MoViPac: Vibrational Spectroscopy with a Robust Meta-Program for Massively Parallel Standard and Inverse Calculations

Thomas Weymuth,[a] Moritz P. Haag,[a] Karin Kiewisch,[a]† Sandra Luber,[a]‡ Stephan Schenk,[a]§ Christoph R. Jacob,*[b] Carmen Herrmann,*[c] Johannes Neugebauer,*[d] and Markus Reiher*[a]

We present the software package MoViPac for calculations of vibrational spectra, namely infrared, Raman, and Raman Optical Activity (ROA) spectra, in a massively parallelized fashion. MoViPac unites the latest versions of the programs Snf and Akira alongside with a range of helpful add-ons to analyze and interpret the data obtained in the calculations. With its efficient parallelization and meta-program design, MoViPac focuses in particular on the calculation of vibrational spectra of very large molecules containing on the order of a hundred atoms. For this purpose, it also offers different subsystem approaches such as Mode- and Intensity-Tracking to selectively calculate specific features of the full spectrum. Furthermore, an approximation to the entire spectrum can be obtained using the Cartesian Tensor Transfer Method. We illustrate these capabilities using the example of a large π-helix consisting of 20 (*S*)-alanine residues. In particular, we investigate the ROA spectrum of this structure and compare it to the spectra of α- and 3_{10}-helical analogs. © 2012 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23036

Introduction

Theoretical vibrational spectroscopy on large molecules usually relies on the harmonic approximation to the potential energy surface,[1] that is, one approximates in the nuclear Schrödinger equation (within the Born–Oppenheimer approximation) the total electronic energy as a quadratic potential. The equation to solve is then (in Hartree atomic units)

$$\left(-\frac{1}{2}\nabla^{(m)\dagger}\nabla^{(m)}+\frac{1}{2}\boldsymbol{R}^{(m)\dagger}\boldsymbol{H}^{(m)}\boldsymbol{R}^{(m)}\right)\chi_{\mathrm{tot}}=E_{\mathrm{tot}}\chi_{\mathrm{tot}}, \qquad (1)$$

where χ_{tot} is the total nuclear wave function, E_{tot} its associated energy, the vector $\nabla^{(m)}$, contains the mass-weighted first derivatives with respect to the nuclear coordinates, $\boldsymbol{R}^{(m)}$ are the mass-weighted Cartesian nuclear coordinates, and $\boldsymbol{H}^{(m)}$ is the mass-weighted Hessian matrix, the elements of which are defined as

$$\boldsymbol{H}_{ij}^{(m)}=\left(\frac{\partial^2 E_{\mathrm{el}}(\boldsymbol{R}^{(m)})}{\partial R_i^{(m)}R_j^{(m)}}\right)_{\mathrm{eq}}, \qquad (2)$$

in which E_{el} is the total electronic energy and for which we explicitly state the parametric dependence of the latter on the nuclear coordinates. The subscript "eq" indicates that the second derivative has to be taken at an equilibirum geometry, that is at a stationary point on the potential energy surface.

More accurate methods for the variational calculation of molecular vibrations including anharmonic effects[2–7] are under constant development but can hardly be applied to truly large molecules containing on the order of 100 atoms and more. Hence, molecular dynamics approaches are often used as a means to include anharmonic effects on the motion of atomic nuclei[8,9] but suffer from the neglect of quantum effects on the nuclear motion. Moreover, they necessitate force-field approximations or

[a] *T. Weymuth, M. P. Haag, K. Kiewisch, S. Luber, S. Schenk, M. Reiher*
Laboratorium für Physikalische Chemie, ETH Zurich, Wolfgang-Pauli-Str. 10, 8093 Zurich, Switzerland
E-mail: markus.reiher@phys.chem.ethz.ch

[b] *C. R. Jacob*
Center for Functional Nanostructures, Karlsruhe Institute of Technology (KIT), Wolfgang-Gaede-Str. 1a, 76131 Karlsruhe, Germany
E-mail: christoph.jacob@kit.edu

[c] *C. Herrmann*
Institut für Anorganische und Angewandte Chemie, University of Hamburg, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany
E-mail: Carmen.Herrmann@chemie.uni-hamburg.de

[d] *J. Neugebauer*
Institute for Physical and Theoretical Chemistry, Technical University Braunschweig, Hans-Sommer-Str. 10, 38106 Braunschweig, Germany
E-mail: j.neugebauer@tu-braunschweig.de

†*Present address:* Amsterdam Center for Multiscale Modeling, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands.

‡*Present address:* Department of Chemistry, Yale University, P.O. Box 208107, New Haven, Connecticut 06520-8107.

§*Present address:* BASF SE, GMC/M – B001, 67056 Ludwigshafen, Germany.

Contract grant sponsor: Swiss National Science Foundation Snf; contract grant number: 200020-132542/1
Contract grant sponsor: Netherlands Organisation for Scientific Research (NWO); contract grant number: 700.59.422
Contract grant sponsor: DFG-Center for Functional Nanostructures
Contract grant sponsor: Hamburgische Forschungs- und Wissenschaftstiftung (Cluster of Excellence Nanospintronics).

require tremendous resources for the generation of trajectories if the potential energy surface is calculated with first-principles methods.[10]

As a consequence, the harmonic approximation remains the standard approach in quantum chemistry.[11] It has the important advantage that intensity expressions for numerous spectroscopic techniques can be easily evaluated within the double-harmonic approximation.[12] Anharmonic corrections are usually considered perturbatively by calculating cubic and quartic force constants for elongations along the normal modes obtained within the harmonic approximation.[13,14] This method is then often referred to as VPT2; note that the first calculations of second order perturbations to harmonic frequencies were done as early as 1933.[15] Often one observes a very good agreement of calculated harmonic frequencies and measured fundamental ones (see, e.g., Refs. [16–19]), which is due to a fortunate error compensation of some density-functional approximations and sufficiently large basis sets.[14] If this is not the case, scaling factors may be used; note that the so-called scaled quantum mechanical force fields are one particular 'flavor' of these techniques.[14,20–25]

In 2002, we presented a seminumerical, massively parallel implementation for quantum chemical calculations of molecular vibrations in the harmonic approximation which has become known as the SNF program.[26] SNF was based on earlier work in the 1990s by Grimme and Marian[27] and its structure has been under constant development. Moreover, offsprings have been developed with special capabilities. One is the AKIRA program which implements the Mode-Tracking algorithm[28] for the efficient and targeted calculation of selected molecular vibrations in large molecules and molecular aggregates.[29] Also, AKIRA has been continuously developed and supplemented by new algorithms like Intensity-Tracking.[30–33] Other offsprings are the anharmonicity program ANF[14] and the resonance Raman program KTRANS.[34]

Here, we present the latest developments of SNF and AKIRA which are now united into one single program package for vibrational spectroscopy called MOVIPAC. Some of these latest developments are applied to a π-helix containing 20 (*S*)-alanine residues as an example. The article is organized as follows: In 'The MOVIPAC Philosophy' section, we explain the basic program philosophy which guides the whole program design. Then, after briefly commenting on some technical aspects of MOVIPAC in section 'Technical Aspects of MOVIPAC', we discuss in more detail the Mode- and Intensity-Tracking algorithms and the Cartesian Tensor Transfer Method (CTTM) as well as the concept of localizing normal modes in sections, 'Mode-Tracking' to 'Localizing Normal Modes', respectively. The example of a large π-helix serves to illustrate some of the capabilities of MOVIPAC in section 'Example: ROA Spectra of π-Helices' (the computational methodologies adopted in the individual calculations presented in this work are explained in the respective sections). Finally, conclusions are drawn in section 'Conclusions'.

The MOVIPAC Philosophy

MOVIPAC sacrifices the possibility to analytically evaluate the quantum chemical force field (i.e., second partial derivatives of the electronic energy with respect to nuclear coordinates) and the spectroscopic intensities (i.e., first derivatives of property tensors with respect to nuclear coordinates), but, instead uses a (semi-)numerical differentiation scheme (i.e., the Hessian matrix is evaluated as the numeric first derivative of analytic energy gradients, while the spectroscopic intensities are obtained as numeric first derivatives of the respective property tensors). This has two disadvantages. First, it is less efficient by a constant, but small factor than the fully analytical evaluation and second, it introduces a numerical error. The latter disadvantage, namely the numerical error, turns out to be negligible in view of the harmonic approximation used (about 1 cm^{-1}) and can be easily reduced by taking advantage of a Bickley finite difference formula with more grid points.[35] The former disadvantage is not severe and turns into more than one advantage: The fact that one needs only analytic geometry gradients and molecular property tensors as raw data, which are provided by almost any multipurpose quantum chemistry program, allows us to calculate spectra in a parallel fashion for many spectroscopic techniques using basically any electronic structure method. We thus circumvent the implementation of analytic property gradients and second geometry derivatives for nonstandard quantum chemical methods and their parallelization at a negligible price. The reduced efficiency resulting in longer calculation times is made up for by the huge number of computing cores available to research groups nowadays. To shed more light onto this issue, we shall provide a reliable comparison of analytical *versus* numerical derivatives in the next section.

Performance of Numerical and Analytical Derivatives

The Raman spectrum of the C_{60} molecule has been calculated with SNF (using the so-called message passing interface, MPI, as parallelization scheme) in conjunction with a new version of TURBOMOLE,[36] that is, TURBOMOLE 6.3.1 and with TURBOMOLE 6.3.1 alone. For the SNF calculations, a serial version of TURBOMOLE was used, while for the TURBOMOLE-only calculations, both an MPI- and a shared-memory parallelized (SMP) version were considered. In all these calculations, density functional theory was applied using the BP86 exchange–correlation functional[37,38] and Ahlrichs' def2-TZVP basis set[39] at all atoms. Advantage was taken of the resolution-of-the-identity technique with the corresponding auxiliary basis set.[40] The high symmetry of Buckminsterfullerene (point group I_h) can be conveniently exploited with both programs, SNF as well as TURBOMOLE. All calculations have been performed in the same computing environment, namely a blade system featuring two dodeca-core AMD Opteron 6174 processors (i.e., a total of 24 cores) and 48 GB of memory. Therefore, these calculations are truly comparable to each other in terms of timings, as the very same properties have been calculated with the same methods (note also that the keywords in the TURBOMOLE control file were the same in all calculations, except the one defining the symmetry—the individual steps during the SNF calculation are performed in C_1

Table 1. Comparison of the total wall-time needed to calculate the Raman spectrum of C_{60} with SNF in conjunction with TURBOMOLE 6.3.1 and TURBOMOLE 6.3.1 alone.

Program (parallelization)	7 cores		12 cores		18 cores	
	IR	Raman	IR	Raman	IR	Raman
SNF (MPI)	2:54	6:26	n/a	n/a	n/a	n/a
TURBOMOLE (MPI)	6:33	6:59	6:32	6:58	6:34	7:01
TURBOMOLE (SMP)	2:24	2:40	2:04	2:21	2:03	2:19

All timings are reported in the format hrs:min.

symmetry, since the distorted structure do no longer exhibit point group I_h) on the same computers.

The individual timings obtained are reported in Table 1. We note that SNF already reaches its maximum performance on seven cores; this is because for C_{60} we only need to calculate six individual steps in the icosahedral point group (namely, two distortions in each spatial dimension for one atom, as all 60 atoms are related to each other by symmetry; on the seventh core a master task is running which does no calculations but only distributes the individual steps and collects the results). Note that one could of course link a parallelized version of TURBOMOLE (or any other quantum mechanical backend) with SNF to speed up the calculation of the individual steps. In fact, when resorting to a SMP version of TURBOMOLE (see below), the entire infrared (IR) spectrum of C_{60} can be calculated in only 1 h and 26 min (when each step is running on three cores), compared to the 2 h and 54 min (Table 1) in the case of a serial TURBOMOLE version.

Next, by looking at the results in Table 1, we note that MPI-parallelized TURBOMOLE is significantly slower than SNF. This is due to the fact that in this TURBOMOLE version, the AOFORCE module, required to calculate the Hessian matrix, is not parallelized. As more than 90% of the total calculation time is spent in the module AOFORCE, there is of course also no speedup of the calculation when additional cores are used. The results in Table 1 nicely illustrate this. When using the SMP version of TURBOMOLE (which uses MPI and shared memory versions of the modules RIDFT and RDGRAD, and multithreaded versions of AOFORCE and EGRAD), however, the Hessian matrix can be calculated in a parallel fashion, which significantly speeds up the calculation. In this case, TURBOMOLE is faster than SNF.

In this context, it is interesting to note that the calculation of the IR spectrum takes only roughly 30 min longer on seven cores with SNF as compared to TURBOMOLE, while the Raman spectrum needs almost 4 h more to be calculated. This may be attributed to the fact that the calculation of the Hessian matrix scales in the same fashion for both analytical and numerical implementations (only the prefactor is somewhat larger in the case of a numerical implementation). For the calculation of the Raman intensities, however, an analytical implementation has a better scaling behavior. In fact, in a recent paper by Rappoport and Furche, it was claimed[41] that analytic gradients of the electric-dipole–electric-dipole polarizability (in short often called polarizability) important for Raman intensities are by a factor of 100 more efficient than the numerical first derivatives as implemented in SNF.[26] Unfortunately, this conclusion is misleading as the authors of that paper compared calculations which are not comparable for the following reason. In our original SNF article,[26] we studied Buckminsterfullerene (C_{60}) with the TURBOMOLE 5.1 program. However, that early version of TURBOMOLE was not able to calculate polarizability tensors with the resolution-of-the-identity density-fitting technique, which increases the efficiency by a factor of about ten. By contrast, the authors of Ref. [41] used this technique. Of course, since more recent TURBOMOLE versions always invoke the resolution-of-the-identity technique,[41] also SNF benefits from it. Indeed, the data in Table 1 clearly shows that the analytical polarizability gradients are by no means 100 times more efficient. The difference between the time needed to obtain the Raman spectra and the IR spectra is 16 min in the case of the SMP version of TURBOMOLE on seven cores. This is precisely the time required by the TURBOMOLE module EGRAD, which is responsible for the analytical derivatives of the polarizability. Compared to this, we may estimate the time needed by SNF for the numerical derivatives to be roughly 386 min. Thus, it can be said that the analytic derivatives are about a factor of 24 more efficient than their numerical counterparts. Furthermore, one should not forget that the calculation of these derivatives is only one part required to obtain the entire Raman spectrum. If one takes into consideration also the time needed to compute the Hessian matrix, one understands that the seminumeric SNF calculation as a whole is only a factor of roughly 2.5 slower than the fully analytic calculation using TURBOMOLE alone.

From the timings in Table 1, we can also see that the maximum performance of the SMP version of TURBOMOLE is reached at roughly 12 cores; using six additional cores does lower the required computing time only very little.

Because of the different scaling behaviors of the calculation of the Hessian matrix and the property derivatives, the former often becomes the bottleneck in a vibrational calculation for large molecules. This is precisely the background for techniques such as Mode- and Intensity-Tracking (see below).

Technical Aspects of MOVIPAC

After having outlined the fundamental philosophy behind MOVIPAC, some remarks on the technical progress of the program package are appropriate. The entire source code base of the SNF program has been reviewed. Numerous modifications have been incorporated. Code duplications and dead code paths have been eliminated. The parser of the output generated by the quantum chemical back-end programs is now more robust due to the use of Regular Expressions. Especially the C code parts have seen large changes to adhere to the coding standards of contemporary Linux distributions. The code now relies on the compiler and operating system (OS) to support the C99 and POSIX-1.2001 standards, respectively, that are fulfilled, for example, by any reasonably recent Linux distribution.[42]

The build system has been revamped from scratch and now uses the GNU build system (autotools)[43,44] that is used by the majority of GNU software. The "autoconf" tool from this collection is used to create a "configure" script that automatically

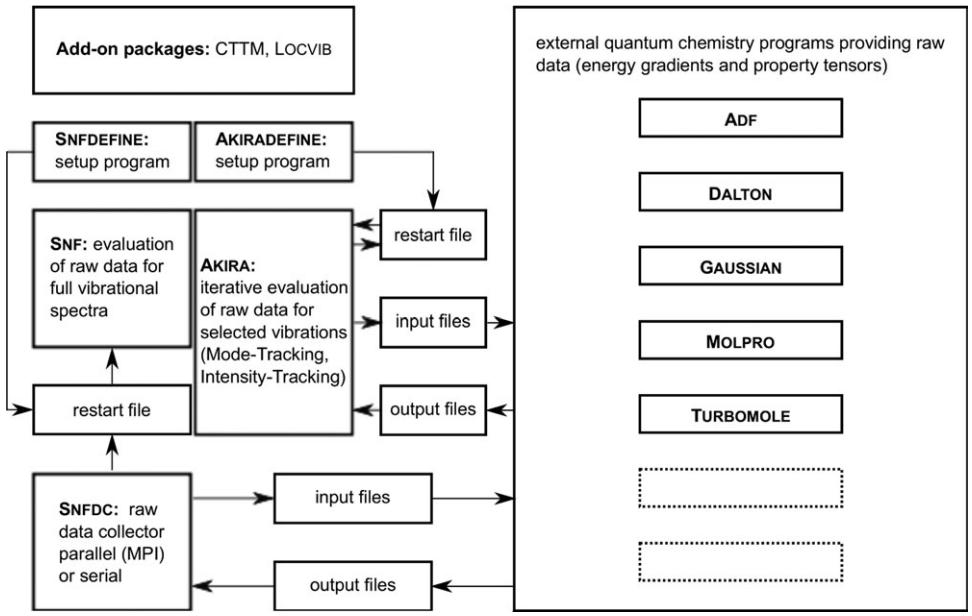

Figure 1. Schematic overview of the meta-structure of MOVIPAC. The dashed boxes on the right hand side illustrate the possibility to straightforwardly implement interfaces to additional quantum chemical programs.

detects features of the OS, for example, 64 *versus* 32 bit, compiler, libraries (including BLAS/LAPACK) or header files. The "automake" tool is used to create the different Makefiles. This combination of tools allows for parallel compilation, out-of-tree builds or installation into nondefault subdirectories. Furthermore, it provides the means for a clean implementation of the Fortran/C interface by automatic detection of the name mangling scheme of the compiler combination used.

By default, a serial version of MOVIPAC is built. A parallel version relying on MPI can be built by passing an appropriate command line option to "configure." The MPI version has received a lot of testing and turned out to work well. It uses a master-slave architecture where one master process is used to distribute tasks to the slave nodes. All MPI implementations that provide compiler/linker wrapper scripts (e.g., "mpif77") are supported.

Although the two main parts of MOVIPAC, SNF and AKIRA, were originally independent programs they are now united in a single meta-program. A schematic overview of MOVIPAC's structure is shown in Figure 1. As can be seen, MOVIPAC features a modular structure, which simplifies additions to it (like, e.g., a new interface to another quantum mechanical program). There are programs to interactively set up the input files necessary for the calculations (SNFDEFINE for common frequency analyses; AKIRADEFINE for Mode- and Intensity-Tracking), and programs to evaluate the resulting output (SNF and AKIRA, respectively). In the case of SNF, there is a dedicated program controlling the actual calculation of all necessary data (SNFDC) by steering the quantum mechanical backend programs via simple file-based communication. For Mode- and Intensity-Tracking calculations, this task is done by the program AKIRA. We would like to put special emphasis on the central restart file, in which the data computed so far is stored. In the case of a computer crash, this data is not lost, and, thus, the calculation can be straightforwardly restarted at the point it crashed.

The MOVIPAC source code is made available on the internet at www.reiher.ethz.ch/software/movipac.

Mode-Tracking

As already mentioned in the "Introduction", MOVIPAC focuses in particular on vibrational calculations for large molecules. Several methods have been developed over the past years to simplify frequency analyses and vibrational spectra calculations for large molecules by introducing additional approximations beyond the harmonic approximation. Examples include partial Hessian diagonalization,[45] mobile block Hessian approaches,[46,47] or Cartesian tensor transfer techniques.[48]

The mode-tracking method[28] implemented in the AKIRA program within MOVIPAC differs from those approaches, since it reduces the computational effort by focusing on one (or a few) vibrations instead of the entire spectrum, without introducing additional approximations other than the harmonic one.

The motivation behind Mode-Tracking is that it is often possible to guess a vibration that is characteristic for a particular spectroscopic phenomenon or a molecular structure. Starting from that guess, the vibration is iteratively refined until a normal mode is found that is similar to the guess vibration (measured by an overlap criterion), but is also an eigenvector of the Hessian (within a given numerical precision).

Mathematically speaking, a Krylov subspace iteration is performed based on Davidson's method.[49,50] It sets out to calculate the Davidson matrix for a user-defined guess, from which a residual vector is obtained after diagonalization that is a measure for the deviation from an exact eigenvector, and thus, provides information on how to improve the guess vector. From this residual and preconditioner, which is an approximation to the inverse of the Hessian (minus the sought-for eigenvalue), a correction vector is generated, which is added as a new basis vector in the iterative subspace diagonalization. The theory has been explained in great detail in the original articles[28,51] and in recent reviews.[29,33,52]

Subspace iteration methods for vibrational problems have been used before in the context of molecular mechanics applications,[53–56] where huge Hessians must be diagonalized and where subspace iteration techniques are a natural means to solve the diagonalization problem. In quantum chemistry, however, the Mode-Tracking idea established new principles[28,57–63]: (1) isolated, structure-characteristic vibrations can be directly targeted, (2) vibrations that involve only a subset of atoms (e.g., in the case of adsorbates on surfaces[64] or in quantum mechanics (QM)/molecular mechanics (MM) partitionings[65]) can be directly

optimized, (3) low-dimensional Hessians also benefit if their entries are time-consuming to calculate.

In our former work,[51] it was shown that the proper construction of a guess vector has an important influence on the convergence characteristics of the algorithm. Often, such guess vectors can be constructed based on intuition (e.g., for localized stretch vibrations[51]) or on the basis of model calculations.[58] In the following, we will demonstrate this feature for the example of the N—H stretch vibrations in the adenine–thymine (AT) base pair shown in Figure 2. This is an important example from the class of supermolecular assemblies, which appear in an increasing number of quantum chemical applications, for example, in explicit solvation studies or in studies on models for protein binding pockets.

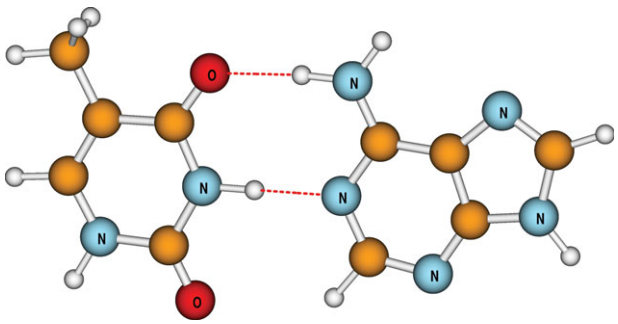

Figure 2. Structure of the AT base pair used as an example for mode-tracking calculations.

Structure optimizations and single-point calculations needed for the mode-tracking calculations as reported in the following have been performed with the program package TURBOMOLE 6.0[36] using the BP86 density functional[37,38] and Ahlrichs' def-TZVP basis set[66] for all atoms.

A straightforward way to construct guess vibrations in supermolecular systems is to start from the vibrations calculated for isolated subsystems. The AT base pair represents a fairly simple example in this context, but the procedure outlined here can easily be extended to several subsystems. In the present case, we first performed frequency analyses for the two building blocks, that is, for adenine and thymine. Both molecules were fully optimized, and the resulting frequencies for the N—H stretch vibrations are listed in the first column of Table 2. For the isolated bases, all N—H stretch vibrations lie between 3500 and 3660 cm^{-1}.

Table 2. Vibrational wavenumbers (in units of cm^{-1}) for the AT base pair.

Isolated	Complex	M.-T. (1)	M.-T. (final)	Assignment
3500.1	2673.6	2668.8	2673.8	T: NH···N
3518.8	3209.9	3318.6	3210.1	A: NH2···O sym
3544.8	3541.4	3541.7	3541.7	T: NH isolated
3552.6	3551.8	3551.9	3552.0	A: NH isolated
3652.1	3588.7	3483.0	3588.6	A: NH2···O asym

The values reported in the column 'complex' refer to the reference, that is, full vibrational calculation; 'M.-T.' stands for Mode-Tracking results in the first (1) and final iteration.

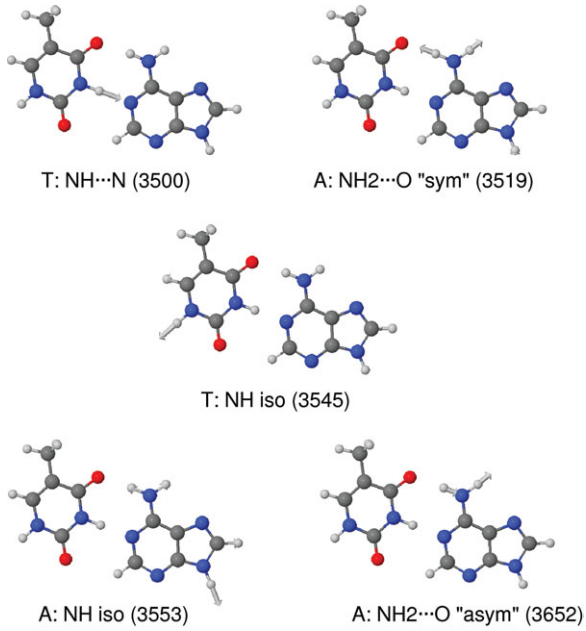

Figure 3. Normal modes obtained for isolated thymine or adenine, respectively, used as guess modes in the Mode-Tracking calculation. The vibrational wavenumber associated with each normal mode is given in brackets (all values in cm^{-1}).

The Hessians of the two constituent molecules were then used as blocks for an approximate Hessian of the base pair, and the normal modes were adopted as guess normal modes in a mode-tracking calculation on the complex. In the present case, each vibration has been tracked separately; results are given in Table 2 for the first and final iteration of the mode-tracking calculation (we report the data with an accuracy of 0.1 cm^{-1} to illustrate the small numerical differences between

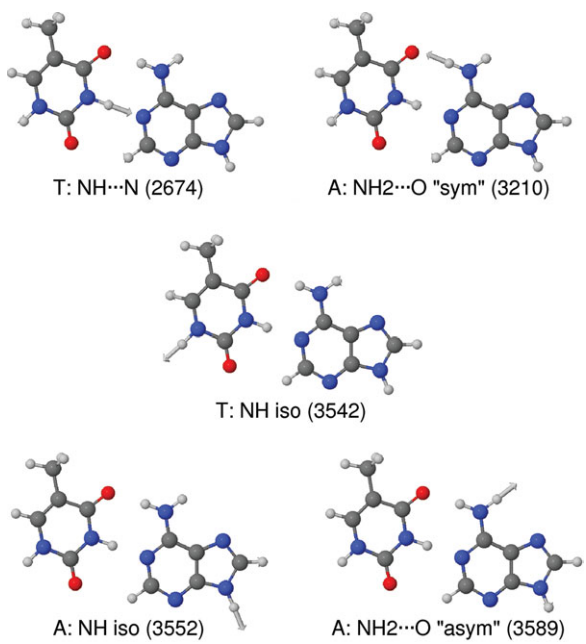

Figure 4. Normal modes for the AT dimer. The vibrational wavenumber associated with each normal mode is given in brackets (all values in cm^{-1}). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Mode-Tracking and full vibrational calculations, even though typical discrepancies between measured fundamental frequencies and calculated harmonic ones are on the order of 10–20 cm^{-1}). The results of the first iteration just reflect the change in wavenumber when assuming that a given monomer normal mode is unchanged upon complex formation. The guessed vibrations are shown in Figure 3, and the normal modes of the complex are displayed in Figure 4.

From Table 2, we can distinguish three different types of vibrations. The first type comprises vibrations which are hardly affected upon base pairing, neither normal mode nor the vibrational frequency are affected. Hence, already the results for the isolated molecules are very close to the results obtained for the complex. The two "isolated" N—H vibrations, that is, the N—H stretch vibrations not involved in hydrogen bonding, belong to this class. They are found at 3545 and 3553 cm^{-1} in the isolated case, and at 3542 and 3552 cm^{-1} in the base pair. The changes in wavenumber are thus of the order of 1–3 cm^{-1}, and already the first mode-tracking iteration essentially leads to the converged result.

The second type involves modes which significantly change in wavenumber, but the normal modes are hardly changed upon base pairing. In our example, the thymine NH···N mode belongs to this class. It decreases by more than 820 cm^{-1} in wavenumber, from 3500 to 2674 cm^{-1}. Interestingly, the largest part of this change is already captured in the first mode-tracking iteration, which is an indication that the exact normal mode is already well approximated by the guess vibration taken from the set of modes of the isolated molecule. This guess deviates by less than 5 cm^{-1} from the converged wavenumber, and thus, overestimates the wavenumber change by only 0.6%. By comparing Figures 3 and 4, it can be seen that these normal modes are essentially unchanged.

This situation changes for the third type of vibrations, for which both the wavenumber and the normal modes change significantly. In the present example, this involves the two N—H stretch modes of the NH_2 group, which can be classified as "symmetric" and "antisymmetric" in the free adenine molecule. But, this symmetry is completely broken by the hydrogen-bonding interaction of one of the two hydrogen atoms with the thymine oxygen atom, as can be seen in Figure 4. Also here, already the results obtained with the guess mode in the first iteration of the mode-tracking calculation indicate significant changes in the wavenumbers, but they still deviate from the final results by about $\pm$100 cm^{-1}. Most of this deviation is corrected in the second iteration, in which an approximation for the other NH_2···O mode is obtained as a by-product. We note in passing that all converged wavenumbers from the mode-tracking calculation agree within 0.4 cm^{-1} with the results from a conventional frequency analysis.

The present example indicates that the construction of guess vibrations based on isolated molecule calculations is a powerful tool when calculating vibrational spectra for molecules in explicit environments. The weaker the interaction with the environment is, the faster the convergence will be. In the present example, all vibrations were converged within four iterations or less, based on a preconditioner constructed from the subsystem Hessians. A more challenging study in a similar spirit has recently addressed the effects of a protein binding pocket on the resonance Raman spectrum of the carotenoid spheroidene in the photosynthetic reaction center of the purple bacterium *Rhodobacter sphaeroides*.[67] Using a similar approach as used here, guess vibrations were constructed for the isolated spheroidene (102 atoms), while subsequently the vibrations have been refined in the binding pocket (507 atoms), thus, demonstrating the power of Mode-Tracking for challenging vibrational problems.

Intensity-Tracking

The Mode-Tracking approach is the method of choice if the calculation of one or several specific normal modes is desirable. It is especially attractive if it is already known what part of the molecule under study produces a characteristic vibrational pattern so that this information can be used for the set up of the starting guess of the iterations. However, often the characteristic pattern of a vibrational spectrum—usually the intense bands in a certain wavenumber range—is of interest, not only one or a few vibrational modes. Therefore, an algorithm which selectively calculates only the intense bands in a spectrum is desirable. For this task, Intensity-Tracking[30–33] has been developed and implemented into the AKIRA program package. This algorithm achieves to produce a spectrum containing the important features and simultaneously saving computational time by avoiding the calculation of normal modes with small intensities.

The Intensity-Tracking algorithm is, like the Mode-Tracking approach, based on the iterative solution of the Hessian eigenvalue problem. The main criterion in an Intensity-Tracking calculation for the selection of molecular collective distortions, which shall be converged to normal modes, is the intensity assigned to these distortions. Thus, the best starting point is a hypothetical mode (i.e., a collective distortion of the positions of atomic nuclei from the equilibrium structure) which carries high intensity. In each iteration, the intensities of the approximate normal modes expanded in a basis of collective molecular distortions are calculated and all normal modes which satisfy the selection criterion are considered for optimization. For each of these selected approximate normal modes, an additional basis vector is added to improve its description as long as its residuum vector is larger than a given convergence criterion. If the residuum vector fulfills the convergence criterion, the normal mode is considered as converged and not refined further. The iterations terminate if all normal modes which fulfill the selection criterion are converged.

Three different intensity-based selection procedures were tested for Intensity-Tracking calculations and are implemented in MOVIPAC[30,31]:

1. selection of a specified number of (approximate) normal modes from the (approximate) normal modes with highest intensity,

2. selection of many (approximate) normal modes so that the sum of their intensities exceeds a given threshold, and

3. selection of all (approximate) normal modes whose relative intensity is higher than a certain threshold.

In principle, Intensity-Tracking can be used for any kind of theoretical vibrational spectroscopy. Ideally, a starting nuclear distortion with high intensity is available as a starting guess. For Resonance Raman spectroscopy, a high-intensity starting guess has been derived[30] in the framework of Heller's short-time approximation[68,69] and successfully applied in Intensity-Tracking iterations.[30] In this case, the direction of the gradient of the electronically excited state leading to the resonance enhancement is taken as the starting nuclear distortion.

Resonance Raman spectroscopy is often applied to probe the environment of aromatic amino acids in proteins. Therefore, we illustrate a Resonance Raman intensity-tracking calculation for the example of tyrosine (see Fig. 5). The optimized structure of tyrosine as well as the gradient of the 5^1A state (5.73 eV), which was chosen as an example, were obtained using the hybrid functional B3LYP and the def-TZVP basis sets as implemented in TURBOMOLE 5.10.

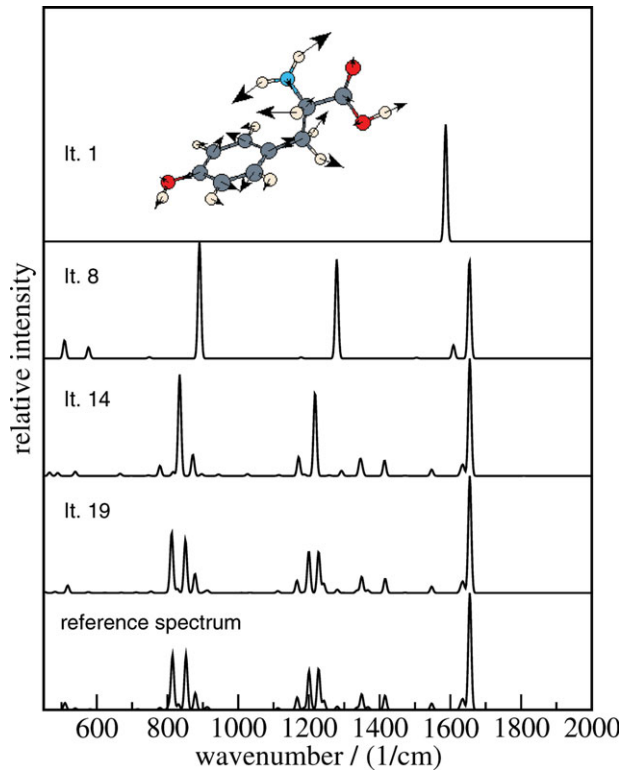

Figure 5. Intensity tracking for the B3LYP/TZVP Resonance Raman spectrum of the 5^1A state of tyrosine. In the first three panels, intermediate approximate spectra are shown for the starting vector, for iteration 8 (22 basis vectors) and iteration 14 (49 basis vectors). The first panel also contains a representation of the starting vector. In the last two panels, the converged intensity-tracking spectrum of iteration 19 (64 basis vectors) and the conventionally calculated reference spectrum are depicted, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The starting vector, in this case the electronically excited-state gradient, constitutes the first approximate spectrum, and thus carries the complete intensity and information on the collective motion of the intense modes. When adding basis vectors in further iterations, modes split up, and the intensity is distributed over the spectrum. Among these new approximate modes, the intense ones are selected for further optimization. In this example calculation, only the five modes with the highest intensity in the wavenumber range between 500 and 2000 cm^{-1} are chosen, since this is the region we are interested in in this Resonance Raman spectrum. The most intense peak appears at its correct position already after a few iterations. Moreover, the approximate spectrum represents a fairly good distribution of the intensity over the spectrum, as can be seen in iteration 8 in this example (Fig. 5). Although new basis vectors are only generated for a small number of intense modes, the description of all modes is improved with the increasing number of basis vectors (cf. iteration 14 in Fig. 5), so that the intensity-tracking calculation finally yields also the finer details of the spectrum, resulting in the overall characteristics of the spectrum to be reproduced. Between the converged intensity-tracking spectrum (iteration 19) and the reference spectrum, which was generated by projecting the excited-state gradient onto normal modes obtained from a SNF calculation, there are only very slight differences in the wavenumbers and intensities of low-intensity modes. It is also interesting to note that the approximated spectrum agrees qualitatively quite well with the reference spectrum (in particular in the spectral region above 1200 cm^{-1}) already in iteration 14 where significantly fewer basis vectors are used. Especially for large molecules, Intensity-Tracking can save a great amount of computational time (see "Example: ROA Spectra of π-Helices" section for an example).

For other vibrational spectroscopies, it is possible to write the intensity expression directly in terms of normal modes. Then, a hypothetical mode with maximum intensity, a so-called intensity-carrying mode (ICM), can be derived by maximizing the intensity with respect to the distortions of a starting mode. In case of IR spectroscopy, this ansatz leads to a matrix eigenvalue equation where the matrix is determined by the Cartesian gradients of the electric-dipole moment components leading to three hypothetical modes with high intensity. These ICMs have been found by Torii et al.[70] and in an alternative derivation by us.[31] In a similar way, the ICMs can be derived for Raman intensities where the Cartesian gradients of the electric-dipole–electric-dipole polarizability tensor are required[71,32] for the eigenvalue equation. This concept has also been extended to Raman Optical Activity (ROA) spectroscopy, where, in addition to the electric-dipole–electric-dipole polarizability tensor, the electric-dipole–magnetic-dipole and electric-dipole–electric-quadrupole polarizability tensors are necessary.[32] Intensity-Tracking using these ICMs has been implemented for IR, Raman, and ROA spectra.[31,32]

IR, Raman, and ROA spectra contain, in contrast with Resonance Raman spectra, usually many intense bands which may lead to a slow convergence of Intensity-Tracking calculations. To allow for a faster convergence, it is useful to combine intensity-based selection criteria with mode-based ones.[31,32] For example, selection criteria leading solely to optimization of high-intensity (approximate) normal modes in a certain wavenumber range have been implemented and applied to IR, Raman, and ROA Intensity-Tracking.[31,32]

CTTM

The CTTM, proposed by Bouř et al.,[48] is an approximate method to construct spectra of very large molecules from smaller fragments. In contrast to the Mode- or Intensity-Tracking approaches, not only a part of the full vibrational spectrum is calculated but an approximation to the entire spectrum is obtained. Although the method yields good results for vibrational frequencies and reasonable spectra in case of IR and Raman spectroscopy, the calculation of ROA spectra by these methods has limitations.[72,73] Nevertheless, it has found wide application and we provide our implementation[72] in the MoViPac package. The CTTM approximates a spectrum by composing the property tensor derivatives from the corresponding quantities obtained from full calculations on smaller fragments. It can therefore be used in cases where the full calculation is not feasible, but one has to keep in mind that the approximation introduces errors depending on the details of the fragmentation which can be non-negligibly large.

The CTTM in the MoViPac package requires subsystem calculations of the Hessian matrix and the property tensor derivatives to be combined to yield these quantities of the full molecule. They are then processed by the Snf routines of the MoViPac package as in every other calculation. Therefore, the key steps in a spectrum construction using the CTTM are

1. the calculation of the Hessian and property derivatives for the fragments with Snf,

2. the composition of the results to obtain the Hessian and property tensor derivatives of the whole molecule, and

3. the calculation of the spectrum with Snf from these quantities.

As the choice of the fragments and the way they are combined heavily influences the result of the calculation,[72] there is no automatic procedure implemented in the MoViPac package. Instead we added a new feature to the Snf program that allows to read in Hessian matrices and property tensor derivatives from external files. The extraction and composition of the raw data from the subsystem calculations is taken care of by an add-on Python script.

To illustrate the CTTM method, we may consider an oligomer of several identical monomers such as, for instance, the π-helix discussed in "Example: ROA Spectra of π-Helices". In this case, a smaller fragment composed of only a few monomers is used for the calculation of the Hessian matrix and property tensor derivatives, and these results are then transferred to the large molecule. The small fragment is moved along the original molecule, and for each position a mapping between atoms of the small fragment and atoms of the large molecule is defined. For each pair of atoms (including also neighboring atoms) defined by this mapping, a rotation and translation is determined to align the atoms of the fragment to those of the large system. These operations are then used to rotate and translate the Hessian matrix elements and property tensor derivatives corresponding to the atom pair. Pair by pair the complete Hessian matrix and property tensor derivatives of the large molecule are constructed. Hence, a band diagonal structure arises for the Hessian matrix of the large molecule. A schematic example is illustrated in Figure 6.

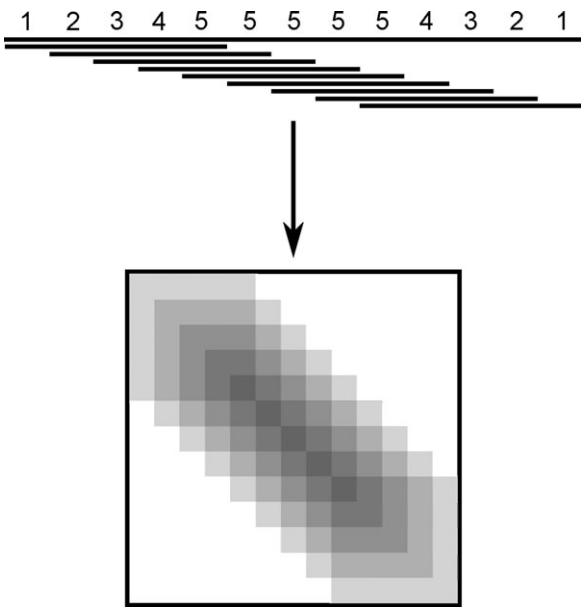

Figure 6. An abstract illustration of the Cartesian Tensor Tranfer Method: A large polymer consisting of 13 monomers is reconstructed by a fragment (here a pentamer). The numbers above the bar show how many overlaps exist for one unit of the large polymer. The overlap with the smallest RMS value is chosen always. An approximation of the Hessian matrix of the large molecule is then constructed with the ones of the fragments. The resulting block diagonal Hessian matrix is shown schematically in the bottom part of the figure.

Localizing Normal Modes

Although MoViPac allows for the efficient calculation of the vibrational spectra of very large molecules, this creates new problems to be solved. For very large molecules, many normal modes contribute to the vibrational spectra. In many cases, individual vibrational transitions are very close in frequency and cannot be resolved in experiment where only few bands are observed. In general, many normal modes contribute to one of these bands and the individual normal modes are often very delocalized. In such cases, the interpretation of the calculated results is very difficult.

For instance, when comparing the calculated vibrational spectra of polypeptides in different conformations to understand the relationship between secondary structure and vibrational spectra, one aims to understand how structural changes affect the positions, intensities, and shapes of the observed bands. However, it becomes impossible to compare all the individual normal modes contributing to each band, in particular because of their delocalization over the whole polypeptide and because the normal modes change significantly when the structure changes.

As a tool for the analysis of calculated vibrational spectra of large molecules, we recently developed a method for obtaining localized modes. This makes it possible to reconstruct a simpler and more intuitive subsystem picture from a full vibrational calculation.[74] To perform such an analysis, one determines a unitary transformation $\boldsymbol{U}$ for the subset of normal modes $\boldsymbol{Q}^{\mathrm{sub}}$ contributing to one band,

$$\tilde{Q}^{\mathrm{sub}}_{i\alpha,p} = \sum_q U_{qp} Q^{\mathrm{sub}}_{i\alpha,q}, \tag{3}$$

where $Q^{\mathrm{sub}}_{i\alpha,q}$ denotes the $\alpha \in \{x, y, z\}$ component of nucleus i in normal mode q, such that the transformed modes $\tilde{\boldsymbol{Q}}^{\mathrm{sub}}$ are as localized as possible. This can be achieved by defining a suitable localization criterion and determining the unitary transformation $\boldsymbol{U}$ through consecutive Jacobi rotations.[74] The resulting localized modes are compact and easy to interpret, since only a few atoms contribute. Furthermore, for large molecules built from similar repeating units such as polypeptides[75] or synthetic polymers,[76] the localized modes obtained for one band are very similar to each other. In a polypeptide, they represent a set of the same type of vibration, but localized on the different amino acid residues.

If the localized modes are very similar, it is sufficient to consider only one representative localized mode. It is then possible to understand the changes in the positions and total intensities of individual bands induced by a structural change by comparing the frequencies and intensities of representative localized modes.[75] In addition, it is possible to extract coupling constants between localized modes, which can be used to explain the band shapes.[74,75] For example, this allows one to directly relate the sign of the amide I couplet in ROA spectra to the secondary structure.[77] Furthermore, the coupling constants often show a simple dependence on the molecular structure that can be parameterized to obtain simpler local-mode models for predicting vibrational spectra of polypeptides and polymers.[78]

The localization of normal modes is implemented in the LocVib add-on package to Snf, which is included in the MoViPac release. This add-on reads in the results of a previous Snf calculation, assists the user with the assignment of the normal modes to vibrational bands and performs the localization. It provides frequencies and intensities of the localized modes as well as the coupling constants and in addition also includes some functionality to perform further analysis.

Example: ROA Spectra of π-Helices

To demonstrate some of the capabilities MoViPac offers for the calculation of large molecules, we present vibrational spectra of a π-helix which consists of 20 (*S*)-alanine residues (compare Fig. 7; note that the structure of this helix is somewhat distorted, and thus, care must be taken in the analysis of the vibrational spectra). In general, a π-helix is characterized by a hydrogen bond of the N—H group of amino acid j to the carbonyl group of the amino acid $j - 5$ five residues earlier in the chain. This is different from α- and 3_{10}-helices, where this hydrogen bond occurs between amino acid residues j and $j - 4$ and j and $j - 3$, respectively, and results in a larger N—C^{α}—C—N torsional angle compared to α- and 3_{10}-helices.

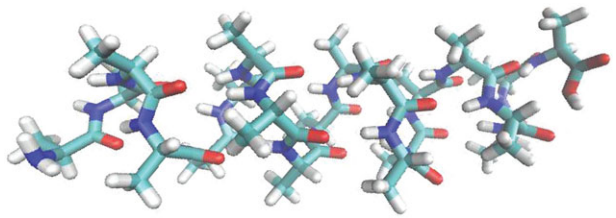

Figure 7. Optimized structure of the π-helix consisting of 20 (*S*)-alanine residues. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

All calculations were performed with the program package Turbomole 5.7.1[36] with the BP86 density functional[37,38] and Ahlrichs' def-TZVP basis set.[66] As has been shown in Refs. [79,80], the ROA intensity differences of organic molecules are not sensitive to the choice of the exchange–correlation functional. We should state here that the (unscaled) harmonic frequencies obtained with the BP86 functional are generally found to be in good agreement with measured fundamental ones, which is due to some kind of error cancellation.[14,16,23] Therefore, this functional is very well suited for vibrational analyses. In special cases, when weak interactions between noncovalently bound molecular moieties cannot be neglected, one can straightforwardly apply empirical dispersion corrections, such as the ones by Grimme and coworkers (see, e.g., Refs. [81–86]). Note, however, that there is evidence that vibrational frequencies are not affected very much by dispersive interactions.[87] Analytic electronic energy gradients and electric-dipole moments were evaluated with Turbomole, and our local version[88] of the escf module[89] of Turbomole was utilized for the polarizability tensor calculations. The resolution-of-the-identity density-fitting technique[90–92] was used in all calculations. The spectra are plotted with a Lorentzian band width at half-maximum height of 15 cm^{-1}. The velocity representation of the electric-dipole operator was used for the $\beta(\boldsymbol{G}')^2$ invariant[88] to ensure gauge invariance. An excitation wavelength of 810 nm was set to ensure that the calculations are obtained far away from any electronic absorption wavelength of the π-helical $(\mathrm{Ala})_{20}$. The molecular structure was plotted with VMD.[93]

As the π-helix model considered in this work contains 203 atoms, 603 vibrational normal modes are found. Such a high number of normal modes is already challenging for the calculation of Raman spectra but even more in the case of ROA spectroscopy where, in addition to the electric-dipole–electric-dipole polarizability tensor, the electric-dipole–magnetic-dipole and electric-dipole–electric-quadrupole polarizability tensors are needed.[12] The massively parallel and restart-friendly possibilities of Snf in combination with the efficient calculation of the ROA polarizability tensors[88] has made such calculations feasible. Using this methodology it has been possible, for instance, to present the first calculated ROA spectra of chiral metal complexes[88,94] (the first experimental ROA spectrum of

a metal complex has been reported very recently[95]), and the ROA spectrum of a rat metallothionein, which is the largest molecule (411 atoms) for which a ROA spectrum has been calculated so far.[96] These results for the ROA spectrum of rat metallothionein have also been used in another study[97] on β-turns in ROA spectra. Furthermore, we laid a strong focus of the investigation of ROA spectra of biomolecules such as sugars[98] and polypeptides,[63,99,100], where we for example studied the dependence of ROA spectra on secondary structure, and also investigated enhancement and de-enhancement effects in Resonance ROA.[101]

The ROA spectrum of the π-helical $(Ala)_{20}$ is shown in Figure 8. The most obvious difference in the calculated ROA spectrum from 1100 to 1800 cm^{-1} compared to the spectra of the α- and 3_{10}-helices (Ref. [77]) are observed in the extended amide III region (i.e., the spectral region between about 1200 to 1300 cm^{-1}), which mainly comprises in phase-combinations of N—H bending and C—N stretching vibrations of the amide group and C^{α}—H bending vibrations. Three positive bands are obtained in case of the π-helix in the wavenumber region from about 1200 to 1300 cm^{-1}, contrary to the spectra of the α- and 3_{10}-helices[77] where alternating negative and positive bands can be observed. The amide I bands between 1600 and 1700 cm^{-1} in the ROA spectrum of the π-helix are similar to the ones obtained for the α-helix,[77] namely a couplet that is negative at lower wavenumbers and positive at higher wavenumbers. This is exactly opposite to the case of the 3_{10}-helix, where in the amide I region a couplet is observed positive at lower wavenumbers and negative at higher wavenumbers, such that this couplet may serve as a signature for 3_{10}-helices.[77]

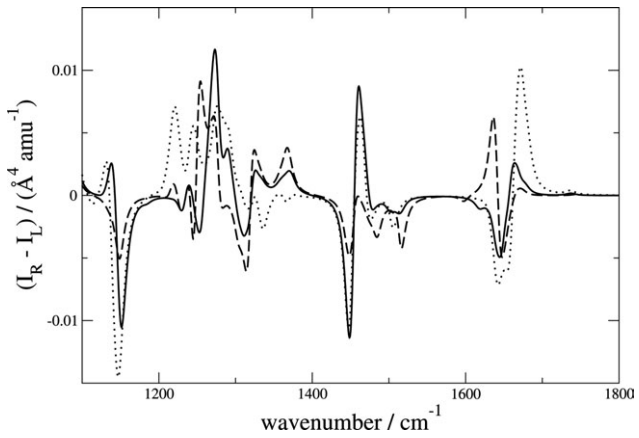

Figure 8. Calculated ROA backscattering spectrum of π-helical $(Ala)_{20}$ (dotted) alongside with the spectra of α- and 3_{10}-helical $(Ala)_{20}$ (straight and dashed lines, respectively). The data for the α- and 3_{10}-helical structures has been taken from Ref. [77].

Besides the amide I region, the spectra of α- and π-helical $(Ala)_{20}$ are also very similar in the ranges between 1400 and 1600 cm^{-1} and between 1100 and 1200 cm^{-1}. However, in the spectral range between 1320 and 1400 cm^{-1}, the sign of the ROA backscattering intensity of the π-helix is exactly the opposite of the one observed in the case of the α- and the 3_{10}-helices. Therefore, based on these data one might propose this spectral feature as a signature specific to π-helices. However, one must emphasize that the validity of this possible signature must be established by carefully analyzing additional experimental and/or theoretical data, in particular because the structure of the π-helix used in this work is somewhat distorted.

An approach which saves computer time is the above-mentioned Intensity-Tracking (compare "Intensity-Tracking" section). Here, the approximations introduced in the calculations can, in contrast to the CTTM, be controlled by the choice of the selection and convergence criteria. An IR Intensity-Tracking calculation of the π-helical $(Ala)_{20}$ is shown in Figure 9. In this calculation, the selection criterion was set to the selection of the five (approximate) normal modes with the highest IR intensity and a convergence threshold of 0.001 Hartree/(amu $\times$ $bohr^2$) for the maximum element of the residuum vector was applied.

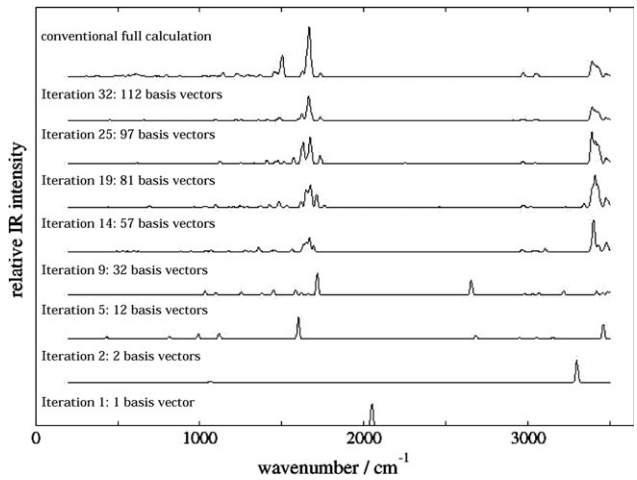

Figure 9. Intensity-Tracking spectra of different iterations of the Intensity-Tracking calculation for the π-helical $(Ala)_{20}$ compared to the conventional full calculation (top).

The first iteration contains as starting guess the ICM with the highest IR intensity (compare Fig. 10). A second basis vector is added in the second iteration leading to two bands in the spectrum (the one around 1000 cm^{-1} shows only a negligibly small IR intensity). As the calculation proceeds, more and more basis vectors are added to improve the description of the selected intense (not yet converged) normal modes. The calculation is converged in iteration 32 with 112 basis vectors, that is, 224 single-points were needed for this calculation as opposed to the full conventional calculation by the SNF program with 1218 single points. The final Intensity-Tracking spectrum contains the most intense bands around 1650 and 3400 cm^{-1} as can be seen in comparison to the full conventional calculation in Figure 9. The differences in the intensity of the bands in the Intensity-Tracking and in the conventional full calculation arise from small deviations in the normal modes which can be diminished by choosing a more strict convergence criterion. To obtain also the bands around 1500 cm^{-1}

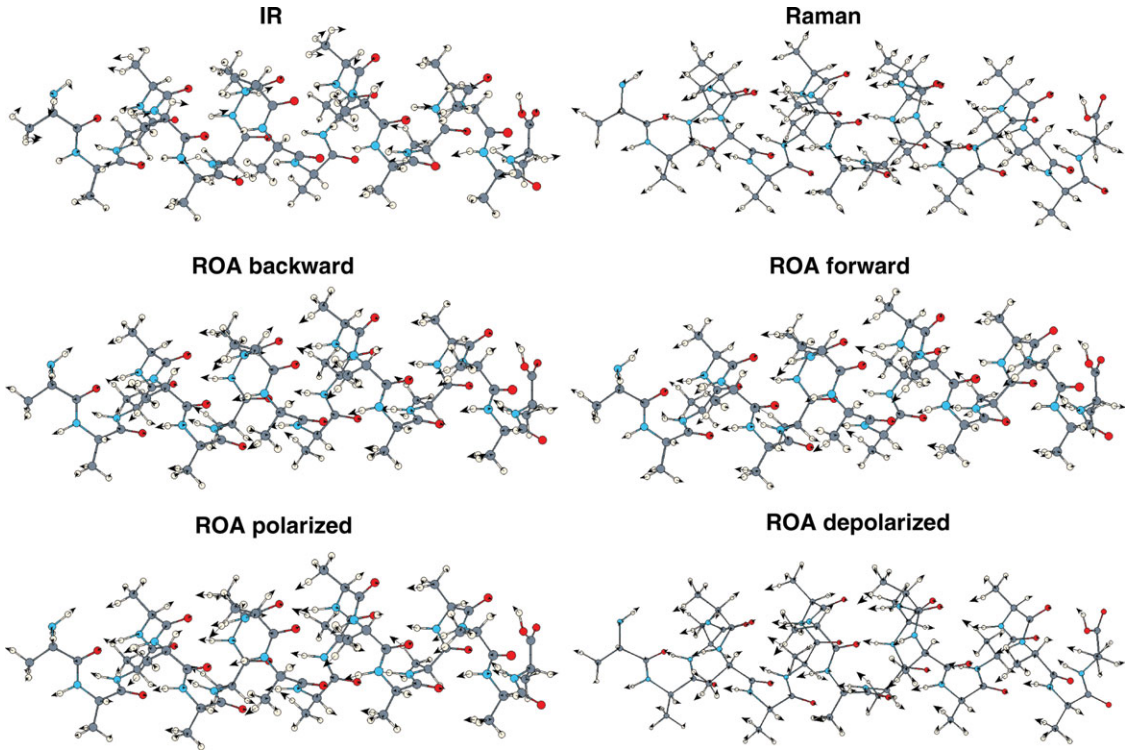

Figure 10. ICMs with the highest (absolute) IR, Raman, and ROA intensity (for the backward, forward, polarized, and depolarized ROA scattering direction) for the π-helical $(Ala)_{20}$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

observed in the full calculation, another intensity-selection criterion could have been chosen and possibly also combined with mode-based ones like a specific wavenumber range.

To demonstrate the ICMs for the model helix, the ones with the highest (absolute) IR, Raman, or ROA intensity are depicted in Figure 10.

The IR ICM contains mainly methyl group, N—H, and C^{α}—H bending vibrations. The Raman ICM features mostly hydrogen stretching vibrations including symmetric stretching vibrations of the hydrogen atoms of the methyl groups, similar to the ICMs for the ROA forward and polarized scattering directions. The ICMs for the ROA backward and depolarized scattering are also dominated by hydrogen stretching vibrations; however, they contain also the corresponding asymmetric hydrogen stretching vibration of the methyl groups.

A very powerful and intuitive concept for the analysis of vibrational spectra for large molecules introduced in 'Localizing Normal Modes' is the localization of normal modes. For instance, such an analysis has already been applied to the investigations of typical spectral signatures of 3_{10}- and α-helices.[75,77] Here, we show as an example a localized mode derived from the amide I normal modes of the π-helical $(Ala)_{20}$, which are in general dominated by carbonyl stretching and N—H bending modes. The decomposition of the ROA intensity of this localized mode into local contributions according to the scheme proposed by Hug[102] is shown on the left-hand side of Figure 11. The size of the contributions of the atomic groups is proportional to the area of the circles (filled circles denote positive and empty circles negative contributions).

As can be seen in Figure 11, the largest positive contribution indeed arises from the carbonyl group whereas negative

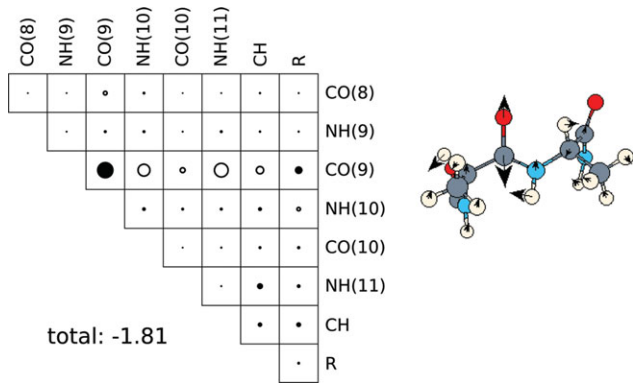

Figure 11. Analysis of a representative amide I localized mode centered on residues eight, nine, 10, and 11 of the π-helical $(Ala)_{20}$: group coupling matrix (left-hand side) and mode picture (right-hand side); 'CO' refers to the carbonyl, 'NH' to the NH, and 'CH' to the C^{α} H—CH_3 groups of the residues, R includes all contributions of atoms not shown; the total ROA intensity of this mode is also given in units of $Å^4$ amu^{-1}. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

ones originate from the N—H groups so that in total a small ROA intensity of -1.81 $Å^4$ amu^{-1} is obtained. The corresponding localized mode is visualized on the right-hand side of Figure 11. This localized mode is akin to the analogous one in the α-helical $(Ala)_{20}$.[77]

Conclusions

In this software update, we have described the new MoViPac vibrational spectroscopy program which unites the latest versions of the Snf and Akira programs for standard, Mode-

Tracking, and Intensity-Tracking calculations combined with analysis tools for the investigation of localized vibrations[74,75] and for the decomposition of vibrational intensities in terms of atomic or functional group contributions by Hug.[102] Special features of MoViPac are

1. its trivial and massive parallelism with load-balancing capabilities that provides automatic parallelization for electronic structure methods for which no parallelized analytic second derivatives of the electronic energy and property gradients are available,

2. its modular structure that has already been interfaced to five standard quantum chemistry program packages (Adf,[103] Dalton,[104] Gaussian,[105] Molpro,[106] and Turbomole[36]) and that can be easily extended to other programs for the production of the raw data,

3. its stable and efficient restart capabilities,

4. its advanced, inverse-quantum-chemical algorithms like Mode- and Intensity-Tracking, and

5. its specialized spectroscopic techniques like ROA, in particular in combination with our local version[88] of Turbomole[36] for the calculation of ROA property tensors, or the constitution of spectra using CTTM.[48]

Another feature of the program convenient for research groups is that it can be run on desktop computers and allows one to exploit a heterogeneous computer cluster in an efficient way. Moreover, the program structure allows one to easily extend the machinery in order to account for additional spectroscopic intensity expressions and new electronic structure methods implemented in quantum chemistry program packages that have not yet been interfaced.

We have applied the methodology implemented in the MoViPac to a large π-helix consisting of 20 (*S*)-alanine residues to illustrate some of its capabilities. First, we compared the ROA spectrum of this π-helix to the spectra of α- and 3_{10}-helical analogs. Major differences are observed in the extended amide I region, that is, between $\sim$1200 and 1300 cm^{-1} and between $\sim$1320 and 1400 cm^{-1}. Based on this comparison, we identified the existence of a possible signature for π-helical structures in the latter wavenumber range. However, the validity of this signature has yet to be established by further studies. Then, the performance of Intensity-Tracking has been demonstrated by successfully reproducing the dominant spectral features of high intensity with less than a fifth of the number of single-point calculations necessary in a conventional full calculation. Finally, the concept of localized modes has been illustrated by decomposing the intensity of a representative amide I vibration into local contributions.

We plan to continuously develop the MoViPac package and shall provide anharmonic corrections based on perturbation theory as well as new interfaces (e.g., to the Molcas environment[107]) in the future.

Acknowledgments

We acknowledge support by working groups 3 and 4 of COST ACTION CODECS.

Keywords: helical structures · parallel calculations · quantum chemistry · subsystem approaches · vibrational spectroscopy

How to cite this article: T. Weymuth, M. P. Haag, K. Kiewisch, S. Luber, S. Schenk, C. R. Jacob, C. Herrmann, J. Neugebauer, M. Reiher, *J. Comput. Chem.* **2012,** *33,* 2186–2198. DOI: 10.1002/jcc.23036

[1] E. B. Wilson, Jr., J. C. Decius, P. C. Cross, Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra; Dover Publications: New York, **1980**.
[2] S. Heislbetz, G. Rauhut, *J. Chem. Phys.* **2010**, *132,* 124102.
[3] D. M. Benoit, *Front. Biosci.* **2009**, *14,* 4229.
[4] M. Sparta, M. B. Hansen, E. Matito, D. Toffoli, O. Christiansen, *J. Chem. Theory Comput.* **2010**, *6,* 3162.
[5] J. M. Bowman, T. Carrington, H.-D. Meyer, *Mol. Phys.* **2008**, *106,* 2145.
[6] N. Matsunaga, G. M. Chaban, R. B. Gerber, *J. Chem. Phys.* **2002**, *117,* 3541.
[7] L. Pele, R. B. Gerber, *J. Chem. Phys.* **2008**, *128,* 165105.
[8] M. Schmitz, P. Tavan, In Modern Methods for Theoretical Physical Chemistry of Biopolymers; E. B. Starikov, J. P. Lewis, S. Tanaka, Eds.; Elsevier Science: Amsterdam, **2006**, p. 159.
[9] J. Jeon, S. Yang, J.-H. Choi, M. Cho, *Acc. Chem. Res.* **2009**, *42,* 1280.
[10] D. Marx, J. Hutter, Ab Initio Molecular Dynamics: Basic Theory And Advanced Methods; Cambridge University Press: New York, **2009**.
[11] C. Herrmann, M. Reiher, *Top. Curr. Chem.* **2007**, *268,* 85.
[12] L. D. Barron, Molecular Light Scattering and Optical Activity, 2nd ed.; Cambridge University Press: Cambridge, **2004**.
[13] W. Schneider, W. Thiel, *Chem. Phys. Lett.* **1989**, *157,* 367.
[14] J. Neugebauer, B. A. Hess, *J. Chem. Phys.* **2003**, *118,* 7215.
[15] A. Adel, D. M. Dennison, *Phys. Rev.* **1933**, *43,* 716.
[16] M. Reiher, G. Brehm, S. Schneider, *J. Phys. Chem. A* **2004**, *108,* 734.
[17] M. Reiher, B. A. Hess, *Adv. Inorg. Chem.* **2004**, *56,* 55.
[18] G. Brehm, M. Reiher, S. Schneider, *J. Phys. Chem. A* **2002**, *106,* 12024.
[19] G. Brehm, M. Reiher, B. L. Guennic, M. Leibold, S. Schindler, F. W. Heinemann, S. Schneider, *J. Raman Spectrosc.* **2006**, *37,* 108.
[20] G. Rauhut, P. Pulay, *J. Phys. Chem.* **1995**, *99,* 3093.
[21] G. Rauhut, P. Pulay, *J. Phys. Chem.* **1995**, *99,* 14572.
[22] L. Yu, C. Greco, M. Bruschi, U. Ryde, L. De Gioia, M. Reiher, *Inorg. Chem.* **2011**, *50,* 3888.
[23] A. P. Scott, L. Radom, *J. Phys. Chem.* **1996**, *100,* 16502.
[24] J. W. Tye, M. Y. Darensbourg, M. B. Hall, *J. Comput. Chem.* **2006**, *27,* 1454.
[25] P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, A. Vargha, *J. Am. Chem. Soc.* **1983**, *105,* 7037.
[26] J. Neugebauer, M. Reiher, C. Kind, B. A. Hess, *J. Comput. Chem.* **2002**, *23,* 895.
[27] B. A. Hess, Numfreq, University of Erlangen-Nürnberg, **2001**. (Based on work by S. Grimme, University of Bonn,1998, with contributions from C. Marian and M. Gastreich.)
[28] M. Reiher, J. Neugebauer, *J. Chem. Phys.* **2003**, *118,* 1634.
[29] C. Herrmann, J. Neugebauer, M. Reiher, *New J. Chem.* **2007**, *31,* 818.
[30] K. Kiewisch, J. Neugebauer, M. Reiher, *J. Chem. Phys.* **2008**, *129,* 204103.
[31] S. Luber, J. Neugebauer, M. Reiher, *J. Chem. Phys.* **2009**, *130,* 064105.
[32] S. Luber, M. Reiher, *ChemPhysChem* **2009**, *10,* 2049.
[33] K. Kiewisch, S. Luber, J. Neugebauer, M. Reiher, *Chimia* **2009**, *63,* 270.
[34] J. Neugebauer, B. A. Hess, *J. Chem. Phys.* **2004**, *120,* 11564.
[35] W. G. Bickley, *Math. Gaz.* **1941**, *25,* 19.
[36] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162,* 165.
[37] A. D. Becke, *Phys. Rev. A* **1988**, *38,* 3098.
[38] J. P. Perdew, *Phys. Rev. B* **1986**, *33,* 8822.

[39] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7,* 3297.
[40] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8,* 1057.
[41] D. Rappoport, F. Furche, *J. Chem. Phys.* **2007**, *126,* 201104.
[42] ISO/IEC JTC1/SC22, International standardization subcommittee for programming languages, their environments and system software interfaces, Available at: http://www.open-std.org/jtc1/sc22/. Accessed on May 30, 2012.
[43] GNU autoconf, Available at: www.gnu.org/software/autoconf. Accessed on May 30, 2012.
[44] GNU automake, Available at: www.gnu.org/software/automake. Accessed on May 30, 2012.
[45] N. A. Besley, K. A. Metcalf, *J. Chem. Phys.* **2007**, *126,* 035101.
[46] A. Ghysels, D. Van Neck, V. Van Speybroeck, M. Waroquier, *J. Chem. Phys.* **2007**, *126,* 224102.
[47] A. Ghysels, V. Van Speybroeck, E. Pauwels, D. Van Neck, B. R. Brooks, M. Waroquier, *J. Chem. Theory Comput.* **2009**, *5,* 1203.
[48] P. Bouř, J. Sopková, L. Bednárová, P. Maloň, T. A. Keiderling, *J. Comput. Chem.* **1997**, *18,* 646.
[49] E. R. Davidson, *J. Comp. Phys.* **1975**, *17,* 87.
[50] C. W. Murray, S. C. Racine, E. R. Davidson, *J. Comput. Phys.* **1992**, *103,* 382.
[51] M. Reiher, J. Neugebauer, *Phys. Chem. Chem. Phys.* **2004**, *6,* 4621.
[52] J. Neugebauer, *Phys. Rep.* **2010**, *489,* 1.
[53] B. R. Brooks, M. Karplus, *Proc. Natl. Acad. Sci. USA* **1985**, *82,* 4995.
[54] B. R. Brooks, D. Janežič, M. Karplus, *J. Comput. Chem.* **1995**, *16,* 1522.
[55] F. Filippone, M. Parrinello, *Chem. Phys. Lett.* **2001**, *345,* 179.
[56] F. Filippone, S. Meloni, M. Parrinello, *J. Chem. Phys.* **2001**, *115,* 636.
[57] J. Neugebauer, M. Reiher, *J. Comput. Chem.* **2004**, *25,* 587.
[58] J. Neugebauer, M. Reiher, *J. Phys. Chem. A* **2004**, *108,* 2053.
[59] A. L. Kaledin, *J. Chem. Phys.* **2005**, *122,* 184106.
[60] M. Reiher, J. Neugebauer, *J. Chem. Phys.* **2005**, *123,* 117101.
[61] A. L. Kaledin, M. Kaledin, J. M. Bowman, *J. Chem. Theory Comput.* **2006**, *2,* 166.
[62] T. B. Adler, N. Borho, M. Reiher, M. A. Suhm, *Angew. Chem. Int. Ed. Engl.* **2006**, *45,* 3440.
[63] C. Herrmann, K. Ruud, M. Reiher, *ChemPhysChem* **2006**, *7,* 2189.
[64] C. Herrmann, M. Reiher, *Surf. Sci.* **2006**, *9,* 1891.
[65] C. Herrmann, J. Neugebauer, M. Reiher, *J. Comput. Chem.* **2008**, *29,* 2460.
[66] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100,* 5829.
[67] J. Boereboom, M. C. van Hemert, J. Neugebauer, *ChemPhysChem* **2011**, *12,* 3157.
[68] E. J. Heller, R. Sundberg, D. Tannor, *J. Phys. Chem.* **1982**, *86,* 1822.
[69] A. B. Myers, *Chem. Rev.* **1996**, *96,* 911.
[70] H. Torii, Y. Ueno, A. Sakamoto, M. Tasumi, *J. Phys. Chem. A* **1999**, *103,* 5557.
[71] H. Torii, *J. Comput. Chem.* **2002**, *23,* 997.
[72] N. S. Bieler, M. P. Haag, C. R. Jacob, M. Reiher, *J. Chem. Theory Comput.* **2011**, *7,* 1867.
[73] S. Yamamoto, P. Bouř, *Collect. Czech. Chem. Commun.* **2011**, *76,* 567.
[74] C. R. Jacob, M. Reiher, *J. Chem. Phys.* **2009**, *130,* 084106.
[75] C. R. Jacob, S. Luber, M. Reiher, *J. Phys. Chem. B* **2009**, *113,* 6558.
[76] V. Liégeois, C. R. Jacob, B. Champagne, M. Reiher, *J. Phys. Chem. A* **2010**, *114,* 7198.
[77] C. R. Jacob, S. Luber, M. Reiher, *Chem. Eur. J.* **2009**, *15,* 13491.
[78] T. Weymuth, C. R. Jacob, M. Reiher, *J. Phys. Chem. B* **2010**, *114,* 10649.
[79] M. Reiher, V. Liégeois, K. Ruud, *J. Phys. Chem. A* **2005**, *109,* 7567.
[80] S. Luber, C. Herrmann, M. Reiher, *J. Phys. Chem. B* **2008**, *112,* 2218.
[81] S. Grimme, *J. Comput. Chem.* **2006**, *27,* 1787.
[82] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2007**, *9,* 3397.
[83] J. Antony, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8,* 5287.
[84] J. Antony, B. Brüske, S. Grimme, *Phys. Chem. Chem. Phys.* **2009**, *11,* 8440.
[85] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132,* 154104.
[86] S. Grimme, J. Antony, T. Schwabe, C. Mück-Lichtenfeld, *Org. Biomol. Chem.* **2007**, *5,* 741.
[87] N. Marom, A. Tkatchenko, M. Scheffler, L. Kronik, *J. Chem. Theory Comput.* **2010**, *6,* 81.
[88] S. Luber, M. Reiher, *Chem. Phys.* **2008**, *346,* 212.
[89] S. Grimme, F. Furche, R. Ahlrichs, *Chem. Phys. Lett.* **2002**, *361,* 321.
[90] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *240,* 283.
[91] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *242,* 652.
[92] Available at: ftp://ftp.chemie.uni-karlsruhe.de/pub/jbasen.
[93] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14,* 33.
[94] S. Luber, M. Reiher, *ChemPhysChem* **2010**, *11,* 1876.
[95] C. Johannessen, L. Hecht, C. Merten, *ChemPhysChem* **2011**, *12,* 1419.
[96] S. Luber, M. Reiher, *J. Phys. Chem. B* **2010**, *114,* 1057.
[97] T. Weymuth, C. R. Jacob, M. Reiher, *ChemPhysChem* **2011**, *12,* 1165.
[98] S. Luber, M. Reiher, *J. Phys. Chem. A* **2009**, *113,* 8268.
[99] C. Herrmann, K. Ruud, M. Reiher, *Chem. Phys.* **2008**, *343,* 200.
[100] C. R. Jacob, S. Luber, M. Reiher, *ChemPhysChem* **2008**, *9,* 2177.
[101] S. Luber, J. Neugebauer, M. Reiher, *J. Chem. Phys.* **2010**, *132,* 044113.
[102] W. Hug, *Chem. Phys.* **2001**, *264,* 53.
[103] G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. F. Guerra, E. J. Baerends, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22,* 931.
[104] T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, K. Ruud, H. Ågren, A. A. Auer, K. L. Bak, V. Bakken, O. Christiansen, S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez, C. Hättig, K. Hald, A. Halkier, H. Heiberg, H. Hettema, D. Jonsson, S. Kirpekar, R. Kobayashi, H. Koch, K. V. Mikkelsen, P. Norman, M. J. Packer, T. B. Pedersen, T. A. Ruden, A. Sanchez, T. Saue, S. P. A. Sauer, B. Schimmelpfenning, K. O. Sylvester-Hvid, P. R. Taylor, O. Vahtras, Dalton, A Molecular Electronic Structure Program, Release 1.2, **2001**.
[105] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09 Revision A.1; Gaussian Inc.: Wallingford CT, **2009**.
[106] R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. Lindh, A. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schtz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, H.-J. Werner, Molpro, A Package of Ab Initio Programs, designed by H.-J.Werner and P. J. Knowles, version 2002.1, **2002**.
[107] F. Aquilante, L. de Vico, N. Ferré, G. Ghigo, P.-A. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitoňák, M. Reiher, B. O. Roos, L. Serrando-Andrés, M. Urban, V. Veryazov, R. Lindh, *J. Comput. Chem.* **2009**, *31,* 224.

Received: 20 March 2012
Revised: 16 May 2012
Accepted: 19 May 2012
Published online on 20 June 2012