## COMMUNICATIONS

# On the influence of the water electrostatic field on the amide group vibrational frequencies

Petr Bouř[a)]
*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610, Praha 6, Czech Republic*

(Received 17 May 2004; accepted 3 September 2004)

For clusters of N-methylacetamide and water molecules the performance of the fixed-charged approximation was tested against continuum and explicit *ab initio* models. The dispersion of the vibrational frequencies when constant electrostatic potential was maintained at the solute atoms was compared to the distribution caused by geometry fluctuations. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1810138]

The role of the electrostatic field of the solvent on physical properties of the solute can be hardly underestimated. For bigger systems, the solvent molecules often cannot be included in the theoretical models individually and must be approximated implicitly. Such approximations span from the simplest dipolar model[1] to dielectric[2] or conductor-like[3] models that include the solute cavity more faithfully. In an alternative approach, solvent is put in regular grids,[4] replaced by reference sites[5] or a continuous distribution.[6] The need to approximate interactions between molecular parts by simple electrostatics is also encountered in *ab initio* modeling of larger molecules.[7]

Many simplified force fields have been proposed for the polar aqueous environment, allowing one to accurately reproduce water structure and thermodynamic properties.[8–11] Similarly, reasonably faithful descriptions of vibrational properties of systems in liquid water were achieved by combinations of molecular dynamics (MD) with quantum electronic structure computations.[12–17] In order to obtain the realistic inhomogeneous linewidths, many MD configurations have to be averaged within a reasonable time. For this purpose, empirical corrections to vacuum vibrational frequencies were proposed, based on solvent electrostatic potential on the solute nuclei,[14,18] or on a projection of the electric field vector onto vibrating bonds.[13] Although in these models the solvent is approximated by point charges only, they provide reasonable band frequencies and inhomogeneous line widths.

However, they were developed for isolated vibrations (the O—H, O—D, or C═O bond stretching) and their reliability for less-localized vibrational modes and against errors stemming from the point-charge approximation was not tested. In this work, spectra of clusters of N-methylacetamide (NMA) with water molecules calculated absorption and with the point charges are analyzed. The NMA molecule was chosen because of the importance of the amide chromophore for the vibrational spectroscopy of peptides and proteins.[19–21] Particularly, answers to the following questions were aimed at: Are the partial charge fits generally applicable for the aqueous systems? What is their performance in comparison with the self-consistent continuum models? Can they be used for any vibrational mode? What is the influence of the charge-transfer on the vibrational properties of hydrogen-bonded group? Finally, what are the consequences for empirical solvent modeling.

In the spirit of previous works,[13,15,18,22] a ''benchmark'' cluster of NMA surrounded by 23 water molecules was created, on the basis of a random MD configuration created by the TINKER[23] program with the Amber[24] force field. The size was chosen so that the first and second hydration spheres incorporating most of the solvent effects[15] were included. Note that the main purpose of this work is the comparison and reliability test of the computational models, while the sampling of clusters chosen from the MD trajectory is not sufficient to give reliable values of frequencies, which can be only roughly related to the experiment. Also, thousands of water configurations have to be averaged in order to obtain a smooth inhomogeneous profile, as was shown in Ref. 15. The absorption spectrum (trace d in Fig. 1) was calculated at the BPW91[25,26]/6-31G** level using the GAUSSIAN[27] and QGRAD[28] programs for the partial optimization in the normal coordinates. Modes with harmonic frequencies within ($-300$,300 $cm^{-1}$) were kept frozen. The atomic polar tensors were set to zero for the water atoms, so that the solvent did not contribute to the spectrum. Following the experimental practice,[21,29] acidic hydrogens were replaced by deuteria in order to simplify the amide I signal. As can be seen in Fig. 1, the frequencies and intensities calculated in vacuum at the same level (trace a) are significantly modified upon the influence of the water shell. There are minor changes in the frequencies of the C–H stretching modes (2970–3160 $cm^{-1}$), while their intensities are redistributed in favor of the lowest-frequency 2974 $cm^{-1}$ band. The N–D stretching (''amide A'') mode shifts from 2620 $cm^{-1}$ in vacuum to 2502 $cm^{-1}$ and its intensity increases drastically. The amide I mode is affected by formation of the hydrogen bonds simi-

a)Electronic mail: bour@uochb.cas.cz

 

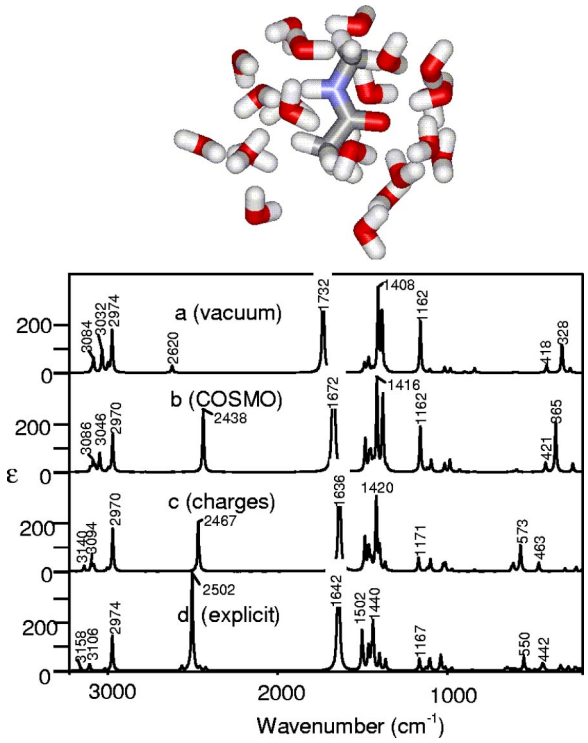


FIG. 1. Calculated absorption spectra of N-methylacetamide in vacuum (a), using the COSMO solvent model (b), the atomic point charges (c), and the explicit water molecules (d). Deuterated system is modeled (with N–D and $D_2O$), geometry of d and c is displayed above.

larly as amide A, which results in a downshift of the vacuum frequency from 1732 to 1642 $cm^{-1}$. As discussed before,[15,21] the amide I band has a paramount importance for protein absorption spectroscopy, and its solvent shift is roughly proportional to the permitivity of the solvent. Experimentally, the absorption maximum was found at 1623–1626 $cm^{-1}$ for water solutions.[21,29] The weak amide II signal calculated in vacuum at 1488 $cm^{-1}$ shifts upward in the cluster (to 1502 $cm^{-1}$) and its increased intensity becomes comparable to the C–H bending signal centered around 1440 $cm^{-1}$. This is roughly in agreement with two absorption maxima of similar intensity (at 1513 and 1494 $cm^{-1}$) observed experimentally in this region.[29] The relatively high deviation between the computed frequency of the C–H bending and experiment can probably be attributed to the lack of the statistical averaging in the cluster, needed for the bulk system (NMA in $D_2O$). The intensity of the amide III mode (1167 $cm^{-1}$, trace d) is weaker than for vacuum, but its frequency is not affected much by the solvent. A very interesting phenomenon can be observed for the ND out of plane bending mode, since its vacuum frequency calculated at 328 $cm^{-1}$ rises to 550 $cm^{-1}$ in the cluster, thus surpassing another mode characteristic for the amide chromophore, composed mainly from the C–C–N and C–N–N bending (418 in vacuum, 442 $cm^{-1}$ in the cluster). The unusually volatility of the NH (ND) out-of-plane wagging mode was noticed also for the formamide molecule, where its vibrational frequency varied within 126–841 $cm^{-1}$ according to the environment of the amino ($NH_2$) chromophore.[30]

Because of the complexity of the changes in the spectrum, it is rather surprising that the simpler point charge

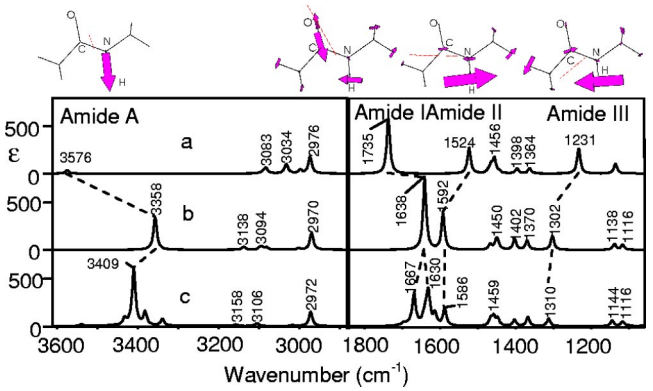


FIG. 2. The amide A, I, II, and III absorption bands of (nondeuterated) NMA calculated in vacuum (a), and for the charge (b), and water (c) clusters. The normal mode displacements are displayed above the spectra, the dashed lines follow relative magnitudes and directions of the transition dipole moments.

model (trace *c* in Fig. 1) reproduces most of these frequency and intensity solvent effects, even more faithfully than the self-consistent continuum COSMO[31] model. (Strictly speaking, MD configurations–averaged value should be compared to COSMO; however, this is not a source of its poorer performance.) For example, the amide A downshift is overestimated by 64 $cm^{-1}$ (trace b), only about 44% of the amide I frequency shift is reproduced, the change of the amide III mode is not indicated, and neither is the switching of the two lowest-frequency modes. On the contrary, the electrostatic field modeled with the point charges reproduces most of the observed changes better. This is good news for the modeling, since the latter approach is not iterative (self-consistent) and neglects the charge transfer occurring during the formation of the hydrogen bonds.[32] The charge distribution can be easily treated by commercial software[27] and the computations are much faster than using explicit molecules.

Extended tests (not shown here) confirmed that the point-charge electrostatic field reproduces the solvent frequency shifts sufficiently well for NMA-water clusters of variable sizes and shapes. For amide A, I, and II modes, an overall (RMS) error of 7.7 $cm^{-1}$ was found between the *ab initio* and point charge model for a set of 20 clusters. Additionally, the size-dependence was tested with clusters of variable size of the water shell (3–12 Å thick). This test confirmed previous findings,[15] that the first and second hydration shell are crucial, although when the thickness of the shell increased from 10–12 Å, the amide A, I, and II frequencies still changed by up to 6, 2, and 7 $cm^{-1}$, respectively.

In Fig. 2, the vacuum, charge, and water-cluster NMA spectra are modeled for natural, nondeuterated amide chromophore. Similarly, as for the deuterated case, the amide A mode shifts down to 3409 $cm^{-1}$, which is slightly overestimated by the charge model, while the C–H modes are less affected. However, the amide mode splits into two bands (at 1667 and 1630 $cm^{-1}$) in the water cluster. This is caused by the coupling of the C═O stretching and H–O–H water bending modes, and corresponds to experimental observations, where a slightly smaller split of about 20 $cm^{-1}$ was reported, centered around 1620–1625 $cm^{-1}$.[21,29] This seems like the biggest handicap of the fixed charge modeling, as it

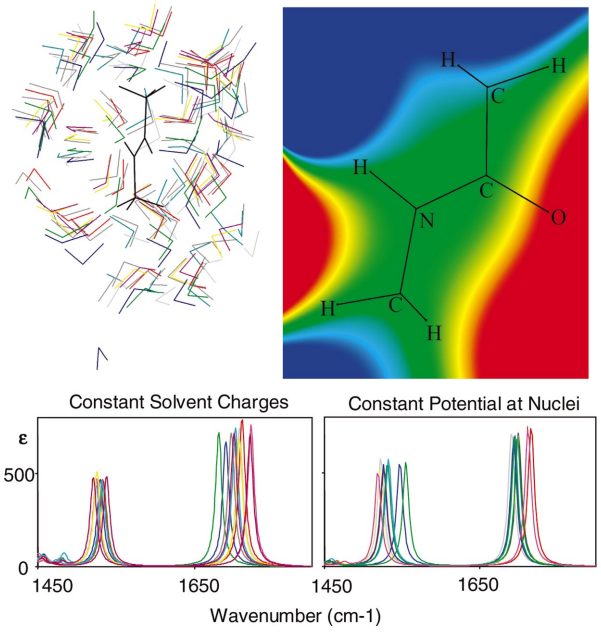


FIG. 3. (Color) Absorption spectra of NMA (deuterated) with 11 randomly selected configurations of the charges. Left: the charges were kept constant ($-0.84$ for O and 0.42 for H). Right: the electrostatic potential at the NMA nuclei was kept constant. The configurations and the averaged potential are displayed at the top. The color scale used in the isopotential plot covers approximately the region of ($-0.1$,...,0.1 a.u.).

cannot account for the resonance vibrational effects. Nevertheless, the average frequency shift is predicted correctly, not only for amide I, but also for the amide II and amide III modes. For the latter two modes, the water causes a blueshift of the vacuum frequencies and the calculated wave numbers (1586 and 1310 $cm^{-1}$) for this cluster nicely correspond to the observations ($\sim$1580 and 1317 $cm^{-1}$).

Next, we want to answer the question whether the electrostatic potential measured at the nuclei of the solute is solely responsible for the vibrational frequency shifts. This was a premise in the aforementioned combined MD-electronic structure approaches.[13–15,18] As the effect of the external field (without solvent) on vibrational properties is still rather a complex phenomenon, including geometry-rearrangement, polarization effect, and direct modification of the molecular potential energy surface, we investigate two extreme cases: (i) when constant partial charges move around the NMA molecule and (ii) when it is the potential at the solute NMA atoms that remains constant under this movement. For the first ''classical'' case, eleven random MD configurations were selected, created with the Tinker software and Amber force field, and separated by 1000 1 ps steps. Then 20 water molecules closest to the NMA were taken into the clusters, replaced by atomic partial charges ($-0.84/0.42$), and their spectra were calculated at the same manner as for the benchmark cluster in Fig. 1. The geometry dispersion of the clusters together with the simulated spectra can be seen at the left-hand side of Fig. 3 where the amide I and II regions are displayed.

The constant potential conditions were more difficult to realize. For comparison with the previous case, a set of electrostatic potentials $\{\varphi_j, j=1,...,NAT\}$ on NMA atoms coming from the water partial charges $\{Q_i\}$ was obtained as the average over the 11 configurations. (The averaged potential, however, was already almost the same as for an infinite number of configurations.) The potential map in Fig. 3, upper-right part, indicates nicely the role of the polar atoms in the amide group of NMA in the rearrangement of the solvent. Since the water molecules form hydrogen bonds to the carbonyl oxygen, their preferred orientation gives a rise to the positive (red in Fig. 3) mean solvent electrostatic field in the vicinity of this atom. Similarly, water oxygens are predominantly oriented towards NH, causing a net negative (blue) solvent field.

Next, for each configuration, new charges $\{q_i\}$ were calculated, providing the same potentials,

$$\varphi_j = \sum_{i=1}^{NC} \frac{q_i}{r_{ij}}, j=1,...,NAT, \tag{1}$$

where $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$ is the distance between nucleus $j$ and the charge $q_i$. A minimal change of the original Tinker charges and a charge conservation were required,

$$S = \tfrac{1}{2}\sum_{i=1}^{NC} (q_i - Q_i)^2 \rightarrow \min, \tag{2}$$

$$\sum_{i=1}^{NC} q_i = 0. \tag{3}$$

Introducing the Lagrange multipliers $\{\lambda, \omega_j, j=1,...,NAT\}$, we obtain

$$q_l + \sum_{j=1}^{NAT} \omega_j \frac{1}{r_{lj}} + \lambda = 0, \quad \text{for } l=1,...,NC. \tag{4}$$

The formulas 1, 3, and 4 form a set of $NC+NAT+1$ linear equations for the same number of variables $\{\lambda, \omega_j, q_i\}$ and can be solved via a matrix inversion. Thus charges providing the averaged potential can be obtained for any geometry. The NMA molecule surrounded by these charges was optimized with modified normal mode optimization technique[28] so that the partial charges were actualized in each step and the potentials at the nuclei did not change.

The dispersion of the absorption spectra caused by the different distribution of the charges while the electrostatic potential at the nuclei is the same can be seen at the right-lower part of Fig. 3. In other words, the dispersion is caused only by the field intensity and its higher-order derivatives. Apparently, the same-charge and same-potential dispersions are roughly comparable in this spectral region (c.f., the two spectra sets in Fig. 3), and the frequency fluctuations within 20–50 $cm^{-1}$ are consistent with the experimental linewidths.[15,21,29] Similar dispersion was observed for other modes and is not shown. For the amide I mode the dispersion under the constant potential is markedly smaller and for nine of the 11 clusters lies even within few $cm^{-1}$. However, the constant-potential variable-geometry condition leads to larger splitting of the amide II signal. This can explain previous successes of the potential/field corrections, as they were applied rather to the localized O–H and C═O modes.[13,15,18] Nevertheless, the large dispersion of the constant-potential spectra in Fig. 3, as well as its mode-

dependence, suggest that such an approach cannot be applied universally. On the other hand, as the electrostatic models seem to reproduce the solvent effect well (Figs. 1 and 2), at least for the extremely polar aqueous environment, the approximation of the solvent by the charged points seems to be a solely convenient alternative to the explicit solvent computations.

We may thus conclude that, for the NMA/$H_2O$ (and probably for similar highly-polar) systems, the point charge approximation faithfully reproduces the influence of the solvent on the solute. This can certainly be conveniently used for speeding up the computations, especially in combined MD/*ab initio* studies where large amounts of the configurations have to be studied in order to obtain realistic spectral profiles or other physical properties. The limitation of the continuum solvent model, although statistically providing correct trends, was found in qualitative agreement with previous works.[4,12,13,15,18] The limit is most probably not given by the neglect of the charge-transfer, since this is also not included in the point charge model but rather by the inherent inability to describe the directional, space-oriented hydrogen bonds and the higher charge densities localized on solvent atoms. Finally, we find that the potential/field empirical fit of the solvent shift of the vibrational frequencies should be used cautiously because of the complicated nonlinear relations between the geometry, frequencies, and solvent field, and perhaps the importance of the higher-order potential derivatives.

The work was supported by the Grant Agency of the Academy of Sciences (A4055104).

[1] M. W. Wong, M. J. Frisch, and K. B. Wiberg, J. Am. Chem. Soc. **113**, 4776 (1991).
[2] V. Barone, M. Cossi, and J. Tomasi, J. Comput. Chem. **19**, 404 (1998).
[3] F. Eckert and A. Klamt, AIChE J. **48**, 369 (2002).
[4] J. Florián and A. Warshel, J. Phys. Chem. B **101**, 5583 (1997).
[5] H. Sato and S. Sakaki, J. Phys. Chem. A **108**, 1629 (2004).
[6] S. Ten-no, F. Hirata, and S. Kato, J. Chem. Phys. **100**, 7443 (1994).
[7] T. E. Exner and P. G. Mezey, J. Phys. Chem. A **108**, 4301 (2004).
[8] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[9] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **114**, 363 (2001).
[10] A. Glättli, X. Daura, and W. F. van Gunsteren, J. Chem. Phys. **116**, 9811 (2002).
[11] H. Yu., T. Hansson, and W. F. Gunsteren, J. Chem. Phys. **118**, 221 (2003).
[12] P. Bouř, Chem. Phys. Lett. **365**, 82 (2002).
[13] S. A. Corcelli, C. P. Lawrence, and J. L. Skinner, J. Chem. Phys. **120**, 8107 (2004).
[14] K. Kwac and M. Cho, J. Chem. Phys. **119**, 2247 (2003).
[15] P. Bouř and T. A. Keiderling, J. Chem. Phys. **119**, 11253 (2003).
[16] S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth, J. Chem. Phys. **120**, 10896 (2004).
[17] C. H. Langley and N. L. Allinger, J. Phys. Chem. A **107**, 5208 (2003).
[18] S. Ham, J. H. Kim, H. Kochan, and M. Cho, J. Chem. Phys. **118**, 3491 (2003).
[19] X. G. Chen, R. S. Stenner, S. A. Asher, N. G. Mirkin, and S. Krimm, J. Phys. Chem. **99**, 3074 (1995).
[20] S. Iuchi, A. Morita, and S. Kato, J. Phys. Chem. B **106**, 3466 (2002).
[21] J. Kubelka and T. A. Keiderling, J. Phys. Chem. A **105**, 10922 (2001).
[22] K. J. Jalkanen, R. M. Nieminen, K. Frimand, J. Bohr, H. Bohr, R. C. Wade, E. Tajkhorshid, and S. Suhai, Chem. Phys. **265**, 125 (2000).
[23] R. V. Pappu, R. K. Hart, and J. W. Ponder, J. Phys. Chem. B **102**, 9725 (1998).
[24] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[25] A. Becke, Phys. Rev. A **38**, 3098 (1988).
[26] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[27] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision A.1, Gaussian, Inc., Pittsburgh, PA, 2003.
[28] P. Bouř and T. A. Keiderling, J. Chem. Phys. **117**, 4126 (2002).
[29] S. Song, S. A. Asher, S. Krimm, and K. D. Shaw, J. Am. Chem. Soc. **113**, 1155 (1991).
[30] C. N. Tam, P. Bouř, J. Eckert, and F. R. Trouw, J. Phys. Chem. **101**, 5877 (1997).
[31] V. Barone, M. Cossi, and J. Tomasi, J. Comput. Chem. **19**, 404 (1998).
[32] J.-H. Guo, Y. Luo, A. Augustsson, J.-E. Rubensson, C. Såthe, C. Ågren, H. Siegbahn, and J. Nordgren, Phys. Rev. Lett. **89**, 137402 (2002).