J. Am. Chem. Soc. 1989, 111, 3417–3421

Appraisal of Solvation Effects on Chemical Functional Groups: Amides and Esters in Terms of Transferable Subgroup Contributions

Giuliano Alagona,[*,†] **Caterina Ghio,**[†] **Juan Igual,**[‡] **and Jacopo Tomasi**[§]

Contribution from the Istituto di Chimica Quantistica ed Energetica Molecolare del C.N.R., Via Risorgimento, 35, I-56126 Pisa, Italy, Departamento de Quimica Fisica, Facultat de Ciencias de Tarragona, Pl. Imperial Tarraco, 1, Tarragona, Spain, and Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Risorgimento, 35, I-56126 Pisa, Italy.

Received August 15, 1988

Abstract: The basic aspects of the solvation of *N*-methylformamide and methyl formate, obtained and analyzed in another paper on the basis of a full ab initio MO SCF procedure, are here examined by making use of a semiclassical model that replaces the full ab initio calculations with a library of chemical group descriptions derived from wave functions at the 4-31G level. The procedure outlined in this paper permits a sequence of approximations of increasing realism, each related to the introduction of additional physical effects, eventually reaching the ab initio level. Even the lowest levels of the procedure, however, are able to describe the essential aspects of the solvation of the two test molecules.

1. Introduction

The scope of the present paper is to reexamine the description of solvent effects on two characteristic chemical functions, the amidic and the esteric linkages, given in another paper[1] (hereafter called paper 1) on a full ab initio basis, using now a semiclassical approach, i.e., quantum mechanical descriptions of chemical subunits interacting in a classical way. It is a specific application, and a test, of a more general research project, addressed to the evaluation and interpretation of molecular properties without resorting to direct ab initio calculations on each molecular system taken into account, but trying to exploit, in a controlled way, the ab initio results obtained for prototype molecules. The final goal is the definition of procedures able to give reliable information even on very large molecular systems and, at the same time, an acceptable interpretation of the information thus ensured.

Exploiting our previous attempts of modeling chemical substitution effects (overviews are reported in ref 2–4), we enlarge the field of application of our semiclassical model to solvation effects, treated on the same foot as chemical substitution effects. Some preliminary material justifying the assimilation of internal (as chemical substitutions) and external effects (as the interaction with a solvent) may be found in a recent paper.[5]

We may say that, after having treated "*true*" *molecules in a model solvent*, we pass now to a second step, in which the "true" molecules are replaced with group models (called transferable group models).

As far as the outline of the solvation model is concerned, we refer the reader to preceding papers,[1–6] limiting ourselves to report here the concepts indispensable to make this paper self-exhaustive.

2. *Outline of the Solvation Method*

Basic features of the solvation method[6] are the operative definition of a solvent reaction operator, $V(\sigma,\mathrm{M})$, for the solute M in a solvent with given macroscopic properties and the methods employed to get a simultaneous solution of the electrostatic and Hartree–Fock (HF) problems, when M is inserted in the solvent (considered as a continuous dielectric medium). In order to do this, one has to solve, iteratively, the following pseudo-HF equation:

$$(H^{\circ}_{\mathrm{M}} + V(\sigma,\mathrm{M}))\Psi'_{\mathrm{M}} = E'_{\mathrm{el}}(\mathrm{M,sol})\Psi'_{\mathrm{M}} \qquad (1)$$

$V(\sigma,\mathrm{M})$ mainly depends on the total charge distribution of M in solution:

$$\Gamma(r_1;\mathrm{M,sol}) = -n\int \Psi'_{\mathrm{M}}(r_1 \ldots r_n)\Psi'^{*}(r_1 \ldots r_n)\,\mathrm{d}r'_1 + \sum_{\alpha}^{\mathrm{nuc}} Z_{\alpha}\delta(r_1 - R_{\alpha}) \qquad (2)$$

via the corresponding electrostatic potential

$$V(\mathrm{M,sol}) = \int (r - r_1)^{-1}\Gamma(r_1;\mathrm{M,sol})\,\mathrm{d}r \qquad (3)$$

$V(\sigma,\mathrm{M})$, in fact, depends on an apparent charge distribution $\sigma_{\mathrm{M}}(\mathrm{s})$ defined over the surface of the solute cavity in the dielectric medium

$$\sigma_{\mathrm{M}}(\mathrm{s}) = -\frac{\epsilon - 1}{4\pi\epsilon}\,\frac{\partial}{\partial n}[V(\mathrm{M,sol}) + V(\sigma,\mathrm{M})] \qquad (4)$$

This implicit definition (implemented in the program via an iterative procedure) has several characteristics of interest for our attempt of using semiclassical models.

The interactions involved are of classical origin: σ_{M} produces a field $F(\sigma,\mathrm{M}) = -\nabla V(\sigma,\mathrm{M})$ that modifies $\Gamma(\mathrm{M})$, and its value depends on $\Gamma(\mathrm{M})$ itself. We are thus free of introducing approximate descriptions in $\Gamma(\mathrm{M})$, allowing for classical modifications to the charge distribution.

The energy of the system (actually a free energy G) is related to the solution of eq 1, modified by a corrective factor that depends again on $\Gamma(\mathrm{M,sol})$ and $V(\sigma,\mathrm{M})$:

$$G_{\mathrm{SCF}}(\mathrm{M,sol}) = E'_{\mathrm{tot}}(\mathrm{M,sol}) - \frac{1}{2}\int \Gamma(\mathrm{M,sol})V(\sigma,\mathrm{M})\,\mathrm{d}\tau \qquad (5)$$

The electrostatic contribution to the free energy of solution is given by

$$\Delta G^{\mathrm{sol}}_{\mathrm{el}}(\mathrm{M}) = G_{\mathrm{SCF}}(\mathrm{M,sol}) - E^{\circ}_{\mathrm{tot}}(\mathrm{M,vac}) \qquad (6)$$

[†] Istituto di Chimica Quantistica ed Energetica Molecolare del C.N.R.
[‡] Departamento de Quimica Fisica, Facultat de Ciencias de Tarragona.
[§] Dipartimento di Chimica e Chimica Industriale, Università di Pisa.

(1) Alagona, G.; Ghio, C.; Igual, J.; Tomasi, J. "An Appraisal of Solvation Effects on Chemical Functional Groups: The Amidic and the Esteric Linkages". *J. Mol. Struct.: THEOCHEM*, in press.

(2) Bonaccorsi, R.; Ghio, C.; Scrocco, E.; Tomasi, J. *Isr. J. Chem.* **1980**, *19*, 109.

(3) Bonaccorsi, R.; Ghio, C.; Tomasi, J. *Int. J. Quantum Chem.* **1984**, *26*, 637.

(4) Alagona, G.; Bonaccorsi, R.; Ghio, C.; Tomasi, J. *J. Mol. Struct.: THEOCHEM* **1986**, *135*, 39.

(5) Alagona, G.; Bonaccorsi, R.; Ghio, C.; Montagnani, R.; Tomasi, J. *Pure Appl. Chem.* **1988**, *60*, 231.

(6) Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

0002-7863/89/1511-3417$01.50/0

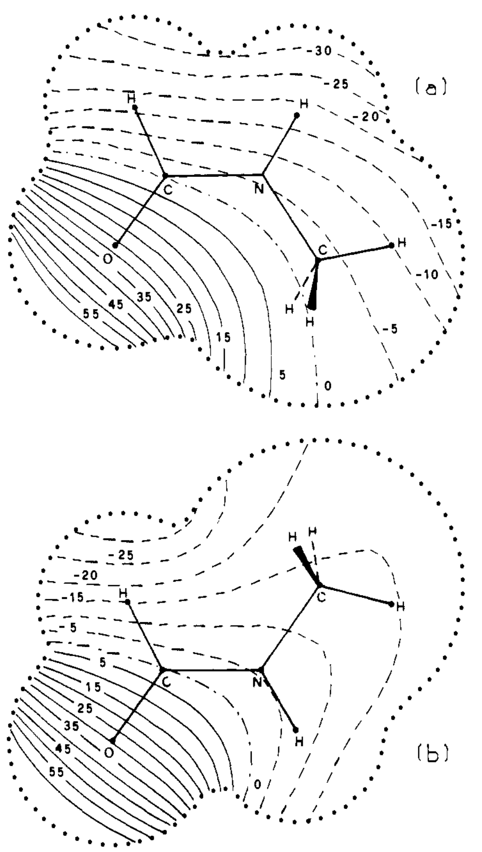


Figure 1. Solvent (ϵ = 78.5) reaction potential $V(\sigma,M)$ in the molecular plane, obtained from a given general library group description (eq 12), at the 4-31G level. The isopotential lines are spaced by 5 kcal/mol. (a) *trans-N*-Methylformamide (**1**-T); (b) *cis-N*-methylformamide (**1**-C).

In view of the semiclassical approach we are here testing, we add an approximation of eq 6 in which the difference between the mean kinetic energy of the electrons of M in solution and in vacuo is neglected:

$$\Delta G_{el}^{sol}(M)^* = \frac{1}{2}\int \Gamma(M,sol)V(\sigma,M)\,d\tau \qquad (7)$$

In the following we shall consider approximate expressions of $\Gamma(M)$, and of $V(\sigma,M)$, still retaining the basic features of the solvation model.

3. Use of Chemical Group Models

Definitions, justifications, and examples of our elaboration of chemical group models can be found in the literature.[2-5] We shall resume here some formal aspects, restricting the numerical applications to the simplest level that seems suitable for approximate evaluations of solvent effects.

The molecular wave function of a molecule M is divided into group contributions, making use of a transformation into Boys localized orbitals.[7] The total charge distribution is written as a function of the groups g as

$$\Gamma(M) = \sum_{g}\gamma(g,M) \qquad (8)$$

The sum of charge distributions $\gamma(g,M)$ is exactly equal to the charge distribution obtained from the molecular wave function.

The examination of a considerable number of charge distributions for g in different molecules M leads to the conclusion that the mutual interactions of the groups inside M produce minor modifications in a prototype of $\gamma(g,M)$, which has no direct reference to M: we indicate this prototype as $\gamma(g,0)$. In addition, these modifications are described, to a good extent, by the action of classical fields.

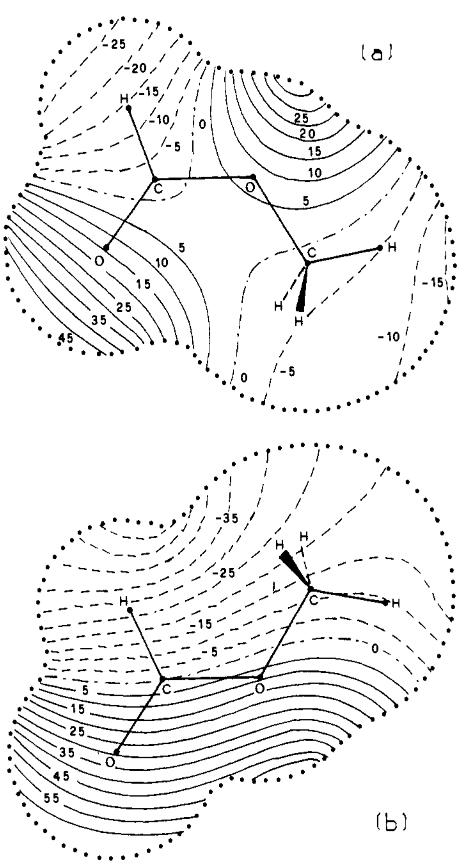


Figure 2. As Figure 1, but referred to (a) *trans*-methyl formate (**2**-T) and (b) *cis*-methyl formate (**2**-C).

To this end, we define the field of the molecular remainder, $F(M/g)$, which collects the fields of all the groups of M, with the exception of g. In first approximation, the field of the molecular remainder may be expressed in terms of the prototype distributions, $F(M/g,0)$.

A more accurate discription of g inside M may be thus obtained; the modification of $\gamma(g,0)$ is symbolically represented by

$$\gamma(g,0) \xrightarrow{F(M/g,0)} \gamma(g,M,p) \qquad (9)$$

Basically, the transformation (9) requires a further partition of g subunits into atomic contributions and the application of semiempirical recipes for the atom and bond polarizations. For instance, assuming g = b_{AB}, i.e., a σ bond between atoms A and B, we have

$$\gamma(b_{AB},0) = -2\lambda^2(r,b_{AB}) + 1\cdot\delta(r - R_A) + 1\cdot\delta(r - R_B) \qquad (10)$$

$$\lambda(r) = c_A h_A(r) + c_B h_B(r) \qquad (11)$$

where $h_A(r)$ and $h_B(r)$ are generalized atomic hybrids and c_A and c_B are numerical coefficients. The shape of the hybrids and the values of the coefficients depend on $F(M/g,0)$. By combining ab initio calculations on different molecules containing g and the application to these molecules of additional external fields F(ext) of different shape and strength, it has been possible to derive some rules on the dependency of the h_A's and of the c_A's on the field acting on g. From these rules we have derived some practical recipes about the most convenient way of measuring the field and of translating this information into a modification of $\gamma(g,0)$. In the b_{AB} group considered as an example, we found that about 90% of the polarization effect is due to changes in the c_A and c_B coefficients and that this change may be monitored by the value of the parallel component of $F(M/g,0)$, measured at the midpoint of A–B. Similar recipes have been obtained for other groups, and also for the effect of $F(M/g,0)$ on the bond lengths and angles. The empirical relationships are, of course, basis set dependent.

---

[7] Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.

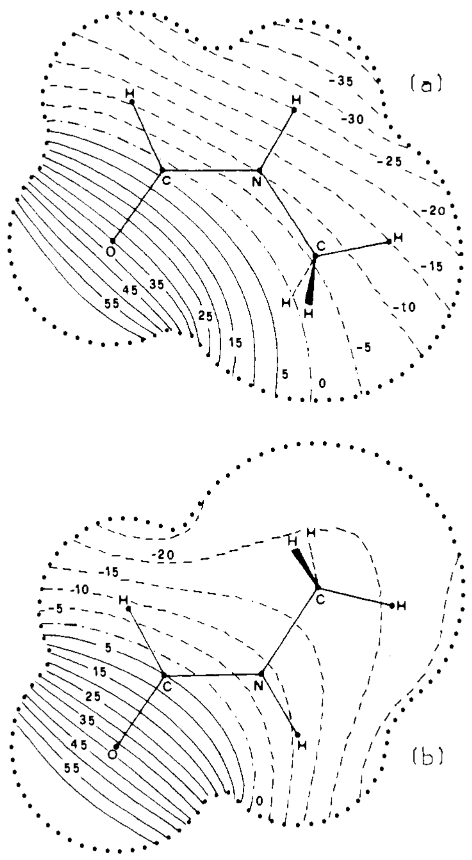


Figure 3. Solvent (ϵ = 78.5) reaction potential $V(\sigma,M)$ in the molecular plane, obtained from a full ab initio calculation at the 4-31G level. The isopotential lines are spaced by 5 kcal/mol. (a) *trans-N*-Methylformamide (**1**-T); (b) *cis-N*-methylformamide (**1**-C).

Practical methods to get $\gamma(g,M,p)$, once $\gamma(g,0)$ and $F(M/g,0)$ are known, have been reported in previous papers.[2,3,5,8] The description of $\gamma(g,M,p)$ may be refined, but this simple procedure accounts for more than 90% of the effect of the molecular remainder.

We have defined thus far two levels of approximation for the molecular charge distribution:

$$\Gamma^0(M) = \sum_g \gamma(g,0) \tag{12}$$

$$\Gamma^p(M) = \sum_g \gamma(g,M,p) \tag{13}$$

From both charge distributions we may define, according to eq 4, a solvent reaction potential $V(\sigma,M,0)$ and $V(\sigma,M,p)$, respectively. Analogous definitions hold for the solvent reaction field $F(\sigma,M,0)$ and $F(\sigma,M,p)$. At a first level of approximation, the action of the electrostatic field is linear. We may thus add $F(\sigma,M,0)$ to $F(M/g,0)$ to get group descriptions modified by the solvent:

$$\gamma(g,0) \xrightarrow{F(M/g,0)\,+\,F(\sigma,M,0)} \gamma(g,M,\sigma) \tag{14}$$

Using $F(\sigma,M,p)$ instead of $F(\sigma,M,0)$ in eq 14, one obtains a description of the group picture in solution slightly different from that given by eq 14. It is possible to adopt here an iterative procedure that takes into account also nonlinear electrostatic effects.

This procedure is the semiclassical counterpart of that given in paper 1 for the evaluation of solvent effects on the molecular wave function of M. It can be noticed that ab initio calculations on M are not necessary, since the library of $\gamma(g,0)$ charge distributions is derived from calculations on simpler molecules.

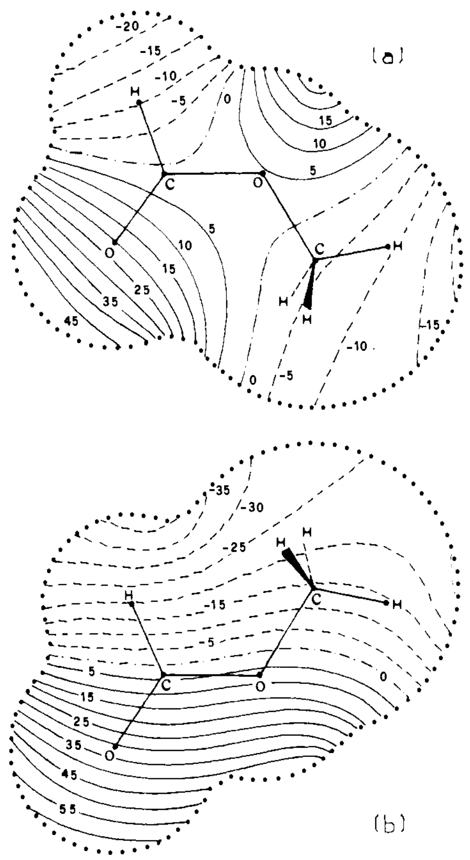


Figure 4. As Figure 3, but referred to (a) *trans*-methyl formate (**2**-T); (b) *cis*-methyl formate (**2**-C).

A more approximate definition of solvent effects makes use of $\Gamma^0(M)$ again and of $V(\sigma,M,0)$, without introducing polarization effects, as done in eq 14. In this case we neglect solute polarization. When solute polarization is neglected, a direct simple expression is obtained for the electrostatic solvation free energy:

$$\Delta G_{el} = \frac{1}{2}\int \Gamma^0(M)V(\sigma,M,0)\,d\tau \tag{15}$$

This is the simplest expression used in the following. Another expression, of similar nature, takes into account the effect of the molecular field, $F(M/g)$, on the group charge distributions:

$$\Delta G_{el} = \frac{1}{2}\int \Gamma^p(M)V(\sigma,M,0)\,d\tau \tag{16}$$

We have thus defined a hierarchy of approximations, starting from a very crude model, which uses only rigid descriptions of groups, taken from a general library without additional manipulations, and reaching, step by step, the full ab initio calculation, in our computational scheme.

This set of approximations may be used to get appraisals of the solvation energy contributions (eq 6 and 7) and information about the effects of the solvent, coupled or not coupled to the effect of chemical substitution. We shall not use here the complete set of approximations on the compounds considered, *N*-methylformamide (**1**) and methyl formate (**2**), because we limit ourselves to show to what extent the lowest levels of the approximations reproduce the outstanding information. The interested reader may find examples of a full application of all the approximation levels for the evaluation of conformational energy surfaces in solution in ref 9 and applications of several approximation levels to the combined effects of chemical substitution and solvent in ref 5.

4. Applications

4.1. Solvent Reaction Potential. Our interpretation of solvent effects is essentially based on the $V(\sigma,M)$ maps. In Figure 1a,b we report the maps for *trans*- and *cis-N*-methylformamide (**1**-T

---

(8) Ghio, C.; Scrocco, E.; Tomasi, J. *Theor. Chim. Acta* **1980**, *56*, 75.

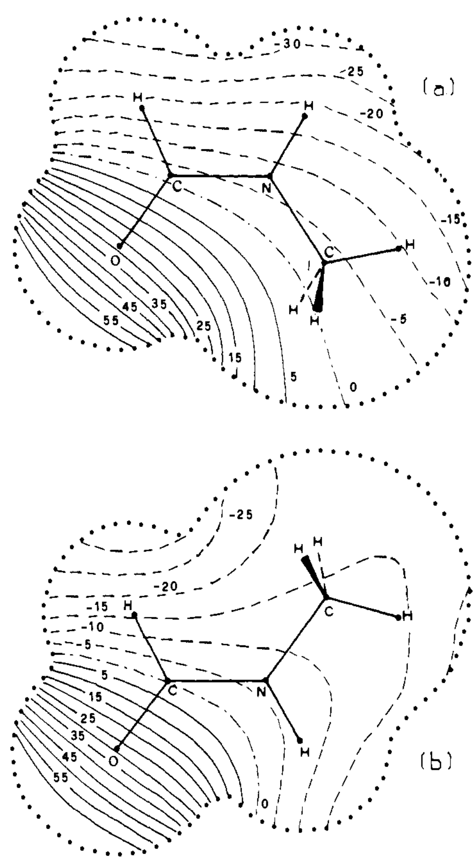


Figure 5. Solvent (ϵ = 78.5) reaction potential $V(\sigma,M)$ in the molecular plane, obtained from a group description modified by the molecular field (eq 13). The isopotential lines are spaced by 5 kcal/mol. (a) *trans-N*-Methylformamide (**1**-T); (b) *cis-N*-methylformamide (**1**-C).

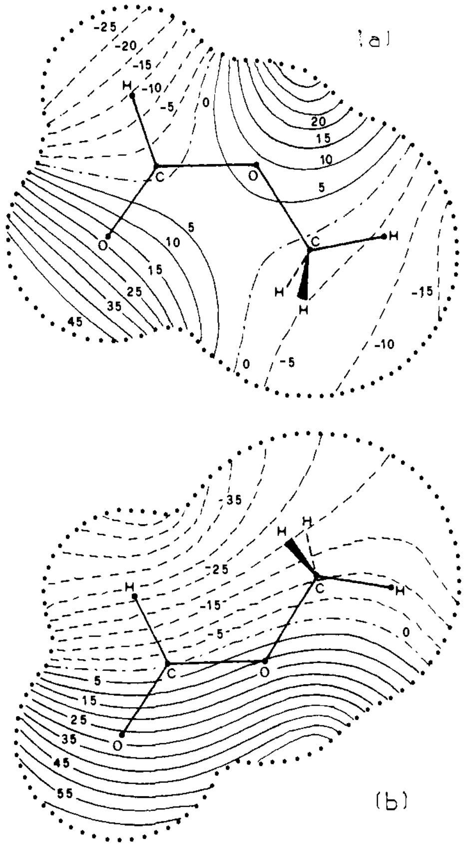


Figure 6. As Figure 5, but referred to (a) *trans*-methyl formate (**2**-T) and (b) *cis*-methyl formate (**2**-C).

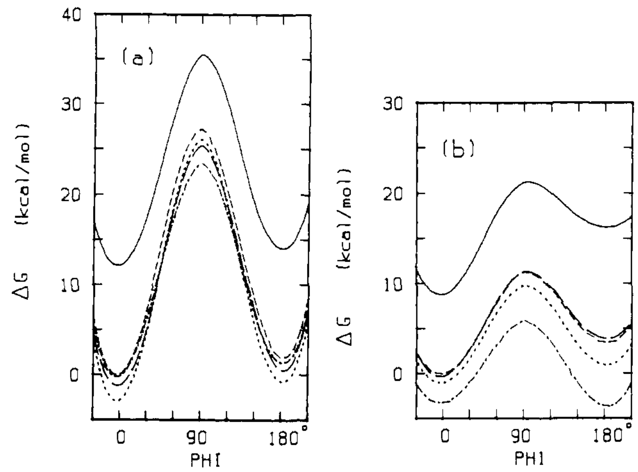
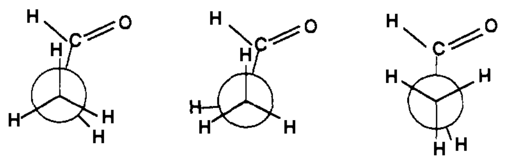


Figure 7. Section of the potential energy hypersurface corresponding to the rotation around the C–X bond. X stands for either N or O, depending on the system considered. The values are relative to the $\Delta G^{\rm sol}$ for the trans conformers taken as zeroes. (—) $\Delta G^{\rm vac}$; (---) $\Delta G^{\rm sol}$; (•••) ΔG° (eq 7); (-•-) $\Delta G^{\rm rig}$ (eq 12); (– – –) $\Delta G^{\rm pol}$ (eq 13). (a) *N*-Methylformamide; (b) methyl formate.

Chart I. View Along the N–C_α Bond Head On



1-TSA **1-TSB** **1-TSpl**

and **1**-C), respectively, in the molecular plane, obtained with the γ(g,0) group descriptions. Analogous maps are reported in Figure 2a,b for *trans*- and *cis*-methyl formate (**2**-T and **2**-C). A comparison with Figures 3a,b and 4a,b, reporting analogous maps computed with the full ab initio procedure, shows that the outstanding characteristics are reproduced in a semiquantitative manner. In Figures 5a,b and 6a,b the maps for the same compounds in the same plane, but obtained with group descriptions modified by the molecular field, are displayed. The application of the solvent field $F(\sigma,M,0)$ produces an even better agreement with $V(\sigma,M)$.

4.2. Trans–Cis Conversion. We report in Figure 7a,b the section of the potential energy surface (PES) corresponding to the rotation around the C–X bond for **1** and **2**, respectively, obtained in different approximations. We report also the corresponding SCF curves in vacuo obtained with geometry optimization at several φ values. The transition state (TS) considered in Figure 7a is that labeled **1**-TSA in Chart I. The lowest value of the in vacuo curve has been taken as a reference energy in both cases (see paper 1 for the actual numerical values of these conformations). It may be noticed that both the energy barrier and the cis–trans energy difference are appreciated in a reasonable way. In particular, the features distinguishing the rotation in **1** with respect to **2** (difference in the barrier height) are reproduced by the rigid γ(g,0) representation.

Here again the consideration of the field due to the molecular remainder, $F(M/g,0)$, improves the description.

The inclusion of the solvent polarizing effects (eq 14) produces potential curves in a fair agreement with those concerning approximation (7) without using, as already remarked, the wave function of M.

4.3. Inversion at the N Atom of 1 at the TS. We have shown in paper 1 that the N atom of **1** loses its planar conformation during the rotation about the C–N bond. In vacuo the PES has

---

(9) Tomasi, J.; Alagona, G.; Bonaccorsi, R.; Ghio, C. In *Modelling of Structures and Properties of Molecules*; Maksic, Z. B., Ed.; Horwood: Chichester, 1987; Chapter 20, p 330.

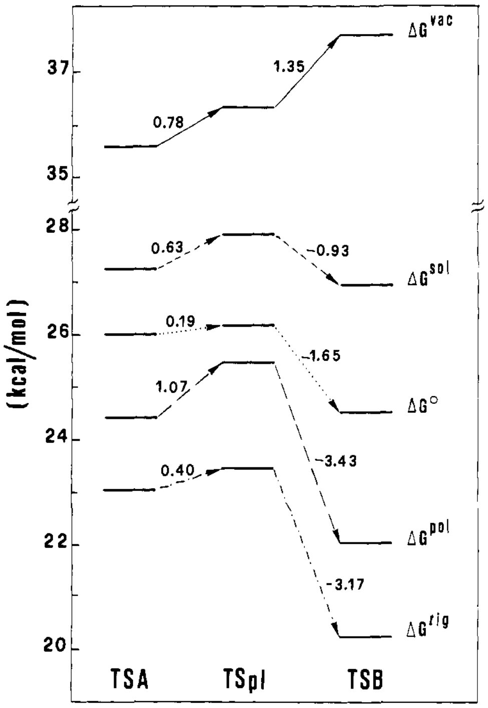


Figure 8. Diagram of the potential energy corresponding to the different pyramidalizations at the N atom for *N*-methylformamide. The values are relative to the ΔG^{sol} for **1**-T taken as zero. (—) ΔG^{vac}; (- - -) ΔG^{sol}; (•••) $\Delta G°$ (eq 7); (-·-); ΔG^{rig} (eq 12); (– – –) ΔG^{pol} (eq 13).

only one well-defined saddle point, **1**-TSA, with a pyramidalization angle $\tau_A = 29.27°$, whereas a second geometry called **1**-TSB, with $\tau_B = 30°$, practically corresponds to an inflection on the surface. In aqueous solution **1**-TSA and **1**-TSB are two distinct and well-characterized saddle points, at $\tau_A = 29.266°$ and $\tau_B = 29.78°$, with comparable energies.

The description in terms of rigid group contributions, $\gamma(g,0)$, is able to reproduce this feature of the PES in solution (see Figure 8). This finding strongly suggests that the appearance of the second TS in *N*-methylformamide is mainly due to *direct* solvent effects and not to secondary effects due to the polarization of the solute under the action of the solvent reaction field.

The use of group contributions polarized either by the molecular remainder (eq 13) or by the solvent (eq 14) simply refines the description provided by the rigid group contributions.

For the sake of clarity the results displayed in Figure 8 are also reported in Table I.

5. Conclusions

We have reported a few examples of the results obtainable with the semiclassical approximation regarding some aspects of the solvation energy, to which others could be added (e.g., the rotation of the Me groups) with similar results. The shape of the semiclassical representations of $V(\sigma,M)$ supports the statement that the interpretation of solvent effects, given in paper 1, could be repeated without noticeable changes also using these approximate tools. The relatively large set of successful applications of the semiclassical reproduction and interpretation of effects due to chemical substitutions seems to be enlarged also to solvation effects. This conclusion is justified also by physical considerations: chemical substitutions are short-range strong interactions, with a considerable weight of nonclassical quantal effects, whereas solvent effects are more classical in nature and reduced only to the classical components in the continuous model for solutions. The description adopted by these models neglects to consider random fluctuations in the solvent, which could induce effects having a more important quantal contribution. A semiclassical modelization of these fluctuations is an open question not yet examined.

Table I. Values of ΔG (kcal/mol) in the Various Approximations for the Conformers of *N*-Methylformamide (**1**) and Methyl Formate (**2**)

conformer	ΔG^{vac}	ΔG^{sol}	$\Delta G°$ [a]	ΔG^{pol} [b]	ΔG^{rig} [c]
1-T	12.16	0.00	-2.84	-1.18	-0.12
1-C	14.10	2.00	-0.69	0.61	1.40
1-TSA	35.58	27.26	25.98	24.40	23.05
1-TSpl	36.36	27.89	26.17	25.47	23.45
1-TSB	37.71	26.96	24.52	22.04	20.28
2-T	8.79	0.00	-1.01	-0.31	-3.21
2-C	16.26	3.50	0.96	3.89	-3.56
2-TS	21.24	11.32	9.84	11.39	5.88

[a] From eq 7. [b] Produced by the charge distribution from eq 13. [c] Produced by the charge distribution from eq 12.

A second point of some relevance for the conclusions is that the lowest approximation level gives valuable semiquantitative results. One could argue whether it is worthwhile to do the extra computational effort necessary to reach the second and the following steps in the approximation. In the present case the answer is negative, but the chemical systems here considered are relatively simple. It is prudent, however, to have at our disposal more effective tools to be employed where necessary. It rests that all the approximations we have considered are decidedly less expensive than the ab initio calculations, especially when the size of M increases.

There are now several efficient methods available, either of empirical or semiempirical origin, able to give reasonable estimates of geometries, relative energies, and properties of isolated molecules with a modest computational cost. A great effort has been spent also in the formulation of transferable expressions of group contributions to the electrostatic potential in vacuo (see ref 10 and 11 for recent additions). There are in the literature some attempts to introduce such simplified and transferable descriptions of electrostatic properties of groups into solvation methods akin to ours,[12,13] but this field is still in its infancy: we are confident that the present approach may represent a valid and profitable contribution to its development.

Acknowledgment. All calculations were carried out on the Gould 32/8705 at ICQEM (Pisa).

Registry No. 1, 123-39-7; **2**, 107-31-3.

---

(10) Chirlian, L. E.; Miller Francl, M. *J. Comput. Chem.* **1987**, *8*, 894.
(11) Vignê-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934.
(12) Abraham, R. J.; Hudson, B. D.; Kermode, M. W.; Mines, J. R. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84*, 1911.
(13) Drummond, M. L. J. *J. Chem. Phys.* **1988**, *88*, 5014, 5021.