# Density Functional Theory Vibrational Frequencies of Amides and Amide Dimers

**Tim M. Watson and Jonathan D. Hirst***

*School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, U.K.*

*Received: January 28, 2002; In Final Form: May 3, 2002*

Infrared spectra of amides and polypeptides can provide detailed information on conformation. An understanding of the amide group in model compounds is a vital step toward a deeper insight into the vibrational spectra of proteins. We show that in contrast to MP2 and the popular B3LYP functional, which overestimate amide I frequencies by 20−80 $cm^{-1}$, the recently developed empirical density functional, EDF1, yields unscaled harmonic vibrational frequencies of monoamides in close agreement with experimental data, even using the relatively small 6-31+G* basis set. New calculations on several hydrogen-bonded amide dimers and the experimental data available for these dimers also support the conclusion that EDF1 yields frequencies in better agreement with experiment than MP2 or B3LYP. We present calculated minimum-energy structures and vibrational spectra of *N*-acetylglycine-*N*′-methylamide and *N*-acetyl-L-alanine-*N*′-methylamide at the EDF1/6-31+G* and B3LYP/6-31+G* levels.

## Introduction

Infrared (IR) spectroscopy is widely used to estimate the secondary structure content of polypeptides and proteins, since many vibrational bands characteristic of the peptide unit are sensitive to their environment.[1] Such studies have benefitted from the advent of techniques such as Fourier self-deconvolution and second-derivative resolution enhancement that allow greater spectral detail to be observed.[2,3] In general, α-helical structure gives rise to a single carbonyl stretch band at ∼1650 $cm^{-1}$. $\beta$-Sheets can give rise to carbonyl stretch bands at ∼1620 $cm^{-1}$ (strong) and at ∼1690 $cm^{-1}$ (weak). Recent experimental developments have expanded the role of IR. Laser temperature jump methods permit time-resolved IR studies to be performed. Recent time-resolved IR studies[4−6] have probed helix−coil interconversion dynamics, an important process in the understanding of protein folding mechanisms. Other studies have investigated the folding of $\beta$-hairpin-type motifs.[7] Whole protein folding studies using time-resolved IR have also been performed using perturbation-induced folding and unfolding methods.[8,9] However, assignments based upon empirical observations of structure−spectrum correlations can, in some circumstances, be misleading.[10] A better understanding of the relationship between protein conformations and their IR spectra would aid in the interpretation of these complex data.

In vibrational spectroscopy of peptides and proteins, a number of key modes are characterized by designations which we will use. The amide A mode is the NH stretch completely located in the NH group. The amide I mode is predominantly the stretching of the CO bond together with an out-of-phase CN stretch component (∼20%). The amide II mode is an out-of-phase combination of NH in-plane bend and CN stretch. The amide III mode is the in-phase combination of NH in-plane bend and CN stretch. There are other mode designations, but amide I−III are the most used in assigning protein structure.[1] Much effort has been put into studying simple molecules that contain the amide unit[1,11−19] (see Figure 1).

The IR spectrum of formamide has been determined by gas-phase and low-temperature argon matrix isolation experiments.[12] Theoretical calculations by Lundell et al.[20] employed second-order Moller−Plesset perturbation theory[21] (MP2) and density functional theory (DFT) using the hybrid functional B3LYP with the split-valence 6-311++G(2d,2p)[22] basis set. The calculated frequencies of the amide I mode were 1778 $cm^{-1}$ (MP2) and 1782 $cm^{-1}$ (B3LYP). Compared to the gas-phase and matrix isolation values of 1755 and 1740 $cm^{-1}$, respectively, there is a 20−40 $cm^{-1}$ discrepancy, despite the relatively large basis set.

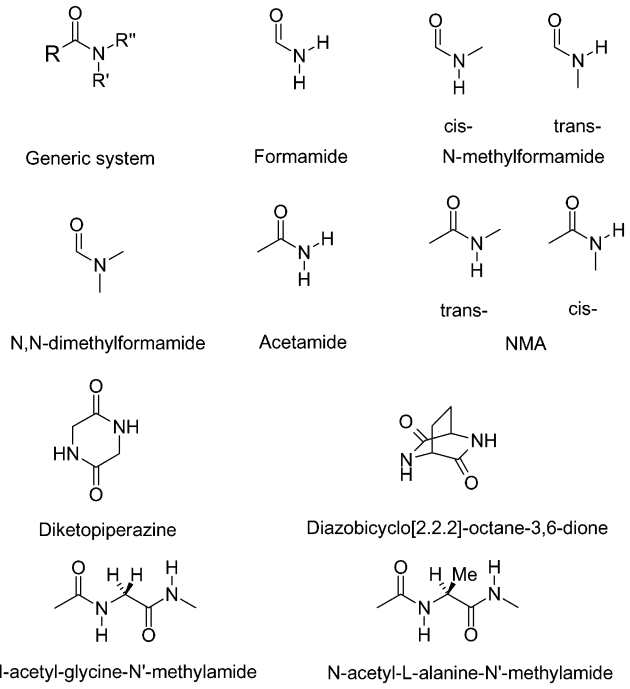


**Figure 1.** Simple amides.

Acetamide has received attention due to debate regarding the CN double bond character and the planarity of the amide system. Kydd and Dunham[23] studied its vapor-phase far-IR, and Kutzelnigg et al.[24] obtained near-IR vapor-phase data. Wong and Wilberg[13] performed MP2 and CISD geometry optimizations using basis sets up to 6-311++G**, although they did not report vibrational data. In a more recent study, Samdal[25] calculated optimized geometries and frequencies with various

10.1021/jp025551l CCC: $22.00 
Published on Web 08/01/2002

methods, from Hartree−Fock (HF) to B3LYP. The calculated amide I frequencies were 1804 $cm^{-1}$ (MP2/cc-pVTZ) and 1762 $cm^{-1}$ (B3LYP/6-311++G**), about 30−70 $cm^{-1}$ greater than the experimental value of 1733 $cm^{-1}$.

Much work has been published on the two major conformations of the secondary amide, *N*-methylacetamide (NMA). The *trans* isomer has long served as a simple model for understanding the nature of the peptide group. Apart from its relevance to models of proteins, it is an interesting molecule in its own right. We do not review recent research on condensed-phase studies. Experimentally, Ataka et al.[26] obtained low-temperature matrix isolated IR spectra of both the *trans* and less stable *cis* isomers. On the theoretical side, Polavarapu et al.[11] reported scaled HF/6-31G* gas-phase studies of both conformers. Herrebout et al.[19] reported gas-phase B3LYP/6-31G* calculations. The amide I frequency was calculated to be 1792 $cm^{-1}$, 85 $cm^{-1}$ greater than the experimental value.

Experimental data for *N*,*N*-dimethylformamide were collected by Stålhandske et al.[15] and compared to their own RHF/6-31G** and DFT (BP86) calculations. The DFT calculations yielded an amide I value of 1715 $cm^{-1}$ compared to the experimental value of 1677 $cm^{-1}$. Experimental matrix isolated data for the *cis* and *trans* isomers of *N*-methylformamide were obtained by Ataka et al.[26] in the same study as *cis/trans*-NMA. To our knowledge no theoretical studies of the vibrational spectra of *N*-methylformamide have been undertaken.

Some cyclic dipeptides have also been previously studied. Hirst and Persson[18] reported MP2 and B3LYP vibrational data, using correlation consistent double- and triple-$\zeta$ basis sets,[27−29] respectively, for planar and boat conformations of diketopiperazine (DKP). For the lowest energy (boat) conformer the calculated amide I frequencies were 1820 $cm^{-1}$ (MP2) and 1771 $cm^{-1}$ (B3LYP), for the strongly IR active bands. A rigid bridged analogue of DKP, diazobicyclo[2.2.2]octane-3,6-dione, has also been studied.[17] The calculated strongly active amide I frequencies were 1827 and 1819 $cm^{-1}$ for MP2/6-31G* and B3LYP/cc-pVDZ, respectively. For both DKP and the bridged DKP, there are no experimental gas-phase IR data, although there are data for the solid phase[30] and solution phase,[17] respectively.

Some studies have been undertaken on hydrogen-bonded dimers of formamide and NMA. Colominas et al.[31] obtained various structures of formamide dimers using Monte Carlo calculations to explore configurational space. These configurations were then optimized by HF, MP2, and B3LYP methods using the 6-31G* basis set, but no vibrational data were calculated. Vargas et al.[32] calculated five optimized formamide dimer and four NMA dimer structures at the MP2/aug-cc-pVTZ and MP2/aug-cc-pVDZ levels, respectively. Computational expense precluded frequency calculations. Both of these studies serve as starting points for our calculations.

Most of the work discussed above has concentrated on small-molecule experimental and theoretical studies. The calculations, while perhaps restricted in scope, have an important role. Without these studies, extension of such calculations to larger systems would be qualitative at best. Indeed, this is one of the reasons that these systems continue to be studied by both theoretical and experimental methods.[10,33−38] A recent study used DFT methods to show that the amide III band of the alanine dipeptide depends on the backbone $\phi$ and $\psi$ dihedral angles, hinting strongly at the possibility of obtaining structural parameters from IR.[39] This shows that in concert with conformational studies of dipeptides[35,40,41] (and larger systems) vibrational calculations can yield valuable information. The computationally cheaper and hence widely used molecular mechanics and semiempirical methods often depend on accurate ab initio calculations for parametrization.[42,43] This requires the development and validation of new methods that yield data in better agreement with experiment.

In many studies, the HF method is used to calculate vibrational frequencies, yet it has long been known that it overestimates these frequencies, sometimes to an alarming degree. The main contributions to the error come from incomplete treatment of electron correlation (i.e., improper dissociation behavior for HF) and the anharmonicity of the vibrations. However, the errors are reasonably uniform in nature and amenable to scaling to bring values in line with experiment.[44] A recent evaluation of scaling factors[45] showed that for HF frequencies a scale factor of 0.897 was optimal for the 6-31+G* basis set; for MP2/6-31G* and B3LYP/6-31G*, 0.943 and 0.961, respectively, are optimal. Such studies have concentrated on molecules with no more than 4 heavy atoms and less than 10 atoms in total. Scaling factors can still give significant discrepancies between theory and experiment. To our knowledge, scale factors have not been tested comprehensively on larger systems, and this approach may not provide an adequate foundation for approaching polypeptide systems in the which splitting of spectral features can be on the order of 10−100 $cm^{-1}$. New studies[46] on aromatic carbonyls have concluded that the recently developed empirical density functional, EDF1,[47] provides better unscaled vibrational frequencies than unscaled B3LYP. Herein, we apply the EDF1 functional to the calculation of the IR spectra of simple amides, hydrogen-bonded homodimers of formamide and NMA, and the two covalent dipeptides *N*-acetylglycine-*N*′-methylamide (Ac-Gly-NHMe) and *N*-acetyl-L-alanine-*N*′-methylamide (Ac-Ala-NHMe).

Since EDF1 is a recent development, we describe its origin and implementation. The functional was devised with several objectives in mind. The first was to increase the emphasis on empirical parametrization, instead of imposing an exact fit to certain limiting conditions. The second objective was to employ a relatively small basis set, 6-31+G*, so that large systems could be tackled. It was hoped that the increased empirical parametrization would absorb some of the limitations of the basis set. The third objective was to determine whether Fock exchange was necessary for good performance. The final optimized functional consists of a Slater component, a linear combination of two Becke88[48] functions (double-Becke) and a reparametrized LYP[49,50] function. Incorporating Fock exchange did not significantly improve performance.

The parametrization was performed against a modified G2[51] database of experimental atomization energies, ionization potentials, electron affinities, and proton affinities; no vibrational data were used. The latter is true of most functionals developed so far, yet DFT calculations routinely give vibrational frequencies in better agreement with experiment than strictly ab initio approaches, such as MP2 with similar size basis sets. This is not fully understood. As has been noted, functionals parametrized for one basis set may not be equally suitable for another.[47] This implies that using any functional with basis sets other than that for which it was parametrized may lead to unreliable results. B3LYP was parametrized using a completely numerical non-basis-set calculation that (in effect) is equivalent to an infinite basis set calculation.[52] However, large-basis calculations can be prohibitively expensive, so somewhat smaller basis sets are generally used. There is, therefore, some ambiguity over which basis set to use and the reliability of the results. With EDF1, 6-31+G* is unambiguously the basis set that should be used and is amenable to application to larger systems.

## Computational Details

Molecular geometries were built using the molecular editor of Spartan '02.[53] All calculations were carried out using the Q-Chem 2.01 ab initio program.[54] Harmonic vibrational frequencies have been computed at the MP2, B3LYP, and EDF1 levels of theory. All reported vibrational frequencies are unscaled. The data presented are the unmodified results of the vibrational calculation. Coupled cluster calculations, although desirable, were too computationally expensive for all but formamide with a moderately sized basis set. Q-Chem default convergence criteria (maximum gradient $(3.0 \times 10^{-4})E_h a_o^{-1}$, maximum atomic displacement $(1.2 \times 10^{-3})a_o$, and maximum energy change $(1.0 \times 10^{-6})E_h$) were used in all cases. All geometries were fully optimized prior to frequency calculations at the same level of theory.

Frequency calculations in the ab initio framework usually yield harmonic values. While it is possible to add anharmonic corrections, this is nontrivial and beyond the scope of the current study. These corrections typically are on the order of a few percent, i.e., $\sim 10-30$ cm$^{-1}$.[55] The anharmonic correction might be anticipated to be larger for the mode(s) involving the hydrogen bond. By examining frequency-shifted transient difference spectra of the transition, Hamm et al. determined the anharmonicity of the amide I band in deuterated NMA to be 16 cm$^{-1}$.[56] As we shall see, this compares well with the remaining differences between the experimental data and the EDF1 calculations.

Basis set superposition error (BSSE) can have considerable effects on the calculated properties of weakly bound systems. Paizs and Suhai suggest that correction for BSSE is mandatory.[57] However, there are conflicting opinions. Rablen et al. studied hydrogen-bonded complexes of small organic molecules and reported that, for the same basis set, BSSE in the DFT calculation is much smaller than at the MP2 level.[58] The counterpoise (CP) method[59] is most commonly used to compensate for BSSE, yet there is some concern that CP correction may overestimate the BSSE.[58] Due to these concerns and the fact that DFT converges more quickly with respect to basis set size than post-HF methods, we have not included BSSE corrections.

For each molecule, formamide, acetamide, *cis/trans-N*-methylformamide, *N,N*-dimethylformamide, and *cis/trans*-NMA, we performed the calculations MP2/6-31+G*, B3LYP/6-31+G*, B3LYP/6-311++G(2d,2p), BLYP/6-31G*, and EDF1 with the basis sets 6-31+G*, 6-31+G**, 6-311+G*, 6-311+G(df,p), and 6-311++G(2d,2p). We performed the BLYP calculations with the 6-31G* basis set, since this combination leads to quite accurate vibrational frequency predictions.[45] In the DFT calculations, the standard SG-1 grid[60] was inadequate, with geometry optimizations unable to converge to structures at potential energy minima. Hence, all reported calculations make use of the 70-point Euler−Maclaurin radial grid[61] combined with the two-dimensional Lebedev grid[62] with 302 angular points. The methyl groups in these molecules have several low-energy torsional conformers that can cause problems for the energy minimization routine; a good example is *trans-N*-methylacetamide. Previous MP2/6-31+G* calculations have shown that the lowest energy conformation is slightly nonplanar, with the methyl groups $\sim 4°$ out of plane.[11] To allow for these effects we have performed all calculations in $C_1$ symmetry.

Having determined that EDF1 provides a reasonable description of the vibrational modes of monomeric amides in the gas phase, we studied homodimers of formamide and NMA. In particular, we were interested in how the 6-31+G* basis set coped with the hydrogen bonding in these systems. We calculated geometries and frequencies using the following methods: MP2/6-31+G*, B3LYP/6-31+G*, B3LYP/6-311++G(2d,2p), EDF1/6-31+G*, EDF1/6-31++G*, and EDF1/6-31+G**.

Subsequently we studied the well-understood Ac-Gly-NHMe and Ac-Ala-NHMe molecules (see Figure 1 for structures). Although there are several variants of the glycine and alanine dipeptide mimics, the methyl-capped systems that we studied are the systems analogous to the single amide protein model, *trans*-NMA. The potential energy surfaces with respect to the dihedral angles $\phi$ and $\psi$ for these molecules have been calculated previously[63,64] using various ab initio methods with basis sets up to 6-31G*. Using the B3LYP/6-31G* minima calculated previously as initial geometries, we have calculated the minimum-energy structures at the B3LYP/6-31+G* and EDF1/6-31+G* levels of theory. For each of the minima we report calculated harmonic vibrational frequencies and important structural details.

## Results and Discussion

**Monomeric Amides.** To gauge how well EDF1 works, we collated known literature gas-phase IR data on a series of simple amides. We then calculated optimized geometries, IR frequencies, and IR intensities. Table 1 contains selected optimized geometric data, and Tables 2−4 show values for the amide I−III bands, respectively. We can readily see that the MP2/6-31+G* calculation consistently overestimates all vibrational frequencies by $20-80$ cm$^{-1}$. The B3LYP functional, while an improvement on MP2, also overestimates vibrational frequencies. The EDF1 functional furnishes the best results, even with the relatively small basis set used.

We define the error to be the difference between the experimental frequency and the unscaled calculated harmonic frequency:

$$\Delta = \nu_{\text{exptl}} - \nu_{\text{calcd}}$$

We have analyzed the errors in these calculations in two ways. The first method considers a given vibrational mode across all of the molecules in the data set. For the amide I−III modes we calculated a mean signed error, the standard deviation (sd), of the errors and the root-mean-square error ($\text{rms}_{\text{mode}}$).

$$\text{rms}_{\text{mode}} = \left[\frac{\sum_{1}^{n_{\text{molecules}}} \Delta^2}{n_{\text{molecules}}}\right]^{1/2}$$

This measures how well given modes are reproduced by each method. Our second analysis compares modes for a single molecule. For instance, for *trans*-NMA we have a total of 19 assigned experimental bands. We calculate the mean, sd, and root-mean-square errors ($\text{rms}_{\text{mol}}$) for all of the assigned modes in the single molecule.

$$\text{rms}_{\text{mol}} = \left[\frac{\sum_{1}^{n_{\text{modes}}} \Delta^2}{n_{\text{modes}}}\right]^{1/2}$$

From Table 5, we see that, for all of the modes studied, the EDF1/6-31+G* calculation has the lowest mean signed error.

**TABLE 1: Selected Optimized Structural Data for the Monomers Studied[a]**

| | $r_{CO}$ (Å) | $r_{CN}$ (Å) | $r_{RC}$ (Å) | $r_{NR'}$ (Å) | $r_{NR''}$ (Å) | $\angle_{NCO}$ (deg) | $\angle_{RCO}$ (deg) | $\angle_{CNR'}$ (deg) | $\angle_{CNR''}$ (deg) | T1 (deg) | T2 (deg) | T3 (deg) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Formamide | | | | | | |
| MP2 | 1.228 | 1.361 | 1.103 | 1.010 | 1.014 | 124.6 | 122.3 | 121.4 | 119.1 | | | |
| B3LYP | 1.219 | 1.362 | 1.106 | 1.010 | 1.013 | 124.8 | 122.6 | 121.6 | 119.5 | | | |
| EDF1 | 1.226 | 1.365 | 1.112 | 1.014 | 1.015 | 125.0 | 122.8 | 121.7 | 119.6 | | | |
| | | | | | | *trans*-*N*-Methylformamide | | | | | | |
| MP2 | 1.233 | 1.357 | 1.103 | 1.012 | 1.454 | 124.0 | 122.6 | 118.6 | 121.1 | 179.2 | | |
| B3LYP | 1.222 | 1.362 | 1.106 | 1.012 | 1.455 | 125.5 | 122.3 | 117.5 | 123.6 | 1.0 | | |
| EDF1 | 1.229 | 1.367 | 1.112 | 1.015 | 1.455 | 125.8 | 122.5 | 117.2 | 124.2 | 0.8 | | |
| | | | | | | *cis*-*N*-Methylformamide | | | | | | |
| MP2 | 1.232 | 1.359 | 1.104 | 1.450 | 1.016 | 124.9 | 122.1 | 124.7 | 115.7 | | −0.1 | |
| B3LYP | 1.222 | 1.361 | 1.108 | 1.452 | 1.015 | 125.0 | 122.3 | 125.2 | 115.7 | | 0.3 | |
| EDF1 | 1.229 | 1.365 | 1.113 | 1.453 | 1.018 | 125.0 | 122.6 | 125.6 | 115.6 | | 0.1 | |
| | | | | | | *N*,*N*-Dimethylformamide | | | | | | |
| MP2 | 1.234 | 1.361 | 1.104 | 1.450 | 1.450 | 125.6 | 121.8 | 121.7 | 120.5 | 5.3 | −7.7 | |
| B3LYP | 1.225 | 1.364 | 1.106 | 1.451 | 1.455 | 125.8 | 121.9 | 121.8 | 120.5 | 0.1 | −0.1 | |
| EDF1 | 1.230 | 1.370 | 1.112 | 1.453 | 1.455 | 125.9 | 122.1 | 121.7 | 120.5 | 0.2 | −0.3 | |
| | | | | | | Acetamide | | | | | | |
| MP2 | 1.232 | 1.371 | 1.511 | 1.010 | 1.013 | 121.8 | 122.5 | 121.2 | 117.0 | | | 152.2 |
| B3LYP | 1.225 | 1.369 | 1.520 | 1.009 | 1.012 | 121.9 | 122.1 | 123.0 | 118.5 | | | 180.0 |
| EDF1 | 1.232 | 1.373 | 1.524 | 1.012 | 1.015 | 121.9 | 122.2 | 123.2 | 118.5 | | | 180.0 |
| | | | | | | *trans*-NMA | | | | | | |
| MP2 | 1.237 | 1.361 | 1.513 | 1.011 | 1.453 | 121.6 | 122.0 | 119.7 | 120.8 | | 180.0 | 180.0 |
| B3LYP | 1.229 | 1.365 | 1.521 | 1.009 | 1.457 | 121.8 | 122.0 | 119.4 | 121.5 | | 180.0 | −179.9 |
| EDF1 | 1.235 | 1.374 | 1.524 | 1.013 | 1.454 | 122.9 | 121.6 | 118.4 | 123.5 | | 0.1 | 179.7 |
| | | | | | | *cis*-NMA | | | | | | |
| MP2 | 1.238 | 1.369 | 1.511 | 1.453 | 1.016 | 121.2 | 122.4 | 125.9 | 113.5 | −172.0 | | 3.7 |
| B3LYP | 1.229 | 1.370 | 1.520 | 1.455 | 1.013 | 121.1 | 122.3 | 127.1 | 113.9 | −170.7 | | 3.7 |
| EDF1 | 1.237 | 1.376 | 1.523 | 1.456 | 1.017 | 120.8 | 122.4 | 127.9 | 113.4 | −167.2 | | 4.4 |

[a] Calculations performed with the 6-31+G* basis set. See Figure 1 for the definition of R, R′, and R″. T1 = torsion angle C−N−C′−H(ip). T2 = torsion angle C−N−C″−H(ip). T3 = torsion angle H(ip)−R−C=O.

**TABLE 2: Calculated Amide I Frequencies Relative to Experimental Data and Absolute Intensities**

| | | exptl − calcd | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MP2/6-31+G* | | B3LYP/6-31+G* | | B3LYP/6-311++G(2d,2p) | | B-LYP/6-31G* | | EDF1/6-31+G* | |
| species | exptl ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| formamide[12] | 1755 | −40 | 443 | −45 | 467 | −26 | 442 | −16 | 352 | −9 | 423 |
| acetamide[14] | 1733 | −47 | 354 | −38 | 414 | −19 | 394 | −7 | 317 | 1 | 375 |
| *trans*-NMF[26] | 1721 | −49 | 362 | −60 | 389 | −42 | 365 | −26 | 269 | −27 | 391 |
| *N*,*N*-DMF[15] | 1677 | −83 | 520 | −87 | 519 | −68 | 430 | −53 | 367 | −52 | 462 |
| *cis*-NMA[11] | 1707 | −56 | 412 | −51 | 446 | −32 | 430 | −21 | 316 | −10 | 393 |
| *trans*-NMA[11] | 1707 | −46 | 399 | −44 | 300 | −32 | 285 | −8 | 209 | −10 | 280 |

**TABLE 3: Calculated Amide II Frequencies Relative to Experimental Data and Absolute Intensities**

| | | exptl − calcd | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MP2/6-31+G* | | B3LYP/6-31+G* | | B3LYP/6-311++G(2d,2p) | | B-LYP/6-31G* | | EDF1/6-31+G* | |
| species | exptl ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| formamide[12] | 1580 | −71 | 54 | −60 | 60 | −44 | 61 | −7 | 44 | −27 | 48 |
| acetamide[14] | 1600 | −63 | 94 | −46 | 96 | −25 | 92 | 12 | 72 | −15 | 82 |
| *trans*-NMF[26] | 1528 | −52 | 102 | −43 | 104 | −31 | 101 | −41 | 66 | −9 | 79 |
| *N*,*N*-DMF[15] | 1507 | −76 | 20 | −49 | 21 | −37 | 21 | −57 | 8 | −18 | 16 |
| *cis*-NMA[11] | 1485 | −86 | 32 | −57 | 31 | −42 | 32 | −67 | 16 | −27 | 28 |
| *trans*-NMA[11] | 1511 | −79 | 217 | −50 | 198 | −16 | 182 | −49 | 106 | −30 | 121 |

Although the sd in errors for the amide I band is slightly higher for EDF1 than the other calculations, they are lower for amide II and amide III; EDF1 does better in this regard than all but B3LYP/6-31+G* for amide II. The $rms_{mode}$ error is lower for EDF1 across all three bands. Table 6 shows data for modes within a given molecule. For all calculations, the largest errors originate from the $\nu$(N−H) and $\nu$(C−H) modes at high wavenumbers (i.e., above ∼2900 $cm^{-1}$). Since these modes are not generally of interest for protein/peptide analysis, in addition to the errors for all vibrational modes with experimental data, we have calculated the errors excluding these high $\nu$(N−H) and $\nu$(C−H) modes. Considering all of the modes, BLYP/6-31G* and EDF1/6-31+G* are the two best performing calculations. BLYP has lower sd's and $rms_{mol}$ errors, although EDF1 has better mean signed errors. On excluding the high-wavenumber modes, EDF1 is clearly the more accurate calculation.

Although increasing the size of the basis set in the B3LYP calculations did reduce the error, there is still significant deviation from experiment. This is important, considering the large increase in computational cost. We also investigated EDF1 with a series of basis sets ranging from 6-31+G* to 6-311++G(2d,2p).

**TABLE 4: Calculated Amide III Frequencies Relative to Experimental Data and Absolute Intensities**

| | | exptl − calcd | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | MP2/6-31+G* | | B3LYP/6-31+G* | | B3LYP/6-311++G(2d,2p) | | B-LYP/6-31G* | | EDF1/6-31+G* | |
| species | exptl ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| formamide[12] | 1255 | −48 | 116 | −23 | 115 | −12 | 112 | −7 | 103 | −1 | 97 |
| acetamide[14] | 1319 | −66 | 122 | −37 | 150 | −23 | 147 | −16 | 161 | −7 | 144 |
| *trans*-NMF[26] | 1207 | −62 | 44 | −24 | 77 | −11 | 8 | 1 | 112 | 0 | 65 |
| *N,N*-DMF[15] | 1388 | −83 | 23 | −55 | 79 | −44 | 88 | −51 | 96 | −18 | 73 |
| *cis*-NMA[11] | 1325 | −66 | 115 | −35 | 144 | −23 | 143 | −14 | 179 | −6 | 131 |
| *trans*-NMA[11] | 1266 | −57 | 57 | −26 | 83 | −82 | 88 | −10 | 113 | −14 | 107 |

**TABLE 5: Errors in Amide I−III Frequencies for All Monomers in the Study**

| | mean error | sd in errors | $rms_{mode}$ |
|---|---|---|---|
| Amide I | | | |
| MP2/6-31+G* | −53 | 15 | 55 |
| B3LYP/6-31+G* | −54 | 18 | 57 |
| B3LYP/6-311++G(2d,2p) | −37 | 17 | 40 |
| B-LYP/6-31G* | −22 | 17 | 27 |
| EDF1/6-31+G* | −18 | 19 | 25 |
| Amide II | | | |
| MP2/6-31+G* | −71 | 12 | 72 |
| B3LYP/6-31+G* | −51 | 6 | 51 |
| B3LYP/6-311++G(2d,2p) | −33 | 11 | 34 |
| B-LYP/6-31G* | −35 | 31 | 45 |
| EDF1/6-31+G* | −21 | 8 | 22 |
| Amide III | | | |
| MP2/6-31+G* | −64 | 12 | 65 |
| B3LYP/6-31+G* | −34 | 12 | 35 |
| B3LYP/6-311++G(2d,2p) | −33 | 27 | 41 |
| B-LYP/6-31G* | −16 | 18 | 23 |
| EDF1/6-31+G* | −3 | 11 | 10 |

**TABLE 6: Errors for All Modes with Experimental Data**

| | all vibrations | | | excluding high $\nu$(CH) | | |
|---|---|---|---|---|---|---|
| | mean error | sd in errors | $rms_{mol}$ | mean error | sd in errors | $rms_{mol}$ |
| Formamide | | | | | | |
| MP2/6-31+G* | −110 | 73 | 129 | −53 | 13 | 54 |
| B3LYP/6-31+G* | −79 | 86 | 92 | −40 | 16 | 43 |
| B3LYP/6-311++G(2d,2p) | −72 | 59 | 90 | −28 | 13 | 30 |
| B-LYP/6-31G* | −34 | 20 | 39 | −33 | 27 | 40 |
| EDF1/6-31+G* | −36 | 36 | 50 | −9 | 13 | 14 |
| Acetamide | | | | | | |
| MP2/6-31+G* | −91 | 92 | 127 | −44 | 31 | 53 |
| B3LYP/6-31+G* | −61 | 82 | 100 | −21 | 40 | 44 |
| B3LYP/6-311++G(2d,2p) | −50 | 91 | 102 | −6 | 51 | 50 |
| B-LYP/6-31G* | −55 | 49 | 73 | −37 | 34 | 49 |
| EDF1/6-31+G* | −33 | 76 | 81 | −4 | 40 | 39 |
| *trans*-*N*-Methylformamide | | | | | | |
| MP2/6-31+G* | −112 | 86 | 139 | −65 | 28 | 70 |
| B3LYP/6-31+G* | −79 | 64 | 100 | −45 | 20 | 49 |
| B3LYP/6-311++G(2d,2p) | −70 | 64 | 94 | −35 | 21 | 40 |
| B-LYP/6-31G* | −59 | 41 | 71 | −45 | 25 | 51 |
| EDF1/6-31+G* | −45 | 59 | 73 | −14 | 21 | 25 |
| *trans*-NMA | | | | | | |
| MP2/6-31+G* | −123 | 64 | 138 | −78 | 18 | 80 |
| B3LYP/6-31+G* | −85 | 46 | 96 | −54 | 19 | 57 |
| B3LYP/6-311++G(2d,2p) | −70 | 40 | 80 | −43 | 17 | 46 |
| B-LYP/6-31G* | −60 | 24 | 64 | −55 | 26 | 61 |
| EDF1/6-31+G* | −48 | 42 | 64 | −22 | 24 | 32 |

Table 7 shows there are small differences in the calculated frequencies, but the calculations are reasonably well converged. Since EDF1 was not parametrized for vibrational frequencies, this is important as it confirms the adequacy of the 6-31+G* basis set for EDF1 frequency calculations. For these unscaled calculations, EDF1 is superior to the more computationally expensive MP2 and B3LYP methods in terms of reproducing experimental frequencies.

**TABLE 7: Convergence of EDF1 Amide I Frequencies with the Basis Set**

| | exptl − calcd ($cm^{-1}$) | | | |
|---|---|---|---|---|
| species | 6-31+G* | 6-311+G* | 6-311+G(df,p) | 6-311++G(2d,2p) |
| formamide | −9 | −6 | 4 | 9 |
| acetamide | 1 | 6 | 8 | 21 |
| *trans*-NMF | −27 | −22 | −19 | −7 |
| *N,N*-DMF | −52 | −47 | −46 | −33 |
| *cis*-NMA | −10 | −4 | −2 | 9 |
| *trans*-NMA | −10 | −4 | −3 | 10 |

**TABLE 8: Selected Geometry Information for Hydrogen-Bonded Dimers[a]**

| | NH---O=C | | | |
|---|---|---|---|---|
| method | $r_{HO}$ (Å) | $r_{NO}$ (Å) | $\angle_{NHO}$ (deg) | $\angle_{COH}$ (deg) |
| Formamide Dimer 1 | | | | |
| MP2 | 1.907 | 2.928 | 172.4 | 122.2 |
| B3LYP | 1.886 | 2.909 | 172.1 | 122.2 |
| EDF1 | 1.922 | 2.949 | 172.8 | 121.1 |
| Formamide Dimer 2 | | | | |
| MP2 | 1.935 | 2.945 | 167.6 | 107.4 |
| B3LYP | 1.911 | 2.922 | 167.0 | 109.6 |
| EDF1 | 1.968 | 2.988 | 171.1 | 110.6 |
| Formamide Dimer 3 | | | | |
| MP2 | 2.030 | 2.979 | 153.6 | 111.3 |
| B3LYP | 2.004 | 2.973 | 157.4 | 114.0 |
| EDF1 | 2.098 | 3.115 | 175.2 | 141.8 |
| Formamide Dimer 4 | | | | |
| MP2 | 1.969 | 2.982 | 172.8 | 124.1 |
| B3LYP | 1.976 | 2.994 | 176.7 | 131.2 |
| EDF1 | 2.076 | 3.096 | 178.1 | 134.2 |
| *cis*-NMA Cyclic Dimer | | | | |
| MP2 | 1.871 | 2.901 | 179.3 | 120.7 |
| B3LYP | 1.876 | 2.905 | 179.9 | 121.9 |
| EDF1 | 1.909 | 2.942 | 179.1 | 121.7 |
| *trans*-NMA Dimer 1 | | | | |
| MP2 | 1.952 | 2.965 | 172.7 | 141.9 |
| B3LYP | 2.021 | 3.004 | 162.4 | 178.4 |
| EDF1 | 2.201 | 3.200 | 166.9 | 164.8 |
| *trans*-NMA Dimer 2 | | | | |
| MP2 | 1.956 | 2.952 | 179.6 | 133.6 |
| B3LYP | 1.995 | 3.010 | 175.0 | 147.5 |
| EDF1 | 2.183 | 3.195 | 172.3 | 148.9 |
| | CH---O=C | | | |
| method | $r_{HO}$ (Å) | $r_{CO}$ (Å) | $\angle_{CHO}$ (deg) | $\angle_{COH}$ (deg) |
| Formamide Dimer 2 | | | | |
| MP2 | 2.304 | 3.249 | 143.0 | 115.2 |
| B3LYP | 2.336 | 3.262 | 140.5 | 115.8 |
| EDF1 | 2.616 | 3.500 | 136.4 | 113.8 |

[a] Reported data were calculated using the 6-31+G* basis set.

**Hydrogen-Bonded Dimers.** We utilized structures calculated by Vargas et al. as starting points for our calculations on formamide dimers.[32] Figure 2 shows the EDF1/6-31+G*-optimized geometries for the four dimers considered. Table 8

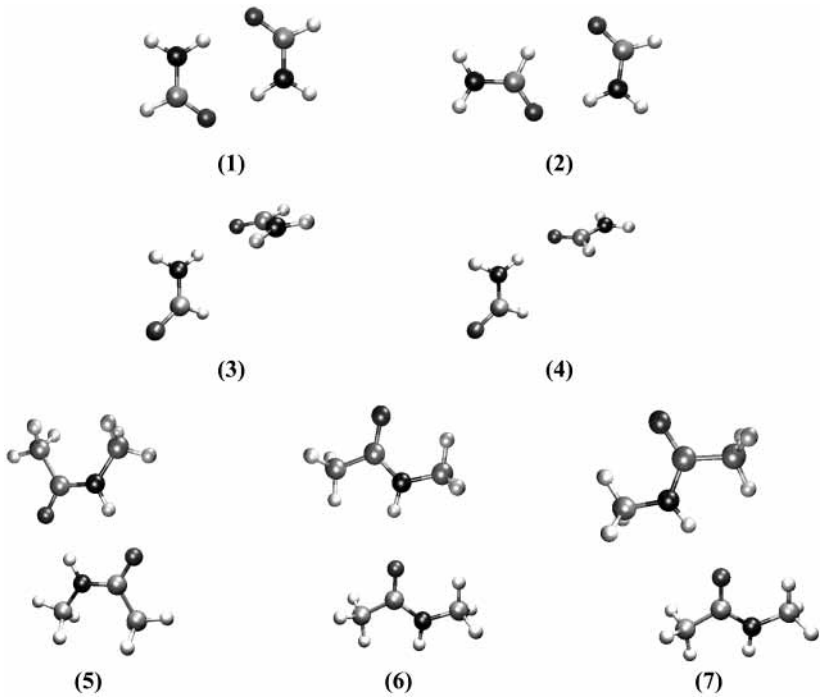


**Figure 2.** Conformations of the formamide and NMA dimers studied: (**1**) formamide dimer 1 ($C_{2h}$), (**2**) formamide dimer 2 ($C_s$), (**3**) formamide dimer 3 ($C_1$), (**4**) formamide dimer 4 ($C_1$), (**5**) cyclic *cis*-NMA dimer, (**6**) *trans*-NMA dimer 1, (**7**) *trans*-NMA dimer 2.

**TABLE 9: Calculated Frequencies and Intensities for the Amide I Mode in the Formamide and NMA Dimers**

| | | MP2/6-31+G* | | B3LYP/6-31+G* | | B3LYP/6-311++G(2d,2p) | | EDF1/6-31+G* | |
|---|---|---|---|---|---|---|---|---|---|
| species | exptl ($cm^{-1}$) | $E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| formamide dimer 1 | | 1796 | 896 | 1786 | 924 | 1765 | 870 | 1752 | 813 |
| | | 1775 | 0 | 1758 | 0 | 1734 | 0 | 1726 | 0 |
| formamide dimer 2 | | 1793 | 702 | 1786 | 789 | 1767 | 731 | 1754 | 653 |
| | | 1762 | 189 | 1750 | 148 | 1729 | 123 | 1724 | 181 |
| formamide dimer 3 | | 1789 | 452 | 1792 | 495 | 1775 | 469 | 1757 | 445 |
| | | 1778 | 415 | 1776 | 463 | 1756 | 448 | 1745 | 490 |
| formamide dimer 4 | | 1792 | 154 | 1790 | 321 | 1771 | 281 | 1757 | 151 |
| | | 1782 | 945 | 1778 | 758 | 1759 | 756 | 1746 | 885 |
| *cis*-NMA cyclic dimer | 1695 | 1763 | 916 | 1743 | 964 | 1722 | 927 | 1705 | 854 |
| | | 1728 | 0 | 1711 | 0 | 1691 | 0 | 1676 | 0 |
| *trans*-NMA dimer 1 | 1686 | 1753 | 14 | 1743 | 12 | 1724 | 7 | 1710 | 28 |
| | | 1741 | 819 | 1735 | 765 | 1716 | 733 | 1702 | 688 |
| *trans*-NMA dimer 2 | 1686 | 1750 | 15 | 1742 | 15 | 1722 | 31 | 1708 | 2 |
| | | 1741 | 784 | 1733 | 780 | 1713 | 720 | 1701 | 712 |

contains selected geometry data, and the detailed geometries are provided in the Supporting Information. As in the monomer calculations a downward trend in amide I vibrational frequencies is evident (Table 9) as we move from MP2 to EDF1 calculations. Since no experimental data are available, it is not possible to quantify the accuracy of the various calculations. However, the trend is similar to that of the monoamides, so we might expect EDF1 to be the more accurate calculation.

Upon hydrogen bond formation, we observe a decrease in predicted amide I IR frequency from the monomer values, in accord with previous experimental studies.[33] The amide I splitting is greater for the cyclic dimers in which there are two hydrogen bonds. For formamide dimers **1**−**4** the splittings are 26, 30, 11, and 12 $cm^{-1}$, respectively. However, only one of the two bands is strongly active. If we consider the cyclic formamide shown in Figure 3, the in-phase stretch is strictly IR inactive due to its $C_{2h}$ symmetry; there is no change in dipole moment, so the mode does not couple to the electromagnetic field. The out-of-phase mode has a large change in dipole moment and hence a nonzero (large) intensity.

We considered three NMA dimers: a cyclic *cis*-NMA system and two configurations of a singly hydrogen bonded *trans*-NMA dimer, as studied by Torii et al.[33] EDF1/6-31+G*-optimized geometries are shown in Figure 3. Tables 9 and 10 show the amide I and II calculated frequencies. For NMA dimers, there are experimental data on the amide I band from matrix isolated infrared studies.[33] The MP2 vibrational frequencies are higher than the experimental data. B3LYP offers some improvement, but requires large basis sets to achieve even reasonable accuracy. EDF1/6-31+G* gives good agreement with experiment. Experimental data for the *cis*- and *trans*-NMA dimers show only one amide I band, and this is predicted by all of the calculations for the same symmetry arguments that applied to the formamide

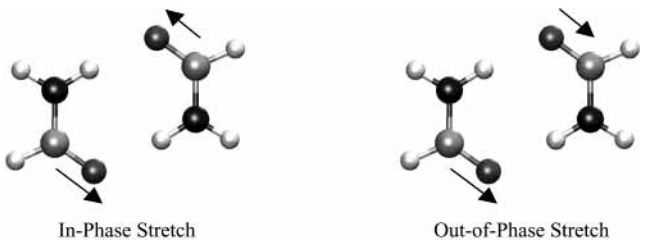


**Figure 3.** In- and out-of-phase stretching modes of the cyclic formamide dimer (**1**).

**TABLE 10: Calculated Frequencies and Intensities for the Amide II Mode in the Formamide and NMA Dimers**

| | MP2/6-31+G* | | B3LYP/6-31+G* | | B3LYP/6-311++G(2d,2p) | | EDF1/6-31+G* | |
|---|---|---|---|---|---|---|---|---|
| species | $E$ (cm$^{-1}$) | $I$ (km mol$^{-1}$) | $E$ (cm$^{-1}$) | $I$ (km mol$^{-1}$) | $E$ (cm$^{-1}$) | $I$ (km mol$^{-1}$) | $E$ (cm$^{-1}$) | $I$ (km mol$^{-1}$) |
| formamide dimer 1 | 1677 | 20 | 1670 | 26 | 1655 | 21 | 1638 | 19 |
| | 1669 | 0 | 1662 | 0 | 1644 | 0 | 1630 | 0 |
| formamide dimer 2 | 1673 | 8 | 1664 | 18 | 1651 | 14 | 1629 | 14 |
| | 1649 | 38 | 1641 | 42 | 1624 | 41 | 1608 | 34 |
| formamide dimer 3 | 1679 | 55 | 1669 | 51 | 1649 | 47 | 1632 | 37 |
| | 1652 | 17 | 1644 | 28 | 1630 | 30 | 1616 | 35 |
| formamide dimer 4 | 1681 | 44 | 1665 | 58 | 1647 | 50 | 1634 | 37 |
| | 1652 | 40 | 1644 | 45 | 1629 | 45 | 1610 | 39 |
| cis-NMA cyclic dimer | 1590 | 0 | 1570 | 0 | 1556 | 0 | 1535 | 0 |
| | 1578 | 51 | 1560 | 32 | 1549 | 35 | 1527 | 26 |
| *trans*-NMA dimer 1 | 1632 | 191 | 1600 | 196 | 1586 | 185 | 1561 | 152 |
| | 1601 | 217 | 1586 | 230 | 1572 | 225 | 1549 | 159 |
| *trans*-NMA dimer 2 | 1648 | 224 | 1611 | 228 | 1595 | 221 | 1566 | 170 |
| | 1616 | 191 | 1587 | 186 | 1573 | 179 | 1549 | 136 |

**TABLE 11: Difference between EDF1/6-31+G* and EDF1/6-31+G** Frequencies**

| | formamide (cm$^{-1}$) | formamide dimer 1 (cm$^{-1}$) | formamide dimer 2 (cm$^{-1}$) | formamide dimer 3 (cm$^{-1}$) | formamide dimer 4 (cm$^{-1}$) |
|---|---|---|---|---|---|
| amide A | 29 | 24 | 29 | 28 | 26 |
| | | 24 | 27 | 16 | 23 |
| | 19 | −4 | 17 | 18 | 15 |
| | | −7 | 7 | −8 | 13 |
| amide I | −3 | −6 | −4 | −6 | −4 |
| | | −7 | −4 | −5 | −4 |
| amide II | −18 | −20 | −17 | −19 | −22 |
| | | −19 | −18 | −18 | −18 |
| $\nu$(CN) | −8 | −6 | −7 | −8 | −10 |
| | | −5 | −7 | −7 | −9 |
| $NH_2$ rock | −8 | −8 | −9 | −1 | −8 |
| | | −8 | −7 | −8 | −9 |

**TABLE 12: Effect of the Basis Set on Calculated Frequencies of Formamide**

| | | exptl − calcd (cm$^{-1}$) | | |
|---|---|---|---|---|
| assignment | exptl (cm$^{-1}$) | EDF1/ 6-31+G* | EDF1/ 6-31++G* | EDF1/ 6-31+G** |
| amide A | 3570 | −86 | −84 | −115 |
| | 3448 | −72 | −71 | −90 |
| $\nu$(CH) | 2855 | −62 | −61 | −58 |
| amide I | 1755 | −9 | −9 | −6 |
| amide II | 1580 | −27 | −26 | −9 |
| $\delta$(CH) | 1390 | 0 | 0 | 5 |
| $\nu$(CN) | 1255 | 1 | 2 | 10 |

dimer. The EDF1 calculation predicts that for the *cis* dimer the allowed transition is at 1705 cm$^{-1}$, in excellent agreement with the observed band at 1695 cm$^{-1}$. Although the experimental data are not structurally resolved in terms of dimer conformation, we again see good agreement between EDF1 calculations and experiment.

The correlation between the EDF1/6-31+G* results and experiment suggests that the basis set adequately reproduces the effects of the hydrogen bond. However, we have investigated this more carefully. Diffuse and polarization functions can be important in describing hydrogen bonds,[65] so we have examined the 6-31++G* and 6-31+G** basis sets. Table 11 shows a selection of calculated vibrational modes that are sensitive to the basis set. Adding polarization functions causes moderate changes in formamide frequencies for modes involving hydrogen atoms; e.g., the amide A mode exhibits ∼20−30 cm$^{-1}$ changes, yet the amide I band is virtually unchanged. The new values are not, however, unilateral improvements on the 6-31+G* results (see Table 12). Although the predicted amide II frequency is much closer to experiment, the amide A and $\nu$(CN) bands are worse. The monomer changes and dimer changes are, in general, very similar; the largest differences occur in the amide A modes. The small change moving from monomer to dimer suggests that the extra polarization functions do not significantly improve the representation of the hydrogen bond. Coincidentally, the majority of the differences occur for modes that are IR inactive. Diffuse functions make negligible differences for the formamide monomer and dimers. From these data we conclude that the 6-31+G* basis set is adequate for describing the hydrogen bond interactions in these dimers.

**Covalent Dipeptides.** Knowledge of small peptide units is the first rung on the ladder to calculation of large polypeptide/protein IR spectra. The next important step is to understand how peptide groups interact when bonded directly together. The simplest systems to study are the covalent dipeptides. Many such systems have been studied at varying levels of theory. To maintain consistency with the monomer model peptide, *trans*-NMA, we have studied the methyl-capped Ac-Gly-NHMe and Ac-Ala-NHMe molecules (see Figure 1). All geometries and frequencies are available as Supporting Information. In accord with previous dipeptide studies,[66] we found three vibrational modes that contain large contributions from the NH in-plane bend and CN stretch. All three are assigned as amide III in Tables 14 and 15. The third calculated band has no experimental data; hence absolute energies are shown in parentheses.

*Ac-Gly-NHMe.* The potential energy surface for Ac-Gly-NHMe has two minima at the EDF1/6-31+G* and B3LYP/6-31+G* levels: the fully extended $C_5$ conformer and the internally hydrogen bonded $C_7$ conformer. Figure 4 shows EDF1/6-31+G*-optimized structures. Previous calculations using HF/TZP identified the $C_5$ conformer to be the global minimum-energy structure.[67] Our calculations, both EDF1/6-31+G* and B3LYP/6-31+G*, show that the $C_7$ structure has lower energy. This makes intuitive sense due to the internal hydrogen bond formation, but the energy difference is less than 2kJ mol$^{-1}$. The B3LYP and EDF1 geometries are in good agreement with ∼2° differences between the $\varphi$ and $\psi$ dihedral angles (Table 13).

Table 14 shows the calculated values of the amide I−III bands alongside experimental data from argon matrix isolated studies.[68] For the amide I band EDF1/6-31+G* is in good agreement with experiment and somewhat better agreement than B3LYP/6-31+G* as was seen previously. For amides II and III, the $C_7$ conformer EDF1/6-31+G* results are in excellent agreement with experiment, but the $C_5$ conformer results are not as good.

*Ac-Ala-NHMe.* Previous work has used a wide range of methods to study the Ac-Ala-NHMe potential energy surface.[35,63,64,67,69] The most recent work[35] uses BLYP/TZVP+ and MP2/aug-cc-pVDZ methods. Different calculations identify

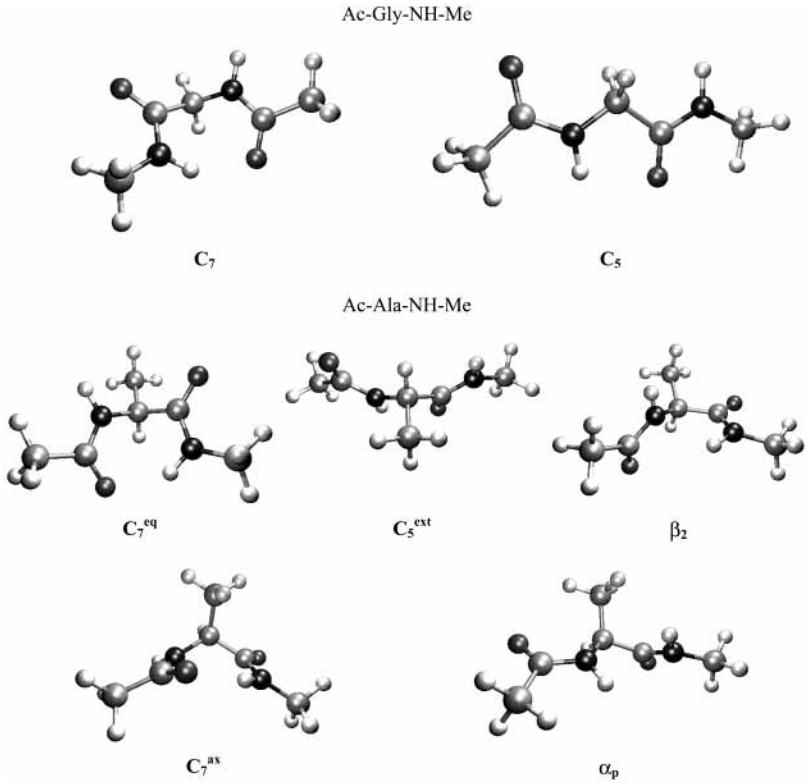


**Figure 4.** Conformations of the dipeptides studied.

**TABLE 13: Geometries and Energies for the Dipeptides Studied**

| species | method | energy (au) | $\phi$ (deg) | $\psi$ (deg) |
|---|---|---|---|---|
| | Ac-Gly-NH-Me | | | |
| $C_7$ | EDF1/6-31+G* | −456.6180044 | −83.1 | 71.3 |
| | B3LYP/6-31+G* | −456.5608526 | −82.1 | 69.1 |
| | HF/TZP[67] | −453.9435886 | −85.4 | 75.5 |
| $C_5$ | EDF1/6-31+G* | −456.5600361 | 180.0 | 180.0 |
| | B3LYP/6-31+G* | −456.6178082 | 180.0 | −179.9 |
| | HF/TZP[67] | −453.9442361 | −179.9 | −179.8 |
| | Ac-Ala-NH-Me | | | |
| $C_7^{eq}$ | EDF1/6-31+G* | −495.9340205 | −83.9 | 76.6 |
| | B3LYP/6-31+G* | −495.8784168 | −83.0 | 74.6 |
| | MP2/aug-cc-pVDZ[35] | −494.5635029 | −82.6 | 75.8 |
| | B3LYP/6-31G* [69] | −495.8551600 | −81.9 | 72.3 |
| $C_5^{ext}$ | EDF1/6-31+G* | −495.9329274 | −146.1 | 151.4 |
| | B3LYP/6-31+G* | −495.8766823 | −154.9 | 159.2 |
| | MP2/aug-cc-pVDZ[35] | −494.56046 | −161.1 | 155.5 |
| | B3LYP/6-31G* [69] | −495.85288 | −157.3 | 165.3 |
| $\beta_2$ | EDF1/6-31+G* | −495.9307948 | −116.0 | 12.9 |
| | B3LYP/6-31+G* | −495.8740482 | −113.7 | 12.4 |
| | MP2/aug-cc-pVDZ[35] | −494.55854 | −82.3 | −9.5 |
| | B3LYP/6-31G* [69] | −495.85009 | −135.9 | 23.4 |
| $C_7^{ax}$ | EDF1/6-31+G* | −495.9298560 | 72.3 | −54.9 |
| | B3LYP/6-31+G* | −495.8744567 | 73.1 | −55.2 |
| | MP2/aug-cc-pVDZ[35] | −494.55987 | 73.7 | −53.7 |
| | B3LYP/6-31G* [69] | −495.85105 | 73.8 | −60.0 |
| $\alpha_p$ | EDF1/6-31+G* | −495.9240066 | −159.8 | −49.9 |
| | B3LYP/6-31+G* | −495.8679065 | −164.7 | −44.1 |
| | MP2/aug-cc-pVDZ[35] | −494.55312 | −164.7 | −38.3 |
| | B3LYP/6-31G* [69] | −495.84423 | −169.4 | −37.8 |

different conformers to be minima on the potential energy surface, although they all agree that the two lowest energy conformers are $C_7^{eq}$ and $C_5^{ext}$. We optimized the previously reported structures using B3LYP/6-31+G* and EDF1/6-31+G*. Table 13 shows the geometries and energies for each conformer located. The ordering of the two lowest energy structures is the same for all methods although the geometries differ slightly. However, the energy order for the $\beta_2$ and $C_7^{ax}$ structures depends on the method used. Large deviations in the $\beta_2$ conformer structure observed by Vargas et al.[35] were ascribed to a flat region in the potential energy surface for which relatively large changes in $\varphi$ and $\psi$ caused little energy change. Table 14 shows vibrational data for the two lowest energy conformations where experimental matrix isolated data are available. For the lowest energy $C_7^{eq}$ conformation, EDF1 performs excellently for amide I and amide II. For amide III there is a discrepancy between theory and experiment. For the $C_5^{ext}$ conformer the amide I value is in good agreement, but the amide II and III results for both B3LYP and EDF1 are a long way from the experimental values. One possible issue is the assignment of experimental bands. This was achieved by a combination of deuteration studies and comparison with model compounds. Since the experimental argon matrix contained a mixture of the two lowest energy conformers, assignments to individual molecules and groups within those modes can be difficult. As the authors acknowledge, the assignments for the amide III bands are tentative.[68] Table 15 contains frequency data for the three conformations without experimental data.

**Timings.** For the molecules studied, EDF1 calculations are more accurate than MP2 and B3LYP. The times taken for the frequency calculations (Table 16) show that EDF1 has another benefit over MP2 and B3LYP. The EDF1 calculation takes on average a quarter of the time required for the MP2 calculation. It is up to 15% faster than the B3LYP calculation for the same basis set and can be up to 4 times faster than B3LYP with a larger basis set. We have not reported geometry optimization times, since the quality of the starting guess geometry affects the time required for the calculation. However, it is the frequency calculation that is the rate-limiting step in these calculations. For comparison, a coupled cluster geometry optimization and frequency calculation on formamide were carried out with the cc-pVDZ basis set. The frequency calculation took 474 h and yielded amide I and II frequencies of 1871 and 1628 $cm^{-1}$, respectively. The agreement with experimental

**TABLE 14: Calculated Frequencies and Intensities for the Amide I−III Modes in the Ac-Gly-NHMe and Ac-Ala-NHMe Dipeptides**

Ac-Gly-NH Me

| | $C_7$ | | | | | $C_5$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | exptl $E$ | B3LYP/6-31+G* | | EDF1/6-31+G* | | exptl $E$ | B3LYP/6-31+G* | | EDF1/6-31+G* | |
| mode | ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| amide I | 1707 | −46 | 460 | −14 | 418 | 1707 | −48 | 108 | −15 | 123 |
| | 1683 | −40 | 154 | −3 | 143 | 1693 | −40 | 461 | −4 | 404 |
| amide II | 1553 | −41 | 206 | −9 | 153 | 1516 | −60 | 123 | −40 | 135 |
| | 1516 | −42 | 166 | −5 | 133 | 1496 | −47 | 457 | −9 | 301 |
| amide III | 1288 | −38 | 22 | −4 | 35 | 1271 | −11 | 17 | 23 | 19 |
| | 1271 | −35 | 92 | 0 | 74 | 1246 | −14 | 93 | 19 | 102 |
| | N/A | (1264) | 27 | (1234) | 29 | N/A | (1253) | 0 | (1227) | 0 |

Ac-Ala-NH Me

| | $C_7^{eq}$ | | | | | $C_5^{ext}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | exptl $E$ | B3LYP/6-31+G* | | EDF1/6-31+G* | | exptl $E$ | B3LYP/6-31+G* | | EDF1/6-31+G* | |
| mode | ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | ($cm^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) | $\Delta E$ ($cm^{-1}$) | $I$ (km $mol^{-1}$) |
| amide I | 1705 | −39 | 460 | −16 | 345 | 1705 | −43 | 81 | −10 | 90 |
| | 1680 | −40 | 121 | −1 | 167 | 1688 | −42 | 450 | −8 | 414 |
| amide II | 1550 | −54 | 219 | −9 | 154 | 1513 | −61 | 121 | −38 | 131 |
| | 1513 | −51 | 149 | 1 | 161 | 1496 | −48 | 387 | −13 | 209 |
| amide III | 1281 | −40 | 52 | 25 | 55 | 1257 | −21 | 16 | 13 | 21 |
| | 1257 | −41 | 75 | 22 | 51 | 1240 | −9 | 59 | 27 | 66 |
| | N/A | (1200) | 2 | (1168) | 6 | N/A | (1201) | 8 | (1165) | 20 |

**TABLE 15: Vibrational Data for the Three Higher Energy Conformers of Ac-Ala-NHMe**

| | | amide I | | amide II | | amide III | | |
|---|---|---|---|---|---|---|---|---|
| | | | | $\beta_2$ | | | | |
| EDF1/6-31+G* | $E$ ($cm^{-1}$) | 1720 | 1714 | 1537 | 1515 | 1235 | 1232 | 1159 |
| | $I$ (km $mol^{-1}$) | 252 | 291 | 136 | 97 | 114 | 58 | 9 |
| B3LYP/6-31+G* | $E$ ($cm^{-1}$) | 1759 | 1746 | 1566 | 1539 | 1274 | 1267 | 1192 |
| | $I$ (km $mol^{-1}$) | 277 | 302 | 196 | 176 | 51 | 117 | 1 |
| | | | | $C_7^{ax}$ | | | | |
| EDF1/6-31+G* | $E$ ($cm^{-1}$) | 1711 | 1683 | 1568 | 1527 | 1283 | 1266 | 1174 |
| | $I$ (km $mol^{-1}$) | 421 | 105 | 194 | 117 | 63 | 63 | 0 |
| B3LYP/6-31+G* | $E$ ($cm^{-1}$) | 1744 | 1720 | 1604 | 1564 | 1321 | 1298 | 1201 |
| | $I$ (km $mol^{-1}$) | 460 | 121 | 219 | 149 | 52 | 75 | 1 |
| | | | | $\alpha_p$ | | | | |
| EDF1/6-31+G* | $E$ ($cm^{-1}$) | 1725 | 1717 | 1530 | 1503 | 1255 | 1241 | 1161 |
| | $I$ (km $mol^{-1}$) | 242 | 255 | 121 | 121 | 81 | 88 | 5 |
| B3LYP/6-31+G* | $E$ ($cm^{-1}$) | 1764 | 1758 | 1562 | 1533 | 1291 | 1274 | 1191 |
| | $I$ (km $mol^{-1}$) | 1264 | 265 | 264 | 100 | 70 | 101 | 5 |

**TABLE 16: Run Times for NMA Dimer Frequency Calculations[a]**

| dimer | MP2/ 6-31+G* (h) | B3LYP/ 6-31+G* (h) | B3LYP/ 6-311++G(2d,2p) (h) | EDF1/ 6-31+G* (h) |
|---|---|---|---|---|
| cyclic *cis*-NMA | 132.5 | 34.8 | 125.5 | 34.4 |
| *trans*-NMA 1 | 110.0 | 34.4 | 123.5 | 29.2 |
| *trans*-NMA 2 | 119.6 | 31.9 | 112.1 | 29.1 |

[a] On a single processor of a Dual Xeon 1.7 GHz Linux box with 1 GB of RAM.

gas-phase data is relatively poor, but this is understandable since it is known that large basis sets are required to realize improved accuracy with coupled cluster methods.[70] Clearly such calculations, especially on larger amides, are beyond current computational resources.

## Conclusion

We calculated the theoretical vibrational frequencies of a series of small monoamides using MP2 and DFT (with the B3LYP and EDF1 functionals). The results were compared to existing experimental data, and it was found that EDF1/6-31+G* calculations gave the most accurate results. MP2 and B3LYP calculations consistently gave poorer results. This conclusion is in agreement with unpublished data from Besley et al., who saw similar improvements in the predicted IR for a range of aromatic aldehydes and ketones.[46] We calculated structures and vibrational data for some hydrogen-bonded formamide and NMA dimers. The experimental data available suggest that EDF1 is again the superior calculation, although further experimental data would be desirable. We investigated how extra p-type polarization functions on hydrogen atoms and extra diffuse functions affected the representation of hydrogen bonding in the dimer systems and found that the 6-31+G* basis set in conjunction with EDF1 is quite adequate. We also calculated minimum-energy structures for the Ac-Gly-NHMe and Ac-Ala-NHMe dipeptides. We found EDF1/6-31+G* and B3LYP/6-31+G* structures to be in agreement with other calculations. Frequency calculations for each of the minima are reported and compared to experimental data where available. As before EDF1/6-31+G* yields harmonic vibrational frequencies in better agreement with experiment than the B3LYP/6-31+G* calculation.

We have presented calculations on isolated monoamides and

dimers and have compared these with gas-phase experimental data. For oligopeptides, the IR spectra are most often obtained in the solution phase. These can vary significantly from isolated gas-phase spectra. The next step is therefore to model solvent effects in these calculations. Several studies have shown that good agreement between experiment and calculation can be achieved by combining implicit and explicit solvent models.[34,71,72] We are currently pursuing this line of research.

**Acknowledgment.** We thank the EPSRC for a studentship to support T.M.W. We thank Professor P. M. W. Gill, Dr. M. W. George, and Dr. N. A. Besley for useful discussions, sharing of unpublished data, and a preprint[46] prior to publication. We thank Dr. A. T. B. Gilbert for useful discussions.

**Supporting Information Available:** Tables of optimized geometries and calculated vibrational frequencies of the monomers and dimers. This information is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

(1) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181.
(2) Wi, S.; Pancoska, P.; Keiderling, T. A. *Biospectroscopy* **1998**, *4*, 93.
(3) Simonetti, M.; Di Bello, C. *Biopolymers* **2001**, *62*, 95.
(4) Huang, C. Y.; Getahun, Z.; Wang, T.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2001**, *123*, 12111.
(5) Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. *J. Phys. Chem. B* **2002**, *106*, 487.
(6) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. *Biochemistry* **1996**, *35*, 691.
(7) Colley, C. S.; Griffiths-Jones, S. R.; George, M. W.; Searle, M. S. *Chem. Commun.* **2000**, 593.
(8) Herberhold, H.; Winter, R. *Biochemistry* **2002**, *41*, 2396.
(9) Colley, C. S.; Clark, I. P.; Griffiths-Jones, S. R.; George, M. W.; Searle, M. S. *Chem. Commun.* **2000**, 1493.
(10) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 12048.
(11) Polavarapu, P. L.; Deng, Z. Y.; Ewig, C. S. *J. Phys. Chem.* **1994**, *98*, 9919.
(12) King, S. T. *J. Phys. Chem.* **1971**, *75*, 405.
(13) Wong, M. W.; Wiberg, K. B. *J. Phys. Chem.* **1992**, *96*, 668.
(14) Knudsen, R.; Sala, O.; Hase, Y. *J. Mol. Struct.* **1994**, *321*, 187.
(15) Stalhandske, C. M. V.; Mink, J.; Sandstrom, M.; Papai, I.; Johansson, P. *Vib. Spectrosc.* **1997**, *14*, 207.
(16) Kang, Y. K. *J. Mol. Struct.: THEOCHEM* **2001**, *546*, 183.
(17) Besley, N. A.; Brienne, M.-J.; Hirst, J. D. *J. Phys. Chem. B* **2000**, *104*, 12371.
(18) Hirst, J. D.; Persson, B. J. *J. Phys. Chem. A* **1998**, *102*, 7519.
(19) Herrebout, W. A.; Clou, K.; Desseyn, H. O. *J. Phys. Chem. A* **2001**, *105*, 4865.
(20) Lundell, J.; Krajewska, M.; Rasanen, M. *J. Phys. Chem. A* **1998**, *102*, 6643.
(21) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(22) Hehre, W. J.; Radom, L.; Schleyer, P. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
(23) Kydd, R. A.; Dunham, A. R. C. *J. Mol. Struct.* **1980**, *69*, 79.
(24) Kutzelnigg, W.; Mecke, R. *Spectrochim. Acta* **1962**, *18*, 549.
(25) Samdal, S. *J. Mol. Struct.* **1998**, *440*, 165.
(26) Ataka, S.; Takeuchi, H.; Tasumi, M. *J. Mol. Struct.* **1984**, *113*, 147.
(27) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.
(28) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(29) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
(30) Cheam, T. C.; Krimm, S. *Spectrochim. Acta, Part A* **1984**, *40*, 481.
(31) Colominas, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem. A* **1999**, *103*, 6200.
(32) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963.
(33) Torii, H.; Tatsumi, T.; Kanazawa, T.; Tasumi, M. *J. Phys. Chem. B* **1998**, *102*, 309.
(34) Kubelka, J.; Keiderling, T. A. *J. Phys. Chem. A* **2001**, *105*, 10922.
(35) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 3213.
(36) Kobko, N.; Paraskevas, L.; del Rio, E.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2001**, *123*, 4348.
(37) Nandini, G.; Sathyanarayana, D. N. *J. Mol. Struct.: THEOCHEM* **2002**, *579*, 1.
(38) Kearley, G. J.; Johnson, M. R.; Plazanet, M.; Suard, E. *J. Chem. Phys.* **2001**, *115*, 2614.
(39) Mirkin, N. G.; Krimm, S. *J. Phys. Chem. A* **2002**, *106*, 3391.
(40) Aleman, C. *J. Phys. Chem. A* **2002**, *106*, 1441.
(41) Csaszar, A. G.; Perczel, A. *Prog. Biophys. Mol. Biol.* **1999**, *71*, 243.
(42) Maple, J. R.; Hwang, M. J.; Jalkanen, K. J.; Stockfisch, T. P.; Hagler, A. T. *J. Comput. Chem.* **1998**, *19*, 430.
(43) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.
(44) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; Defrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A.; Hout, R. F.; Hehre, W. J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1981**, 269.
(45) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
(46) Besley, N. A.; Colley, C. S.; Yang, J.; George, M. W.; Gill, P. M. W. Manuscript in preparation.
(47) Adamson, R. D.; Gill, P. M. W.; Pople, J. A. *Chem. Phys. Lett.* **1998**, *284*, 6.
(48) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(49) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(50) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.
(51) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.
(52) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(53) Spartan '02, Wave function Inc., Irvine, CA, 2001.
(54) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y. H.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W. M.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Voorhis, T.; Oumi, M.; Hirata, S.; Hsu, C. P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532.
(55) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: New York, 1999.
(56) Hamm, P.; Lim, M. H.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123.
(57) Paizs, B.; Suhai, S. *J. Comput. Chem.* **1998**, *19*, 575.
(58) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782.
(59) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1990**, *19*, 553.
(60) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506.
(61) Murray, C. W.; Handy, N. C.; Laming, G. J. *Mol. Phys.* **1993**, *78*, 997.
(62) Lebedev, V. I. *Zh. Vychisl. Mat. Mat. Fiz.* **1975**, *15*, 48.
(63) Jalkanen, K. J.; Suhai, S. *Chem. Phys.* **1996**, *208*, 81.
(64) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2000**, *256*, 15.
(65) Del Bene, J. E.; Jordan, M. J. T. *J. Mol. Struct.: THEOCHEM* **2001**, *573*, 11.
(66) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; K., G.; Mroz, P. A.; Kozlowski, P. M. *J. Phys. Chem. B* **2002**, *106*, 4294.
(67) Bohm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.
(68) Grenie, Y.; Avignon, M.; Garrigou-Lagrange, C. *J. Mol. Struct.* **1975**, *24*, 293.
(69) Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. *Chem. Phys.* **2001**, *263*, 203.
(70) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.
(71) Han, W. G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587.
(72) Knapp-Mohammady, M.; Jalkanen, K. J.; Nardi, F.; Wade, R. C.; Suhai, S. *Chem. Phys.* **1999**, *240*, 63.