Available online at www.sciencedirect.com
SCIENCE DIRECT
Chemical Physics 323 (2006) 87–101
Chemical Physics
www.elsevier.com/locate/chemphys

Reducing the vibrational coupling network in *N*-methylacetamide as a model for ab initio infrared spectra computations of peptides

M. Bounouar, Ch. Scheurer *

Lehrstuhl für Theoretische Chemie, Technische Universität München, Lichtenbergstraße 4, 85748 Garching, Germany

Received 24 April 2005; accepted 16 August 2005
Available online 6 October 2005

Abstract

The *N*-methylacetamide (NMA) molecule is studied as a model to develop scalable algorithms for the simulation of multi-dimensional infrared (IR) spectra of peptides and proteins. Anharmonic vibrational frequencies for the fundamental transitions of NMA in the $trans_{ct}$ conformation are computed using the vibrational self-consistent field (VSCF) method based on potential energy points from second-order Møller–Plesset (MP2) ab initio computations. Dual level schemes, employing coupling potentials evaluated at a lower level of theory, are successfully applied to reduce the total computational cost. Especially, the semi-empirical PM3 method is found to perform well for the evaluation of coupling terms. New selection schemes are introduced that reduce the number of coupling potentials needed in the VSCF procedure without significant loss in accuracy for the frequencies of the spectroscopically relevant amide-II, amide-I, and NH stretch (amide-A,B) modes.

1. Introduction

The recent success in the experimental realization of coherent multi-dimensional infrared (IR) spectroscopy provides a powerful new tool to study structure and dynamics of proteins and possibly other biomolecules with a temporal resolution down to the sub-picosecond regime [1–3]. This allows us to study transient effects in systems far from their equilibrium state [4,5] prepared by any of the ultrafast initiation techniques (photo-switches, pH jump, local electric field, etc.) developed recently [6]. These experiments will deepen our understanding of fundamental biochemical processes in the realm of protein folding and function.

Infrared absorption spectra of peptides and proteins are dominated by vibrational bands that can be described in terms of oscillators localized on each repetitive unit and their mutual couplings [1,7]. The most extensively studied bands, denoted amide-A and amide-B (essentially the NH stretch) in the 3000–3500 cm^{-1} region and amide-I and amide-II between 1500 and 1700 cm^{-1}, are spectrally well separated from the remaining spectrum and exhibit a strong dependence on the structural motifs present in the investigated biomolecule. The amide-I vibrational mode, which involves mainly the CO stretch coordinate with contributions from the CN stretch and NH bend motions [7], has experimentally been the most important mode due to its large transition dipole moment and because it appears to be mostly decoupled from the remaining vibrations in proteins.

Conventional one-dimensional linear IR absorption experiments yield broad unresolved bands which contain only sufficient information to qualitatively or semi-quantitatively assign a relative ratio of secondary structure motifs (i.e., α-helix versus β-sheet) contained in a particular protein. Multidimensional IR spectroscopy has the potential to disentangle the congested vibrational spectra of biomolecules to some extent similar to multidimensional NMR [8–11] but with significantly higher temporal resolution. In nonlinear multidimensional spectra the structural and dynamical information is typically present in terms of diagonal- and cross-peak locations, shapes, and intensities and

* Corresponding author.
E-mail addresses: mehdi.bounouar@ch.tum.de (M. Bounouar), christoph.scheurer@ch.tum.de (Ch. Scheurer).

0301-0104/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.chemphys.2005.08.037

their respective temporal evolution. The structure dependence of the coupling between the amide-I modes of two adjacent peptide planes has for example been used experimentally to determine the φ, ψ-backbone dihedral angles in a small peptide molecule [12,13]. The inversion of the spectroscopic data was based on a two-dimensional coupling strength surface obtained from quantum chemical calculations with the φ, ψ-backbone dihedral angles as variables [14]. The interpretation of analogous data for larger peptides or proteins in terms of a dynamical model of the biomolecule requires extensive theoretical modeling.

Several models and computational strategies have recently been introduced that generally aim at representing either only the amide-I or a slightly larger vibrational subspace of the vibrational modes of the protein [8,15–21]. The vibrational couplings are often classified as either *through-space* or *through-bond* [11,14,22–25]. Through-space interactions are typically electrostatic in nature and have been successfully described using multipole or even dipole–dipole interaction models [7,17,26]. The through-bond interactions are typically comparatively short ranged and need to be computed employing ab initio quantum chemical methods [14,27]. This applies to interactions between nearest neighbor peptide groups and especially to vibrations within a single peptide moiety. A detailed analysis of the anharmonic couplings present within a single peptide plane is called for in view of the wide-spread use of vibrational model Hamiltonians for the amide bands that are based on the assumption of a separation of a set of spectroscopically relevant vibrations from the remaining protein modes.

In this work, we study the anharmonic vibrational potential energy surface (PES) of *N*-methylacetamide (NMA) in the vicinity of the $trans_{ct}$ minimum as the simplest model which exhibits most features of the local vibrational system of a single peptide plane in proteins [7]. We will investigate in particular the through-bond coupling network in NMA based on the computation of vibrational wave functions within the framework of the vibrational self consistent field method (VSCF) which was developed predominantly by the groups of Bowman and Gerber over the past two decades [28–30]. The VSCF method yields approximate frequencies and wavefunctions for the anharmonic vibrational system that can also be used to simulate multi-dimensional IR spectra by sum-over-states or propagation techniques [31]. NMA has been studied recently using the correlation-corrected vibrational self-consistent field (CC-VSCF) method [32–34], but a systematic approach to the analysis and reduction of the vibrational coupling network within NMA that aims at scalability of ab initio VSCF computations to larger peptides and proteins, while retaining the spectroscopically relevant contributions has not been presented so far. The feasibility of the generation of ab initio PES for the simulation of vibrational spectra of larger peptides and proteins relies on a systematic procedure to reduce the number of coupling terms in the PES expansion. Because of the close relationship of the VSCF to the MCTDH method [35], any PES model that is computationally well suited for VSCF will also be usable in the MCTDH context for the computation of time-dependent multidimensional spectra.

The structure of the current paper is as follows. Section 2.1 will shortly review the VSCF method and the PES expansion as it is used in the remaining sections of this work. Section 2.2 provides details about the underlying quantum chemical calculations. General features of the vibrational spectrum of NMA are discussed in Section 3.1, and some of the problems encountered within the standard VSCF approach are indicated. Dual-level computations which reduce the computational cost of calculating coupling potentials are presented in Section 3.2 and compared to a recently suggested method which is based on scaled semi-empirical potentials [36]. Finally, Sections 3.3 and 3.4 treat the problem of partially decoupling the NMA modes while, at the same time, retaining the spectroscopically relevant coupling terms.

2. Computational methods

2.1. Vibrational SCF theory

In the following we outline the key aspects of the VSCF method very briefly as it applies to our current problem. For a recent review on VSCF see [37] and citations therein.

In the VSCF approach the vibrational Schrödinger equation is approximated by a set of coupled effective single mode equations based on the variational treatment of a Hartree Ansatz $\Psi_{\mathbf{n}}(Q_1,\ldots,Q_N)=\prod_{k=1}^{N}\psi_k^{(n)}(Q_k)$ for the vibrational wave function of N coupled anharmonic vibrations [37]. The set of VSCF equations has to be solved iteratively until self-consistency is achieved. The effective one-dimensional Schrödinger equations can be solved by any of several standard grid based techniques [38–40]. Anharmonic vibrational wave-functions and frequencies were obtained using our own implementation of the VSCF procedure [37,41] which is based on the collocation approach with equidistant or Gauss–Hermite grid points and quadrature weights [39]. The latter sometimes result in increased numerical stability for soft strongly anharmonic modes. Correlation corrected vibrational frequencies were computed using the CC-VSCF implementation in GAMESS-US [34,42].

The one-dimensional mean-field potential $\bar{V}_{\mathbf{n},j}(Q_j)$ for each mode Q_j is given by the integral over all remaining degrees of freedom:

$$\bar{V}_{\mathbf{n},j}(Q_j)=\left\langle\prod_{l\neq j}^{N}\psi_l^{(n)}(Q_l)\middle|V(Q_1,\ldots,Q_N)\middle|\prod_{l\neq j}^{N}\psi_l^{(n)}(Q_l)\right\rangle. \quad (1)$$

The evaluation of the multidimensional integrals involving the potential function $V(Q_1,\ldots,Q_N)$ poses the main computational difficulty for large systems. As first suggested in the context of the VSCF method by Carter [43], it is possible to express the PES in terms of a hierarchical expansion:

$$V(Q_1,\ldots,Q_N) = \sum_i^N V_i^{(1)}(Q_i) + \sum_{i<j}^N V_{ij}^{(2)}(Q_i,Q_j) + \sum_{i<j<k}^N V_{ijk}^{(3)}(Q_i,Q_j,Q_k) + \cdots + \sum_{i<j<\cdots<r<s}^N V_{ij\cdots rs}^{(n)}(Q_i,Q_j,\ldots,Q_r,Q_s) + \cdots \quad (2)$$

For typical molecular problems quadruple ($n = 4$) interaction potentials have a negligible influence on the vibrational spectrum [43]. Most generally used models for the description of multidimensional vibrational spectroscopy in proteins do not include any three-mode couplings, i.e., they assume $V^{(n)} = 0$ for all $n > 2$. In the following, we will also include only pairwise interactions between modes, approximating the potential energy surface of the system by an expansion in terms of anharmonic diagonal $V^{(1)}$ and pair-coupling $V^{(2)}$ contributions. In the current context triple and higher contributions ($n \geqslant 3$) to the PES expansion will not be included since our focus is on the development of an approximate method that scales to larger peptide systems. To achieve linear scaling with molecular size it is unavoidable to neglect or approximate even many of the more important pair-couplings and thus only the inclusion of the strongest triple-couplings would be justified. A recent detailed analysis of the effect of triple contributions on high level VSCF computations for small molecules [44] justifies the neglect of many triple contributions $V_{ijk}^{(3)}$ and shows that the triple couplings which contribute most to the anharmonic frequencies involve modes which also exhibit large mutual pair-couplings $V_{ij}^{(2)}$, $V_{ik}^{(2)}$, and $V_{jk}^{(2)}$. It seems thus justified to first analyze the pair-coupling network in NMA before proceeding to approximations involving triple contributions. We are currently investigating models for the extension of our PES expansion by few selected approximate triple-couplings based on a similar pre-screening criterion as in [44]. The use of curvilinear coordinates might also reduce the size of many pair- and most triple-couplings in general. Work in this direction is currently performed in our group [45].

As we are primarily interested in studying the NMA molecule as a model for larger peptides and proteins, overall rotational contributions to the vibrational Hamiltonian are also neglected. This might influence the accuracy of the vibrational transition frequencies in NMA similar to the effects observed in recent high-accuracy computations on formaldehyde and 1,2,5-oxadiazole [44]. We expect the approximations in the treatment of the potential expansion and the choice of quantum mechanical methods for the computation of the PES in this work to have a larger effect on the vibrational spectrum of NMA than the neglected rotational corrections. The Watson correction and Coriolis coupling contributions to the vibrational Hamiltonian for non-rotating ($J = 0$) polyatomic molecules are certainly negligible for our conclusions for the treatment of biomolecules as both contributions scale with the inverse of the moments of inertia tensor of the whole molecule.

2.2. Quantum chemical computations

Geometry optimization and harmonic normal mode analysis were performed using the GAMESS-US program [42], with second-order Møller–Plesset perturbation theory (MP2) employing DZP [46] and cc-pVDZ [47] basis sets. A comparative computation of the anharmonic spectrum based on perturbation theory at the MP2/DZP level was obtained with the Gaussian 03 program [48]. Test calculations with different DFT methods, which are in wide use to successfully simulate IR linear absorption spectra, resulted in harmonic frequencies that were lower than the MP2 frequencies and thus closer to the experimentally observed values. The explicit inclusion of anharmonic effects at the DFT level then resulted in true anharmonic frequencies that were systematically too low. We attributed this effect to an artificial softening of the DFT PES curves which mimics anharmonic effects at the harmonic level. Appropriate scalings of the anharmonic DFT PES curves are currently being investigated and might provide a computationally more efficient method in the future than the current MP2 model.

All potentials were first evaluated on an eight point grid and then interpolated to the final 16 point grids which were used in the collocation treatment. Pair-potentials were evaluated on 8×8 point direct product grids and also interpolated to 16×16 grids by two-dimensional cubic spline interpolation. None of the interpolation and reduction techniques presented in [44] were applied as this would have introduced an independent second source of approximation errors that would have been difficult to separate out. The energy evaluations for the coupling potentials were performed at the MP2/DZP level and, in dual-level computations, also at the HF/STO3-G level or using the semi-empirical PM3 method. Dual level computations will be identified by a short-hand notation of the form D/C where D describes the method used to obtain the diagonal potentials and C the coupling potentials, respectively.

3. Results and discussion

The following results are based on a minimum energy structure of the experimentally more stable *trans* conformer of NMA [49] which was obtained by a standard minimization procedure using MP2 correlation treatment and a DZP basis set. The cartesian normal modes obtained from this structure form the basis for the expansion of the anharmonic PES according to Eq. (2) for all further computations.

3.1. Vibrational spectrum of N-methylacetamide

The harmonic and anharmonic vibrational frequencies of NMA obtained using the non-correlated VSCF method

as well as CC-VSCF at the MP2/DZP level (see Table 1) are in general agreement with an earlier study by Gregurick et al. [32]. The observed frequencies have been taken from the work of Ataka et al. [49] on *cis*-/*trans*-NMA spectra in rare-gas matrices. The possible presence of a matrix shift of up to 40 cm^{-1} introduces an additional uncertainty, which was already pointed out earlier [32]. We are comparing our computations to the 1D spectra because these are the only experimental results that have been obtained in a matrix. All 2D experiments were performed only on selected vibrational bands and in solution where the influence of the solvent is significant [2,50,51] and would have to be included in the model which would introduce an additional source of approximations. For modes 5, 7, and 14, the data used by Gregurick is inconsistent with the original data by Ataka (as shown in Table 1) and discrepancies were also found between their assignment and the more recent work by Herrebout et al. [52]. Our assignment of modes 5, 6, and 7 is based on the comparison of isotope shifts in the spectrum of NMAD (CH_3–CO–ND–CH_3) [49] with the computed anharmonic frequencies for the deuterated compound. The mode which is experimentally found at 439 cm^{-1} exhibits a red-shift and can be most likely identified with mode 5 which involves the N–H out-of-plane bend motion, while modes 6 and 7 are largely unaffected by the isotope substitution. We also find a pronounced isotope effect (≈500 cm^{-1} red shift) on the NH stretch vibration which matches the experimental findings and a transition at 1398 cm^{-1} in the computed spectrum which corresponds to the new dominant band at this frequency that emerges in the NMA spectrum upon *N*-deuteration [49].

The deviations of the computed VSCF from the experimental vibrational frequencies in Table 1 and Fig. 1 exhibit several interesting features. The changes of the root mean square deviation (RMSD) in Table 1 going from harmonic frequencies (RMSD = 133 cm^{-1}), to the inclusion of only diagonal anharmonic potentials (144 cm^{-1}), and finally to the full PES expansion up to pair contributions (55 cm^{-1}) show that a balanced description of diagonal and coupling contributions is of importance in simplified models. On the other hand, the inclusion of vibrational correlation in CC-VSCF has a rather small influence RMSD = 48 cm^{-1} compared to the anharmonic effects captured already by an

Table 1
Vibrational frequencies of NMA based on a MP2/DZP vibrational potential

Assignment	Mode Nr.	Harmonic	Diagonal	VSCF, MP2/DZP			CC-VSCF, MP2/DZP	Observed
				PM3(pp)	STO-3G(pp)	DZP(pp)	DZP(pp)	
NH s	30	3752	3609	3567	3545	3525	3542	3498
NCH_3 as	29	3261	3328	3107	3047	3076	3076	2973
CCH_3 as	28	3248	3316	3089	3021	3057	3032	3008
CCH_3 as	27	3245	3283	3082	3064	3088	3057	3008
NCH_3 as	26	3225	3274	3078	3055	3078	3038	2973
CCH_3 ss	25	3133	3099	2958	3013	3021	2999	2958
NCH_3 ss	24	3124	3085	2962	2995	3003	3009	2915
Amide I	23	1780	1775	1764	1747	1752	1749	1707
Amide II	22	1585	1590	1567	1543	1542	1536	1511
NCH_3 ab	21	1548	1548	1514	1506	1504	1500	1472
CCH_3 ab	20	1526	1525	1483	1493	1486	1484	1432
NCH_3 ab	19	1520	1520	1483	1476	1475	1473	1446
CCH_3 ab	18	1509	1508	1466	1473	1472	1471	1432
NCH_3 sb	17	1484	1487	1462	1463	1456	1454	1419
CCH_3 sb	16	1438	1441	1404	1421	1413	1410	1370
Amide III	15	1310	1314	1297	1287	1283	1284	1266
NCH_3 r CN s NH ipb	14	1208	1211	1212	1196	1190	1186	1168
NCH_3 r	13	1174	1188	1187	1154	1156	1152	–
NCH_3 s CCH_3 r	12	1134	1134	1128	1117	1116	1113	1089
CCH_3 r	11	1074	1082	1072	1063	1061	1051	1037
CCH_3 r CC s NCH_3 s	10	1020	1031	1006	1018	1014	1009	980
NCH_3 r CCH_3 s CO–NH s	9	890	892	887	890	886	883	857
CO ipb CC s	8	634	636	641	638	636	710	658
CO opb CCH_3 r	7	624	630	637	636	631	715	619
CNC d CO ipb CCH_3 r	6	429	433	428	441	438	434	429
CO NH opb	5	364	630	614	560	575	501	439
CNC d	4	270	278	283	291	284	275	279
CNC d NH opb	3	154	203	243	230	215	150	–
CCH_3 d	2	69	381	310	354	352	312	–
NCH_3 d	1	53	347	334	337	337	327	–
rmsd		133	144	63	49	55	48	

Symbols: (pp) = pair-potential, s = stretch, as = asymmetric stretch, ss = symmetric stretch, d = deformation, r = rocking, ab = asymmetric bend, sb = symmetric bend, ipb = in-plane bend, opb = out-of-plane bend.

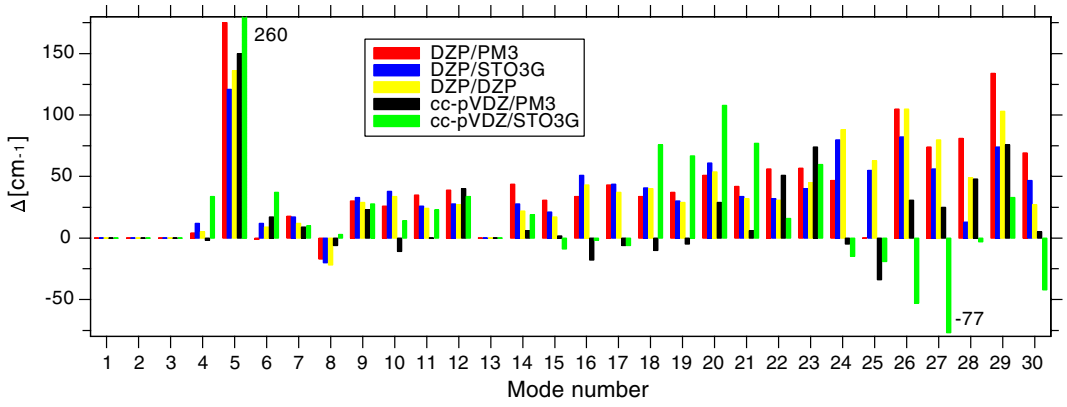

Fig. 1. Difference between VSCF computed and measured frequencies for *N*-methylacetamide (data from Tables 1 and 2). The labels indicate which level of theory was used for the diagonal and the coupling contributions, respectively (see also Section 3.2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

uncorrelated VSCF treatment. The importance of this correction which accounts for correlation between vibrational modes is expected to be reduced even further with increasing system size [34].

The computed frequencies of the methyl group bending (16–21) and stretching (24–29) modes are systematically about 40 cm^{-1} (bends) to 80 cm^{-1} (stretches) too high as can be seen from Fig. 1 (DZP/DZP computation, yellow bars). The neglect of vibrational correlation can most likely be excluded as a reason as the CC-VSCF results in Table 1 indicate. The deviations have been attributed earlier to an insufficient basis set in the ab initio method [32]. Preliminary results from computations at the CCSD(T)/cc-pVTZ level of theory [42] indicate, that the harmonic force constants are approximately by a factor of 0.95 smaller than for the MP2/DZP calculations, while the relative changes in the diagonal third and fourth order anharmonicites are similar in size yet not as uniform [45]. These scaling factors can account for most of the observed deviations. A further source of errors could be in the neglect of triple coupling potentials within the individual methyl groups. These would also partially explain the correlated changes in computed frequencies that occur, e.g., when switching from the DZP/DZP (yellow in Fig. 1) to the cc-pVDZ/STO-3G (green) potentials, where the bend motions shift collectively to more positive frequencies while the stretches are all lowered.

The description of methyl torsional modes (1, 2) in cartesian normal coordinates provides some difficulties. The representative diagonal potential for one of the methyl torsions as a function of the displacement along the respective mode is given in Fig. 2(a). The second derivative of the PES near the minimum is obviously very small for this low–frequency mode and a grid size based on this derivative, as the one in Fig. 2(a) from GAMESS-US [42], yields large displacements along the cartesian normal mode coordinates, with energies dominated by the CH bond stretch at large amplitudes. This explains the large energy values as high as 0.45 Hartree and introduces strong couplings with the high frequency CH stretching modes 24–29. For comparison, the largest energy difference between the *cis–trans* and *trans–cis* methyl conformers does not exceed 8×10^{-4} Hartrees ($\sim$175 cm^{-1}) and an individual methyl rotation barrier can be estimated (without re-optimisation) to be below 0.002 Hartree (500 cm^{-1}). Large positive anharmonic shifts of almost 300 cm^{-1} and a rather poor description of the PES by only a few grid points near the minimum are the result. Perturbation theory yields negative anharmonic shifts of -10 cm^{-1} for mode 2 and -77 cm^{-1} for mode 1, respectively. This shifts mode 1 to a negative frequency. The true values most likely are between these extremes, yet, as no experimental frequencies are available for these modes, a definite answer is currently impossible.

The large couplings to the CH stretch vibrations can also lead to convergence problems in the VSCF procedure. This has been countered before (e.g., in GAMESS-US [42]) by rescaling the pair-coupling potentials at those points where they exceed the energy values of the single mode potentials, a procedure with little physical significance that will also shift the methyl frequencies noticably depending on the rescaling applied. In the present case, it is probably a reasonable assumption that the methyl torsions are almost decoupled and neglect all coupling potentials involving either of these modes. The results are shown in Fig. 3 for several of the potentials from Fig. 1. Most frequencies exhibit some improvement which is also manifested in the lowered RMSD values (DZP/DZP: 58 cm^{-1}, cc-pVDZ/PM3: 44 cm^{-1}, cc-pVDZ/STO3-G 27 cm^{-1}). A systematical improvement is visible especially for the methyl stretch modes 24–29 which is to be expected from the discussion above. Mode 5 also benefits from the decoupling, which is due to a pronounced anharmonic coupling network involving mode 3, which is discussed in detail below.

The preceding discussion of the methyl torsion PES in Fig. 2(a) indicates that appropriate grids have to be chosen carefully for vibrations with substantial anharmonicites that also exhibit large amplitude motions due to a small effective mass. Further modes with this characteristic are

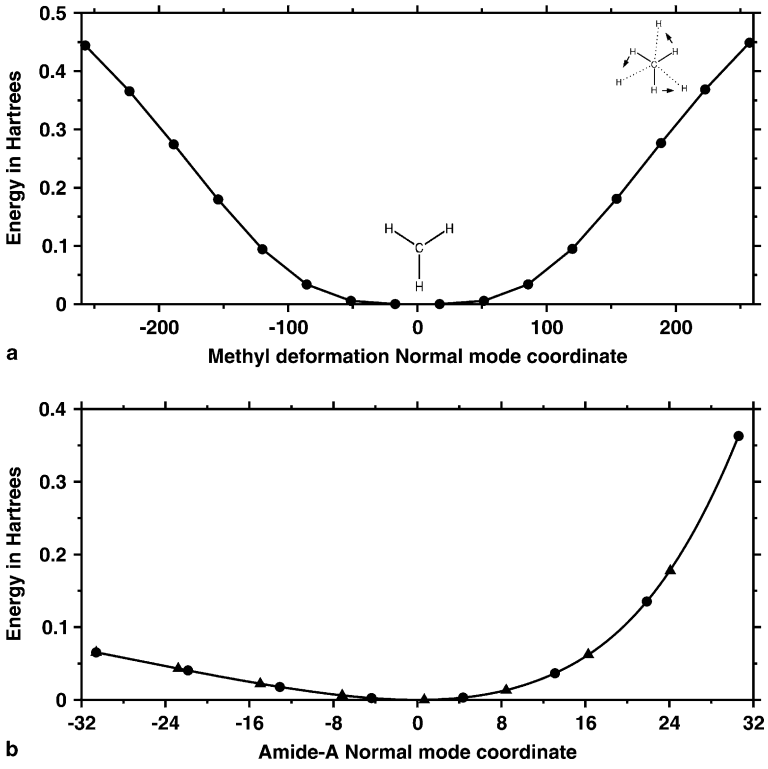

Fig. 2. (a) Diagonal potential for the methyl torsional mode with displacements shown at the equilibrium and the maximally deformed geometry. (b) Diagonal potential for the NH stretch (amide-A) normal mode with the default grid (dots) and an adapted grid (triangles).

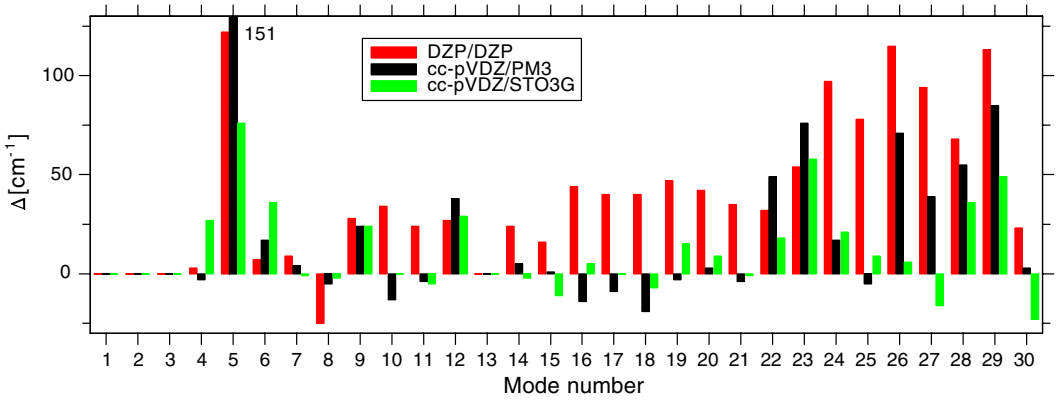

Fig. 3. Difference between VSCF computed and measured frequencies for *N*-methylacetamide for the case where the methyl torsion modes 1 and 2 are decoupled (see also Section 3.1). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

mode 5 and the NH stretch mode 30. Fig. 2(b) presents the diagonal PES along the NH stretch mode with dots indicating grid points for a default grid based on the harmonic approximation of the one-dimensional potential at equilibrium [42]. An improved centered equidistant grid is indicated by triangles. It is based on the explicit evaluation of the diagonal potential and a relative energy threshold with respect to the harmonic frequency to establish the grid range. DFT computations with small basis sets or semi-empirical methods can be used to obtain the PES values efficiently. Grids based on this adaptive approach eliminate most VSCF convergence problems and yield a stable description of strongly anharmonic stretch modes like the NH stretch. Standard grid shrinking techniques employed in combination with moderate extrapolation has also been found to stabilize the VSCF procedure in such cases [44].

The CO NH out-of-plane bend mode 5, where the VSCF frequency deviates by almost 30% relative to the experimental value, is the most problematic mode in the current model. Typical deviations in VSCF frequencies are usually on the order of 3–5%. As mentioned above, the major motion in this mode is a large out-of-plane excursion of the hydrogen atom out of the CNO plane, leading to a pronounced deuterium isotope effect. Its description was

already considerably improved on adapted grids, removing for example serious VSCF convergence problems. The accuracy is limited though by the use of cartesian normal mode coordinates and the neglect of contributions beyond pair couplings in the PES expansion. Possibilities to improve the description of modes with dominant contributions of bend or torsional internal coordinates, like mode 5 or the methyl torsions, based on curvilinear coordinates are currently investigated.

Besides the NMA transition frequencies, also some anharmonicities are available from multidimensional condensed phase IR experiments. Hamm et al. [53] found an experimental value of 16 cm^{-1} for the anharmonicity of the amide I vibration. The current VSCF computations systematically result in smaller values. Using the semi-empirical PM3 pair-potential a value of 5.5 cm^{-1}, with the STO3-G pair potential 9.5 cm^{-1}, and for the full ab initio MP2/DZP 7 cm^{-1} were found for this anharmonicity based on the computation of first and second excited states in the amide-I mode. The corresponding anharmonicity from a perturbational calculation including up to fourth order force constants is 11 cm^{-1}. Diagonal quadratic, cubic, and quartic force-constants obtained from the amide-I diagonal potential ($0.120E_h/(ua_0^2)$, $0.117E_h/(u^{\frac{3}{2}}a_0^3)$, and $0.155E_h/(u^2a_0^4)$) are in very good agreement with those from the perturbative treatment ($0.120E_h/(ua_0^2)$, $0.115E_h/(u^{\frac{3}{2}}a_0^3)$, and $0.150E_h/(u^2a_0^4)$). The discrepancies between VSCF and perturbative anharmonic frequencies are likely due to the different representation of the PES which contains higher order diagonal and pair anharmonicities in VSCF but, on the other hand, triple contributions in the perturbative treatment. Third order force constants involving amide-I and two further different modes are at least one to two orders of magnitude smaller than the anharmonicities already included in the PES expansion used for the VSCF calculation.

The remaining differences with the experiment are likely partially due to a misrepresentation of the anharmonic potentials by the MP2/DZP method, similar, but less pronounced than for the CH vibrations. A further source of deviations is the influence of hydrogen bonding with water in the condensed phase which is not captured by our reduced system. The hydrogen bonding of NMA with water molecules in solution is known to be complex and induces effects that are not understood in detail [54–57]. Further theoretical as well as experimental studies on NMA/water complexes will be needed to elucidate these solvation effects in detail [50].

3.2. Dual level calculations

It has been noted earlier that, while the computation of the harmonic transition frequencies and low-order diagonal anharmonicities require sophisticated ab initio methods to reach near spectroscopic accuracy, it is possible to obtain reasonably accurate inter-mode coupling potentials from simpler methods like semi-empirical or charge interaction models [26]. This fact has been widely used in the simulation of multi-dimensional IR spectra where the experimentally more easily accessible transition frequencies have been used in combination with relatively simple models for the description of the mode couplings.

The potential energy expansion in Eq. (2) allows a straightforward implementation of such a dual level scheme. A potential $V^{(n)}$ is by construction zero at all points at which at least one coordinate is zero. It is therefore possible to combine, e.g., diagonal potentials from higher level ab initio calculations with coupling potentials from a lower level computation. This scheme can also be applied to higher order coupling potentials [44] and even multi level schemes are possible.

The results of dual level computations are shown in Tables 1 and 2 and Fig. 1. The diagonal potentials were obtained from MP2 computations with either DZP (Table 1) or cc-pVDZ (Table 2) basis sets, respectively. For the lower level methods the semi-empirical PM3 model has been chosen due to its encouraging performance in the evaluation of coupling potentials in multidimensional IR spectroscopy [26], as well as a minimal basis (STO-3G) set Hartree Fock (HF) scheme, due to the fact that linear scaling ab initio computations for large biomolecules are performed at similar levels of theory.

The results are encouraging considering the approximately two orders of magnitude reduction in computational effort compared to the accuracy obtained. The overall RMSD from the experimental frequencies in all cases is not significantly different from the value of 55 cm^{-1} obtained with couplings at the MP2/DZP level (yellow bars in Fig. 1). With respect to the RMSD values PM3 as well as HF/STO-3G couplings perform similarly well and which method yields the better result depends on the diagonal potentials. The individual deviations from the experimental frequencies given in Fig. 1 though indicate that PM3 exhibits the more systematic behavior and resembles in general more the results from the DZP/DZP computation. The comparison with the results from a computation in which the methyl rotational modes 1 and 2 have been decoupled (Fig. 3) shows that the HF/STO-3G couplings suffer more severely from the incorrect description of these lowest frequency modes.

For comparison, we have included in Table 3 computations based on a recently suggested rescaling method for semi-empirical potential expansions [36]. These are not dual level computations since only the harmonic frequencies of a higher level method are used to determine a scaling factor (ratio of the harmonic frequencies at both levels) for the individual normal mode coordinate. This leads, in essence, to a tightening or softening of the whole semi-empirical potential depending on the curvature at the potential energy minimum. A combination of rescaled coupling potentials with MP2/DZP diagonal potentials yields a RMSD of 50 cm^{-1} comparable to the full MP2/DZP potential (55 cm^{-1}). The good performance of the fully scaled PM3 potential with a RMSD of 35 cm^{-1} must therefore be

Table 2
Vibrational frequencies of NMA based on a MP2/cc-pVDZ vibrational potential (for symbols, see Table 1)

Assignment	Mode Nr.	Harmonic	Diagonal	VSCF, MP2/cc-pVDZ		Observed
				PM3(pp)	STO-3G(pp)	
NH s	30	3676	3536	3503	3456	3498
NCH_3 as	29	3225	3208	3049	3006	2973
CCH_3 as	28	3224	3218	3056	3005	3008
CCH_3 as	27	3188	3252	3033	2931	3008
NCH_3 as	26	3164	3214	3004	2920	2973
CCH_3 ss	25	3094	3057	2924	2939	2958
NCH_3 ss	24	3078	3032	2910	2900	2915
Amide I	23	1802	1797	1781	1767	1707
Amide II	22	1577	1583	1562	1527	1511
NCH_3 ab	21	1509	1508	1478	1549	1472
NCH_3 ab	20	1498	1499	1461	1540	1446
CCH_3 ab	19	1488	1488	1441	1513	1432
CCH_3 ab	18	1468	1468	1426	1508	1432
NCH_3 sb	17	1438	1441	1409	1413	1419
CCH_3 sb	16	1387	1390	1352	1368	1370
Amide III	15	1281	1285	1268	1257	1266
NCH_3 r CN s NH ipb	14	1178	1180	1174	1187	1168
NCH_3 r	13	1152	1164	1162	1160	–
NCH_3 s CCH_3 r	12	1134	1136	1129	1123	1089
CCH_3 r	11	1046	1053	1037	1060	1037
CCH_3 r CC s NCH_3 s	10	985	995	969	994	980
NCH_3 r CCH_3 s CO–NH s	9	881	883	880	885	857
CO ipb CC s	8	649	652	652	661	658
CO opb CCH_3 r	7	610	616	628	629	619
CNC d CO ipb CCH_3 r	6	434	448	446	466	429
CO NH opb	5	413	605	589	699	439
CNC d	4	286	293	277	313	279
CNC d NH opb	3	148	188	199	257	–
CCH_3 d	2	56	335	325	377	–
NCH_3 d	1	41	278	291	290	–
rmsd		104	107	42	67	

due to the scaling of the diagonal potentials and it is questionable whether this indicates more than a fortuitous cancellation of errors. The known weakness in the semi-empirical description of the rotational barrier of the C–N bond [58,59], which is only corrected by adding a classical molecular mechanics potential energy term in PM3, provides another strong point against the use of this method for the evaluation of diagonal potentials.

3.3. Coupling potential norms and global selection schemes

A common assumption in models for the amide modes is to treat them as effectively decoupled from the remaining vibrational degrees of freedom of the peptide. For localized modes located at well separated sites of a larger biomolecule this approximation can be justified using models of interacting charge densities or classical force-fields. The level of decoupling from vibrations within the same peptide unit though is more difficult to establish.

The PES expansion Eq. (2) lends itself to a study of inter-mode couplings by neglecting certain coupling potentials $V_{ij}^{(2)}$ and measuring the effect on the computed spectrum. In the simplest case, all coupling potentials can be sorted according to the Frobenius norm of their matrix representation on the respective two-dimensional grids. The VSCF computation can then be simplified by neglecting all coupling potentials with a norm below a certain threshold ζ [60]. An appropriate definition of a *coupling strength* or a *coupling potential norm* similar to the simple Frobenius norm is central to any approach that selects coupling potentials without a priori knowledge of their influence on the spectrum. A reasonable definition of such a norm should account for the accessibility of the coupling potential, i.e. include the wave function of the system as a weighting factor. This leads naturally to definitions based on some integral of the form $\langle\psi_i^{(k)}\psi_j^{(l)}|V_{ij}^{(2)}|\psi_i^{(k)}\psi_j^{(l)}\rangle$. These integrals contribute directly to the total energy of the vibrational system if the $\psi_i^{(k)}$ are the converged VSCF single particle functions [44]. A related quantity has been used before to estimate coupling strengths that mediate IVR processes [32].

The evaluation of the above mentioned integrals presupposes the knowledge of the full VSCF wave function which is not available before the PES expansion has been defined. We will use instead the eigenfunctions of the corresponding one-dimensional sub-problems,

$$\left(T_i + V_i^{(1)}\right)\varphi_i^{(k)} = \varepsilon_i^{(n)}\varphi_i^{(k)} \tag{3}$$

Table 3
Vibrational frequencies of NMA based on a MP2/DZP vibrational potential compared to the PM3 scaling method of [36] (for symbols, see Table 1)

Assignment	Harmonic		Scal	VSCF, MP2/DZP		PM3(Scal)	Observed
	MP2/DZP	PM3		DZP(pp)	PM3(pp,Scal)		
NH s	3752	3395	1.11	3525	3540	3490	3498
NCH_3 as	3261	3061	1.07	3076	3098	3055	2973
CCH_3 as	3248	3085	1.05	3057	3074	3042	3008
CCH_3 as	3245	3082	1.05	3088	3033	3013	3008
NCH_3 as	3225	3041	1.06	3078	3033	3009	2973
CCH_3 ss	3133	3179	0.99	3021	2920	2918	2958
NCH_3 ss	3124	3132	1.00	3003	2935	2910	2915
Amide I	1780	1915	0.93	1752	1766	1759	1707
Amide II	1585	1485	1.07	1542	1555	1535	1511
NCH_3 ab	1548	1382	1.12	1504	1502	1481	1472
CCH_3 ab	1526	1401	1.09	1486	1474	1459	1432
NCH_3 ab	1520	1370	1.11	1475	1467	1453	1446
CCH_3 ab	1509	1376	1.10	1472	1448	1455	1432
NCH_3 sb	1484	1348	1.10	1456	1449	1465	1419
CCH_3 sb	1438	1397	1.03	1413	1389	1392	1370
Amide III	1310	1251	1.05	1283	1288	1296	1266
NCH_3 r CN s NH ipb	1208	1056	1.14	1190	1212	1220	1168
NCH_3 r	1174	984	1.19	1156	1180	1197	–
NCH_3 s CCH_3 r	1134	1135	1.00	1116	1124	1134	1089
CCH_3 r	1074	1014	1.06	1061	1067	1057	1037
CCH_3 r CC s NCH_3 s	1020	991	1.03	1014	979	1026	980
NCH_3 r CCH_3 s CO–NH s	890	879	1.01	886	888	894	857
CO ipb CC s	634	584	1.09	636	639	643	658
CO opb CCH_3 r	624	648	0.96	631	635	582	619
CNC d CO ipb CCH_3 r	429	428	1.00	438	429	433	429
CO NH opb	364	459	0.79	575	580	493	439
CNC d	270	272	0.99	284	282	280	279
CNC d NH opb	154	166	0.92	215	236	175	–
CCH_3 d	69	58	1.20	352	260	173	–
NCH_3 d	53	94	0.57	337	163	482	–
rmsd	133	92		55	50	35	

to define the norms used in the following:

$$\left\|V_{ij}^{(2)}\right\|_{kl} = \left|\langle\varphi_i^{(k)}\varphi_j^{(l)}|V_{ij}^{(2)}|\varphi_i^{(k)}\varphi_j^{(l)}\rangle\right|. \quad (4)$$

The coupling potentials and the corresponding norms $\|V_{ij}^{(2)}\|_{00}$ are shown in Fig. 4 for the MP2/DZP PES. The dominance of the couplings within the methyl groups is clearly visible. The strongest couplings are found between each of the pairs of methyl asymmetric stretch modes (27/28 and 26/29) followed immediately by their couplings to the low frequency rotational modes 1 and 2 within the same methyl group. The couplings involving the corresponding symmetric stretch modes 24 and 25 are much less pronounced. A pronounced coupling network is also found between the stretch modes of each methyl group (24, 26, 29 for NCH_3 and 25, 27, 28 for CCH_3) and the bend modes within the same methyl group (17, 19, 21 for NCH_3 and 16, 18, 20 for CCH_3).

The strongest couplings involving any of the spectroscopically important amide modes (amide-II: 22, amide-I: 23, and NH stretch: 30) all involve the NH stretch (or amide-A) mode. It is strongly coupled to mode 5 which is predominantly an NH out-of-plane bend motion of the hydrogen atom. This in turn couples strongly to mode 3 (CNC d NH opb), forming a coupling network between modes 3, 5, and 30 together with the coupling between modes 3 and 30. The NH stretch is also the most important coupling partner of mode 15 which is the amide-III mode. The most pronounced coupling among the amide modes 22, 23 and 30 is between the NH stretch and the amide-II mode (30/23) which has also been observed experimentally in a two-color 2D IR experiment recently [61]. The strongest coupling involving the otherwise mostly decoupled amide-I mode is the one with amide-II (24/23). In a recent dual frequency 2D-IR experiment of NMA in DMSO this coupling was directly observed with a coupling strength of 27 cm^{-1} and a mixed mode anharmonicity of 3.5 cm^{-1} [62] which is in contrast to earlier calculations stating that no coupling between the amide-I and amide-II modes exists [32].

Contour plots of the three coupling potentials between modes 22, 23 and 30 with contour lines at ±5, ±15, …, ±95 cm^{-1} (dashed lines for negative values) are presented in Fig. 5. The lower frequency mode always varies along the horizontal axis and the higher frequency mode along the vertical axis, respectively. All coordinates are mass weighted cartesian normal modes in the range $[-20, 20]$. These ranges can be translated into corresponding values

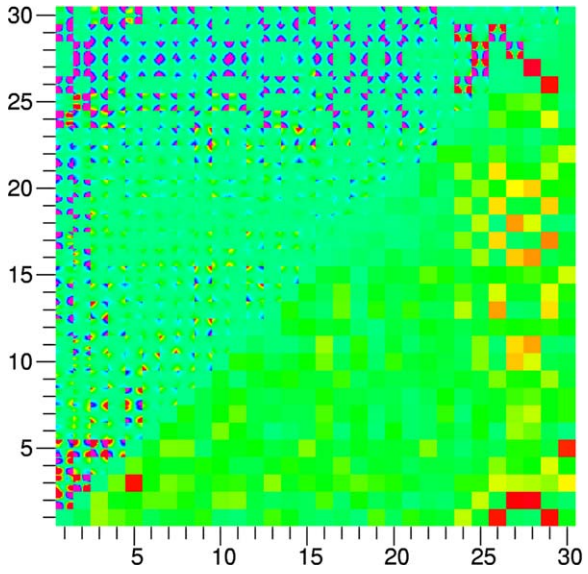

Fig. 4. Upper left triangle: MP2/DZP pair-potential terms $V_{ij}^{(2)}(Q_i, Q_j)$. The potential energy values have been clipped at $\pm 0.015 E_h$ and are color coded with a rainbow palette ranging from violet ($-0.015 E_h$) to red ($+0.015 E_h$). Lower right triangle: relative norms of the coupling potentials $\|V_{ij}^{(2)}\|_{00}$ as defined in Eq. (4). The maximum norm has been scaled to 1 and the color coding on the interval $[0, 1]$ ranges from green to red.

for the dominant bond stretches involved in each mode. For amide-II the CN bond length ranges between 1.3704 ± 0.0584 Å, for amide-I the CO bond length is 1.2363 ± 0.0911 Å, and the NH stretch is in the range 1.0073 ± 0.2556 Å. To gain insight into which parts of the coupling surface are accessible to the system, the squared vibrational ground state product functions $(\varphi_i^{(0)} \varphi_j^{(0)})^2$ which provide the weighting in Eq. (4) are also included in the plots (red contour lines at 30%, 50%, 70%, and 90% of the maximum value). While amide-I is decoupled from the NH stretch motion, we can clearly see its coupling to amide-II which is on the same order of magnitude as reported in the experiment and also exhibits the qualitatively correct anti-correlation that was observed from the shape of the experimental cross peaks [62].

To reduce the number of coupling potentials in the VSCF computation and to study the influence of vibrational coupling semi-quantitatively, a threshold ζ can be applied to the norms given in Fig. 5 and all coupling potentials with a norm $\|V_{ij}^{(2)}\|_{00} < \zeta$ are omitted from the PES expansion [60]. The influence on the total vibrational spectrum can be described by the change of the RMSD of the vibrational frequencies ω_i^{ζ} with respect to the frequencies from the full PES expansion ω_i^{full}. We also divide the RMSD at every value ζ by the RMSD for the diagonal system, corresponding to $\zeta = 0$ (solution of Eq. (3)), to obtain a measure $\Delta(\zeta)$ ranging in $[0, 1]$:

$$\Delta(\zeta) = \frac{\sqrt{\sum_i \left(\omega_i^{\text{full}} - \omega_i^{\zeta}\right)^2}}{\sqrt{\sum_i \left(\omega_i^{\text{full}} - \omega_i^{\text{diag}}\right)^2}}. \tag{5}$$

The error function $\Delta(\zeta)$ evaluated at different threshold values ζ is shown in Fig. 6(b) side by side with the cumulative distribution function of the coupling norms from Fig. 4. The distribution is smooth and continuous for coupling potentials with a relative norm $\zeta \leqslant 0.1$. This is also reflected in a smooth behavior of the error function $\Delta(\zeta)$ up to this value. Above this value it is possible to identify individual couplings and coupling networks which are the same couplings that dominate in Fig. 4. The function $\Delta(\zeta)$ exhibits jumps and discontinuities which can be attributed to these coupling networks. This behavior explains why in [60] the threshold could not be set lower than 0.1 or a value at which about 50% of all couplings were included.

As we can see from Fig. 6 neglecting the pair-potentials that have a norm less than 0.01 reduces the number of pair potentials by about 150 out of a total of 435 pairs with an error Δ of 4%. This deviation rises to 20% when neglecting all mode pairs below 0.05 which implies the neglect of 20% of all mode couplings. These seem to be relatively modest savings in terms of neglected coupling potentials but we have to keep in mind that NMA represents effectively only a single peptide plane and the couplings that are neglected

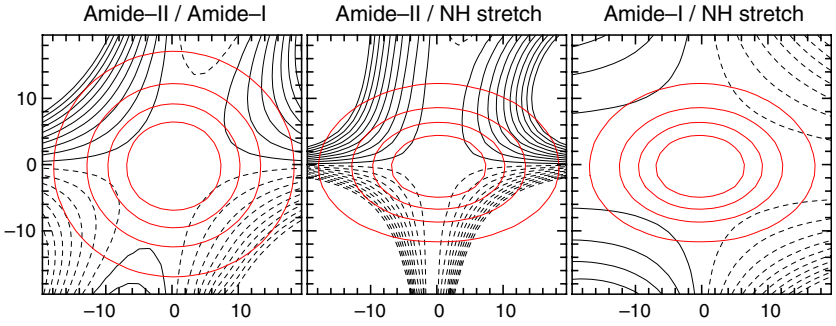

Fig. 5. Detail contour plots of the pair–potentials between the amide-II, amide-I, and NH stretch modes. Contour lines are at $\pm 5, \pm 15, \ldots, \pm 95$ cm^{-1} with dashed lines indicating negative contours. The lower frequency modes vary along the horizontal axis in each plot and mass weighted cartesian normal modes are used as coordinates. The grid ranges translate into a CN stretch of 1.3704 ± 0.0584 Å for amide-II, a CO stretch of 1.2363 ± 0.0911 Å for amide-I, and a NH stretch of 1.0073 ± 0.2556 Å. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

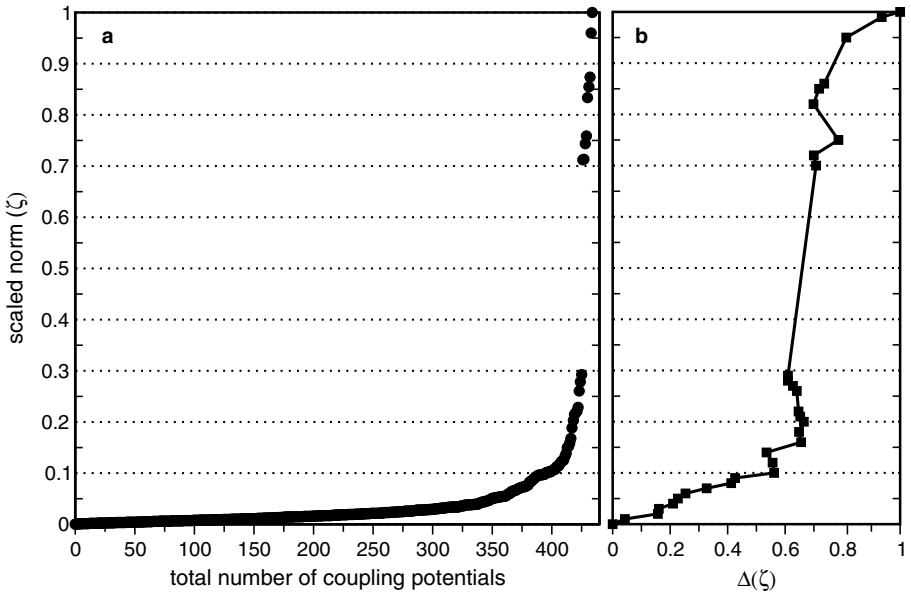

Fig. 6. (a) Distribution of the relative norms of the coupling potentials from Fig. 4 as a function of the total number of coupling potentials. (b) Relative RMSD error of the computed frequencies given by Eq. (5) as a function of varying threshold value ζ.

are couplings within this unit. A further problem lies in the use of a single global threshold parameter which forces us to either include a significant portion of the continuous tail of the distribution function or, alternatively, to loose a major fraction of the weaker couplings, potentially decoupling individual modes completely. Other selection strategies which are more successful in this respect will be introduced in Section 3.4.

So far only the norm defined by Eq. (4) with the ground state wave functions, i.e. $k = l = 0$, was employed. Alternatively, higher excited state wave functions of the one-dimensional sub-spaces of Eq. (3) can be chosen. A comparison of the norms based on the ground and first excited state wave functions (k, $l = 0$ or 1) shows that the norm using the first excited state throughout ($k = l = 1$) assigns the most significant measure to the different coupling potentials with the improved selection schemes discussed in Section 3.4. This becomes clear if we consider the nodal structure of the product functions $\varphi_i^{(1)}\varphi_j^{(1)}$. The product functions possess nodal lines close to the axes $Q_i = 0$ and $Q_j = 0$ where also the coupling potential $V_{ij}^{(2)}$ is zero by construction and it exhibits one extremum in each quadrant of the Q_i–Q_j-plane, where the coupling potential has its most significant contributions. On the contrary, the corresponding ground state product function $\varphi_i^{(0)}\varphi_j^{(0)}$ is maximal where $V_{ij}^{(2)}$ is zero, and it is quickly decaying in the regions where the coupling potential contributes to the effective potential (compare Fig. 5). The mixed product functions with $k = 0$, $l = 1$ or $k = 1$, $l = 0$ present an intermediate case. The relative norms $\|V_{ij}^{(2)}\|_{11}$ for the MP2/DZP coupling

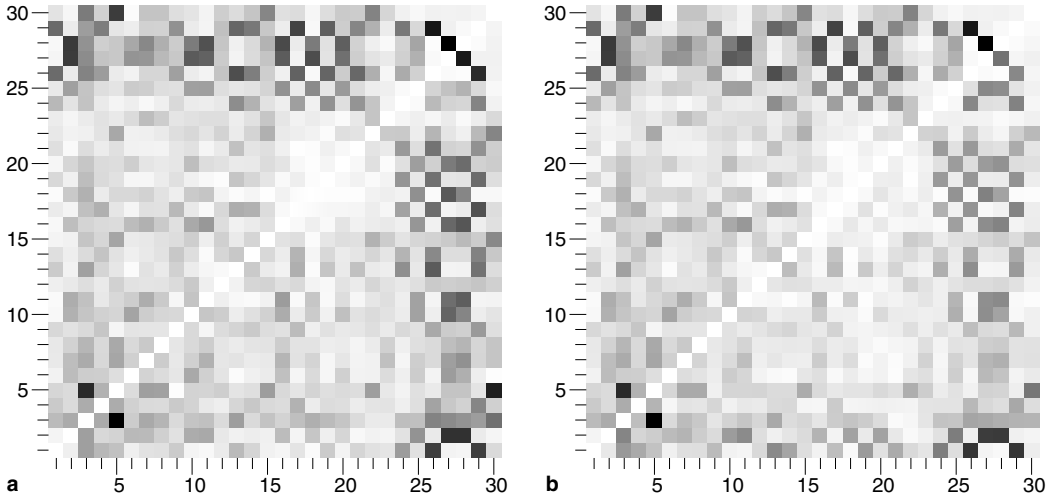

Fig. 7. Comparison of different norms for the coupling potentials. The upper left triangle contains the norm over the DZP potentials as defined in Eq. (4) for $k = l = 1$. The lower right triangle contains the approximate norms N_{ij} from Eq. (6) for (a) the MP2/DZP and (b) the PM3 potentials, respectively.

potentials are shown in Fig. 7(a) in the upper left triangle using a gray scale ranging from 0 (white) to 1 (black) with respect to the largest norm. The general pattern resembles the one from Fig. 4 closely and the important changes are not in the largest, most prominent entries but rather in the relative changes of the intermediate norms.

The computation of the $\|V_{ij}^{(2)}\|_{kl}$ norm according to Eq. (4) requires the values of the coupling potential on a full Q_i–Q_j-grid for the evaluation of the integral. As this is the most computationally demanding operation, it is necessary to approximate the integrals by simpler expressions. The norms $\|V_{ij}^{(2)}\|_{11}$ can be reasonably approximated as

$$\left\|V_{ij}^{(2)}\right\|_{11} \approx N_{ij} = |(b_i - a_i)(b_j - a_j)| \times \sum_{(x,y)\in\{a_i,b_i\}\times\{a_j,b_j\}} \varphi_i^{(1)}(x)\varphi_j^{(1)}(y)V_{ij}^{(2)}(x,y). \tag{6}$$

Here, a_i and b_i are the positions of the extrema of the corresponding one-dimensional wave function $\varphi_i^{(1)}$. In this approximation only four evaluations of the coupling potential are needed to compute N_{ij}. The resulting relative norms for the MP2/DZP coupling potentials are presented in the lower right triangle of Fig. 7(a). The fact that Fig. 7(a) is symmetric with respect to its diagonal indicates that the N_{ij} provide a very accurate approximation to the full norms $\|V_{ij}^{(2)}\|_{11}$. This is confirmed by tests with several of the selection schemes discussed in Section 3.4 in which the reduced norm can be substituted for the full integrals without any effect on the result. A further reduction of the computational cost can be achieved using PM3 to evaluate the coupling potentials at the required set of points. The comparison of the approximate norms N_{ij}^{PM3} with the full norms $\|V_{ij}^{(2)}\|_{11}$ in Fig. 7(b) shows that the agreement is still very good. Noticable deviations occur only for the couplings involving CH stretch modes. This is most likely due to a known weakness of the PM3 method which affects bond stretches involving hydrogen atoms [58,59]. A reasonable measure for the coupling strength can thus be obtained at a dramatically reduced computational cost involving only four PM3 single point energy evaluations compared to, e.g., 64 MP2/DZP computations for the case of the full integrals of Eq. (4) on an 8×8 grid.

3.4. Advanced mode coupling selection schemes

The shortcomings of a global threshold scheme as it was applied in Fig. 6 were already mentioned above. This problem will even become more pronounced in larger molecules, like peptides and proteins, where it might be necessary to include couplings with a smaller norm in a localized sub-unit of the system while long-range couplings with comparable norms have a smaller effect. Furthermore, one is frequently not interested in all vibrational frequencies of a certain molecule but rather in only a selected subset which is of special spectroscopical relevance. This applies in particular to biopolymers in which one is typically interested only in some specific vibrational bands, usually the amide bands, within a certain spectral range. We will thus investigate whether it is possible to reduce the number of coupling potentials in the PES expansion Eq. (2) further without sacrificing the accuracy of a selected subset of all vibrational modes. For the purpose of this paper we select the amide-II, amide-I, and the NH stretch (amide-A,B) vibrations as our set of target modes. Very similar results as the ones discussed below have been obtained though as well, using any single one or pair of these modes as the target set.

The problem of finding the optimal set of a fixed number n out of a total of N coupling potentials is a combinatorial optimization problem for which an exhaustive search scales very unfavorably as $\binom{N}{n}$, with the added difficulty that the best value for n is not known a priori. It is possible though to start with the set of target modes and all their mutual couplings and gradually add more coupling potentials to other modes to this set. This approach leads in essence to algorithms of a build-up clustering type which differ in the particular selection strategy taken to enlarge the set of modes (and couplings). The coupling norm matrices introduced in Section 3.3 take on the role of the distance metric in this case. The results discussed below do not differ whether the full norm integrals of Eq. (4) or the reduced expression Eq. (6) are used. In the following, the MP2/DZP coupling potentials are used throughout. Slight, typically insignificant, changes in the selection process are observed when the reduced norms are based on the PM3 coupling potentials. The deviations in frequencies based on the PM3 coupling potentials are typically also small and on the order of magnitude that would be expected from the discussion in Section 3.1.

Several selection strategies have been examined for our model system, which each depend on a number of choices:

1. The next coupling potential to be included is either selected as: (a) the globally largest coupling or, alternatively, (b) as the largest one of the couplings which connect to a mode that was already included before.
2. When a coupling potential has been selected for inclusion, then either: (a) only this potential is added or, alternatively, (b) the mode to which this newly selected coupling connects is examined and all (or some subset) of this new mode's couplings are also included.
3. A drop-tolerance scheme can be applied if case (b) is used in the preceeding item.

The schemes can be judged based on how quickly the frequencies of the target modes converge to the values of the full computation (including all coupling potentials) in terms of the number of necessary coupling potentials and, additionally, on the smoothness of the approach to this limit. As outlined above, schemes which use a global selection criterion 1(a) are inferior as their convergence is slower and more coupling potentials need to be included

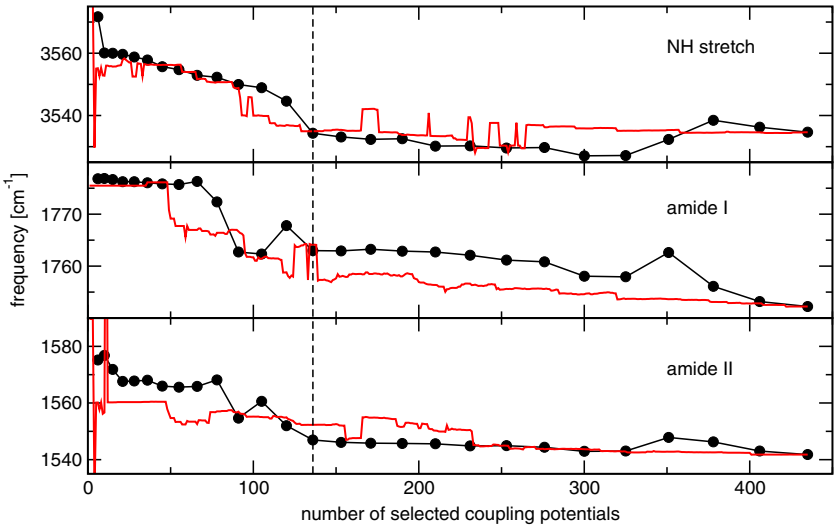

Fig. 8. Convergence of the amide transition frequencies as a function of the number of included pair potentials for different selection schemes (see text). The black line with dots corresponds to the data given in Table 4.

than for choice 1(b) in which the set of modes and couplings that has already been selected is taken into account. Fig. 8 presents the convergence behavior of the transition frequencies of the three selected amide modes for two schemes based on choice 1(b) without a drop-tolerance. The black line with dots represents a selection based on scheme 2(b) with one new mode added at each dot and the red line represents a scheme adding a single potential at each step according to option 2(a). The approach to the limit is considerably smoother for the preferable mode based scheme 2(b) and a stop criterion can be formulated based on the relative change between consecutive steps which stops at the vertical dashed line.

The modes, and their respective couplings, that have been included at this stage are given in Table 4 above the horizontal gap. As the fourth column indicates, a total of 136 coupling potentials have to be included to find frequencies for the three target modes which are stable and close to the converged results for the full PES expansion. The first mode added in this scheme is the problematic mode 5, which indicates its importance for the peptide vibrations, followed by the amide-III mode 15. Furthermore most of the NCH_3 modes are included (stretch 24, 26, 29; bend 17, 21; rocking 13, 14) through their coupling to amide-II (22/26) while only two of the asymmetric CCH_3 modes (bend 19, stretch 28) need to be included. The remaining modes are CNC deformation modes (3, 4) or involve methyl as well as peptide plane motions (9). It is interesting to see that the coupling potential $V^{(2)}_{27,28}$ with the largest relative norm (1.0) is not included at this stage, and that the inclusion of mode 27 does not influence the computed frequencies noticably. Additionally, a number of coupling potentials with comparatively large relative norms between 0.5 and 0.6 are not included and have little influence on the frequencies while the inclusion of, e.g., modes 9, 4, and 3 via much smaller coupling potentials lead to clearly visible changes (see Fig. 8 just left of the dashed line). This demonstrates why global threshold schemes will fail if not a large fraction of all possible coupling potentials is included. We can also see that the changes in the last five points of

Table 4
Convergence of amide frequencies with a mode based selection scheme (see text)

Mode added	Coupling	Relative norm	Total number	Amide-II	Amide-I	NH stretch
5	30	0.46	6	1575	1777	3572
15	30	0.39	10	1577	1777	3560
26	22	0.26	15	1572	1777	3560
29	26	0.91	21	1568	1776	3560
17	29	0.61	28	1568	1776	3559
13	26	0.54	36	1568	1776	3558
24	29	0.52	45	1566	1776	3556
19	29	0.52	55	1566	1776	3555
21	26	0.49	66	1566	1776	3553
14	29	0.42	78	1568	1772	3552
9	29	0.31	91	1555	1763	3550
4	29	0.27	105	1561	1762	3549
3	30	0.25	120	1552	1768	3545
28	15	0.24	136	1547	1763	3534
27	28	1.00	153	1546	1763	3533
16	28	0.58	171	1546	1763	3532
18	27	0.58	190	1546	1763	3533
20	28	0.52	210	1546	1763	3530
11	28	0.52	231	1545	1762	3530
10	27	0.48	253	1545	1761	3530
25	18	0.37	276	1544	1761	3530
7	28	0.31	300	1543	1758	3527
2	28	0.30	325	1543	1758	3527
6	27	0.25	351	1548	1763	3532
1	29	0.21	378	1546	1756	3538
8	28	0.21	406	1543	1753	3536
12	29	0.19	435	1542	1752	3535

The first column gives the index i of the newly added mode, the second column the index j of the coupling potential $V^{(2)}_{ij}$ leading to the selection of mode i, the third column the relative norm of the coupling potential, the fourth column the total number of coupling potentials included at this step, and the remaining columns the computed transition frequencies for the respective mode in units of cm^{-1}.

the black curve in Fig. 8 is due to the inclusion of modes 1 and 2 which seems more likely to be an artifact than any real improvement in the description of the amide modes, as seen from the discussion in Section 3.1.

It is possible to reduce the number of required coupling potentials even further by applying thresholds to the sets of selected modes. As it is clear from the above discussion a global threshold cannot be appropriate. A relative threshold τ though, that selects for each mode only those coupling potentials that have norms larger than τ times the largest coupling in which this mode is involved, yields a stable scheme. Applying, e.g., a relative threshold of $\tau = 0.3$ to the set of 136 couplings discussed above, reduces the total number of couplings to 82 with minor changes in the frequencies (amide-II: $+1\ \mathrm{cm}^{-1}$, amide-I: $+2\ \mathrm{cm}^{-1}$, NH s: $+4\ \mathrm{cm}^{-1}$).

A dual threshold scheme proves to be even more efficient. Here, a larger threshold τ is chosen for all coupling potentials involving at least one of the modes of the target set, while all remaining couplings are selected based on a smaller threshold τ'. With values of $\tau = 0.5$ and $\tau' = 0.25$ this scheme yields a further reduction of the number of coupling potentials down to only 62 with very similar results as the single threshold scheme (amide-II: $+2\ \mathrm{cm}^{-1}$, amide-I: $+3\ \mathrm{cm}^{-1}$, NH s: $+3\ \mathrm{cm}^{-1}$). Both of these threshold methods give comparable results to within $\pm 1\ \mathrm{cm}^{-1}$ if either the full norms according to Eq. (4) ($k = l = 1$) or one of the reduced norms of Eq. (6) with MP2/DZP or PM3 evaluation of the potential energy points are employed. We also found that the use of thresholds in combination with a mode based selection scheme is less sensitive to small variations in the chosen thresholds than for a scheme which only relies on thresholding.

4. Conclusions

We have performed anharmonic vibrational frequency calculations for *N*-methylacetamide using the VSCF method based on a many-body expansion of the PES up to second order in terms of cartesian normal modes. We were able to show that reasonable agreement with experimental data obtained in rare gas matrices can be achieved using ab initio potential energy points evaluated at the MP2/DZP level. The assignment of a previous theoretical study [32] of NMA was slightly corrected based on the computation of isotope shifts in *N*-deuterated NMA. The existence of the couplings between the amide-II mode and the NH stretch and between amide-II and amide-I, which were recently observed in two-dimensional IR two-color experiments, was theoretically confirmed [2]. We could also verify the experimentally observed anti-correlation in the coupling between amide-II and amide-I, as well as the order of magnitude for this coupling.

It was shown that dual level computations in which the diagonal anharmonic potentials along a single vibrational mode are evaluated using a higher level ab initio method than for the coupling potentials provides an efficient route to the computation of PES expansions in the VSCF framework. Especially, the use of the semi-empirical PM3 method seems promising in this respect. We are currently investigating whether the computational effort can be further decreased using interpolation of fewer points based on appropriately chosen analytical model functions for pair- and triple-coupling potentials. This would also allow us to incorporate the few most important triple-coupling potentials at a reasonable computational cost. A comparison of the dual level computation with a recently suggested scaling procedure for semi-empirical PES expansions shows a comparable level of accuracy. In the scaling procedure though, two independent normal mode computations at the higher level of theory and at the semi-empirical level have to be matched to obtain the scaling coefficients for all modes. Particularly for larger molecules with denser spectra this might turn out to be problematic considering that optimized geometries as well as normal modes can exhibit noticable differences between ab initio and PM3 computations. The quality of scaled diagonal anharmonic surfaces based on only the second derivative information at the equilibrium geometry remains to be more thoroughly investigated in general. This might provide an economic route to more precise diagonal potentials from DFT computations combined with few high-level quantum chemical (e.g., CCSD(T)/cc-pVTZ) energy points.

The systematic comparison of different schemes for the selection of relevant pair-coupling potentials yielded an algorithm that allowed us to reduce the number of couplings required for the computation of the amide-II, amide-I, and NH stretch mode to 62. The frequencies obtained with this reduced set of couplings and the corresponding 17 vibrational modes are in good agreement with computations based on the full set of 435 couplings and 30 normal modes in NMA. The relevance of an individual coupling in the selection process could be estimated based on only four evaluations of the coupling potential with the simple semi-empirical PM3 model. For modes belonging to distant sites in larger molecules the selection scheme might even be based on coupling energy evaluations at the level of molecular mechanical force fields or charge interaction models. Work along this line is in progress. It is noteworthy that a meaningful reduction of the anharmonic pair-couplings in the PES expansion is already possible in a small molecule with rather pronounced couplings like NMA. We expect the reduction effects to be even more substantial for larger peptides and proteins, a pre-requisite for ab initio based VSCF computations in biomolecules, where our algorithm allows us to iteratively construct and refine a model Hamiltonian which only includes those degrees of freedom that are important for an accurate description of the few bands that are usually observed in 2D IR experiments.

The description of strongly anharmonic modes and, especially, modes exhibiting large amplitude motions, can still pose problems in the VSCF framework. In NMA these are mostly the methyl group vibrations, but also the NH

stretch and NH out-of-plane vibrations. In the case of the methyl group bends and stretches it might prove necessary to compute diagonal potentials at a higher level of theory than MP2/DZP and to include triple coupling potentials while for the NH stretch mode already the use of an adapted grid that accounts for the anharmonicity was successful. Modifications of the PES expansion that result in a more appropriate description of "soft" modes like the methyl torsions and the NH out-of-plane bend vibration are currently investigated. In proteins and peptides the situation is somewhat simpler than in NMA as methyl groups are not directly attached to the peptide plane but only occur in sidechains. Yet, the recent suggestion to use the C^{α}–H^{α} vibration as a secondary structure sensitive probe [63] shows that a detailed study of CH vibrations in oligo-peptides is necessary. It will thus be interesting to see how the coupling networks in small peptides which possess only a single H^{α} instead of methyl groups like NMA change in comparison to our present analysis.

Acknowledgment

Financial support from the Emmy Noether program of the DFG is gratefully acknowledged.

References

[1] S. Woutersen, P. Hamm, J. Phys.: Condens. Matter 14 (2002) R1035.
[2] I.V. Rubtsov, J.P. Wang, R.M. Hochstrasser, Proc. Natl. Acad. Sci. USA 100 (2003) 5601.
[3] S. Woutersen, R. Pfister, P. Hamm, Y.G. Mu, D.S. Kosov, G. Stock, J. Chem. Phys. 117 (2002) 6833.
[4] J. Bredenbeck, J. Helbing, A. Sieg, T. Schrader, W. Zinth, C. Renner, R. Behrendt, L. Moroder, J. Wachtveitl, P. Hamm, Proc. Natl. Acad. Sci. USA 100 (2003) 6452.
[5] J. Bredenbeck, J. Helbing, P. Hamm, J. Am. Chem. Soc. 126 (2004) 990.
[6] M. Volk, Eur. J. Org. Chem. 14 (2001) 2605.
[7] S. Krimm, J. Bandekar, Adv. Protein Chem. 38 (1986) 181.
[8] C. Scheurer, A. Piryatinski, S. Mukamel, J. Am. Chem. Soc. 123 (2001) 3114.
[9] C. Scheurer, S. Mukamel, J. Chem. Phys. 115 (2001) 4989.
[10] C. Scheurer, S. Mukamel, J. Chem. Phys. 116 (2002) 6803.
[11] C. Scheurer, S. Mukamel, Bull. Chem. Soc. Jpn. 75 (2002) 989.
[12] S. Woutersen, P. Hamm, J. Phys. Chem. B 104 (2000) 11316.
[13] P. Hamm, S. Woutersen, Bull. Chem. Soc. Jpn. 75 (2002) 985.
[14] H. Torii, M. Tasumi, J. Raman Spectrosc. 29 (1998) 81.
[15] A.M. Moran, S.-M.M. Park, J. Dreyer, S. Mukamel, J. Chem. Phys. 118 (2003) 3651.
[16] A.M. Moran, S.-M.M. Park, S. Mukamel, J. Chem. Phys. 118 (2003) 9971.
[17] S.Y. Cha, S.H. Ham, M.H. Cho, J. Chem. Phys. 117 (2002) 740.
[18] M.H. Cho, J. Chem. Phys. 118 (2003) 3480.
[19] J.H. Choi, S.Y. Ham, M. Cho, J. Phys. Chem. B 107 (2003) 9132.
[20] S. Ham, J.H. Kim, H. Lee, M.H. Cho, J. Chem. Phys. 118 (2003) 3491.
[21] D. Abramavicius, S. Mukamel, J. Chem. Phys. 120 (2004) 8373.
[22] S. Gnanakaran, R.M. Hochstrasser, J. Am. Chem. Soc. 123 (2001) 12886.
[23] S. Ham, S. Cha, J.-H. Choi, M. Cho, J. Chem. Phys. 119 (2003) 1451.
[24] W. Barber-Armstrong, T. Donaldson, H. Wijesooriya, R.A.G.D. Silva, S.M. Decatur, J. Am. Chem. Soc. 126 (2004) 2339.
[25] R. Huang, J. Kubelka, W. Barber-Armstrong, R.A.G.D. Silva, S.M. Decatur, T.A. Keiderling, J. Am. Chem. Soc. 126 (2004) 2346.
[26] P. Hamm, S. Wouterson, Bull. Chem. Soc. Jpn. 75 (2002) 985.
[27] A. Moran, S. Mukamel, Proc. Natl. Acad. Sci. USA 101 (2004) 506.
[28] J.M. Bowman, J. Chem. Phys. 68 (1978) 608.
[29] G.D. Carney, L.I. Sprandel, C.W. Kern, Adv. Chem. Phys. 37 (1978) 305.
[30] R.B. Gerber, M.A. Ratner, Chem. Phys. Lett. 68 (1979) 195.
[31] S. Mukamel, Principles of Nonlinear Optical Spectroscopy, Oxford University Press, New York, 1995.
[32] S.K. Gregurick, G.M. Chaban, R.B. Gerber, J. Phys. Chem. A 106 (2002) 8696.
[33] R.B. Gerber, B. Brauer, S.K. Gregurick, G.M. Chaban, Phys. Chem. Commun. 5 (2002) 142.
[34] J.O. Jung, B.R. Gerber, J. Chem. Phys. 105 (1996) 10332.
[35] M.H. Beck, A. Jäckle, G.A. Worth, H.-D. Meyer, Phys. Rep. 324 (2000) 1.
[36] B. Brauer, G.M. Chaban, R.B. Gerber, Phys. Chem. Chem. Phys. 6 (2004) 2543.
[37] R.B. Gerber, J.O. Jung, in: P. Jensen, P.R. Bunker (Eds.), Computational Molecular Spectroscopy, John Wiley and Sons Ltd., 2000, p. 365.
[38] J.C. Light, I.P. Hamilton, J.V. Lill, J. Chem. Phys. 82 (1985) 1400.
[39] W. Yang, A.C. Peet, Chem. Phys. Lett. 153 (1988) 98.
[40] C.C. Marston, C.G. Balint-Kurti, J. Chem. Phys. 91 (1989) 3571.
[41] G.M. Chaban, J.O. Jung, R.B. Gerber, J. Chem. Phys. 111 (1999) 1823.
[42] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.J. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, et al., J. Comput. Chem. 14 (1993) 1347.
[43] S. Carter, J.M. Bowman, L.B. Harding, Spectrochim. Acta A 53 (1997) 1179.
[44] G. Rauhut, J. Chem. Phys. 121 (2004) 9313.
[45] M. Bounouar, C. Scheurer, 2005, in preparation.
[46] T.H. Dunning, P.J. Hay, Methods of Electronic Structure Theory, Plenum, NY, 1977.
[47] T.H. Dunning, J. Chem. Phys. 90 (1989) 1007.
[48] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, et al., Gaussian 03, Revision B. 04, Gaussian, Inc., Wallingford, CT, 2004.
[49] S. Ataka, H. Takeuchi, J. Mol. Struct. 113 (1984) 147.
[50] M.F. DeCamp, L. DeFlores, J.M. McCracken, A. Tokmakoff, K. Kwac, M. Cho, J. Phys. Chem. B 109 (2005) 11016.
[51] M.T. Zanni, M.C. Asplund, R.M. Hochstrasser, J. Chem. Phys. 114 (2001) 4579.
[52] W.A. Herrebout, K. Clou, H.O. Desseyn, J. Phys. Chem. A 105 (2001) 4865.
[53] P. Hamm, M. Lim, R.M. Hochstrasser, J. Phys. Chem. B 102 (1998) 6123.
[54] T. Köddermann, R. Ludwig, Phys. Chem. Chem. Phys. 6 (2004) 1867.
[55] K. Kwac, M.H. Cho, J. Chem. Phys. 119 (2003) 2247.
[56] K. Kwac, M.H. Cho, J. Chem. Phys. 119 (2003) 2256.
[57] J.H. Kim, M.H. Cho, Bull. Korean Chem. Soc. 24 (2003) 1061.
[58] T. Clark, Theochem-J. Mol. Struct. 530 (2000) 1.
[59] P. Winget, C. Selcuki, A.H.C. Horn, B. Martin, T. Clark, Theor. Chem. Acc. 110 (2003) 254.
[60] D.M. Benoit, J. Chem. Phys. 120 (2004) 562.
[61] I.V. Rubtsov, J. Wang, R.M. Hochstrasser, J. Phys. Chem. A 107 (2003) 3384.
[62] I.V. Rubtsov, J. Wang, R.M. Hochstrasser, Proc. Natl. Acad. Sci. USA 100 (2003) 5601.
[63] N.G. Mirkin, S. Krimm, J. Phys. Chem. A 108 (2004) 10923.