Theor Chem Acc (2000) 105:7–14
DOI 10.1007/s002140000180

Theoretical
Chemistry Accounts

*Regular article*

# The inclusion of electron correlation in intermolecular potentials: applications to the formamide dimer and liquid formamide

**Steve Brdarski**[1], **Per-Olof Åstrand**[2], **Gunnar Karlström**[1]

[1] Department of Theoretical Chemistry, Center for Chemistry and Chemical Engineering, University of Lund, P.O. Box 124, 22100 Lund, Sweden
[2] Condensed Matter Physics and Chemistry Department, Riso National Laboratory, P.O. Box 49, 4000 Roskilde, Denmark

Received: 20 March 2000 / Accepted: 18 April 2000 / Published online: 18 August 2000


**Abstract.** A test of the quality of the electrostatic properties and polarizabilities used in the nonempirical molecular orbital (NEMO) potential is carried out for formamide by calculating the molecular dipole moment and polarizability at the second-order Møller–Plesset (MP2) level of theory. The molecular dipole moment is 11% lower at the MP2 level than at the Hartree–Fock (HF) level, whereas the isotropic part of the polarizability is increased by 36% by adding electron correlation and using a considerably larger basis set. The atomic charges, dipole moments and polarizabilities obtained at the HF level are rescaled to get the correct molecular properties at the MP2 level. The potential minimum for the cyclic dimer of formamide is −17.50 kcal/mol with the MP2-scaled properties and is significantly lower than other potentials give. Two intermolecular potentials are constructed and used in subsequent molecular dynamics simulations: one with the regular NEMO potential and the other with the rescaled MP2 properties. A damping of the electrostatic field at short intermolecular distances is included in the present NEMO model. The average energies for liquid formamide are lower for the MP2-scaled model and are in good agreement with experimental results. The lowering of the simulation energy for the MP2-scaled potential indicates the strong dispersive interactions in liquid formamide.

**Key words:** Dispersion interaction – Hydrogen bonds – Molecular dynamics – Polarizable potentials

## 1 Introduction

A key issue underlying the success of simulation methods is the accuracy of the potential functions describing the interactions in the systems modeled. The demand of a detailed description of the interaction between atoms in molecules and the requirement that the potential be fast to evaluate put severe restrictions on the potential model to be used [1–5]. Especially in molecular simulations, Monte Carlo and molecular dynamics (MD), efficient calculations of potentials and forces are of the utmost importance. Historically, intermolecular potentials have mostly been evaluated with Lennard-Jones types of potential functions.

$$E_{AB} = 4\varepsilon_{AB}\left[\left(\frac{\sigma}{\tau}\right)^{12} - \left(\frac{\sigma}{\tau}\right)^{6}\right] \quad (1)$$

This model is very simple. $\varepsilon$ is the depth of the potential well and $\sigma$ is the size of the atom. However, the model is not sufficient to describe the interaction between two water molecules because water is a polar molecule with a large permanent dipole moment and the Lennard-Jones model does not take into account the effects of the electrostatics and polarization of the interacting species. It has proved necessary to include this effect, especially when water is involved in the simulations, and nowadays there are many potential models which take electrostatic and polarization effects into account [1–7].

Even if the interaction potential is modeled reliably, the accuracy of the atomic properties also has to be considered. The charge distribution and response properties are important features of interacting molecules. The charge distribution is normally analyzed according to a generalized Mulliken population analysis scheme, which in some sense is arbitrary. It depends strongly on the basis set used and must be carefully balanced for the atoms in the molecule [8–10]. In spite of the problems, the population analysis is a useful interpretive device when used properly. The potential model used in this work is the nonempirical molecular orbital (NEMO) force field model [5, 11–13]. The charge distribution in the NEMO model is represented by a multicenter multipole expansion

*Correspondence to*: G. Karlström,
e-mail: gunnar.kalstrom@teokem.lu.se

(MME) distributed over the atoms in the molecules [14], while the response properties of the molecules considered are described with atomic polarizabilities [15, 16]. These atomic properties are extracted from Hartree–Fock (HF) Self-consistent-field (SCF) wave functions of each of the monomers and are known to give too large a dipole moment owing to the lack of antibonding configurations and too low polarizabilities owing to the lack of electron correlation and limited basis sets [17–20]. The deficiency in the polarizability has a major effect on the dispersion energy, whereas the polarity overestimates the electrostatic interaction. In the calculation of intermolecular interaction energies with the NEMO model, these errors seem to cancel each other. Nonetheless, for some systems we have observed difficulties in describing the interactions accurately [5], and the problems are most probably because the electrostatics is not modeled accurately enough. The question then arises of how large are the errors in the NEMO potential by describing the molecular properties with HF SCF monomer wave functions and with the basis sets used? We have seen from large calculations on the water molecule that the polarizability is increased by 35% when the basis set is increased from 24 contracted basis functions to 523 at the HF level of theory [18]. In practical calculations such large basis sets are not feasible for larger molecules on today's computers, but it is seen to be necessary to include large basis sets if a detailed description of the charge and polarizability distributions is desired. The remaining part of the interaction energy, the exchange repulsion, is evaluated as described elsewhere [5].

In this work, we investigate the quality of the charge and polarizability distributions used in the NEMO model by calculating the molecular dipole moment and polarizability at the second-order Møller–Plesset (MP2) level of theory [21] and with a larger basis set than has been used previously in the NEMO approach. Two potentials are constructed from these calculations and are used in subsequent MD simulations: one with atomic properties from the standard NEMO model and the other where the MME and the polarizabilities are scaled using ratios of the values obtained from the MP2 and the SCF calculations. An appropriate test system is the formamide molecule since formamide is a highly polar molecule with large polarizability and one expects electrostatic interactions to play a dominant role. The N—H···O=C bond of formamide makes it a suitable model molecule in the study of breaking and forming peptide linkages occuring in biological systems. Formamide is also one of the molecules we had difficulties with previously [5].

This article is organized as follows. The next section describes the NEMO model used in this work. A damping term in the electrostatic and induction energies at short interatomic distances is included. In Sect. 3 we give the computational details, both for the construction of the potentials and the MD simulations. The results and discussions are presented in Sect. 4, and finally in Sect. 5 the conclusions are given.

## 2 Interaction potential

### *2.1 The NEMO model*

The basis of the NEMO method is the partitioning of the interaction energy at the HF level of theory into first- and second-order perturbation terms [11, 22]. Formally the interaction energy is divided into physically recognizable terms that are supposed to add up to the SCF energy according to

$$\Delta E_{\mathrm{SCF}} = E_{\mathrm{ele}} + E_{\mathrm{ind}} + E_{\mathrm{rep}} \ , \tag{2}$$

where $E_{\mathrm{ele}}$ is the electrostatic energy, $E_{\mathrm{ind}}$ is the induction energy and $E_{\mathrm{rep}}$ is the remaining part, which mainly consists of the exchange repulsion energy. The advantage of this energy partitioning scheme is, apart from the physical interpretation, that each part can be modified and systematically improved upon, and this can be done independently of the other parts. To obtain the total interaction energy, a dispersion term according to London [23, 24] is added to Eq. (2).

$$E_{\mathrm{tot}} = \Delta E_{\mathrm{SCF}} + E_{\mathrm{disp}} \tag{3}$$

This term describes the energy contribution to the intermolecular interactions from the correlated motion of the electrons not available in the HF formalism. In the NEMO model, each of these energy contributions is calculated from properties obtained from the SCF wave function of the monomers. The electrostatic energy is evaluated as a sum of interactions over the expansion centers, while the induction energy is calculated by letting the multipoles interact with atomic polarizabilities. The MME in NEMO is truncated at the quadrupole moment level. The advantage of using MME lies in the drastic reduction in detail by using point multipoles instead of covering space by basis functions as in quantum-chemical methods. For a full description of the procedure used to obtain an MME, we refer to Refs. [14, 25].

The repulsion energy is modeled with an exponential function as

$$E_{\mathrm{rep}} = \sum_{ij} a_{ij} \mathrm{e}^{-b_{ij} r_{ij}} \ , \tag{4}$$

with $a_{ij} = q_i^v q_j^v \kappa_i \kappa_j$ and $b_{ij} \sim 1/(\beta_i + \beta_j)$ and where $\kappa$ and $\beta$ are fitted repulsion parameters and $q^v$ is the valence charge of the atoms. The short-range repulsion term used in previous versions of the NEMO model [5] has not been used here. The dispersion is modeled with a London-type formula as [11, 12]

$$E_{\mathrm{disp}}^{\mathrm{Lon}} = -\sum_{mn}^{\mathrm{atoms}} f_{mn} \frac{CE_{12}}{4} \sum_{ijkl}^{3} \alpha_{ij}^{m} \alpha_{kl}^{n} T_{ik} T_{jl} \ , \tag{5}$$

where $\alpha$ is a component of the atomic polarizability tensor, $T_{ik}$ a component of the interaction tensor, $\nabla\nabla(1/r)$, $C$ a constant with the value 1.89 calibrated from an MP2 calculation [26], $E_{12}$ the average molecular excitation energy and $f_{mn}$ a site–site damping function, introduced according to Tang and Toennies [27] in order to estimate the effects from the overlap of the charge distributions. For more details on the recent develop-

ments of the NEMO model, we refer to Ref. [5]. A new set of repulsion parameters is constructed and is shown in Table 1. The repulsion parameters are taken from previous work [28] and were reoptimized using isotropic polarizabilities. The use of isotropic polarizabilities is required by the scheme chosen for the inclusion of damping of the electrostatic field. We want the damping to be equal in all directions and the procedure is described later. The damping was not included in the fitting of the repulsion parameters. As seen from Table 1 in the previous work [28] compared with Table 1 in this article, the use of isotropic polarizabilities does not change the parameters very much. The interaction sites in the energy expressions are the centers of charge of the atoms in the molecule. The motivation for the choice of these centers instead of the atomic centers normally used is given in Ref. [5]. The advantage of the NEMO approach, in its present version, is that the parameters used need to be computed only once for a given monomer and that no additional fitting has to be done. In that sense, the NEMO model is generally applicable to all kinds of system.

### *2.2 Inclusion of a damping term in the electrostatic and induction energies*

The atomic induced dipole moment, $\mu_{p,\alpha}^{\text{ind}}$, is given as

$$\mu_{p,\alpha}^{\text{ind}} = \alpha_{p,\alpha\beta}\left(E_{p,\beta}^{\text{perm}} + \sum_{q\neq p}^{N} T_{pq,\beta\gamma}^{(2)}\mu_{q,\gamma}^{\text{ind}}\right) \ , \tag{6}$$

where $\alpha_{p,\alpha\beta}$ is the atomic polarizability, $E_{p,\beta}^{\text{perm}}$ is the electric field from the permanent electrostatic moments of the other molecules and $\sum_{q\neq p}^{N} T_{pq,\beta\gamma}^{(2)}\mu_{q,\gamma}^{\text{ind}}$ is the electric field from the induced dipole moments of the surrounding molecules. Thus, the atomic induced dipole moments are calculated in a self-consistent procedure which may diverge at short intermolecular distances. Here, this problem is avoided by including a damping term according to Thole [29], which was originally adopted in a method for calculating molecular polarizabilities from a set of interacting atoms. The Thole model has been investigated in more detail recently [30–32] and it has also been applied to simulations of water clusters and ice [33]. Thole modified the so-called interaction tensors as [29]

$$T_{pq}^{(0)} = \frac{v_{pq}^4 - 2v_{pq}^3 + 2v_{pq}}{R_{pq}} \ , \tag{7}$$

$$T_{pq,\alpha}^{(1)} = \frac{-(4v_{pq}^3 - 3v_{pq}^4)R_{pq,\alpha}}{R_{pq}^3} \ , \tag{8}$$

**Table 1.** Repulsion parameters with isotropic polarizabilities. The error of the fit as defined by Eq. (24) in Ref. [5] is 4.05

| | C | O | N | H |
|---|---|---|---|---|
| $\beta$ (au) | 0.4897 | 0.2797 | 0.4101 | 0.0668 |
| $\kappa$ (kcal/mol) | 0.3229 | 12.8868 | 1.8366 | 753.4707 |

and

$$T_{pq,\alpha\beta}^{(2)} = \frac{3v_{pq}^4 R_{pq,\alpha}R_{pq,\beta}}{R_{pq}^5} - \frac{(4v_{pq}^3 - 3v_{pq}^4)\delta_{\alpha\beta}}{R_{pq}^3} \ , \tag{9}$$

where $R_{pq,\alpha}$ is a component of the distance vector between atoms $p$ and $q$ and $v_{pq} = R_{pq}/s_{pq}$ if $R_{pq} < s_{pq}$. Otherwise, $v_{pq} = 1$ and the regular expressions are obtained. Thole defined $s_{pq}$ as

$$s_{pq} = 1.662(\alpha_p\alpha_q)^{\frac{1}{6}} \ , \tag{10}$$

The parameter 1.662 is optimized from a set of molecules to describe the molecular polarizability and contains a parameterization of only atomic polarizabilities. This factor is reparameterized in the NEMO model since we also have atomic charges and dipole moments, which give a large electric field. Therefore, $s_{pq}$ is multiplied by a factor of 1.4, which was obtained from a calibration of the $Cl^-$-$H_2O$ complex. The damping suggested by Thole is especially suitable for MD simulations since the modified interaction tensors are the only modifications required to obtain the interaction energies and forces for the electrostatic and induction terms. Furthermore, the Thole damping does not add any additional atomic parameters to the model. Instead the steep $r^{-n}$-repulsion parameters used previously [5, 11] to stop the molecules coming too close to each other are not required anymore; thus the number of total potential parameters has been reduced.

## 3 Computational details

The geometry for the planar formamide monomer used in this work is taken from the work of Kitano and Kuchitsu [34] and is kept rigid. A pyramidal structure with the amino group about 15° off the plane has been observed from electron diffraction experiments [35]. However, MP2 calculations give only a small energetic preference for the pyramidal structure [36]. The choice between the planar and the nonplanar structure could be important for vibrational analysis but it does not play a role in this investigation.

The coordinates for the atomic centers and for the charge centers of the molecule are shown in Table 2. The atomic properties – charges, dipoles, second moments and polarizabilities – are given in Table 3. Note that the second moments in Table 3 are explicitly used in the evaluation of the repulsion energy, but are modeled with dipoles for the description of the electrostatic interaction. To compare how well the SCF calculations represent the atomic properties with the NEMO basis set, the molecular dipole moment and the polarizability of formamide were calculated at the MP2 level of theory within the finite-field perturbation scheme [17]. The standard NEMO model [5] uses the smaller atomic natural orbitals (ANO-S) basis set [37]. For the carbon, nitrogen and oxygen atoms, a [10/6/3] primitive basis was contracted to (4/3/2). For hydrogen atoms bonded to polar atoms, such as nitrogen and oxygen, a [6/4] set contracted to (3/2) was used and for hydrogen atoms bonded to nonpolar atoms, such as carbon, we used [6/3] contracted to (3/1). The basis set used for formamide in the MP2 calculations was the larger contracted atomic natural orbitals (ANO-L) basis set [38], H (4s3p2d), C (5s4p3d2f) and O (5s4p3d2f), which gives 351 primitive basis functions contracted to 207 (NEMO gives 274 contracted to 108). The basis set was chosen to match ANO-6 in Ref. [18] and is probably not large enough to saturate the polarizability, but we expect that the size of the basis set is good enough for our purposes. We applied an electric field in the $\pm x$, $y$, and $z$-directions of 0.001 au. All ab initio calculations were performed using the MOLCAS package [39].

**Table 2.** Internal coordinates (Å) for the atomic center and center of charge of the atoms in formamide

| | Atomic center | | | Center of charge | | |
|---|---|---|---|---|---|---|
| | *x* | *y* | *z* | *x* | *y* | *z* |
| $H_1$ | 0.000 | −0.689 | −1.370 | 0.000 | −0.592 | −1.225 |
| C | 0.000 | −0.253 | −0.330 | 0.000 | −0.264 | −0.366 |
| O | 0.000 | −0.946 | 0.664 | 0.000 | −0.944 | 0.668 |
| N | 0.000 | 1.115 | −0.330 | 0.000 | 1.116 | −0.329 |
| $H_2$ | 0.000 | 1.625 | −1.223 | 0.000 | 1.577 | −1.145 |
| $H_3$ | 0.000 | 1.610 | 0.571 | 0.000 | 1.591 | 0.500 |

**Table 3.** Atomic charges, dipoles, second moments and isotropic polarizabilities (au) for formamide. $\alpha$ is the isotropic polarizability. The properties were obtained from the standard nonempirical molecular orbital (NEMO) procedure

| | *q* | $\mu_x$ | $\mu_y$ | $\mu_z$ | $\Omega_{xx}$ | $\Omega_{yy}$ | $\Omega_{zz}$ | $\alpha$ |
|---|---|---|---|---|---|---|---|---|
| $H_1$ | 0.057 | 0.000 | −0.195 | −0.291 | 1.049 | 1.656 | 3.996 | 2.234 |
| C | 0.411 | 0.000 | 0.088 | 0.281 | 4.019 | 1.647 | 3.332 | 2.999 |
| O | −0.499 | 0.000 | −0.034 | −0.054 | 3.534 | 8.170 | 7.478 | 6.394 |
| N | −0.656 | 0.000 | −0.011 | −0.005 | 5.868 | 7.351 | 4.558 | 5.926 |
| $H_2$ | 0.327 | 0.000 | 0.071 | −0.117 | 0.690 | 2.089 | 2.335 | 1.705 |
| $H_3$ | 0.359 | 0.000 | 0.027 | 0.101 | 0.570 | 2.123 | 1.763 | 1.485 |

Two potentials were constructed for the simulations: one with atomic properties from the NEMO procedure and the other where the atomic properties were scaled by the ratio of the MP2 and SCF molecular properties. Hereafter we refer to the potential with atomic properties from the MEMO procedure as model 1 and the potential with MP2-scaled properties as model 2. The density of liquid formamide is 1.1334 g/cm$^3$ [40], which corresponds to 216 formamide molecules in a cubic box with a side of length 24.2 Å. However, with the experimental density we obtained too high an energy in a simulation for model 1. To overcome this, the length of the side of the box was gradually increased until the energy decrease leveled out. We finally chose a length of 26.0 Å, which is the breakpoint for where the energy decrease began to level out. Two simulation boxes were used: one with a length, $L$, of 24.2 Å, giving the experimental density, and one with $L = 26.0$ Å. Both potentials were tested in the two different boxes. We carried out MD simulations of 216 formamide molecules enclosed in a cubic box with periodic boundary conditions. The molecular cutoff was 12.0 Å for $L = 24.2$ Å and 12.5 Å for $L = 26.0$ Å. The simulation times were 50 ps in all cases, expect for model 1 in the box with $L = 26.0$ Å, where we simulated for 100 ps. A time step of 1 fs was used, and the systems were well equilibrated around their average energies, more than 30 ps. MD simulations were carried out using the MOLSIM [41] package, and the equations of motion were solved by using the velocity Verlet algorithm with quaternions adopted for the rotational motion [42]. The damping described previously is now implemented in the MOLSIM package. The temperature was kept around 300 K by using the velocity scaling procedure due to Berendsen et al. [43] with a time constant of 0.01 ps.

## 4 Results and discussions

### *4.1 Atomic properties and potentials*

The molecular dipole moment and the diagonal terms of the polarizability tensor of formamide are given in Table 4 for the two models. We have included the SCF results with the large basis set from model 2 for comparison.

**Table 4.** Molecular dipole moment, $\vec{\mu}$ (au) and diagonal elements of the polarizability tensor (au), $\alpha_{ii}$ for models 1 and 2. $\bar{\alpha}$ is one-third of the trace of the polarizability tensor

| | $\mu_x$ | $\mu_y$ | $\mu_z$ | $\lvert\vec{\mu}\rvert$ (D) |
|---|---|---|---|---|
| $\vec{\mu}^{\text{mod1}}_{\text{SCF}}$ | 0.000 | 1.28 | −1.07 | 4.24 |
| $\vec{\mu}^{\text{mod2}}_{\text{SCF}}$ | 0.000 | 1.27 | −1.06 | 4.21 |
| $\vec{\mu}^{\text{mod2}}_{\text{MP2}}$ | 0.000 | 1.20 | −0.89 | 3.79 |
| | $\alpha_{xx}$ | $\alpha_{yy}$ | $\alpha_{zz}$ | $\bar{\alpha}$ |
| $\hat{\alpha}^{\text{mod1}}_{\text{SCF}}$ | 15.73 | 23.04 | 23.46 | 20.74 |
| $\hat{\alpha}^{\text{mod2}}_{\text{SCF}}$ | 18.89 | 30.26 | 27.61 | 25.59 |
| $\hat{\alpha}^{\text{mod2}}_{\text{MP2}}$ | 20.64 | 34.78 | 28.68 | 28.28 |

The molecular dipole moment calculated at the SCF level of theory does not change much with increases basis set, but by including correlation effects in the MP2 calculations, we observe a decrease in the molecular dipole moment of 11% compared to model 1. $|\vec{\mu}|^{\text{mod2}}_{\text{MP2}} = 3.79$ D is in good agreement with experimental results [44, 45], $|\vec{\mu}|_{\text{exp}} = 3.71$ D, and other theoretical calculations [7, 46–48]. The isotropic polarizability is 23% higher at the SCF level in model 2 compared to model 1, and 36% higher at the MP2 level. The structures of three different potential minima of the formamide dimer and the corresponding potential surfaces for models 1 and 2 are shown in Fig. 1.

The energies for the optimized geometries are given in Table 5 and the intermolecular distances in Table 6. The structures for minima 1 and 2 are more or less the same for both models. The molecules in minimum 1 are pushed away from each other in model 2 compared to model 1 owing to reduction in the electrostatic interaction. Minimum 3 in Fig. 1c behaves differently for the two models. In model 1 the minimum is the planar linear structure, but in model 2 one of the molecules is translated along the axis perpendicular to the plane for model 1. The structure has, to our knowledge, not been observed before. The cyclic structure in Fig. 1a is the global minimum for the dimer and has been predicted by others [7, 46–49]. The energy in the potential minimum for model 2 is lower than what other groups using analytical potential functions have found for the cyclic dimer [7, 46, 47]. Cabaleiro-Lago and Rios [7] used a similar potential to NEMO and found a minimum energy of −11.13 kcal/mol. Their potential function was designed from ab initio MP2 computed molecular properties and intermolecular perturbation theory calculations involving the 6-311G** basis set. Gao et al. [47] predicted the cyclic bonding energy to be −14.0 kcal/mol, lower than Cabaleiro-Lago and Rios but still 3.5 kcal/mol higher than our energy. Their intermolecular potential is derived from ab initio calculations on bimolecular complexes using the 6-31G(d) basis set and from thermodynamic properties of liquid formamide. The 6-311G** basis set is comparable to the NEMO basis, while 6-31G(d) is smaller. Large ab initio calculations on the cyclic dimer complex have been done by Suhai [48] and Colominas et al. [49]. Suhai made extensive MP2 and fourth-order Møller–Plesset calculations using several partly highly polarized atomic basis sets of increasing size. The largest basis set used in his

MP2 calculations consisted of 189 uncontracted basis functions and 138 contracted for the dimer complex. With this basis set he found the interaction energy to be −13.04 kcal/mol. The energy is corrected for the basis set superpositions errors (BSSE) by the counterpoise correction [50]. An even larger basis set was used by Colominas et al. [49], who used an augmented correlation consistent polarized valence triple-zeta (aug-cc-pVTZ) basis set consisting of approximately 260 contracted basis functions for the dimer complex. The interaction energy (BSSE corrected) for the cyclic dimer was now lowered to −14.50 kcal/mol at the MP2 level of theory. The trend in the extensive ab initio calculations by Suhai and Colominas et al. shows that a larger basis set gives a lower interaction energy. It appears as if the bottom of the potential well is not reached with the basis set used at the MP2 level of theory. From ab initio calculations on the water dimer [18] using very large basis sets, the interaction energy did not converge at the MP2 level of theory until more than 800 basis functions were used, while the SCF energy converged after 200 contracted basis functions. The slow convergence at the MP2 level of theory was attributed to the problem of describing the dispersion interaction [18]. The basis set of Colominas et al., aug-cc-pVTZ, is of the same size as ANO-3 in Ref. [18]. The interaction energy for the water dimer with the ANO-3 basis is 0.5 kcal/mol higher than the MP2 limit (Fig. 2 in Ref. [18]). On the basis of this and the arguments in Ref. [18], a rough estimate of the interaction energy for the cyclic formamide at the MP2 limit would probably be about 1 − 1.5 kcal/mol lower in energy than the results of Colominas et al. This would give an interaction energy of about −16 kcal/mol, which is closer to our interaction energy for model 2. The basis sets used by Cabaleiro et al. and Gao et al. are probably too small to describe the dispersive interactions accurately. It is noted that the NEMO potential is general and nonempirical, while both Cabaleiro et al. and Gao et al. tuned their potentials from additional calculations for or from experimental properties of formamide.

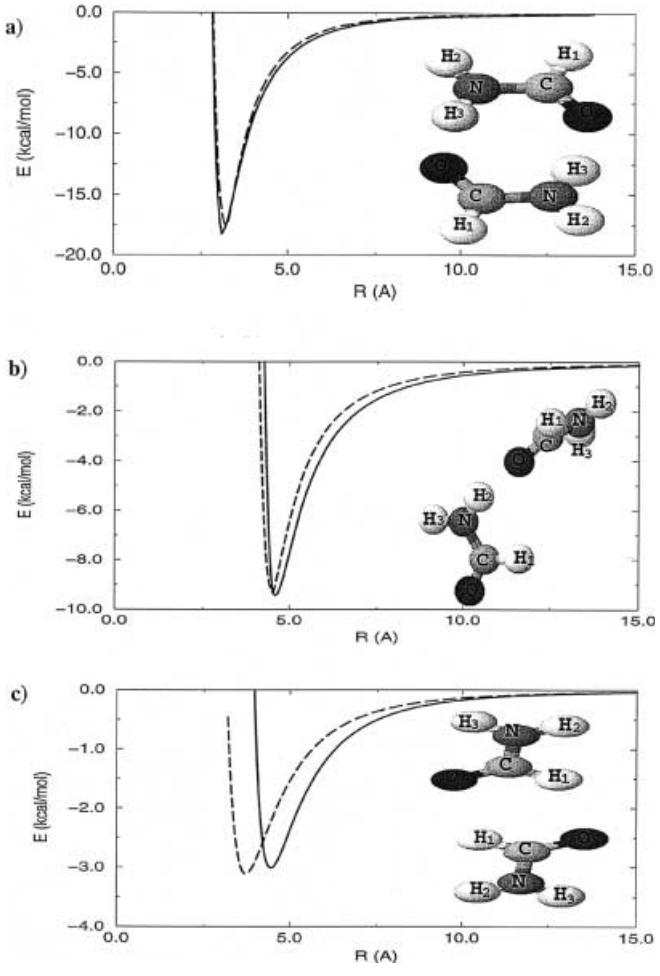


**Fig. 1a–c.** Potential surfaces and structures for three minima. **a** A cyclic structure (minimum 1), **b** a linear and twisted structure (minimum 2) and **c** a linear structure (minimum 3). The distances in the potential surfaces are between the center of mass of the molecules. The *solid lines* corresponds to model 1 and the *dashed lines* to model 2

### *4.2 Simulation results*

The simulation potential for model 1 is constructed from the atomic properties in Table 3 and for model 2 by multiplying the charges and dipoles in Table 3 by 0.89 and the polarizabilities by 1.36. The second moments of the charge distribution are not changed since the repulsion parameters in Eq. (4) are fitted from SCF energies. The average energies for liquid formamide with the two potentials in the two boxes are shown in Table 7. The potential energy from the simulation with model 2 is

**Table 5.** Energies (kcal/mol) for the different minima in Fig. 1. In **A** we compare the effect of using atomic properties from model 2 in structures optimized with atomic properties from model 1. In **B** the structures and energies are optimized with atomic properties from model 2

| | A | | | | | | B | | |
|---|---|---|---|---|---|---|---|---|---|
| | Minima 1 | | Minima 2 | | Minima 3 | | Minimum 1 | Minimum 2 | Minimum 3 |
| | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 | | | |
| $U_{tot}$ | −18.17 | −16.19 | −9.43 | −8.87 | −3.01 | −2.98 | −17.50 | −9.33 | −3.12 |
| $U_{ele}$ | −23.10 | −17.32 | −11.14 | −8.76 | −3.47 | −2.59 | −15.60 | −8.11 | −2.75 |
| $U_{ind}$ | −5.62 | −4.83 | −2.30 | −2.30 | −0.28 | −0.31 | −4.18 | −2.12 | −0.41 |
| $U_{rep}$ | 15.84 | 16.31 | 6.14 | 6.34 | 1.71 | 1.80 | 10.24 | 4.42 | 2.76 |
| $U_{disp}$ | −5.30 | −10.34 | −2.13 | −4.16 | −0.98 | −1.88 | −7.96 | −3.52 | −2.72 |

**Table 6.** Atom–atom distances and damping radius, $s$, (Å) for the shortest distances in Fig. 1. In model 1 the structures are optimized with parameters from standard NEMO, in model 2 with second-order Møller–Plesset perturbation theory scaled parameters

| | Minima 1 | | Minima 2 | | Minima 3 | | $s$ | |
|---|---|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| C—C | 3.88 | 4.01 | 4.64 | 4.56 | 3.77 | 3.32 | 1.66 | 1.75 |
| C—N | 3.59 | 3.72 | 4.02 | 4.09 | 5.56 | 4.27 | 1.77 | 1.87 |
| C—O | 3.39 | 3.49 | 3.54 | 3.52 | 3.50 | 3.14 | 1.79 | 1.89 |
| N—N | 3.80 | 3.92 | 4.79 | 4.80 | 6.47 | 5.40 | 1.86 | 1.96 |
| N—O | 2.72 | 2.84 | 2.83 | 2.92 | 4.72 | 3.89 | 1.87 | 1.97 |
| O—O | 3.29 | 3.36 | 4.75 | 4.73 | 3.64 | 3.41 | 1.88 | 1.98 |
| O—$H_1$ | 4.45 | 4.59 | 3.21 | 3.10 | 2.67 | 2.81 | 1.76 | 1.86 |
| O—$H_2$ | 3.50 | 3.63 | 1.83 | 1.94 | 4.95 | 4.13 | 1.74 | 1.84 |
| O—$H_3$ | 1.73 | 1.85 | 3.59 | 3.73 | 5.56 | 4.55 | 1.74 | 1.84 |

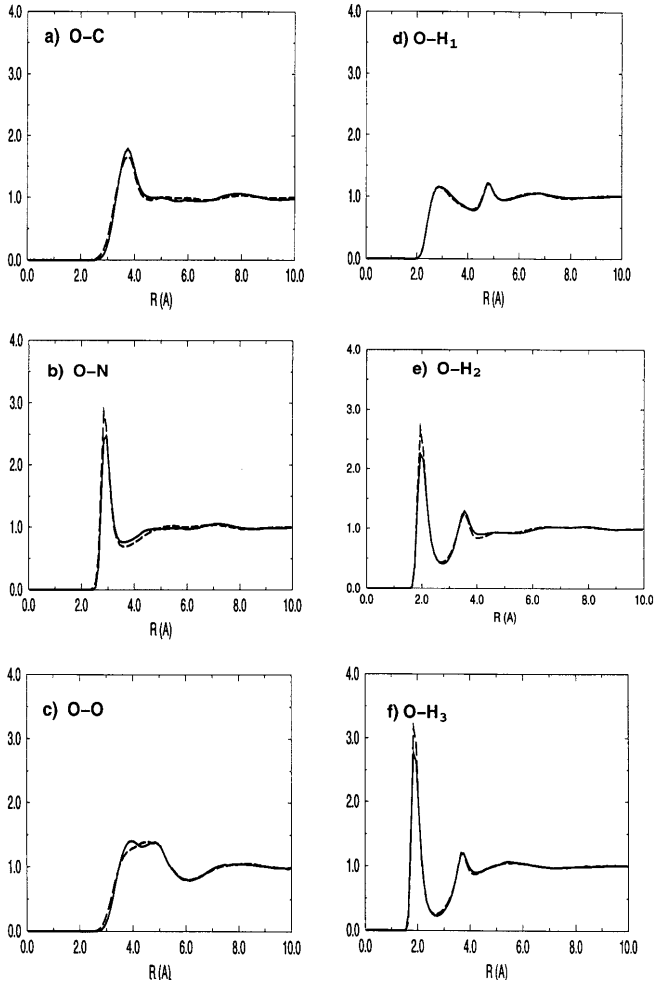


**Fig. 2a–f.** Radial distribution functions for oxygen in the simulation box with $L = 24.2$ Å. The *solid lines* correspond to model 1 and the *dashed lines* to model 2

in good agreement with experimental results (−14.13 kcal/mol [51]) for both boxes, especially for the box with $L = 24.2$ Å. Even though model 1 has a lower potential energy for the dimer structures in Fig. 1, model 2 gives a lower potential energy in the simulations and it also gives a lower energy for the experimental density. Since the magnitude of the electrostatic energy for the dimers in model 2 is decreased and the magnitude of the dispersion energy is increased compared to model 1, dispersion forces must be important in the liquid. This is clear from the two-body contribution to the energy in Table 7 for $L = 24.2$ Å. The two-body energy consists of both repulsion and dispersion energies and the scaling of the atomic properties affects the dispersion more than the electrostatic contribution. Note that the repulsion model is the same for models 1 and 2. The increased induced dipole moment for model 2 in Table 7 also reflects the fact that dispersive interactions are important for liquid formamide. The lower potential energies for model 2 in Table 7 justifies the scaling of the multipoles and polarizabilities from MP2 results and indicates the necessity of using large basis sets when dealing with highly polarizable systems.

**Table 7.** Simulation energies per molecule (kcal/mol) and induced dipole moment ($q$ Å) for formamide in the two simulations with the two side lengths. $U_{\text{two-body}}$ consist of repulsion and dispersion energies

| | $L = 24.2$ Å | | $L = 26.0$ Å | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 1 | Model 2 |
| $U_{\text{pot}}$ | −10.5 | −13.8 | −11.2 | −13.3 |
| $U_{\text{kin}}$ | 1.8 | 1.8 | 1.8 | 1.8 |
| $U_{\text{ele}}$ | −12.5 | −9.7 | −11.7 | −9.3 |
| $U_{\text{pol}}$ | −3.0 | −3.4 | −2.6 | −3.1 |
| $U_{\text{two-body}}$ | 5.0 | −0.6 | 3.2 | −0.8 |
| $\mu_{\text{ind}}$ | 0.21 | 0.29 | 0.20 | 0.27 |

The structure in the liquid is characterized by the radial distribution functions (RDFs) and the running coordination number (RCN). Figure 2 shows the RDF for oxygen interacting with the other atoms in formamide, and Fig. 3 shows the RCN for the formamide molecules. Both figures are for the simulation box with experimental density ($L = 24.2$ Å). It is striking how similar the liquid structure is for the two models. Both potentials yield fundamentally the same structure for liquid formamide. One could expect a shift of the peaks for model 2 to longer distances since the overall distances for minimum 2 in Table 5 are about 0.1 Å longer than for minimum 1. The strong first peaks in Fig. 2e and f are in excellent agreement with experimental diffraction studies [52, 53] of the O—H contact, which has a peak at 1.9 Å. Another peak that has been characterized by diffraction experiments is the O—N RDF, which occurs between 2.9–3.05 Å [52, 53]. Our simulation produces a value of 2.95 Å for model 1 and 2.85 Å for

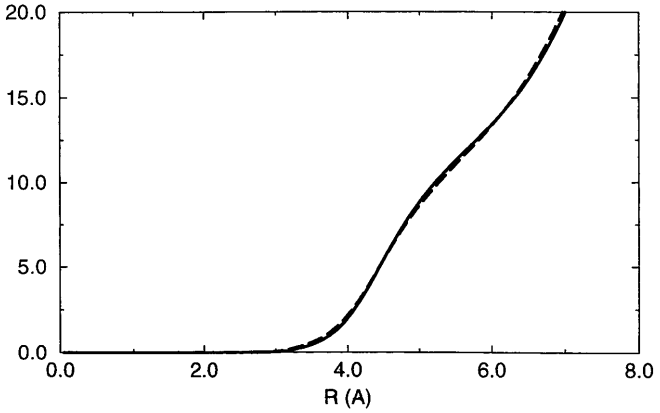


**Fig. 3.** Running coordination number between the center of mass of the formamide molecules in the simulation box with $L = 24.2$ Å. The *solid line* corresponds to model 1 and the *dashed line* to model 2

model 2. Unfortunately, the full individual RDFs are not yet available from experiments, but the other peaks in Fig. 2 are in good agreement with simulations carried out by other groups [46, 47, 54]. A similar picture as for the cyclic formamide dimer was observed for the urea dimer with a previous NEMO potential [12]. The cyclic urea dimer has two N—H···O=C bonds, the same as the cyclic formamide dimer. The potential minimum obtained with the NEMO potential used was much lower than other groups had reported. Both formamide and urea have large polarizabilities on the oxygens and hydrogens, which gives large contributions to the induction and dispersion energies. Especially the cyclic structure with two N—N···O=C bonds is sensitive to errors in the polarizabilities, since at such short distances the induction and dispersions energies change rapidly and even small errors will translate into significant contributions to the total energy. Despite the deficiency in describing the polarizabilities accurately enough in the standard NEMO model, the model gives fairly good interaction energies for the dimer complexes as seen in this work and for the urea dimer [12]. The main effect of scaling the atomic properties in seen in the simulation of liquid formamide. The results indicate the importance of using a large basis set and include correlation effects to describe the polarizabilities for polar systems. Obviously, the kinds of effects discussed here are valid for any potential model and will have large effects on the modeling of interactions in peptides and proteins.

## 5 Conclusions

A test of the reliability of the atomic properties in the NEMO model, obtained at the HF level of theory, was carried out by scaling the atomic charges, dipoles and polarizabilities of formamide by results from MP2 calculations using a larger basis set than normally used in the NEMO model. The molecular dipole moment at the MP2 level is decreased by 11% compared to the SCF results and the polarizability is increased by 36%. The NEMO model is augmented with a damping procedure of the electric field at short interatomic distances. The potential minimum for the cyclic dimer complex with the two models is significantly lower than that obtained by other groups using analytical potential functions. The discrepancy is due to strong dispersive interactions between formamide molecules, which require a large basis set to be described well. The MP2-scaled model gives a potential minimum for the cyclic dimer of $-17.50$ kcal/mol, while the most accurate ab initio calculation gives $-14.50$ kcal/mol [49]. However, trends in the ab initio calculations show that the bottom of the potential well is not reached yet, and a rough estimate of the potential minimum will give an interaction energy of $-16$ kcal/mol at the MP2 limit. MD simulations give an average potential energy of $-13.8$ kcal/mol for the MP2-scaled model and $-10.5$ kcal/mol for the standard NEMO model at the experimental density, while experimental results give $-14.13$ kcal/mol. The lowering of the potential energy for the MP2-scaled model justifies the scaling of the atomic properties and indicates the importance of the dispersion interactions in liquid formamide. The results show that the cancellation of errors, which the NEMO model to some extent relies on, is not always valid, especially for highly polarizable systems. Despite the deficiency in describing the polarizabilities accurately enough in the standard NEMO model, it gives fairly good interaction energies for the dimer complexes as seen in this work and for the urea dimer [12]. The main effect of scaling the atomic properties is not seen in the dimer energies, but in the simulation of liquid formamide. The kinds of effect observed here are valid for any potential model and will have large effects on the modeling of interactions in peptides and proteins.

## References

1. Karim OA (1991) Chem Phys Lett 184: 560
2. Perera L, Berkowitz ML (1991) J Chem Phys 95: 1954
3. Perera L, Berkowitz ML (1993) J Chem Phys 99: 4236
4. Dang LX, Rice JE, Caldwell J, Kollman PA (1991) J Am Chem Soc 113: 2481
5. Brdarski S, Karlström G (1998) J Phys Chem A 102: 8182
6. Sprik M (1991) J Phys Chem 95: 2283
7. Cabalerio-Lago EM, Rios MA (1998) J Chem Phys 110: 6782
8. Stone AJ, Alderton M (1985) Mol Phys 56: 1047
9. Cioslowski J (1989) J Am Chem Soc 111: 8333
10. Åstrand P-O, Ruud K, Mikkelsen KV, Helgaker T (1998) J Phys Chem A 102: 7686
11. Wallqvist A, Karlström G (1989) Chem Scr 29A: 131
12. Åstrand P-O, Wallqvist A, Karlström G (1994) J Chem Phys 100: 1262
13. Engkvist O, Åstrand P-O, Karlström G (1996) J Phys Chem 100: 6950
14. Karlström G (1981) In: van Duijnen PTh, Nieuwpoort WC (eds) Proceedings of the 5th Seminar on Computational Methods in Quantum Chemistry. Laboratory of Chemical Physics, University of Gronningen, Gronningen, The Netherlands, P353, published by Max-Plank-Institut für Physik und Astrophysik; Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, 857 40 Garching bei München, Deutschland
15. Karlström G (1982) Theor Chim Acta 60: 535
16. Åstrand P-O, Karlström G (1992) Mol Phys 77: 143
17. Kellö V, Roos B-O, Sadlej A-J (1998) Theo Chim Acta 74: 185
18. Schütz M, Brdarski S, Widmark P-O, Lindh R, Karlström G (1997) J Chem Phys 107: 4597

19. Morita A, Kato S (1999) J Chem Phys 110: 11987
20. Szabo A, Ostlund NS (1989) Modern quantum chemistry. McGraw-Hill, New York
21. Møller C, Plesset MS (1934) Phys Rev 46: 618
22. Buckingham AD (1967) Adv Chem Phys 12: 107
23. London F (1930) Z Phys 63: 245
24. London F (1930) Z Phys Chem B 11: 221
25. Stone AJ (1981) Chem Phys Lett 83: 233
26. Karlström G (1980) Theor Chim Acta 55: 233
27. Tang KT, Toennies JP (1984) J Chem Phys 80: 3726
28. Brdarski S, Karlström G, Linse P (1999) J Phys Chem A (submitted)
29. Thole BT (1981) Chem Phys 59: 341
30. de Vries AH, van Duijnen PTh, Zijlstra RWJ, Swart M (1997) J Electron Spectrose Relat Phenom 86: 49
31. van Duijnen PTh, Swart M (1998) J Phys Chem A 102: 2399
32. Jensen L, Åstrand P-O, Sylvester-Hvid KO, Mikkelsen KV (2000) J Phys Chem A104, 1563
33. Burnham CJ, Li J, Xantheas SS, Leslie M (1999) J Chem Phys 110: 4566
34. Kitano M, Kuchitsu K (1974) Bull Chem Soc Jpn 47: 67
35. Schultz G, Hargittai J (1993) J Phys Chem 97: 4966
36. Kwiatkowski JS, Leszczynski J (1993) J Mol Struct 297: 277
37. Pierloot K, Dumez B, Widmark P-O, Roos BO (1995) Theor Chim Acta 90: 87
38. Widmark P-O, Malmqvist P-Å, Roos BO (1990) Theor Chim Acta 77: 291
39. Andersson K, Blomberg MRA, Fülscher MP, Karlstöm G, Lindh R, Malmqvist P-Å, Neogrády P, Olsen J, Roos BO, Sadlei AJ, Schütz M, Seijo L, Serrano-Andrés L, Siegbahn PEM, Widmark PO (1997) MOLCAS, version 4.0. Department of Theoretical Chemistry, Lund University, Sweden
40. (1969) Handbook of chemistry and physics. Chemical Rubber Company, Cleveland, Ohio, Page C-303
41. Linse P, Wallqvist A, Åstrand P-O, Nymand TM, Lobaskin V (1999) MOLSIM 2.6.8. Lund University, Sweden
42. Allen MP, Tildesley DJ (1997) Computer simulation of liquids. Clarendon, Oxford
43. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR (1984) J Chem Phys 81: 3684
44. Kurland RJ, Wilson EB (1957) J Chem Phys 27: 585
45. Staley RH, Hardling LB, Goddard WA, Beauchamp JL (1975) Chem Phys Lett 36: 589
46. Sagarik KP, Ahlrichs R (1987) J Chem Phys 86: 5117
47. Gao J, Pavelites JJ, Habibollazadeh D (1996) J Chem Phys 100: 2689
48. Suhai S (1995) J Chem Phys 103: 7030
49. Colominas C, Luque FJ, Orozco M (1999) J Phys Chem 103: 6200
50. Boys SF, Bernardi F (1970) Mol Phys 19: 553
51. Bacon GG (1963) Neutron diffraction. Clarendon Press, Oxford
52. Kalman E, Serke I, Palinkas G, Zeidlen MD, Weissman FJ, Bertagnolli H, Chieuz PZ (1983) 2 Naturforsch A Phys Sci 38: 231
53. Ohtaki H, Funaki A, Rode BM, Reibnegger GJ (1983) Bull Chem Soc Jpn 56: 2116
54. Jorgensen WL, Swenson CJ (1985) J Am Chem Soc 107: 569