

Numerical computation of the electrostatic interaction energy between methanol and the dyad water-imidazole

Dominique Dehareng, Georges Dive, Josette Lamotte-Brasseur,
and Jean-Marie Ghuysen

Service de Microbiologie, Institut de Chimie, B6, Université de Liège, B-4000 Sart Tilman, Belgium

(Received October 13, 1988; revised February 20/Accepted February 28, 1989)

Summary. The electrostatic interaction energy between methanol and the dyad water-imidazole has been computed numerically at three levels of approximation from 3D grids of the charge density of one partner and the electrostatic potential of the other. The minimum positions and energy values thus obtained compare well with those calculated analytically. The numerical procedure is especially interesting for the prediction of the stable conformers.

Key words: Electrostatic interaction energy

1. Introduction

Understanding at the most fundamental level how proteins and their ligands interact is one of the foundation stones of the developing modern biosciences. Such an interaction is defined by the energy it generates and this interaction energy, E_{int} , can be evaluated at different levels of approximation depending on the size of the system and the distance between the interacting partners. At large distances, the most important component of E_{int} is the electrostatic interaction energy E_{int}^{ES} [1–3]. This paper shows how and under which conditions, the E_{int}^{ES} value for the non-covalent interaction between the dyad water-imidazole (which mimicks a portion of a large biomolecule such as an enzyme active site) and methanol (representing a small organic ligand) can be computed accurately by numerical procedures which integrate the electrostatic potential map of one partner and the charge density map of the other [4, 5].

2. Expression of the electrostatic interaction energy E_{int}^{ES}

Given two interacting molecules A and B, at point R defined by its position vector \mathbf{r} , A develops a force field that combines the electrostatic potentials generated by the electrons, $V_A^{\text{el}}(\mathbf{r})$, and the nuclei, $V_A^{\text{N}}(\mathbf{r})$. The total electrostatic potential of A is

$$V_A(\mathbf{r}) = V_A^{\text{el}}(\mathbf{r}) + V_A^{\text{N}}(\mathbf{r}) = V_A^{\text{el}}(\mathbf{r}) + \sum_{\alpha} Z_{\alpha}/|\mathbf{r} - \mathbf{r}_{\alpha}| \quad (1)$$

where Z_{α} and \mathbf{r}_{α} are the electric charge and vector position of nucleus α .

In turn, B can be characterized by its charge density, $\gamma_B(\mathbf{r})$, which is the sum of the electronic and nuclear terms:

$$\gamma_B(\mathbf{r}) = \rho_B^{\text{el}}(\mathbf{r}) + \sum_{\beta} Z_{\beta} \delta(\mathbf{r} - \mathbf{r}_{\beta}) \quad (2)$$

where ρ_B^{el} is the electronic charge density and $\delta(\mathbf{r} - \mathbf{r}_{\beta})$ is the Dirac delta function.

The electrostatic interaction energy between A and B is given by

$$\begin{aligned} E_{\text{int}}^{ES} = & \int d\mathbf{r} V_A^{\text{el}}(\mathbf{r}) \rho_B^{\text{el}}(\mathbf{r}) + \sum_{\alpha} Z_{\alpha} \int d\mathbf{r} \rho_B^{\text{el}}(\mathbf{r}) / |\mathbf{r} - \mathbf{r}_{\alpha}| \\ & + \sum_{\beta} Z_{\beta} \int d\mathbf{r} \rho_A^{\text{el}}(\mathbf{r}) / |\mathbf{r} - \mathbf{r}_{\beta}| + \sum_{\alpha} \sum_{\beta} Z_{\alpha} Z_{\beta} / |\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}|. \end{aligned} \quad (3)$$

Regrouping Eq. (3) gives

$$E_{\text{int}}^{ES} = \int d\mathbf{r} V_A(\mathbf{r}) \rho_B^{\text{el}}(\mathbf{r}) + \sum_{\beta} Z_{\beta} V_A(\mathbf{r}_{\beta}). \quad (4)$$

3. Calculation of the electrostatic interaction energy E_{int}^{ES}

3.1. Analytical calculation

E_{int}^{ES} was calculated analytically without any approximation at the chosen SCF *ab initio* level, using the GAUSS82 program implemented on a Data General MV7800 computer. The chosen level was STO3G-W1s, i.e. the STO3G basis set without the core orbitals but after proper adjustment of the nuclei charges. That the core electrons do not play any essential role in the computation of E_{int}^{ES} for medium range interactions (i.e., for distances between the two partners larger than 2.5 Å [4]) was already known [6]. In addition, the validity of this assumption was verified by comparing the interaction energy between formamide (A) and water (B) calculated at the SCF *ab initio* STO3G level,

$$E_{\text{int}}^{a.i.} = E^{a.i.}(\text{AB}) - E^{a.i.}(\text{A}) - E^{a.i.}(\text{B}),$$

and the electrostatic energy, E_{int}^{ES} , calculated analytically as indicated above. For this purpose, H₂O was rotated around formamide for distances O (water)—N (formamide) and O (water)—C (formamide) equal to 4 and 4.5 Å, respectively (Fig. 1). The shapes of $E_{\text{int}}^{a.i.}$ and E_{int}^{ES} as a function of the rotations were

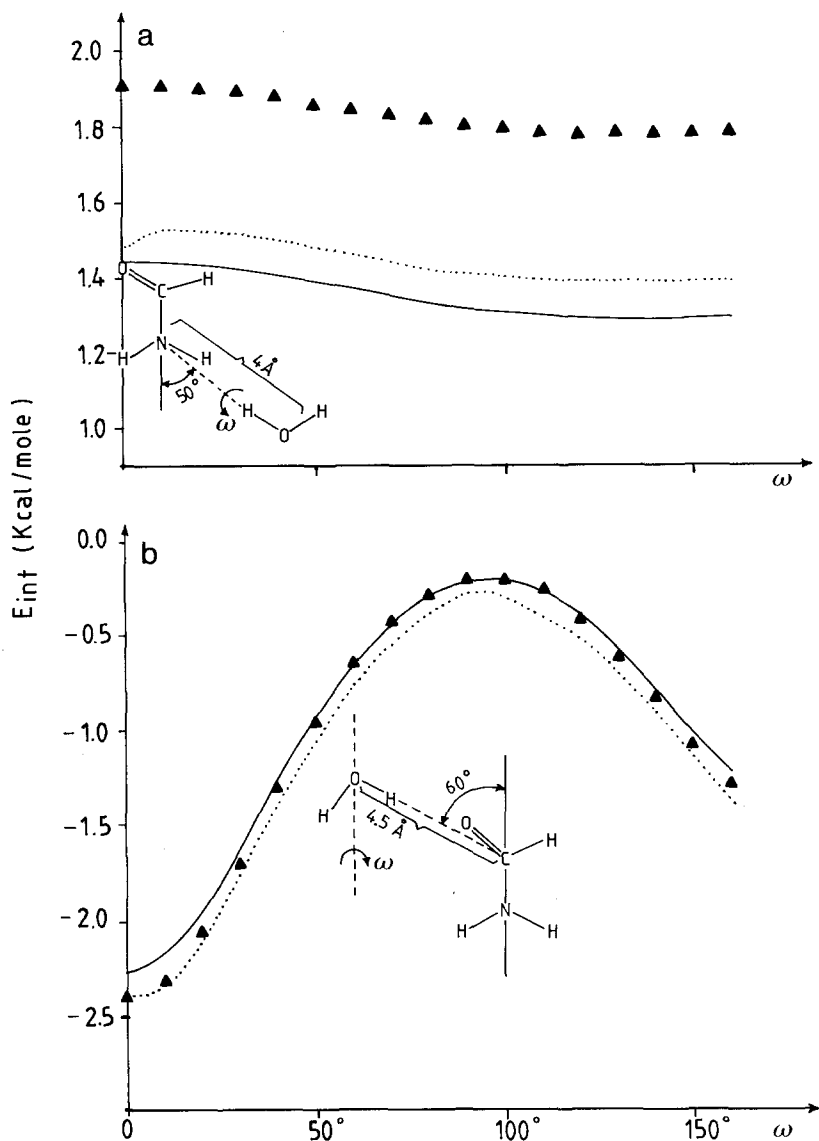


Fig. 1. Calculation of the interaction energy between a formamide and a water molecule. \blacktriangle , $E_{int}^{a.i.}$ (see text) at the STO-3G level; —, E_{int}^{ES} calculated analytically at the STO3G-W1s level; \cdots , E_{int}^{ES} calculated numerically at the STO3G-W1s level

superimposable, and the minima occurred at the same position. Differences of 0.5 kcal/mole were observed when the distances between the nuclei of H_2O and those of formamide became smaller than 2.5 Å (Fig. 1a). The variation of $E_{int}^{a.i.}$, analytical (W1s) E_{int}^{ES} and numerical (W1s) E_{int}^{ES} has also been studied for a few rotations with the 6-31G basis set, and the absolute energy values were found to be about twice as high as those calculated at the STO-3G level. As in the earlier

case, the greatest difference between $E_{\text{int}}^{a.i.}$ and the analytic (W1s) E_{int}^{ES} is of the order of 0.5 kcal/mol for the same intersystem distances. The difference between the numerically and analytically calculated E_{int}^{ES} is greater than in the STO-3G energy curves, but never by more than 0.3 kcal/mol. Despite these differences, the general conclusions emerging from both basis sets are the same: the positions of the $E_{\text{int}}^{a.i.}$ extrema are very well reproduced by the electrostatic component. Moreover, the analytical and numerical curves of E_{int}^{ES} are superimposable.

3.2. Numerical calculation

The 3D grids of the electrostatic potential $V_A(\mathbf{r})$ of partner A and the charge density $\gamma_B(\mathbf{r})$ of partner B [see Eq. (4)] were calculated at three levels of approximation. Level (a) is the STO3G-W1S level, i.e. neglecting the 1s orbitals; the reason for this is that when the 1s orbitals are used in the calculations, the electrostatic potential and the charge density vary too abruptly in the neighbourhood of the nuclei thus giving rise to erroneous numerical values. Level (b) is the same as level (a), but using density matrices obtained from CNDO deorthogonalized MO coefficients (see approximation IV [6]), and level (c) is level (b) with two additional approximations regarding the electrostatic potential: (i) using the net charges calculated by Mulliken population analysis, and (ii) replacing the mono-electronic integrals by $\gamma_{\alpha H}$ (see approximation I of [6]).

In each case, partner B and its charge density were translated and rotated in the frame of partner A. Given that the mesh points of the rotated density grid did not coincide with the mesh points of the electrostatic potential grid, the values of the electrostatic potential at each point of the density grid were estimated by a 3D linear interpolation from the eight nearest apex points of the surrounding cube; the 3D numerical integration was then performed using the Gill-Miller algorithm [7]. Since Eq. (3) has four components that have large values of opposite signs, the final result is, relatively, very small and so is imprecise even if the 3D integration of the first term is obtained with precision. In order to avoid this inherent difficulty, the integration was performed using Eq. (4), with the second term of Eq. (4) being calculated analytically.

4. Computational methods

The electrostatic interaction energy is numerically computed by 3D integration (Eq. (4)) from the V_A and ρ_B^{el} 3D grids calculated at the three levels mentioned above.

Level (a). The V_A and ρ_B^{el} 3D grids were calculated analytically by the programs set GAUSS82 on a Data General MV7800 (0.8 Mips), at the STO3G-W1s level.

Level (b). The V_A and ρ_B^{el} 3D grids at the CNDO level (approximation IV of [6]) were calculated analytically using the subroutines LINK301 and LINK604 [8] of the program set GAUSS70 (which, for LINK604, is 2.5 to 3 times faster than GAUSS82) adapted on an attached processor FPS164 (11 Mflops) and coupled

with a Digital VAX 11/780 computer (1 Mips); such a combination works about 6.5 times faster than the same run of GAUSS70 on the Data General MV7800. Thus, a 3D electrostatic potential of ($117 \times 119 \times 51$) points of water-imidazole was obtained in 30700 seconds, and a 3D electronic charge density grid of ($21 \times 25 \times 7$) points for methanol was calculated in 45 seconds.

Level (c). A program (constructed by GD) based on the calculation of the bielectronic integrals of the CNDO scheme was used. It allows one to calculate a 3D electrostatic potential of ($121 \times 161 \times 101$) points in 3060 seconds. This program is also executed on the attached processor FPS164.

4.1. Comparison of analytical with numerical computation times

The analytic electrostatic interaction energy is computed by our modified LINK604 of GAUSS82. The averaged time needed to *analytically* calculate one value of $E_{\text{int}}^{\text{ES}}$ at the STO3G-W1s level is 103 seconds, while the time needed to calculate it *numerically* is 25 seconds (with a stepsize of the 3D grids of 0.25 Å, irrespective of the level of approximation used to determine the density and potential grids). This great difference justifies the choice of a numerical procedure for the computation of $E_{\text{int}}^{\text{ES}}$, given the many values that have to be calculated to identify the stable conformations.

5. Results

The internal coordinates of the interacting triad methanol-water-imidazole were those of the optimized geometry at the SCF STO3G level projected on the plane of H₂O (Fig. 2, Table 1). The distance C (methanol)—O (water) was fixed at

Table 1. Geometric parameters of the frozen conformations of methanol and the water-imidazole dyad (Fig. 2); angles are in degrees, lengths in Å

Water-imidazole dyad			
$R(\text{O}_1\text{-H}_2)$	0.98921	$R(\text{O}_1\text{-H}_3)$	0.9825
$R(\text{H}_2\text{-N}_4)$	1.81073	$R(\text{C}_5\text{-N}_4)$	1.31772
$R(\text{N}_6\text{-C}_5)$	1.37868	$R(\text{C}_7\text{-N}_6)$	1.39152
$R(\text{C}_8\text{-C}_7)$	1.34703	$R(\text{H}_9\text{-C}_8)$	1.07929
$R(\text{H}_{10}\text{-C}_7)$	1.07945	$R(\text{H}_{11}\text{-N}_6)$	1.02178
$R(\text{H}_{12}\text{-C}_5)$	1.08426	$\alpha[\text{H}_3\text{O}_1\text{H}_2]$	102.831
$\alpha[\text{O}_1\text{H}_2\text{N}_4]$	177.462	$\alpha[\text{H}_2\text{N}_4\text{C}_5]$	123.536
$\alpha[\text{N}_6\text{C}_5\text{N}_4]$	111.546	$\alpha[\text{O}_7\text{N}_6\text{C}_5]$	106.844
$\alpha[\text{N}_6\text{C}_7\text{C}_8]$	105.953	$\alpha[\text{H}_9\text{C}_8\text{C}_7]$	128.023
$\alpha[\text{H}_{10}\text{C}_7\text{N}_6]$	121.904	$\alpha[\text{H}_{11}\text{N}_6\text{C}_5]$	126.450
$\alpha[\text{H}_{12}\text{C}_5\text{N}_4]$	126.307		
Methanol			
$R(\text{C}_1\text{-O}_2)$	1.42900	$R(\text{O}_2\text{-H}_3)$	0.99400
$R(\text{C}_1\text{-H}_{4,5,6})$	1.094		
$\alpha[\text{H}_3\text{O}_2\text{C}_1]$	104.7	$\alpha[\text{H}_{4,5,6}\text{C}_1\text{O}_2]$	109.4

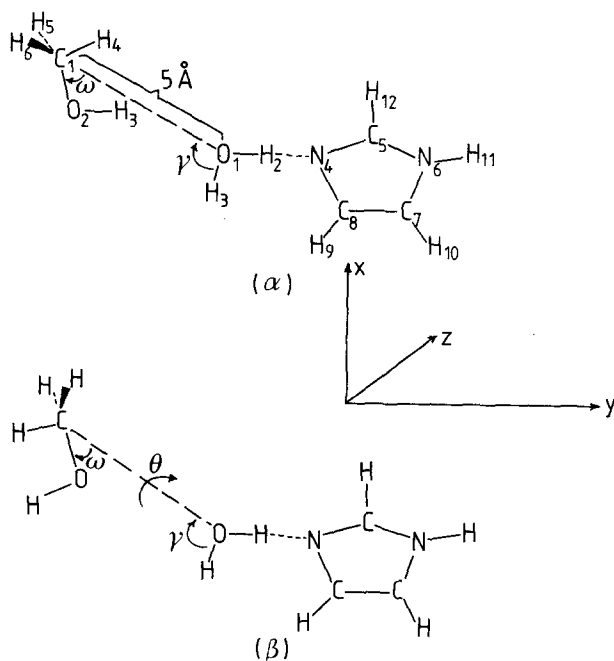


Fig. 2. Orientations of the methanol molecule in the water-imidazole frame

5 Å, two conformers (α and β) were investigated and the various rotations of the system were defined by the coordinates γ , ω and θ ; coordinate γ defines the rotation of methanol around O (H_2O) in the water-imidazole plane, ω defines the rotation of methanol around its carbon atom, again in the water-imidazole plane and θ defines the torsional angle $\text{O}_2\text{-C}_1\text{-O}_1\text{-H}_2$.

Figures 3 and 4 compare the $E_{\text{int}}^{\text{ES}}$ values obtained using the analytical procedure with those obtained using the numerical procedures (a), (b) and (c). The electrostatic potential was that of the dyad water-imidazole and the charge density was that of methanol. The calculations were performed for the two configurations α and β : (i) at fixed ω values of $+15^\circ$ and -15° with γ varying from 0 to 180° , (ii) at a fixed γ value of 140° with ω varying from 0 to 360° , and (iii) at a fixed γ value of 140° and a fixed ω value of -50° , with θ varying from -180° to 180° . The stepsize, i.e. the smallest distance between two mesh points in the 3D grids, was 0.5 \AA and 0.25 \AA for Figs. 3 and 4, respectively.

With a 0.5 \AA stepsize (Fig. 3), the shapes of the curves generated by the numerical procedures (a), (b) and (c) were similar to each other but they differed markedly from that generated by the analytical procedure. This discrepancy was attributed to the large value selected for the stepsize. With a 0.25 \AA stepsize, the computation time needed for the numerical procedure (a) became prohibitive; comparison with the analytical procedure was therefore restricted to the numerical procedures (b) and (c). Remarkably, and as shown by the curves in Fig. 4, the positions of the maxima and minima coincided well and the observed variations in the values of these extrema did not exceed 1 kcal/mole .

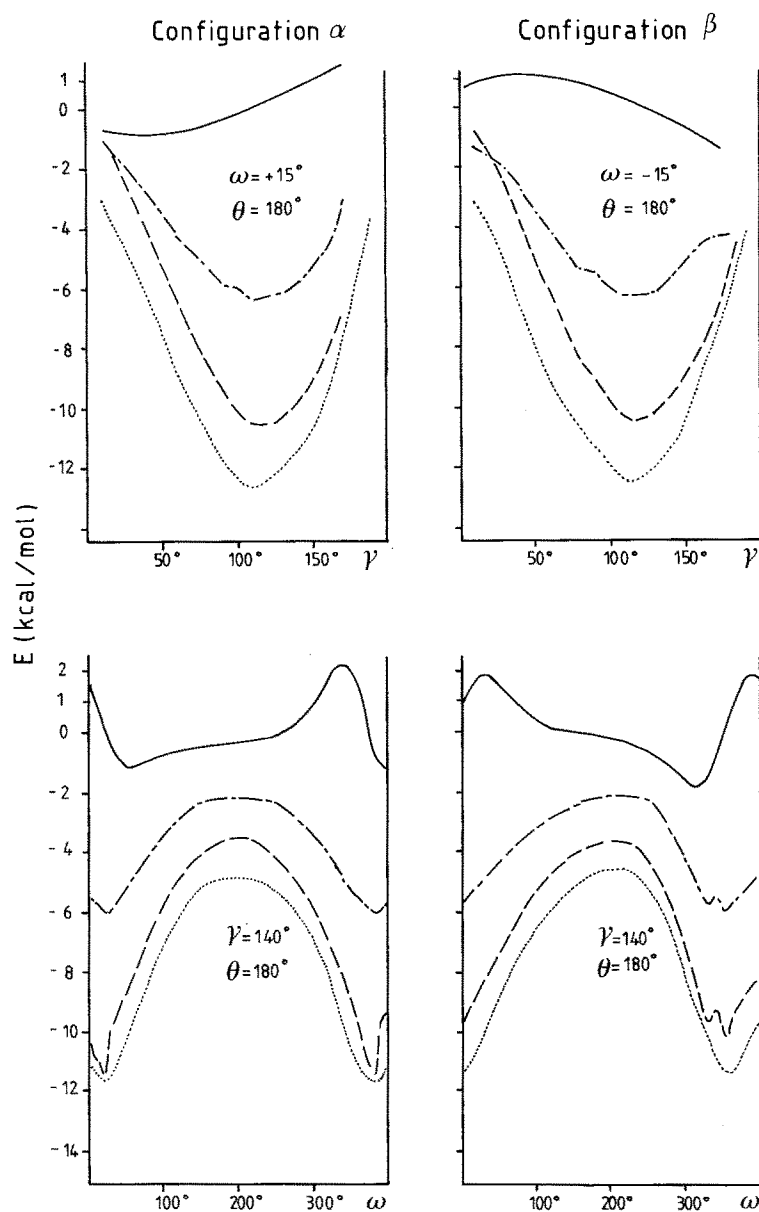


Fig. 3. Calculation of the electrostatic energy (in kcal/mole) between methanol and the water-imidazole dyad using a 0.5 Å stepsize. —, analytical STO3G-W1s procedure; ····, numerical STO3G-W1s procedure (a); ---, numerical CNDO procedure (b); ·-·-, numerical CNDO procedure (c); for further information, see text. The chosen configurations (see Fig. 2) and the values of the coordinates γ , ω and θ are as indicated

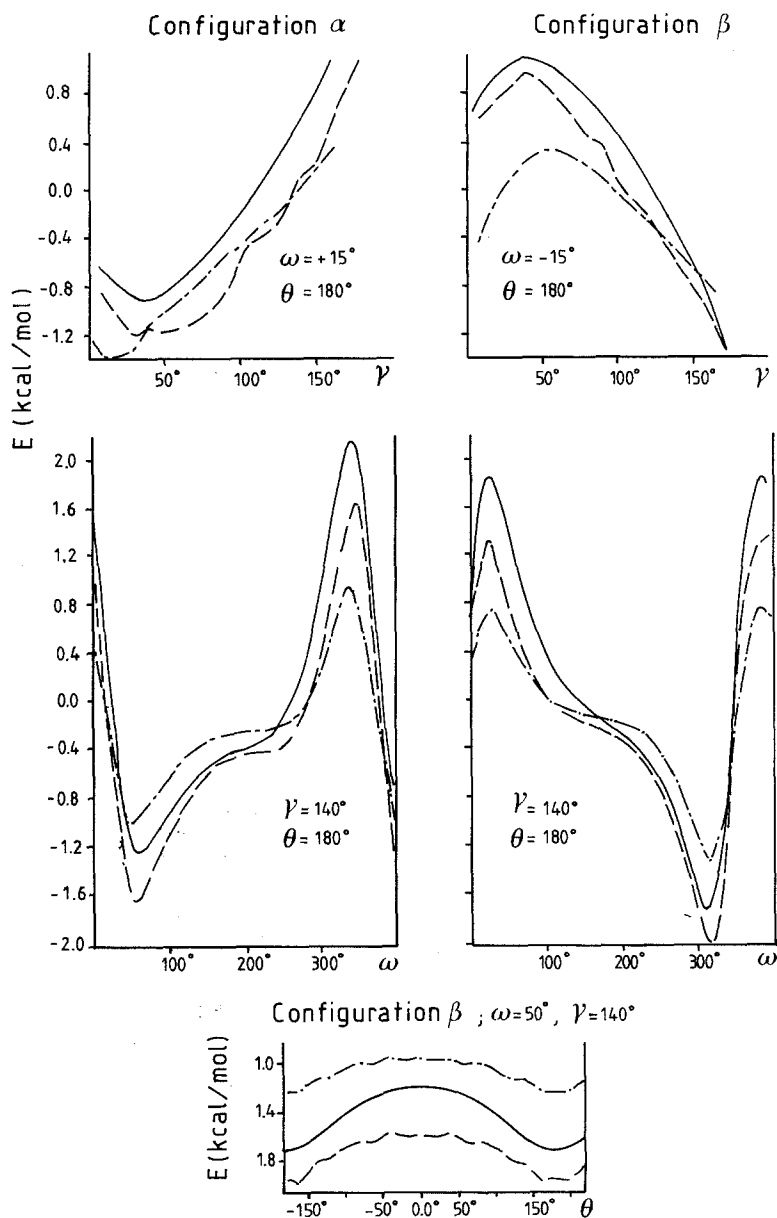


Fig. 4. Calculation of the electrostatic energy (in kcal/mole) between methanol and the water-imidazole dyad using a 0.25 \AA stepsize. For symbols, configurations and coordinates see Fig. 3

In a second series of experiments, the $E_{\text{int}}^{\text{ES}}$ values were computed using, as above, the electrostatic potential of the dyad water-imidazole and the charge density of methanol, and conversely, the electrostatic potential of methanol and the charge density of the dyad water-imidazole; calculations were made with both the numerical procedures (b) and (c). The stepsize was 0.25 \AA in the O_x and O_z

directions and 0.20 Å in the O_y direction (referred to as $0.25 \times 0.20 \times 0.25$) for procedure (c). This same grid, and a grid with 0.25 Å steps in all three directions (referred to as 3×0.25) were used for procedure (b).

These calculations revealed that the selected stepsize had little effect on the $E_{\text{int}}^{\text{ES}}$ values when the electrostatic potential of the dyad water-imidazole was used for the computation of Eq. (4), with the observed energy differences not exceeding 0.1 kcal/mole. Conversely, differences as large as 1.5 kcal/mole were observed when the electrostatic potential was that of methanol. This different behaviour was related to the fact that in the interacting region the electrostatic potential of methanol changed signs during rotations whereas that of the dyad water-imidazole varied smoothly and remained negative. However, with a stepsize of $0.25 \times 0.20 \times 0.25$, it was not important whether Eq. (4) was solved by procedure (b) or procedure (c), or which of the two partners, methanol or the dyad water-imidazole, was used to compute the electrostatic potential or the charge density: the observed differences in the minima and maxima of the $E_{\text{int}}^{\text{ES}}$ values ranged between 0.2 and 0.45 kcal/mole, and the largest difference observed between the positions of the minima did not exceed 10° .

6. Discussion

Defining the interaction energy between a biological macromolecule such as an enzyme or a receptor and its ligands allows one to define, from the positions and the values of the minima, the most stable conformers and the heat of formation of these conformers, and hence to predict the outcome of the reactions. As long as the distances between the nuclei of the interacting partners remain above 2.5 Å, the interaction energy is nearly quantitatively expressed by the electrostatic interaction energy. Knowing the spatial disposition of the atoms of a model enzyme or receptor active site, one can calculate its electrostatic potential on a chosen 3D grid, and from this, the electrostatic interaction energy for a given ligand with distinct orientations and at varying distances.

The methodological developments described above show that, given a well-calibrated stepsize, the electrostatic interaction energy of the trimer methanol-water-imidazole can be computed numerically using Eq. (4) at the CNDO levels of approximation (b) and (c). These procedures yield consistent minimum positions and energy values that compare well with those calculated analytically. These results are encouraging since approximation (c) is known to generate electrostatic potential wells that are too shallow. One should note, however, that a shallow electrostatic potential well varies smoothly in the volume defined by the 3D charge density grid, i.e. it has many values of the same order as that of the extremum. Hence, the results of a 3D integration [as well as that of the sum in Eq. (4)] with such a smooth and shallow potential calculated at level (c) may be similar to those obtained at level (b) with a deeper and more abruptly varying potential well.

Two other comments deserve attention. They relate to the CPU time. As already mentioned (Sect. 3), the numerical procedure is about 4 times faster than

the analytical one. Given the large number of E_{int}^{ES} values that must be computed to identify the stable conformers, this gain of time is appreciable. In addition, the question of whether the potential or the density to be used is that of the large or the small partner is also dictated by CPU time considerations. Calculation of the potential of the small partner may be faster, but then the dimension of the 3D grid needed to include the 3D grid of the density of the large partner is, of course, much wider. Moreover, the time that is needed to numerically integrate the product of the potential and the electronic density depends mainly on the dimension of the 3D density grid, and thus is much longer when the large partner is used to calculate the density. In the case of the methanol-(water-imidazole) system studied here, the process is 3 times faster when the charge density used for the computation is that of methanol and the electrostatic potential is that of the water-imidazole dyad.

Further work is needed before the approach presented here can be applied to a macromolecular system. In particular, formation of non-covalent complexes between an enzyme or a receptor and their ligands are likely to involve interactions occurring at distances below 2.5 Å. A polarization term [3] must then be introduced in the expression of the interaction energy.

Acknowledgements. This work was supported in part by the Fonds de la Recherche Scientifique Médicale (contract n°3.4507.83), an Action concertée with the Belgian Government (convention 86/91-90), a convention with the Région wallonne (C2/C16/Conv.246/20428), the Fonds de Recherche de la Faculté de Médecine ULg and a contract with the EEC (BAP-0197-B).

References

1. Pullman A, Pullman B (1975) *Ann Rev Biophys* 7:505
2. Bonaccorsi R, Ghio C, Tomasi J (1984) *Int J Quantum Chem* 26:637
3. Kitaura K, Morokuma K (1976) *Int J Quantum Chem* 10:325
4. Scrocco E, Tomasi J (1973) *Topics Curr Chem* 42:130
5. Weinstein H, Osman R, Green JP, Topiol S (1981) In: Politzer P, Truhlar DG (eds) *Chemical applications of atomic and molecular electrostatic potentials*. Plenum Press, New York
6. Giessner-Prettre C, Pullman A (1972) *Theor Chim Acta* 25:83
7. Gill PE, Miller GF (1972) *Comp J* 15:80
8. Peeters D, Sana M. Program DENPOT, University of Louvain, Louvain-la-Neuve, Belgium, QCPE 360, Quantum Chemistry Program Exchange, Department of Chemistry, Indiana University, Bloomington, IN 47405, USA