A.: FEBS Letters **170**, 191 (1984)