Theoret. Chim. Acta (Berl.) 66, 1–20 (1984)

THEORETICA CHIMICA ACTA

Theoretical studies of molecular conformation. Derivation of an additive procedure for the computation of intramolecular interaction energies. Comparison with *ab initio* SCF computations

Nohad Gresh, Pierre Claverie, and Alberte Pullman

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

An additive procedure (SIBFA) is developed for the rapid computation of conformational energy variations in very large molecules. The macromolecule is built out of constitutive molecular fragments and the intramolecular energy is computed as a sum of interaction energies between the fragments. The electrostatic and the polarization components are calculated using multicenter multipole expansions of the *ab initio* SCF electron density of the fragments. The repulsion component is obtained as a sum of bond and lone pair interactions.

Tests of the procedure on a series of model compounds containing ether oxygens and pyridine-like nitrogens are reported and compared with the results of corresponding *ab initio* SCF calculations. The resulting methodology is compatible with the simultaneous computation of intermolecular interactions.

Key words: Macromolecular conformations, computation of — intramolecular interactions — SIBFA (Sum of Interactions Between Fragments computed *Ab initio*) method.

1. Introduction

The theoretical investigation of the binding specificities of large biomolecules often requires the simultaneous calculation of both inter- and intramolecular interactions, the two dominant factors governing the binding selectivities being

the *inter*molecular interaction energy between the macromolecule and its partner, and the *intra*molecular variations in energy due to the conformational rearrangements of the interacting entities necessary to attain an optimal mutual fitting. Outstanding examples of such situations are the problems of the selective recognition of a drug molecule by a protein or a nucleic acid receptor and the selective complexation of cations by ionophores.

When dealing with relatively small molecules, it is possible to calculate both the inter- and the intramolecular energies using the *ab initio* SCF method which is known to provide results in fair agreement with available experimental data, both in conformational [1] and intermolecular interaction computations [2]. The utilization of this method is unfortunately untractable for *very* large systems of biological interest. Hence the incentive for the elaboration of simplified procedures aimed at a satisfactory reproduction of *ab initio* computations in model cases and applicable to larger systems.

We have recently elaborated such a procedure for the fast computation of intermolecular interactions [3], based on the use of additive components of the binding energy, the expressions of which are fitted in such a way as to satisfactorily reproduce the results of *ab initio* SCF supermolecule computations on small representative systems (these theoretical results being themselves in satisfactory agreement with the experimental ones) [3–5]. Distinctive features of the procedure are the construction of the large molecule out of subunits, the *ab initio* SCF wave function of which is preliminarily determined, the computation of the electrostatic and polarization contribution to the binding using a multicenter multipole expansion of the *ab initio* SCF electron density distribution of the entities in interaction and the computation of the repulsion contribution by means of a sum of bond-bond interactions.

It is the purpose of the present paper to show how the procedure can be successfully extended to allow the computation of variations in conformational energies *within* a given molecule.

In this paper we give the principles of the method and examples of application to systems containing carbon, hydrogen, nitrogen and ether-like oxygens. The accuracy of the computed variations in energy will be compared with the corresponding ones obtained by *ab initio* SCF calculations on test molecules, specifically: choline, ethanolamine, 1,2-dihydroxyethane, dimethoxymethane, methylethylether, histamine, and 1,3-diaminopropane.

2. Standpoint and methodology

According to the technique developed originally for the computation of the electrostatic potential of macromolecules [6, 7] and utilized subsequently in the study of intermolecular interactions [8], the large molecule is built out of constitutive molecular fragments separated by single bonds. We shall take advantage of this feature *to calculate the variation in the intramolecular energy upon a conformational change as the variable part of the sum of interactions between the*

fragments, using as the expression of the interfragment interaction energy a sum of five contributions:

$$E = E_{MTP} + E_{pol} + E_{rep} + E_{dl} + E_{tor}. \quad (1)$$

As in our *intermolecular* calculations [8], E_{MTP} denotes the electrostatic interaction energy between the multipolar expansions of the *ab initio* electron density of the fragments, E_{pol} is the corresponding polarization contribution, E_{rep} is the repulsion interaction, and E_{dl} is a dispersion-like contribution. These contributions are computed using the same principles as in our intermolecular procedure, with some necessary modifications indicated below. E_{tor} is a transferable torsional energy contribution, calibrated for elementary rotations around single bonds.

The *ab initio* SCF conformational energy calculations which serve us as tests are performed using Melius-Topiol pseudopotentials [9–10], together with the Least-Squares Fit basis set (LSQF) discussed in Ref. [11], with a dzeta exponent of 1.2 on the C—H hydrogens and 1.5 on the N—H ammonium hydrogens. This procedure was shown previously to provide results on conformational energy computations in close agreement with those of the corresponding full electron computations, themselves in agreement with experimental results [11]. The multipolar expansions of the charge distribution of the constituent fragments of the investigated molecule, required to compute the electrostatic and the polarization contributions, are accordingly also derived from *ab initio* SCF computations on the fragments using the same pseudopotentials and the same basis set. The input geometry of the fragments is the same as in the larger molecule. The overlap multipolar expansions of the fragments are simplified according to the procedure of Ref. [12], which splits every multipole located on the center of a non-bonded pair of atoms between the two centers closest to it, either atom or bond barycenter.

The analytical formulas utilized to compute the first four contributions of expression (1) are the same as in our intermolecular procedure [3] with the *same* set of constants, supplemented by the new ones necessary in the repulsion contribution (vide infra).

Aside from the inclusion of E_{tor}, two adaptations were necessary to render the procedure apt to calculate energy variations *within* a molecule. The first one stemmed from the necessity to dispose of an adequate representation of the multipolar expansion along the chemical bonds which are at the junction between two successive fragments. The problems raised by the existence of superposed multipolar expansions on the junctions were discussed in Ref. [13] and a solution was proposed through the use of localized molecular orbitals. We adopt here a solution in keeping with the use of the simplified multipolar expansion of Ref. [12].

The second adaptation concerns the short-range contribution, E_{rep}. This contribution involves a range of interatomic distances considerably shorter than those encountered in intermolecular computations, so that a more precise representation covering this situation is needed: this is solved here by introducing explicitly bond–lone pair and lone pair–lone pair contributions in E_{rep} aside from bond–bond contributions.

We thus present in what follows the three modifications required to calculate intramolecular energies within our former methodology:

a) Redistribution of the multipolar expansions along the junctional bonds $X-Y$ and consistent attribution of polarizabilities on the junctions.
b) Computation of E_{rep} as a sum of bond–bond, bond–lone pair, and lone pair–lone pair interactions.
c) Explicit incorporation of a torsional energy contribution, E_{tor}.

2.1. Multipoles and polarizabilities on the junctions

Let us consider the case of the two overlapping "fictive" bonds, A_1-H_1 and A_2-H_2, belonging to two successive fragments, and which, in the large molecule, are superposed to represent the single bond A_1-A_2 (see Fig. 1). B_1 and B_2 denote, respectively, the barycenters of bonds A_1-H_1 and A_2-H_2. A_3 is the middle of A_1-A_2. We have followed an approach closely similar to that adopted in [12] for relocating the multipoles pertaining to a non-bonded pair of atoms on the two centers closest to it.

The monopoles, dipoles and quadrupoles located on a given center F (H or B) on a "fictive" bond, will be divided between the two centers closest to it (thus, in the disposition of Fig. 1, H_1 onto A_2 and A_3, B_1 onto A_1 and A_3, H_2 onto A_1 and A_3 and B_2 onto A_3 and A_2) as follows: denoting A_i and A_j the two centers closest to F, one defines two weights λ_i and λ_j inversely proportional to the respective distances A_iF and A_jF as:

$$\lambda_i = (1/|A_iF|)/[(1/|A_iF|)+(1/|A_jF|)] = |A_jF|/[|A_iF|+|A_jF|] \tag{2}$$

and the analog λ_j. Then, each multipole M located at F will be split into two parts λ_iM and λ_jM, and each part (i) will be replaced by its multipole expansion around the corresponding center A_i. The choice (2) for the weights λ_i is by no means unique, but it insures that $\lambda_i = 1$ (and $\lambda_j = 0$) whenever $|A_iF| = 0$, a logical requirement to avoid relocating a fraction of M on other centers A_j when it happens to be located precisely at A_i.

According to the above conventions, the center A_i will receive from F:
a monopole:

$$q_i = \lambda_i q_F, \tag{3}$$

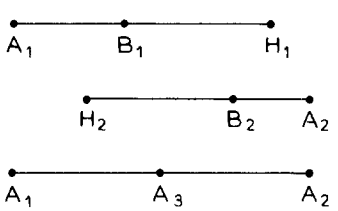


Fig. 1. The merging of bonds A_1H_1 and A_2H_2 into bond A_1A_2

a dipole μ_i generated by the monopole and the dipole of F, with the components:

$$(\mu_i)_x = \lambda_i[(\mu_F)_x + q_F(\overrightarrow{A_iF})_x] \tag{4}$$

and analogs for y and z, a quadrupole Q_i generated by the monopole, dipole and quadrupole of F, with components:

$$(Q_i)_{xy} = \lambda_i[(Q_F)_{xy} + q_F(\overrightarrow{A_iF})_x(\overrightarrow{A_iF})_y + (\mu_F)_x(\overrightarrow{A_iF})_y + (\mu_F)_y(\overrightarrow{A_iF})_x] \tag{5}$$

and analogs.

Corresponding terms will be located on A_j by using λ_j and the subscript j everywhere in (3) to (5).

These distributions are added to those preexisting on centers A_i and A_j and the final quadrupolar tensor is diagonalized to yield the values along its principal axes. Once this is done, the whole molecule is consistently represented by a set of multipoles located at all atoms and at some center on each chemical bond (whether it belongs to a fragment or joins two fragments).

As concerns the polarizabilities of the atoms A_i, A_j of a junction bond and of its center, they are computed as for any bond according to the previously defined partitioning of the appropriate bond polarizabilities [3].

At this stage, the molecule appears as a set of fragments *complemented with the centers (such as A_3) pertaining to the junction bonds*, each of these junction centers carrying a set of multipoles (M) and a polarizability α. For practical convenience, each of these junction centers is finally shared equally between the two attached fragments: specifically, labeling β and γ these two fragments, we replace the center A_3 $(M; \alpha)$ by two superimposed centers A_3^β $(M/2; \alpha/2)$ and A_3^γ $(M/2; \alpha/2)$, and we then consider A_3^β (resp. A_3^γ) as belonging to fragment β (resp. γ). The whole molecule now appears exactly as the union of individualized fragments on which one may distinguish "normal" centers (atoms and bond centers not belonging to any junction), junction atoms (atomic ends of junction bonds) and shared junction centers defined as above. This division dictates the rules for the practical computation of the variable part of E_{MTP} and E_{pol} in the intramolecular energy upon conformational changes:

(a) When two junction atoms are the endpoints of the *same* junction bond, their electrostatic interaction is suppressed in E_{MTP} and the electrostatic field created by each one onto the other atom is also suppressed in E_{pol}.
(b) When we consider a shared junction center, say A_k^β, belonging to fragment β and to the junction between fragments β and γ, the electrostatic interaction terms between A_k^β and *all* the centers A_l^γ $(l = 1, \ldots, N_\gamma)$ of fragment γ are suppressed, and similarly for the electric fields created by A_k^β on all the A_l^γ's and by all the A_l^γ's on A_k^β.

The purpose of these rules is to suppress interaction terms which are quite large but can be considered as approximately constant upon rotation around the junctions. *Mutatis mutandis*, this is analogous to the neglect of 1–2 and 1–3 interactions in the methods using classical atom–atom potentials [14–15].

Taking the preceding rules into account the MTP and polarization terms in Eq. (1) may be conveniently expressed by the following formulas:

$$E_{\mathrm{MTP}} = \sum_{\beta<\gamma}\sum E_{\mathrm{MTP}}(\beta, \gamma) \tag{6}$$

with

$$E_{\mathrm{MTP}}(\beta, \gamma) = \sum_{i=1}^{N_\beta}\sum_{j=1}^{N_\gamma} E_{\mathrm{MTP}}(A_i^\beta, A_j^\gamma) F_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta, A_j^\gamma) \tag{7}$$

and

$$E_{\mathrm{pol}} = \sum E_{\mathrm{pol}}^\beta \quad \text{with} \quad E_{\mathrm{pol}}^\beta = -\tfrac{1}{2}\sum_{i=1}^{N_\beta} \alpha(A_i^\beta)|\mathscr{E}(A_i^\beta)|^2 \tag{8}$$

and

$$\mathscr{E}(A_i^\beta) = \sum_{\gamma\neq\beta}\sum_{j=1}^{N_\gamma} \mathscr{E}(A_j^\gamma \to A_i^\beta) F_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta, A_j^\gamma) \tag{9}$$

where $E_{\mathrm{MTP}}(A_i^\beta, A_j^\gamma)$ denotes the electrostatic interaction between the set of multipoles on fragment β and the set of multipoles on fragment γ, $\mathscr{E}(A_j^\gamma \to A_i^\beta)$ denotes the electric field created at the point A_i^β by the set of multipoles at A_j^γ, and finally the factor:

$$F_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta, A_j^\gamma) = [1 - \Delta_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta, A_j^\gamma)][1 - \delta_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta)][1 - \delta_{\beta\gamma}^{\mathrm{junc}}(A_i^\gamma)] \tag{10}$$

accounts for the rules (a) and (b), by:

$$\Delta_{\beta\gamma}^{\mathrm{junc}}(A_i^\beta, A_j^\gamma) = 1 \tag{11}$$

if A_i^β and A_j^γ are the endpoints of the bond joining the fragments β and γ, and 0 otherwise

$$\delta_{\beta\gamma}^{\mathrm{junc}}(A_k^\eta) = 1 \tag{12}$$

if A_k^η is a junction center belonging to the bond joining the fragments β and γ, and 0 otherwise.

As a final remark to this section, let us point out that the decomposition into fragments used for the evaluation of conformational energy changes *could conceivably be different* from the one used for calculating the multipoles: since the final multipolar distribution has always the same form (set of multipoles on the atoms and on suitable centers pertaining to the chemical bonds), one might argue that such a distribution could be obtained at once for the whole molecule when it is not too large (or for fragments as large as can be computed), and then for computing conformational changes, subfragments could be defined *ad libitum* provided single bonds were chosen as junctions. Since, however, the large fragments would be computed in a given conformation, their multipoles could be characteristic of this conformation. Using them in subfragments for computing another conformation would necessarily introduce a bias. The simplest way to treat all conformations on the same footing is to compute the multipoles for the fragments which are used in the calculation of the interfragment interactions.

2.2. *The repulsion contribution*

For atoms carrying lone-pair electrons, each lone pair is explicitly introduced by a fictitious atom L, located at the barycenter of the appropriate hybrid. Thus, for an sp^3 oxygen, two fictitious atoms L_1 and L_2 are placed so as to complete a tetrahedron with the single bonds. Using classical formulas [16], the distance d_L from O to the barycenter of the hybrid

$$\chi_a = C_s 2s_a + C_p 2p_{\sigma a} \tag{13}$$

reduces to:

$$\langle \chi_a | z | \chi_a \rangle = 2C_s C_p \langle 2s_a | z | 2p_{\sigma a} \rangle = \frac{320}{\sqrt{3}} \frac{C_s C_p (\zeta_{sa})^{5/2} (\zeta_{pa})^{5/2}}{(\zeta_{sa} + \zeta_{pa})^6} a_H \tag{14}$$

in terms of the Slater exponents ζ_{sa}, ζ_{pa} and of the Bohr radius a_H.

Adopting $\zeta_{sa} = \zeta_{pa} = 2.24$ [17] and the standard coefficients $C_s = 0.50$, $C_p = 0.866$, we obtain $d_L = 0.295$ Å.

For pyridine-like nitrogens, the fictitious atom is located along the external bisector of the valence angle centered on N. A distance of 0.37 Å between N and L was computed with the sp^2 hybrid coefficients $C_s = 0.577$, $C_p = 0.816$ and a Slater exponent of 1.92.

The repulsion contribution of expression (1) is then written as a sum of interfragment contributions:

$$E_{\text{rep}} = \sum_{\beta \neq \gamma} \sum E_{\text{rep}}^{\beta\gamma} + \sum_{\beta J} \sum E_{\text{rep}}^{\beta J} (1 - \delta_{\beta J}) + \sum_{J \neq K} \sum E_{\text{rep}}^{JK} (1 - \delta_{JK}) \tag{15}$$

in which β and γ denote fragments (excluding the junctions), and J, K denote junctional bonds:

$\delta_{\beta J} = 1$ if J shares an atom with β
$\delta_{\beta J} = 0$ otherwise
$\delta_{JK} = 1$ if J and K share a common atom
$\delta_{JK} = 0$ otherwise.

The general expression for a repulsion term can be written:

$$\begin{aligned} E_{\text{rep}} = C_1 & \sum_{PQ}^{\text{bonds of } \beta} \sum_{RS}^{\text{bonds of } \gamma} \text{rep}\,(PQ, RS) \\ + C_2 \Bigg\{ & \sum_{PQ}^{\text{bonds of } \beta} \sum_{L\gamma}^{\text{lone pairs of } \gamma} \text{rep}\,(PQ, X_\gamma L_\gamma) \\ + & \sum_{L_\beta}^{\text{lone pairs of } \beta} \sum_{RS}^{\text{bonds of } \gamma} \text{rep}\,(X_\beta L_\beta, RS) \\ + & \sum_{L_{\beta(\beta>\gamma)}}^{\text{lone pairs of } \beta} \sum_{L_\gamma}^{\text{lone pairs of } \gamma} \text{rep}\,(X_\beta L_\beta, X_\gamma L_\gamma) \Bigg\} \end{aligned} \tag{16}$$

where X_β, X_γ designate the atoms of β and γ which carry lone pairs and L_β, L_γ are the corresponding lone pair centers. The formulas given and justified in [3] for the *PQ*, *RS* bond–bond repulsion terms are generalized in a straightforward way to include lone pairs: each time a lone pair is involved, *PQ* or *RS* is to be replaced by $X_\beta L_\beta$ or $X_\gamma L_\gamma$ in the expression below to be inserted in (15).

$$\text{rep}\,(PQ, RS) = [M_{PR}\, e^{-\alpha d_{PR}} + M_{PS}\, e^{-\alpha d_{PS}} + M_{QR}\, e^{-\alpha d_{QR}} + M_{QS}\, e^{-\alpha d_{QS}}]^2 \qquad (17)$$

with

$$d_{PR} = r_{PR}/4(W_P W_R)^{1/2} \qquad (18)$$

$$M_{RP}^2 = \frac{K_{PR}}{\nu_P \nu_R}\left(1 - \frac{Q_P}{N_P^{\text{val}}}\right)\left(1 - \frac{Q_R}{N_R^{\text{val}}}\right) \qquad (19)$$

and analogs, where W_P denotes the effective radius of atom *P*, N_P^{val} the number of valence electrons, ν_P the total number of bonds and lone pairs emanating from *P*, Q_P the monopole of *P* in the multipole expansion. The K_{PR} values between two atoms of atomic numbers Z_P and Z_R are the products of atomic $K(Z_P)$ and $K(Z_R)$ values derived as in Ref. [18]. They are the same as those utilized for the dispersion-like contribution. When *P* is a lone pair fictitious atom, the corresponding values of $K(Z_P)$ and ν_P are set equal to unity, as well as that of the term $(1 - Q_P/N_P^{\text{val}})$.

The values of α and C_1 are consistently equal to 12.35 and 30 322.0 respectively, as in the intermolecular procedure. The values of the effective radii, *W*, of atoms which carry no lone pairs are the same as in the intermolecular procedure.

The representation of a lone pair by a "pseudo-bond" $X_A L_A$ necessitates the definition of a proper value of the effective radius of L_A, and also of center X_A. Moreover, a value of C_2 distinct from that of C_1 is required. We have already observed that the distances involved in intramolecular interactions, which are imposed by the molecular structure, are much shorter than in intermolecular interactions. The partial projection of heteroatom *P* into a center located 0.3–0.4 Å away from it may result into further shortened interatomic distances involving this atom and neighboring sites. When $C = 30\,322.0$, the evolution of E_{rep} in intermolecular interaction studies matches properly that of its *ab initio* supermolecule counterpart ΔE_{rep} in a wide range of interatomic distances encompassing the Van der Waals minimum and beyond [3, and unpublished data], but its evolution becomes much steeper than that of E_{rep} for smaller distances. The adoption of a smaller value for C_2 than for C_1 corrects for this behavior.

The values of C_2, as well as those of W_{XA} and W_{LA} for sp^3 oxygens, were determined so as to retrieve the main characteristics of the conformational behavior of the choline molecule, obtained in *ab initio* calculations (see below). This yields $C_2 = 16\,068.0$, (for distances in Å and energies in kcal/mole), $W_{XO} = 1.285$ Å and $W_{LO} = 0.970$ Å. Recalling that the original values of the effective radii on N and O are 1.70 Å and 1.50 Å respectively, we have adopted for the

values of W_{XN} and W_{LN} on the pyridine-like nitrogen and its associated lone pair fictitious atom the values 1.42 Å and 1.00 Å respectively.

2.3. The torsional energy contribution E_{tor}

The conformational changes investigated in the present study result from rotations around C—C and C—O ether bonds. We have incorporated for each such bond a three-fold torsional energy contribution:

$$E_{tor} = \tfrac{1}{2} V_0 (1 + \cos 3\varphi) \tag{12}$$

with $V_0 = 2.3$ and 1.6 kcal/mole respectively, calibrated so as to reproduce the values of the staggered to eclipsed torsional barrier in ethane and dimethylether, namely 3.0 and 2.5 kcal/mole, calculated in our *ab initio* pseudopotential procedure. Note that in both molecules part of the torsional barrier is inherently contained in the other contributions to the binding, essentially in E_{rep}. The values of V_0 are assumed to be transferable for treating the torsional contributions around all C—C and C—O ether bonds (C being a saturated carbon). This assumption will be justified by the practical results.

As a matter of convenience, we shall use in the text the abbreviation SIBFA (Sum of Interactions Between Fragments computed *Ab initio*) to designate briefly the procedure.

3. Conformational calculations

We have adopted, for all the molecules studied below, standard bond lengths and valence angles, and assumed rigid rotation. The same bond lengths and valence angles are used on the constitutive fragments as on the molecule. The terminal methyl and ammonium hydrogens, when present, are fixed in a staggered conformation. The torsion angles are varied by 20° increments.

3.1. Choline

As mentioned above, choline is the molecule which served for fitting the values of C_2, W_{XO} and W_{LO} using as a guideline *ab initio* results. For this purpose, our SCF calculations on this molecule concentrate on the rotations around the carbon–carbon bond (see figure 2) which are decisive in commanding the overall character of the conformation in molecules [19, 20, 21, 22] of the choline family. For this investigation we have fixed the terminal hydroxyl in a *trans* configuration and kept the three methyl groups of the onium head staggered with respect to the C—N bonds.

The SCF results (full curve of Fig. 2) indicate that the intrinsically preferred conformation is a *gauche* conformation along the C—C bond ($\tau = 80°$), stabilized by an attractive interaction between the hydroxyl oxygen and the positive charge

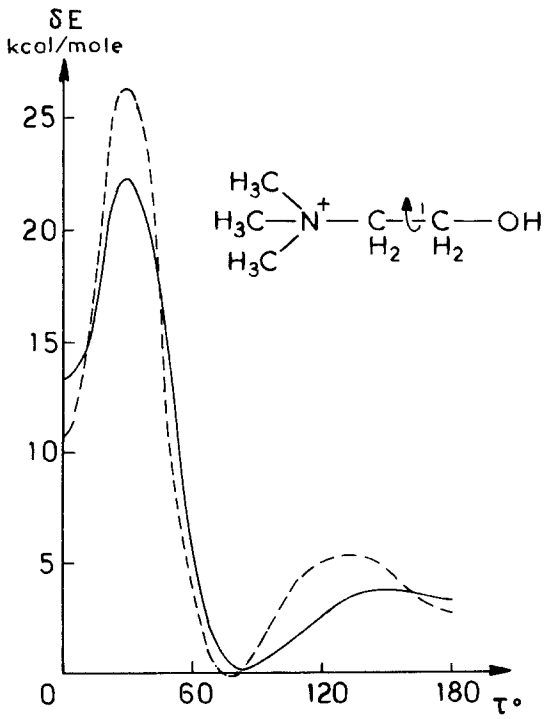


Fig. 2. Choline. Evolutions of the *ab initio* and the SIBFA energy curves for torsions along the bond NC—CO. Energy of the most stable conformer taken as energy zero

of the trimethylammonium moiety (see also Refs. [20, 23]). The *trans* conformation is 3.2 kcal/mole above the energy minimum, whereas the eclipsed conformation ($\tau = 0°$) is 13 kcal/mole above it. A steep energy peak is encountered for $\tau = 30°$, owing to a close contact between one trimethylammonium hydrogen and the oxygen.

In the SIBFA computations choline was built out of the three constitutive fragments tetramethylammonium, methane and water appropriately placed with the same bond lengths and valence angles as in the global molecule. Based on the SCF results, we have attempted to obtain a *gauche* to *trans* energy difference, δ_1, close to 3 kcal/mole, while maintaining a *gauche* to *eclipsed* energy difference, δ_2, as close as possible to 13 kcal/mole. The best compromise was obtained for $C_2 = 16068.0$, $W_{OX} = 1.285$, and $W_{OL} = 0.970$. The minimum energy conformation is at $\tau = 75°$, very close to the SCF value. The evolution of the intramolecular energy as a function of τ is represented in Fig. 2 (dotted line), showing that the qualitative characteristics of the corresponding *ab initio* curve are preserved in a very satisfactory manner. In all subsequent calculations the C_2, W_{XO} and W_{LO} values were conserved.

3.2. Ethanolamine

As in choline, the terminal hydroxyl hydrogen of ethanolamine was fixed in a *trans* configuration. The fragments used in building the molecule are monomethyl-

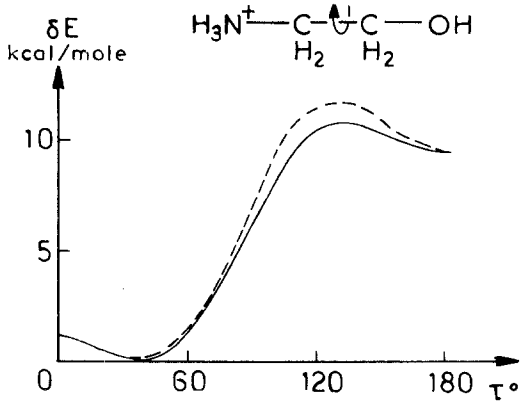


Fig. 3. Ethanolamine. Evolutions of the *ab initio* and the SIBFA energy curves for torsions along the bond NC—CO. Energy of the most stable conformer taken as energy zero

ammonium, methane and water with the bond lengths and valence angles of ethanolamine. The evolutions of the *ab initio* and SIBFA energies as a function of the C—C torsion are very closely related as shown in Fig. 3. The intrinsically preferred conformation corresponds to $\tau = 40°$ for both procedures, a structure notably more folded than the preferred conformation of choline. The *gauche* to *trans* energy separation is much larger than in choline (9.3 kcal/mole in both procedures) while the *gauche* to eclipsed barrier is considerably smaller than in choline (1.3 kcal/mole in both procedures). The energy maximum is at $\tau = 130°$ for both procedures ($\delta_{SCF} = 10.7$ kcal/mole, $\delta_{SIBFA} = 11.5$ kcal/mole).

3.3. 1,2 dihydroxyethane

In this molecule, the hydrogen atom of one of the hydroxyl groups is fixed in a *trans* conformation, whereas the torsional angle CC—OH defining the conformation of the other hydroxyl group is fixed, successively at $\varphi = 180°$, 120°, 60°, 0° and torsional variations along the C—C bond are generated. For the first two values of φ, the conformational behavior is expected to be governed by the mutual repulsions between the two hydroxyl oxygens which, in the SIBFA procedure, exert themselves through the electrostatic contribution E_{MTP} and the short-range repulsive contribution E_{rep} involving both chemical bonds and lone pairs. This molecule thus provides a rather stringent test, as to the extent to which the method can numerically translate such repulsions when confronted to the *ab initio* results. On the other hand for $\varphi = 60°$ and $\varphi = 0°$, the conformational behavior is governed by an attractive interaction between the oxygen of the "*trans*" hydroxyl group and the hydrogen of the other, which is expected to favor the prevalence of a *gauche* conformation. It is, nevertheless, considerably less stabilizing than that involving an oxygen atom and a cationic moiety, as in the related molecule of ethanolamine. Hence another possibility of testing the ability of SIBFA to account for such cases involving a reduced attraction between fragments.

In the present case, the molecule was built out of methane and water fragments in the appropriate geometries. The compared evolutions of the *ab initio* and SIBFA curves are shown in Figs. 4(a–d), as a function of the torsional angle $\tau = OC—CO$.

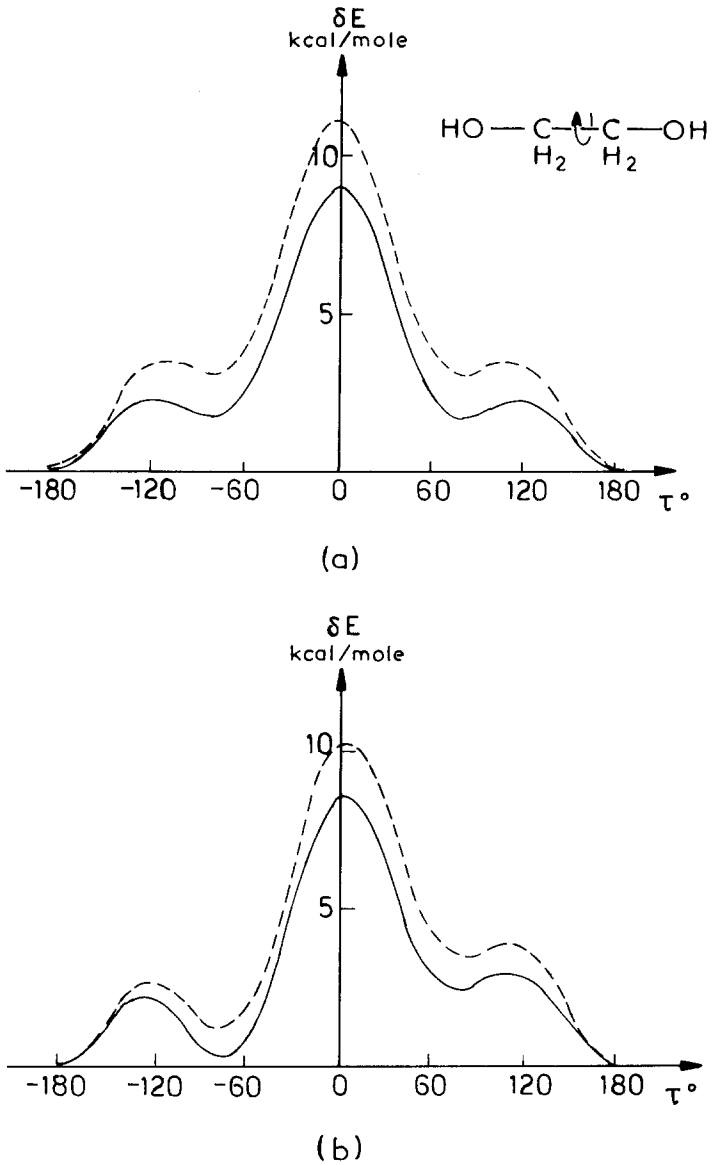
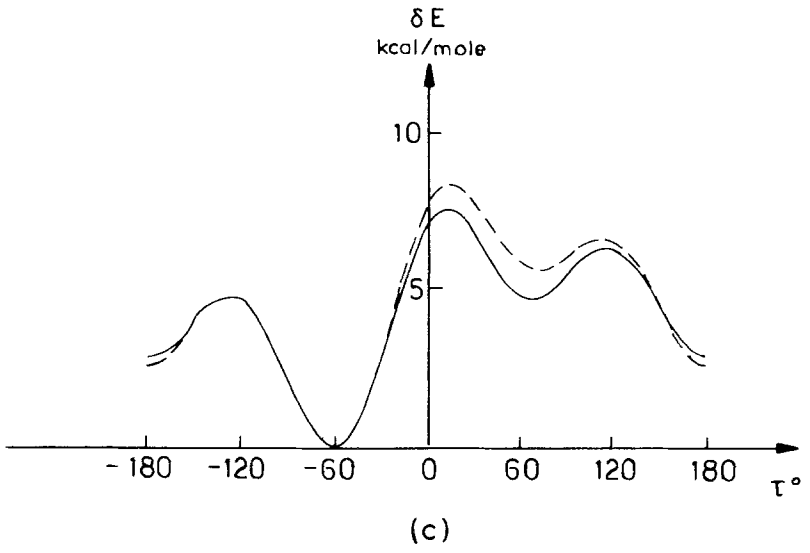


Fig. 4. 1,2 dihydroxyethane. Evolutions of the *ab initio* and the SIBFA energy curves for torsions along the bond OC—CO. Energy of the most stable conformer taken as energy zero. One of the torsion angles CC—OH is fixed at the value 180°, the other (φ) is prefixed in four different values. a) $\varphi = 180°$, b) $\varphi = 120°$, c) $\varphi = 60°$, d) $\varphi = 0°$

3.3.1. $\varphi = 180°$

The eclipsed conformation ($\tau = 0°$) is separated from the *trans*, minimum energy conformation, by a high energy difference ($\delta_{SCF} = 9.1$ kcal/mole, $\delta_{SIBFA} = 11.4$ kcal/mole). Two local minima (at $\tau = \pm 80°$) and two local maxima (at



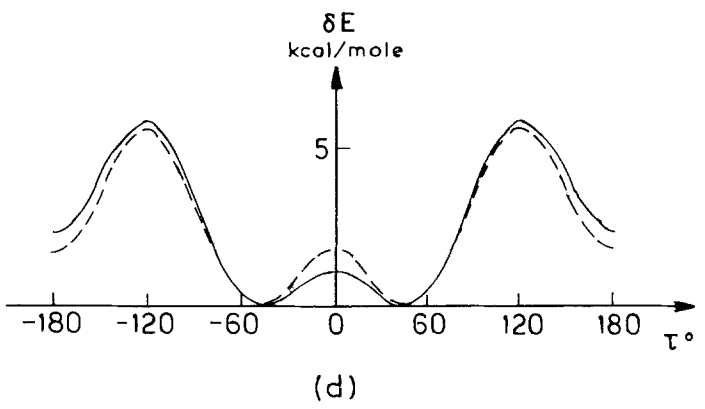


Fig. 4 (*cont.*)

$\tau = \pm 120°$) appear in both *ab initio* and SIBFA curves with energies in very reasonable agreement.

3.3.2. $\varphi = 120°$

The values of the *trans* to eclipsed energy difference are reduced with respect to the $\varphi = 180°$ situation ($\delta_{SCF} = 8.6$ kcal/mole; $\delta_{SIBFA} = 10.7$ kcal/mole). The global energy maximum is shifted to $\tau = 5°$, in both methods. The SCF and SIBFA curves are dissymetrical with respect to $\tau = 0°$ so that the two local minima and the two local maxima are no longer degenerate. In fact, both the left side local minimum ($\tau = -80°$) *and* the left side local maximum ($\delta = -120°$) are more stable energetically than the right side local minimum ($\tau = 80°$). On the other hand, the energy difference separating the right side local maximum and the right side local

minimum is twice as small as the corresponding difference between left side minimum and maximum. All the salient features of the *ab initio* curve are thus adequately reproduced by its SIBFA counterpart.

3.3.3. $\varphi = 60°$

The intrinsically preferred conformation is *gauche* ($\tau = -60°$), stabilized by an H-bonding interaction between the *trans* hydroxyl oxygen and the hydrogen of the other hydroxyl group ($d_{O-H} = 2.15$ Å). The *trans* conformation of dihydroxyethane is now a local minimum ($\delta_{SCF} = 2.8$ kcal/mole; $\delta_{SIBFA} = 2.5$ kcal/mole). Another local minimum occurs for $\tau = 70°$ ($\delta_{SCF} = 4.7$ kcal/mole; $\delta_{SIBFA} = 5.8$ kcal/mole). The global maximum is shifted to $\tau = 20°$ ($\delta_{SCF} = 7.5$ kcal/mole; $\delta_{SIBFA} = 8.2$ kcal/mole). Two local maxima are found at $\tau = 120°$ and $\tau = -120°$.

3.3.4. $\varphi = 0°$

The intrinsically preferred conformation, again *gauche*, occurs for a smaller absolute value of τ ($\pm 45°$). The *trans* conformation corresponds to a local minimum ($\delta_{SCF} = 2.3$ kcal/mole; $\delta_{SIBFA} = 1.6$ kcal/mole). The eclipsed conformation now has a stability comparable to the *trans* conformation ($\delta_{SCF} = 1.2$ kcal/mole; $\delta_{SIBFA} = 2.1$ kcal/mole), whereas the global maximum occurs for $\tau = 120°$ ($\delta_{SCF} = 5.9$ kcal/mole; $\delta_{SIBFA} = 5.6$ kcal/mole).

3.4. Dimethoxymethane

The conformation of dimethoxymethane (Fig. 5) has been investigated before both by semi-empirical [22, 25] and *ab initio* SCF [24, 25, 26] computations. When standard valence angles are adopted, the molecule manifests an intrinsic preference for the *gauche–gauche* conformation, attributed to the "anomeric effect" [see, e.g. Ref. (24)]: in the *trans–trans* conformation, each lone pair on one oxygen, eclipses a corresponding lone pair on the other oxygen, which maximizes their mutual repulsions.

In procedures using empirical potentials this kind of effect requires generally the introduction of special terms (see for instance [27, 28]). Thus this molecule

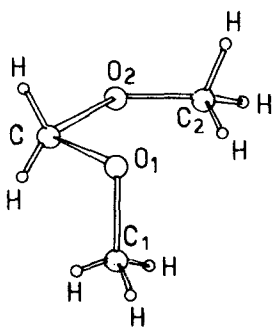


Fig. 5. Representation of dimethoxymethane in its *gauche–gauche* conformation

Table 1. Values of the *ab initio* and SIBFA energy differences δ and MTP and repulsion components between the *gg*, *gt* and *tt* conformations of dimethoxymethane. Energies in kcal/mole

	δ_{SCF}	δ	δ_{MTP}	$\delta_{mono-mono}$	δ_{lp-lp}	$\delta_{lp-bond}$	$\delta_{bond-bond}$
gg	0.0	0.0	0.0	0.0	0.0	2.1	0.9
gt	2.9	2.5	2.6	1.2	0.3	1.1	0.0
tt	6.9	5.8	5.7	2.5	0.5	0.0	0.1

provides an interesting test for our procedure and we will more particularly examine to what extent can the existence of the anomeric effect be accounted for and what are the respective role of E_{MTP} and E_{rep}. For that purpose, we investigate the three model conformations, *gauche–gauche* (*gg*), *gauche–trans* (*gt*), and *trans–trans* (*tt*) to compare the *ab initio* and the SIBFA conformational results. In the SIBFA computations, the molecule is built out of the constitutive fragments methane and water in the appropriate geometries. A *gauche* conformation of one methyl group (say C_1) corresponds to a dihedral angle $C_1—O_1—C_2—O_2$ equal to 60° and a *trans* conformation to 180°. For these values, our three-fold torsional energy contribution is zero. The results are reported in Table 1, together with the values of the energy differences between the corresponding values of E_{MTP} and E_{rep}. It is seen that the stability order $gg > gt > tt$ is correctly accounted for by the SIBFA procedure, the computed values of the conformational energy differences, δ, being themselves close to the corresponding *ab initio* values. The values of δ are also close to the corresponding values of δ_{MTP}, a feature which indicates that E_{MTP} plays a decisive role in dictating the intrinsic conformational preference for a *gauche–gauche* conformation. It is also worth noting that the sole monopole–monopole component of E_{MTP} is insufficient to account for such a preference, (compare $\delta_{mono-mono}$ and δ_{MTP} in Table 1).

As concerns the behavior of E_{rep} in the three conformations, Table 1 shows that the lone pair–lone pair component of E_{rep} is more unfavorable for *tt* than for *gg*. δ_{lp-lp}, which amounts to 0.5 kcal/mole, is, however, overcompensated by the lone-pair–bond component, which disfavors *gg* over *tt* by 2 kcal/mole. The bond–bond repulsion term has a more complicated behavior, as it is the result of, on the one hand, the mutual repulsions between the $—OCH_3$ bonds, disfavouring *gg*, and on the other hand, the repulsions of these bonds with the central methylene group, which disfavors *tt*. As a result the bond–bond repulsion term is the smallest in the *gt* conformation. It thus appears that the origin of the anomeric effect is not the short-range part of the lone pair–lone pair repulsion but rather the electrostatic repulsion of the electron densities located on the two oxygens and their associated bonds. This analysis indicates the potential usefulness of our computational procedure in delineating the individual energy contributions of the intramolecular energy which dictate the intrinsic conformational preferences of a molecule. It also underlines the necessity of using an extended multipolar expansion rather than monopoles only.

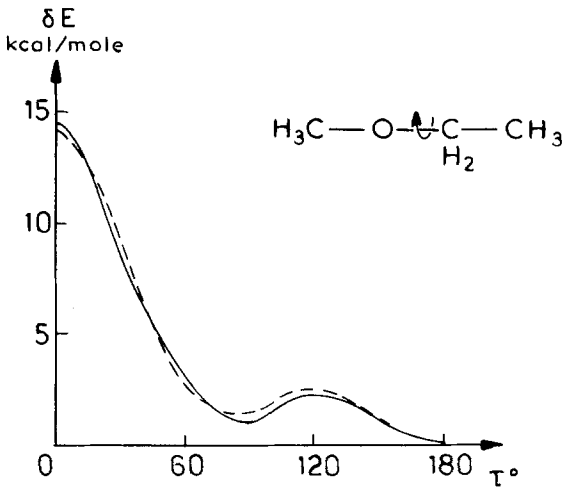


Fig. 6. Methylethylether. Evolutions of the *ab initio* and the SIBFA energy curves for torsions along the bond CO—CC. Energy of the most stable conformer taken as energy zero

3.5. Methylethylether

The conformation of methylethylether has been studied by previous investigators using *ab initio* SCF [25, 29], PCILO [25], as well as molecular mechanics [27]. We have concentrated on torsions along the O—C bond, indicated in Fig. 6. In the SIBFA computations, the molecule is built out of methane and water in the appropriate geometry. The *trans* conformation corresponds to the conformational minimum. The eclipsed conformation, (global energy maximum) is separated from it by a high energy difference in both curves. A small local minimum is found for $\tau = 80°$, a result in agreement with the conclusion of a gas-phase electron diffraction study [38], indicating the existence of a local minimum corresponding to a *gauche* conformation for a dihedral angle of 84°, and an energy at 1.23 kcal/mole above the *trans* minimum. A local maximum similar in the two curves occurs at $\tau = 120°$.

3.6. Diprotonated 1,3 diamino propane

1,3-diamino propane is a prototype for the diamine and polyamine molecules, known to interact with a high affinity with DNA [31, 32]. Its conformational behavior is clearly governed by the electrostatic repulsions between the two charged cationic ends, and it is important to probe to what an extent the peculiar characters of the SCF conformational curves are retrieved for this type of interactions. We have investigated variations of the dihedral angle $\tau_1 = N_1C_1—CC_2$, for two typical values $\tau_2 = 180°$ of the other dihedral angle τ_2. In the SIBFA computations, the molecule was built out of the constitutive fragments methane and monomethylammonium in the appropriate geometries. The evolution of the *ab initio* and the SIBFA curves are shown in Fig. 7 where the two lower curves pertain to $\tau_2 = 180°$ and the two upper ones to $\tau_2 = 120°$.

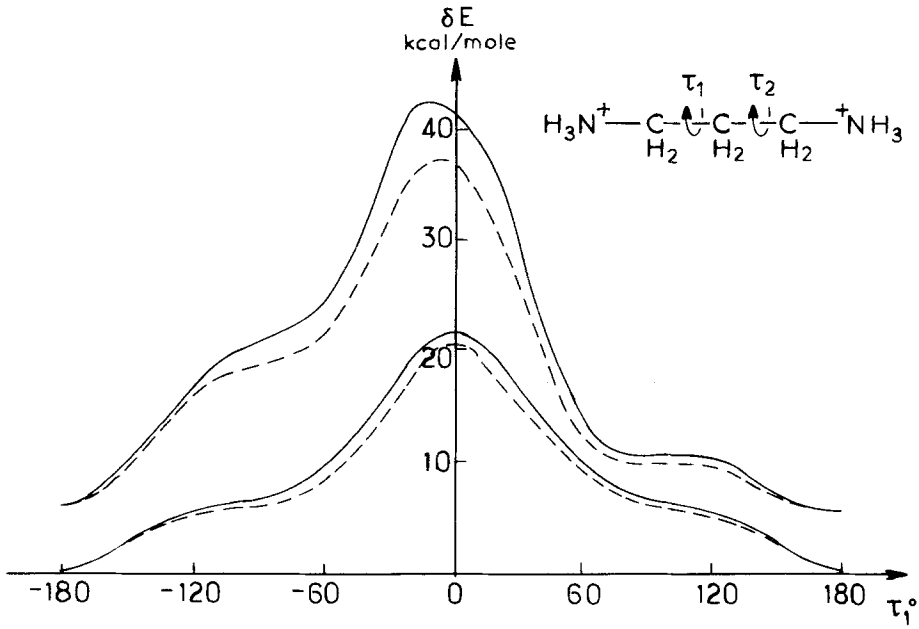


Fig. 7. 1,3 diaminopropane. Evolution of the *ab initio* and the SIBFA energy curves for torsions along one of the bonds NC—CC (τ_1). The other torsional angle CC—CN (τ_2) is fixed at $\tau_2 = 180°$ (two lower curves) and $\tau_2 = 120°$ (two upper curves). Energy of the most stable conformer taken as energy zero

$\tau_2 = 180°$

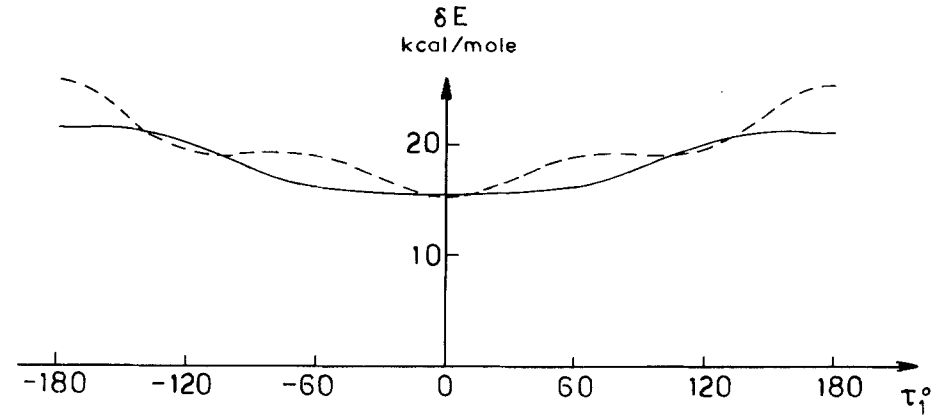
The eclipsed conformation is separated by a high energy difference from the *trans* minimum ($\delta_{SCF} = 21.9$ kcal/mole; $\delta_{SIBFA} = 20.9$ kcal/mole). The existence of an inflexion point is noted in the vicinity of $\tau_1 = \pm 100°$.

$\tau_2 = 120°$

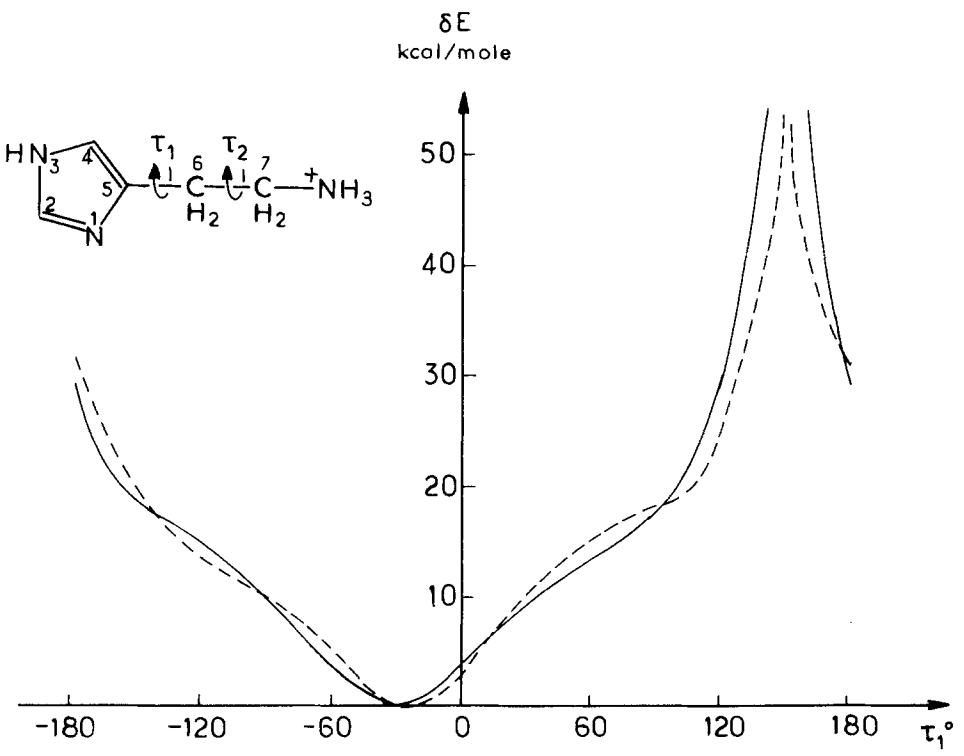
The eclipsed conformation is separated by a notably increased energy difference from the *trans* minimum ($\delta_{SCF} = 41.7$ kcal/mole; $\delta_{SIBFA} = 37.3$ kcal/mole). The global maximum is shifted to $\tau_1 = -10°$. A marked dissymmetry is noted with respect to $\tau_1 = 0°$: a plateau region exists in the right-hand side ($80° \leq \tau_1 \leq 120°$; $\delta_{SCF} = 10.4$ kcal/mole; $\delta_{SIBFA} = 9.4$ kcal/mole) while two inflexion points are present in the left-hand side, for $\tau_1 = -60°$ and $\tau_2 = -120°$. All the singularities present in the two *ab initio* conformational curves are fully reproduced in their SIBFA counterparts.

3.7. Histamine

Histamine is a molecule of outstanding pharmacological interest. Its conformation as a function of the angle τ_1 and τ_2 (Fig. 8b) was previously investigated in this laboratory, using the *ab initio* SCF procedure with a STO3G basis set, in conjunction with the PCILO procedure [33]. For consistency with our other SCF computations, we have recalculated the conformational energy of monoprotonated histamine in the framework of the pseudopotential SCF procedure described before.



(a)



(b)

Fig. 8. Histamine. Evolutions of the *ab initio* and the SIBFA energy curves for torsions along the bond $N_1C_5—C_6C_7$. The torsional angle $C_5C_6—C_7N$ (τ_2) is fixed at $\tau_2 = 180°$ (Fig. 8a) and $\tau_2 = 60°$ (Fig. 8b). Energy of the most stable conformer taken as energy zero

As in Ref. [33], the value of τ_2 was pre-fixed at $\tau_2 = 60°$ or $\tau_2 = 180°$, and the evolution of the energy as a function of τ_1 was studied. In the SIBFA computations, the molecule is built out of the constitutive fragments imidazole, methane and monomethylammonium in the appropriate geometries. The *ab initio* and the SIBFA curves are given in Figs. 8a ($\tau_2 = 180°$) and 8b ($\tau_2 = 60°$).

As in Ref. [37], the most stable *ab initio* conformation found occurs for $\tau_1 = -30°$ and $\tau_2 = 60°$, a folded conformation stabilized by a close approach between the pyridine-like nitrogen of the imidazole and the terminal ammonium group. The preferential stabilization of this conformation could not be accounted for by EHT computations [34]. In the SIBFA computations, we have tentatively retained, for the torsions along $C_5—C_6$, our three-fold ethane-like rotational barrier, the maximum of the barrier occurring when $C_4—C_5$ is eclipsed by either of the three bonds $C_6—C_7$ or C_6H.

The characteristics of both *ab initio* curves are closely reproduced by their SIBFA counterparts. The SIBFA minimum for $\tau_2 = 60°$, is at $\tau_1 = -25°$, very close to the *ab initio* minimum. The $\tau_2 = 60°$ curves present a steep maximum for $\tau_1 = 150°$, owing to a steric interaction between one ammonium hydrogen and the imidazole $C_4—H$ hydrogen. The $\tau_2 = 180°$ curves display, on the other hand, a shallow behavior in the whole range of τ_1 values. The energy difference separating the minimum at $\tau_1 = 60°$ from the minimum at $\tau_1 = 180°$ is the same (15.5 kcal/mole) in the two methods.

4. Concluding remarks

This study shows that it is possible to reproduce with a very encouraging accuracy, the results of *ab initio* SCF conformational energy computations on model compounds by a procedure calculating intramolecular interactions as a sum of interactions between fragments. Two important conditions appear necessary in order to ensure such an agreement. One is an appropriate representation of the electrostatic contribution by a multicenter multipolar expansion of the electron distribution in the fragments. This was exemplified in the analysis of the case of dimethoxymethane. The other one is the computation of E_{rep} as a sum of bond–bond, bond–lone pair and lone pair–lone pair interactions, in order to ensure a proper directionality of this contribution, especially when the distances between interacting centers on the subunits are small. An interesting feature of the procedure is the possibility to explicate the interplay of the energy contributions, in the differential stabilization of differing competing conformations of a molecule. With respect to our intermolecular procedure [8], only a minimal calibration was necessary, due to the introduction of C_2 and W_X and W_{LX}.

The investment in computer time, required to perform the *ab initio* computations on the building blocks, is not prohibitive, on account of the rather reduced size of the fragments. The computer time is roughly proportional to $4N^2$, N being the total number of atoms of the system. This increase as a function of N is appreciably slower than that occurring in quantum-mechanical semi-empirical

procedures, for which it is proportional to N^3–N^4. Furthermore no integral storage problems arise.

The present procedure hopefully will allow us to treat within a unified approach, and in an automatic fashion, the interplay of both the *inter-* and *intra*molecular interactions which commands the binding specificities of large molecules of biological interest.

Acknowledgment. The programmation of the SIBFA procedure as well as the corresponding calculations have been performed on a VAX/750 Computer, thanks to the support of the National Foundation for Cancer Research (U.S.A.) to which the authors express their gratitude.

References

1. Payne, P. W., Allen, L. C.: Applications of electronic structure theory, p. 29, Schaeffer, III, H., Ed. New York: Plenum Press, 1977
2. Intermolecular Interactions: From diatomics to biopolymers, Pullman, B., Ed. New York: Wiley, 1978
3. Gresh, N., Claverie, P., Pullman, A.: Int. J. Quantum Chem. Symp. **13**, 243 (1979)
4. Gresh, N., Pullman, B.: Theoret. Chim. Acta (Berl.) **52**, 67 (1979)
5. Gresh, N.: Biochim. Biophys. Acta **597**, 345 (1980)
6. Pullman, A., Zakrzewska, K., Perahia, D.: Int. J. Quantum Chem. **16**, 395 (1979)
7. Pullman, A., Pullman, B.: Quart. Rev. Biophys. **14**, 289 (1981)
8. Gresh, N., Etchebest, C., De La Luz Rojas, O., Pullman, A.: Int. J. Quantum Chem., Quantum Biol. Symp. **8**, 109 (1981)
9. Melius, C., Goddard III, W.: Phys. Rev. **A10**, 1528 (1974)
10. Topiol, S., Moskowitz, J., Melius, C.: J. Chem. Phys. **68**, 2364 (1978)
11. Gresh, N., Pullman, A.: Theoret. Chim. Acta (Berl.) **49**, 283 (1978)
12. Vigné-Maeder, F., Claverie, P.: J. Chem. Phys., to be submitted
13. Etchebest, C., Lavery, R., Pullman, A.: Theoret. Chim. Acta (Berl.) **62**, 17 (1982)
14. Hopfinger, A.: Conformational properties of macromolecules. New York: Wiley, 1973
15. Simon, Z.: Quantum biochemistry and specific interactions. Kent: Abacus Press, 1976
16. Hamilton, W.: J. Chem. Phys. **26**, 1016 (1956)
17. Clementi, E., Raimondi, D.: J. Chem. Phys. **38**, 2686 (1963)
18. Caillet, J., Claverie, P.: Acta Cryst. Section **A31**, 448 (1975)
19. Genson, D., Christoffersen, R.: J. Am. Chem. Soc. **95**, 362 (1973)
20. Pullman, A., Port, G.: Theoret. Chim. Acta (Berl.) **32**, 77 (1973)
21. Kier, L.: Mol. Pharmacol. **3**, 487 (1967)
22. Beveridge, D., Radna, R.: J. Am. Chem. Soc. **93**, 3759 (1971)
23. Pullman, B., Courrière, P.: Mol. Pharmacol. **8**, 612 (1972)
24. Gorenstein, D., Kar, D.: J. Am. Chem. Soc. **99**, 672 (1977)
25. Bendl, J., Pretsch, E.: J. Comput. Chem. **3**, 580 (1982)
26. Jeffrey, G., Pople, J., Binkley, J., Vishveshwara, S.: J. Am. Chem. Soc. **100**, 371 (1977)
27. Govil, G.: Biopolymers **15**, 2303 (1976)
28. Lipari, G., Tosi, C.: Theoret. Chim. Acta (Berl.) **50**, 169 (1978)
29. Burkert, U.: J. Comput. Chem. **1**, 285 (1980)
30. Oyanagi, K., Kuchitsu, K.: Bull. Soc. Chem. Japan, **51**, 2237 (1978)
31. Suwalsky, M., Traub, W., Schmueli, D., Subirana, J.: J. Mol. Biol. **42**, 363 (1969)
32. Minyat, E., Ivanov, V., Kritzyn, A., Minchenkova, L., Schyolkina, A.: J. Mol. Biol. **128**, 397 (1978)
33. Pullman, B., Port, G.: Mol. Pharmacol. **10**, 360 (1974)
34. Kier, L.: J. Med. Chem. **11**, 441 (1968)

Received February 27, 1984