Finding a needle in a haystack: direct determination of vibrational signatures in complex systems†

Carmen Herrmann, Johannes Neugebauer and Markus Reiher*

Received (in Montpellier, France) 22nd December 2006, Accepted 6th March 2007
First published as an Advance Article on the web 11th April 2007
DOI: 10.1039/b618769m

Vibrational spectroscopy is a powerful tool to investigate the structure and dynamics of molecular systems. When large molecules are studied, quantum chemical calculations are used to interpret the spectra. In many cases, experimentally driven questions are related to specific regions of a vibrational spectrum, so that an assignment is only required for a subset of vibrations. This holds true in particular for biomolecules, inorganic compounds which are stabilized by many bulky ligands, or other extended systems. In standard quantum chemical calculations of the vibrational spectrum, all normal modes and frequencies of the molecule under study are determined. However, the selective calculation of only relevant information can be made much more efficient by using mode-selective techniques as provided by the mode-tracking algorithm. A critical point for the performance of the mode-tracking scheme is the preparation of a guess vibration, which is then iteratively refined. The guess defines the scientific problem which is to be studied. Various examples are presented to highlight this aspect and the features of the mode-tracking algorithm in general.

1 Introduction

The properties of macroscopic systems such as biological cells and molecular materials are determined by the properties of their molecular building blocks and the interactions between them.[1] In order to describe these molecular building blocks with high accuracy and transferability, first-principles quantum chemical calculations on comparatively large systems are required.[2,3] However, as the system size is enlarged, the calculation as well as the interpretation of the results of these calculations get more and more demanding. The standard procedure of quantum chemistry is first to calculate the total

Laboratorium für Physikalische Chemie, ETH Zürich, Hönggerberg, Wolfgang-Pauli-Str. 10, CH-8093 Zürich, Switzerland. E-mail: markus.reiher@phys.chem.ethz.ch; Fax: +41-44-63-31594; Tel: +41-44-63-34308

† The HTML version of this article has been enhanced with colour images.

Carmen Herrmann studied chemistry at the University of Erlangen-Nuremberg. She earned her PhD under the supervision of Prof. Dr Markus Reiher at the universities of Bonn and Jena, and finally at ETH Zurich, with funding by a Doktorandenstipendium of the Fonds der Chemischen Industrie. Her work focussed on the extraction of chemical relevant information on local properties (such as partial spins and local molecular vibrations). She is now working as a postdoc in Prof. Reiher's group at ETH Zurich.

Johannes Neugebauer obtained his PhD from the university of Erlangen-Nuremberg in 2003, where he studied efficient methods for the theoretical calculation of vibrational spectra under the supervision of Prof. Bernd Artur Hess funded by a Kekulé fellowship of the Fonds der Chemischen Industrie (FCI). From 2003 to 2006 he was a postdoctoral fellow with Prof. Evert Jan Baerends at the Vrije Universiteit Amsterdam. During that time, he developed methods for modeling environmental effects on molecular properties employing subsystem formulations within density functional theory. Further aspects of his work in Amsterdam concerned the vibrational structure of electronic spectra and time-dependent density functional theory. In 2006, he was offered an Emmy Noether grant by the Deutsche Forschungsgemeinschaft (DFG). Since March 2006, he is working at the Laboratory of Physical Chemistry at ETH Zurich as an independent researcher, funded by a Liebig-Stipendium of the FCI. His main focus is on theoretical methods for chromophore-specific spectroscopy, in particular on the development of subsystem approaches for chromophores in complex environments, and on couplings between electronic and vibrational degrees of freedom in resonance Raman and absorption spectroscopy.

Markus Reiher is Professor for Theoretical Chemistry at the Laboratory for Physical Chemistry at ETH Zurich. He earned his PhD in theoretical chemistry from the University of Bielefeld with Professor Juergen Hinze, PhD, in 1998. After completing the work for his habilitation in theoretical chemistry at the University of Erlangen-Nuremberg with Professor Dr. Bernd Artur Hess in 2002 he worked as a Privatdozent at the University of Bonn (2003–2005). During this time he was also the representative of the vacant Chairs of Theoretical Chemistry at Erlangen (2003/2004) and at Bonn (2004/2005). In December 2004 he was offered a position as Full Professor for Theoretical Chemistry at the University of Groningen but decided to accept a position as Professor for Physical Chemistry (designation: Theory) at the University of Jena in April 2005. Markus Reiher was a research fellow during short-time stays in Tel Aviv (2000), Budapest (2001), Tromsø (2003/2004), and Lund (2006). His awards include the ADUC prize 2004, the Emmy-Noether habilitation award 2003 of the University of Erlangen-Nuremberg, and in 2005 the Dozentenstipendium of the Fonds der Chemischen Industrie. Current research interests of his group comprise topics from relativistic quantum chemistry, bioinorganic and coordination chemistry, theoretical spectroscopy and the foundations of chemistry.

wave function or electron density of the molecular system under study, and then to extract chemically relevant information from it. This information may be, *e.g.*, molecular vibrations associated with high spectroscopic intensities, such as carbonyl C=O stretching modes, electronic excitation energies, or local electronic properties such as partial charges and partial spins.

From the point of view of theory applied to actual chemical problems, the standard procedure should be turned upside down, *i.e.*, the pre-determination of the relevant information should precede the calculation of the wave function or electron density, which would allow for an efficient and focused calculation of the relevant information only. Such "inverted" approaches are currently being developed, for example, for the targeted design of drug candidates,[4] as well as for the design of molecules with specific tailored properties such as the polarizability.[5] In a similar spirit, a method for the calculation of wave functions of solids with predefined energy eigenvalues has been derived[6] and applied to energy-selective band-structure calculations.[7] Here, we review our efforts in the field of inverse quantum chemical calculations for the direct targeting of molecular properties, in particular of molecular vibrations.

Our work concentrates on the selective calculation of vibrational normal modes and harmonic frequencies with predefined properties using the mode-tracking algorithm.[8] Because of the large number of special vibrational spectroscopic techniques such as resonance-Raman, Raman optical activity or difference infrared spectroscopy, and the size of many interesting molecules (like *e.g.* large biomolecules), it is often difficult to interpret recorded spectra based on experience and empirical rules of thumb alone. It might even be impossible to assign important vibrational peaks uniquely. In such cases, quantum chemical calculations of vibrational spectra can help to make reliable connections between molecular structure and spectra.[9] Our mode-tracking protocol is particularly suited for the efficient calculation of local vibrations, such as those of small ligands bound to metal centers of enzyme active sites. The vibrational properties of such systems may serve as a tool to elucidate bonding of the small ligands to the active sites (see refs. 10–12 for some recent examples).

This work is organized as follows: In order to provide the necessary theoretical foundations, the Born–Oppenheimer approximation, some important methods to calculate the electronic potential energy surface, and the standard approach to quantum chemical vibrational analyses are briefly outlined in Section 2. In the following section, the selective calculation of vibrational frequencies and normal modes employing the mode-tracking algorithm is presented, including a comparison to other subsystem-oriented approaches as well as a description of the computational implementation of the algorithm. The next three sections provide illustrative examples for applications of the mode-tracking protocol to large molecular systems. In Section 4, the selective calculation of NO stretching modes in ruthenium-based NO and NOH complexes is studied as an example for a prototypical application of the mode-tracking protocol to local vibrations. In Section 5, mode-tracking applied to vibrations of a flexible thiophenolate adsorbate attached to an Ag_{51} cluster is reviewed, while Section 6 shows how the breathing mode of helical decaalanine was obtained very efficiently. Finally, an outlook on possible future developments is given in Section 7.

2 Theoretical basics of vibrational spectroscopy

The basis of the quantum chemical description of molecular vibrations of molecules with more than a few atoms is provided by the Born–Oppenheimer approximation,[13] which allows for a separate description of the motion of atomic nuclei and electrons.

2.1 Born–Oppenheimer approximation

The nuclear and the electronic degrees of freedom are coupled in the time-independent nonrelativistic Schrödinger equation

$$\hat{H}_{\mathrm{mol}}\Psi_I=[\hat{T}_{\mathrm{K}}+\hat{T}_{\mathrm{e}}+\hat{V}_{\mathrm{nuc,nuc}}+\hat{V}_{\mathrm{nuc,e}}+\hat{V}_{\mathrm{e,e}}]\Psi_I=E_I\Psi_I, \quad (1)$$

(where I labels the state of the system) by the electron–nucleus interaction Hamiltonian $\hat{V}_{\mathrm{nuc,e}}$, whereas the contributions to the total molecular Hamiltonian $\hat{H}_{\mathrm{mol}}$ which describe the kinetic energy of the nuclei ($\hat{T}_{\mathrm{K}}$), the kinetic energy of the electrons ($\hat{T}_{\mathrm{e}}$), the interaction of the nuclei ($\hat{V}_{\mathrm{nuc,nuc}}$), and the interaction of the electrons ($\hat{V}_{\mathrm{e,e}}$) refer to either nuclear or electronic coordinates only and thus do not prevent a separation of the Schrödinger equation into a nuclear and an electronic part. The Born–Oppenheimer approximation, which is closely related to the adiabatic approximation, decouples the motion of nuclei and electrons. Both terms are not

rigorously used in the literature and are sometimes employed as synonyms. In both cases, the wave function $\Psi_I(\boldsymbol{R}, \boldsymbol{r})$ is separated into an electronic and a nuclear part. The nuclear part $\chi_a(\boldsymbol{R})$ depends on the $3M$ coordinates $\boldsymbol{R}$ of all M nuclei, and the electronic part $\Psi_{e,i}(\boldsymbol{r}; \boldsymbol{R})$ depends explicitly on the $3N$ coordinates $\boldsymbol{r}$ of all N electrons, and parametrically on the nuclear coordinates (which is expressed by the semicolon),

$$\Psi_I(\boldsymbol{R}, \boldsymbol{r}) = \chi_a(\boldsymbol{R}) \cdot \Psi_{e,i}(\boldsymbol{r}; \boldsymbol{R}). \quad (2)$$

The indices a and i label nuclear and electronic energy levels, respectively, and are joined in the composite total index I. In most quantum chemical calculations, as in Sections 5 and 6, the action of the nuclear kinetic energy operator $\hat{T}_K$ onto the electronic wave function is neglected. As a result, after inserting eqn (2) into eqn (1), multiplying from the left with $\Psi^*_{e,i}(\boldsymbol{r}; \boldsymbol{R})$ and integrating over all electronic coordinates, the nuclear Schrödinger equation,

$$[\hat{T}_K + E_{e,i}(\boldsymbol{R})]\chi_a(\boldsymbol{R}) = E_{tot}\chi_a(\boldsymbol{R}), \quad (3)$$

is obtained, which describes the nuclei as moving on the potential energy surface (PES) defined by $E_{e,i}(\boldsymbol{R})$ of the electronic state i. The PES may be obtained pointwise by solving the electronic Schrödinger equation,

$$[\hat{T}_e + \hat{V}_{nuc,nuc} + \hat{V}_{nuc,e} + \hat{V}_{e,e}]\Psi_{e,i}(\boldsymbol{r}; \boldsymbol{R}) = E_{e,i}(\boldsymbol{R})\Psi_{e,i}(\boldsymbol{r}; \boldsymbol{R}) \quad (4)$$

for different positions of the nuclei, *i.e.*, for different molecular structures. This provides the potential in which the nuclei are moving. The calculation of the electronic PES will be sketched in the following paragraph, while the standard approach for calculating eigenfunctions and eigenvalues of the nuclear Hamiltonian will be presented in Section 2.2.

In order to compute the PES, a variety of first-principles methods has been developed, which differ in terms of accuracy and efficiency.[3,14] First-principles calculations differ from purely empirical and semi-empirical methods in that although they employ experimental data, these are either only natural constants (in pure *ab initio* methods) or used in a very general form. The latter applies to Kohn–Sham density functional theory (KS-DFT)[15–19] with approximate density functionals, which is of particular importance for the applications presented here since it provides an efficient description of electron correlation. Parameters for the approximate density functionals may be obtained from fitting to experimental data or to data from parameter-free, high-accuracy *ab initio* methods. In contrast to genuine semi-empirical methods, however, the density functionals obtained in such a way yield high-quality results even for molecules which are structurally very different from the ones which have been used in the test set of the fitting procedure. In addition, the number of parameters in well-established functionals is astonishingly small.

It should be noted that in most experimental setups, time-averaged properties of a bulk of molecules are measured, which may be accounted for by describing a suitable ensemble of molecules within a molecular dynamics (MD) framework and under periodic boundary conditions.[20,21]

2.2 Calculation of vibrational normal modes and frequencies: standard approach

Once it is clear how to solve the electronic Schrödinger eqn (4) for a given nuclear configuration using the methods discussed in the previous section, all relevant points on the PES can, in principle, be calculated, although this soon becomes unfeasible. Knowledge of the PES is necessary in order to solve the nuclear Schrödinger eqn (3). The most simple expression for the nuclear wave function $\chi_a(\boldsymbol{R})$ is a product of functions which each only depend on one nuclear coordinate. This product would be an exact *ansatz* for the solution of eqn (3) if the individual coordinates were not coupled through $E_{e,i}(\boldsymbol{R})$ in the Hamiltonian in eqn (3). The first step to achieve such a decoupling is to express the dependence of the PES on Cartesian nuclear coordinates as a Taylor series around the nuclear equilibrium structure $\boldsymbol{R} \equiv 0$,

$$E_{e,i}(\boldsymbol{R}) = E_{e,i}(0) + \sum_{A=1}^{3M} \left(\frac{\partial E_{e,i}}{\partial R_A}\right)_0 R_A + \frac{1}{2}\sum_{A=1}^{3M}\sum_{B=1}^{3M} \left(\frac{\partial^2 E_{e,i}}{\partial R_A \partial R_B}\right)_0 R_A R_B + \ldots, \quad (5)$$

and to truncate it after the second-order term (harmonic approximation). It is most convenient to treat vibrational motions in terms of mass-weighted coordinates, which are related to non-mass-weighted (n.m.w.) ones as $R_A = M_A^{1/2} \cdot R_A^{(n.m.w.)}$, where A refers to the x, y or z coordinate of nucleus A and M_A is its mass. If not mentioned explicitly otherwise, all nuclear coordinates in the following are mass-weighted ones. Then, considering that the energy gradient is zero at the equilibrium (=minimum) structure and choosing the equilibrium energy as zero reference energy, only the sum over second derivatives in eqn (5) remains. In order to achieve a decoupled description of the nuclear vibrational motions, the equation is subsequently transformed to a set of $3M$ mass-weighted normal coordinates $\boldsymbol{q}_A$, which makes all mixed second derivatives of the electronic energy $(\partial^2 E_{e,i}/\partial \boldsymbol{q}_A \partial \boldsymbol{q}_B)_{A \neq B}$ vanish. Hence, a quantum chemical vibrational analysis comprises the calculation and diagonalization of the (mass-weighted) Hessian $\boldsymbol{H} = (\partial^2 E_{e,i}/\partial R_A \partial R_B)$,[22]

$$\boldsymbol{Q}^T \boldsymbol{H} \boldsymbol{Q} = \boldsymbol{H}^Q, \quad (6)$$

in order to achieve that $(\partial^2 E_{e,i}/\partial \boldsymbol{q}_A \partial \boldsymbol{q}_B)_{A \neq B} = 0$. The columns of the matrix $\boldsymbol{Q}$ which diagonalizes $\boldsymbol{H}$ are the (mass-weighted) normal modes, which describe the collective motions of the nuclei.

The nuclear kinetic energy operator $\hat{T}_K = \sum_A \partial^2/\partial \boldsymbol{q}_A{}^2$, is also diagonal with respect to these coordinates, so that the product

$$\chi_a(\{\boldsymbol{q}_I\}) = \psi_1(\boldsymbol{q}_1) \cdot \psi_2(\boldsymbol{q}_2) \cdots \psi_{3M}(\boldsymbol{q}_{3M}) \quad (7)$$

is an exact solution of the nuclear Schrödinger equation in harmonic approximation, which is now separable into $3M$ independent equations of harmonic oscillator form with a mass equal to unity (since the masses are absorbed in the mass-weighted coordinates). The vibrations of the molecule are thus described as a set of decoupled harmonic oscillators. The eigenfunctions of the harmonic oscillator equation are

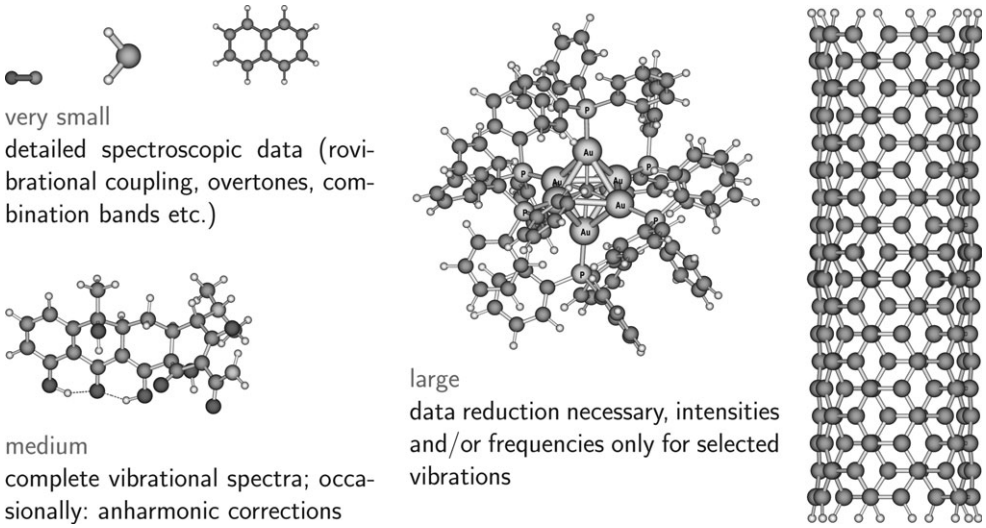


Fig. 1 Overview of molecule size classifications and the feasibility of present-day quantum chemical methods to calculate their spectra.

well-known,[23,24] and their eigenvalues (the angular frequencies ω_I) are equal to the square root of the second energy derivatives. Of course, since $\omega = 2\pi\nu$ and $\tilde{\nu} = \nu/c$, vibrational frequencies ν and wavenumbers $\tilde{\nu}$ can be easily calculated from the eigenvalues ($\partial^2/E_{e,i}/\partial \boldsymbol{q}_A \partial \boldsymbol{q}_A$) of the Hessian matrix.

Truncating the Taylor series expansion of the PES in eqn (5) is mandatory for molecules with about five atoms or more for reasons of computational feasibility. For smaller molecules, the Taylor expansion can be circumvented by calculating the PES explicitely pointwise. If, for example, the PES of a triatomic molecule is calculated using 10 points along each of the three internal coordinates, the total electronic energy needs to be evaluated at $10^3 = 1000$ different molecular structures (=points on the PES). By adding only one atom, the dimensionality of the problem rises by a factor of 1000, since the system now has six independent internal coordinates, so $10^6 = 1\,000\,000$ points would have to be evaluated. Even if this increase can be damped by special techniques tailored for four-atomic molecules, the general exponential increase can hardly be avoided. This consideration shows why large molecules need to be treated employing additional approximations in order to reduce the dimensionality. In practice, large systems will have to be treated within the harmonic approximation. For medium-sized molecules, it is routinely possible to capture the deviations of the harmonic approximation from the true PES either by a variational procedure, or by calculating anharmonic corrections—*i.e.*, higher terms in the Taylor series—in a second step after solving the vibrational problem in harmonic approximation (see, *e.g.*, ref. 25). An overview of typical molecular sizes and the capabilities of present-day quantum chemical methods to describe their vibrational spectra is given in Fig. 1.

3 Inverse calculation of pre-selected vibrations *via* mode-tracking

Especially when large molecules are investigated, the interesting parts of the vibrational spectrum may be localized on a certain fragment of the molecule. Conventional quantum chemical calculations of vibrational spectra, however, yield all vibrational modes of the molecule through eqn (6). When first-principles quantum chemical methods such as DFT with sufficiently large basis sets shall be employed, the calculation of the full Hessian matrix for systems with 100 atoms or more requires enormous computational resources. Fig. 2 shows an example of a large molecule with bulky substituents and its calculated infrared spectrum. This complex contains 182 atoms, which corresponds to 546 degrees of freedom and 540 vibrational modes. Bulky substituents are often ignored in quantum chemical calculations, although their effect on vibrational spectra can be significant[26] as highlighted in the infrared spectrum depicted in Fig. 2

3.1 Mode-tracking: basic idea

If only those parts of the vibrational spectrum were calculated with a mode-selective technique which are required for a particular scientific purpose (*e.g.*, for the structural characterization), the effort would be heavily reduced and even larger molecules would become accessible to first-principles methods. Aiming at a target-oriented solution of problems, it actually seems unreasonable first to calculate much more information than necessary associated with a huge computational effort just to discard the major part of the data in a second step. Moreover, the latter step may require to find "a needle in a haystack" if a particular vibration among hundreds of others must be identified by visual inspection of the whole set of modes $\{\boldsymbol{q}_\mu\}$. In contrast to this, a mode-specific technique yields "the needle without the haystack". In order not to loose accuracy in such a mode-specific approach, a necessary condition is that no (or only well-controlled) approximations are introduced in addition to those already made in the calculation of full vibrational spectra. "Well-controlled" means that the size of the error introduced is known. For normal modes, all of this is achieved by the mode-tracking protocol,[8] which makes the selected calculation of exact normal modes possible based on preconstructed initial guess vibrations. Various applications of mode-tracking are documented in refs. 27–35. The

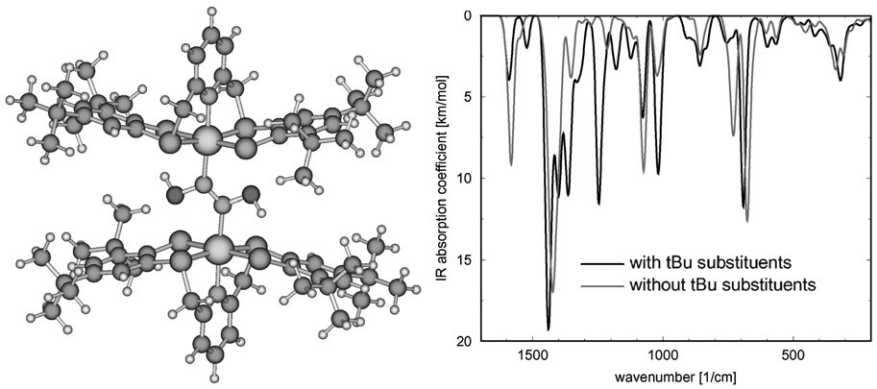


Fig. 2 Calculated infrared spectrum of a $(NOH)_2$ iron–sulfur metal complex, calculated with and without inclusion of the bulky *tert*-butyl substituents (BP86/RI/SVP).[26]

numerical stability and efficiency of the algorithm has been analyzed in detail in ref. 29. In Fig. 3, the principle of data reduction in vibrational calculations is schematically shown. An efficient implementation of approach A for first-principles methods which provide analytical gradients can be found in ref. 22, while approach B is highlighted in ref. 36. The mode-tracking protocol[8] implements the idea of approach C.

Mode-tracking has been shown to work for various types of local subsystems—for a flexible adsorbate on a rigid metal cluster[30] (see middle right panel of Fig. 4), for carbon chains in dinuclear polyynediyl rhenium complexes[27] (lower right panel), for OH fragments in weakly interacting oligomers[32] (lower left panel), as well as for calculating the effects of the periphery of a large transition metal complex on local vibrations[28] and other large aggregates[37,38] (middle left panel), for local stretching vibrations of small ligands in large aggregates,[38] and for ligand vibrations in CO-myoglobin[39] (upper left panel). It has also been successfully applied to the selective calculation of vibrations which involve atoms in the whole system, and which are characterized by, *e.g.*, a certain symmetry, or by breathing motions (see upper right panel of Fig. 4).

Two of these examples, a thiophenolate adsorbate on a silver cluster (middle right panel) and helical decaalanine (upper right panel) will be reviewed in further detail after discussing some algorithmic and practical aspects of the mode-tracking scheme.

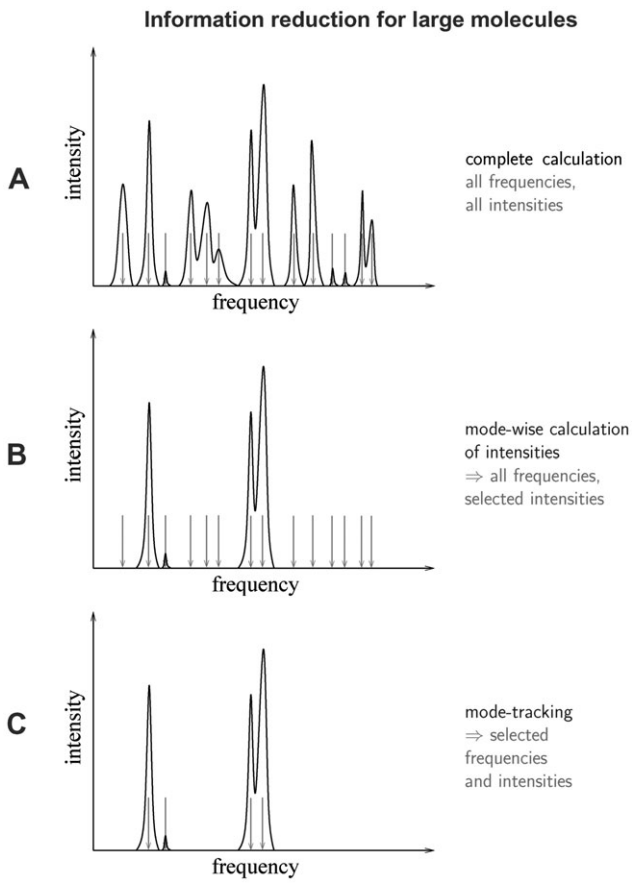


Fig. 3 Illustration of the central goal of mode-tracking compared to standard techniques: calculate only the relevant information, *i.e.*, only selected vibrational frequencies and intensities.

3.2 A closer look: mode-tracking as a projection onto nuclear displacement basis vectors

Since mode-tracking projects onto a relevant subset of vibrations, it is helpful to formulate this action explicitely by a projection operator. The vibrational normal mode vectors $\{\boldsymbol{q}_p\}$ of a molecule may be projected onto an orthonormal basis of nuclear displacement vectors $\{\boldsymbol{a}_j\}$,

$$\hat{P}_{\{\boldsymbol{a}_j\}}\boldsymbol{q}_p \equiv \sum \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{q}_p = \sum C_{jp}\boldsymbol{a}_j, \tag{8}$$

with the projector $\hat{P}_{\{\boldsymbol{a}_j\}}$ onto the basis set $\{\boldsymbol{a}_j\}$. The scalar product $\boldsymbol{a}_j^{\mathrm{T}}\boldsymbol{q}_p$ is abbreviated as the expansion coefficient C_{jp} (the superscript T denotes the transpose of a vector). Since the basis set is orthonormal, $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j = \delta_{ij}$, $\hat{P}_{\{\boldsymbol{a}_j\}}$ is idempotent,

$$\hat{P}^2_{\{\boldsymbol{a}_j\}} = \left(\sum_j \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}}\right)^2 = \sum_{ij} \boldsymbol{a}_i\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}} = \sum_{ij} \boldsymbol{a}_i\delta_{ij}\boldsymbol{a}_j^{\mathrm{T}} = \sum_j \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}} = \hat{P}_{\{\boldsymbol{a}_j\}}, \tag{9}$$

and symmetric,

$$\hat{P}^{\mathrm{T}}_{\{\boldsymbol{a}_j\}} = \left(\sum_j \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}}\right)^{\mathrm{T}} = \sum_j \boldsymbol{a}_j\boldsymbol{a}_j^{\mathrm{T}} = \hat{P}_{\{\boldsymbol{a}_j\}}\,. \tag{10}$$

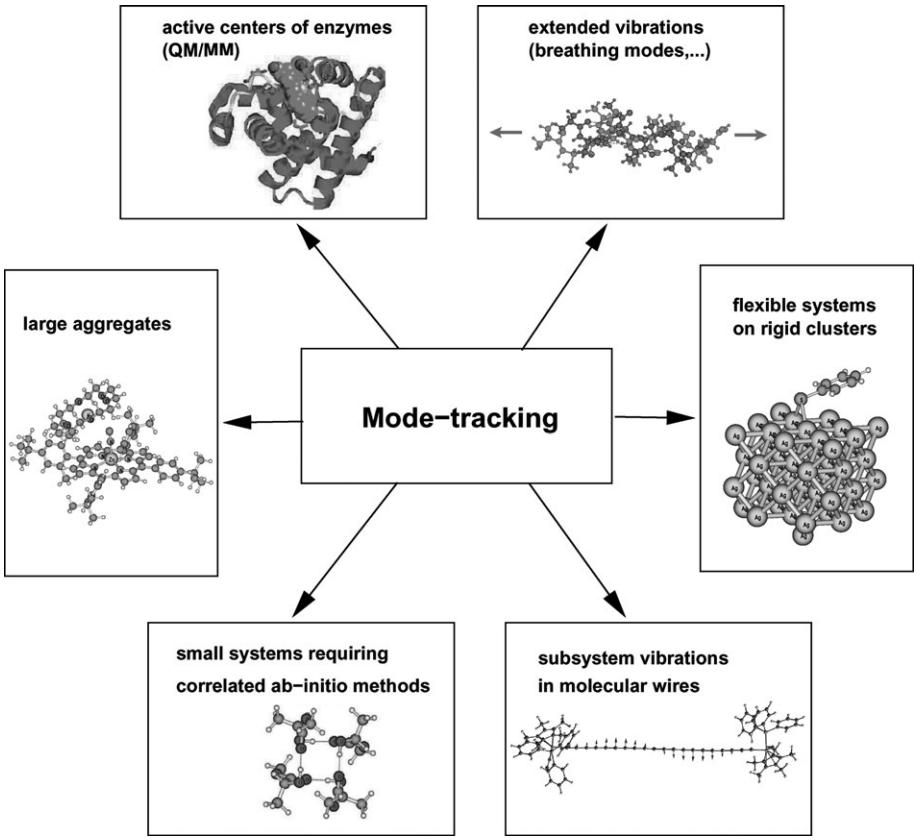


Fig. 4 Some examples of the range of applicability of the mode-tracking scheme.

Since the nuclear displacement vectors are real, it follows from eqn (10) that the projection operators are Hermitian. If the vectors $\{\boldsymbol{a}_j\}$ form a complete basis, the projector $\hat{P}_{\{\boldsymbol{a}_j\}}$ will be equal to the unity operator,

$$\hat{P}_{\{\boldsymbol{a}_j\}}\boldsymbol{q}_p = \hat{1}\boldsymbol{q}_p = \boldsymbol{q}_p, \tag{11}$$

so all normal modes are eigenvectors of $\hat{P}_{\{\boldsymbol{a}_j\}}$ as one would require. If, however, a basis set $\{\boldsymbol{b}_j\}$ is incomplete with respect to the vector space spanned by the normal mode vectors (*e.g.* because the number of orthonormal basis vectors is lower than the number of normal modes), then the normal modes will not necessarily be eigenvectors of $\hat{P}_{\{\boldsymbol{b}_j\}}$. It is nonetheless possible that the small basis spans a subspace of the space spanned by $\{\boldsymbol{q}_p\}$ which contains one or several normal mode vectors $\{\boldsymbol{q}_s\}$, so that these normal mode vectors are eigenvectors of $\hat{P}_{\{\boldsymbol{b}_j\}}$,

$$\hat{P}_{\{\boldsymbol{b}_j\}}\boldsymbol{q}_s = \boldsymbol{q}_s. \tag{12}$$

If one is interested in one special normal mode $\boldsymbol{q}_s$, the goal is to construct a basis set of distortion vectors $\{\boldsymbol{b}_j\}$ which is as small as possible, but in which $\boldsymbol{q}_s$ can be expanded either exactly or to a well-defined accuracy. The most obvious basis set of this kind would consist of only one vector, namely $\boldsymbol{q}_s$ itself. Since $\boldsymbol{q}_s$ is not known exactly from the outset, a workaround has to be found, which is provided by the mode-tracking algorithm. Mode-tracking may be considered as a method to construct an orthonormal basis set which is as small as possible, starting from one initial guess vector and expanding it iteratively until the resulting basis set spans a subspace containing the desired vibrational coordinate $\boldsymbol{q}_s$. Only in unfortunate cases, this set has to be complete in a sense that a full vibrational analysis must be performed. However, the mode-tracking algorithm will automatically account for these cases.

The construction of such an algorithm may be based on the following considerations: Using the definition of the projection operator in eqn (8), the projection of a normal mode vector $\boldsymbol{q}_s$ onto a small basis set $\{\boldsymbol{b}_j\}$ which is complete with respect to $\boldsymbol{q}_s$ can be reformulated in matrix form as

$$(\hat{P}_{\{\boldsymbol{b}_j\}}\,\boldsymbol{Q})_{As} = Q_{As} = \sum_j C_{js}B_{Aj} = \sum_j B_{Aj}C_{js} = (\boldsymbol{BC})_{As}, \tag{13}$$

(where B_{Aj} denotes the Ath element of the basis vector $\boldsymbol{b}_j$, and accordingly Q_{As} the Ath element of the vector $\boldsymbol{q}_s$, and the sum over j is running over all basis vectors) or

$$\boldsymbol{Q} = \boldsymbol{BC} \Leftrightarrow \boldsymbol{Q}^{\mathrm{T}} = (\boldsymbol{BC})^{\mathrm{T}} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}. \tag{14}$$

This result can be used to replace $\boldsymbol{Q}$ and $\boldsymbol{Q}^{\mathrm{T}}$ in eqn (6); note, however, that $\boldsymbol{Q}$ is now a rectangular matrix (at least as long as it is not complete with respect to all normal modes),

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{HQ} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{HBC} = \boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{H}}\boldsymbol{C} = \boldsymbol{H}^{(Q)}, \tag{15}$$

where the small Davidson matrix $\tilde{\boldsymbol{H}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{HB}$ has been introduced. For an illustration of the matrix multiplications involved, see Fig. 5.

For a given set of basis vectors, it is possible to calculate $\tilde{\boldsymbol{H}}$ without knowing the full Hessian matrix $\boldsymbol{H}$. For this, a matrix $\boldsymbol{\Sigma}^{(i)} = \boldsymbol{HB}^{(i)}$ is defined (where an index for the ith iteration has

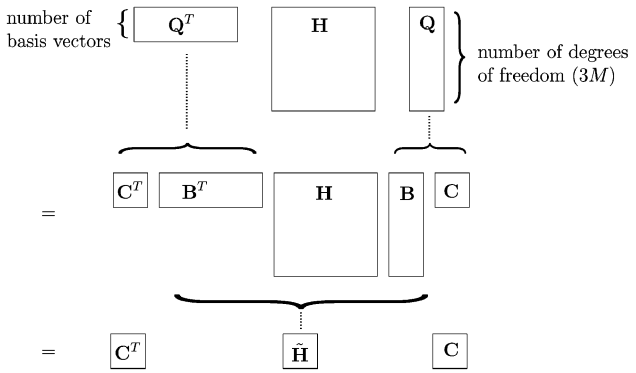


Fig. 5 Schematic illustration of the matrix multiplications involved in eqn (15). Some (but not necessarily all) of the columns of the matrix $\boldsymbol{Q}$ are eigenvectors $\boldsymbol{q}_s$ of the Hessian (*i.e.*, those vectors which can be expanded exactly in the basis set $\{\boldsymbol{b}_j\}$).

been introduced in order to indicate that the basis set is extended in each iteration), whose elements $\Sigma_{Aj}^{(i)}$ are evaluated as

$$\Sigma_{Aj}^{(i)} = \sum_{B=1}^{3M} \left(\frac{\partial^2 E_{e,i}}{\partial R_A \partial R_B}\right)_0 B_{Bj}^{(i)} = \left(\frac{\partial^2 E_{e,i}}{\partial R_A \partial \boldsymbol{b}_j}\right)_0 = \left[\frac{\partial}{\partial \boldsymbol{b}_j}\left(\frac{\partial E_{e,i}}{\partial R_A}\right)\right]_0 . \quad (16)$$

The $B_{Bj}^{(i)}$ are the coefficients for the Cartesian basis vector R_B in the basis vectors $\boldsymbol{b}_j$, and $E_{e,i}$ denotes the total electronic energy. The components of $\boldsymbol{\Sigma}^{(i)}$ can thus be calculated numerically as directional derivatives of the gradient components along the basis vectors $\boldsymbol{b}_j$. Therefore, provided that a suitable, small set of basis vectors has been found, it is possible to calculate exact normal modes by constructing $\tilde{\boldsymbol{H}}$ and diagonalizing it.

The practical problems the mode-tracking algorithm has to solve are therefore (1) to check whether a small basis set is appropriate, that is whether it spans a subspace containing at least the desired normal mode vector(s), and if this is not the case, (2) to enlarge the insufficient basis set in a fashion which makes it converge to an appropriate basis set as efficiently as possible. This leads to an iterative algorithm which consists of alternating "basis set check" and "basis set expansion" steps until convergence is achieved (see Fig. 6).

The convergence check may be based on several criteria, such as the maximum component of the residuum vectors, which are defined in Fig. 6. In order to construct an additional basis vector $\boldsymbol{b}_{i+1}$ in each iteration, a preconditioner $\boldsymbol{X}^{(i)}$ is applied,

$$\boldsymbol{b}_{i+1} = \boldsymbol{X}^{(i)} \boldsymbol{r}_\mu, \quad (17)$$

which determines an optimal new basis vector per tracked normal mode $\boldsymbol{q}_\mu^{(i)}$ from the corresponding residuum vector $\boldsymbol{r}_\mu^{(i)}$.[29] A schematic flow chart for the mode-tracking algorithm is given in Fig. 7.

3.3 Special features of mode-tracking compared to other subsystem-oriented approaches

Subspace iteration techniques come immediately to one's mind if *huge* Hessian matrices are to be diagonalized as in molecular mechanics applications,[40–43] as these techniques were developed for such purposes. Iterative refinement schemes have also been applied to large fractions of a vibrational spectrum of a large molecule (usually the upper or lower part of the spectrum is then obtained[40,41]) instead of the selected, narrow sections of the spectrum as in mode-tracking. These approaches are based on techniques[44–49] which make explicit reference to the form of the converged eigenvectors,[34] which is not needed in the mode-tracking protocol. The mode-tracking protocol comprises three additional points: (1) subspace iteration techniques can be fruitfully employed if the matrix to be diagonalized is comparatively small but its entries are difficult to calculate; (2) depending on the local nature of the desired eigenvector, its targeted calculation can be highly efficient so that only a very small number of iterations is actually required; and (3) any vector from any region of the spectrum of eigenvalues can be targeted.

In contrast to other approaches which reduce the computational effort by concentrating on selected vibrations, the mode-tracking technique converges a set of initial guess vectors to *exact* eigenvectors of the full Hessian provided the

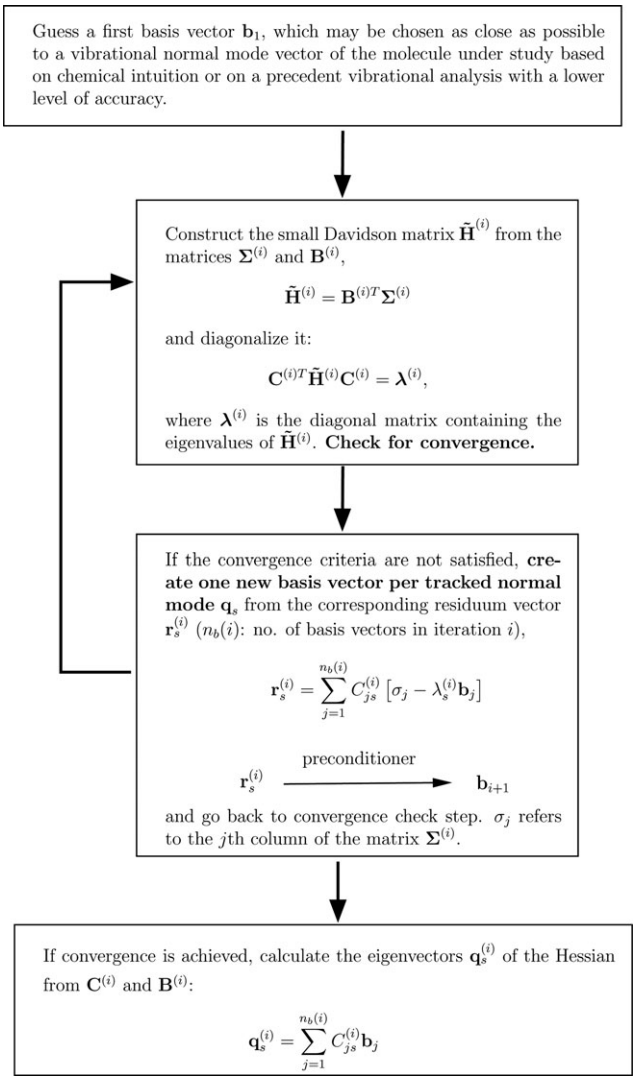


Fig. 6 Scheme of the mode-tracking algorithm. For further technical details on the algorithm, the reader is referred to ref. 29.

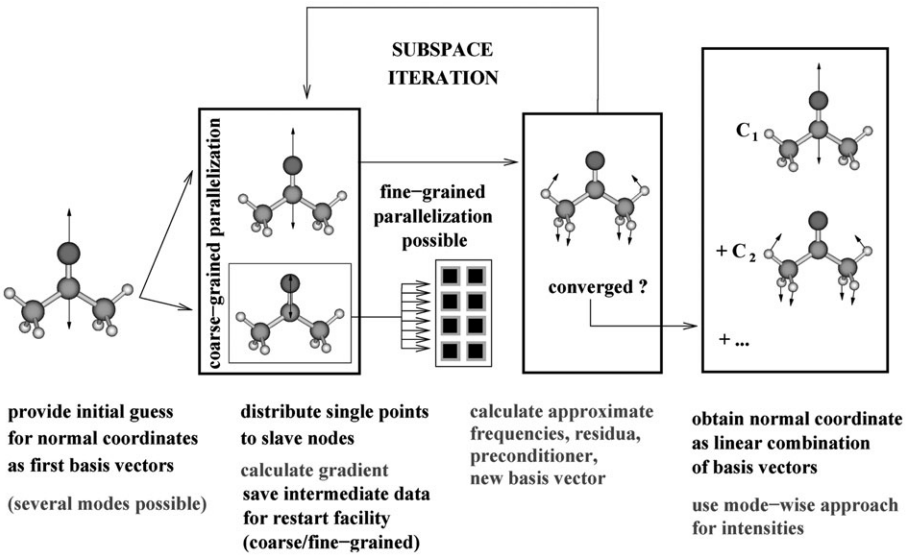


Fig. 7 Flow-chart for mode-tracking calculations.

convergence criteria are chosen sufficiently tight. One example for such approaches is the partial Hessian vibrational analysis technique,[50] where the problem is simplified by diagonalizing only a subsystem block of the Hessian while discarding couplings to the remaining degrees of freedom. A recent application is reported in ref. 51. For systems with a repeated structural motif, such as the amide bond in peptides, vibrational modes which are dominated by motions of the atoms belonging to this structural motif may also be calculated selectively in an approximate fashion by transfer of molecular properties from small model systems to larger molecules of interest. A general algorithm relying on the transfer in Cartesian coordinates instead of internal ones has been developed by Bouř and co-workers,[52–54] and refined for arbitrary polypeptides in solution by Choi *et al.*[55–57] This scheme may also be applied to the property tensors determining the spectral intensities, and has accordingly been applied to infrared, Raman, VCD and ROA spectroscopy.[52]

3.4 Algorithmic program design

In order to make mode-tracking applicable to actual scientific problems, an efficient and reliable implementation of the algorithm is required. This is provided by the program AKIRA,[8,58] which is designed to calculate the entries in the $\boldsymbol{\Sigma}^{(i)}$ matrix (see eqn (16)) in a semi-numerical fashion, *i.e.*, as numerical derivatives of analytical gradients,

$$\Sigma_{Ij}^{(i)} = \frac{\partial}{\partial \boldsymbol{b}^j}\left(\frac{\partial E_{\mathrm{e},i}}{\partial R_I}\right)_{(\boldsymbol{R}_{eq}^{\mathrm{n.m.w.}}=0)} = \frac{\left(\frac{\partial E_{\mathrm{e},i}}{\partial R_I}\right)_{\left(\boldsymbol{R}_{eq}^{\mathrm{n.m.w.}}+s_R\Delta\boldsymbol{R}_j^{\mathrm{n.m.w.,\,norm}}\right)} - \left(\frac{\partial E_{\mathrm{e},i}}{\partial R_I}\right)_{\left(\boldsymbol{R}_{eq}^{\mathrm{n.m.w.}}-s_R\Delta\boldsymbol{R}_j^{\mathrm{n.m.w.,\,norm}}\right)}}{2s_{bj}\left|\Delta\boldsymbol{b}_j^{\mathrm{norm}}\right|}, \quad (18)$$

where s_R is the step-size in terms of the normalized (non-mass-weighted) Cartesian displacement vectors $\Delta\boldsymbol{R}_j^{\mathrm{n.m.w.,norm}}$ corresponding to the normalized steps $\Delta\boldsymbol{b}_j^{\mathrm{norm}}$ along the mass-weighted basis vectors, and s_{bj} is a dimensionless basis-vector specific step size equal to

$$s_{bj} = s_R\left[\sum_{I=1}^{3M}(B_{Ij}^{(i)})^2/m_I\right]^{-1/2}\left[\frac{[\text{unit of length}]}{[\text{unit of mass}]^{1/2}}\right], \quad (19)$$

with m_I denoting the mass of the atom with which coordinate I is associated. For a flow-chart of mode-tracking calculations with AKIRA, see Fig. 8.

AKIRA is a meta-program which may be connected to any quantum chemistry program package providing energy gradients. At the moment, interfaces to TURBOMOLE,[59] DALTON,[60]

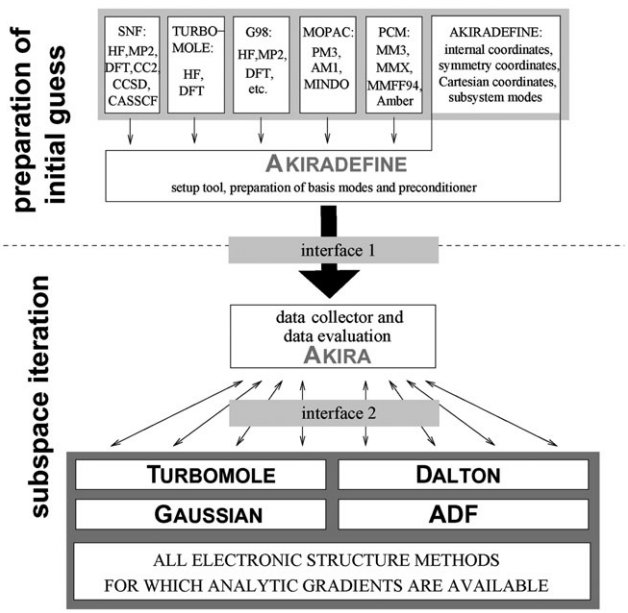


Fig. 8 Hierarchical structure of the mode-tracking program AKIRA;[8] http://www.theochem.ethz.ch/software/akira.

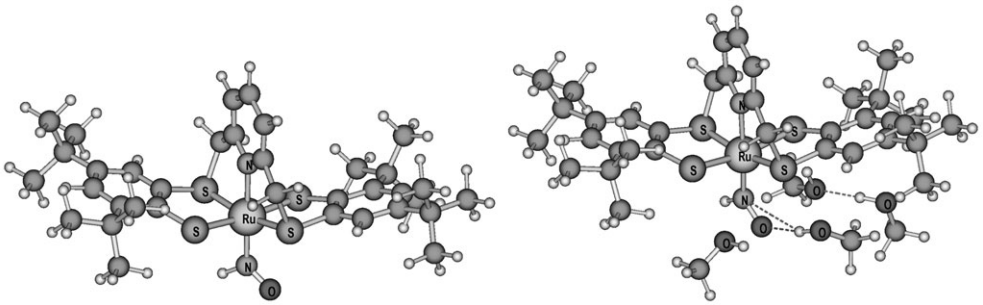


Fig. 9 Optimized structures of the $[Ru(NOH)(`py^{R}S_4')]^+$ cation in its isolated form (left) and with four methanol molecules attached (DFT–BP86 density functional, TZVP basis set).

ADF,[61,62] and GAUSSIAN[63] are available for the raw data calculation (see Fig. 8).

In the context of mode-tracking, the term "raw data" denotes the molecular energy gradients. The advantages of a semi-numerical implementation have been discussed in ref. 22 for the SNF program package, which is designed for full vibrational calculations. In particular, raw data from almost every quantum chemical program—and thus from any electronic structure method for which analytical gradients (for geometric structure optimizations) are implemented—can be processed, and easy and efficient restart facilities are available. Furthermore, coarse- and fine-grained parallelization in computer clusters is easily possible and available, and the numerical error may be controlled *via* the number of grid points used for the numerical differentiation. As a consequence, various existing program packages can be docked and driven, and calculations are possible for any molecular size for which a (partial) structure optimization is feasible.

The decisive initial guess modes for the mode-tracking algorithm may be obtained in various ways, *e.g.*, by specifying certain internal coordinates like the stretching of a bond, by starting with all possible displacements of a certain atom along Cartesian coordinates, or by performing a full frequency calculation with a method of low computational cost (and thus of low accuracy). The Hessian obtained from such a calculation may also be employed as a guess for the exact Hessian for preconditioning (see ref. 29 for more details on this issue, as well as for a discussion on how to define initial guess normal modes).

4 Tracking the NO stretching mode in [Ru]–NO and [Ru]–HNO complexes

As a first example for applications of the mode-tracking algorithm, nitroxyl (HNO) complexes shall be studied. These compounds have been of interest for a long time due to their role as a reactive intermediate in various reactions, among others in the biological reduction of nitrate to N_2 and ammonia (see introduction of ref. 64 for further examples and for references). However, HNO complexes are difficult to isolate and characterize unambiguously. $[Ru(NOH)(`py^{R}S_4')]Br$ ($`py^{R}S_4'$ = 2,6-bis(2-mercapto-3,5-di-*tert*-butylphenylthio) dimethylpyridine(2−); see Fig. 9) was the first hydride reduction product of a NO complex ever characterized.[64]

Since the NO stretching mode is very sensitive to structural changes, vibrational spectroscopy is a convenient tool for monitoring the reduction of NO complexes to their HNO analoga. The wavenumbers associated with the N–O stretching mode may range from 1410 to 1901 cm^{-1} (see, *e.g.*, the discussion in ref. 64 as well as the results reported in ref. 65). Therefore, a comparison to calculated wavenumbers is often needed in order to identify this vibrational mode unambiguously in a measured vibrational spectrum. The prediction of the NO stretching wavenumber is a task optimally suited for mode-tracking because of the local nature of the vibrational mode which is sought for. Mode-tracking will first of all lead to a reduction of the computer time required, but also has the additional advantage of directly providing the NO stretching mode, which may otherwise be cumbersome to find among the results of a full vibrational analysis within the broad frequency range of interest.

Mode-tracking calculations were carried out (after structure optimization) with the AKIRA program[58] for the isolated $[Ru–HNO]^+$ and $[Ru–NO]^+$ complexes as well as for a $[Ru–HNO]^+$ complex with four methanol molecules attached in order to mimic solvent effects on the molecular vibrations (the reduction of $[Ru–NO]^+$ to $[Ru–HNO]^+$ was carried out in methanol). As an initial guess, pure NO bond stretching modes defined *via* internal coordinates were employed, using a convergence criterion of 5.0×1.0^{-4} a.u. as the maximum component of the residuum vector and maximum overlap with the initial guess vector for root-homing. The mode-tracking algorithm on the two isolated systems converged within three iterations. This corresponds to six single-point calculations only per mode-tracking calculation, two at each iteration for the numerical evaluation of energy derivative w.r.t. the new basis vector introduced in the iteration (provided a three-point formula is used for numerical differentiation). For a full numerical vibrational analysis using a three-point formula, $91 \times 6 = 336$ and $90 \times 6 = 330$ single-point calculations would have been required (or 168 and 165 coupled-perturbed Kohn–Sham calculations in an analytical frequency analysis) so that mode-tracking provides a reduction of CPU time to less than 2%. The wavenumbers of the NO stretching modes were predicted to be 1510 and 1943 cm^{-1}, respectively. A full vibrational calculation on the $[Ru–HNO]^+$ system yielded a wavenumber of 1509 cm^{-1} for the NO stretching mode in excellent agreement with the mode-tracking results. The corresponding vibrational wavenumber shifts measured in THF

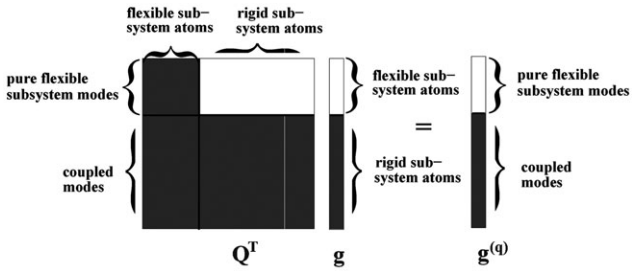


Fig. 10 Schematic representation of the transformation of the (mass-weighted) Cartesian gradient vector $\boldsymbol{g}$ from a Cartesian to a normal-mode basis ($\boldsymbol{g}^{(q)}$) by the transpose of the (mass-weighted) normal-mode matrix $\boldsymbol{Q}$. The white areas denote vanishing vector and matrix entries, the grey areas non-vanishing ones. Whether the "coupled modes" are located on the rigid subsystem (in which case the lower left part of $\boldsymbol{Q}$ would also be white) or delocalized, does not affect our reasoning.[30]

solution were reported as 1378 and 1877 cm^{-1} in ref. 64 and thus considerably lower than the ones calculated for the isolated complexes. Accordingly, for the $[Ru–HNO]^+$ complex with four methanol molecules, the mode-tracking algorithm converged within 25 iterations to a N–O stretching mode with a wavenumber of 1371 cm^{-1}. These calculated data are sufficient for an assignment of the NO stretching band in the measured spectra.

5 Flexible subsystems in rigid environments: direct targeting of subsystem vibrations

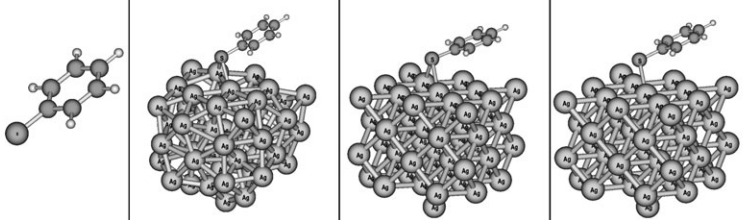
Mode-tracking is also well suited for selectively calculating the vibrational frequencies and normal modes of a flexible subsystem which is attached to an environment that may be considered as rigid. For such systems, there will be many vibrations which involve mainly the atoms of the flexible subsystem, so that *all* calculated normal modes of the free subsystem provide a high-quality initial guess for the composite system.

One important example is an adsorbate molecule attached to a metal cluster. We investigated[30] whether the mode-tracking protocol can be selectively applied to the adsorbate vibrations only, whether it reproduces the results of a complete calculation of all vibrational modes, and how vibrational calculations for partially optimized structures compare to those for fully optimized structures. Three variants of thiophenolate attached to a Ag_{51} cluster were studied:

- a fully optimized thiophenolate–Ag_{51} cluster (structure **1** in Fig. 11), for which the harmonic approximation is valid,
- a partially relaxed adsorbate-metal cluster system where the silver cluster was frozen during the structure optimization of the adsorbate (structure **2** in Fig. 11), and
- a thiophenolate anion which was optimized in its isolated form and placed on an unrelaxed Ag(111) surface fragment without further optimization (structure **3** in Fig. 11).

The frozen clusters in structures **2** and **3** were constructed as a fragment of an Ag(111) surface using the experimental parameters of elemental silver. The metal fragment in the completely relaxed model **1** lost its regular structure after optimization. The three structures do not differ much with respect to the orientation of the adsorbate on the surface. Potential differences in their vibrational spectra will therefore most likely be explained by the influence of the structural changes in the metal cluster (when comparing **1** and **2**) or by the nonzero adsorbate gradient components (when comparing **2** and **3**).

In principle, the calculation of vibrational frequencies and normal modes within the harmonic approximation requires that the molecule under study is in an energy minimum with respect to all nuclear degrees of freedom (see eqn (5)). However, a molecular structure optimization of the whole system might lead to a metal cluster structure which deviates much from the one of an extended crystal surface. A partial optimization of the molecular structure reproduces the experimental arrangement of the atoms better than a fully relaxed surface. It can be shown that the harmonic approximation is justified for



	free	fully optimized **1**		partially optimized **2**	non-optimized **3**
mode	full	full	mode-tracking	mode-tracking	mode-tracking
1	160	241	241	225	215
11	802	858	858	857	857
27	3046	3100	3102	3099	3091

Fig. 11 Selected vibrational wavenumbers in units of cm^{-1} for thiophenolate in its free form and attached to an Ag_{51} cluster, calculated by a full vibrational analysis or by mode-tracking. Structure optimizations carried out with DFT–BP86/RI, SV(P) (Ag)/TZVP (else); see ref. 30 for details.

partly optimized structures if only those modes are interpreted which are centered on the optimized molecular subsystem.[30]

This is because for those normal modes which involve solely motions of the relaxed subsystem, only the gradient components involving coordinates of the subsystem atoms need to be zero in order to preserve the validity of the harmonic approximation. Note that in ref. 30, in order to point out the connection to the actual motions of the atoms in Cartesian space, the non-mass-weighted normal-mode matrix was employed in the equation illustrated by Fig. 10. From a formal point of view, however, it is more intuitive to use the mass-weighted normal-mode matrix, which has therefore been done in this review. In practical applications, *e.g.* in the case of an adsorbed molecule on a metal cluster under study here, normal modes which only involve the optimized adsorbate do not necessarily exist. If the coupling of the adsorbate vibrations to the surface is weak, however, the considerations made above will nonetheless be of practical relevance. Indeed, except for the five lowest thiophenolate modes, the block of the matrix shown in white in Fig. 10 contained only elements close to zero for the actually calculated transpose of the normal mode matrix, $\boldsymbol{Q}^{\mathrm{T}}$, so that the 25 highest-wavenumber vibrational modes of the adsorbed thiophenolate can be interpreted within the harmonic approximation in all three cases.

As an initial guess for the mode-tracking algorithm, the 36 vibrational, rotational and translational normal modes obtained with the SNF vibrational spectroscopy program[22] for the small free (isolated) thiophenolate anion were chosen. The vibrational frequencies of four of the 30 highest-frequency vibrational modes, which are mainly located on the thiophenolate adsorbate, are reported in Fig. 11 and compared to a full frequency calculation for the fully relaxed system. As a reference, the vibrational frequencies of free thiophenolate are also given.

The wavenumber shifts with respect to the free thiophenolate were within 20 to 50 cm^{-1} for 12 of the 30 modes (where the partially optimized structure **2** was taken for comparison). The mode-tracking vibrational analysis of **1** (column 4 in the table in Fig. 11) yielded frequencies very close to those of the full calculation of the vibrational spectrum (column 3) for most normal modes. The difference usually did not exceed 3 cm^{-1}. Modes 11 and 27 are excellent examples for this and illustrate in addition that the differences between the three different structures were far below the wavenumber shifts of the thiophenolate modes induced by the presence of the silver cluster. At very low wavenumbers, the influence of the model (**1**, **2** or **3**) was larger, but still sufficiently below the induced wavenumber shifts.

The vibrational analysis of all three models yielded very similar normal modes. Furthermore, the vibrational normal modes of the thiophenolate, except for modes 7, 11 and 28, were not changed qualitatively by the adsorption. As an example, the modes 12 and 27, whose frequencies changed by more than 50 cm^{-1} after adsorption when taking the partially optimized structure **2** for comparison, are depicted in Fig. 12 for the free thiophenolate as well as for structure **2**.

To summarize, our results suggested that for the surface–adsorbate cluster study[30] reviewed here, with a given adsorbate binding site and structure, the structure of the surface did not necessarily have to reproduce the experimental one exactly

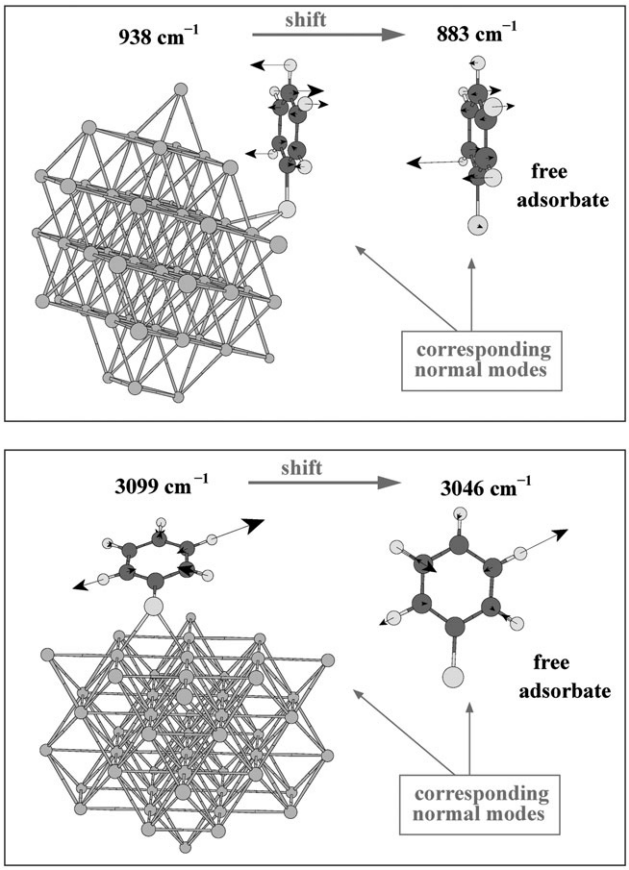


Fig. 12 Normal modes 12 and 27 of the free thiophenolate and of the partially optimized structure **2** (for more details, see also ref. 30).

in order to describe the effects of adsorption on a metal surface within the harmonic approximation. Furthermore, the error introduced by the lack of structure optimization of the adsorbate in structure **3** was not crucial in this case.

In all three cases, the use of the mode-tracking algorithm led to a significant reduction of the computational cost.[30] For example, convergence was achieved within three iterations with 78 single-point calculations for the partially optimized structure **2**, and within one iteration, which corresponds to 72 single-point calculations, for the unrelaxed structure **3**. Convergence within one iteration as in the latter case indicates that the guess normal modes of the free subsystem are already very close to the exact ones.

6 Low-frequency helical breathing modes in decaalanine

While the examples discussed so far referred to a local stretching vibration and in the second case to a complete set of all adsorbate vibrations, we now consider an example of a vibrational Raman optical activity (ROA) study where we used mode-tracking also as a *localization* technique.[31] ROA is the chiral analogue to Raman spectroscopy.[66] Instead of an unpolarized incident light beam, right and left circularly

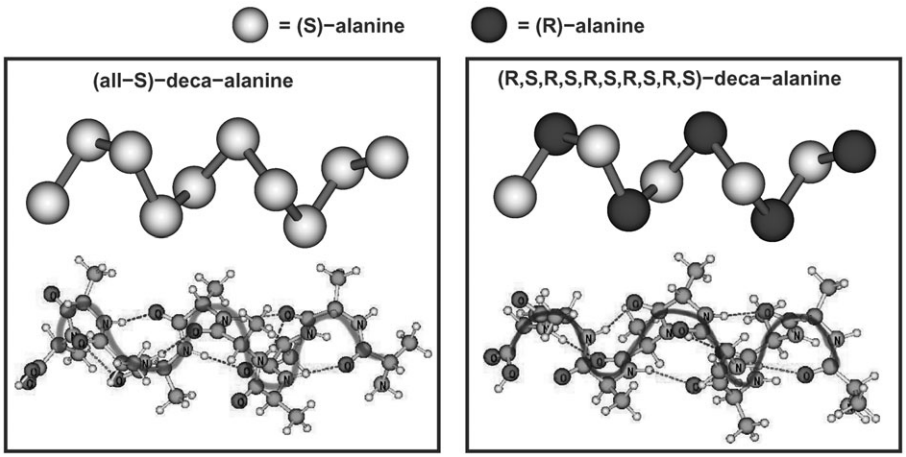


Fig. 13 Symbolical and ball-and-stick representation of both diastereomers of decaalanine (see ref. 31 for details).

polarized light is used, and the scattering intensity *differences* between both are plotted in the spectra. For achiral molecules, these differences are zero, but for chiral molecules, they appear as positive or negative peaks in the spectrum. ROA is especially valuable for the investigation of peptides, since as a Raman technique, it allows for measurements in the natural environment of most peptides, *i.e.*, in aqueous media. Furthermore, from experimental evidence it is known that ROA is especially sensitive to the peptide secondary structure. In order to test the limits of this sensitivity, we investigated whether the ROA patterns associated with a certain backbone conformation are retained regardless of the local chirality of the individual amino acids.[31] This comparison was carried out for two right-handed helical decaalanines with different local configurations of the amino acid building blocks. The helices under study were (all-*S*)-decaalanine (see Fig. 13; the label *S* is chosen according to the Cahn–Ingold–Prelog notation and corresponds to the biologically relevant L configuration) and (*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*)-decaalanine.

The calculation of ROA spectra of molecules with 50 to 100 atoms or more is still a computational challenge. Decaalanine, which consists of 103 atoms, is the largest molecule for which quantum chemical ROA studies have been carried out at present. A calculation of an extended wavenumber range pushes current computational resources to their limit, so that massive parallelization was required to calculate the necessary property tensors and their derivatives. For the calculation of vibrational normal modes and frequencies, second derivatives of the electronic energy w.r.t. nuclear coordinates are needed, and for the ROA intensities, first derivatives w.r.t. nuclear coordinates of three generalized molecular polarizability tensors are required.[67–71] These are the electric dipole–electric dipole polarizability (which is also needed for the calculation of conventional Raman spectra), the electric dipole–magnetic dipole polarizability, and the electric dipole–electric quadrupole polarizability.

Just as the amide modes, for which calculated ROA intensities have been reported in ref. 31, the low-frequency breathing modes of the helical (all-*S*) and (*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*) diastereomers of decaalanine were expected to reflect the global chirality of the two right-handed helices rather than the local chirality of the individual amino acids. As can be seen in Fig. 14, these modes solely consist of helix backbone motions. Breathing modes of helical peptides have not been recorded (or rather assigned) experimentally yet because of their low wavenumber.

Note that these lower-frequency regions of peptides largely couple because of near-degeneracies of vibrational states and are therefore delicate motions not fully accessible within the harmonic approximation. Nevertheless it was interesting to investigate a normal coordinate which describes a breathing mode. These modes were not obtained by a standard full vibrational analysis, from which one would usually get rather unstructured modes in this frequency region, but were calculated using the mode-tracking technique,[8,29] which works in this case as a localization technique on a comparatively pure breathing mode due to the suitably set up guess vibration. The relation between the breathing modes obtained by mode-tracking and the normal modes calculated in the standard way subject to a mixture with (near-) degenerate vibrations can be checked *via* evaluation of the square of the scalar product of the corresponding mass-weighted normal modes. We found, for instance, in the case of (all-*S*)-decaalanine, that the breathing mode contributed to 49.5% to a vibration found at lower wavenumber in a full vibrational analysis, whereas this was reduced to 28.3% in the corresponding (*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*) case.[31]

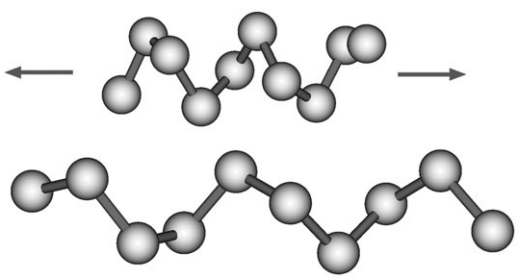


Fig. 14 Schematic representation of the helical breathing mode of (all-*S*)-decaalanine (see ref. 31 for more details).

Indeed, the two breathing modes, which are predicted to show up at 65 and 120 cm^{-1} for the (all-*S*) and the (*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*,*R*,*S*), configurations, respectively, show backscattering intensity differences of 1.31×10^{-2} and 3.99×10^{-2} $Å^4$ a.m.u.$^{-1}$, respectively, at an excitation wavelength of 488.88 nm. These intensity differences are small compared to those of the amide modes, because the Raman scattering intensity of this mode is small. If circular intensity differences (CIDs) $(I^R - I^L)/(I^R + I^L)$ are calculated, however, the CIDs of the breathing modes (4.32×10^{-3} and 2.55×10^{-3}) would be comparable in magnitude to those of the amide modes (ranging from -5.87×10^{-3} to 4.40×10^{-3} and from -9.33×10^{-3} to 7.78×10^{-3}, respectively). Despite the different individual amino acid configurations, the intensity differences of both modes have the same sign, and magnitudes that differ not very much from each other in view of the range of -22.47×10^{-2} to 54.30×10^{-2} covered by the intensity differences associated with the amide vibrational modes. The breathing mode ROA signal is therefore apparently dominated by the axial chirality of the helix.

7 Perspective and outlook

Problems which may be solved by investigating the properties of certain subsystems of molecules or molecular aggregates occur quite frequently in chemistry.[72,73] Such subsystems may be, *e.g.*, solute molecules, the central region of a metal complex with bulky ligands, adsorbates on clusters or surfaces, or the active centers of enzymes. The selective calculation of information on the relevant subsystem only instead of all information on the whole system poses a challenge for quantum chemistry. Nonetheless, developing algorithms for such selective calculations is more than worth the effort, since—given a system of certain size—this may reduce the computer-time considerably, and for given computational resources and computer time, it may make much larger systems accessible to calculations. Such selective procedures are also appealing from an intellectual point of view: while the conventional procedure is first calculating all data and then selecting an important subset of data, an inverse quantum chemistry approach like mode-tracking reverses this order. The mode-tracking algorithm[8] is an excellent means to deal with molecular subsystems in the context of vibrational spectroscopy—and in general cases whenever the calculation of a desired property can be mapped onto an eigenvalue problem. In the future, the range of applications will be extended, among others, to the binding of substrates to the active sites of enzymes. A new interface has now been established which connects the mode-tracking program AKIRA to the Quantum Mechanics/Molecular Mechanics (QM/MM) module of the ADF[61,62] quantum chemistry program package. It will be used to investigate the influence of the myoglobin enzyme surroundings onto carbon monoxide binding to its iron-porphyrin active site.[39] Furthermore, the principle of mode-tracking may be applied to other properties whose calculation requires a matrix diagonalization as mentioned above.

Acknowledgements

C. H. gratefully acknowledges funding by a Chemiefonds-Doktorandenstipendium of the Fonds der Chemischen Industrie (FCI), J. N. by a Liebig-Habilitationsstipendium of the FCI. This work was supported by the Swiss National Science Foundation (project 200021-113479).

References

1 N. Laurendeau, *Statistical Thermodynamics: Fundamentals and Applications*, Cambridge University Press, Cambridge, 2005.
2 F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, New York, 1998.
3 C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons, New York, 2nd edn, 2004.
4 O. A. von Lilienfeld, R. D. Lins and U. Rothlisberger, *Phys. Rev. Lett.*, 2005, **95**, 153002.
5 J. Wang, J. Chen and R. M. Hochstrasser, *J. Phys. Chem. B*, 2006, **110**, 7545–7555.
6 L.-W. Wang and A. Zunger, *J. Chem. Phys.*, 1994, **100**, 2394–2397.
7 A. Franceschetti and A. Zunger, *Nature*, 1999, **402**, 60–63.
8 M. Reiher and J. Neugebauer, *J. Chem. Phys.*, 2003, **118**, 1634–1641.
9 C. Herrmann and M. Reiher, Atomistic Approaches in Modern Biology. From Quantum Chemistry to Molecular Simulations, *Top. Curr. Chem.*, 2007, **268**, 85–132.
10 C. E. Immoos, F. Sulc, P. J. Farmer, K. Czarnecki, D. F. Bocian, A. Levina, J. B. Aitken, R. S. Armstrong and P. A. Lay, *J. Am. Chem. Soc.*, 2005, **127**, 814–815.
11 M. T. Kieber-Emmons, J. Annaraj, M. S. Seo, K. M. V. Heuvelen, T. Tosha, T. Kitagawa, T. C. Brunold, W. Nam and C. G. Riordan, *J. Am. Chem. Soc.*, 2006, **128**, 14230–14231.
12 T. Kitagawa, A. Dey, P. Lugo-Mas, J. B. Benedict, W. Kaminsky, E. Solomon and J. A. Kovacs, *J. Am. Chem. Soc.*, 2006, **128**, 14448–14449.
13 M. Born and R. Oppenheimer, *Ann. Phys.*, 1927, **84**, 457.
14 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons, Chichester, UK, 2000.
15 P. Hohenberg and W. Kohn, *Phys. Rev. Sect. B*, 1964, **136**, 864–871.
16 W. Kohn and L. J. Sham, *Phys. Rev. Sect. A*, 1965, **140**, 1133–1138.
17 R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, International Series of Monographs on Chemistry, Oxford Science Publications, New York, 1989, vol. 16.
18 H. Eschrig, *The Fundamentals of Density Functional Theory*, Edition am Gutenbergplatz, Leipzig, 2003.
19 R. M. Dreizler and E. K. U. Gross, *Density Functional Theory. An Approach to the Quantum Many-Body Problem*, Springer-Verlag, Berlin, 1990.
20 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press USA, New York, 1989.
21 D. Marx and J. Hutter, in *Modern Methods and Algorithms of Quantum Chemistry, Proceedings, NIC Series*, ed. J. Grotendorst, John von Neumann Institute for Computing, Jülich, 2000, vol. 1, pp. 329–477, see also www.fz-juelich.de/nic-series/Volume1.
22 J. Neugebauer, M. Reiher, C. Kind and B. A. Hess, *J. Comput. Chem.*, 2002, **23**, 895–910.
23 E. B. Wilson, Jr, J. C. Decius and P. C. Cross, *Molecular Vibrations*, Dover, New York, 1955.
24 P. W. Atkins and J. de Paula, *Atkins' Physical Chemistry*, Oxford University Press, Oxford, 7th edn, 2001.
25 J. Neugebauer and B. A. Hess, *J. Chem. Phys.*, 2003, **118**, 7215–7225.
26 G. Moritz and M. Reiher, 2003, unpublished results.
27 J. Neugebauer and M. Reiher, *J. Phys. Chem. A*, 2004, **108**, 2053–2061.
28 J. Neugebauer and M. Reiher, *J. Comput. Chem.*, 2004, **25**, 587–597.
29 M. Reiher and J. Neugebauer, *Phys. Chem. Chem. Phys.*, 2004, **6**, 4621–4629.
30 C. Herrmann and M. Reiher, *Surf. Sci.*, 2006, **9**, 1891–1900.

31 C. Herrmann, K. Ruud and M. Reiher, *ChemPhysChem*, 2006, **7**, 2189–2196.
32 (*a*) T. B. Adler, N. Borho, M. Reiher and M. A. Suhm, *Angew. Chem., Int. Ed.*, 2006, **45**, 3440–3445; (*b*) T. B. Adler, N. Borho, M. Reiher and M. A. Suhm, *Angew. Chem.*, 2006, **118**, 3518–3523.
33 A. L. Kaledin, *J. Chem. Phys.*, 2005, **122**, 184106.
34 M. Reiher and J. Neugebauer, *J. Chem. Phys.*, 2005, **123**, 117101.
35 A. L. Kaledin, M. Kaledin and J. M. Bowman, *J. Chem. Theory Comput.*, 2006, **2**, 166–174.
36 M. Reiher, J. Neugebauer and B. A. Hess, *Z. Phys. Chem.*, 2003, **217**, 91–103.
37 M. Helmreich, F. Hampel, N. Jux and M. Reiher, unpublished results.
38 B. Kirchner and M. Reiher, in *Analytical Methods in Supramolecular Chemistry*, ed. C. Schalley, Wiley-VCH, Weinheim, 2007.
39 C. Herrmann, J. Neugebauer and M. Reiher, in preparation.
40 F. Filippone and M. Parrinello, *Chem. Phys. Lett.*, 2001, **345**, 179–182.
41 F. Filippone, S. Meloni and M. Parrinello, *J. Chem. Phys.*, 2001, **115**, 636–642.
42 B. R. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. USA*, 1985, **82**, 4995–4999.
43 B. R. Brooks, D. Janežič and M. Karplus, *J. Comput. Chem.*, 1995, **16**, 1522–1542.
44 D. Neuhauser, *J. Chem. Phys.*, 1990, **93**, 2611–2616.
45 D. A. McCormack, G.-J. Kroes and D. Neuhauser, *J. Chem. Phys.*, 1998, **109**, 5177–5186.
46 R. Santra, J. Breidbach, J. Zobeley and L. S. Cederbaum, *J. Chem. Phys.*, 2000, **112**, 9243–9252.
47 S.-W. Huang and J. Tucker Carrington, *J. Chem. Phys.*, 2000, **112**, 8765–8771.
48 V. A. Mandelshtam and H. S. Taylor, *J. Chem. Phys.*, 1995, **102**, 7390–7399.
49 Y. Huang, W. Zhu, D. J. Kouri and D. K. Hoffman, *Chem. Phys. Lett.*, 1993, **206**, 96–102.
50 J. D. Head, *Int. J. Quantum Chem.*, 1997, **65**, 827.
51 N. A. Besley and K. A. Metcalf, *J. Chem. Phys.*, 2007, **126**, 035101.
52 P. Bouř, J. Sopková, L. Bednárová, P. Maloň and T. A. Keiderling, *J. Comput. Chem.*, 1997, **18**, 646–659.
53 P. Bouř, J. Kubelka and T. A. Keiderling, *Biopolymers*, 2000, **53**, 380–395.
54 P. Bouř, J. Kubelka and T. A. Keiderling, *Biopolymers*, 2002, **65**, 45–59.
55 S. Ham, S. Cha, J.-H. Choi and M. Cho, *J. Chem. Phys.*, 2003, **119**, 1451.
56 J.-H. Choi, S. Ham and M. Cho, *J. Phys. Chem. B*, 2003, **107**, 9132–9138.
57 J.-H. Choi and M. Cho, *J. Chem. Phys.*, 2004, **120**, 4383–4392.
58 J. Neugebauer, C. Herrmann and M. Reiher, http://www.theochem.ethz.ch/software/akira.
59 R. Ahlrichs *et al.*, http//www.cosmologic.de/turbomole.html.
60 DALTON, a molecular electronic structure program, release 2.0 2005, see http://www.kjemi.uio.no/software/dalton/dalton.html.
61 Amsterdam Density Functional Program; Theoretical Chemistry, Vrije Universiteit Amsterdam. http://www.scm.com.
62 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. J. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.
63 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, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision C.02)*, Gaussian, Inc., Wallingford, CT, 2004.
64 D. Sellmann, T. Gottschalk-Gaudig, D. Häussinger, F. W. Heinemann and B. A. Hess, *Chem. Eur. J.*, 2001, **7**, 2099–2103.
65 D. Sellmann, N. Blum, F. W. Heinemann and B. A. Hess, *Chem. Eur. J.*, 2001, **7**, 1874–1880.
66 L. D. Barron, *Molecular Light Scattering and Raman Optical Activity*, Cambridge University Press, Cambridge, 2nd edn, 2004.
67 P. W. Atkins and L. D. Barron, *Mol. Phys.*, 1969, **16**, 453–466.
68 A. D. Buckingham and M. B. Dunn, *J. Chem. Soc. A*, 1971, 1988–1991.
69 L. D. Barron and A. D. Buckingham, *Mol. Phys.*, 1971, **20**, 1111–1119.
70 L. D. Barron, M. P. Boogard and A. D. Buckingham, *J. Am. Chem. Soc.*, 1973, **95**, 603–605.
71 T. Helgaker, K. Ruud, K. L. Bak, P. Jørgensen and J. Olsen, *Faraday Discuss.*, 1994, **99**, 165.
72 M. Reiher, *Found. Chem.*, 2003, **5**, 23–41.
73 M. Reiher, *Found. Chem.*, 2003, **5**, 147–163.