

Dissertation zur Erlangung des Doktorgrades
der Fakultät Chemie und Pharmazie
der Ludwig–Maximilians–Universität München

Development and Application
of Efficient *Ab Initio*
Molecular Dynamics Simulations
of Ground and Excited States

Laurens Dirk Marga Peters

aus

München

2019

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Christian Ochsenfeld betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 3.4.2020

(Laurens Peters)

Dissertation eingereicht am: 13.12.2019

1. Gutachter: Prof. Dr. Christian Ochsenfeld

2. Gutachterin: Prof. Dr. Regina de Vivie-Riedle

Mündliche Prüfung am: 12.2.2020

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Dissertation unterstützt und motiviert haben.

Zuerst möchte ich **Prof. Dr. Christian Ochsenfeld** danken, welcher mir die Möglichkeit gegeben hat meine Doktorarbeit in seinem Arbeitskreis anzufertigen und die Fragestellung dieser Arbeit ausgearbeitet hat. Für die Unterstützung und die konstruktiven Besprechungen und Ratschläge möchte ich mich herzlich bedanken.

Des Weiteren möchte ich mich bei **Prof. Dr. Regina de Vivie-Riedle** für die Anfertigung des Zweitgutachtens bedanken.

Ein besonderer Dank gilt meinem Betreuer **Dr. Jörg Kussmann** für die hilfreichen Diskussionen und für das Feedback und die Hilfe bei der Programmierarbeit.

Zudem möchte ich mich bei meinen Kooperationspartnern **Johannes Dietschreit** und **Beatriz von der Esch** für die schöne und spannende Zusammenarbeit bedanken. Dank geht auch an **Sophia Peters** für die experimentellen IR-Spektren und **Thomas Schnappinger** für die hilfreichen Diskussionen.

Des Weiteren möchte ich mich bei dem kompletten **AK Ochsenfeld** für die schöne Zeit und das gute Arbeitsklima bedanken. Besonders hervorheben möchte ich dabei die IT-Sessions mit **Phillip Taenzler** und **Gökçen Savaşçi** sowie die Lauftreffs mit unter anderem **Sigurd Vogler**, **Matthias Beuerle**, **Henryk Laqua** und **Travis Thompson**.

Zu guter Letzt möchte ich mich bei meinen **Freunden**, meiner **Familie** und vor allem bei **Sophia Peters** für das Verständnis, den Rückhalt und die Unterstützung über die gesamte Dauer meines Studiums und der Promotion bedanken.

Summary

Ab initio molecular dynamics reflect the movement of nuclei on a potential energy surface generated by *ab initio* methods. These simulations give access to an entire series of chemically relevant properties, such as vibrational spectra and free energies, and have thus become indispensable in, for example, biochemistry and materials sciences. They are, however, computationally demanding, due to the expensive quantum-chemical calculations that are required at every step. In order to overcome some of the limitations, this thesis presents steps towards efficient but still accurate *ab initio* molecular dynamics simulations, combining recent progress in different fields of computational chemistry. The time-consuming two-electron integral evaluations are conducted on graphics processing units. Their massively parallel architecture leads to speed-ups (with respect to calculations on central processing units) and strong scaling is observed. Expensive electronic structure calculations are circumvented using parametrized methods, such as the corrected small basis set Hartree-Fock method or simplified time-dependent density functional theory. From the field of molecular dynamics, the extended Lagrangian method is adopted to stabilize the trajectories and to accelerate the convergence of the self-consistent field algorithm. Finally, couplings between electronic states are approximated from a finite differences approach to avoid the time-consuming analytical evaluations at the time-dependent density functional theory level. As a result of these approaches, large molecular systems become accessible at comparably low computational cost. This is demonstrated for several illustrative applications. Excited-state dynamics are used to explore the relaxation pathway of the rhodopsin protein and four newly designed rotary molecular motors using the same Schiff base motif. Ground-state simulations deliver vibrational spectra of medium-sized molecules and liquid water. They are used in addition to determine free energy differences of molecular transformations, for which a novel scheme is introduced delivering deeper insights into the underlying process.

List of Publications

This is a cumulative dissertation, comprising four articles in peer-reviewed journals (**I**, **II**, **IV**, and **V**) and one manuscript in preparation (**III**). In the following, all articles are listed together with the author's contribution to each of them.

- I** **L. D. M. Peters**, J. Kussmann, C. Ochsenfeld,
"Efficient and Accurate Born-Oppenheimer Molecular Dynamics
for Large Molecular Systems",
J. Chem. Theory Comput., **13**, 5479 (2017).
Contribution by the Author: *Most of the concept and implementation,
all calculations and analyses, as well as writing the manuscript*
- II** **L. D. M. Peters**, J. Kussmann, C. Ochsenfeld,
"Non-adiabatic molecular dynamics on graphics processing units:
performance and application to rotary molecular motors",
J. Chem. Theory Comput., **15**, 6647 (2019).
Contribution by the Author: *Concept, most of the implementation,
all calculations and analyses, as well as writing the manuscript*
- III** **L. D. M. Peters**, J. Kussmann, C. Ochsenfeld,
"Combining Graphics Processing Units, Simplified Time-Dependent Density
Functional Theory, and Finite-Difference Couplings to Accelerate
Non-adiabatic Molecular Dynamics",
in preparation.
Contribution by the Author: *Most of the concept and implementation,
all calculations and analyses, as well as writing the manuscript*
- IV** **L. D. M. Peters**, J. C. B. Dietschreit, J. Kussmann, C. Ochsenfeld,
"Calculating free energies from the vibrational density of states function:
validation and critical assessment",
J. Chem. Phys., **150**, 194111 (2019).
Contribution by the Author: *Conjoint concept with J. C. B. Dietschreit,
most of the implementation, half of the calculations and analyses, as well as
writing of the manuscript*
- V** J. C. B. Dietschreit, **L. D. M. Peters**, J. Kussmann, C. Ochsenfeld,
"Identifying free energy hot-spots in molecular transformations",
J. Phys. Chem. A, **123**, 2163 (2019).
Contribution by the Author: *Conjoint concept with J. C. B. Dietschreit,
parts of the implementation, half of the calculations and analyses, as well as
writing of the manuscript*

Publications unrelated to this thesis:

- VI** A. Hoffmann, J. Stanek, B. Dicke, **L. Peters**, B. Grimm-Lebsanft, A. Wetzels, A. Jesser, M. Bauer, M. Gnida, W. Meyer-Klaucke, M. Rübhausen, S. Herres-Pawlis,
"Implications of Guanidine Substitution on Copper Complexes as Entatic-State Models",
Eur. J. Inorg. Chem., **2016**, 4731 (2016).
- VII** J. Stanek, N. Sackers, F. Fink, M. Paul, **L. Peters**, R. Grunzke, A. Hoffmann, S. Herres-Pawlis,
"Copper Guanidinoquinoline Complexes as Entatic State Models of Electron-Transfer Proteins",
Chem. Eur. J., **23**, 15738 (2017).
- VIII** B. von der Esch, J. C. B. Dietschreit, **L. D. M. Peters**, C. Ochsenfeld,
"Finding reactive configurations: A machine learning approach for estimating transition barriers applied to Sirtuin 5",
J. Chem. Theory Comput., **15**, 6660 (2019).

Contents

1	Introduction	1
2	Theoretical Basis	5
2.1	Notation	5
2.2	<i>Ab Initio</i> Molecular Dynamics (AIMD)	6
2.2.1	The Schrödinger Equation and the Born-Oppenheimer Approximation	6
2.2.2	Ground-State Born-Oppenheimer Molecular Dynamics (BOMD)	7
2.2.3	Non-Adiabatic Molecular Dynamics (NAMD)	8
2.3	Energy, Gradients, and Couplings Calculations	11
2.3.1	The Hartree-Fock (HF) Method	11
2.3.2	Density Functional Theory (DFT)	13
2.3.3	Time-Dependent Density Functional Theory (TDDFT)	14
2.3.4	Gradients and Couplings at the TDDFT Level	18
2.4	Efficient Molecular Dynamics Simulations	22
2.4.1	Extended-Lagrangian BOMD (XL-BOMD)	22
2.4.2	Corrected Small Basis Set HF (HF-3c)	22
2.4.3	Simplified TDDFT Schemes	23
2.4.4	The Hammes-Schiffer-Tully (HST) Model	24
2.4.5	Molecular Dynamics on Graphics Processing Units (GPUs)	25
2.5	Molecular Dynamics Applications	26
2.5.1	Simulation of Vibrational Spectra	26
2.5.2	Calculation of Free Energies	27
2.5.3	Investigation of Relaxation Pathways	30
3	Publications	31
3.1	Publication I	31
3.2	Publication II	79
3.3	Manuscript III	107
3.4	Publication IV	145
3.5	Publication V	173
4	Conclusions and Outlook	191

Chapter 1

Introduction

Verifying, explaining, or predicting experimental observables (e.g., reaction energies, rate constants, or spectra) are the main objectives of modern computational chemistry. Since the early 2000s, molecular dynamics (MD) have become more and more popular for these purposes. They give access to new or more accurate properties than static calculations and are, therefore, a powerful tool when investigating biochemical reactions, light-matter interactions, or processes in liquids or solid-state materials.

MD methods simulate the movement of atoms in molecules, liquids, or gas mixtures. To do this efficiently, the Born-Oppenheimer approximation [1] is applied, separating the motion of electrons and nuclei. The latter are treated as classical particles propagating on the potential energy surface (PES) and obeying Newton's equations of motion. The PES itself is determined by the electronic structure or the applied force field, known as *ab initio* MD (AIMD) or molecular mechanical MD (MMMD), respectively. The choice of the method to calculate the PES heavily depends on the investigated system and the problem at hand. Due to their neglect of electrons and the approximate (classical) treatment of bonds, MMMD can be applied to large molecular systems, such as enzymes (see, e.g., applications of the AMBER [2] program package). However, the quality of the results mainly depends on the parametrization of the force field and whether it is suitable to describe the system and the regarded property [3]. Standard MMMD fails, for example, to describe chemical reactions [4]. In these cases, AIMD [5, 6] is more suitable. Here, one usually refers to the Hartree-Fock (HF) method [7–10] or density functional theory (DFT) [11–13] to solve the electronic Schrödinger equation [14]. Methods including electron correlation (e.g., Møller-Plesset perturbation theory [15] or Coupled Cluster [16, 17]) can also be used.

MD simulations are normally conducted on a single adiabatic PES, resulting in so-called Born-Oppenheimer MD (BOMD). This adiabatic approximation, however, breaks down when multiple electronic states are involved in the described process (e.g., photochemistry [18], dynamics close to conical intersections [19], or radiationless transitions [20]). For these purposes, non-adiabatic MD (NAMD) using, for example, Ehrenfest dynamics [21], *ab initio* multiple spawning [22, 23], or the fewest-switches surface hopping algorithm [24–27] have to be applied. In the latter, the system is able to change the adiabatic PES during one trajectory. As the hops from one state to another are determined stochastically, observables have to be extracted from an ensemble of trajectories. NAMD calculations require an

electronic structure method that delivers excited-state energies and properties (e.g., time-dependent DFT (TDDFT) [28, 29] or the complete active space self-consistent field (CASSCF) method [30]).

MD simulations provide, as aforementioned, access to a huge number of experimentally observable properties. The calculation of free energies [31], for example, requires the sampling of energies along MD trajectories, yielding reaction energies, binding energies, or rate constants. Infrared (IR) and Raman spectra can be determined directly from sampled dipole and quadrupole moments [6, 32–34], respectively. Vibrational averaging allows for a more realistic description of nuclear magnetic resonance (NMR) shifts [35] and hyperfine coupling constants, as well as g-tensor values [36] in comparison to static calculations. MD simulations can even be used to explore reaction paths in so-called nanoreactors [37, 38] or (in case of NAMD) relaxation pathways of excited molecules [39–49]. This broad range of applications explains why MD methods are nowadays applied in nearly all fields of chemistry.

They are, however, computationally more expensive than single point calculations. This is due to the large number of time steps required to sample a sufficiently large region of configuration space, involving (at least) one energy and forces calculation per step. The problem is even more severe for large molecular systems, where not only the cost per step but also the necessary number of steps increases drastically. As a consequence, efficient methods for (1) the simulation and (2) energy and properties calculations are still very much needed to explore larger systems with higher accuracy.

This dissertation is designed as cumulative comprising four publications (**I**, **II**, **IV**, and **V**) and one manuscript (**III**). In these, ways to accelerate AIMD calculations are investigated and applied to chemically relevant systems and problems. The following approaches are used:

1. Approximate electronic structure theories (HF-3c [50] and simplified TDDFT [51, 52]), which deliver good static energies (and some properties) at comparably low computational cost,
2. the extended Lagrangian method [53], which reduces the number of necessary self-consistent field (SCF) cycles in HF or DFT calculations during MD simulations and leads to more stable trajectories,
3. the Hammes–Schiffer–Tully approach [25] to efficiently calculate couplings between electronic states from finite differences, and
4. efficient routines for the rate-determining two-electron integral evaluations on graphics processing units (GPUs) in addition to central processing units (CPUs) using the FermiONs++ program package [54–56].

The latter have, so far, only been applied to ground-state energy and forces calculations. However, encouraging results from other groups indicated that dynamics and excited state calculations can also be accelerated using GPUs [22, 23, 57–64]. Regarding applications:

1. The simulation of IR spectra of medium-sized molecules and liquid water,
2. the relaxation pathways of excited rotary molecular motors and the excited rhodopsin protein, as well as

3. several free energy calculations,

are presented. Additionally, a novel way to map free energy changes on atoms or regions (so-called free energy hot-spots) is introduced that also allows for a detailed analysis of the underlying processes. In Chapter 2 the underlying theory is recapitulated. The publications form the main part of this work and are presented in Chapter 3, followed by a conclusion and outlook in Chapter 4. A brief introduction to the publications is given below.

Publication **I** presents the first AIMD implementation using the FermiONs++ program package and its efficient two-electron integral evaluations on GPUs. To be able to simulate large systems, the extended Lagrangian scheme and the HF-3c method are used. The resulting method is extremely efficient (e.g., MD simulations of β -carotene cost only a few seconds per step using four GPUs) and scales well with the computational resources (strong scaling). Simulated IR spectra of β -carotene, paclitaxel, and liquid water show satisfying agreement with experimental results.

In publication **II**, the scope of applications of the GPU-based integral routines is extended, enabling the calculation of excited-state energies and properties at the TDDFT level of theory. Similar to ground-state calculations, decent speed-ups (with respect to CPU calculations) and strong scaling are observed. The performance and scaling of the method does not only depend on the size of the molecule, but also on the locality of the excitation. With these efficient routines at hand, NAMD simulations using GPUs become accessible. The photo-induced rotations of four newly designed rotary molecular motors are investigated as a first illustrative example.

The performance of the GPU-based NAMD simulations is further improved in manuscript **III**, where simplified TDDFT schemes and the Hammes–Schiffer–Tully model are used in addition. Here, hybrid QM/MM NAMD simulations of the relaxation of rhodopsin are discussed.

Publication **IV** returns to ground-state simulations, focusing on the calculation of free energies and free energy differences. The main purpose of this work is to investigate the use of the density of states integration (DSI) method (first mentioned by Berens *et al.* in 1983 [65]) for free energy calculations applying AIMD simulations. DSI is compared to more established free energy methods, e.g., exponential averaging [66] and Bennett’s acceptance ratio (BAR) [67, 68], using three numerical and two chemical examples. The overall conclusion is that DSI is en par with BAR in many cases, but requires longer simulation times (or even fails), when the systems contain low-frequency or non-harmonic modes.

A unique feature of DSI is its ability to map vibrational free energies to atoms, functional groups, or residues. This is further illustrated in publication **V**, where these so-called free energy hot-spots are investigated for two prototypical examples: the free energy of binding of an inhibitor to a protein and the anomeric effect. Apart from delivering useful insights into the underlying processes, the resulting hot-spots are also in good agreement with chemical intuition.

Chapter 2

Theoretical Basis

This chapter outlines the theory of *ab initio* molecular dynamics (AIMD) (see Section 2.2) and the electronic structure theory necessary to conduct these simulations (see Section 2.3). Moreover, ways to accelerate the calculations (see Section 2.4) and fields of application (see Section 2.5) are presented as well.

2.1 Notation

Throughout this work, i, j, k, \dots denote occupied, a, b, c, \dots virtual, and p, q, r, \dots arbitrary molecular orbitals (φ). μ, ν, κ, \dots will be used for basis functions (χ). The capital letters I, J, \dots represent electronic states, while A, B, \dots stand for the nuclei. \mathbf{R} and \mathbf{r} are the nuclear and electronic coordinates, respectively, and t is the time. For the sake of simplicity, matrices in the atomic orbital (AO) and molecular orbital (MO) basis share the same symbol but with different indices (e.g., $P_{\mu\nu}$ vs. P_{pq}). When referring to equations, figures, and algorithms, the abbreviations eq. (#), fig. (#), and alg. (#) are used, respectively. δ_{pq} is the Kronecker delta. Superscript $*$, T , and \dagger symbolize the complex conjugate, the transpose, and the conjugate transpose of a matrix or vector, respectively, while \Re stands for the real part of a complex number. $\langle \dots | \dots \rangle$ denotes the Dirac notation and $(\dots | \dots)$ the Mulliken notation of the two electron integrals. Superscript x (e.g., \mathbf{h}^x) denotes derivatives with respect to the nuclear coordinates, while superscript (x) (e.g., $\mathbf{h}^{(x)}$) states that integral derivatives are contracted with unperturbed density matrices.

2.2 *Ab Initio* Molecular Dynamics (AIMD)

2.2.1 The Schrödinger Equation and the Born-Oppenheimer Approximation

The time-dependent Schrödinger equation [14]

$$\hat{H}\Psi(\mathbf{R}, \mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{R}, \mathbf{r}, t) , \quad (2.1)$$

is ubiquitous in quantum chemistry. It describes the time-dependent behaviour of a system, represented by the wavefunction $\Psi(\mathbf{R}, \mathbf{r}, t)$ of electrons and nuclei. i is the imaginary unit and \hbar is the reduced Planck constant. The Hamilton operator (\hat{H}) and the total wavefunction ($\Psi(\mathbf{R}, \mathbf{r}, t)$) can be separated into contributions arising from the electrons (el) and/or nuclei (nuc)

$$\hat{H} = \hat{T}_{\text{el}} + \hat{T}_{\text{nuc}} + \hat{V}_{\text{el}} + \hat{V}_{\text{nuc}} + \hat{V}_{\text{el,nuc}} = \hat{H}_{\text{el}} + \hat{T}_{\text{nuc}} , \quad (2.2)$$

$$\Psi(\mathbf{R}, \mathbf{r}, t) = \sum_J \psi_J^{\text{nuc}}(\mathbf{R}, t) \psi_J^{\text{el}}(\mathbf{r}, \bar{\mathbf{R}}) . \quad (2.3)$$

\hat{T} and \hat{V} are operators for kinetic terms and potentials, respectively. Two other assumptions are made here [69]: (1) The total wavefunction is a sum of multiple stationary states J and (2) the electronic wavefunction does not explicitly depend on t . There is, however, an implicit dependence on \mathbf{R} ($\bar{\mathbf{R}}$) and thus on t . The variables of the wavefunctions will be omitted from now on to ease the reading of equations. Inserting eq. (2.3) into eq. (2.1) and multiplying with $\langle \psi_I^{\text{el}} |$ from the left leads to

$$[\hat{T}_{\text{nuc}} + E_I] \psi_I^{\text{nuc}} + \sum_J \hat{C}_{IJ} \psi_J^{\text{nuc}} = i\hbar \frac{\partial}{\partial t} \psi_I^{\text{nuc}} , \quad (2.4)$$

with E_I being a solution of the time-independent electronic Schrödinger equation

$$\hat{H}_{\text{el}} \psi_I^{\text{el}} = E_I \psi_I^{\text{el}} . \quad (2.5)$$

Here, it is assumed that the set of electronic wavefunctions is orthonormal ($\langle \psi_I^{\text{el}} | \psi_J^{\text{el}} \rangle = \delta_{IJ}$). \hat{C}_{IJ} describes the coupling between two states I and J and is defined as

$$\hat{C}_{IJ} = 2 \langle \psi_I^{\text{el}} | \nabla_{\mathbf{A}} \psi_J^{\text{el}} \rangle \nabla_{\mathbf{A}} + \langle \psi_I^{\text{el}} | \nabla_{\mathbf{A}} \nabla_{\mathbf{A}} \psi_J^{\text{el}} \rangle = 2\boldsymbol{\tau}_{IJ} \nabla_{\mathbf{A}} + \langle \psi_I^{\text{el}} | \nabla_{\mathbf{A}} \nabla_{\mathbf{A}} \psi_J^{\text{el}} \rangle . \quad (2.6)$$

$\nabla_{\mathbf{A}}$ is the derivative with respect to \mathbf{R} . $\boldsymbol{\tau}_{IJ}$ is the non-adiabatic coupling vector (NACV). It populates the non-diagonal elements ($I \neq J$) of \hat{C}_{IJ} and mainly depends on the energy difference and the wavefunction overlap of I and J (see fig. (2.1)). The second term of eq. (2.6) is a small scalar contribution defining the diagonal elements ($I = J$) of \hat{C}_{IJ} .

Solving eq. (2.4) is the main objective of AIMD, calculating the time-dependent movement of the atoms in molecular systems. In the next subsections, dynamics at the two special cases of fig. (2.1) will be discussed: (1) Born-Oppenheimer molecular dynamics (BOMD), where the coupling between the electronic state is assumed to be zero ($\hat{C}_{IJ} = 0$) and (2) non-adiabatic molecular dynamics (NAMD), which is able

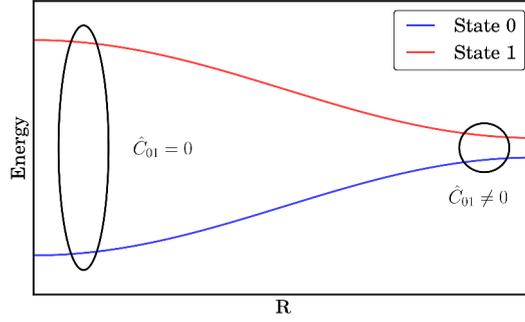


Figure 2.1: Simplified dependence of \hat{C}_{IJ} on the molecular structure (\mathbf{R}): large energy differences lead to small (negligible) couplings, while small energy gaps lead to large couplings.

to describe systems with multiple electronic states ($\hat{C}_{IJ} \neq 0$). The first method is normally applied to ground-state simulations, whereas the second method is used for excited-state processes, such as radiationless transitions.

2.2.2 Ground-State Born-Oppenheimer Molecular Dynamics (BOMD)

In the Born-Oppenheimer approximation [1], $\hat{T}_{\text{nuc}} |\psi_J^{\text{el}}\rangle$, and thus all \hat{C}_{IJ} , are assumed to be zero. This and setting $I = 0$ (ground state) reduces eq. (2.4) to

$$[\hat{T}_{\text{nuc}} + E_0] \psi_0^{\text{nuc}} = i\hbar \frac{\partial}{\partial t} \psi_0^{\text{nuc}}. \quad (2.7)$$

E_0 is electronic energy of the ground state. This equation can be used to calculate the time-dependent change of the expectation value of an arbitrary observable, following the Ehrenfest theorem [70]

$$\frac{d}{dt} \langle \hat{O} \rangle = \frac{d}{dt} \langle \psi_0^{\text{nuc}} | \hat{O} | \psi_0^{\text{nuc}} \rangle = \frac{i}{\hbar} \langle [\hat{T}_{\text{nuc}} + E_0, \hat{O}]_- \rangle + \langle \frac{\partial}{\partial t} \hat{O} \rangle. \quad (2.8)$$

$[\hat{A}, \hat{B}]_-$ is the commutator ($[\hat{A}, \hat{B}]_- = \hat{A}\hat{B} - \hat{B}\hat{A}$). If this is applied to the position ($\hat{\mathbf{R}}$) and momentum ($\hat{\mathbf{p}}$) operator, one obtains

$$\frac{d}{dt} \langle \hat{\mathbf{R}} \rangle = \frac{\langle \hat{\mathbf{p}} \rangle}{m}, \quad (2.9)$$

$$\frac{d}{dt} \langle \hat{\mathbf{p}} \rangle = \langle -\nabla_{\mathbf{A}} E_0 \rangle, \quad (2.10)$$

given the fact that both operators show no explicit time-dependence. m is the mass of the corresponding nucleus. Combining eqs. (2.9) and (2.10) and assuming that the nuclei are classical particles (semi-classical approximation), leads to

$$m \frac{d^2}{dt^2} \mathbf{R} = -\nabla_{\mathbf{A}} E_0. \quad (2.11)$$

Thus, solving eq. (2.7) in the classical limit becomes equivalent to solving Newton's equation of motion. For this purpose, there exists an entire series of propagators,

which are used in BOMD simulations. One of the most prominent ones, the Velocity-Verlet propagator [71, 72], is outlined in alg. (1), which describes a pseudo-code for BOMD simulations.

Algorithm 1 Pseudo-code of Born-Oppenheimer molecular dynamics.

Initial Values for \mathbf{R} and \mathbf{p} at $t = 0$

Initial Calculation of E_0 and $\nabla_A E_0$

while $t < t_{\max}$ **do**

$$(1) \mathbf{p}(t + \frac{\Delta t}{2}) = \mathbf{p}(t) - \frac{\Delta t}{2} \cdot \nabla_A E_0(\mathbf{R}(t))$$

$$(2) \mathbf{R}(t + \Delta t) = \mathbf{R}(t) + \frac{\Delta t}{m} \cdot \mathbf{p}(t + \frac{\Delta t}{2})$$

$$(3) \text{Calculate } E_0(\mathbf{R}(t + \Delta t)), \nabla_A E_0(\mathbf{R}(t + \Delta t))$$

$$(4) \mathbf{p}(t) = \mathbf{p}(t + \frac{\Delta t}{2}) - \frac{\Delta t}{2} \cdot \nabla_A E_0(\mathbf{R}(t + \Delta t))$$

end while

Alg. (1) shows that at each step of the simulation, the electronic energy and its derivative with respect to the nuclear coordinates needs to be calculated "on the fly". The accuracy of the propagator depends on the step size (Δt). For *ab initio* BOMD simulations a Δt between 0.1 fs and 0.5 fs is recommended so that even fast movements (e.g., the O-H stretching vibration) are captured accurately. Applying molecular mechanical (MM) methods to larger systems, one tends to freeze bonds with a high frequency by means of algorithms like SHAKE [73]. This allows for larger time steps. Additionally, one usually applies thermostats and barostats (e.g., see refs. [74] or [75]) to set the mean temperature or pressure of the MD to a specific value.

2.2.3 Non-Adiabatic Molecular Dynamics (NAMD)

Trajectory surface hopping [24, 25] is, similar to BOMD, a semi-classical method, which is applied to cases where the coupling \hat{C}_{IJ} cannot be neglected. The nuclei are treated again as classical particles, which propagate on the potential energy surfaces (PESs). The PESs are given by the adiabatic electronic states I , the electronic states, which are a solution to the electronic time-independent Schrödinger equation (eq. (2.5)). To account for couplings between these states, the system is allowed to change its PES during a single trajectory. This is done stochastically, meaning that the system "hops" from state I to J , when a calculated hopping probability g_{IJ} exceeds a random number (ξ) between zero and one. Observables are then calculated from an ensemble of trajectories using different series of ξ . Fig. (2.2) illustrates the idea of this procedure.

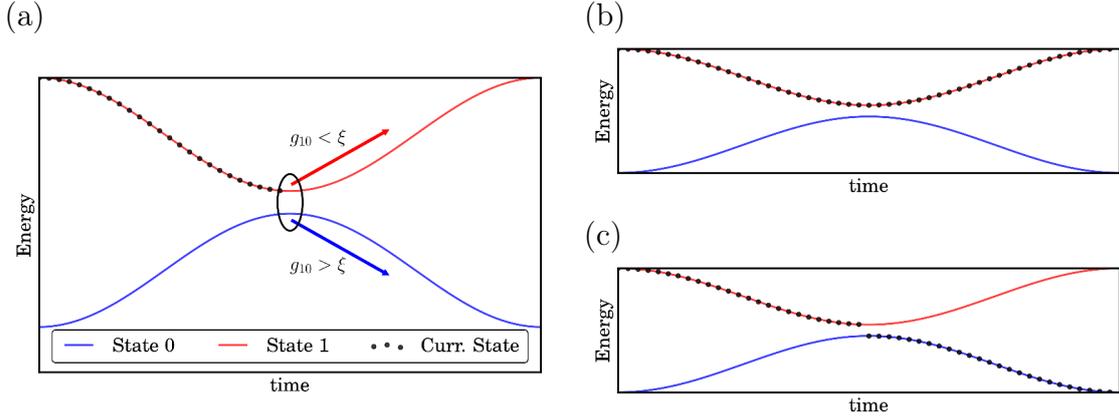


Figure 2.2: (a) Concept of surface hopping dynamics: depending on the random number ξ , the system remains on its PES or hops to state 0. The ratio of the resulting trajectories ((b) and (c)) calculated from an ensemble of independent simulations with different ξ reflects the probability g_{10} , and thus the behavior of the system.

In order to calculate g_{IJ} , the following ansatz [24] for the electronic wavefunction

$$\Psi_{\text{el}}(\mathbf{r}, \bar{\mathbf{R}}, t) = \sum_J c_J(t) \psi_J^{\text{el}}(\mathbf{r}, \bar{\mathbf{R}}) , \quad (2.12)$$

is inserted into the electronic time-dependent Schrödinger equation

$$\hat{H}_{\text{el}} \Psi_{\text{el}}(\mathbf{r}, \bar{\mathbf{R}}, t) = i\hbar \frac{\partial}{\partial t} \Psi_{\text{el}}(\mathbf{r}, \bar{\mathbf{R}}, t) . \quad (2.13)$$

Please note that the time-dependency is shifted to the state amplitudes (\vec{c}), which define the "shares" of the adiabatic states in the current electronic wavefunction. The parameters of the wavefunctions will again be omitted. Multiplying with $\langle \psi_I^{\text{el}} |$ from the left leads to

$$\langle \psi_I^{\text{el}} | \hat{H}_{\text{el}} | \sum_J c_J \psi_J^{\text{el}} \rangle = i\hbar \sum_J c_J \langle \psi_I^{\text{el}} | \frac{\partial}{\partial t} \psi_J^{\text{el}} \rangle + i\hbar \sum_J \langle \psi_I^{\text{el}} | \psi_J^{\text{el}} \rangle \frac{\partial}{\partial t} c_J . \quad (2.14)$$

This can be simplified to

$$\sum_J [E_J \delta_{IJ} - i\hbar Q_{IJ}] c_J = i\hbar \frac{\partial}{\partial t} c_I . \quad (2.15)$$

The non-adiabatic coupling (Q_{IJ}) can be calculated from the NACV by applying the chain rule

$$Q_{IJ} = \langle \psi_I^{\text{el}} | \frac{\partial}{\partial t} \psi_J^{\text{el}} \rangle = \dot{\mathbf{R}} \cdot \boldsymbol{\tau}_{IJ} . \quad (2.16)$$

To approximately solve eq. (2.15), \vec{c} is propagated using a unitary propagator [40]

$$\vec{c}(t + \delta t) \approx \mathbf{U}(t, t + \delta t) \vec{c}(t) , \quad (2.17)$$

$$\mathbf{U}(t, t + \delta t) = \exp \left(\frac{\mathbf{H}(t) + \mathbf{H}(t + \delta t)}{2} \delta t \right) , \quad (2.18)$$

$$H_{IJ}(t) = \exp \left[\frac{i}{\hbar} \int_0^t (E_I - E_J) dt' \right] Q_{IJ}. \quad (2.19)$$

E_I and Q_{IJ} thereby depend on t implicitly through \mathbf{R} . The chosen time step for this propagation (δt) is usually 1000 times smaller than Δt [25]. E_I and Q_{IJ} are, therefore, interpolated between t and $t + \Delta t$. With \vec{c} at hand, g_{IJ} becomes accessible

$$g_{IJ}(t, t + \Delta t) = -\frac{1}{a_{II}} \int_t^{t+\Delta t} dt' \frac{\partial a_{II}}{\partial t} = -\frac{2}{a_{II}} \int_t^{t+\Delta t} dt' \Re \{ a_{JI}^* H_{JI} \}, \quad (2.20)$$

with \mathbf{a} being defined as

$$\mathbf{a} = \mathbf{c} \mathbf{c}^\dagger. \quad (2.21)$$

Hops from I to J are only possible in case of $g_{IJ} > \xi$ and when the system has enough nuclear velocity along the corresponding $\boldsymbol{\tau}_{IJ}$ (not frustrated). After a successful hop, \mathbf{p} is rescaled along $\boldsymbol{\tau}_{IJ}$ in order to conserve the total energy. Alg. 2 presents the pseudo-code of a NAMMD simulation.

Algorithm 2 Pseudo-code of non-adiabatic molecular dynamics.

Initial Values for \mathbf{R} and \mathbf{p}

Initial Calculation of E_I , $\nabla_A E_I$, and \mathbf{Q}

while $t < t_{\max}$ **do**

(1) **Update Nuclear Coordinates:** $\mathbf{R}(t) \rightarrow \mathbf{R}(t + \Delta t)$

(2) **Calculate** $E_I(t + \Delta t)$, $\nabla_A E_I(t + \Delta t)$, and $\mathbf{Q}(t + \Delta t)$

(3) **Determine** $\vec{c}(t + \Delta t)$ and $\mathbf{g}(t, t + \Delta t)$

if $\xi < g_{IK}$ for random number $\xi \in [0; 1]$ and hop is not frustrated **then**

(4) **Change State:** $I \rightarrow K$

(5) **Recalculate** $E_K(t + \Delta t)$, $\nabla_A E_K(t + \Delta t)$, and $\mathbf{Q}(t + \Delta t)$, if necessary

(6) **Rescale** \mathbf{p} along $\boldsymbol{\tau}_{IK}$

end if

end while

Alg. 2 illustrates that NAMMD is computationally more expensive than BOMD. It requires energy calculations of multiple states, gradient calculations of the ground or an excited state, and the calculation of the couplings at each step of the simulation. Additionally, the resulting observables are only meaningful when calculated as an average over multiple, independent trajectories.

2.3 Energy, Gradients, and Couplings Calculations

In the previous section, it was shown that MD simulations require energies, gradients, and (in case of NAMD) excited-state energies and properties. In this section, their calculation will be outlined, starting with the independent particle models Hartree-Fock (HF) and density functional theory (DFT) for ground-state energies and gradients. Then, the calculation of excited-state energies from linear response DFT, also called time-dependent DFT (TDDFT), is presented. In the last part, excited-state gradients and couplings at the TDDFT level of theory will be discussed.

2.3.1 The Hartree-Fock (HF) Method

In the HF method [7, 8], a Slater determinant [76] of molecular orbitals (MOs)

$$\psi_0^{\text{el}} = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \varphi_i(\mathbf{r}_1) & \dots & \varphi_k(\mathbf{r}_1) \\ \vdots & \ddots & \vdots \\ \varphi_i(\mathbf{r}_N) & \dots & \varphi_k(\mathbf{r}_N) \end{pmatrix}, \quad (2.22)$$

is used as the wavefunction in the electronic time-independent Schrödinger equation

$$\hat{H}_{\text{el}} \psi_0^{\text{el}} = E_0 \psi_0^{\text{el}}. \quad (2.23)$$

N is the number of electrons. The resulting expression for the ground-state energy

$$E_0 = \sum_i \langle \varphi_i | \hat{h} | \varphi_i \rangle + \frac{1}{2} \sum_{ij} \langle \varphi_i \varphi_j | | \varphi_i \varphi_j \rangle + V_{\text{nuc,nuc}}, \quad (2.24)$$

can be, due to the variational principle, minimized with respect to $\{\varphi_i\}$ [77]. This leads to the HF equations

$$\hat{F} |\varphi_i\rangle = \varepsilon_i |\varphi_i\rangle. \quad (2.25)$$

\hat{F} is the Fock operator, consisting of the one-electron operator (\hat{h}), the Coulomb (\hat{J}_j), and the exchange (\hat{K}_j) operator

$$\hat{F} = \hat{h} + \sum_j \hat{J}_j - \hat{K}_j, \quad (2.26)$$

$$\hat{h} = -\frac{\nabla_i^2}{2} - \sum_A \frac{Z_A}{r_{1A}}, \quad (2.27)$$

$$\hat{J}_j = \int d\mathbf{r}_2 \varphi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \varphi_j(\mathbf{r}_2), \quad \hat{K}_j = \int d\mathbf{r}_2 \varphi_j^*(\mathbf{r}_2) \frac{\hat{P}_{12}}{r_{12}} \varphi_j(\mathbf{r}_2). \quad (2.28)$$

∇_i is the derivative with respect to the electronic coordinates and Z_A is the nuclear charge. r_{1A} and r_{12} are electron-nuclear and electron-electron distances, respectively. \hat{P}_{12} is the permutation operator switching the orbitals of electron 1 and 2. The HF equations have to be solved iteratively, as the Fock operator acting on one orbital (φ_i) includes the interaction of all other orbitals (φ_j) as a mean field.

To do this efficiently, the MOs are written as a linear combination of atomic orbitals (AOs)

$$|\varphi_i\rangle = \sum_{\nu} C_{\nu i} |\chi_{\nu}\rangle , \quad (2.29)$$

where \mathbf{C} is the coefficient matrix. If eq. (2.29) is inserted into eq. (2.25), one obtains the Roothaan-Hall equations [9, 10]

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} , \quad (2.30)$$

with \mathbf{S} being the overlap matrix ($S_{\mu\nu} = \langle\mu|\nu\rangle$) and $\boldsymbol{\varepsilon}$ the orbital energies. \mathbf{F} is the Fock matrix

$$F_{\mu\nu} = \langle\mu|\hat{F}|\nu\rangle = h_{\mu\nu} + \sum_{\kappa\lambda} P_{\kappa\lambda}(\mu\nu||\kappa\lambda) , \quad (2.31)$$

depending on \mathbf{C} through the density matrix (\mathbf{P})

$$P_{\mu\nu} = \sum_i C_{\mu i}^* C_{\nu i} . \quad (2.32)$$

Eq. 2.30 is solved iteratively in the so-called self-consistent field (SCF) algorithm. The final density can then be used to calculate the HF energy

$$E_0 = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\kappa\lambda} P_{\mu\nu} P_{\kappa\lambda}(\mu\nu||\kappa\lambda) + V_{\text{nuc,nuc}} . \quad (2.33)$$

The equation for the HF gradients is as follows [77]:

$$E_0^x = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu}^x + \frac{1}{2} \sum_{\mu\nu\kappa\lambda} P_{\mu\nu} P_{\kappa\lambda}(\mu\nu||\kappa\lambda)^x - \sum_{\mu\nu} W_{\mu\nu} S_{\mu\nu}^x + V_{\text{nuc,nuc}}^x . \quad (2.34)$$

\mathbf{W} is the energy-weighted density matrix. Please note that no derivatives of the density or the coefficients appear in eq. (2.34). They are circumvented by exploiting the orthonormality of the MOs

$$[\mathbf{C}^\dagger \mathbf{S} \mathbf{C}]^x = \mathbf{C}^{x,\dagger} \mathbf{S} \mathbf{C} + \mathbf{C}^\dagger \mathbf{S}^x \mathbf{C} + \mathbf{C}^\dagger \mathbf{S} \mathbf{C}^x = 0 . \quad (2.35)$$

The HF method covers about 99 % of the exact ground-state energy in a given basis. The remaining part (the electron correlation) is a result of the approximate mean-field treatment of the electron-electron interaction. It can be calculated using perturbation theory (e.g., MP2 [15]), Coupled-Cluster methods [16, 17], or Configuration Interaction [78].

2.3.2 Density Functional Theory (DFT)

A different ansatz than the discussed wavefunction-based methods was stimulated by Hohenberg and Kohn [11], who were able to prove that the energy of a system can be written as a functional of the electron density

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu} \chi_{\mu}^*(\mathbf{r}) \chi_{\nu}(\mathbf{r}) . \quad (2.36)$$

The exact form of this functional, however, remains unknown, as only the Coulomb energy (E_{H}) and interaction of the electrons with an external potential (including the potential of the nuclei, E_{ext}) can easily be written as density functionals. To tackle this problem, Kohn and Sham [12] reintroduced orbitals (so-called Kohn-Sham (KS) orbitals), leading to the following energy expression

$$E_0 = E_{\text{T}}[\{\varphi_i\}] + E_{\text{ext}}[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] , \quad (2.37)$$

with

$$E_{\text{T}}[\{\varphi_i\}] = - \sum_i \int d\mathbf{r} \varphi_i^*(\mathbf{r}) \frac{\nabla_i^2}{2} \varphi_i(\mathbf{r}) , \quad (2.38)$$

$$E_{\text{ext}}[\rho] = \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) , \quad (2.39)$$

$$E_{\text{H}}[\rho] = \int \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1) \rho(\mathbf{r}_2)}{r_{12}} , \quad (2.40)$$

$$E_{\text{xc}}[\rho] = \Delta E_{\text{T}}[\rho] + E_{\text{x}}[\rho] + E_{\text{c}}[\rho] . \quad (2.41)$$

E_{T} is the kinetic energy calculated from the KS orbitals. E_{xc} incorporates exchange and correlation contributions, while correcting E_{T} to the exact solution ($\Delta E_{\text{T}}[\rho] = E_{\text{T}}^{\text{exact}} - E_{\text{T}}[\{\varphi_i\}]$). Minimizing the KS energy expression leads to the KS equations

$$\hat{h}_{\text{s}}[\rho] |\varphi_i\rangle = \varepsilon_i |\varphi_i\rangle , \quad (2.42)$$

with the one-particle Kohn-Sham operator

$$\hat{h}_{\text{s}}[\rho](\mathbf{r}) = -\frac{\nabla_i^2}{2} + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}[\rho](\mathbf{r}) + V_{\text{xc}}[\rho](\mathbf{r}) , \quad (2.43)$$

$$V_{\text{H}}[\rho](\mathbf{r}) = \int d\mathbf{r}_2 \frac{\rho(\mathbf{r}_2)}{r_{12}} , \quad (2.44)$$

$$V_{\text{xc}}[\rho](\mathbf{r}) = \frac{\partial E_{\text{xc}}[\rho]}{\partial \rho} . \quad (2.45)$$

As such, the KS equations are very similar to the HF equations. Both describe the system as a sum of independent particles and are solved similarly. The main difference is, however, that a KS matrix ($\langle \mu | \hat{h}_{\text{s}}[\rho] | \nu \rangle$) is built instead of the Fock matrix and that the KS reference system (consisting of the KS orbitals) is ought to reproduce the fully interacting system. Due to the complex functional form of E_{xc} , V_{xc} is obtained via numerical integration on a grid.

The results of DFT depend heavily on the chosen exchange-correlation functional (E_{xc}), for which a hierarchy exist. Climbing the so-called Jacob's ladder [79], the number of elements on which the functional depends increases, leading to (roughly) better results. Examples are local density approximation (LDA) functionals, which depend solely on the density ($E_{xc}[\rho]$), generalized gradient approximation (GGA) functionals, which incorporate contributions from the density gradient ($E_{xc}[\rho, \nabla_A \rho]$), and hybrid functionals, which add an amount (c_x) of exact HF exchange ($E_{xc}[\rho, \nabla_A \rho, \{\varphi_i\}]$). DFT energies and properties are typically better than HF (in some cases even comparable to post-HF calculations), while featuring the same complexity. This is mainly a result of the implicit form of electron correlation in V_c . Additionally, DFT does not require very large basis sets, which are inevitable in post-HF methods.

2.3.3 Time-Dependent Density Functional Theory (TDDFT)

Static and Dynamic Linear Response DFT

Properties at the DFT level of theory can be derived from linear response theory. For static (time-independent) properties, this is done by expanding the KS Hamiltonian, the KS orbitals, and the orbital energies as follows [80, 81]:

$$\hat{h}_s = \hat{h}_s^0 + \lambda \hat{h}_s^1 + \dots, \quad (2.46)$$

$$\varphi_i(\mathbf{r}) = \varphi_i^0(\mathbf{r}) + \lambda \varphi_i^1(\mathbf{r}) + \dots, \quad (2.47)$$

$$\varepsilon_i = \varepsilon_i^0 + \lambda \varepsilon_i^1 + \dots. \quad (2.48)$$

Superscript 0 and 1 thereby denote unperturbed and perturbed (first order) quantities. If this is inserted into the KS equations (see eq. (2.42)), one obtains a set of equations for the different orders of λ . In linear response theory, the equation of order 1, also called Sternheimer equation, is considered

$$\hat{h}_s^1 \varphi_i^0 = (\varepsilon_i^0 - \hat{h}_s^0) \varphi_i^1. \quad (2.49)$$

Here, ε_i^1 has been set to zero, as it has no influence on φ_i^1 . Hence, the response of a system (φ_i^1) can be calculated from an effective potential, which incorporates the applied potential (V_{appl}) and the response potential (V_{scf})

$$\hat{h}_s^1 = V_{\text{eff}}[\rho^1](\mathbf{r}_1) = V_{\text{appl}}(\mathbf{r}_1) + V_{\text{scf}}[\rho^1](\mathbf{r}_1) =$$

$$V_{\text{appl}}(\mathbf{r}_1) + V_{\text{H}}[\rho^1](\mathbf{r}_1) + V_{\text{xc}}[\rho^1](\mathbf{r}_1) = V_{\text{appl}}(\mathbf{r}_1) + \int \left[\frac{1}{r_{12}} + f_{\text{xc}}[\rho^0](\mathbf{r}_1, \mathbf{r}_2) \right] \rho^1(\mathbf{r}_2) d\mathbf{r}_2, \quad (2.50)$$

with

$$f_{\text{xc}}[\rho^0](\mathbf{r}_1, \mathbf{r}_2) = \frac{\partial V_{\text{xc}}[\rho^0](\mathbf{r}_1)}{\partial \rho(\mathbf{r}_2)} = \frac{\partial^2 E_{\text{xc}}[\rho^0]}{\partial \rho(\mathbf{r}_1) \partial \rho(\mathbf{r}_2)}, \quad (2.51)$$

$$\rho^1(\mathbf{r}) = \sum_i [\varphi_i^1(\mathbf{r})]^* \varphi_i^0(\mathbf{r}) + [\varphi_i^0(\mathbf{r})]^* \varphi_i^1(\mathbf{r}). \quad (2.52)$$

Please note that the terms for exact exchange (in case of hybrid functionals) have been neglected here. In 1984, Runge and Gross [82] delivered a formal justification for DFT being suitable for calculating time-dependent properties. In these cases, the Hamiltonian and the wavefunction of the time-dependent KS equations

$$\hat{h}_s[\rho_t](t)\varphi_i(\mathbf{r}, t) = i\hbar \frac{\partial \varphi_i(\mathbf{r}, t)}{\partial t} , \quad (2.53)$$

with

$$\hat{h}_s[\rho_t](t) = -\frac{1}{2}\nabla_i^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{appl}}(\mathbf{r}, t) + V_{\text{H}}[\rho_t](\mathbf{r}) + V_{\text{xc}}[\rho_t](\mathbf{r}) , \quad (2.54)$$

$$\varphi_i(\mathbf{r}, t) = \exp\left\{-\frac{i}{\hbar}\varepsilon_i t\right\}\varphi_i(\mathbf{r}) , \quad (2.55)$$

are considered, using the adiabatic approximation [83]

$$V_{\text{xc}}[\rho](\mathbf{r}, t) = V_{\text{xc}}[\rho_t](\mathbf{r}) = \frac{\partial E_{\text{xc}}[\rho_t]}{\partial \rho_t(\mathbf{r})} , \quad (2.56)$$

which states that the exchange-correlation potential only depends on time implicitly through the change of the electron density. The only explicit time-dependency is observed in the applied potential, while the potential of the nuclei (V_{ext}) is assumed to be time-independent. To obtain the dynamic response equations [28, 29], \hat{h}_s and φ_i are expanded

$$\hat{h}_s(t) = \hat{h}_s[\rho_0](0) + \lambda(V_{\text{appl}}(t) + V_{\text{scf}}[\rho_t^1]) + \dots ,$$

$$\varphi_i(\mathbf{r}, t) = \exp\left\{-\frac{i}{\hbar}(\varepsilon_i^0 + \lambda\varepsilon_i^1 + \dots)t\right\}\{\varphi_i^0(\mathbf{r}, 0) + \lambda\varphi_i^1(\mathbf{r}, t) + \dots\} , \quad (2.57)$$

with the Fourier transform

$$\varphi_i^1(\mathbf{r}, \pm\omega) = \int dt \varphi_i^1(\mathbf{r}, t) \exp\left\{-\frac{i}{\hbar}(\pm\omega)t\right\} . \quad (2.58)$$

ω is the frequency. Please note that unperturbed quantities are time- or frequency-independent. The linear response equation ($\lambda = 1$) is

$$(V_{\text{appl}}(\pm\omega) + V_{\text{scf}}[\rho_\omega^1])\varphi_i^0 = (\varepsilon_i^0 \pm \omega - \hat{h}_s[\rho_0])\varphi_i^1(\pm\omega) , \quad (2.59)$$

with

$$\rho_\omega^1(\mathbf{r}) = \sum_i [\varphi_i^1(\mathbf{r}, -\omega)]^* \varphi_i^0(\mathbf{r}) + [\varphi_i^0(\mathbf{r})]^* \varphi_i^1(\mathbf{r}, \omega) . \quad (2.60)$$

Again, the response of the system ($\varphi_i^1(\pm\omega)$) to a frequency-dependent perturbation ($V_{\text{appl}}(\pm\omega)$) is calculated via the response potential ($V_{\text{scf}}[\rho_\omega^1]$).

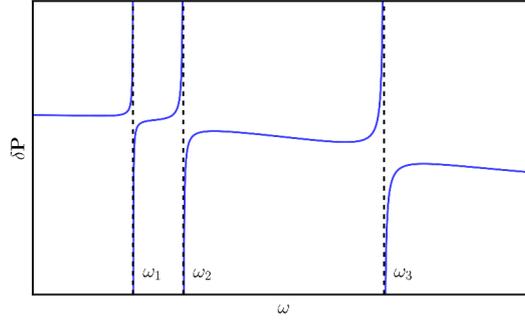


Figure 2.3: Frequency-dependent response ($\delta\mathbf{P}(\omega)$) of a system involving three excited states. Note that $\delta\mathbf{P}(\omega)$ features poles at the excitation energies (ω_I).

The Random Phase Approximation (RPA) and the Tamm-Dancoff Approximation (TDA)

Eq. (2.60) consists of two sets of equations one for $+\omega$ and one for $-\omega$. To transform it into an algebraic form, the response orbitals ($\varphi_i^1(\mathbf{r}, \pm\omega)$) in eqs. (2.59) and (2.60) are expanded in a series of virtual unperturbed KS orbitals

$$\varphi_j^1(\mathbf{r}, +\omega) = \sum_b X_{jb}(\omega)\varphi_b(\mathbf{r}) , \quad (2.61)$$

$$\varphi_j^1(\mathbf{r}, -\omega) = \sum_b Y_{bj}(\omega)\varphi_b(\mathbf{r}) . \quad (2.62)$$

The equation of $+\omega$ is then multiplied with $|\varphi_a\rangle$ from the right and the equation of $-\omega$ is multiplied with $\langle\varphi_a|$ from the left. The results can be concatenated to a single equation [28, 29]

$$\left(\omega \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} - \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \right) \begin{pmatrix} \mathbf{X}(\omega) \\ \mathbf{Y}(\omega) \end{pmatrix} = \begin{pmatrix} V_{\text{appl}}(\omega) \\ V_{\text{appl}}^*(\omega) \end{pmatrix} , \quad (2.63)$$

with

$$A_{iajb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + D_{iajb} , \quad (2.64)$$

$$B_{iajb} = D_{iabj} , \quad (2.65)$$

using

$$D_{pqrs} = (pq|rs) + (pq|f_{\text{xc}}[\rho^0]|rs) = (pq|rs) + f_{pqrs}^{\text{xc}} . \quad (2.66)$$

Eq. (2.63) describes, as already mentioned above, the response of the density ($\delta\mathbf{P}(\omega)$, in form of the transition densities $\mathbf{X}(\omega)$ and $\mathbf{Y}(\omega)$) to an applied, frequency-dependent perturbation ($V_{\text{appl}}(\omega)$). It roughly takes the following form

$$\delta\mathbf{P}(\omega) = \sum_I \frac{f_I}{\omega_I^2 - \omega^2} , \quad (2.67)$$

which is plotted in fig. (2.3).

When ω hits an excitation energy of the system (ω_I), $\delta\mathbf{P}$ goes to infinity. Assuming an arbitrary small perturbation for this case, eq. (2.63) reads

$$\left(\omega_I \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} - \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \right) \begin{pmatrix} \infty \\ \infty \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (2.68)$$

The first part must thus be zero, leading to the random phase approximation (RPA) or TDDFT equation [28, 84, 85]

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} = \omega_I \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}. \quad (2.69)$$

The transition matrices are usually normalized

$$\mathbf{X}_I^\dagger \mathbf{X}_I - \mathbf{Y}_I^\dagger \mathbf{Y}_I = \pm 1. \quad (2.70)$$

To solve the TDDFT equation, eq. (2.69) is converted to an eigenvalue problem of type [29]

$$\mathbf{\Omega} \vec{u} = \omega_I^2 \vec{u}, \quad (2.71)$$

with

$$\mathbf{\Omega} = (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} (\mathbf{A} + \mathbf{B}) (\mathbf{A} - \mathbf{B})^{\frac{1}{2}}, \quad (2.72)$$

$$\vec{u}_I = (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} (\mathbf{X}_I + \mathbf{Y}_I), \quad (2.73)$$

and using

$$(\mathbf{A} + \mathbf{B})_{iajb} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + 2(ia|jb) + 2f_{iajb}^{\text{xc}} - c_x [(ja|ib) + (ab|ij)], \quad (2.74)$$

$$(\mathbf{A} - \mathbf{B})_{iajb} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + c_x [(ja|ib) + (ab|ij)]. \quad (2.75)$$

Eqs. (2.74) and (2.75) now contain the more general form, which is also suitable for hybrid density functionals. $\mathbf{\Omega}$ is diagonalized iteratively via the Davidson algorithm [86], as not all excitation energies $\omega = (\omega_1, \omega_2, \dots)$ are required but only the first few of them. In the Tamm-Dancoff approximation (TDA) [87], \mathbf{B} in eq. (2.69) is neglected, leading to the following equation:

$$\mathbf{A} \mathbf{X}_I = \omega_I \mathbf{X}_I. \quad (2.76)$$

Eqs. (2.69) and (2.76) are identical to the time-dependent HF (TDHF) and configuration interaction singles (CIS) equations, respectively, when c_x is set to one and the exchange-correlation terms are neglected.

Despite its shortcomings [29, 88] (e.g., incapability to describe two-electron excitations), TDDFT has become a widely used method for the calculation of excitation energies, delivering good results at comparably low computational cost [89, 90]. Applying range-corrected DFT functionals, it is even possible to describe charge-transfer excitations [91]. TDA results are often close to TDDFT results, while requiring less computation time (fewer two-electron integral evaluations are needed). Additionally, it was shown that TDA features less convergence problems close to conical intersections involving the ground state [44, 92].

2.3.4 Gradients and Couplings at the TDDFT Level

Lagrangian Formalism

In general, excited-state properties are calculated by taking the derivative with respect to the nuclei of a function G' , which is defined in the following sections for gradients ($G' = G_I$) and couplings ($G' = G_{0I}, G_{IJ}$). A direct calculation of G'^x would, however, incorporate the expensive calculation of \mathbf{C}^x . This can be circumvented by calculating the Lagrangian L' instead [93–96]

$$L' = G' + \sum_{ia} Z'_{ia} F_{ia} - \sum_{pq} W'_{pq} (S_{pq} - \delta_{pq}) , \quad (2.77)$$

with \mathbf{Z}' and \mathbf{W}' being Lagrange multipliers forcing the MOs to be results of the KS equations and orthonormal, respectively

$$\frac{\partial L'}{\partial Z'_{ia}} = F_{ia} = 0 , \quad (2.78)$$

$$\frac{\partial L'}{\partial W'_{pq}} = S_{pq} - \delta_{pq} = 0 . \quad (2.79)$$

The derivative

$$L'^x = G'^{(x)} + \sum_{ia} Z'_{ia} F_{ia}^{(x)} - \sum_{pq} W'_{pq} (S_{pq}^{(x)} - \delta_{pq}) , \quad (2.80)$$

contains only integral derivatives, when \mathbf{Z}' and \mathbf{W}' are calculated from the following equation

$$\frac{\partial L'}{\partial C_{\mu p}} = 0 , \quad (2.81)$$

leading to

$$Q'_{pq} + \sum_{ia} Z'_{ia} \sum_{\mu} \frac{\partial F_{ia}}{\partial C_{\mu p}} C_{\mu q} = \sum_{rs} W'_{rs} \sum_{\mu} \frac{\partial S_{rs}}{\partial C_{\mu p}} C_{\mu q} , \quad (2.82)$$

$$Q'_{pq} = \sum_{\mu} \frac{\partial G'}{\partial C_{\mu p}} C_{\mu q} . \quad (2.83)$$

It is possible to separate \mathbf{Q} into four submatrices

$$Q'_{ij} + H_{ij}^+[\mathbf{Z}'] = W'_{ij} , \quad (2.84)$$

$$Q'_{ia} + \epsilon_a Z'_{ia} + H_{ia}^+[\mathbf{Z}'] = W'_{ia} , \quad (2.85)$$

$$Q'_{ai} + \epsilon_i Z'_{ia} = W'_{ai} , \quad (2.86)$$

$$Q'_{ab} = W'_{ab} , \quad (2.87)$$

using

$$H_{pq}^+[\mathbf{V}] = \sum_{rs} \{2(pq|rs) + 2f_{pqrs}^{\text{xc}} - c_x[(ps|rq) - (pr|sq)]\} V_{rs} . \quad (2.88)$$

The conditional equation of \mathbf{Z}' is obtained by exploiting that $W'_{ia} = W'_{ai}$

$$(\varepsilon_a - \varepsilon_i)Z'_{ia} + H_{ia}^+[\mathbf{Z}'] = -(Q'_{ia} - Q'_{ai}) , \quad (2.89)$$

$$\sum_{jb} (\mathbf{A} + \mathbf{B})_{iajb} Z'_{jb} = -U'_{ia} . \quad (2.90)$$

This is known as the Z-vector equation. Having \mathbf{Z}' at hand, \mathbf{W}' is determined via eqs. (2.84)-(2.87). Both multipliers are then used in eq. (2.80) to obtain the desired derivative. TDA derivatives are, again, obtained by setting \mathbf{B} to zero.

In the following, the calculation of the TDDFT excited-state gradients (ω_I) and NACVs ($\boldsymbol{\tau}_{0I}$ and $\boldsymbol{\tau}_{IJ}$) are briefly presented and discussed. The equations for the relaxed differences density matrices ($\mathbf{P}_I, \mathbf{P}_{0I}, \mathbf{P}_{IJ}$), energy-weighted difference density matrices ($\mathbf{W}_I, \mathbf{W}_{0I}, \mathbf{W}_{IJ}$), and two-particle difference density matrices ($\boldsymbol{\Gamma}_I, \boldsymbol{\Gamma}_{0I}, \boldsymbol{\Gamma}_{IJ}$), as well as the right hand sides of the Z-vector equations ($\mathbf{U}_I, \mathbf{U}_{0I}, \mathbf{U}_{IJ}$) are listed in the appendix of publication **II**.

Excited-State Gradients

In the case of the excited-state gradient calculation, G' is defined as follows [93]:

$$G_I = \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} - \omega_I \left[\begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} - 1 \right] . \quad (2.91)$$

Finding a stationary point of G_I with respect to $(\mathbf{X}_I \mathbf{Y}_I)$ and ω_I is equivalent to solving the RPA equations with orthonormal transition densities

$$\frac{\partial G_I}{\partial (\mathbf{X}_I \mathbf{Y}_I)} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} - \omega_I \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} = 0 , \quad (2.92)$$

$$\frac{\partial G_I}{\partial \omega_I} = \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} - 1 = \mathbf{X}_I^T \mathbf{X}_I - \mathbf{Y}_I^T \mathbf{Y}_I - 1 = 0 . \quad (2.93)$$

It can be used to calculate the change of the excitation energy with respect to the nuclear coordinates, using the ansatz discussed above

$$\begin{aligned} \omega_I^{\text{x}} &= L_I^{(\text{x})} = G_I^{(\text{x})} + \sum_{ia} Z_{ia}^I F_{ia}^{(\text{x})} - \sum_{pq} W_{pq}^I S_{pq}^{(\text{x})} \\ &= \sum_{\mu\nu} \{h_{\mu\nu}^{\text{x}} P_{\mu\nu}^I - S_{\mu\nu}^{\text{x}} W_{\mu\nu}^I + V_{\mu\nu}^{\text{xc}(\text{x})} P_{\mu\nu}^I\} \\ &\quad + \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^{\text{x}} \Gamma_{\mu\nu\kappa\lambda}^I + \sum_{\mu\nu\kappa\lambda} f_{\mu\nu\kappa\lambda}^{\text{xc}(\text{x})} R_{\mu\nu}^I R_{\kappa\lambda}^I . \end{aligned} \quad (2.94)$$

\mathbf{R}_I is equal to $(\mathbf{X}_I + \mathbf{Y}_I)$ or \mathbf{X}_I in case of RPA or TDA, respectively. In order to calculate \mathbf{P}_I , the Z-vector equation (eq. (2.90)) has to be solved iteratively.

Couplings between the Ground and an Excited State

Calculating the NACVs including the ground state is significantly simpler, as G' reduces to [94]

$$G_{0I} = \sum_{pq} \langle \varphi_p | \varphi_q \rangle \gamma_{pq}^{0I} \quad (2.95)$$

$$\gamma_{0I} = \mathbf{L}_I . \quad (2.96)$$

\mathbf{L}_I is $(\mathbf{X}_I - \mathbf{Y}_I)$ or \mathbf{X}_I in case of TDDFT or TDA, respectively. The reason for this is that the perturbation (x) can be shifted to the ground-state determinant. Consequently, the orthonormality of the transition densities and the validity of the RPA equations upon the perturbation does not have to be enforced, as demonstrated in the previous section. The NACVs can then be derived as follows:

$$\begin{aligned} \tau_{0I} &= L_{0I}^{(x)} = G_{0I}^{(x)} + \sum_{ia} Z_{ia}^{0I} F_{ia}^{(x)} - \sum_{pq} W_{pq}^{0I} S_{pq}^{(x)} \\ &= \sum_{\mu\nu} [h_{\mu\nu}^x P_{\mu\nu}^{0I} - S_{\mu\nu}^x W_{\mu\nu}^{0I} + V_{\mu\nu}^{xc(x)} P_{\mu\nu}^{0I}] \\ &\quad + \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^x \Gamma_{\mu\nu\kappa\lambda}^{0I} + \sum_{\mu\nu} S_{\mu\nu}^{[A]x} \gamma_{\mu\nu}^{0I} , \end{aligned} \quad (2.97)$$

with $\mathbf{S}^{[A]x}$ being an anti-symmetric overlap derivative

$$S_{\mu\nu}^{[A]x} = \langle \chi_\mu^x | \chi_\nu \rangle - \langle \chi_\mu | \chi_\nu^x \rangle . \quad (2.98)$$

Here, the iterative solution of the Z-vector equation is not necessary, as \mathbf{P}_{0I} can be calculated directly from the transition densities. The term including $\mathbf{S}^{[A]x}$ introduces translational variance into the NACV. Setting this term to zero is equivalent to incorporating electron-translational factors (ETFs) [97].

Couplings between two Excited States

Calculating NACVs involving two excited states is more complicated than NACVs with the ground state. Here, G' is defined as follows [95, 96]:

$$G_{IJ} = \frac{1}{\omega_{JI}} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_J \\ \mathbf{Y}_J \end{pmatrix} + \sum_{pq} \langle \varphi_p | \varphi_q \rangle \gamma_{pq}^{IJ} , \quad (2.99)$$

$$\gamma_{IJ} = \frac{1}{2} \begin{pmatrix} -(\mathbf{R}_I^T \mathbf{L}_J + \mathbf{L}_I^T \mathbf{R}_J)_{ij} & 0 \\ 0 & (\mathbf{R}_I \mathbf{L}_J^T + \mathbf{L}_I \mathbf{R}_J^T)_{ab} \end{pmatrix} . \quad (2.100)$$

ω_{JI} is equal to $\omega_J - \omega_I$. The first part of the equation forces the transition densities to be solutions of the RPA equations and to be orthonormal

$$\begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_J \\ \mathbf{Y}_J \end{pmatrix} = \omega_J \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_J \\ \mathbf{Y}_J \end{pmatrix} = \omega_J (\mathbf{X}_I^T \mathbf{X}_J - \mathbf{Y}_I^T \mathbf{Y}_J) = 0 , \quad (2.101)$$

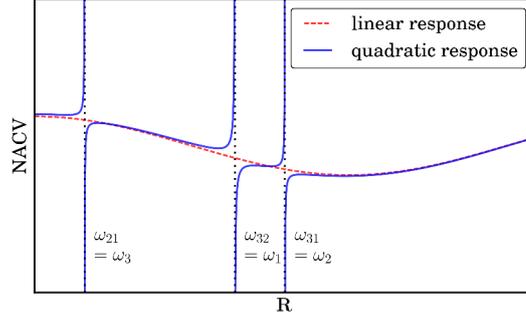


Figure 2.4: Schematic comparison of NACVs calculated from linear and quadratic response. Note that the quadratic response features unphysical poles at $\omega_J - \omega_I = \omega_K$.

$$\begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_J \\ \mathbf{Y}_J \end{pmatrix} = \omega_I \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}^T \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{pmatrix} \mathbf{X}_J \\ \mathbf{Y}_J \end{pmatrix} = \omega_I (\mathbf{X}_I^T \mathbf{X}_J - \mathbf{Y}_I^T \mathbf{Y}_J) = 0. \quad (2.102)$$

The second part calculates the NACVs using the pseudo-wavefunction approach. If G_{IJ} is used in the Lagrangian presented above (see eq. (2.77)), the following equation for the NACVs can be derived [95–101]

$$\begin{aligned} \boldsymbol{\tau}_{IJ} = L^{(x)} &= G_{IJ}^{(x)} + \sum_{ia} Z_{ia}^{IJ} F_{ia}^{(x)} - \sum_{pq} W_{pq}^{IJ} S_{pq}^{(x)} \\ &= \frac{1}{\omega_{JI}} \left\{ \sum_{\mu\nu} [h_{\mu\nu}^x P_{\mu\nu}^{IJ} - S_{\mu\nu}^x W_{\mu\nu}^{IJ} + V_{\mu\nu}^{\text{xc}(x)} P_{\mu\nu}^{IJ}] \right. \\ &\quad \left. + \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^x \Gamma_{\mu\nu\kappa\lambda}^{IJ} + \sum_{\mu\nu\kappa\lambda} f_{\mu\nu\kappa\lambda}^{\text{xc}(x)} R_{\mu\nu}^I R_{\kappa\lambda}^J \right\} \\ &\quad + \sum_{\mu\nu} S_{\mu\nu}^{[A]x} \gamma_{\mu\nu}^{IJ}. \end{aligned} \quad (2.103)$$

Neglecting the last term is, again, equivalent to applying ETFs. Calculating $\boldsymbol{\tau}_{IJ}$ is computationally more time-consuming than the calculation of the other excited-state properties. In contrast to $\boldsymbol{\tau}_{0I}$, the \mathbf{Z} -vector equation has to be solved iteratively and the algorithm requires more two-electron integral contractions than in case of ω_I .

NACVs between excited states have also been derived from quadratic response theory. However, the resulting couplings feature unphysical poles, when the energy difference between two states approaches the excitation energy of another state ($\omega_J - \omega_I = \omega_K$) [95, 96, 102–104], as demonstrated in fig. (2.4). For this reason, the use of the presented pseudo-wavefunction approach is recommended.

2.4 Efficient Molecular Dynamics Simulations

In the last sections, it was shown that AIMD simulations are computationally demanding. Therefore, techniques, approximations, and algorithms to accelerate these calculations while maintaining the necessary accuracy are essential when investigating large molecular systems. This section presents different approaches enabling efficient AIMD simulations: The extended-Lagrangian scheme, the corrected small basis set HF method, simplified TDDFT schemes, the Hammes–Schiffer–Tully approach, and the use of GPUs. They address different parts of the algorithm and can, thus, easily be combined.

2.4.1 Extended-Lagrangian BOMD (XL-BOMD)

The main problem of AIMD simulations using HF or DFT is that both the energy and the electron density need to be determined self-consistently at every step of the MD simulation. This is circumvented in Car-Parrinello MD (CPMD) [105], where the electronic structure is propagated along with the nuclei and close to the ground state. This approach has, however, some major disadvantages [106]. The time step has to be chosen smaller than in standard BOMD, the choice of the fictitious electron mass is not arbitrary, and the system is only close to the ground state.

The extended-Lagrangian BOMD (XL-BOMD) method of Niklasson *et al.* [53, 107–112] combines the best of CPMD and BOMD. Instead of the electron density itself, an auxiliary density is propagated close to the ground-state density [53]

$$\mathbf{P}^{\text{aux}}(t + \Delta t) = 2\mathbf{P}^{\text{aux}}(t) - \mathbf{P}^{\text{aux}}(t - \Delta t) + \kappa[\mathbf{P}(t) - \mathbf{P}^{\text{aux}}(t)] . \quad (2.104)$$

If $\mathbf{P}^{\text{aux}}(t + \Delta t)$ is used as a guess for the SCF at $t + \Delta t$, the number of necessary SCF cycles is greatly reduced. The problem with this approach is, however, that errors arising in the calculation of \mathbf{P} , e.g., from loose convergence criteria of the SCF, are also propagated and amplified throughout the trajectory. For this reason, a dissipative force

$$\mathbf{F}^{\text{diss}} = \alpha_K \sum_{k=0}^K c_k \mathbf{P}^{\text{aux}}(t - k\Delta t) , \quad (2.105)$$

is added to eq. (2.104). The parameters κ , α_K , and c_K have been optimized for different orders K [53]. The XL-BOMD scheme does not only reduce the number of necessary SCF cycles, it also enables stable BOMD with a fixed number of SCF cycles. In publication **I**, only two SCF steps (diagonalizations) are performed at every step of the simulations. Using this approach it is not possible to describe dynamics involving the dissociations of bonds. The number of diagonalizations can even be reduced to one, but this requires additional terms in the forces calculations [111]. In all other works (**II–V**), the extended-Lagrangian scheme is simply used to accelerate the SCF convergence.

2.4.2 Corrected Small Basis Set HF (HF-3c)

As discussed in Section 2.2.1, the HF method is nowadays mostly used as a starting point for wavefunction-based methods, like MP2 or Coupled Cluster. However, it

was shown already in the past that HF combined with a minimal basis set can lead to (surprisingly) good results regarding small molecule geometries or interaction energies [113, 114]. This inspired Sure and Grimme to use HF with the minimal basis set MINIX as a starting point for the corrected small basis set HF (HF-3c) method [50]. Three correction terms with nine empirical parameters are then added subsequently to counteract the major shortcomings of the HF/MINIX calculation: The lack of dispersion, the basis set superposition error, and the erroneous bond lengths:

$$E_0^{\text{HF-3c}} = E_0^{\text{HF/MINIX}} + E_{\text{disp}}^{\text{D3(BJ)}} + E_{\text{BSSE}}^{\text{gCP}} + E_{\text{SRB}} . \quad (2.106)$$

$E_{\text{disp}}^{\text{D3(BJ)}}$ is the D3 correction with Becke-Johnson damping [115–117], $E_{\text{BSSE}}^{\text{gCP}}$ is the geometrical counterpoise correction [118], and E_{SRB} is the short-ranged term to correct the bond lengths. All three contributions depend solely on the geometry.

The HF-3c method has been benchmarked regarding geometries, interaction energies, and vibrational frequencies [50]. It outperforms pure semi-empirical methods, without the need of atom-type dependent corrections, e.g., for hydrogen bonds. Due to its comparably low computational cost, it even allows for calculating small proteins, which normally require the use of highly parametrized force field methods. In publication **I**, it is used for efficient AIMD simulations. The resulting IR spectra are of good quality and are comparable to experimental spectra, when a scaling factor is applied. This indicates that the entire PES is described reasonably well with HF-3c, so that it is used for the free energy calculations in publications **IV** and **V** as well.

2.4.3 Simplified TDDFT Schemes

In TDDFT and TDA calculations, the evaluation of the two-electron integrals (see eqs. (2.74) and (2.75)) is the most expensive step. For this reason, these integrals are approximated in the simplified TDA/TDDFT approaches by Grimme and coworkers (sTDA or sTDDFT) [51, 52]. The approach uses the Mataga-Nishimoto-Ohno-Klopman damped Coulomb operators to calculate Coulomb (J') and exchange (K')

$$J'_{pqrs} = \sum_{AB} q_{pq}^A \left(\frac{1}{r_{AB}^{\beta_s} + (c_x \eta_{AB})^{-\beta_s}} \right)^{\frac{1}{\beta_s}} q_{rs}^B , \quad (2.107)$$

$$K'_{pqrs} = \sum_{AB} q_{pq}^A \left(\frac{1}{r_{AB}^{\alpha_s} + \eta_{AB}^{-\alpha_s}} \right)^{\frac{1}{\alpha_s}} q_{rs}^B . \quad (2.108)$$

The transition/charge density monopoles (q) are obtained from a Löwdin population analysis [119]. η_{AB} is the mean of the chemical hardness of the corresponding atoms, while α_s and β_s are defined as:

$$\alpha_s = \alpha^{(1)} + c_x \alpha^{(2)} , \quad (2.109)$$

$$\beta_s = \beta^{(1)} + c_x \beta^{(2)} . \quad (2.110)$$

$\alpha/\beta^{(1)/(2)}$ are global fit parameters. With J' and K' at hand, the simplified matrices \mathbf{A}' and \mathbf{B}' are built

$$A'_{iajb} = \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i) + s_{\mathbf{K}} K'_{iajb} - J'_{ijab} , \quad (2.111)$$

$$B'_{iajb} = s_k K'_{iabj} - c_x K'_{iba j} . \quad (2.112)$$

s_k is 2 or 0 for singlet-singlet or singlet-triplet excitations. These are then incorporated in the eigenvalue problems (see eqs. (2.71) and (2.76)), which deliver the excitation energies by diagonalizing $\mathbf{\Omega}'$ or \mathbf{A}' in case of sTDDFT or sTDA, respectively. To avoid the diagonalization of the entire matrix, the single excitation space is truncated [51].

As such, sTDA is able to predict UV- and CD-spectra with an error of 0.2-0.3 eV [51], while being 100 to 1000 times faster than conventional TDDFT. sTDDFT improves upon sTDA regarding the transition dipole moments [52]. The simplified schemes have also been expanded towards range-corrected functionals [120] and have been used to calculate polarizabilities and hyperpolarizabilities [121]. In manuscript **III**, it is shown that sTDDFT and sTDA are also suitable for the calculation of excited-state properties and NAMD simulations.

2.4.4 The Hammes–Schiffer–Tully (HST) Model

In publication **II**, the non-adiabatic couplings (Q_{IJ}) are determined from the NACVs (see eq. (2.16)). They can, however, also be calculated via a finite-differences approach in the so-called Hammes–Schiffer–Tully (HST) model [25]

$$Q_{IJ}(t) \approx \frac{1}{4\Delta t} \{ 3O_{IJ}(t, t - \Delta t) - 3O_{IJ}(t - \Delta t, t) - O_{IJ}(t - \Delta t, t - 2\Delta t) + O_{IJ}(t - 2\Delta t, t - \Delta t) \} , \quad (2.113)$$

with

$$O_{IJ}(t_1, t_2) = \langle \phi_I(t_1) | \phi_J(t_2) \rangle . \quad (2.114)$$

Q_{IJ} is thus determined solely from wavefunction overlaps of two different states at two different times. In TDDFT, they can be approximated by using the transition densities

$$O_{0I}(t_1, t_2) = \sum_{ia} \gamma_{ia}^{0I}(t_1) S_{ia}(t_1, t_2) ,$$

$$O_{IJ}(t_1, t_2) = \sum_{pq} \gamma_{pq}^{IJ}(t_1) S_{pq}(t_1, t_2) + \sum_{ia} X_{ia}^I(t_1) X_{ia}^J(t_2) - \sum_{ia} Y_{ia}^I(t_1) Y_{ia}^J(t_2) . \quad (2.115)$$

γ_{0I} and γ_{IJ} are defined in eqs. (2.96) and (2.100). $S_{pq}(t_1, t_2)$ is the overlap between two MOs at different times

$$S_{pq}(t_1, t_2) = \langle \varphi_p(t_1) | \varphi_q(t_2) \rangle . \quad (2.116)$$

The HST model is widely used in NAMD simulations at the TDDFT but also at other levels of theory [39, 41]. In case of TDDFT, it greatly reduces the computation time, as only overlap calculations and linear algebra are required. Additionally, it was shown by Plasser *et al.* [45] that the finite-difference couplings are more stable in the vicinity of a conical intersection. In manuscript **III**, the HST model is also applied to accelerate NAMD simulations at the (s)TDDFT level of theory.

2.4.5 Molecular Dynamics on Graphics Processing Units (GPUs)

The use of GPUs in addition to CPUs has become a major trend in computational chemistry [54–60]. The reason for this is obvious: a GPU with its thousands of threads easily outperforms a CPU with only 12 or 20 threads, especially for operations that can be parallelized. In quantum chemistry, they are usually applied to the two-electron integral evaluations and their derivatives

$$J(\mathbf{M})_{\mu\nu} = \sum_{\kappa\lambda} M_{\kappa\lambda}(\mu\nu|\kappa\lambda) , \quad (2.117)$$

$$K(\mathbf{M})_{\mu\nu} = \sum_{\kappa\lambda} M_{\kappa\lambda}(\mu\kappa|\nu\lambda) , \quad (2.118)$$

$$J^x(\mathbf{N}, \mathbf{M}) = \sum_{\mu\nu\kappa\lambda} N_{\mu\nu} M_{\kappa\lambda}(\mu\nu|\kappa\lambda)^x , \quad (2.119)$$

$$K^x(\mathbf{N}, \mathbf{M}) = \sum_{\mu\nu\kappa\lambda} N_{\mu\nu} M_{\kappa\lambda}(\mu\kappa|\nu\lambda)^x . \quad (2.120)$$

\mathbf{M} and \mathbf{N} are arbitrary density matrices. In nearly all calculations, these are the most time-consuming steps. For an efficient evaluation of the Coulomb (J) terms, the shellpair data is rearranged before it is sent to the GPUs [58]. The contraction itself is done using the J-engine algorithm [122, 123], which, although designed for CPUs, is highly suitable for massively parallel architectures. In case of the exchange (K) terms, the shellpair-rearrangement can be combined with a pre-selective screening (preLinK) [54, 55]

$$\sum_{\kappa\lambda} \sqrt{(\mu\kappa|\mu\kappa)} \times |M_{\kappa\lambda}| \times \sqrt{(\nu\lambda|\nu\lambda)} \geq \vartheta_{\text{pre}} , \quad (2.121)$$

$$K(\mathbf{M})_{\mu\nu} \times N_{\mu\nu} \geq \vartheta_{\text{pre}}^{\nabla} . \quad (2.122)$$

Shellpairs ($\mu\nu$) with no significant contribution (i.e., smaller than the given thresholds ϑ_{pre} and $\vartheta_{\text{pre}}^{\nabla}$) are discarded. This screening allows for an unhampered calculation on the GPUs, while leading to an $\mathcal{O}(N)$ scaling behavior with respect the system size.

The use of GPUs significantly reduces the prefactor of these calculations, while additionally showing a strong scaling. The latter means that by increasing the number of GPUs, computation time can be reduced even further. The program package TeraChem is a pioneer in this field, featuring GPU-based energy [59], forces [60], and excited-states calculations [61–64] that have been applied, e.g., to BOMD [60], NAMD [62], and *ab initio* multiple spawning [22, 23]. In this work, the multi-architecture program package FermiONs++ [54–56] with its efficient preLinK scheme and OpenCL support is in focus. In publications **I** and **II** and manuscript **III**, it is used for ground and excited-state MD simulations, showing remarkable speed-ups and a good scalability with the computational resources.

2.5 Molecular Dynamics Applications

As already mentioned in the introduction, AIMD simulations give access to a large number of experimental properties. In this section, three of them are briefly introduced and discussed: The calculation of vibrational spectra, free energy calculations, and the investigation of relaxation pathways.

2.5.1 Simulation of Vibrational Spectra

Vibrational spectra are usually calculated within the harmonic approximation via the second derivative of E_0 with respect to the nuclear coordinates at a minimum energy geometry ($\nabla_A E_0^{\min} = 0$) [124]. For non-linear systems, this yields the $N_F - 6$ eigenmodes (f) of the system, with frequency ν_f and N_F being the number of degrees of freedom of the system. The experimentally observable infrared (IR) and Raman spectra additionally require an intensity $I(\nu_f)$, which is calculated from the change of the dipole and quadrupole moments, respectively, along the eigenmodes.

Alternatively, vibrational spectra can be calculated from MD simulations. IR spectra can, for example, be obtained as the following Fourier transform [6, 32–34]

$$I(\nu) \propto \int dt \langle \dot{\boldsymbol{\mu}}(\tau) \dot{\boldsymbol{\mu}}(t + \tau) \rangle_{\tau} \exp\{-i2\pi\nu t\} . \quad (2.123)$$

$\dot{\boldsymbol{\mu}}$ is the derivative of the dipole moments with respect to the time, $\langle a(\tau)a(t + \tau) \rangle_{\tau}$ is the autocorrelation function. The mass-weighted Fourier transform of the velocity autocorrelation function

$$D(\nu) = 2\beta \sum_A m_A \int dt \langle \mathbf{v}_A(\tau) \mathbf{v}_A(\tau + t) \rangle_{\tau} \exp\{-i2\pi\nu t\} , \quad (2.124)$$

yields the vibrational density of state function ($D(\nu)$), which can be interpreted as a vibrational spectrum [125]. β is equal to $\frac{1}{k_B T}$ with k_B being the Boltzmann constant and T the absolute temperature of the system.

The MD approach to vibrational spectra has some advantages over the standard, derivative-based approach. First of all, it requires only first-order derivatives of E_0 to calculate a spectrum. Secondly, it yields an entire spectrum and not only discrete values for ν_f , featuring band shapes and even anharmonicities of the modes (see fig. (2.5)). This contrasts with derivative-based methods that assume band shapes to be arbitrary Gaussian or Lorentzian functions while anharmonic effects have to be included via higher order derivatives or perturbation theory [126–128]. Finally, the approach gives access to temperature-dependent spectra and even spectra of certain subsystems using Voronoi dipole moments [129] or the density matrix of the subsystem

$$\mathbf{P}_{\text{sub}} = (\mathbf{S}^{-\frac{1}{2}})^T (\mathbf{S}_{\text{sub}}^{\frac{1}{2}})^T \mathbf{P} (\mathbf{S}_{\text{sub}}^{\frac{1}{2}}) (\mathbf{S}^{-\frac{1}{2}}) . \quad (2.125)$$

\mathbf{S}_{sub} is only build from basis functions, which are part of the subsystem. Subsystem spectra can be used to investigate bulk materials (e.g., solvents) or solvated molecules. Publication **I** presents IR spectra of large organic molecules calculated from efficient AIMD simulations, which are in good agreement with experimental data. Even bulk properties of liquid water (the liquid water band) can be repro-

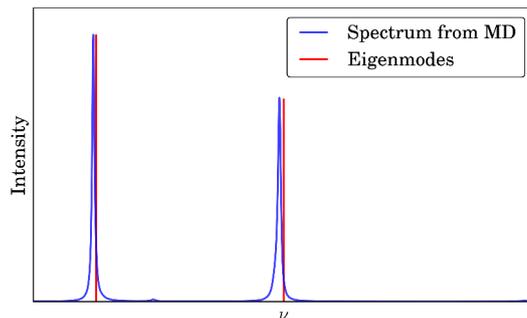


Figure 2.5: Comparison of a spectrum ($D(\nu)$) calculated from MD simulations and eigenmodes obtained from energy derivatives. The eigenmodes are discrete values at ν_f , whereas the MD-based spectrum is (nearly) continuous.

duced. $D(\nu)$ can, additionally, be used in free energy calculations, which are discussed in the next section.

2.5.2 Calculation of Free Energies

Overview over Free Energy Methods

The free energy (A) is one of central quantities linking theoretical calculations and experiments [31, 130–135]. Free energies of reactants, intermediates, transition states, and products give access to reaction pathways, binding energies, rate constants, etc. It consists of a translational, a rotational, and a vibrational contribution

$$A = E_0^{\min} + A_{\text{trans}} + A_{\text{rot}} + A_{\text{vib}} . \quad (2.126)$$

A_{trans} and A_{rot} can be derived solely from the geometry of the system, so that A_{vib} and the electronic energy at the minimum energy geometry (E_0^{\min}) are the targets of quantum-mechanical calculations. For small molecules, the standard approach is to calculate A_{vib} from the eigenmodes (see previous section), known as the normal mode analysis (NMA) [136–138]

$$A_{\text{vib}}^{\text{CL}} = \beta^{-1} \sum_f \ln[\beta h \nu_f] , \quad (2.127)$$

$$A_{\text{vib}}^{\text{QM}} = \beta^{-1} \sum_f \ln \left[\frac{1 - \exp(-\beta h \nu_f)}{\exp(-\frac{1}{2}\beta h \nu_f)} \right] . \quad (2.128)$$

h is the Planck constant. Here, it is assumed that the eigenmodes of the system are classical (CL) or quantum-mechanical (QM) harmonic oscillators. For larger molecules, this ansatz becomes prohibitively expensive, as the time-consuming second-order derivative is required for all conformations of the system. Instead, the free energy is calculated from sampled energies along MD trajectories [66]

$$E_0^{\min} + A_{\text{vib}} = -\beta^{-1} \ln \langle \exp\{-\beta E_0\} \rangle , \quad (2.129)$$

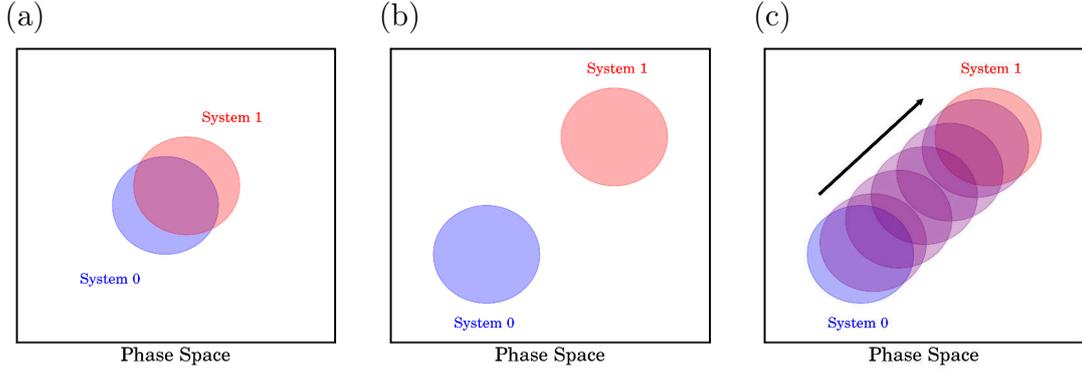


Figure 2.6: Prototypical examples of phase space overlaps of two systems (0 and 1) showing a good overlap in (a) and no overlap in (b). A simulation of 0 will deliver a good estimate for $\Delta A(0 \rightarrow 1)$ in case of (a), but a bad estimate in case of (b). For the latter, it is necessary to divide the overall transformation ($0 \rightarrow 1$) into smaller steps (c) to increase the overlap.

with $\langle x \rangle$ being the ensemble average of the property x . As normally only free energy differences between two systems 0 and 1 are considered, eq. (2.129) can be transformed to

$$\Delta A(0 \rightarrow 1) = A(1) - A(0) = -\beta^{-1} \ln \langle \exp\{-\beta[E_0(1) - E_0(0)]\} \rangle_0, \quad (2.130)$$

with $\langle x \rangle_0$ being the ensemble average over configurations sampled from system 0. This is known as the exponential averaging (EXP) method [66]. The quality of $\Delta A(0 \rightarrow 1)$ mainly depends on the sampling of both PESs ($E_0(1)$ and $E_0(0)$) using a simulation of 0 [31]. This is presented in fig. (2.6), where different cases of phase space distributions of 0 and 1 are discussed. As demonstrated in fig. (2.6c), alchemical transformations ease the sampling and lead to better estimates for $\Delta A(0 \rightarrow 1)$ [66, 130, 139]. The Bennett acceptance ratio (BAR) method [67, 68] is able to determine the "best" free energy from two EXP calculations ($\Delta A(0 \rightarrow 1)$ and $\Delta A(1 \rightarrow 0)$).

In publication **IV**, the different methods presented are compared to the density of states integration (DSI) method, which was introduced by Berens *et al.* in 1983 [65]. Here, the free energy and free energy differences are determined from $D(\nu)$ (see eq. (2.124)) as follows:

$$A = E_0 + \beta^{-1} \int_0^\infty d\nu D(\nu) W_A(\nu), \quad (2.131)$$

$$\Delta A(0 \rightarrow 1) = [E_0(1) - E_0(0)] + \beta^{-1} \int_0^\infty d\nu [D(\nu, 1) - D(\nu, 0)] W_A(\nu). \quad (2.132)$$

W_A is the classical or quantum-mechanical weighting function:

$$W_A^{\text{CL}}(\nu) = \ln[\beta h \nu], \quad (2.133)$$

$$W_A^{\text{QM}}(\nu) = \ln \left[\frac{1 - \exp(-\beta h \nu)}{\exp(-\frac{1}{2} \beta h \nu)} \right]. \quad (2.134)$$

DSI features some advantages over NMA. As it calculates the system as a (nearly infinite) sum of harmonic oscillators instead of $N_F - 6$, it is able to correctly describe anharmonic modes (see fig. (2.5)). Additionally, it only requires first-order energy derivatives. There are also some advantages when comparing to BAR. DSI does not require alchemical transformations, since the free energy difference is calculated from the absolute free energies. Quantum-corrected free energies are accessible without further calculations, which is not the case when applying EXP or BAR. Moreover, contributions of different atoms or groups to the free energy can easily be determined, as illustrated in the next section. Having said that, there are also some drawbacks: it fails (just like NMA) in case the system features non-harmonic modes and the occurrence of low-frequency modes leads to a slow convergence of the method with the simulation time for larger systems. The latter might be the reason why Berens *et al.* [65] solely used the method to determine quantum corrections ($W_A = W_A^{\text{QM}} - W_A^{\text{CL}}$ in eq. (2.131)), which can be neglected for low-frequency modes.

Free Energy Hot-Spots

As the $D(\nu)$ is written as a sum over all atoms

$$D(\nu) = \sum_B D(\nu, B) , \quad (2.135)$$

it is possible to calculate the vibrational free energy contribution of atom B as follows:

$$A_{\text{vib}}(B) = \beta^{-1} \int_0^\infty d\nu D(\nu, B) W_A(\nu) . \quad (2.136)$$

The atomic $A_{\text{vib}}(B)$ can then be summed up to any meaningful ensemble of atoms, e.g., residues or groups. This straightforward approach is not possible using the other free energy methods, as this would require a fragmentation of E_0 or the eigenmodes f [140–142]. In publication **V**, this method is discussed and used to illustrate the anomeric effect and the binding of an inhibitor to a protein. However, great care has to be taken when interpreting these so-called hot-spots, as only the vibrational free energy is considered and slow modes might not be described accurately, as discussed in the previous section. Nevertheless, the method offers some useful insights into the process that causes the free energy to change, which agree with chemical intuition.

2.5.3 Investigation of Relaxation Pathways

One of the main fields of application of NAMD simulations is the investigation of relaxation pathways of excited molecules. As already discussed above, this has to be done by analyzing the entire ensemble of trajectories, as surface hopping is a stochastic algorithm. Useful measures of this are the mean occupancy and the mean population (see, for example ref. [39, 41, 49]). The population of a state is defined by a_{II} (see eq. (2.21)), while the occupation is either 0 or 1, depending on the PES the trajectory is currently propagating on. Especially the latter can be used to determine the lifetime of an excited state. To study the geometric effects of the relaxation, it requires sampling of geometrical features (e.g., bond lengths, dihedrals) along the trajectories.

In case of the rotary molecular machines studied in publication **II** and the rhodopsin protein analyzed in manuscript **III**, the relaxation pathways are straightforward. The initial π - π^* excitation into the S_1 state allows the rotation around a former double bond (marked blue in fig. (2.7)), which is induced by steric repulsion or the environment.

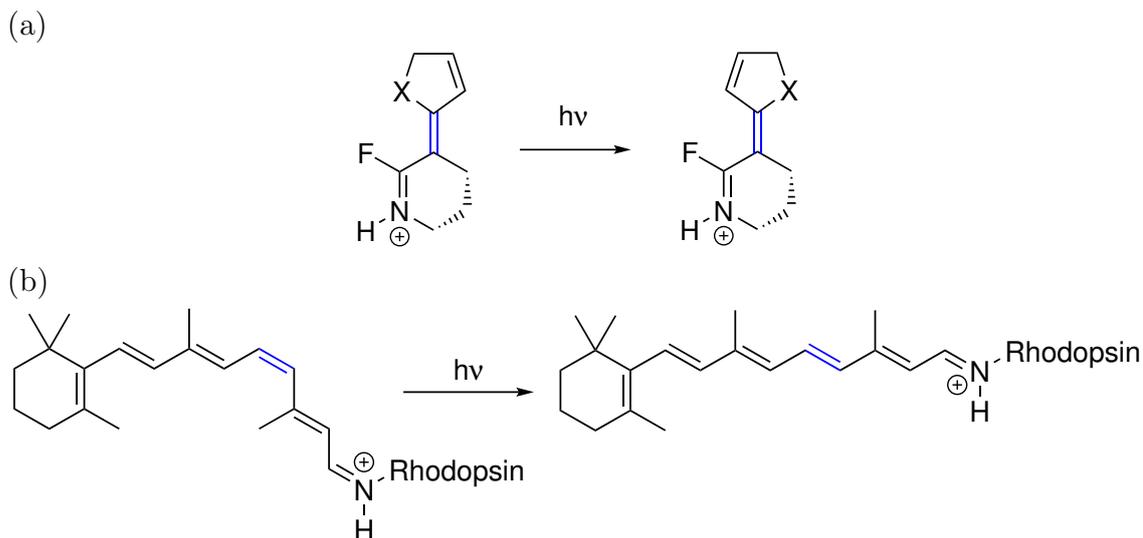


Figure 2.7: Photo-induced relaxation mechanisms studied in publication **II** (a) and manuscript **III** (b). Both molecules show a rotation around a double bond (marked blue). X is equal to either CH_2 , NH , S , or O .

Chapter 3

Publications

3.1 Publication I: Efficient and Accurate Born–Oppenheimer Molecular Dynamics for Large Molecular Systems

L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Efficient and Accurate Born–Oppenheimer Molecular Dynamics for Large
Molecular Systems",
J. Chem. Theory Comput., **13**, 5479 (2017)

Abstract: An efficient scheme for the calculation of Born–Oppenheimer molecular dynamics (BOMD) simulations is introduced. It combines the corrected small basis set Hartree–Fock (HF-3c) method by Sure and Grimme [*J. Comput. Chem.* **2013**, 43, 1672], extended Lagrangian BOMD (XL-BOMD) by Niklasson et al. [*J. Chem. Phys.* **2009**, 130, 214109], and the calculation of the two electron integrals on graphics processing units (GPUs) [*J. Chem. Phys.* **2013**, 138, 134114; *J. Chem. Theory Comput.* **2015**, 11, 918]. To explore the parallel performance of our strong scaling implementation of the method, we present timings and extract, as its validation and first illustrative application, high-quality vibrational spectra from simulated trajectories of β -carotene, paclitaxel, and liquid water (up to 500 atoms). We conclude that the presented BOMD scheme may be used as a cost-efficient and reliable tool for computing vibrational spectra and thermodynamics of large molecular systems including explicit solvent molecules containing 500 atoms and more. Simulating 50 ps of maitotoxin (nearly 500 atoms) employing time steps of 0.5 fs requires 3 weeks on 12 CPUs (Intel Xeon E5 2620 v3) with 24 GPUs (AMD FirePro 3D W8100).

Reprinted with permission from:

L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Efficient and Accurate Born–Oppenheimer Molecular Dynamics for Large
Molecular Systems",
J. Chem. Theory Comput., **13**, 5479 (2017)

Copyright 2017 American Chemical Society.

<https://pubs.acs.org/doi/10.1021/acs.jctc.7b00937>

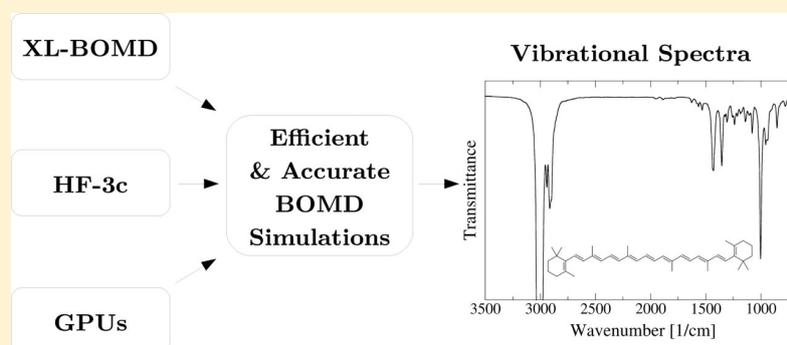
Efficient and Accurate Born–Oppenheimer Molecular Dynamics for Large Molecular Systems

Laurens D. M. Peters,^{†,‡} Jörg Kussmann,^{†,‡} and Christian Ochsenfeld^{*,†,‡,§}

[†]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany

[‡]Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstr. 5-13, D-81377 München, Germany

§ Supporting Information



ABSTRACT: An efficient scheme for the calculation of Born–Oppenheimer molecular dynamics (BOMD) simulations is introduced. It combines the corrected small basis set Hartree–Fock (HF-3c) method by Sure and Grimme [*J. Comput. Chem.* **2013**, *43*, 1672], extended Lagrangian BOMD (XL-BOMD) by Niklasson et al. [*J. Chem. Phys.* **2009**, *130*, 214109], and the calculation of the two electron integrals on graphics processing units (GPUs) [*J. Chem. Phys.* **2013**, *138*, 134114; *J. Chem. Theory Comput.* **2015**, *11*, 918]. To explore the parallel performance of our strong scaling implementation of the method, we present timings and extract, as its validation and first illustrative application, high-quality vibrational spectra from simulated trajectories of β -carotene, paclitaxel, and liquid water (up to 500 atoms). We conclude that the presented BOMD scheme may be used as a cost-efficient and reliable tool for computing vibrational spectra and thermodynamics of large molecular systems including explicit solvent molecules containing 500 atoms and more. Simulating 50 ps of maitotoxin (nearly 500 atoms) employing time steps of 0.5 fs requires ~ 3 weeks on 12 CPUs (Intel Xeon E5 2620 v3) with 24 GPUs (AMD FirePro 3D W8100).

1. INTRODUCTION

The simulation of the time-dependent behavior of molecular systems via ab initio molecular dynamics (AIMD) has become a powerful tool for investigating molecular properties. It can be used not only for sampling potential energy surfaces, but also for the prediction of experimental spectra and thermodynamic properties.^{1,2}

The key assumption of AIMD is the separation of the electronic and nuclear degrees of freedom. The electronic structure is calculated quantum mechanically, whereas the nuclei are treated as classical particles, obeying Newton's equations of motion. In the first efficient and applicable AIMD scheme of Car and Parrinello in 1985,³ the electrons were fictitiously propagated along with the nuclei, keeping the system close to its ground state and avoiding the expensive calculation of the electronic structure. With progresses in the fields of electronic structure theory and computer technology, Born–Oppenheimer molecular dynamics (BOMD)^{4,5} became again more popular, in which the nuclei are moving on the

potential energy surface of the electronic ground state. At every time step of these simulations, the electronic Schrödinger equation is solved, using, for example, Hartree–Fock (HF)⁶ or density functional theory (DFT).^{7,8}

BOMD simulations of large molecular systems are still challenging.⁹ Observables are usually calculated as means or integrals of properties, so that many MD steps are required to obtain accurate results. Consequently, a huge number of minimizations of the electronic wave function and determinations of the gradient of the electronic ground state are required.

The main objective of the present work is to combine three recent developments from the fields of AIMD, electronic structure theory, efficient screening methods, and computer technology to an efficient and accurate BOMD scheme: (1) The corrected small basis set HF (HF-3c) method by Sure and

Received: September 5, 2017

Published: October 25, 2017

Grimme,¹⁰ which is a cost-efficient scheme to obtain reasonably accurate interaction energies and geometries (comparable to large basis set DFT calculations) at the cost of one HF calculation with a minimal basis set, (2) the extended Lagrangian BOMD (XL-BOMD) method by Niklasson et al.,¹¹ which reduces the number of necessary self-consistent field cycles, while still conserving the total energy of the system, and (3) efficient methods for calculating Coulomb and exchange terms on graphics processing units (GPUs) within our FERMIONS++ program package.^{12–14}

We start with a brief review of the three methods and the calculation of vibrational spectra. The performance and the parallel efficiency of the resulting method are analyzed subsequently. To illustrate the new possibilities of our method, we simulate vibrational spectra of several molecular systems ranging from β -carotene and paclitaxel as representatives for biomolecules to a bulk of water molecules containing up to 500 atoms.

2. THEORY

2.1. Corrected Small Basis Set Hartree–Fock Method.

In the corrected small basis set Hartree–Fock (HF-3c) method,¹⁰ three correction terms (including nine empirical parameters) are added to a Hartree–Fock energy calculated with a minimal basis set MINIX.

$$E_{\text{total}}^{\text{HF-3c}} = E_{\text{total}}^{\text{HF/MINIX}} + E_{\text{disp}}^{\text{D3(BJ)}} + E_{\text{BSSE}}^{\text{gCP}} + E_{\text{SRB}} \quad (1)$$

The first term introduces the dispersion energy, using the D3 correction scheme¹⁵ with Becke–Johnson damping.^{16,17} The second term is the geometrical counterpoise correction for the basis set superposition error (BSSE),¹⁸ whereas the last one is a short-ranged term, which tackles the bond length errors of the small basis set. As the HF-3c method delivers good geometries, vibrational frequencies, and interaction energies of large molecular systems,¹⁰ we expect it to yield reasonably accurate potential energy surfaces for molecular dynamics trajectories, as illustrated later in this work in Section 5.

2.2. Extended Lagrangian Born–Oppenheimer Molecular Dynamics. Every step of a standard Born–Oppenheimer molecular dynamics (BOMD) simulation comprises the calculation of the ground-state energy and its derivative with respect to the nuclear coordinates, using, for example, self-consistent field (SCF) methods. To reduce the number of necessary SCF cycles, it is common practice to use a linear combination of converged densities (\mathbf{P}) of previous time steps as a guess for the SCF procedure.^{19,20} Under incomplete SCF convergence (which is always the case as a certain convergence criteria is introduced), these algorithms are not time-reversible and errors within the calculation of \mathbf{P} are propagated throughout the trajectory.²¹ Both major shortcomings have been tackled by Niklasson and co-workers in a series of publications.^{11,21–26} In the resulting extended Lagrangian BOMD (XL-BOMD) method,¹¹ an auxiliary density (\mathbf{P}^{aux}) is propagated (in the spirit of the method of Car and Parrinello³) along with the nuclei and close to the ground-state density.

$$\mathbf{P}^{\text{aux}}(t + \delta t) = 2\mathbf{P}^{\text{aux}}(t) - \mathbf{P}^{\text{aux}}(t - \delta t) + \kappa(\mathbf{P}(t) - \mathbf{P}^{\text{aux}}(t)) \quad (2)$$

When $\mathbf{P}^{\text{aux}}(t)$ is used as an initial guess for the SCF calculation at time t , the overall MD scheme becomes time-

reversible. In order to cancel out the error propagation, a dissipative force term is added to the right-hand side of eq 2.¹¹

$$\mathbf{F}^{\text{diss}} = \alpha \sum_{k=0}^K c_k \mathbf{P}^{\text{aux}}(t - k\delta t) \quad (3)$$

For optimized values of α , κ , and c_k for different orders K , the reader is referred to ref 11. As a consequence, XL-BOMD simulations are (even if the number of SCF cycles is kept constant to three or four) energy-conserving and the resulting trajectories are very similar to those obtained from fully converged time-reversible BOMD simulations.^{24,25} Reducing the number of Fock matrix builds and diagonalizations per step leads to a significant acceleration of the calculation, enabling accurate BOMD simulations of large molecular systems.

2.3. Graphics Processing Units. Using graphics processing units (GPUs) in addition to central processing units (CPUs) has provided a major leap in the performance of quantum chemical calculations throughout the past decade.^{12–14,27–33} Key in this area is the efficient evaluation and contraction of the two-electron integrals, for which a rearrangement of the shell-pair data is necessary.²⁹ This leads, in combination with the J-engine^{34,35} for the Coulomb terms and a preselective screening method (PreLinK) for the exchange calculation,^{12,14} to a large speedup of self-consistent field and gradient calculations of both small and large molecular systems, particularly for small l -quantum numbers. The integral routines show a good parallel efficiency (strong scaling). GPUs have also been used successfully for accelerating BOMD simulations.³¹

2.4. Vibrational Spectra. In modern quantum chemistry, vibrational spectra (and thermodynamics) are usually obtained from vibrational frequency calculations, which are calculated as the second derivative of the energy with respect to the nuclear coordinates at a minimum energy geometry.³⁶ While even linear-scaling methods are available for large systems,^{37–39} the approach assumes the potential around the minimum structure to be harmonic, so that anharmonic effects must be included via scaling factors,^{40,41} vibrational self-consistent field,⁴² or vibrational perturbation theory.^{43,44}

Alternatively, vibrational spectra can also be extracted from ab initio molecular dynamics (AIMD) simulations.^{2,45–47} Infrared (IR) spectra are, for example, obtained as the Fourier transform of the autocorrelation of the time derivative of the dipole moment ($\dot{\boldsymbol{\mu}}$):

$$A(\omega) \propto \int \langle \dot{\boldsymbol{\mu}}(\tau) \dot{\boldsymbol{\mu}}(t + \tau) \rangle_{\tau} e^{-i\omega t} dt \quad (4)$$

where $A(\omega)$ denotes the intensity at frequency ω .⁴⁷ The presented approach has been used successfully to predict IR spectra of small molecules using density functional theory⁴⁷ (for further examples, the reader is referred to ref 2). It features three advantages for the calculation of spectra:

- (1) the anharmonicity of the vibrations is taken into account intrinsically, since they are determined using the calculated, nonharmonic potential energy surface;
- (2) the method requires only first-order derivatives (nuclear gradients), facilitating the applicability to large molecular systems and even to excited states,⁴⁸ and
- (3) influences of temperature⁴⁷ and solvents⁴⁹ either via continuum models or explicit solvent molecules can be included.

Table 1. XL-BOMD Timings of β -Carotene, Paclitaxel, and Maitotoxin at the HF-3c Level of Theory (Three SCF Cycles Per Step) Calculated on up to Six Nodes^a

number of nodes	XL-BOMD Timings (s)								
	β -Carotene			Paclitaxel			Maitotoxin		
	SCF	forces	step	SCF	forces	step	SCF	forces	step
1	1.00	1.29	2.58	4.65	7.62	12.7	37.5	31.7	70.9
2	0.69	0.89	1.88	2.65	4.14	7.22	22.7	18.0	42.7
3	0.58	0.69	1.56	1.98	3.06	5.48	16.7	12.8	31.4
4	0.53	0.59	1.42	1.73	2.42	4.57	14.3	10.3	26.6
6	0.47	0.51	1.28	1.41	1.91	3.75	10.4	7.38	19.6

^aEach node contains two Intel Xeon E5 2620 v3 (12 threads) CPUs and four AMD FirePro 3D W8100 GPUs. The systems consist of 96 atoms for β -carotene, 113 atoms for paclitaxel, and 492 atoms for maitotoxin.

In order to obtain an IR spectrum of an explicitly solvated molecule, its dipole moment must be determined from the electron density of the entire system (\mathbf{P}). This property can be calculated approximately from the density matrix of the solvated molecule (\mathbf{P}^{Sub}), which is formed via a Löwdin-like projection⁵⁰ of \mathbf{P} .

$$\mathbf{P}^{\text{Sub}} = (\mathbf{S}^{-1/2})^T (\mathbf{S}_{\text{Sub}}^{-1/2})^T \mathbf{P} \mathbf{S}_{\text{Sub}}^{-1/2} \mathbf{S}^{-1/2} \quad (5)$$

\mathbf{S}_{Sub} is built from the overlap matrix \mathbf{S} , using only basis functions located on the solvated molecule.

3. COMPUTATIONAL AND EXPERIMENTAL DETAILS

All calculations have been performed using the FERMIONS++ program package.^{12,14} FERMIONS++ was compiled using the GNU C++ compiler v4.8 with “-O3”, the Intel Math Kernel Library (MKL), and MVAPICH2 for parallel calculations. Routines for the calculation on graphics processing units (GPUs) have been compiled with the Nvidia Cuda compiler (in the case of Nvidia GPUs) or with the OpenCL C compiler (in the case of AMD GPUs). In addition, gCP v2.02,¹⁸ DFTD3 v3.1,^{15,16} and the LibXC library v3.0.0⁵¹ were used.

BOMD simulations were calculated with the extended Lagrangian formalism (see eqs 2 and 3 with $K = 9$) and the Velocity Verlet propagator^{52,53} for the movement of the nuclei. Energies and gradients were calculated (if not stated otherwise) at the HF-3c level of theory, performing only three SCF cycles per step (involving three Fock matrix builds and two diagonalizations). The integral and the PreLinK threshold were set to 10^{-10} and 10^{-3} , respectively, which are expected to provide μH accuracy.

Timings have been obtained as averages of 100 XL-BOMD steps. The calculations were performed on 1–6 nodes, each containing two Intel Xeon E5 2620 v3 (12 threads) CPUs and four AMD FirePro 3D W8100 GPUs. The electron–nuclear attraction and the two electron integrals were evaluated exclusively on GPUs. All other operations (including the linear algebra) were performed on CPUs. The only exception is the calculation of the exchange kernels of maitotoxin for which the hybrid CPU/GPU engine³² was used. For all cases, the CPU and GPU batch sizes have been optimized prior to the MD simulation.

Vibrational spectra were calculated by sampling the dipole moments of the system (or, in the case of the water spheres of the central water molecule, using the projection of eq 5) after an equilibration time and applying eq 4. The spectra of β -carotene and paclitaxel were obtained as means of five independent trajectories of 15 ps (including the equilibration time of 100 fs) using step sizes of 0.1 and 0.2 fs, respectively.

For the generation of the liquid water spectra, four different spheres (3 Å, 6 Å, 8 Å, and 10 Å around the central water molecule) were cut out of a TIP3P water box generated with AmberTools.⁵⁴ For each sphere, one trajectory of 20 ps (including the equilibration time of 5 ps) was calculated using a step size of 0.5 fs. The temperature was set to 298 K, using a Nosé–Hoover chain thermostat.^{55–57} In all cases, exponential damping and zero shifting were applied to generate the final spectra.

Experimental spectra of β -carotene (Alfa Aesar, 99%), paclitaxel (Alfa Aesar, 99.5%), and water (deionized) have been measured in the present work as averages of 20 scans with 1 cm^{-1} resolution, using a Thermo Fischer Nicolet 6700 FT-IR apparatus.

4. PERFORMANCE

Three molecular systems have been used to investigate the performance of the new BOMD scheme: β -carotene ($\text{C}_{40}\text{H}_{56}$), paclitaxel ($\text{C}_{47}\text{H}_{51}\text{NO}_{14}$), and maitotoxin ($\text{C}_{164}\text{H}_{258}\text{O}_{68}\text{S}_2$). The sulfate substituents of maitotoxin have been saturated with one proton to obtain an uncharged molecule. The computation times of the self-consistent field calculation, the nuclear forces calculation, and the overall time step during the BOMD simulation of the three example molecules are listed in Table 1. Figure 1 presents the speedup and the parallel efficiency of the SCF and the nuclear forces calculation.

The SCF and forces calculations within the implemented BOMD scheme show a good parallel efficiency for all examples, including medium-sized and large molecular systems. It ranges from 0.35 for the SCF calculation of β -carotene to 0.72 for the forces calculation of maitotoxin, both distributed on six nodes. Here, we want to stress that all calculations have been performed using serial linear algebra routines and that the evaluation of the Coulomb integrals, even on one node, is extremely fast (up to 20 times faster than the evaluation of the exchange integrals). Therefore, the presented parallel efficiency is mainly a result of the strong scaling evaluation of the exchange integrals,^{12,14} which explains (1) its dependency on the system size (larger molecules show a higher efficiency) and (2) the fact that the forces calculations are slightly more efficient than the SCF calculations.

The presented BOMD routine, despite the lack of distributed linear algebra routines, is very efficient. A BOMD simulation of maitotoxin, which is known as the largest, nonbiopolymer natural product,⁵⁸ requires ~ 3 weeks on six nodes, calculating 100 000 time steps (up to 50 ps). An equivalent simulation of β -carotene can be performed within 3 days on one node and 1.5 days on six nodes, indicating that speedups can even be

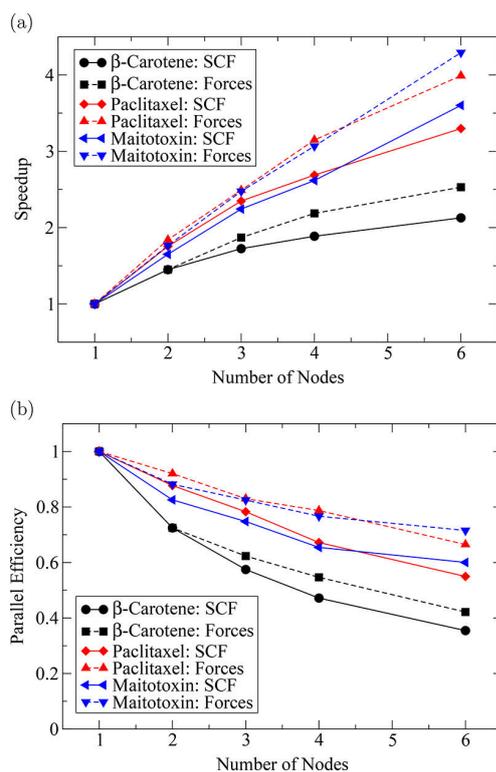


Figure 1. (a) Speedup and (b) parallel efficiency of SCF and forces calculations during the XL-BOMD simulation at the HF-3c level of theory (three SCF cycles per step) of the three test molecules on up to six nodes.

observed for subsecond time steps. We are currently working on an efficient linear algebra parallelization for future work.

5. ILLUSTRATIVE CALCULATIONS

As a first illustrative application, we have calculated XL-BOMD simulations of the natural product β -carotene, the anticancer drug paclitaxel, and a bulk of water molecules. In each step of these simulations, only three SCF cycles were performed, since this does not affect the energy conservation of the simulation (see Figures S1 and S2 in the Supporting Information) and the resulting IR spectrum (see Figure S3 in the Supporting Information), while yielding a speedup of 2–2.5, in comparison to BOMD simulations with full SCF convergence. The extension of the step size from 0.1 fs to 0.5 fs also has no significant effect on the spectrum (see Figures S3, S9, and S10 in the Supporting Information). For proof that the sampling during these simulations is sufficient, the reader is referred to Figures S4–S8 in the Supporting Information, where spectra of the investigated systems are compared for different simulation times and trajectory numbers.

Figure 2 shows the simulated spectra of β -carotene and paclitaxel, together with experimental data (see the previous section for experimental and computational details). The calculated spectra are in good agreement with the experimental spectra, when a scaling factor of $\gamma = 0.81$ is applied. The difference of γ to the reported scaling factor of harmonic vibrational frequencies at the HF-3c level of theory (0.86)¹⁰ may be due to the different approach for obtaining the spectrum (via an MD simulation) or due to the fact that a larger test set has been used to determine the latter value. Yet,

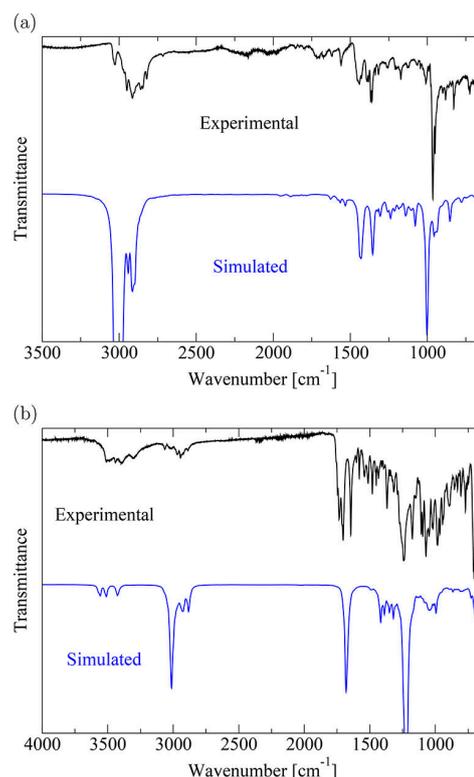


Figure 2. Experimental and simulated IR spectra of (a) β -carotene and (b) paclitaxel obtained from XL-BOMD at the HF-3c level of theory (three SCF cycles per step). The simulated spectra have been scaled with a factor of $\gamma = 0.81$.

there are two differences between the simulated and the experimental data:

- (1) the intensity of the C–H stretching modes (~ 3000 cm⁻¹) is overestimated and
- (2) some deformation vibrations (especially in the case of paclitaxel) do not appear in the simulated spectrum.

The reason for the first shortcoming is that the HF-3c level is less adequate for describing the change of the dipole moment during these vibrations. This is demonstrated in Figure S11 in the Supporting Information, where we compare the IR spectra of ethylene simulated at the HF-3c and B3LYP⁵⁹-D3¹⁵/def-SV(P)⁶⁰ levels of theory. The intensities of the C–H stretching modes are significantly larger at the HF-3c level of theory. The second observation may originate from the fact that we compare experimental solid-state IR measurements with gas-phase calculations. The positions, relative intensities, and shapes of the other peaks are described remarkably well with our simulated spectra.

To reproduce a spectrum of liquid water, four different spheres (3 Å, 6 Å, 8 Å, and 10 Å around the central water molecule containing 5, 41, 92, and 171 water molecules, respectively) were simulated (see the previous section for computational details). The resulting IR spectra are compared to an experimental spectrum of liquid water in Figure 3a (also measured experimentally in the present work). The IR spectrum of the central water molecule changes significantly, when the size of the water sphere increases. The peak of the bending vibration (~ 1635 cm⁻¹) is blue-shifted, while the two peaks of the stretching vibrations merge into one red-shifted

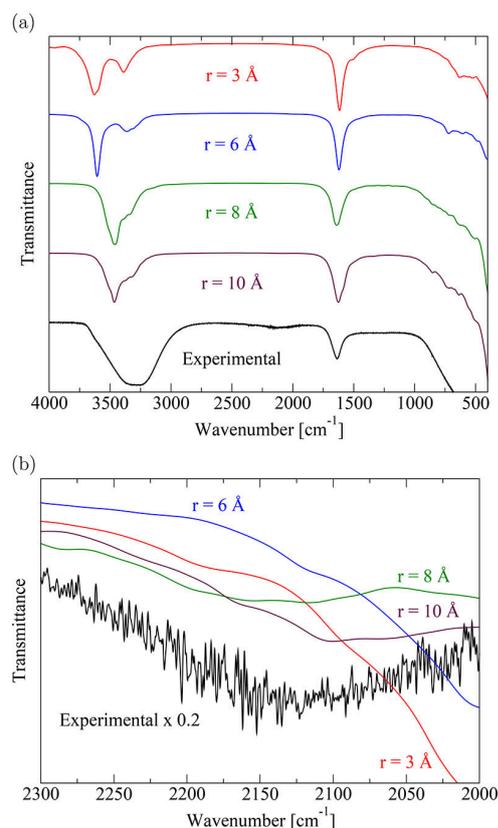


Figure 3. (a) Experimental IR spectrum of liquid water and simulated IR spectra of the central water molecule in four different water spheres (with radii of 3 Å, 6 Å, 8 Å, and 10 Å) obtained from XL-BOMD at the HF-3c level of theory (three SCF cycles per step). The simulated spectra have been scaled with a factor of $\gamma = 0.81$. (b) Expanded view of the region described in panel (a), to illustrate the appearance of the association band at 2150 cm^{-1} . [Note that the transmittance of the experimental spectrum has been scaled with 0.2.]

peak with a large intensity. Both effects can be explained by the larger number of hydrogen bonds in the simulated system. The spectrum of the largest sphere, as a result of this, is in good agreement with the experimental data. Even the liquid water or association band⁶¹ at 2150 cm^{-1} (see Figure 3b) and the rotations at 650 cm^{-1} are visible. The only shortcoming is the bad description of the low-frequency region of the O–H stretching mode ($\sim 3000\text{ cm}^{-1}$), which may again be a result of the use of the HF-3c for the calculation of the dipole moments. This shows that the introduced BOMD scheme is (in combination with the projector of eq 5) capable of describing explicit solvent effects on the vibrational spectrum, when the number of explicit solvent molecules is sufficiently large. For this purpose, a quantity of ~ 100 water molecules seems to be sufficient, since the spectrum of the 8 Å water sphere (98 molecules) does not differ from a spectrum obtained from the simulation of a 10 Å water sphere (see Figure 3a).

6. CONCLUSION

We have presented the combined use of the corrected small basis set Hartree–Fock (HF-3c) method and graphics processing units (GPUs) for extended Lagrangian Born–Oppenheimer molecular dynamics (XL-BOMD) simulations. The resulting scheme is efficient and features a good parallel

efficiency (strong scaling), enabling accurate molecular dynamics simulations of large molecular systems at comparably low computational cost. The method has been used successfully to simulate infrared spectra of medium-sized organic molecules, which are in good agreement with experimental data, when a scaling factor $\gamma = 0.81$ is introduced. Since the simulation seems to capture the potential energy surface remarkably well, it may be suitable not only for the prediction of vibrational spectra but also for the calculation of various other properties (e.g., free energies). This includes also the computation of properties of liquids and solvated molecules.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00937.

Additional figures and molecular geometries (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: christian.ochsenfeld@cup.uni-muenchen.de.

ORCID

Christian Ochsenfeld: 0000-0002-4189-6558

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge Sophia Schwarz and Prof. Oliver Trapp for their help with the experimental IR spectra. Financial support was provided by the SFB 749 “Dynamik und Intermediate molekularer Transformationen” (DFG) and the DFG Cluster of Excellence (EXC 114) “Center for Integrative Protein Science Munich” (CIPSM).

■ REFERENCES

- (1) Marx, D.; Hutter, J. In *Modern Methods and Algorithms of Quantum Chemistry-Proceedings*, Second Edition; Grotendorst, J., Ed.; NIC Series, Vol. 3; John von Neumann Institute for Computing (NIC): Jülich, Germany, 2000; pp 329–477.
- (2) Kirchner, B.; di Dio, P. J.; Hutter, J. Real-world predictions from ab initio molecular dynamics simulations. *Top. Curr. Chem.* **2011**, *307*, 109–154.
- (3) Car, R.; Parrinello, M. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
- (4) Leforestier, C. Classical trajectories using the full ab initio potential energy surface $\text{H}^+ + \text{CH}_4 \rightarrow \text{CH}_4 + \text{H}^+$. *J. Chem. Phys.* **1978**, *68*, 4406–4410.
- (5) Warshel, A.; Karplus, M. Semiclassical trajectory approach to photoisomerization. *Chem. Phys. Lett.* **1975**, *32*, 11–17.
- (6) Roothaan, C. C. J. New developments in molecular orbital theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- (7) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864–B871.
- (8) Kohn, W.; Sham, L. J. Quantum density oscillations in an inhomogeneous electron gas. *Phys. Rev.* **1965**, *137*, A1697–A1705.
- (9) Mniszewski, S. M.; Cawkwell, M. J.; Wall, M. E.; Mohd-Yusof, J.; Bock, N.; Germann, T. C.; Niklasson, A. M. N. Efficient parallel linear scaling construction of the density matrix for Born–Oppenheimer molecular dynamics. *J. Chem. Theory Comput.* **2015**, *11*, 4644–4654.
- (10) Sure, R.; Grimme, S. Corrected small basis set Hartree–Fock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- (11) Niklasson, A. M. N.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C. J.; Holmström, E.; Zheng, G.; Weber,

- V. Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009**, *130*, 214109.
- (12) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.
- (13) Maurer, S. A.; Kussmann, J.; Ochsenfeld, C. Communication: A reduced scaling J-engine based reformulation of SOS-MP2 using graphics processing units. *J. Chem. Phys.* **2014**, *141*, 051106.
- (14) Kussmann, J.; Ochsenfeld, C. Preselective screening for linear-scaling exact exchange-gradient calculations for graphics processing units and general strong-scaling massively parallel calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
- (15) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (16) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (17) Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.* **2005**, *123*, 154101.
- (18) Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **2012**, *136*, 154101.
- (19) Pulay, P.; Fogarasi, G. Fock matrix dynamics. *Chem. Phys. Lett.* **2004**, *386*, 272–278.
- (20) Herbert, J. M.; Head-Gordon, M. Accelerated, energy-conserving Born-Oppenheimer molecular dynamics via Fock matrix extrapolation. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3269–3275.
- (21) Niklasson, A. M. N.; Tymczak, C. J.; Challacombe, M. Time-reversible Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.* **2006**, *97*, 123001.
- (22) Niklasson, A. M. N.; Tymczak, C. J.; Challacombe, M. Time-reversible ab initio molecular dynamics. *J. Chem. Phys.* **2007**, *126*, 144103.
- (23) Niklasson, A. M. N. Extended Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.* **2008**, *100*, 123004.
- (24) Cawkwell, M. J.; Niklasson, A. M. N. Energy conserving, linear scaling Born-Oppenheimer molecular dynamics. *J. Chem. Phys.* **2012**, *137*, 134105.
- (25) Souvatzis, P.; Niklasson, A. M. N. Extended Lagrangian Born-Oppenheimer molecular dynamics in the limit of vanishing self-consistent field optimization. *J. Chem. Phys.* **2013**, *139*, 214102.
- (26) Odell, A.; Delin, A.; Johansson, B.; Bock, N.; Challacombe, M.; Niklasson, A. M. N. Higher-order symplectic integration in Born-Oppenheimer molecular dynamics. *J. Chem. Phys.* **2009**, *131*, 244106.
- (27) Yasuda, K. Accelerating density functional calculations with graphics processing unit. *J. Chem. Theory Comput.* **2008**, *4*, 1230–1236.
- (28) Yasuda, K. Two-electron integral evaluation on the graphics processor unit. *J. Comput. Chem.* **2008**, *29*, 334–342.
- (29) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222–231.
- (30) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 2. direct self-consistent-field implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015.
- (31) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.
- (32) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU integral engine for strong-scaling ab initio methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
- (33) Kussmann, J.; Ochsenfeld, C. Employing OpenCL to accelerate ab initio calculations on graphics processing units. *J. Chem. Theory Comput.* **2017**, *13*, 2712–2716.
- (34) Reza Ahmadi, G.; Almlöf, J. The Coulomb operator in a Gaussian product basis. *Chem. Phys. Lett.* **1995**, *246*, 364–370.
- (35) White, C.; Head-Gordon, M. A J matrix engine for density functional theory calculations. *J. Chem. Phys.* **1996**, *104*, 2620–2629.
- (36) Wilson, E. B. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill: New York, 1955.
- (37) Burant, J. C.; Strain, M. C.; Scuseria, G. E.; Frisch, M. J. Analytic energy gradients for the Gaussian very fast multipole method (GvFMM). *Chem. Phys. Lett.* **1996**, *248*, 43–49.
- (38) Burant, J. C.; Strain, M. C.; Scuseria, G. E.; Frisch, M. J. Kohn-Sham analytic energy second derivatives with the Gaussian very fast multipole method (GvFMM). *Chem. Phys. Lett.* **1996**, *258*, 45–52.
- (39) Kussmann, J.; Luenser, A.; Beer, M.; Ochsenfeld, C. A reduced-scaling density matrix-based method for the computation of the vibrational Hessian matrix at the self-consistent field level. *J. Chem. Phys.* **2015**, *142*, 094101.
- (40) Scott, A. P.; Radom, L. Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.* **1996**, *100*, 16502–16513.
- (41) Tantirungrotechai, Y.; Phanasant, K.; Roddecha, S.; Surawatanawong, P.; Sutthikhum, V.; Limtrakul, J. Scaling factors for vibrational frequencies and zero-point vibrational energies of some recently developed exchange-correlation functionals. *J. Mol. Struct.: THEOCHEM* **2006**, *760*, 189–192.
- (42) Hrenar, T.; Werner, H.-J.; Rauhut, G. Accurate calculation of anharmonic vibrational frequencies of medium sized molecules using local coupled cluster methods. *J. Chem. Phys.* **2007**, *126*, 134108.
- (43) Neugebauer, J.; Hess, B. A. Fundamental vibrational frequencies of small polyatomic molecules from density-functional calculations and vibrational perturbation theory. *J. Chem. Phys.* **2003**, *118*, 7215–7225.
- (44) Kumarasiri, M.; Swalina, C.; Hammes-Schiffer, S. Anharmonic effects in ammonium nitrate and hydroxylammonium nitrate clusters. *J. Phys. Chem. B* **2007**, *111*, 4653–4658.
- (45) Futrelle, R.; McGinty, D. Calculation of spectra and correlation functions from molecular dynamics data using the fast Fourier transform. *Chem. Phys. Lett.* **1971**, *12*, 285–287.
- (46) Gaigeot, M.-P.; Martinez, R.; Vuilleumier, R. Infrared spectroscopy in the gas and liquid phase from first principle molecular dynamics simulations: application to small peptides. *Mol. Phys.* **2007**, *105*, 2857.
- (47) Thomas, M.; Brehm, M.; Fligg, R.; Vohringer, P.; Kirchner, B. Computing vibrational spectra from ab initio molecular dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.
- (48) Galindo, J. F.; Fernandez-Alberti, S.; Roitberg, A. E. Electronic excited state specific IR spectra for phenylene ethynylene dendrimer building blocks. *J. Phys. Chem. C* **2013**, *117*, 26517–26528.
- (49) Thomas, M.; Brehm, M.; Kirchner, B. Voronoi dipole moments for the simulation of bulk phase vibrational spectra. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207–3213.
- (50) Löwdin, P.-O. On the nonorthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **1950**, *18*, 365–375.
- (51) Marques, M. A. L.; Oliveira, M. J. T.; Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **2012**, *183*, 2272–2281.
- (52) Verlet, L. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **1967**, *159*, 98–103.
- (53) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.
- (54) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin,

V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; University of California: San Francisco, CA, 2010.

(55) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695–1697.

(56) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519.

(57) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nose-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.

(58) Nicolaou, K. C.; Frederick, M. O. On the structure of maitotoxin. *Angew. Chem., Int. Ed.* **2007**, *46*, 5278–5282.

(59) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(60) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian-basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(61) Giuffrida, S.; Cottone, G.; Cordone, L. The water association band as a marker of hydrogen bonds in trehalose amorphous matrices. *Phys. Chem. Chem. Phys.* **2017**, *19*, 4251–4265.

Efficient and Accurate Born-Oppenheimer Molecular Dynamics for Large Molecular Systems - Supporting Information

Laurens D. M. Peters,^{†,‡} Jörg Kussmann,^{†,‡} and Christian Ochsenfeld^{*,†,‡}

*Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU),
Butenandtstr. 7, D-81377 München, Germany, and Center for Integrated Protein Science
(CIPSM) at the Department of Chemistry, University of Munich (LMU),
Butenandtstr. 5–13, D-81377 München, Germany*

E-mail: christian.ochsenfeld@cup.uni-muenchen.de

*To whom correspondence should be addressed

[†]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany

[‡]Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstr. 5–13, D-81377 München, Germany

1 Supplementary Figures

1.1 Energy Conservation

Energy Conservation: 0.1 fs

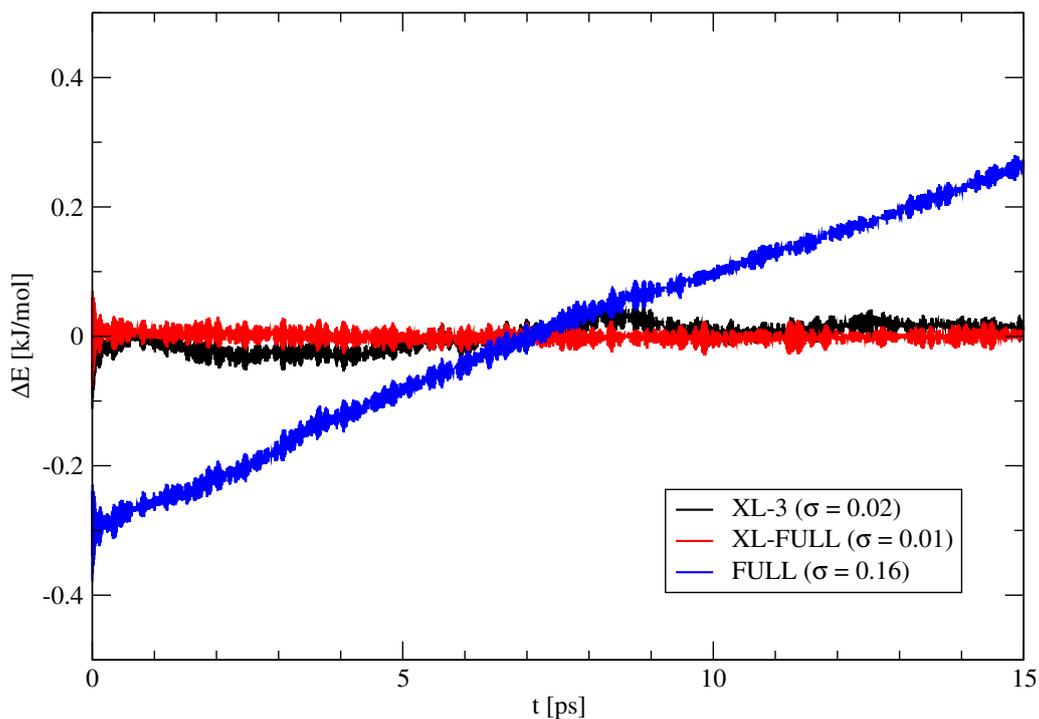


Figure S1: Fluctuation of the total energy during an XL-BOMD simulation performing three SCF cycles per step (XL-3), an XL-BOMD simulation with full SCF convergence (XL-FULL), and a standard BOMD simulation (FULL) of β -carotene (NVE ensemble) at HF-3c level of theory using a step size of 0.1 fs. The average total energy of the individual trajectories is set to zero and the standard deviation σ is given in kJ/mol.

Energy Conservation: 0.5 fs

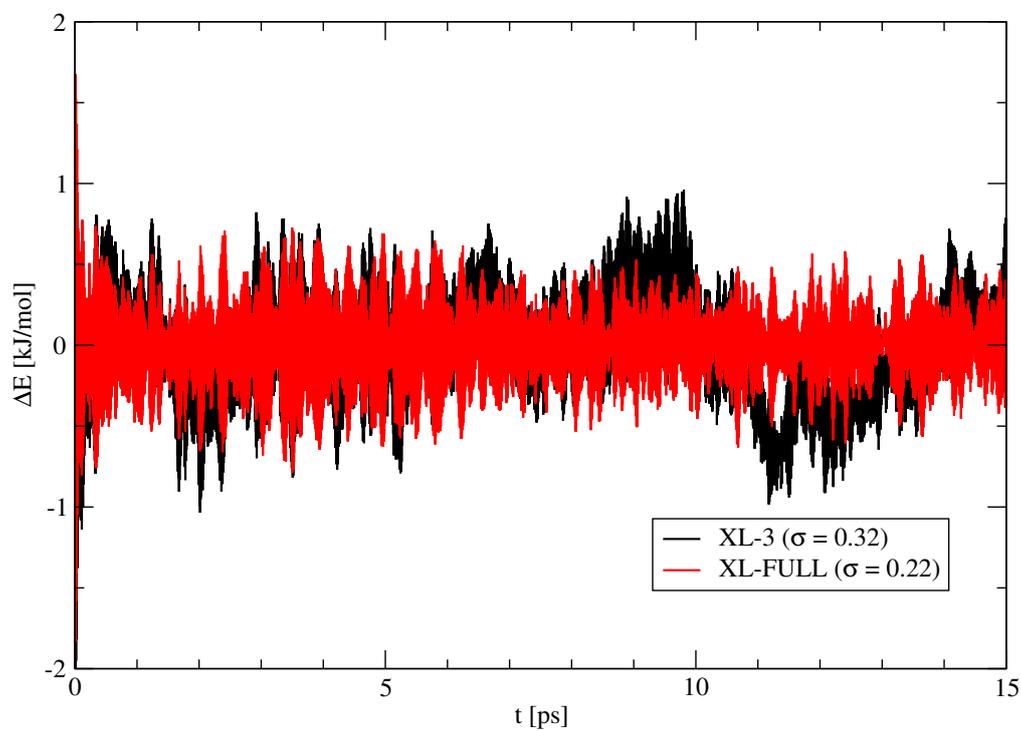


Figure S2: Fluctuation of the total energy during an XL-BOMD simulation performing three SCF cycles per step (XL-3) and with full SCF convergence (XL-FULL) of β -carotene (NVE ensemble) at HF-3c level of theory using a step size of 0.5 fs. The average total energy of the individual trajectories is set to zero and the standard deviation σ is given in kJ/mol.

1.2 β -Carotene

β -Carotene: Spectra Comparison

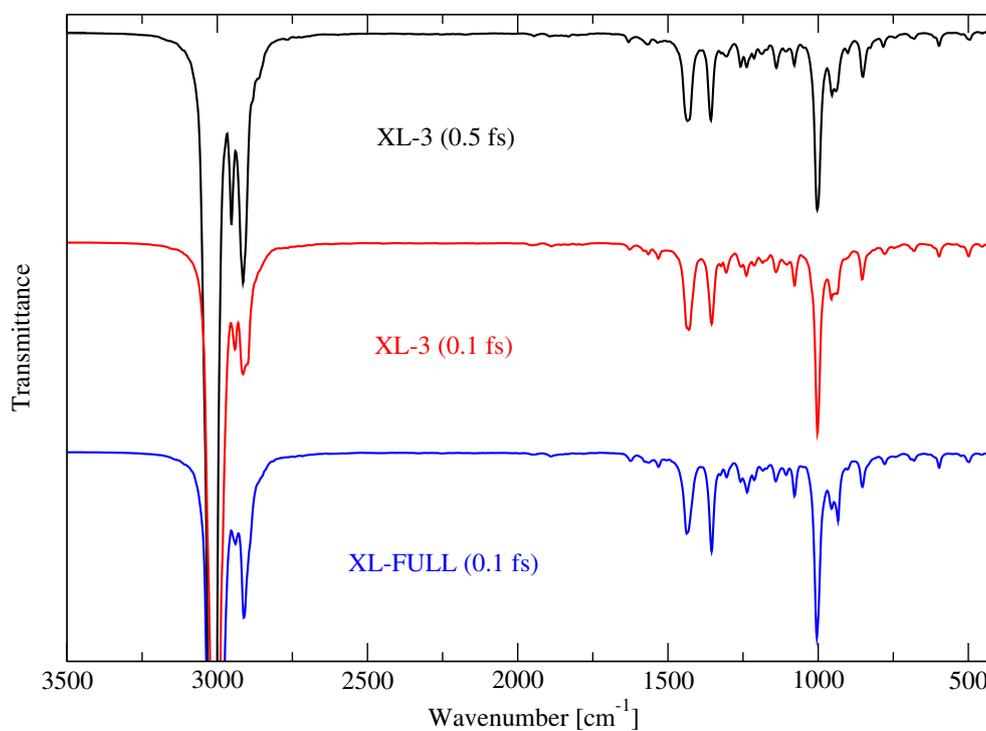


Figure S3: Simulated IR spectra of β -carotene obtained from XL-BOMD at HF-3c level of theory using different step sizes (0.1 fs and 0.5 fs) and convergence criteria (three SCF cycles and full convergence). Five trajectories of 15 ps (including the equilibration time of 100 fs) were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

β -Carotene: Spectra Convergence

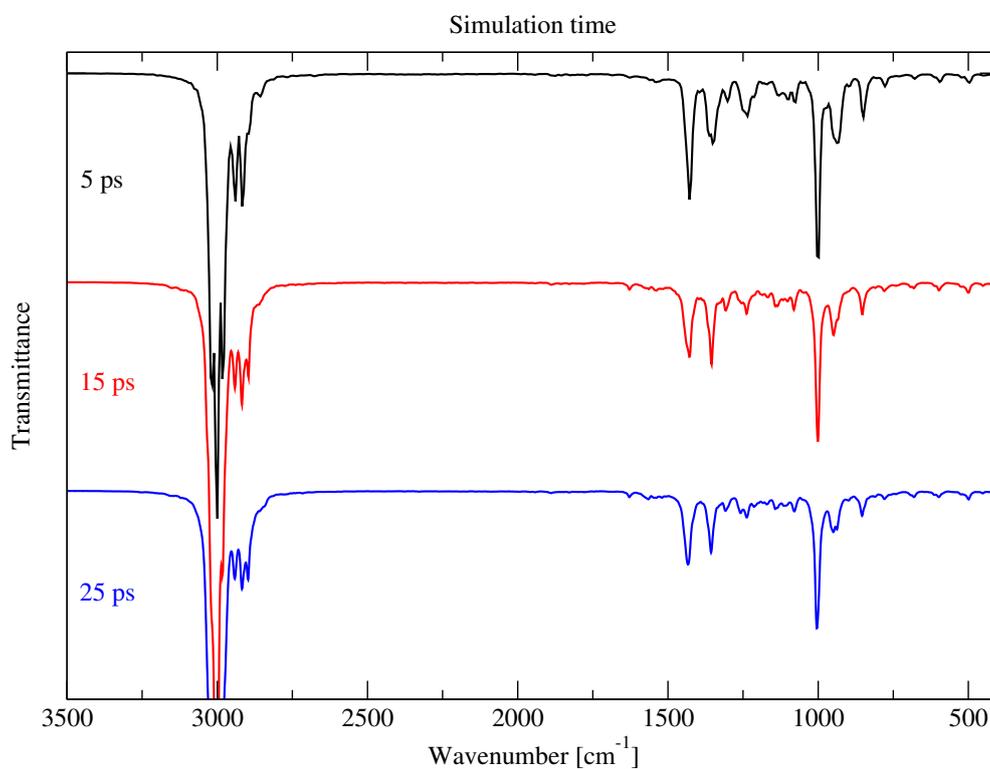


Figure S4: Simulated IR spectra of β -carotene obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) with different simulation times (5 ps, 15 ps, and 25 ps). One trajectory, a step size of 0.1 fs, and an equilibration time of 100 fs were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

β -Carotene: Spectra Convergence

Number of Trajectories

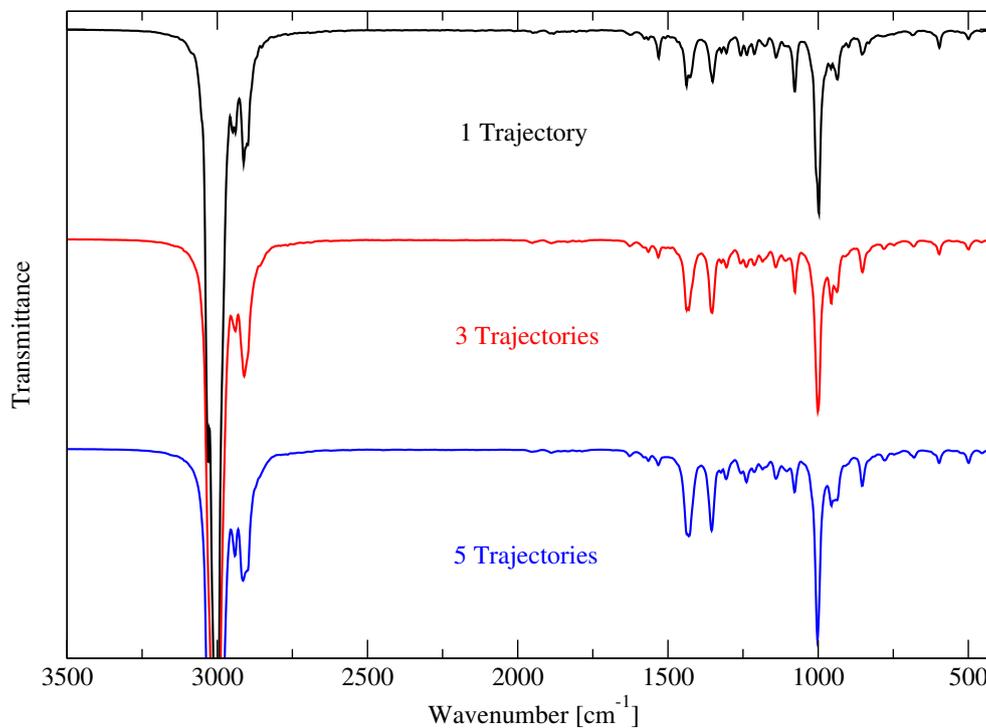


Figure S5: Simulated IR spectra of β -carotene obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) with a different number of trajectories (one, three, and five). A step size of 0.1 fs and a total simulation time of 15 ps (including the equilibration time of 100 fs) were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

1.3 Paclitaxel

Paclitaxel: Spectra Convergence

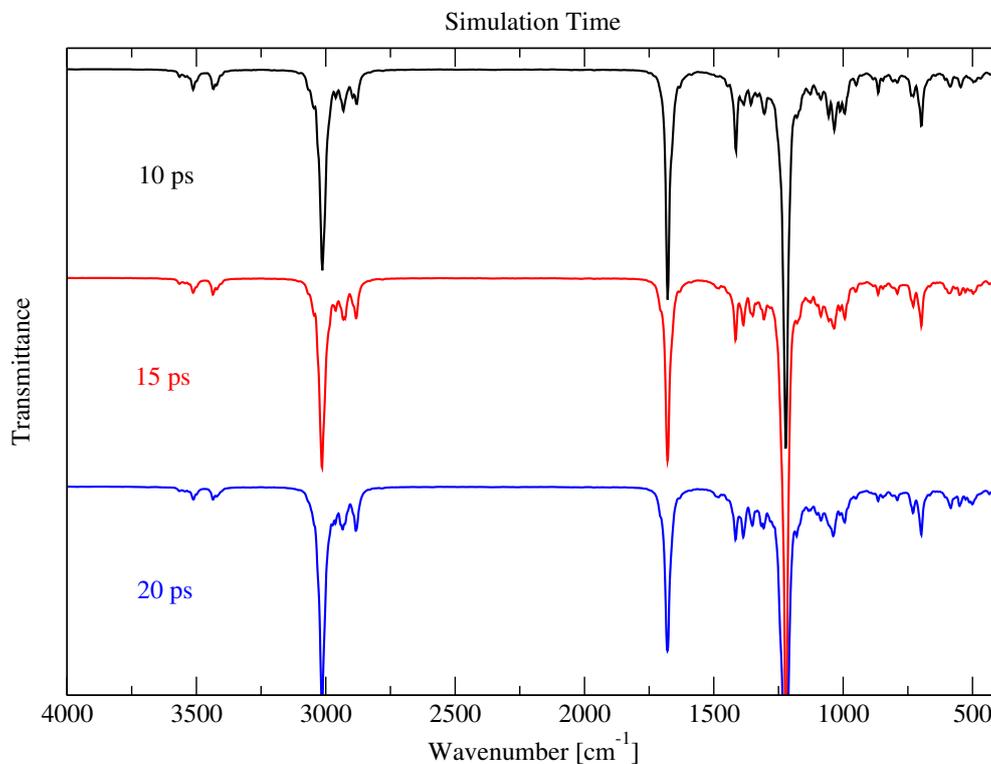


Figure S6: Simulated IR spectra of paclitaxel obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) with different simulation times (10 ps, 15 ps, and 20 ps). One trajectory, a step size of 0.2 fs, and an equilibration time of 100 fs were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

Paclitaxel: Spectra Convergence

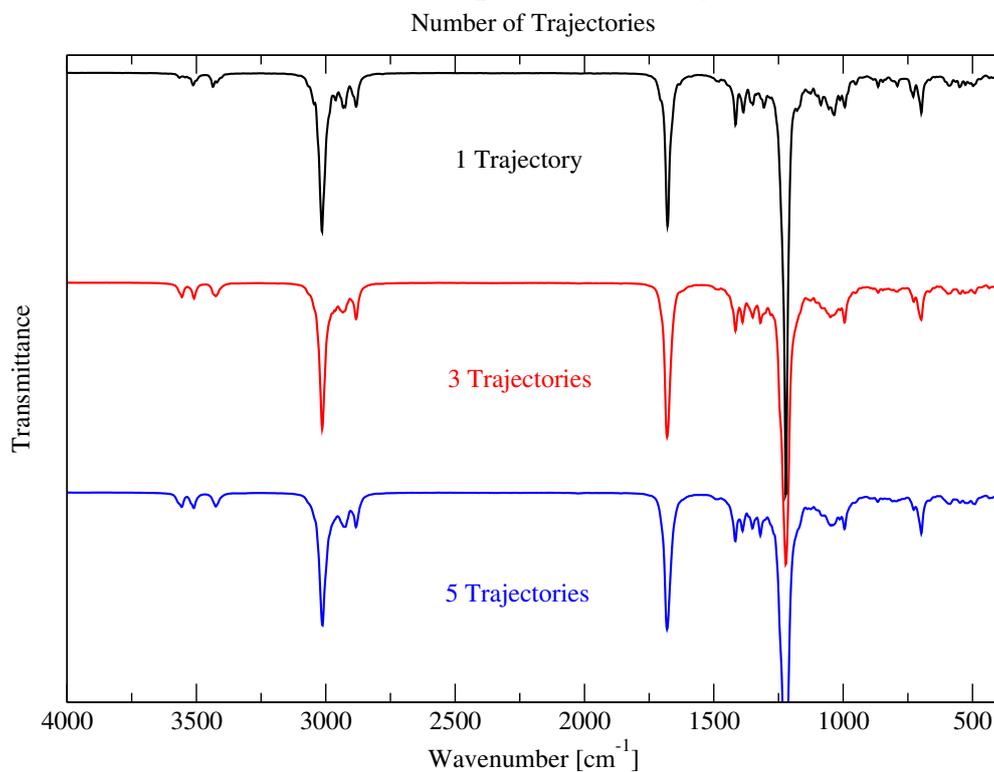


Figure S7: Simulated IR spectra of paclitaxel obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) with a different number of trajectories (one, three, and five). A step size of 0.2 fs and a total simulation time of 15 ps (including the equilibration time of 100 fs) were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

1.4 Liquid Water

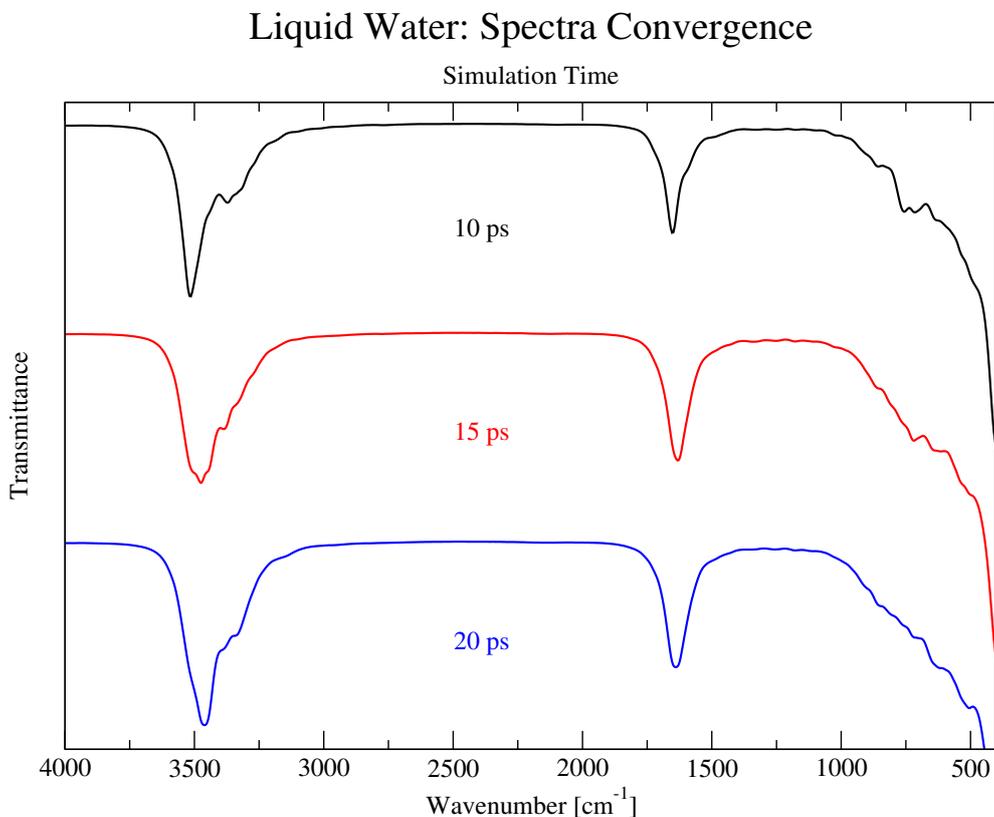


Figure S8: Simulated IR spectra of the central water molecule in a water sphere (with a radius of 8 Å) obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) with different simulation times (10 ps, 15 ps, and 20 ps). One trajectory, a step size of 0.5 fs, and an equilibration time of 5 ps were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

Liquid Water: Spectra Comparison

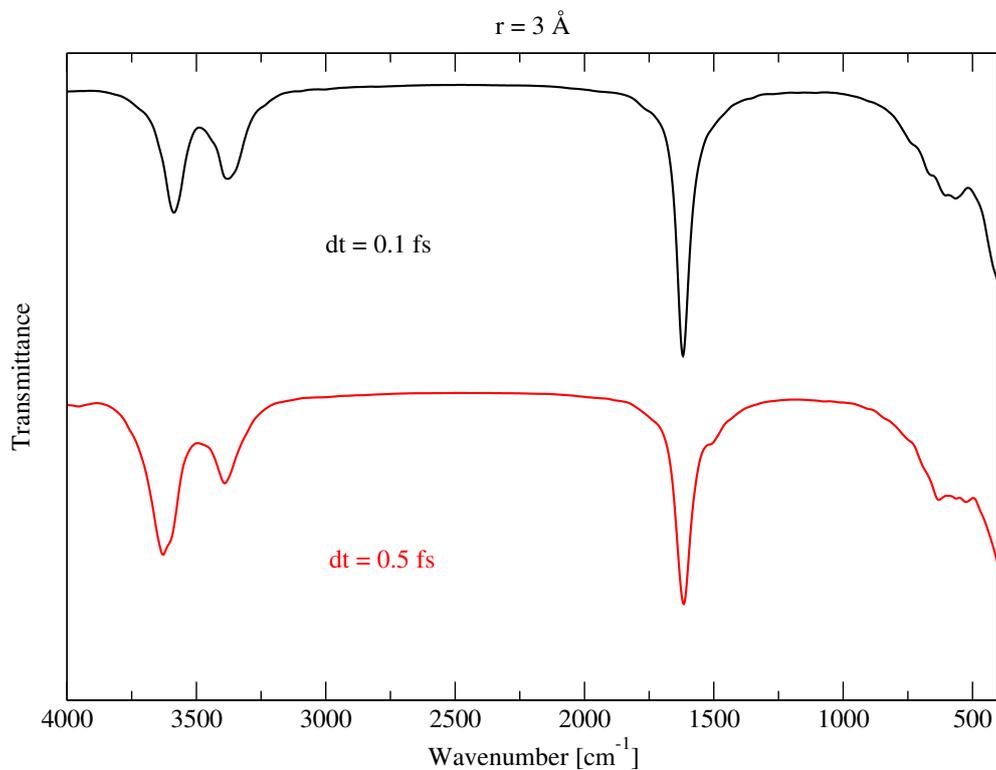


Figure S9: Simulated IR spectra of the central water molecule in a water sphere (with a radius of 3 \AA) obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) using different step sizes (0.1 fs and 0.5 fs). One trajectory of 20 ps (including the equilibration time of 5 ps) and a step size of 0.5 fs were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

Liquid Water: Spectra Comparison

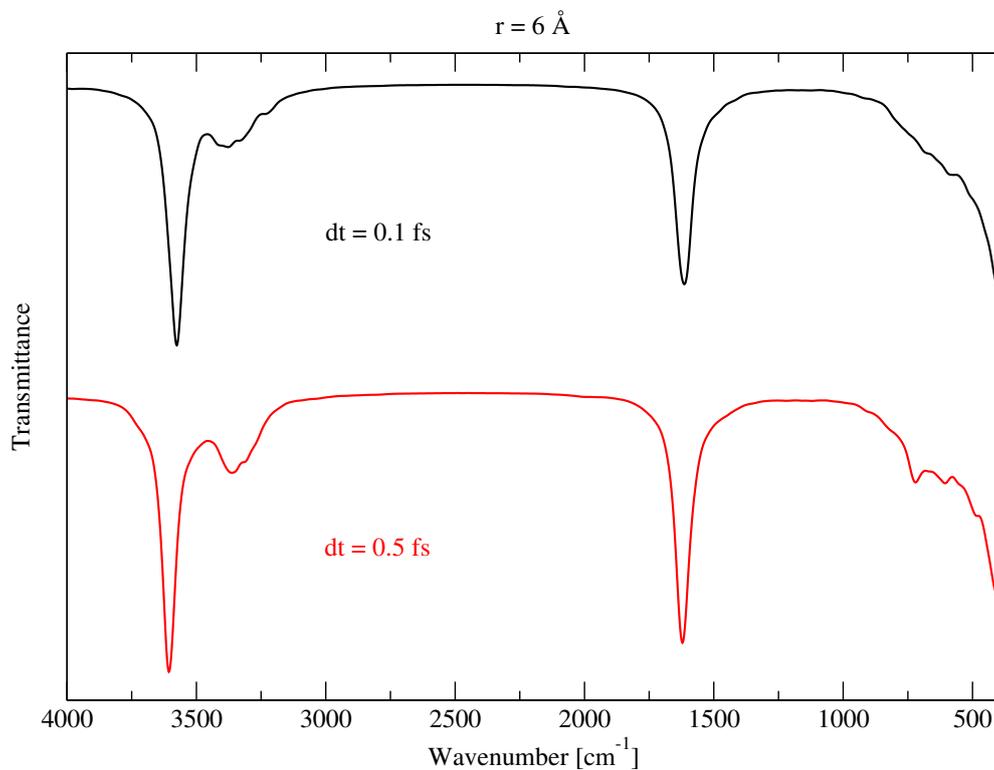


Figure S10: Simulated IR spectra of the central water molecule in a water sphere (with a radius of 6 \AA) obtained from XL-BOMD at HF-3c level of theory (three SCF cycles per step) using different step sizes (0.1 fs and 0.5 fs). One trajectory of 20 ps (including the equilibration time of 5 ps) and a step size of 0.5 fs were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$.

1.5 Ethylene

Ethylene: Spectra Comparison

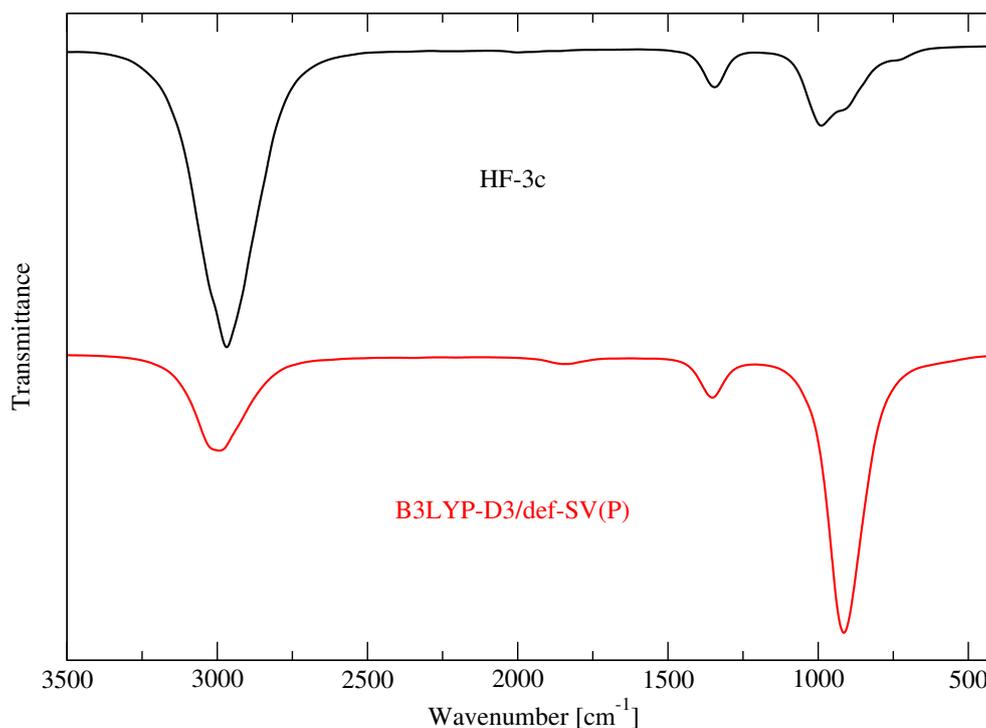
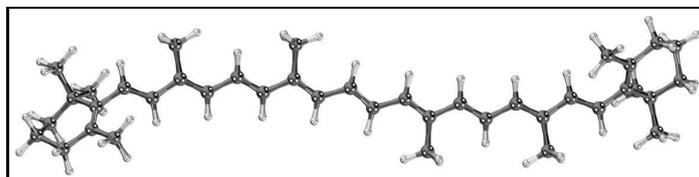


Figure S11: Simulated IR spectra of ethylene obtained from XL-BOMD at HF-3c (three SCF cycles per step) and B3LYP-D3/def-SV(P) (full convergence in every step) level of theory. Five trajectory of 20 ps (including the equilibration time of 100 fs) and a step size of 0.1 fs were used and the temperature of the NVT simulation was set to 298 K. The spectra have been scaled with a factor of $\gamma = 0.81$ (HF-3c) and $\gamma = 0.95$ (B3LYP-D3/def-SV(P)).

2 Initial Structures

2.1 β -Carotene

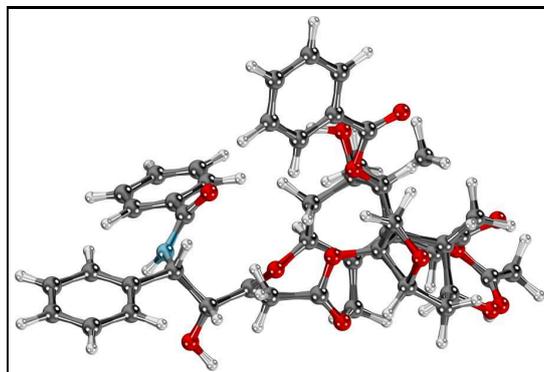


Coordinates [\AA]:

```
C 0.01244966 -0.42928908 0.04938462
C 0.25694583 -0.43389983 1.32333423
H 0.80970298 -0.28283097 -0.63988329
C 1.57539685 -0.22403964 1.93349768
H -0.53786117 -0.60671269 2.01204765
C 1.68442711 -0.25902534 3.23543119
C 2.72185836 0.01995495 0.99057906
C 2.89660959 -0.07528164 4.02087759
H 0.80459132 -0.44013154 3.81072342
C 2.90148529 -0.12864583 5.32390651
H 3.80126783 0.10814699 3.49259898
C 4.08114496 0.04659731 6.17250506
H 1.98854943 -0.31339217 5.84458834
C 3.94591694 -0.03184354 7.47168759
C 5.38490051 0.30881343 5.46917011
C 4.98975573 0.11269254 8.47374789
H 2.97346156 -0.22205664 7.86712681
C 4.75676414 0.01739065 9.75670721
H 5.98242057 0.30408958 8.13710412
C 5.80055182 0.16193509 10.75881521
H 3.76407822 -0.17401305 10.09330911
C 5.66526878 0.08335494 12.05798727
H 6.77300307 0.35224069 10.36343880
C 6.84482933 0.25861318 12.90670809
C 4.36147293 -0.17909237 12.76118244
C 6.84960752 0.20489440 14.20973263
H 7.75777439 0.44348084 12.38612474
C 8.06159452 0.38849397 14.99542290
H 5.94489698 0.02098991 14.73775408
C 8.17036136 0.35382427 16.29743605
H 8.94152889 0.56935303 14.42028468
C 9.48883298 0.56281043 16.90768797
C 7.02338004 0.10971895 17.23969081
C 9.73295867 0.54727418 18.18169284
H 10.28416201 0.72590581 16.21803422
C 11.06120476 0.72819700 18.79547559
H 8.94196899 0.34534635 18.86582590
C -1.32135296 -0.66399516 -0.53784295
C -2.37685793 -0.02855646 -0.13907568
C -3.76393411 -0.30562870 -0.66414710
```

C -2.36366575 1.08122317 0.88255602
C -3.85371127 -1.63138248 -1.40568617
H -4.44964910 -0.28957317 0.15922351
H -4.05406233 0.49629721 -1.31612795
C -2.67740829 -1.73522646 -2.36206214
H -3.82915915 -2.44046371 -0.70735996
H -4.77960143 -1.69562060 -1.93904233
C -1.34655007 -1.74037552 -1.61090859
H -2.74459486 -2.62124339 -2.95665895
H -2.69689570 -0.89818562 -3.02733163
C -1.11863104 -3.10438555 -0.94426923
C -0.23123710 -1.51118562 -2.64039565
C 11.37446890 -0.30533495 19.86546757
C 12.87076420 -0.33186051 20.17067393
C 10.58305861 0.02420473 21.13923389
C 10.96720774 -1.70278826 19.38139431
C 13.41657609 1.07122914 20.37583455
H 13.04111199 -0.93802852 21.03486443
H 13.38287631 -0.78846985 19.35036607
C 13.23375466 1.86966181 19.09215894
H 12.89697005 1.55041198 21.17765351
H 14.45270818 1.03621791 20.64262400
C 11.86950740 1.68728563 18.47063020
H 13.39177486 2.91203223 19.28011332
H 13.97637810 1.57573270 18.37421548
C 11.55890494 2.73926194 17.43400372
H -3.02277306 1.86595334 0.56670389
H -2.71846016 0.72769396 1.83236949
H -1.38409431 1.48594128 1.02025675
H 10.50980932 2.81043603 17.24280101
H 11.91468738 3.69386041 17.76646175
H 12.05415155 2.51703988 16.50721856
H 10.79281381 -0.70154676 21.89765209
H 10.84234964 0.99275939 21.51078026
H 9.53186681 0.01098359 20.94407047
H 11.25698427 -2.43503419 20.10635194
H 9.91017861 -1.77007131 19.23747766
H 11.44970916 -1.93118657 18.45375190
H 0.73393404 -1.68244323 -2.21571765
H -0.35907143 -2.19018780 -3.45777011
H -0.26834710 -0.51146907 -3.02091069
H -0.18915524 -3.10266972 -0.41360827
H -1.90006570 -3.32464231 -0.24830972
H -1.08948182 -3.87731763 -1.68413800
H 2.52815695 0.88722450 0.38847724
H 3.64160598 0.16946145 1.51183264
H 2.84026574 -0.81563215 0.32728943
H 5.61822577 -0.50302096 4.80672053
H 5.31355059 1.20164164 4.87746529
H 6.19220116 0.42449860 6.15837000
H 4.12809036 0.63255104 13.42384101
H 3.55420529 -0.29464451 12.07192569
H 4.43285861 -1.07207309 13.35265609
H 7.20872878 -0.76812420 17.82896602
H 6.91562628 0.93733601 17.91465225
H 6.10103317 -0.02295765 16.71850102

2.2 Paclitaxel



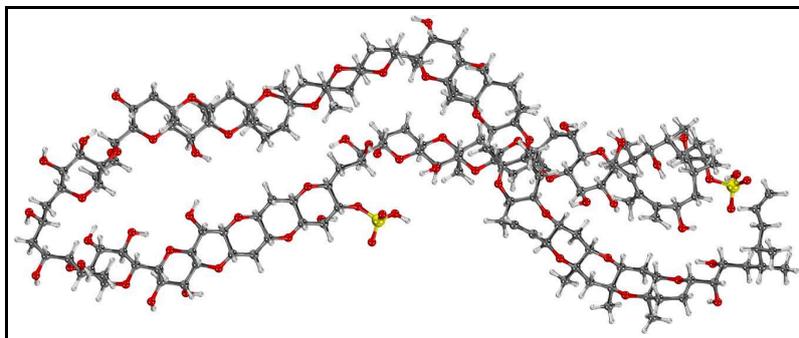
Coordinates [\AA]:

```
O 15.68894 23.55459 9.30853
O 15.81287 29.65719 8.45683
O 13.40468 23.13785 8.04487
O 13.66559 31.43847 8.92988
O 11.84316 25.10056 6.75053
O 11.27957 22.92183 4.48568
O 15.33904 27.93590 11.43104
O 14.89455 24.61996 2.26223
O 16.96517 23.15749 4.02856
O 18.19509 25.47139 4.66759
O 14.27744 27.96112 8.23079
O 14.19171 20.98492 8.01749
O 12.29558 26.86046 5.36613
O 17.56628 25.65774 2.46106
N 14.56699 30.07902 11.30982
C 14.68174 29.22284 8.56648
C 15.45143 24.38602 8.15052
C 13.50110 30.02776 9.09556
C 14.50265 23.55220 7.20651
C 13.97563 24.22371 5.91584
C 13.34539 29.74221 10.59931
C 12.43400 24.18275 5.79537
C 11.84834 24.25174 4.36145
C 15.49332 29.13246 11.64411
C 12.71374 24.40802 3.13631
C 14.15346 24.71850 3.47199
C 14.69562 23.79842 4.58623
C 16.22023 24.00886 4.53087
C 16.81442 25.35470 4.96225
C 16.55976 25.70865 6.39811
C 15.88346 26.84571 6.71155
C 15.33940 27.00498 8.10391
C 14.75248 25.68359 8.61465
C 16.82938 24.68844 7.50189
C 17.81199 25.26213 8.54280
C 17.53019 23.39882 7.06519
C 14.47601 22.30914 4.27596
```

C 15.59993 27.96092 5.75969
C 11.71549 22.82509 5.85431
C 13.43435 21.82880 8.41849
C 12.34736 21.56724 9.38923
C 11.33985 22.48099 9.67784
C 10.33828 22.15278 10.58677
C 10.35451 20.93427 11.24191
C 11.35932 20.02687 10.96377
C 12.34265 20.33518 10.03119
C 12.17664 30.47797 11.22155
C 10.95065 30.57629 10.56476
C 9.87791 31.25309 11.13399
C 10.00793 31.83076 12.37913
C 11.21090 31.74058 13.04578
C 12.27968 31.06453 12.48174
C 11.81973 26.40896 6.38046
C 11.13872 27.18481 7.46687
C 16.71749 29.66167 12.28018
C 16.66427 30.72002 13.18879
C 17.81868 31.13753 13.83690
C 19.03548 30.55095 13.53529
C 19.10009 29.51933 12.61263
C 17.94140 29.06596 11.99062
C 18.42320 25.60374 3.32230
C 19.89529 25.74309 3.07987
H 15.16627 23.88279 10.05888
H 12.61120 29.71350 8.54163
H 15.06941 22.69588 6.86985
H 14.19722 25.27422 6.01922
H 13.14478 28.67651 10.73243
H 14.76787 31.06382 11.43566
H 11.06762 25.00235 4.23170
H 14.21517 25.75778 3.80641
H 15.67245 25.21900 2.28643
H 16.27859 26.06971 4.35425
H 16.16746 27.33282 8.71910
H 11.31303 23.45050 9.20152
H 9.53106 22.84162 10.79023
H 9.57923 20.69386 11.95893
H 11.37607 19.07270 11.47077
H 13.11452 19.60724 9.81988
H 10.82137 30.13843 9.58713
H 8.93238 31.34348 10.61877
H 9.17891 32.35272 12.83531
H 11.31381 32.21170 14.00998
H 13.20619 31.03173 13.03773
H 15.74161 31.22954 13.42928
H 17.77621 31.93087 14.57134
H 19.94155 30.92231 13.99551
H 20.06383 29.08383 12.38096
H 18.00275 28.26321 11.26730
H 12.31093 25.22577 2.52791
H 12.65197 23.52449 2.50430
H 14.68841 25.73408 9.70223
H 13.71185 25.65042 8.29542
H 18.78044 25.48765 8.09474

H 17.96953 24.56421 9.36730
H 18.49572 23.60736 6.60120
H 17.69632 22.73454 7.91451
H 16.95213 22.81777 6.36140
H 14.87233 21.66911 5.06649
H 13.43326 22.04626 4.14831
H 14.97357 22.01508 3.35062
H 16.02277 27.80580 4.77422
H 14.52635 28.09181 5.63565
H 16.04339 28.88920 6.11321
H 10.87952 22.80916 6.55984
H 12.34095 21.95832 6.06644
H 11.10038 28.24006 7.20347
H 10.12125 26.82392 7.61594
H 11.68846 27.07682 8.40192
H 17.48064 26.17473 9.01138
H 20.08808 25.86160 2.01213
H 20.27627 26.62537 3.59684
H 20.41636 24.85112 3.42576
H 13.77125 31.61639 7.98284

2.3 Maitotoxin



Coordinates [Å]:

```
C -17.209000 43.906300 18.540800
C -17.784900 42.918200 19.561500
C -16.743600 42.635000 20.656700
O -15.559700 42.093600 20.036000
C -14.931200 42.948700 19.076000
C -15.918500 43.342600 17.961300
C -17.171600 41.577900 21.683500
O -17.641500 40.404800 21.014200
O -18.985700 43.506200 20.032300
C -19.814900 42.668600 20.827200
C -19.456600 42.819600 22.326200
O -20.515700 42.305801 23.175400
C -21.820100 42.805701 22.889600
C -22.217500 42.409501 21.468600
C -21.256300 43.084501 20.508700
O -23.563700 42.788601 21.168300
C -24.472900 42.116701 22.046700
C -24.280400 42.621301 23.492700
C -22.833700 42.264801 23.904900
C -25.922100 42.266101 21.582000
C -26.821100 41.548001 22.592100
C -26.557100 42.052201 24.012100
O -25.178900 41.920901 24.371800
O -28.195600 41.752901 22.264500
C -29.062601 40.950302 23.082599
C -28.896201 41.315501 24.572900
C -27.424301 41.262501 24.993500
C -30.508100 41.122402 22.568599
O -31.360101 40.312502 23.374299
O -29.397301 42.633301 24.796300
C -24.482900 44.146901 23.637500
O -26.057900 41.633501 20.308000
C -18.206300 41.989200 22.736400
C -19.197400 44.295300 22.708300
C -13.681400 42.210300 18.499800
C -14.023900 40.774200 18.069800
O -15.328800 44.300200 17.088000
O -16.927900 45.155600 19.179800
```

C -30.651200 40.747002 21.077399
O -30.076100 39.451402 20.869099
C -32.087900 40.694402 20.544699
C -32.982400 41.927602 20.744499
C -32.529600 43.179002 19.957399
C -33.578900 44.299202 20.078199
C -34.946000 43.755802 19.663999
C -35.284600 42.515602 20.497699
O -34.291700 41.513702 20.303698
C -36.647800 41.970202 20.079898
C -37.686500 43.082702 20.170499
C -37.223700 44.295702 19.354599
O -35.953200 44.761202 19.814999
O -38.944600 42.609502 19.689999
C -39.969900 43.596102 19.847099
C -39.638200 44.877602 19.073399
C -38.260300 45.409702 19.466700
C -41.304500 43.007202 19.363899
C -42.195000 44.095502 18.742199
C -41.960300 45.427802 19.458400
O -40.613800 45.900602 19.292200
O -43.559500 43.692602 18.792199
C -44.448800 44.638802 18.191300
C -44.357900 45.990401 18.921300
C -42.914600 46.501402 18.937600
C -45.865500 44.008802 18.152699
O -46.699200 44.831702 17.332100
C -48.022800 44.347602 17.112200
C -48.750800 44.041501 18.423799
C -47.909500 43.196601 19.383799
C -46.506000 43.785201 19.533499
O -50.042100 43.469202 18.189899
C -50.055900 42.308502 17.351499
C -49.479700 42.652903 15.972099
C -48.054400 43.176602 16.130699
C -51.498900 41.764302 17.291098
O -52.338100 42.688302 16.592899
O -49.466400 41.492703 15.150098
O -45.738000 42.873601 20.315299
O -48.554500 43.192801 20.659799
O -45.182200 46.953302 18.272101
O -42.850300 47.644401 19.793201
O -41.963100 42.425701 20.493399
O -33.645500 44.753801 21.432600
O -31.284700 43.651002 20.469100
S -30.153400 44.305402 19.475800
O -29.015100 44.531502 20.365000
C -52.136500 41.501301 18.666898
C -51.291000 40.686401 19.653398
O -52.017800 40.564101 20.879298
C -50.959900 39.268601 19.161497
C -49.811200 38.638201 19.977197
O -48.633800 39.403101 19.695698
C -47.439500 38.963201 20.334698
C -47.150700 37.489601 20.046297
C -48.367500 36.595801 20.300597

C -49.601600 37.166601 19.600697
O -45.988500 37.033700 20.742297
C -45.984301 37.251200 22.158097
C -46.149300 38.747499 22.466598
C -47.429500 39.277400 21.828198
C -44.667401 36.673799 22.730297
O -43.560601 37.408899 22.172397
C -42.317701 36.939999 22.694197
C -42.029801 35.512299 22.174596
C -43.130201 34.621299 22.786195
C -44.534201 35.155099 22.462796
O -40.750901 35.049298 22.663796
C -39.670301 35.926698 22.325296
C -39.838101 37.307098 22.984997
C -41.162401 37.925798 22.452498
C -38.309301 35.319498 22.663996
C -37.150701 36.285898 22.276996
C -37.436301 37.657397 22.942697
O -38.743201 38.152798 22.583298
C -36.426500 38.752297 22.606198
C -35.127700 38.727496 23.414598
C -34.426701 37.374096 23.710897
O -33.021300 37.736296 23.741697
C -32.115801 36.646095 23.837396
C -32.213901 35.745696 22.598195
C -33.648501 35.192097 22.592195
C -34.708701 36.313497 22.615496
O -31.296801 34.640996 22.737894
C -29.934501 35.044395 22.915594
C -29.763401 35.940494 24.142195
C -30.691301 37.145895 24.049696
C -29.031502 33.826795 23.077793
C -27.587001 34.281895 23.304493
C -27.480001 35.309394 24.452794
O -28.415601 36.385094 24.229795
C -26.060101 35.965293 24.609194
O -25.817601 36.618595 23.342595
C -24.490101 37.029095 22.995895
C -23.513801 35.873094 23.209893
C -23.555301 35.384393 24.649893
C -24.976901 34.900393 24.945093
O -22.195801 36.238895 22.814294
C -22.046802 36.524096 21.394094
C -23.070602 37.586197 20.940295
C -24.487401 37.315296 21.476895
O -23.137602 37.645898 19.509595
C -22.304902 38.656399 18.953696
C -22.858602 39.069700 17.584097
C -22.035502 40.256801 17.102798
C -20.547903 39.892701 17.000397
O -20.070002 39.376599 18.257996
C -20.799802 38.283599 18.891295
C -19.803502 41.223001 16.735698
C -20.370903 41.959502 15.511599
C -21.907703 42.079903 15.616300
O -22.507903 40.792303 15.865299

C -22.580503 42.722904 14.390801
C -21.992804 42.300606 13.074500
C -20.902704 42.848806 12.511300
C -20.096803 43.994005 13.084201
O -19.442004 44.610806 11.956602
C -18.493703 45.650206 12.261202
C -17.429003 45.057804 13.195301
C -18.052302 44.594503 14.498801
C -19.109903 43.526004 14.189000
O -16.352702 45.983704 13.396202
C -15.637103 46.105305 12.165502
C -16.456803 46.843707 11.085403
C -17.794404 46.055007 10.919803
C -14.232403 46.679304 12.326301
C -13.555303 46.492706 10.970001
C -14.337904 47.170507 9.831402
O -15.729704 46.771408 9.829202
O -12.205103 46.943305 10.964301
C -11.495004 46.424406 9.825600
C -12.178805 46.829308 8.500901
C -13.702905 46.637608 8.525701
C -10.024304 46.894206 9.876400
O -9.972904 48.319506 9.801502
C -14.191203 48.707207 9.832404
C -16.781403 48.320206 11.380705
C -19.235903 46.847805 12.889204
O -19.778702 43.255802 15.430100
C -20.236703 38.929502 15.834895
C -20.459403 36.972399 18.164194
C -20.229901 38.214997 20.332295
C -20.579601 37.029396 21.255194
C -22.215802 35.194397 20.621092
C -24.104001 38.321694 23.738496
C -26.176100 37.011992 25.740995
O -25.043201 34.484692 26.305092
C -31.955501 36.481697 21.265695
C -34.824700 36.883195 25.119796
O -35.968201 35.662797 22.795195
C -37.082501 36.356199 20.732496
C -39.911701 37.233597 24.531397
O -41.057101 38.209799 21.057898
C -42.096601 35.361900 20.636896
O -45.498801 34.436199 23.227395
O -46.207901 38.946699 23.873798
O -50.749100 36.414001 19.988496
O -48.108900 35.293301 19.772596
C -9.254604 46.421504 11.127699
O -9.840103 46.966403 12.313900
C -7.777203 46.845704 11.152999
C -6.896004 46.351705 9.983098
C -5.464504 46.866104 10.198998
C -6.963904 44.812505 9.738696
C -6.189705 44.423706 8.467995
C -6.496404 43.977604 10.952495
C -6.751004 42.506004 10.782593
C -5.796504 41.569904 10.824692

C -12.500600 42.225300 19.513400
O -12.144800 43.583400 19.790700
C -11.228600 41.438701 19.081900
C -10.145800 41.544801 20.165299
C -10.678100 41.890601 17.709500
C -9.555500 40.985601 17.175600
C -9.291900 41.150001 15.659400
O -9.202701 42.538701 15.346000
S -7.767301 43.333101 15.360999
O -6.894201 42.545601 14.492799
C -10.403500 40.495501 14.799500
O -10.490600 39.111801 15.163900
C -10.240101 40.579001 13.256100
C -8.944601 39.907801 12.778299
C -10.342601 41.993701 12.649400
C -11.613001 42.785101 13.011100
O -11.470201 44.118801 12.514500
C -12.888201 42.196101 12.437200
C -13.145001 42.189501 11.116600
C -13.863601 41.662001 13.391100
C -14.440401 42.672100 14.342701
C -14.200401 40.357001 13.372200
C -15.157000 39.654500 14.282600
O -14.468200 38.585100 14.918800
O -7.370201 43.347901 16.767299
O -8.108201 44.649601 14.823499
O -29.919200 43.283603 18.456700
O -30.770800 45.530903 18.973400
H -11.497604 45.328406 9.897099
H -11.966504 47.879708 8.272902
H -11.749705 46.251909 7.673200
H -14.159305 47.116910 7.650502
H -13.932205 45.568109 8.421800
H -13.543304 45.405506 10.798300
H -15.460604 45.080905 11.799700
H -13.681602 46.135703 13.103701
H -14.229502 47.721704 12.657603
H -18.490304 46.604008 10.271303
H -17.583204 45.136208 10.352002
H -17.018903 44.144205 12.737200
H -20.803903 44.717605 13.499602
H -18.564703 42.631804 13.854199
H -17.284002 44.174602 15.158200
H -18.475102 45.428103 15.072102
H -9.523505 46.511107 8.980999
H -9.322104 45.332404 11.212198
H -7.727703 47.942904 11.196100
H -7.341103 46.522102 12.107898
H -7.262604 46.841806 9.071299
H -8.012805 44.558905 9.536696
H -5.432004 44.157303 11.144495
H -7.035204 44.290203 11.853296
H -7.780604 42.194004 10.629593
H -6.049205 40.522304 10.702991
H -4.755304 41.828103 10.982492
H -6.428905 43.400207 8.160694

H -6.455105 45.080607 7.632796
H -5.107005 44.478806 8.618795
H -5.462404 47.956404 10.309099
H -5.013403 46.443503 11.102097
H -4.818404 46.630805 9.348497
H -10.762503 46.663003 12.355300
H -10.417203 48.665405 10.594902
H -14.857703 49.157208 9.088305
H -14.434003 49.149806 10.799104
H -13.173403 49.042307 9.615604
H -16.081902 48.767305 12.090605
H -16.768803 48.916107 10.461905
H -17.770402 48.488507 11.796205
H -18.553402 47.547704 13.377705
H -19.823103 47.378206 12.131305
H -19.952902 46.558504 13.661004
H -20.552104 42.419007 11.574900
H -22.469404 41.462606 12.571899
H -22.527303 43.813704 14.484002
H -23.650103 42.478205 14.381401
H -22.159102 42.732102 16.465301
H -20.098303 41.361703 14.637498
H -22.165902 41.033000 17.871899
H -18.725803 41.053001 16.612897
H -19.880102 41.872700 17.616999
H -23.914402 39.355701 17.673398
H -22.857303 38.241001 16.868596
H -22.408202 39.560198 19.572197
H -20.463503 39.350304 14.852396
H -19.178403 38.645502 15.845695
H -20.817503 38.008602 15.881494
H -19.519003 37.057800 17.609493
H -20.310002 36.126599 18.833493
H -21.246803 36.672800 17.467494
H -22.780301 38.563796 21.341696
H -20.221801 37.323795 22.253394
H -19.929301 36.182696 20.997993
H -19.133001 38.249897 20.264494
H -20.486901 39.147497 20.854596
H -25.142501 38.165996 21.243897
H -24.943401 36.476197 20.933295
H -23.838302 35.010395 22.616793
H -25.123802 34.004693 24.333792
H -22.842101 34.562092 24.792092
H -23.227101 36.151092 25.360094
H -23.229202 34.790597 20.698092
H -22.027602 35.294898 19.552992
H -21.546802 34.425597 21.024091
H -23.159901 38.748694 23.391696
H -24.888101 39.076294 23.620497
H -23.974400 38.174793 24.809596
H -26.762600 36.627192 26.583295
H -25.203600 37.283792 26.156995
H -26.647500 37.932793 25.384796
H -27.754701 34.817093 25.395193
H -30.020701 35.389194 25.057494

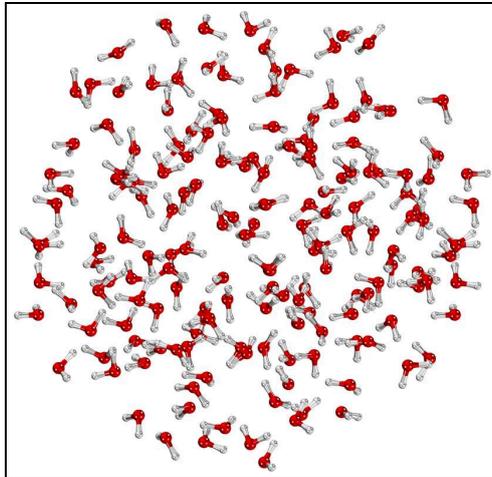
H -29.586801 35.563796 22.014994
H -29.373601 33.212495 23.920292
H -29.095702 33.186396 22.190392
H -27.223201 34.728195 22.370893
H -26.972402 33.399895 23.506492
H -30.626200 37.751794 24.962096
H -30.376201 37.819695 23.245596
H -32.353301 36.059495 24.734095
H -34.661001 36.793098 21.636096
H -33.809502 34.546698 21.717894
H -33.794501 34.516196 23.446494
H -35.308500 39.245296 24.367098
H -34.425600 39.388697 22.886398
H -36.894500 39.732097 22.775299
H -36.193501 38.751598 21.536298
H -37.409200 37.557997 24.027997
H -39.664301 36.040299 21.239796
H -38.192001 34.353698 22.154095
H -38.254201 35.060397 23.729296
H -41.373800 38.886298 22.936698
H -42.437501 36.898498 23.784097
H -44.677801 36.841998 23.815897
H -44.750101 34.938400 21.411896
H -43.031802 33.582399 22.444795
H -42.995501 34.556798 23.875195
H -25.916901 34.085192 26.450892
H -34.400101 35.899995 25.344895
H -34.443600 37.566594 25.887697
H -35.904300 36.794195 25.251396
H -32.640001 37.317798 21.102096
H -30.946401 36.893797 21.199896
H -32.054002 35.785998 20.424295
H -40.270400 38.768799 20.936498
H -40.020600 38.240196 24.951298
H -40.736501 36.615996 24.894797
H -39.012301 36.810296 24.984997
H -41.334201 35.937700 20.109197
H -41.925901 34.316100 20.354795
H -43.058501 35.674701 20.221496
H -36.994001 35.348699 20.309296
H -37.959901 36.839799 20.294097
H -36.214901 36.906099 20.361797
H -46.808301 36.691099 22.616197
H -45.296600 39.314300 22.073598
H -48.299900 38.848600 22.338598
H -47.522000 40.358700 21.988699
H -46.625900 39.550901 19.889298
H -46.904200 37.396601 18.980097
H -45.324401 34.610998 24.167996
H -50.063900 38.719200 21.040997
H -49.489400 37.062401 18.513997
H -48.572100 36.460300 21.367397
H -46.310001 39.900998 24.030399
H -47.299800 34.964301 20.200996
H -50.572200 35.483201 19.768096
H -50.650600 39.297402 18.111398

H -51.857000 38.642601 19.234997
H -50.375100 41.234901 19.889598
H -53.108500 41.010101 18.520698
H -52.401300 42.459401 19.135999
H -52.181500 41.462601 21.213298
H -51.513500 40.838502 16.706398
H -49.438600 41.522902 17.800098
H -52.285100 43.541002 17.057999
H -50.091100 43.419703 15.481699
H -47.399200 42.359602 16.454999
H -47.646900 43.485603 15.159899
H -48.564700 45.168902 16.624800
H -48.945900 44.997101 18.928900
H -45.741500 43.029402 17.672599
H -46.583800 44.723101 20.094700
H -47.848100 42.151201 19.067099
H -49.093100 41.747603 14.289098
H -44.146200 44.789902 17.144800
H -47.985600 42.690800 21.268399
H -44.850400 43.254200 20.427899
H -44.706600 45.897301 19.956000
H -42.615100 46.843902 17.939400
H -42.129400 45.312401 20.537700
H -41.902800 44.199202 17.687599
H -39.608400 44.661602 17.996699
H -40.055700 43.826101 20.918999
H -41.159300 42.196302 18.640699
H -37.981100 46.262202 18.835200
H -38.282000 45.813702 20.487700
H -37.139200 44.025802 18.292399
H -37.788400 43.356201 21.230299
H -34.880400 43.483602 18.600699
H -35.339900 42.764901 21.566099
H -36.936800 41.122001 20.710498
H -36.606399 41.559802 19.062098
H -33.316400 45.171102 19.471400
H -32.432500 42.907803 18.898199
H -33.037800 42.151402 21.816299
H -45.067500 47.797501 18.741601
H -41.920900 47.930201 19.823001
H -41.381400 41.726901 20.839099
H -32.752600 45.051702 21.678400
H -32.594100 39.822102 20.982398
H -32.056500 40.449602 19.473099
H -30.056400 41.437502 20.474899
H -30.816800 42.165702 22.697699
H -28.771000 39.896802 22.970399
H -30.095600 39.272102 19.913499
H -32.280801 40.542802 23.169399
H -29.482301 40.632401 25.197699
H -27.290101 41.669901 26.002800
H -27.081101 40.222301 25.030000
H -26.860700 43.101201 24.103600
H -26.608201 40.471801 22.518300
H -24.246400 41.040201 21.998000
H -26.205200 43.314701 21.442800

H -22.606600 42.644601 24.909500
H -22.724700 41.175801 23.996700
H -21.808900 43.895101 22.994900
H -22.118400 41.323201 21.335400
H -19.734000 41.620600 20.515000
H -21.503400 42.824001 19.471900
H -21.371900 44.175901 20.547300
H -17.674600 42.531000 23.531000
H -18.553800 41.063900 23.219700
H -16.278500 41.248500 22.231300
H -16.465100 43.547900 21.196300
H -18.020800 41.993400 19.021000
H -14.597100 43.859100 19.589200
H -16.183200 42.468400 17.358100
H -17.930100 44.123400 17.744800
H -29.298001 42.828001 25.743700
H -26.994000 41.703801 20.052000
H -16.948000 40.129500 20.390000
H -24.362700 44.446701 24.684900
H -23.784400 44.728901 23.030400
H -25.484400 44.469401 23.340100
H -19.146800 44.409500 23.797400
H -18.263300 44.672300 22.280900
H -20.000600 44.955100 22.366200
H -17.757500 45.463400 19.584000
H -15.156500 45.104900 17.606500
H -13.365200 42.776100 17.616200
H -12.852300 41.801300 20.462300
H -11.502300 40.378601 19.009099
H -11.492000 41.901101 16.982300
H -10.309100 42.921401 17.774800
H -8.635000 41.235901 17.715899
H -9.767500 39.935301 17.404999
H -8.343600 40.643901 15.442199
H -11.368300 40.937001 15.061800
H -11.051501 39.979401 12.820700
H -9.460301 42.586301 12.914900
H -10.275501 41.914201 11.555000
H -14.916900 40.738000 17.441200
H -13.215400 40.330500 17.481100
H -14.199000 40.129300 18.937600
H -9.735900 42.558301 20.227899
H -10.552500 41.280701 21.147399
H -9.317300 40.858601 19.963599
H -11.904400 44.008600 18.950200
H -8.917101 39.863001 11.683999
H -8.058901 40.456802 13.108999
H -8.867200 38.880601 13.148399
H -11.261200 38.736401 14.704500
H -11.687801 42.889801 14.097500
H -10.713101 44.517201 12.977700
H -13.729801 39.709401 12.632400
H -15.572400 40.298200 15.060000
H -15.982000 39.236400 13.697300
H -14.068701 41.785701 10.714200
H -12.435701 42.598301 10.404100

H -14.535201 43.653500 13.866001
H -15.443301 42.400800 14.682901
H -13.794300 42.780400 15.218600
H -15.110100 38.135400 15.494400
H -28.299214 44.914616 19.871816
H -7.983093 43.877328 17.263857

2.4 Liquid Water ($r = 10\text{\AA}$)



Coordinates [\AA]:

```
O -0.184000 -0.030000 -0.122000
H 0.773000 -0.030000 -0.122000
H -0.424000 0.896000 -0.122000
O 2.584000 5.688000 4.201000
H 1.696000 6.030000 4.099000
H 2.868000 5.488000 3.309000
O 4.760000 4.271000 5.235000
H 4.867000 5.172000 5.542000
H 3.849000 4.223000 4.944000
O 2.943000 8.565000 5.184000
H 2.823000 9.291000 4.572000
H 2.633000 7.795000 4.707000
O 5.409000 1.908000 8.566000
H 4.966000 2.689000 8.235000
H 5.968000 2.231000 9.273000
O 8.525000 1.529000 3.447000
H 7.901000 1.525000 4.172000
H 8.438000 0.662000 3.050000
O 0.073000 2.264000 9.147000
H -0.215000 2.835000 9.859000
H 0.818000 1.785000 9.510000
O 0.260000 6.206000 0.357000
H 0.178000 7.150000 0.499000
H -0.091000 6.064000 -0.522000
O 3.910000 4.823000 8.384000
H 3.570000 5.301000 9.140000
H 3.148000 4.370000 8.024000
O 8.122000 5.942000 2.415000
H 8.244000 5.900000 3.363000
H 7.737000 6.804000 2.259000
O 7.134000 8.252000 1.813000
H 6.566000 7.804000 1.187000
H 7.843000 8.611000 1.280000
```

O 1.225000 3.745000 7.300000
H 0.940000 3.380000 8.138000
H 1.163000 3.013000 6.686000
O 0.718000 8.881000 1.260000
H 1.588000 9.163000 1.543000
H 0.227000 9.694000 1.145000
O 3.783000 0.612000 5.853000
H 3.654000 0.604000 6.801000
H 2.959000 0.950000 5.502000
O 5.814000 4.406000 2.663000
H 6.663000 4.843000 2.597000
H 5.447000 4.718000 3.490000
O 3.408000 8.092000 1.820000
H 3.488000 7.145000 1.932000
H 4.300000 8.422000 1.926000
O 5.641000 6.729000 0.057000
H 5.440000 5.866000 0.418000
H 5.470000 6.642000 -0.881000
O 4.439000 3.829000 0.045000
H 4.742000 3.836000 -0.863000
H 5.179000 3.479000 0.541000
O 7.658000 2.773000 1.030000
H 7.310000 2.095000 0.450000
H 7.859000 2.312000 1.845000
O 0.212000 6.511000 3.369000
H -0.691000 6.774000 3.194000
H 0.696000 6.782000 2.589000
O 2.755000 5.568000 1.279000
H 3.464000 5.146000 0.793000
H 2.039000 5.635000 0.647000
O 2.402000 1.622000 0.319000
H 3.125000 2.163000 0.637000
H 2.675000 1.348000 -0.557000
O 1.827000 1.661000 3.912000
H 2.585000 1.774000 3.339000
H 1.365000 2.498000 3.867000
O 4.200000 2.188000 2.453000
H 4.729000 1.400000 2.323000
H 4.832000 2.867000 2.686000
O 6.393000 1.925000 5.507000
H 5.764000 1.224000 5.674000
H 5.854000 2.685000 5.291000
O 1.790000 3.697000 -6.902000
H 1.610000 2.800000 -6.622000
H 1.388000 4.244000 -6.228000
O 2.408000 6.193000 -8.294000
H 3.089000 6.457000 -7.675000
H 2.093000 5.353000 -7.960000
O 4.884000 1.880000 -4.941000
H 4.886000 2.397000 -5.746000
H 4.055000 1.403000 -4.964000
O 4.566000 3.591000 -6.934000
H 4.908000 4.435000 -7.228000
H 3.628000 3.739000 -6.817000
O 7.641000 7.930000 -1.814000
H 7.545000 6.990000 -1.964000

H 8.484000 8.152000 -2.209000
O 2.772000 4.342000 -3.661000
H 2.378000 5.189000 -3.453000
H 2.084000 3.704000 -3.470000
O 9.346000 4.425000 -0.591000
H 8.656000 4.232000 0.044000
H 10.089000 4.709000 -0.058000
O 0.073000 2.264000 -9.628000
H -0.215000 2.835000 -8.915000
H 0.818000 1.785000 -9.264000
O 2.207000 8.882000 -4.138000
H 1.849000 9.547000 -3.551000
H 1.913000 8.050000 -3.767000
O 0.782000 2.371000 -3.461000
H 0.068000 2.571000 -2.856000
H 1.202000 1.595000 -3.089000
O 0.935000 9.913000 -1.719000
H 1.439000 9.334000 -1.147000
H 0.196000 10.197000 -1.181000
O 2.967000 8.056000 -0.824000
H 3.008000 7.758000 0.085000
H 3.822000 8.457000 -0.983000
O 4.908000 6.635000 -2.856000
H 4.378000 6.145000 -3.484000
H 5.577000 7.064000 -3.390000
O 5.395000 3.602000 -2.777000
H 5.301000 2.984000 -3.501000
H 4.815000 4.328000 -3.004000
O 8.161000 2.159000 -5.744000
H 7.713000 2.181000 -6.589000
H 7.528000 1.761000 -5.147000
O 7.671000 5.328000 -2.481000
H 8.261000 5.020000 -1.793000
H 6.853000 4.856000 -2.325000
O 1.588000 6.598000 -2.902000
H 0.693000 6.564000 -2.562000
H 2.127000 6.796000 -2.136000
O 2.092000 0.872000 -8.553000
H 2.739000 0.322000 -8.994000
H 2.512000 1.127000 -7.732000
O 1.845000 1.028000 -5.811000
H 1.496000 1.484000 -5.045000
H 1.514000 0.134000 -5.728000
O 9.357000 2.219000 -2.289000
H 9.296000 2.811000 -1.539000
H 10.297000 2.112000 -2.434000
O 2.517000 0.556000 -2.536000
H 3.065000 0.374000 -3.299000
H 2.582000 -0.235000 -2.000000
O 6.519000 1.203000 -0.832000
H 6.237000 1.707000 -1.596000
H 6.901000 0.406000 -1.198000
O 0.640000 -6.097000 2.707000
H 0.263000 -5.608000 1.975000
H 0.897000 -5.424000 3.337000
O 1.826000 -5.338000 7.401000

H 1.419000 -5.197000 6.547000
H 2.718000 -5.622000 7.199000
O 6.098000 -1.123000 5.514000
H 5.923000 -2.040000 5.300000
H 5.244000 -0.765000 5.753000
O 4.679000 -6.527000 4.086000
H 3.905000 -7.067000 3.926000
H 5.409000 -7.046000 3.748000
O 0.718000 -9.894000 1.260000
H 1.588000 -9.611000 1.543000
H 0.227000 -9.080000 1.145000
O 7.102000 -3.298000 2.643000
H 7.093000 -2.565000 2.028000
H 6.182000 -3.545000 2.736000
O 2.703000 -8.076000 3.388000
H 2.096000 -7.336000 3.397000
H 3.178000 -7.986000 2.562000
O 10.088000 -3.235000 1.306000
H 10.189000 -3.438000 0.376000
H 9.188000 -3.489000 1.511000
O 8.730000 -0.716000 6.000000
H 8.984000 -0.754000 5.078000
H 7.781000 -0.843000 5.993000
O 4.713000 -8.115000 1.137000
H 5.480000 -8.227000 1.699000
H 4.791000 -8.811000 0.484000
O 4.209000 -0.533000 8.923000
H 4.946000 -1.081000 9.194000
H 4.594000 0.333000 8.784000
O 1.094000 -4.466000 4.798000
H 0.168000 -4.224000 4.788000
H 1.560000 -3.638000 4.676000
O 7.320000 -1.360000 0.746000
H 8.113000 -1.071000 0.294000
H 6.800000 -1.791000 0.067000
O 7.837000 -6.140000 0.610000
H 7.733000 -6.754000 1.336000
H 7.670000 -6.663000 -0.173000
O 9.823000 -1.014000 2.649000
H 10.718000 -0.675000 2.671000
H 9.855000 -1.736000 2.021000
O 3.462000 -5.641000 0.878000
H 3.900000 -5.098000 1.533000
H 3.947000 -6.466000 0.887000
O 0.677000 -2.852000 8.330000
H 1.007000 -3.720000 8.097000
H 1.106000 -2.646000 9.160000
O 5.273000 -4.521000 5.990000
H 4.973000 -4.645000 6.890000
H 5.167000 -5.380000 5.581000
O 1.204000 -0.945000 6.409000
H 1.036000 -1.578000 7.107000
H 0.338000 -0.739000 6.059000
O 2.811000 -2.252000 3.854000
H 2.285000 -1.871000 3.150000
H 2.735000 -1.627000 4.575000

O 5.410000 -0.303000 2.729000
H 6.142000 -0.587000 2.182000
H 5.669000 -0.542000 3.619000
O 4.571000 -4.187000 2.734000
H 3.942000 -3.718000 3.282000
H 4.686000 -5.032000 3.168000
O 4.155000 -4.012000 -8.162000
H 4.458000 -4.903000 -8.341000
H 4.899000 -3.455000 -8.390000
O 5.347000 -5.287000 -4.948000
H 5.964000 -4.612000 -5.230000
H 4.486000 -4.878000 -5.040000
O 5.802000 -5.679000 -2.276000
H 4.905000 -5.455000 -2.029000
H 5.828000 -5.556000 -3.225000
O 4.630000 -8.808000 -3.701000
H 4.490000 -7.935000 -3.332000
H 3.748000 -9.150000 -3.846000
O 6.428000 -2.866000 -1.492000
H 6.399000 -2.461000 -2.359000
H 6.264000 -3.795000 -1.655000
O 0.637000 -5.997000 -5.244000
H 0.801000 -6.466000 -6.062000
H 1.103000 -6.504000 -4.579000
O 2.211000 -6.621000 -3.085000
H 1.827000 -7.262000 -2.486000
H 2.523000 -5.918000 -2.515000
O 2.207000 -9.893000 -4.138000
H 1.849000 -9.227000 -3.551000
H 1.913000 -10.724000 -3.767000
O 1.069000 -2.415000 -2.474000
H 1.179000 -3.356000 -2.615000
H 0.418000 -2.150000 -3.123000
O 9.061000 -4.144000 -1.074000
H 8.740000 -4.863000 -0.530000
H 8.387000 -3.468000 -1.001000
O 3.215000 -1.348000 -0.662000
H 4.017000 -1.027000 -0.249000
H 2.700000 -1.709000 0.060000
O 0.935000 -8.861000 -1.719000
H 1.439000 -9.440000 -1.147000
H 0.196000 -8.577000 -1.181000
O 7.989000 -7.454000 -2.185000
H 7.223000 -6.885000 -2.270000
H 8.170000 -7.746000 -3.078000
O 6.418000 -1.415000 -4.007000
H 5.880000 -0.690000 -4.326000
H 6.671000 -1.892000 -4.797000
O 9.274000 -0.179000 -3.748000
H 9.200000 0.724000 -3.441000
H 8.372000 -0.498000 -3.776000
O 0.677000 -2.852000 -10.445000
H 1.007000 -3.720000 -10.677000
H 1.106000 -2.646000 -9.614000
O 1.166000 -1.985000 -7.707000
H 0.729000 -1.805000 -6.874000

H 1.734000 -1.228000 -7.848000
O 3.451000 -4.574000 -1.645000
H 3.408000 -4.945000 -0.764000
H 3.582000 -3.637000 -1.505000
O 7.152000 -3.018000 -6.044000
H 6.911000 -2.751000 -6.931000
H 8.092000 -2.848000 -5.989000
O 3.290000 -1.628000 -4.526000
H 2.696000 -1.760000 -3.787000
H 3.282000 -2.464000 -4.991000
O 2.921000 -4.262000 -5.575000
H 2.195000 -4.882000 -5.642000
H 2.982000 -3.868000 -6.445000
O 9.333000 -0.235000 -0.596000
H 10.259000 -0.471000 -0.553000
H 9.318000 0.592000 -1.077000
O -4.362000 3.394000 6.706000
H -3.541000 2.903000 6.749000
H -4.329000 3.983000 7.459000
O -8.151000 5.416000 1.898000
H -9.090000 5.577000 1.983000
H -7.832000 6.134000 1.352000
O -5.931000 9.945000 2.412000
H -5.492000 9.107000 2.266000
H -5.219000 10.577000 2.498000
O -6.344000 7.038000 4.893000
H -6.943000 7.762000 4.710000
H -5.796000 6.976000 4.111000
O -7.875000 2.638000 4.021000
H -8.689000 2.139000 3.946000
H -8.069000 3.478000 3.605000
O -5.161000 7.010000 2.318000
H -4.273000 7.322000 2.143000
H -5.190000 6.137000 1.926000
O -2.889000 8.321000 1.786000
H -2.973000 8.752000 0.935000
H -2.170000 8.781000 2.218000
O -3.132000 3.819000 3.872000
H -3.761000 3.923000 3.158000
H -3.627000 4.039000 4.662000
O -7.132000 2.823000 6.728000
H -7.255000 2.841000 5.779000
H -6.184000 2.870000 6.848000
O -5.654000 1.540000 2.586000
H -5.374000 0.697000 2.941000
H -6.428000 1.772000 3.099000
O -4.827000 4.096000 1.649000
H -4.408000 3.909000 0.808000
H -5.133000 3.243000 1.959000
O -1.328000 9.008000 5.346000
H -1.704000 8.241000 4.914000
H -2.058000 9.390000 5.834000
O -2.563000 6.491000 4.586000
H -2.067000 6.258000 5.371000
H -2.761000 5.652000 4.170000
O -1.134000 5.127000 7.120000

H -1.244000 4.973000 8.058000
H -0.373000 4.599000 6.877000
O -0.171000 3.729000 2.986000
H -0.966000 3.663000 3.515000
H 0.033000 4.664000 2.975000
O -3.739000 2.059000 9.436000
H -3.298000 2.003000 8.588000
H -4.008000 2.974000 9.507000
O -1.875000 1.690000 7.447000
H -1.937000 0.746000 7.300000
H -1.096000 1.798000 7.992000
O -5.903000 0.439000 8.141000
H -5.300000 1.011000 8.616000
H -6.644000 1.003000 7.918000
O -2.760000 0.841000 3.415000
H -2.565000 1.740000 3.682000
H -3.547000 0.610000 3.908000
O -0.072000 5.410000 -5.504000
H -0.777000 5.799000 -6.022000
H 0.193000 6.105000 -4.902000
O -1.058000 4.779000 -8.905000
H -2.005000 4.813000 -8.773000
H -0.733000 5.593000 -8.520000
O -4.665000 6.884000 -5.953000
H -5.078000 7.710000 -5.699000
H -5.398000 6.299000 -6.145000
O -1.140000 5.936000 -2.071000
H -1.755000 6.444000 -2.599000
H -1.460000 5.036000 -2.128000
O -4.078000 5.996000 -3.475000
H -4.928000 5.697000 -3.154000
H -4.237000 6.240000 -4.386000
O -6.983000 7.496000 -1.375000
H -6.657000 6.663000 -1.715000
H -6.195000 8.019000 -1.227000
O -9.428000 4.425000 -0.591000
H -10.118000 4.232000 0.044000
H -8.685000 4.709000 -0.058000
O -3.262000 1.132000 -2.505000
H -2.938000 0.400000 -1.979000
H -3.227000 0.813000 -3.406000
O -6.498000 4.863000 -5.711000
H -6.130000 4.043000 -5.383000
H -7.358000 4.923000 -5.295000
O -2.331000 3.734000 -4.644000
H -2.834000 4.475000 -4.305000
H -1.516000 4.122000 -4.960000
O -4.496000 8.442000 -1.088000
H -3.914000 7.687000 -1.176000
H -4.320000 8.973000 -1.865000
O -6.284000 4.945000 -2.248000
H -6.973000 4.283000 -2.313000
H -5.564000 4.499000 -1.803000
O -1.938000 8.155000 -3.794000
H -2.169000 7.926000 -4.695000
H -1.677000 9.075000 -3.840000

O -3.499000 3.381000 -1.016000
H -2.622000 3.389000 -0.633000
H -3.487000 2.644000 -1.627000
O -1.874000 7.318000 -6.606000
H -2.824000 7.284000 -6.491000
H -1.693000 8.228000 -6.841000
O -2.787000 1.067000 -5.196000
H -2.150000 1.680000 -4.830000
H -2.640000 1.103000 -6.141000
O -9.417000 2.219000 -2.289000
H -9.478000 2.811000 -1.539000
H -8.477000 2.112000 -2.434000
O -6.780000 1.876000 -2.649000
H -6.166000 2.194000 -3.311000
H -6.235000 1.391000 -2.030000
O -6.039000 0.439000 -6.323000
H -6.913000 0.773000 -6.525000
H -6.148000 -0.033000 -5.497000
O -5.132000 2.646000 -4.938000
H -5.460000 1.832000 -5.318000
H -4.183000 2.603000 -5.062000
O -3.739000 2.059000 -9.338000
H -3.298000 2.003000 -10.186000
H -4.008000 2.974000 -9.267000
O -5.654000 0.619000 -0.254000
H -5.767000 0.938000 0.641000
H -4.837000 1.020000 -0.550000
O -3.680000 -5.257000 5.545000
H -3.790000 -6.200000 5.670000
H -3.683000 -4.895000 6.431000
O -5.931000 -8.830000 2.412000
H -5.492000 -9.668000 2.266000
H -5.219000 -8.197000 2.498000
O -7.822000 -3.890000 6.380000
H -7.382000 -3.148000 5.966000
H -7.883000 -3.648000 7.304000
O -5.155000 -0.903000 3.736000
H -5.315000 -1.193000 4.635000
H -5.320000 -1.679000 3.201000
O -5.452000 -3.395000 2.296000
H -6.146000 -3.397000 1.637000
H -4.710000 -3.819000 1.866000
O -5.887000 -1.768000 6.371000
H -5.431000 -2.549000 6.685000
H -5.548000 -1.058000 6.915000
O -1.860000 -3.599000 4.514000
H -2.503000 -4.254000 4.788000
H -2.247000 -3.194000 3.738000
O -3.471000 -4.707000 1.051000
H -3.697000 -5.424000 1.644000
H -3.861000 -4.956000 0.214000
O -3.889000 -6.790000 2.713000
H -4.177000 -6.329000 3.501000
H -3.094000 -7.251000 2.981000
O -8.686000 -3.235000 1.306000
H -8.585000 -3.438000 0.376000

H -9.586000 -3.489000 1.511000
O -2.037000 -2.626000 9.252000
H -1.912000 -1.803000 9.724000
H -1.170000 -2.837000 8.904000
O -2.402000 -2.133000 2.170000
H -2.687000 -2.546000 1.355000
H -2.600000 -1.204000 2.050000
O -8.952000 -1.014000 2.649000
H -8.056000 -0.675000 2.671000
H -8.920000 -1.736000 2.021000
O -0.732000 -4.741000 0.692000
H -1.563000 -4.674000 1.163000
H -0.555000 -3.851000 0.389000
O -1.276000 -7.997000 3.196000
H -0.514000 -7.427000 3.302000
H -1.261000 -8.564000 3.967000
O -1.224000 -7.703000 0.353000
H -0.950000 -6.795000 0.225000
H -1.413000 -7.767000 1.289000
O -3.229000 -1.006000 6.980000
H -2.633000 -1.328000 6.303000
H -2.972000 -1.479000 7.771000
O -1.207000 -0.693000 5.176000
H -1.032000 -1.379000 4.532000
H -1.905000 -0.168000 4.783000
O -2.510000 -3.823000 -7.033000
H -3.194000 -3.153000 -7.041000
H -1.971000 -3.625000 -7.798000
O -4.974000 -1.979000 -8.318000
H -4.961000 -2.382000 -9.186000
H -5.527000 -1.204000 -8.422000
O -3.975000 -1.844000 -5.265000
H -4.452000 -1.037000 -5.458000
H -4.659000 -2.496000 -5.112000
O -0.992000 -4.892000 -3.036000
H -0.866000 -5.130000 -3.954000
H -0.950000 -5.723000 -2.564000
O -7.130000 -1.189000 -4.466000
H -6.616000 -1.278000 -3.664000
H -7.973000 -0.843000 -4.173000
O -8.023000 -4.637000 -3.259000
H -8.008000 -5.569000 -3.477000
H -8.730000 -4.551000 -2.620000
O -1.812000 -7.997000 -3.912000
H -2.537000 -7.404000 -4.109000
H -1.284000 -8.004000 -4.710000
O -8.589000 -2.996000 -6.165000
H -7.945000 -2.790000 -5.487000
H -8.227000 -3.757000 -6.618000
O -5.542000 -4.193000 -4.243000
H -5.207000 -3.606000 -3.565000
H -6.462000 -4.324000 -4.013000
O -2.037000 -2.626000 -9.523000
H -1.912000 -1.803000 -9.050000
H -1.170000 -2.837000 -9.870000
O -6.723000 -5.240000 -0.715000

H -7.329000 -5.217000 -1.455000
H -7.027000 -5.972000 -0.178000
O -5.293000 -1.512000 -2.104000
H -4.340000 -1.597000 -2.154000
H -5.440000 -0.848000 -1.430000
O -3.643000 -5.723000 -1.628000
H -2.902000 -5.386000 -2.131000
H -4.406000 -5.554000 -2.181000
O -4.969000 -7.672000 -3.092000
H -4.554000 -7.517000 -2.243000
H -5.903000 -7.546000 -2.930000
O -9.500000 -0.179000 -3.748000
H -9.574000 0.724000 -3.441000
H -10.402000 -0.498000 -3.776000
O -1.972000 -2.483000 -3.419000
H -2.572000 -2.424000 -4.162000
H -1.843000 -3.423000 -3.289000
O -3.287000 -5.758000 -5.321000
H -4.101000 -5.328000 -5.059000
H -2.880000 -5.146000 -5.934000
O -0.362000 -0.999000 -5.243000
H -0.856000 -0.602000 -5.961000
H -0.971000 -1.624000 -4.850000
O -2.678000 -0.357000 -7.965000
H -3.138000 0.395000 -8.339000
H -3.307000 -1.076000 -8.023000
O -9.442000 -0.235000 -0.596000
H -8.515000 -0.471000 -0.553000
H -9.457000 0.592000 -1.077000
O -2.874000 -1.008000 -0.772000
H -2.081000 -1.503000 -0.979000
H -2.581000 -0.329000 -0.164000

3.2 Publication II: Nonadiabatic Molecular Dynamics on Graphics Processing Units: Performance and Application to Rotary Molecular Motors

L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Nonadiabatic Molecular Dynamics on Graphics Processing Units:
Performance and Application to Rotary Molecular Motors",
J. Chem. Theory Comput., **15**, 6647 (2019)

Abstract: Nonadiabatic molecular dynamics (NAMD) simulations of molecular systems require the efficient evaluation of excited-state properties, such as energies, gradients, and nonadiabatic coupling vectors. Here, we investigate the use of graphics processing units (GPUs) in addition to central processing units (CPUs) to efficiently calculate these properties at the time-dependent density functional theory (TDDFT) level of theory. Our implementation in the FermiONS++ program package uses the J-engine and a preselective screening procedure for the calculation of Coulomb and exchange kernels, respectively. We observe good speed-ups for small and large molecular systems (comparable to those observed in ground-state calculations) and reduced (down to sublinear) scaling behavior with respect to the system size (depending on the spatial locality of the investigated excitation). As a first illustrative application, we present efficient NAMD simulations of a series of newly designed light-driven rotary molecular motors and compare their S_1 lifetimes. Although all four rotors show different S_1 excitation energies, their ability to rotate upon excitation is conserved, making the series an interesting starting point for rotary molecular motors with tunable excitation energies.

Reprinted with permission from:

L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Nonadiabatic Molecular Dynamics on Graphics Processing Units:
Performance and Application to Rotary Molecular Motors",
J. Chem. Theory Comput., **15**, 6647 (2019)

Copyright 2019 American Chemical Society.

<https://pubs.acs.org/doi/10.1021/acs.jctc.9b00859>



Nonadiabatic Molecular Dynamics on Graphics Processing Units: Performance and Application to Rotary Molecular Motors

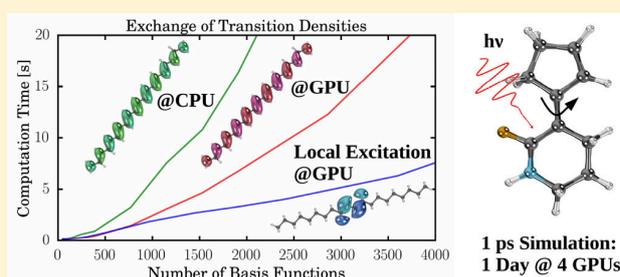
Laurens D. M. Peters,[†] Jörg Kussmann,[†] and Christian Ochsenfeld^{*,†,‡,§}

[†]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandstr. 7, D-81377 München, Germany

[‡]Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart, Germany

Supporting Information

ABSTRACT: Nonadiabatic molecular dynamics (NAMD) simulations of molecular systems require the efficient evaluation of excited-state properties, such as energies, gradients, and nonadiabatic coupling vectors. Here, we investigate the use of graphics processing units (GPUs) in addition to central processing units (CPUs) to efficiently calculate these properties at the time-dependent density functional theory (TDDFT) level of theory. Our implementation in the FermiONs++ program package uses the J-engine and a preselective screening procedure for the calculation of Coulomb and exchange kernels, respectively. We observe good speed-ups for small and large molecular systems (comparable to those observed in ground-state calculations) and reduced (down to sublinear) scaling behavior with respect to the system size (depending on the spatial locality of the investigated excitation). As a first illustrative application, we present efficient NAMD simulations of a series of newly designed light-driven rotary molecular motors and compare their S_1 lifetimes. Although all four rotors show different S_1 excitation energies, their ability to rotate upon excitation is conserved, making the series an interesting starting point for rotary molecular motors with tunable excitation energies.



1. INTRODUCTION

Nonadiabatic processes such as electronic excitations,¹ radiationless transitions,² and electron transfer³ are of key importance in chemistry and biology. One of the most prominent examples in these fields is the rhodopsin protein, whose chromophore undergoes a photoisomerization when exposed to light.⁴ This energy conversion of light to mechanical motion has inspired chemists to design synthetic light-driven rotary molecular motors,^{5,6} for which Bernard Feringa was awarded the Nobel Prize in chemistry in 2016 together with Jean Pierre Sauvage and Sir James Fraser Stoddart.⁷ The description of these nonadiabatic processes is an ongoing challenge in modern quantum chemistry (e.g., ref 8). This is mainly due to the fact that (1) excited states have to be taken into account and (2) the dynamics of the nuclei need to be considered.

To tackle the first challenge, several quantum-chemical methods have been developed. Examples include the complete active space self-consistent field (CASSCF) method,⁹ the algebraic-diagrammatic construction (ADC(2)),¹⁰ several coupled cluster methods (e.g., CC2),¹¹ and time-dependent density functional theory (TDDFT).^{12,13} The latter serves (despite its well-known limitations using today's functionals^{13,14}) as a good compromise between effort and accuracy.^{15,16} As a consequence of this, not only excited-state energies but also excited-state gradients¹⁷ and nonadiabatic

coupling vectors (NACVs)^{18–29} have been implemented to provide access to molecular properties at the TDDFT level of theory. Having excited-state energies and properties at hand, trajectory surface hopping (TSH)^{30,31} is a straightforward way to conduct nonadiabatic molecular dynamics (NAMD) simulations including several electronic states.

Despite many advances in the field of NAMD^{32–34} and their broad field of application,^{35–44} it remains difficult or even impossible to investigate large molecular systems. The reason for this is that TSH requires, because of its stochastic nature, a large set of independent trajectories including many time steps, which involve expensive (even at the TDDFT level of theory) excited-state energy and property calculations. One way to accelerate NAMD simulations is the use of exciton models.⁴⁵ Another approach is the use of graphics processing units (GPUs) in addition to central processing units (CPUs) for the calculations. It was shown that this leads to significant speed-ups for ground-state energy and forces evaluations^{46–53} and was also successfully applied to excited-state energies and properties^{45,54–56} as well as ab initio multiple spawning^{57,58} and NAMD^{59,60} simulations.

In this work, we present efficient NAMD simulations with TDDFT energies, gradients, and NACVs calculated on hybrid

Received: August 29, 2019

Published: November 25, 2019

CPU/GPU architectures using the FermiONs++ program package.^{50–52} We start with a brief summary of the theory behind TSH as well as TDDFT energies, gradients, and NACVs in Section 2 and discuss their implementation on GPUs, featuring the hybrid CPU/GPU engine⁵² for the two-electron integrals as well as the preselective screening procedure for the evaluation of exchange kernels.^{50,51} Computational details are given in Section 3. In Section 4, we discuss the accuracy and the performance of our GPU-based excited-state routines, investigating the scaling of our integral evaluations and contractions. We show timings for selected molecules containing more than 500 atoms. As a prototypical application, we investigate the photoinduced rotation of four newly designed rotary molecular motors via NAMD in Section 5, followed by a conclusion and an outlook.

2. THEORY

All equations in this section use the standard notation for orbitals: i, j, k, \dots denote occupied, a, b, c, \dots denote virtual, and p, q, r, \dots denote arbitrary molecular orbitals, while μ, ν, λ, \dots denote basis functions. I, J, \dots denote different electronic states. v_{xc} and f_{xc} are the first- and second-order exchange-correlation functional derivatives, respectively. c_x is the amount of exact exchange. \mathbf{F} and \mathbf{S} are the Kohn–Sham and the overlap matrices, respectively. \mathbf{P} is the ground-state density, and \mathbf{h} is the one-electron core Hamiltonian matrix. The two-electron integrals are written using Mulliken notation (\dots). Superscript ξ denotes derivatives with respect to the nuclear coordinates, and $*$ and \dagger symbolize complex conjugation and adjunction, respectively. For the sake of simplicity, we use the same symbol for matrices in the atomic orbital (AO) and molecular orbital (MO) bases with different indices (e.g., $P_{\mu\nu}$ vs P_{pq}).

2.1. Trajectory Surface Hopping. In nonadiabatic molecular dynamics (NAMD), the electronic time-dependent wave function of the system is assumed to be a linear combination of the time-independent electronic wave functions of the individual electronic states weighted by the state amplitudes (\vec{c})

$$\Psi(\mathbf{x}, \{\mathbf{r}\}, t) = \sum_I c_I(t) \Phi_I(\mathbf{x}(t), \{\mathbf{r}\}) \quad (1)$$

where \mathbf{x} represents nuclear coordinates and $\{\mathbf{r}\}$ is the set of electronic coordinates. Having the state energies (ω_I , see Section 2.2) and nonadiabatic coupling vectors (NACVs, $\tau_{I \rightarrow J}^\xi$, see Section 2.4) at hand, it is possible to propagate the state amplitudes along with the nuclei during the molecular dynamics (MD) simulation using a unitary propagator:³⁶

$$\begin{aligned} \vec{c}(t + \delta t) &\approx \mathbf{U}(t, t + \delta t) \vec{c}(t) \\ \mathbf{U}(t, t + \delta t) &= \exp\left(\frac{\mathbf{H}(t) + \mathbf{H}(t + \delta t)}{2} \delta t\right) \\ H_{IJ} &= \exp\left[i \int_0^t (\omega_{IJ}) dt'\right] (\hat{\mathbf{x}} \cdot \tau_{I \rightarrow J}^\xi) \end{aligned} \quad (2)$$

Here, ω_{IJ} is the energy difference between two states ($\omega_J - \omega_I$). The applied δt for the propagation of \vec{c} is normally 3 orders of magnitude smaller than the time step of the MD simulation (Δt).³¹ In the fewest-switches surface hopping algorithm,^{30,31} \vec{c} is used to calculate the probability $g_{I \rightarrow J}^{t \rightarrow t + \Delta t}$ of the system switching from its current state I to another state J at every time step:

$$\begin{aligned} g_{I \rightarrow J}^{t \rightarrow t + \Delta t} &= -\frac{2\Delta t}{a_{II}} \int_t^{t + \Delta t} dt \mathcal{R}\{a_{JI}^* H_{JI}\} \\ \mathbf{a} &= \mathbf{c} \dagger \end{aligned} \quad (3)$$

If a random number between zero and one exceeds g , then the MD simulation is continued on the potential energy surface of state J and the nuclear velocity is rescaled along $\tau_{I \rightarrow J}^\xi$. Observables (e.g., lifetimes of states or relaxation pathways) can be drawn from ensembles of trajectories using a different set of random numbers. In the following sections, we will briefly summarize the calculation of the necessary ingredients of TSH (excited-state energies, gradients, and NACVs) at the time-dependent density functional theory (TDDFT) level of theory.

2.2. Excited-State Energies. Excitation energies can be calculated from linear response TDDFT by solving the TDDFT or the random phase approximation (RPA) equation^{12,61,62}

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{bmatrix} = \omega_I \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{bmatrix} \quad (4)$$

$\mathbf{A} \pm \mathbf{B}$ are the orbital rotation Hessians

$$\begin{aligned} (\mathbf{A} + \mathbf{B})_{ijab} &= \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + 2(ialj) + 2f_{iajb}^{xc} \\ &\quad - c_x [(jalib) + (ablij)] \end{aligned} \quad (5)$$

$$(\mathbf{A} - \mathbf{B})_{ijab} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + c_x [(jalib) + (ablij)] \quad (6)$$

with ϵ_p being the energy of orbital p . \mathbf{X}_I and \mathbf{Y}_I are the transition densities for excitation and de-excitation, respectively. Neglecting \mathbf{B} is known as the Tamm–Dancoff approximation (TDA) of TDDFT,⁶³ simplifying eq 4 to

$$\mathbf{A}\mathbf{X}_I = \omega_I \mathbf{X}_I \quad (7)$$

Please note that eqs 4 and 7 are equivalent to the time-dependent Hartree–Fock (TDHF) and configuration interaction singles (CIS) equations, respectively, when c_x is set equal to 1 and all exchange-correlation terms are neglected. TDA is computationally less demanding than TDDFT because fewer two-electron integrals have to be evaluated. It may also be more suitable for NAMD simulations because it is more stable (regarding its convergence) and thus typically delivers better results than TDDFT in the vicinity of conical intersections involving the ground state.^{39,64}

TDDFT is nowadays widely used for the calculation of excitation energies, being a good compromise between accuracy and computational cost.^{15,16} Shortcomings are, however, the inability to calculate double excitations and the poor description of charge-transfer excitations.^{13,14} To tackle the latter, range-corrected functionals (e.g., ω B97⁶⁵) have been introduced.

2.3. Excited-State Energy Gradients. To determine the energy gradient of an excited state I , one has to calculate the change in the excitation energy with respect to the nuclear coordinates (ω_I^ξ). The final equation in the AO basis has been derived by Furche et al.¹⁷

$$\begin{aligned} \omega_I^\xi &= \sum_{\mu\nu} \{h_{\mu\nu}^\xi P_{\mu\nu}^I - S_{\mu\nu}^\xi W_{\mu\nu}^I + v_{\mu\nu}^{xc(\xi)} P_{\mu\nu}^I\} \\ &\quad + \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^\xi \Gamma_{\mu\nu\kappa\lambda}^I + \sum_{\mu\nu\kappa\lambda} f_{\mu\nu\kappa\lambda}^{xc(\xi)} R_{\mu\nu}^I R_{\kappa\lambda}^I \end{aligned} \quad (8)$$

The calculations of the relaxed difference density matrix (\mathbf{P}_I), the energy-weighted difference density matrix (\mathbf{W}_I), and the two-particle difference density matrix (Γ_I) are shown in the Appendix. The first two require an iterative solution of a Z-vector equation.¹⁷ \mathbf{R}_I is $\mathbf{X}_I + \mathbf{Y}_I$ or \mathbf{X}_I in the case of TDDFT or TDA, respectively.

2.4. Nonadiabatic Coupling Vectors. The NACV between two states (I and J) is defined as

$$\tau_{I \rightarrow J}^\xi = \langle \Phi_I | \Phi_J^\xi \rangle \quad (9)$$

It thus describes the change of the overlap of the wave functions of I and J with respect to the nuclear coordinates. The first formulation of the NACV involving the ground state ($\tau_{0 \rightarrow I}^\xi$) derived by Chernyak and Mukamel¹⁸ was corrected for finite basis set effects by Send et al.²³ in 2010:

$$\begin{aligned} \tau_{0 \rightarrow I}^\xi &= \sum_{\mu\nu} [h_{\mu\nu}^\xi P_{\mu\nu}^{0I} - S_{\mu\nu}^\xi W_{\mu\nu}^{0I} + v_{\mu\nu}^{xc(\xi)} P_{\mu\nu}^{0I}] \\ &+ \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^\xi \Gamma_{\mu\nu\kappa\lambda}^{0I} + \sum_{\mu\nu} S_{\mu\nu}^{[A]\xi} \gamma_{\mu\nu}^{0I} \end{aligned} \quad (10)$$

Equations for the density matrices (\mathbf{P}_{0I} , \mathbf{W}_{0I} , and Γ_{0I}) are again given in the Appendix. They can be calculated directly from the TDDFT transition densities so that no Z-vector equation needs to be solved. $S_{\mu\nu}^{[A]\xi}$ denotes an antisymmetric overlap derivative

$$S_{\mu\nu}^{[A]\xi} = \langle \chi_\mu^\xi | \chi_\nu^\xi \rangle - \langle \chi_\nu^\xi | \chi_\mu^\xi \rangle \quad (11)$$

whereas γ_{0I} is defined as

$$\gamma_{0I} = \mathbf{L}_I \quad (12)$$

with \mathbf{L}_I being $\mathbf{X}_I - \mathbf{Y}_I$ or \mathbf{X}_I in the case of TDDFT or TDA, respectively.

The calculation of NACVs between two excited states has been tackled by many publications showing expressions based on linear and quadratic response at different levels of theory.^{25,26,29,66–71} The main conclusion is that the use of quadratic response theory leads to unphysical poles when the energy difference between excited states matches the excitation energy of another state.^{66–70} The use of linear response theory or the so-called pseudowave function approach^{25–27,29} is therefore recommended. The resulting equation in the AO basis is as follows:

$$\begin{aligned} \tau_{I \rightarrow J}^\xi &= \frac{1}{\omega_{JI}} \left\{ \sum_{\mu\nu} [h_{\mu\nu}^\xi P_{\mu\nu}^{IJ} - S_{\mu\nu}^\xi W_{\mu\nu}^{IJ} + v_{\mu\nu}^{xc(\xi)} P_{\mu\nu}^{IJ}] \right. \\ &+ \left. \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^\xi \Gamma_{\mu\nu\kappa\lambda}^{IJ} + \sum_{\mu\nu\kappa\lambda} f_{\mu\nu\kappa\lambda}^{xc(\xi)} R_{\mu\nu}^I R_{\kappa\lambda}^J \right\} + \sum_{\mu\nu} S_{\mu\nu}^{[A]\xi} \gamma_{\mu\nu}^{IJ} \end{aligned} \quad (13)$$

Equations for the density matrices (\mathbf{P}_{IJ} , \mathbf{W}_{IJ} , and Γ_{IJ}) can again be found in the Appendix. γ_{IJ} is defined as

$$\gamma_{IJ} = \frac{1}{2} \begin{pmatrix} -(\mathbf{R}_I^T \mathbf{L}_J + \mathbf{L}_I^T \mathbf{R}_J)_{ij} & 0 \\ 0 & (\mathbf{R}_I \mathbf{L}_J^T + \mathbf{L}_I \mathbf{R}_J^T)_{ab} \end{pmatrix} \quad (14)$$

It was shown by Fatehi et al.²⁵ that the antisymmetric overlap derivatives (eq 11) introduce translational variance into the NACV calculations. Neglecting these terms in eqs 10 and 13 is equivalent to adding electron-translational factors (ETFs).

However, trajectory surface hopping (TSH) simulations using NACVs with and without ETFs lead to nearly identical results for larger molecular systems.⁷²

2.5. Graphics Processing Units. Graphics processing units (GPUs) significantly accelerate quantum-chemical calculations because they allow for an efficient evaluation and contraction of two-electron integrals (eq 15) and their derivatives with respect to the nuclear coordinates (eq 16).^{46–52}

$$\begin{aligned} J(\mathbf{M})_{\mu\nu} &= \sum_{\kappa\lambda} M_{\kappa\lambda} (\mu\nu|\kappa\lambda) \\ K(\mathbf{M})_{\mu\nu} &= \sum_{\kappa\lambda} M_{\kappa\lambda} (\mu\kappa|\nu\lambda) \end{aligned} \quad (15)$$

$$\begin{aligned} J^\xi(\mathbf{N}, \mathbf{M}) &= \sum_{\mu\nu\kappa\lambda} N_{\mu\nu} M_{\kappa\lambda} (\mu\nu|\kappa\lambda)^\xi \\ K^\xi(\mathbf{N}, \mathbf{M}) &= \sum_{\mu\nu\kappa\lambda} N_{\mu\nu} M_{\kappa\lambda} (\mu\kappa|\nu\lambda)^\xi \end{aligned} \quad (16)$$

J and K denote the Coulomb and exchange parts, respectively, and \mathbf{M} and \mathbf{N} denote general density matrices. For the Coulomb part, large speed-ups can be observed when the J-engine^{73,74} is applied to the rearranged shell-pair data.⁴⁷ To see comparable speed-ups for the exchange part, we apply an additional preselective screening (preLinK):^{50,51}

$$\sum_{\kappa\lambda} \sqrt{(\mu\kappa|\mu\kappa)} \times |M_{\kappa\lambda}| \times \sqrt{(\nu\lambda|\nu\lambda)} \geq \vartheta_{\text{pre}} \quad (17)$$

$$K(\mathbf{M})_{\mu\nu} N_{\mu\nu} \geq \vartheta_{\text{pre}}^V \quad (18)$$

It determines the significant shell pairs (those with an expected value above the given thresholds ϑ_{pre} and ϑ_{pre}^V) before their distribution to and calculation on the GPUs. The two-electron integral evaluation can be done even more efficiently when the workload is spread among both GPUs and CPUs.⁵²

In ground-state calculations, \mathbf{M} and \mathbf{N} are solely the ground-state density (\mathbf{P}), whereas for excited states \mathbf{M} and \mathbf{N} can additionally be \mathbf{P}_I , \mathbf{R}_I , or \mathbf{L}_I . However, the discussed procedure (shell pair rearrangement, J-engine, and preLinK) is still valid and can easily be adapted to excited-state routines when keeping in mind that \mathbf{M} is not always equal to \mathbf{N} and may be nonsymmetric.

3. COMPUTATIONAL DETAILS

3.1. General Remarks. The FermiONs++ program package^{50–52} was used for all calculations presented in this work. It was compiled using the Intel compiler (2019),⁷⁵ and the Intel Math Kernel Library (MKL). Routines on AMD GPUs were compiled with the AMD APPSDK compiler. In addition, LibXC library v4.0.1^{76,77} was used. In all calculations, we used the gm5 grid⁷⁸ (with the modified Becke weighting scheme described in ref 78) and tight thresholds for the SCF convergence ($\vartheta_{\text{SCF}} = 10^{-7}$ using the FP-commutator), the integrals ($\vartheta_{\text{INT}} = 10^{-10}$), and the Z-vector equation convergence ($\vartheta_Z = 10^{-5}$). Throughout all calculations, we neglected the symmetry of the molecules and solely calculated singlet excitations. Excited-state energies, gradients, and nonadiabatic coupling vectors (with electron transition factors) at the TDDFT level of theory were calculated using eqs 4, 8, 10, and 13, respectively.

Table 1. Mean Absolute Errors (MAE, in Atomic Units) of Excitation Energies (ω_I), Gradients (ω_I^{\ddagger}), and Nonadiabatic Coupling Vectors ($\tau_{I \rightarrow J}$) of Protonated Formaldimine (I) and the Schiff Base of Retinal (II) Calculated at the PBE0/def2-SVP Level of Theory on GPUs, Employing the preLinK Scheme and Using Different Thresholds for preLinK (ϑ_{pre}), the preLinK Gradient ($\vartheta_{\text{pre}}^{\text{V}}$), and the TDDFT Convergence (ϑ_{TDDFT})^a

screening thresholds and convergence criteria								
ϑ_{pre}	10^{-3}	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}
$\vartheta_{\text{pre}}^{\text{V}}$	10^{-11}	10^{-11}	10^{-11}	10^{-11}	10^{-11}	10^{-9}	10^{-10}	10^{-11}
ϑ_{TDDFT}	10^{-5}	10^{-5}	10^{-6}	10^{-6}	10^{-7}	10^{-7}	10^{-7}	10^{-7}
I								
MAE(ω_I)	2.2×10^{-4}	2.2×10^{-4}	8.8×10^{-5}	8.8×10^{-5}	2.3×10^{-6}	2.3×10^{-6}	2.3×10^{-6}	2.3×10^{-6}
MAE(ω_I^{\ddagger})	5.8×10^{-5}	5.8×10^{-5}	3.7×10^{-5}	3.7×10^{-5}	6.5×10^{-6}	6.7×10^{-6}	6.6×10^{-6}	6.5×10^{-6}
MAE($\tau_{0 \rightarrow 1}$)	5.7×10^{-4}	5.7×10^{-4}	1.2×10^{-4}	1.2×10^{-4}	9.2×10^{-5}	8.8×10^{-5}	9.1×10^{-5}	9.2×10^{-5}
MAE($\tau_{1 \rightarrow 2}$)	1.0×10^{-3}	9.1×10^{-4}	4.7×10^{-4}	4.7×10^{-4}	2.8×10^{-4}	9.2×10^{-4}	5.5×10^{-4}	2.8×10^{-4}
II								
MAE(ω_{II})	1.9×10^{-4}	1.9×10^{-4}	9.6×10^{-5}	9.6×10^{-5}	1.1×10^{-5}	1.1×10^{-5}	1.1×10^{-5}	1.1×10^{-5}
MAE(ω_{II}^{\ddagger})	2.8×10^{-5}	2.8×10^{-5}	1.2×10^{-5}	1.2×10^{-5}	3.1×10^{-6}	3.8×10^{-6}	3.2×10^{-6}	3.1×10^{-6}
MAE($\tau_{0 \rightarrow 1}$)	1.7×10^{-3}	1.7×10^{-3}	7.5×10^{-4}	7.5×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	2.0×10^{-4}	2.0×10^{-4}
MAE($\tau_{1 \rightarrow 2}$)	8.9×10^{-4}	8.9×10^{-4}	3.8×10^{-4}	3.8×10^{-4}	1.1×10^{-4}	1.2×10^{-4}	1.1×10^{-4}	1.1×10^{-4}

^aA calculation on CPUs with $\vartheta_{\text{TDDFT}} = 10^{-7}$ and without CFMM or preLinK is used as a reference. Throughout this work, we will use accurate ($\vartheta_{\text{pre}} = 10^{-3}$, $\vartheta_{\text{pre}}^{\text{V}} = 10^{-10}$, and $\vartheta_{\text{TDDFT}} = 10^{-5}$) or tight thresholds ($\vartheta_{\text{pre}} = 10^{-4}$, $\vartheta_{\text{pre}}^{\text{V}} = 10^{-10}$, and $\vartheta_{\text{TDDFT}} = 10^{-6}$).

3.2. Preparation and Calculation of Systems I, II, III_n, IV_n, V, VI, and VII. To illustrate the performance and/or accuracy of our implementation on GPUs, we use protonated formaldimine (I), the Schiff base of retinal (II), a series of linear polyethynes (III_n) and dialkylethenes (IV_n), a motorized nanocar (without “wheels”, V),⁷⁹ and one (VI) and three (VII) pores of a covalent organic framework⁸⁰ as example molecules. The structures of I, II, and V have been optimized at the PBE0^{81–84}/def2-SVP^{85,86} level of theory, while III_n and IV_n have not been optimized to maintain their linear structures. VI and VII have been prepared according to ref 80. All structures are available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>.

TDDFT energies and properties of I, II, III_n, IV_n, V, VI, and VII were calculated at the PBE0/def2-SVP or PBE/def2-TZVP level of theory. We used tight thresholds for preLinK ($\vartheta_{\text{pre}} = 10^{-4}$), the preLinK gradient ($\vartheta_{\text{pre}}^{\text{V}} = 10^{-10}$), and the TDDFT convergence ($\vartheta_{\text{TDDFT}} = 10^{-6}$). For I, II, III_n, IV_n, V, VI, and VII, four, six, five, three, six, seven, and seven states were taken into account, respectively.

When investigating the accuracy, we use a calculation performed on CPUs with tight thresholds ($\vartheta_{\text{TDDFT}} = 10^{-7}$) as a reference. To allow for a fair comparison, we employed the continuous fast multipole method (CFMM)^{87,88} and the LinK scheme^{89,90} for Coulomb and exchange kernels (and their derivatives) on CPUs when comparing CPU and GPU performance. Coulomb and exchange kernels on GPUs were calculated with the J-engine^{73,74} and the preLinK scheme,^{50,51} respectively. Timings of integral evaluations and entire routines were determined as an average over five independent calculations on two Intel Xeon CPU E5 2640 v4 @ 2.20 GHz (20 threads) CPUs and four AMD FirePro 3D W8100 GPUs. The scaling behavior is determined as the slope of the corresponding log–log plots (Supporting Information) using the timings of III_n and IV_n with $n = 40, 50, 75, 100$. The parallel efficiency is determined as the ratio between the measured and ideal speed-ups.

3.3. Preparation and Calculation of the Rotary Molecular Motors. The structures of the four rotary molecular machines (C, N, S, and O) have been optimized

at the $\omega\text{B97}^{65}/\text{def2-SVP}$ level of theory. Excited-state properties and timings were calculated at the TDA ($\omega\text{B97}/\text{def2-SVP}$) level of theory using accurate thresholds ($\vartheta_{\text{pre}} = 10^{-3}$, $\vartheta_{\text{pre}}^{\text{V}} = 10^{-10}$, and $\vartheta_{\text{TDDFT}} = 10^{-5}$). NAMD simulations were conducted at the same level of theory. The propagation of the nuclei was calculated using the Velocity Verlet algorithm.^{91,92} The extended Lagrangian method⁹³ for the extrapolation of ground-state density was used to accelerate SCF convergence. Transition densities and relaxed difference densities of the previous step were used as guesses for the TDA and Z-vector equation, respectively.

Twenty initial geometries and velocities (available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>) were drawn from a 5 ps ground-state NVT simulation (200 fs equilibration, 0.2 fs step size, velocity rescaling thermostat⁹⁴) at the same level of theory. From each initial condition, five independent (different series of random numbers) NAMD simulations were conducted for 1 ps (0.2 fs step size) without equilibration, thermostat, or decoherence correction, starting at the first excited singlet state (S_1). At every step of the simulation, the overall rotation and translation of the molecule was removed. Three excited states were taken into account, whereas only the coupling vector from the ground state to the first excited state was calculated and used for the propagation of the state amplitudes (eq 2) and the calculation of the hopping probability (eq 3).

4. PERFORMANCE

4.1. Accuracy and Thresholds. We start with an analysis of the errors of excited-state energies and properties introduced by their calculation on GPUs and the use of the preLinK scheme^{51,52} for integrals and integral derivatives.

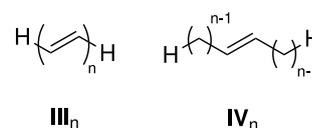


Figure 1. Structures of the linear polyethynes (III_n) and dialkylethenes (IV_n).

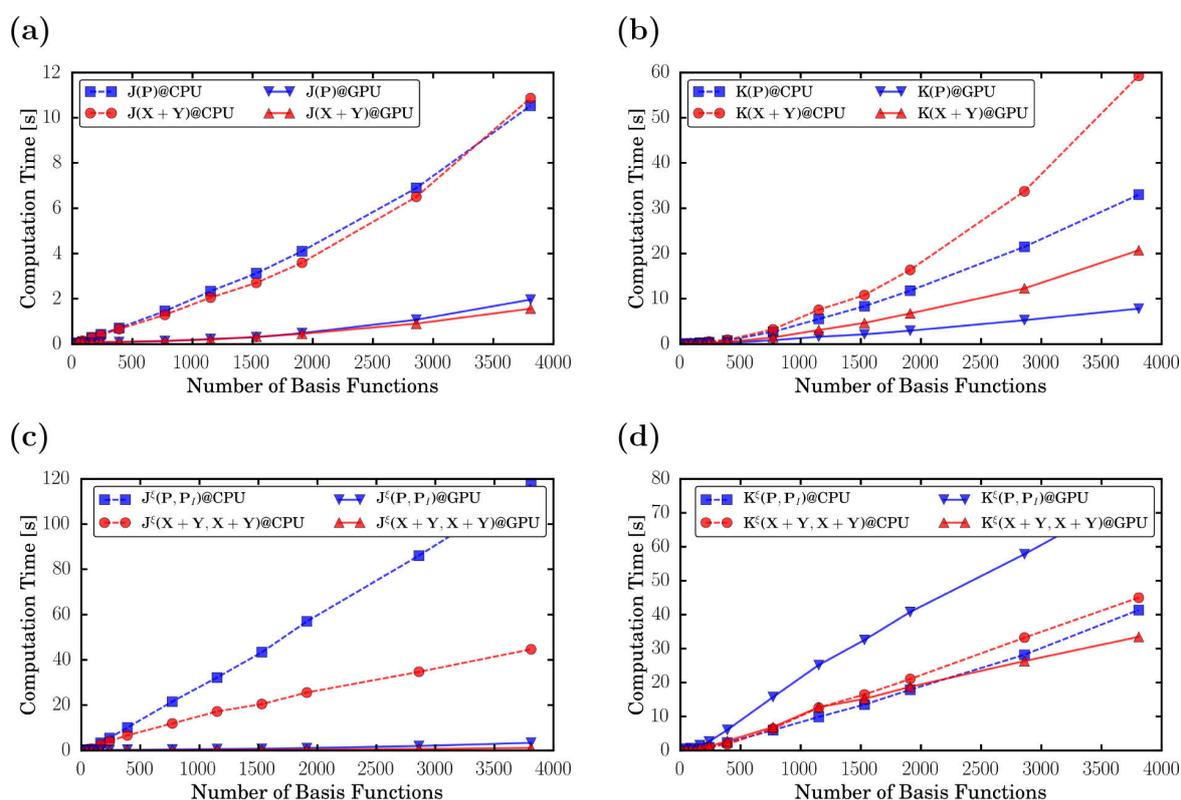


Figure 2. Timings of (a, c) Coulomb and (b, d) exchange integral evaluations (a, b; eq 15) and their derivatives with respect to the nuclear coordinates (c, d; eq 16) of polyethylene (III_n) with $n = 1-100$ calculated at the PBE0/def2-SVP level of theory on CPUs (dashed lines) and GPUs (solid lines). For details on the calculations and the computational setup, see Sections 3.1 and 3.2.

Therefore, we compare calculations of the smallest Schiff base (protonated formalimine, CH_2NH_2^+ , **I**) and the Schiff base of retinal ($\text{C}_{20}\text{H}_{30}\text{N}^+$, **II**) with GPUs and preLinK to calculations on CPUs without prescreening or the continuous fast multipole method (CFMM). (For computational details, see Section 3.2.) The mean absolute errors for different thresholds are listed in Table 1.

The choice of the preLinK threshold (ϑ_{pre}) for excited-state calculations should always depend on the applied convergence threshold for the TDDFT equation (ϑ_{TDDFT}). If the chosen threshold is too loose, then the iterative solution of the TDDFT equation does not converge (especially for large molecular systems), while a ϑ_{pre} that is too tight does not improve the result when ϑ_{TDDFT} is not adjusted accordingly (Table 1). The latter is also attributed to the small effect of preLinK on these relatively small systems.^{50,51} In our calculations with the FermiONS++ program package, the “ideal” ϑ_{pre} is 2 orders of magnitude larger than ϑ_{TDDFT} . A tightening of these two parameters (left side of Table 1) systematically leads to smaller errors. However, the errors in the coupling vectors of the large system do not fall below 10^{-4} a.u., marking the numerical limit of these second- to third-order properties. The use of the tight thresholds ($\vartheta_{\text{TDDFT}} = 10^{-6}$, $\vartheta_{\text{pre}} = 10^{-4}$) should thus, in general, be sufficient because all errors are below 10^{-3} a.u. A looser preLinK threshold ($\vartheta_{\text{pre}} = 10^{-3}$ a.u.), which is expected to give μH accuracy for ground-state properties, is still accurate for excited-state properties and could be used in extensive application calculations such as NAMD simulations, where observables are determined as ensemble averages.

The effect of the preLinK gradient threshold ($\vartheta_{\text{pre}}^{\nabla}$) on the error is not straightforward (right side of Table 1). However, we recommend it to be at least as tight as the integral threshold ($\vartheta_{\text{INT}} = 10^{-10}$). Because the difference between $\vartheta_{\text{pre}}^{\nabla} = 10^{-11}$ and $\vartheta_{\text{pre}}^{\nabla} = 10^{-10}$ is negligibly small, we apply the latter throughout this work.

4.2. Scaling with the System Size. To compare timings on CPUs and GPUs and to investigate the effective scaling behavior of the excited-state integral routines on GPUs, we use integral timings of linear polyethylenes (III_n) and dialkylethenes (IV_n) (for structures, see Figure 1). In Figure 2, we compare CPU and GPU timings of Coulomb (Figure 2a,c) and exchange (Figure 2b,d) calculations at the PBE0/def2-SVP level of theory. We show contractions of the ground-state density ($J/K(P)$), the transition density of the first excited state ($J/K(X_1 + Y_1)$) (Figure 2a,b), the integral derivatives involving the ground state and the relaxed difference density of the first excited state ($J^c/K^c(P, P_i)$), and the transition density of the first excited state ($J^c/K^c(X_1 + Y_1, X_1 + Y_1)$) (Figure 2c,d). The same integrals are shown in Figure 3, where we compare the timings of III_n and IV_n on GPUs. The effective scaling behaviors of the integral evaluations are shown in the Supporting Information. Definitions of the integral contractions are given in eqs 15 and 16. Details of the calculations are given in Sections 3.1 and 3.2.

For the Coulomb and exchange kernels of the transition densities, we observe similar speed-ups (up to a factor of 5) as for the ground-state kernels (Figure 2a,b). The accelerations of the gradient kernels stem (nearly exclusively) from their Coulomb part, which is extremely efficient on GPUs (Figures

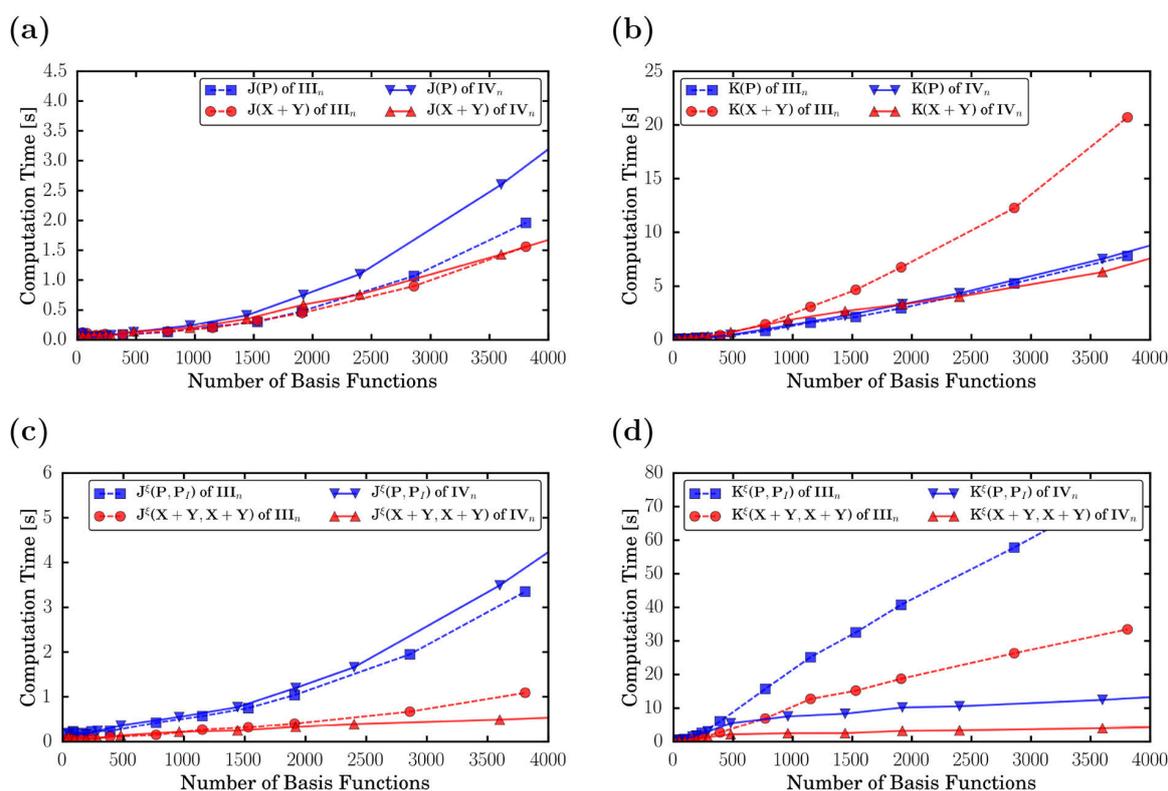


Figure 3. Timings of (a, c) Coulomb and (b, d) exchange integral evaluations (a, b; eq 15) and their derivatives with respect to the nuclear coordinates (c, d; eq 16) of polyethine (III_n , solid line) and dialkylethene (IV_n , dashed line) with $n = 1-100$ calculated at the PBE0/def2-SVP level of theory on GPUs. For details on the calculations and the computational setup, see Sections 3.1 and 3.2.

2c). We also see minor speed-ups for $K^c(\text{X}_1 + \text{Y}_1, \text{X}_1 + \text{Y}_1)$ (Figures 2d). The evaluation of $K^c(\text{P}, \text{P}_i)$ is significantly slower because the GPU routine has to be called twice. This is due to the fact that the analytical exchange evaluation on GPUs^{48–51} does not exploit the full symmetry of the two-electron integral. Please note that this problem can be solved by applying a seminumerical exchange scheme, which is currently developed for GPUs in our group.⁹⁵ For the largest molecule (III_{100}), a speed-up of two for the entire determination of an excited-state gradient involving calculations of the ground-state and excited-state energy and gradient is observed. Here, it should be stressed that the linear algebra and the evaluation of the exchange-correlation kernels are performed entirely on CPUs.

When comparing the effective scaling behavior of the Coulomb integrals of III_n (Supporting Information), one can see that the scaling of the GPU integrals is slightly larger than the scaling of the CPU integrals. This is due to the formal quadratic scaling of the J-engine employed on GPUs, in comparison to the (asymptotically linear scaling) CFMM method. This larger scaling behavior is, however, irrelevant as the prefactor of the routines is greatly reduced (Figure 2a,c). The scaling of the exchange integrals is slightly reduced by exploiting the preLinK method on GPUs. We observe ~ 1.5 and ~ 1.0 scalings for the integrals and the integral derivatives of these system sizes, respectively.

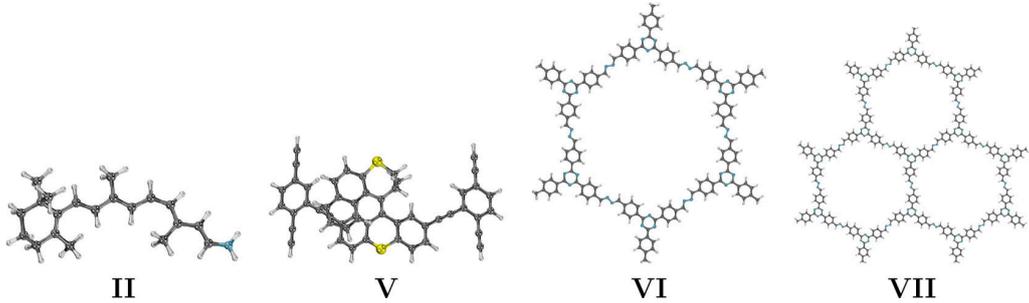
The excitation in III_n is delocalized over the entire molecule, leading to similar scaling behavior for the ground-state and excited-state properties. To show the performance of a system with local excitation, we also investigate IV_n , where only one

double bond instead of a conjugated system is excited. This leads to massive speed-ups (Figure 3) of the excited-state exchange integrals and reduces their scaling significantly (Supporting Information). For the exchange integral derivatives, we even observe sublinear scaling behavior as a result of the preLinK screening. The effect on the Coulomb integrals is smaller, mainly because of the fact that these routines take only a few seconds even for the largest investigated molecules.

4.3. Example Calculations. To demonstrate the applicability of our excited-state properties routines in the FermiONs++ program package,^{50–52} we study four molecules of interest in modern excited-state research and show their timings at the PBE0/def2-SVP and PBE/def2-TZVP levels of theory on GPUs (Table 2). We investigate the Schiff base of retinal (**II**), a model system of the chromophore in rhodopsin,⁴ a nanocar⁷⁹ (**V**) using a rotary molecular motor, and one (**VI**) and three (**VII**) pores of a covalent organic framework,⁸⁰ which catalyze the formation of hydrogen from water when exposed to light. Details of the calculations and the computational setup are again listed in Sections 3.1 and 3.2.

Table 2 shows that with the presented implementation on GPUs, excited-state properties and dynamics become accessible even for large systems and when applying DFT methods with exact exchange (e.g., PBE0) or triple- ζ basis sets. In our examples, the calculations of excited-state properties at the PBE/def2-TZVP level of theory are even faster than at the PBE0/def2-SVP level of theory mainly due to the fast Coulomb contractions discussed above. The only exception is $t(\omega_1^{\xi})$ of **VII**, for which the Z-vector equation converges

Table 2. Molecular Structures and Computational Times (t (s)) of the Schiff Base of Retinal (II), a Motorized Nanocar (V), and One (VI) and Three (VII) Pores of a Covalent Organic Framework, Calculating Ground-State Energies (E_0) and Gradients (E_0^{\ddagger}), Excited-State Energies (ω_1) and Gradients (ω_1^{\ddagger}), and Nonadiabatic Coupling Vectors ($\tau_{I \rightarrow J}$) at the PBE0/def2-SVP and PBE/def2-TZVP Levels of Theory on GPUs^a



	II	V	VI	VII
formula	C ₂₀ H ₃₀ N ⁺	C ₃₁ H ₂₈ S ₂	C ₁₄₄ H ₁₀₂ N ₃₀	C ₃₁₂ H ₂₁₃ N ₆₉
N _{atoms}	51	81	276	594
	PBE0/def2-SVP			
$t(E_0)$ /step	1.3	4.4	17.8	67.7
$t(\omega_1)$ /(step and state)	1.6	6.3	34.4	179.5
$t(E_0^{\ddagger})$	12.36	54.5	140.9	334.3
$t(\omega_1^{\ddagger})$	65.4	247.9	668.4	1392.6
$t(\tau_{0 \rightarrow 1})$	24.3	123.4	375.2	914.8
$t(\tau_{1 \rightarrow 2})$	82.7	349.4	959.0	1858.0
	PBE/def2-TZVP			
$t(E_0)$ /step	1.8	1.5	51.5	272.6
$t(\omega_1)$ /(step and state)	2.1	1.7	30.2	95.3
$t(E_0^{\ddagger})$	6.0	5.2	65.3	241.9
$t(\omega_1^{\ddagger})$	52.5	47.9	576.7	1942.9
$t(\tau_{0 \rightarrow 1})$	12.3	10.1	124.4	377.0
$t(\tau_{1 \rightarrow 2})$	56.1	52.7	523.2	1060.3

^aPlease note that $t(E_0)$ and $t(\omega_1)$ are given per step (and state). For details on the calculations and the computational setup, see Sections 3.1 and 3.2.

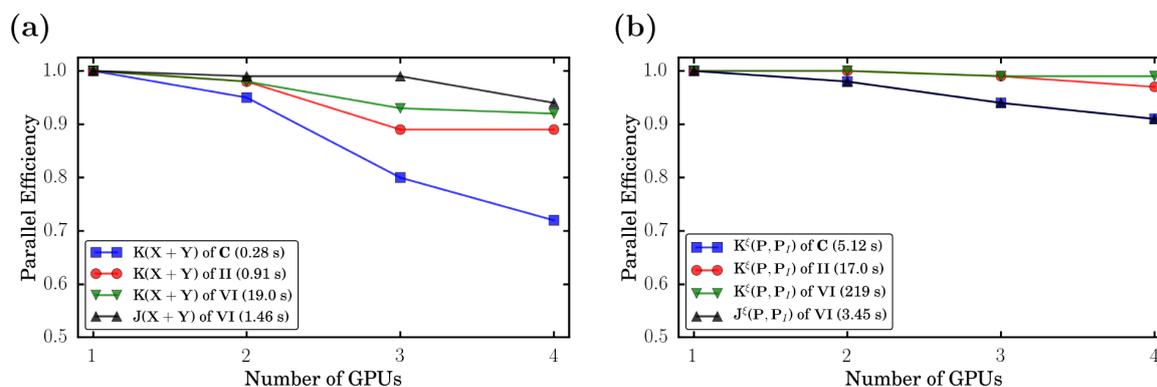


Figure 4. Parallel efficiency of selected GPU integral routines for the calculations of C, II, and VI. The computational time of the integral evaluation using four GPUs is given in parentheses. For details on the calculations and the computational setup, see Section 3.

slowly at the PBE/def2-TZVP level of theory (seven instead of four iterations).

When comparing CPU and GPU timings of the entire calculations, we observe speed-ups of three and eight for VI and VII, respectively. This clearly shows that the use of GPUs becomes more attractive with increasing system size, where the integral evaluations dominate the overall computational time. However, even for relatively small molecule II we already obtain a speed-up of two. These speed-ups exclusively stem from the Coulomb and exchange calculations as the linear

algebra and the evaluation of the exchange-correlation kernels are performed entirely on CPUs. The entire calculations of II and VII at the PBE0/def2-SVP level of theory take ~5 min and ~5 h, respectively, enabling NAMD simulations of II and the calculation of excited-state properties of VII on a reasonable time scale and yet good accuracy.

4.4. Scaling with the Computational Resources. In Figure 4, we show the parallel efficiency of $J/K(X_I + Y_I)$ (Figure 4a) and $J^z/K^z(P, P_I)$ (Figure 4b) for II, VI, and rotary molecular machine C (C₁₀H₁₃FN), which will be investigated

in the next section using up to four GPUs. The Coulomb integrals of **II** and **C** are not analyzed because their computational time is too short.

Figure 4 again shows the suitability of GPUs for large molecular systems. For the time-consuming integral evaluations of **VI**, we observe a nearly perfect scaling of >0.9 . This also indicates that adding even more GPUs will still lead to decent speed-ups of the calculations. The parallel efficiency observed in the case of the smaller molecules is lower but still remarkable considering that some computational times are <1 s. This strong scaling makes the use of GPUs also attractive for small to medium-sized molecules, as shown in the rotary molecular machine example presented in this work.

5. ILLUSTRATIVE EXAMPLES: ROTARY MOLECULAR MACHINES

As a prototypical example, we investigate the properties and the dynamics of four newly designed rotary molecular machines. For structures and definitions, see Figure 5. Similar to previous ab initio studies,^{41–44} our machines contain a $C=N^+$ motif, which is also present in the chromophore of rhodopsin. Upon excitation, the molecule should rotate around the central C–C double bond. The fluoride substituent should accelerate this rotation because of the steric repulsion, whereas the puckering of the six-membered ring should influence the direction of the rotation.^{43,44} Here, we want to investigate the influence of the atom or group X adjacent to the central double bond on the light-driven rotation of the molecule. We use CH_2 , NH, S, and O as X denoted as **C**, **N**, **S**, and **O**, respectively.

When calculating the excited-state properties and dynamics of the rotary molecular machines, we switch to the Tamm–Dancoff approximation (TDA).⁶³ As discussed in Section 2.2, this accelerates the calculation and leads to more stable trajectories close to conical intersections.^{39,64} A comparison of the excited-state energies and properties at the TDA and RPA levels of theory is presented in the Supporting Information, where we show mean differences as well as plots of the difference density, the first excited-state gradient, and the NACV between the ground state and the first excited state. We observe an average difference below 10^{-2} a.u., with the state ordering and the shape of relaxed difference densities not being affected. The comparison of TDA and CASCF energies (Supporting Information) also proves the suitability of TDA for the investigated problem. Trends between the systems and the energies close to the conical intersection agree remarkably well. Moreover, we apply the looser accurate thresholds ($\vartheta_{pre} = 10^{-3}$, $\vartheta_{pre}^v = 10^{-10}$, and $\vartheta_{TDDFT} = 10^{-5}$) because they introduce only a maximum average error of a few 10^{-5} a.u. (Supporting

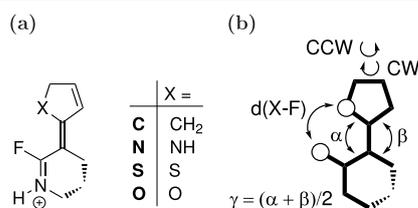


Figure 5. (a) Structures of rotary molecular machines **C**, **N**, **S**, and **O** and (b) definitions of the X–F distance ($d(X-F)$), the dihedral angle (γ), and the direction of rotation: clockwise (CW) and counter-clockwise (CCW).

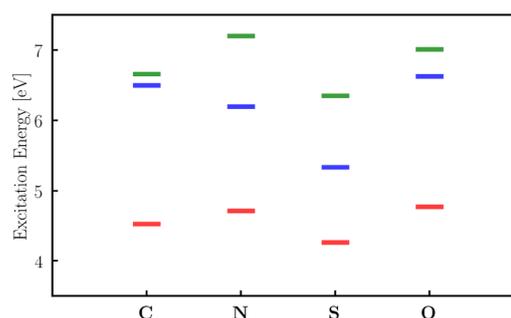


Figure 6. Singlet excitation energies of the four rotary molecular machines (**C**, **N**, **S**, and **O**) calculated at the TDA (wb97/def2-SVP) level of theory. For details on the calculations, see Section 3.3.

Information). Additional details on the calculations are listed in Section 3.3.

In Figure 6, we show the excitation energies of **C**, **N**, **S**, and **O**. Changing the substituent has an influence on the bright S_1 state of the rotary molecular motor. While the difference density plots look similar for all molecules, the excitation energies increase ($S < C < N < O$) and the directions of the gradients and NACVs change slightly. This is due to the different electronegativities of the substituents ($C \approx S < N < O$) and the fact that S, N, and O atoms have free electron pairs.

To study their effect on the light-induced rotation, we conducted NAMD simulations (105 trajectories for each rotor). In Table 3, we list the percentage of the trajectories that showed a rotation (η) as well as the ratio of clockwise and counterclockwise rotations (r). The direction of the rotation has been determined using the dihedral γ . (See the Supporting Information for plots.) The time-dependent occupation of the S_1 state (calculated as an average over all trajectories) and $d(X-F)$ for the trajectory with the fastest rotation of **C**, **N**, **S**, and **O**, respectively, are shown in Figure 7.

For all system, we observed (1) the expected rotation around the central C–C bond, which can be detected using either γ or $d(X-F)$ and (2) the relaxation from S_1 to the ground state (S_0). Movies of the rotations are available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>. The decay of the S_1 population (Figure 7a) seems to correlate with the rotational speed of the corresponding rotary molecular machine (Figure 7b). **C** shows the fastest rotation and the fastest decay, followed by **S** and **O**, which behave nearly identically. The rotation of **N** is the slowest, so is the decay of the S_1 occupancy. The trend in the rotational speeds ($N < S < O < C$) can easily be explained. **O** rotates slower than **C** because the nuclear repulsion between X and F is smaller. In the case of **S**, the repulsion of the S atom should be similar to that of the CH_2 moiety, but the time for the rotation increases because of its larger mass. The rotation of **N** is even slower because of the hydrogen bond between the NH moiety and the F atom.

However, this trend cannot be observed when looking at the rotational efficiency (η , see Table 3). Although **C** shows the largest number of trajectories featuring a rotation (70%), the η of **S** and **O** ($\sim 55\%$) is significantly smaller than the η of **N** (61%). The reason for this might be the slightly different coupling vectors between the ground state and the first excited state (Supporting Information). In contrast to refs 43 and 44, we do not observe an influence of the puckering of the ring on the rotation. The ratios of CW and CCW rotations (r) are

Table 3. Efficiency (η , Determined from the Percentage of Rotating Molecules in the NAMD Simulations) of the Four Rotary Molecular Machines (C, N, S, and O) and Ratio [$r = n(\text{CW})/n(\text{CCW})$] of CW and CCW Rotations

	C	N	S	O
η	70%	61%	54%	56%
r	0.68	0.78	1.19	1.68

close to 1 for all four rotors, and no general trend is visible in the γ plots (Supporting Information).

Changing X thus has an interesting effect on the rotary molecular machines. While the excitation energy changes significantly, the ability of the molecule to rotate around the central C–C bond is (nearly) preserved. These systems may thus be a good starting point for the design of a series of molecular rotors with tunable excitation energy. Our GPU-based routines aide these investigations by accelerating NAMD simulations even for these small systems. Because of the strong scaling (shown in Figure 4), we observe speed-ups in Coulomb and exchange integral evaluations (even when their computing time is shorter than 1 s), leading to a total speed-up of two with respect to a calculation entirely on CPUs. One trajectory took ~ 1 day on two Intel Xeon CPU E5 2640 v4 @ 2.20 GHz (20 threads) CPUs and four AMD FirePro 3D W8100 GPUs.

6. CONCLUSIONS

Throughout this work, we have examined the use of graphics processing units (GPUs) for the evaluation of two-electron integrals in excited-state energies and property (gradients and nonadiabatic coupling vectors) calculations at the time-dependent density functional theory (TDDFT) level of theory. Similar to ground-state calculations, we observe that the use of GPUs along with the J-engine^{73,74} and the preLinK scheme^{50,51} leads to decent speed-ups of the integral calculations while additionally showing a reduced scaling behavior depending on the locality of the examined excitation. These speed-ups may even become larger with our currently developed seminumerical exchange scheme.⁹⁵ By using GPUs, nonadiabatic molecular dynamics become more efficient for large but also small molecules (as the a result of the strong scaling) without a loss of accuracy or the introduction of further assumptions. As a first example, we investigated a series of newly designed rotary molecular machines showing that one

can tune the excitation energy of these systems without losing their ability to rotate by changing the heterocycle. For future applications, we consider extending the current implementation toward decoherence corrections⁹⁶ and triplet states.⁹⁷

NOTATION

All equations in this section use the same notation as in Section 2. g_{xc} are the third order exchange correlation functional derivatives. The linear transformations H^+ and H^- are defined as follows:

$$H_{pq}^+[\mathbf{V}] = \sum_{rs} \{2(pqrs) + 2f_{pqrs}^{xc} - c_x[(psrq) - (prlsq)]\} V_{rs}$$

$$H_{pq}^-[\mathbf{V}] = \sum_{rs} \{c_x[(psrq) - (prlsq)]\} V_{rs}$$
(19)

EXCITED-STATE ENERGY GRADIENTS

The unrelaxed difference density matrix \mathbf{T} is calculated as follows:

$$\mathbf{T}_I = \frac{1}{2} \begin{pmatrix} -(\mathbf{R}_I^T \mathbf{R}_I + \mathbf{L}_I^T \mathbf{L}_I)_{ij} & 0 \\ 0 & (\mathbf{R}_I \mathbf{R}_I^T + \mathbf{L}_I \mathbf{L}_I^T)_{ab} \end{pmatrix}$$
(20)

To obtain the Lagrangian multiplier \mathbf{Z}_I it is necessary to solve the following Z-vector equation:

$$\sum_{jb} (\mathbf{A} + \mathbf{B})_{jab} Z_{jb}^I = -U_{ia}^I$$
(21)

$$U_{ia}^I = \sum_b \{R_{ib}^I H_{ab}^+[\mathbf{R}_I] + L_{ib}^I H_{ab}^-[\mathbf{L}_I]\} - \sum_j \{R_{ja}^I H_{ji}^+[\mathbf{R}_I] + L_{ja}^I H_{ji}^-[\mathbf{L}_I]\} + H_{ia}^+[\mathbf{T}_I] + 2 \sum_{jbkc} \{g_{iajbkc}^{xc} R_{jb}^I R_{kc}^I\}$$
(22)

Solving this Z-vector equation thus leads to the calculation of \mathbf{Z}_I , which is used to determine the relaxed difference density matrix (\mathbf{P}_I):

$$\mathbf{P}_I = \mathbf{T}_I + \mathbf{Z}_I$$
(23)

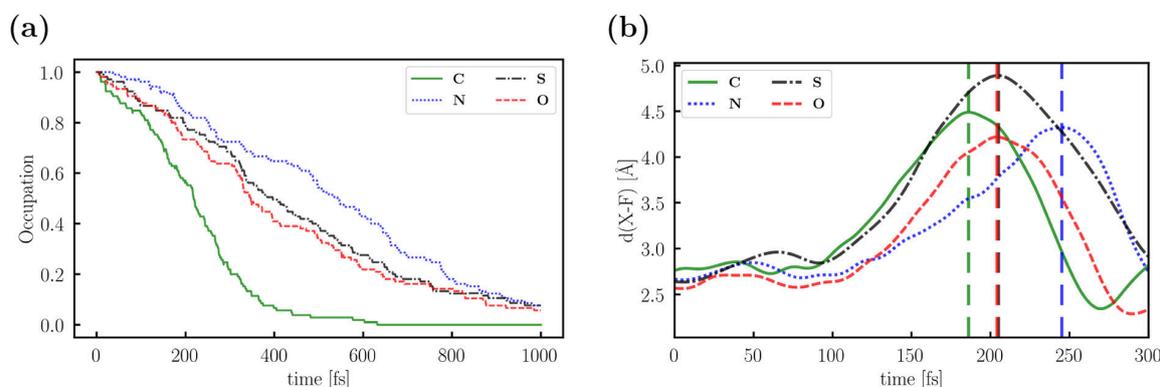


Figure 7. (a) Time-dependent decay of the occupancy of the S₁ state determined as a mean of all nonadiabatic molecular dynamics simulations. (b) Change in $d(\text{X-F})$ during one selected trajectory of C, N, S, and O, respectively, showing the rotation of the molecule around the central C–C bond. The maxima (180° rotation) are marked by the vertical dashed lines.

Using the calculated matrices, the energy-weighted difference density matrix (\mathbf{W}_I) can be formed:

$$\begin{aligned} W_{ij}^I &= \sum_a \omega_i \{R_{ia}^I L_{ja}^I + L_{ia}^I R_{ja}^I\} \\ &\quad - \sum_a \epsilon_a \{R_{ia}^I R_{ja}^I + L_{ia}^I L_{ja}^I\} \\ &\quad + H_{ij}^+[\mathbf{P}_I] + 2 \sum_{kcl} \{g_{ijkcl}^{xc} R_{kc}^I R_{ld}^I\}, \\ W_{ab}^I &= \sum_i \omega_i \{R_{ia}^I L_{ib}^I + L_{ia}^I R_{ib}^I\}, \\ &\quad + \sum_i \epsilon_i \{R_{ia}^I R_{ib}^I + L_{ia}^I L_{ib}^I\} \\ W_{ia}^I &= \sum_j \{R_{ja}^I H_{ji}^+[\mathbf{R}_I] + L_{ja}^I H_{ji}^-[\mathbf{L}_I]\} + \epsilon_i Z_{ia}^I \end{aligned} \quad (24)$$

The calculation of the effective two-particle difference density matrix is

$$\begin{aligned} \Gamma_{\mu\nu\kappa\lambda}^I &= \frac{1}{2} \{2P_{\mu\nu}^I D_{\kappa\lambda} + 2R_{\mu\nu}^I R_{\kappa\lambda}^I - c_x [P_{\mu\lambda}^I P_{\kappa\nu} + P_{\mu\kappa}^I P_{\lambda\nu}] \\ &\quad + R_{\mu\lambda}^I R_{\kappa\nu}^I + R_{\mu\kappa}^I R_{\lambda\nu}^I - L_{\mu\lambda}^I L_{\kappa\nu}^I + L_{\mu\kappa}^I L_{\lambda\nu}^I\} \end{aligned} \quad (25)$$

NONADIABATIC COUPLING VECTORS BETWEEN THE GROUND AND AN EXCITED STATE

First the Lagrangian multipliers \mathbf{Z}_{0I} and \mathbf{W}_{0I} need to be determined. In order to obtain \mathbf{Z}_{0I} the following Z-vector equation has to be solved:

$$\sum_{jb} (\mathbf{A} + \mathbf{B})_{jab} Z_{jb}^{0I} = L_{ia}^I \quad (26)$$

An iterative solution of eq 26 is not necessary. \mathbf{Z}_{0I} is equal to \mathbf{R}_I and the relaxed difference density matrix (\mathbf{P}_{0I}) is calculated as follows:

$$\mathbf{P}_{0I} = \mathbf{Z}_{0I} = \frac{1}{\omega_I} \mathbf{R}_I \quad (27)$$

The calculation of \mathbf{W}_{0I} is presented in eq 28.

$$\begin{aligned} W_{ij}^{0I} &= H_{ij}^+[\mathbf{P}_{0I}], \\ W_{ab}^{0I} &= 0, \\ W_{ia}^{0I} &= \frac{1}{2} L_{ia}^I + \epsilon_i P_{ia}^{0I} \end{aligned} \quad (28)$$

Γ_{0I} is defined as

$$\Gamma_{\mu\nu\kappa\lambda}^{0I} = \frac{1}{2} (2P_{\mu\nu}^{0I} P_{\kappa\lambda} - c_x [P_{\mu\lambda}^{0I} P_{\kappa\nu} + P_{\mu\kappa}^{0I} P_{\lambda\nu}]) \quad (29)$$

NONADIABATIC COUPLING VECTORS BETWEEN TWO EXCITED STATES

The unrelaxed difference density matrix (\mathbf{T}_{IJ}) is calculated as follows:

$$\begin{aligned} T_{ij}^{IJ} &= -\frac{1}{4} (\mathbf{R}_I^T \mathbf{R}_J + \mathbf{R}_J^T \mathbf{R}_I + \mathbf{L}_I^T \mathbf{L}_J + \mathbf{L}_J^T \mathbf{L}_I)_{ij}, \\ T_{ab}^{IJ} &= \frac{1}{4} (\mathbf{R}_I \mathbf{R}_J^T + \mathbf{R}_J \mathbf{R}_I^T + \mathbf{L}_I \mathbf{L}_J^T + \mathbf{L}_J \mathbf{L}_I^T)_{ab} \end{aligned} \quad (30)$$

To obtain the Lagrangian multiplier \mathbf{Z}_{IJ} it is necessary to solve the following Z-vector equation:

$$\sum_{jb} (\mathbf{A} + \mathbf{B})_{jab} Z_{jb}^{IJ} = -U_{ia}^{IJ} \quad (31)$$

$$\begin{aligned} U_{ia}^{IJ} &= \frac{1}{2} \sum_b \{R_{ib}^I H_{ab}^+[\mathbf{R}_J] + L_{ib}^I H_{ab}^-[\mathbf{L}_J]\} \\ &\quad + \frac{1}{2} \sum_b \{R_{ib}^J H_{ab}^+[\mathbf{R}_I] + L_{ib}^J H_{ab}^-[\mathbf{L}_I]\} \\ &\quad - \frac{1}{2} \sum_j \{R_{ja}^I H_{ji}^+[\mathbf{R}_J] + L_{ja}^I H_{ji}^-[\mathbf{L}_J]\} \\ &\quad - \frac{1}{2} \sum_j \{R_{ja}^J H_{ji}^+[\mathbf{R}_I] + L_{ja}^J H_{ji}^-[\mathbf{L}_I]\} + H_{ia}^+[\mathbf{T}_{IJ}] \\ &\quad + 2 \sum_{jbc} \{g_{iajbc}^{xc} R_{jb}^I R_{kc}^J\} \end{aligned} \quad (32)$$

\mathbf{Z}_{IJ} is used to calculate the relaxed difference density matrix (\mathbf{P}_{IJ}):

$$\mathbf{P}_{IJ} = \mathbf{T}_{IJ} + \mathbf{Z}_{IJ} \quad (33)$$

The calculation of \mathbf{W}_{IJ} is

$$\begin{aligned} W_{ij}^{IJ} &= -\frac{1}{4} \sum_a \epsilon_a \{R_{ia}^I R_{ja}^I + L_{ia}^I L_{ja}^I + \\ &\quad R_{ia}^J R_{ja}^J + L_{ia}^J L_{ja}^J\} \\ &\quad + H_{ij}^+[\mathbf{P}_{IJ}] + 2 \sum_{kcl} \{g_{ijkcl}^{xc} R_{kc}^I R_{ld}^J\} \\ &\quad - \frac{1}{2} \sum_a \omega_i \{R_{ia}^I L_{ja}^I + L_{ia}^I R_{ja}^I\} \\ &\quad - \frac{1}{2} \sum_a \omega_j \{R_{ia}^J L_{ja}^J + L_{ia}^J R_{ja}^J\}, \\ W_{ab}^{IJ} &= \frac{1}{4} \sum_i \epsilon_i \{R_{ia}^I R_{ib}^I + L_{ia}^I L_{ib}^I + \\ &\quad R_{ia}^J R_{ib}^J + L_{ia}^J L_{ib}^J\} \\ &\quad + \frac{1}{2} \sum_i \omega_i \{R_{ia}^I L_{ib}^I + L_{ia}^I R_{ib}^I\} \\ &\quad + \frac{1}{2} \sum_i \omega_j \{R_{ia}^J L_{ib}^J + L_{ia}^J R_{ib}^J\}, \\ W_{ia}^{IJ} &= \frac{1}{2} \sum_j \{R_{ja}^I H_{ji}^+[\mathbf{R}_J] + L_{ja}^I H_{ji}^-[\mathbf{L}_J]\} \\ &\quad + \frac{1}{2} \sum_j \{R_{ja}^J H_{ji}^+[\mathbf{R}_I] + L_{ja}^J H_{ji}^-[\mathbf{L}_I]\} + \epsilon_i Z_{ia}^{IJ} \end{aligned} \quad (34)$$

Γ_{IJ} is defined as

$$\Gamma_{\mu\nu\kappa\lambda}^{IJ} = \frac{1}{2} \{ 2P_{\mu\nu}^{IJ} P_{\kappa\lambda} + 2R_{\mu\nu}^I R_{\kappa\lambda}^J - c_x [P_{\mu\lambda}^{IJ} P_{\kappa\nu} + P_{\mu\kappa}^{IJ} P_{\lambda\nu}] + R_{\mu\lambda}^I R_{\kappa\nu}^J + R_{\mu\kappa}^I R_{\lambda\nu}^J - L_{\mu\lambda}^I L_{\kappa\nu}^J + L_{\mu\kappa}^I L_{\lambda\nu}^J \} \quad (35)$$

■ ASSOCIATED CONTENT

5 Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.9b00859>.

log–log plots of the integral timings for the determination of the scaling behavior; validation of the use of the TDA and accurate thresholds, with excited-state energies, relaxed difference density plots, and plots of excited-state gradients and nonadiabatic coupling vectors at the TDA and RPA levels of theory; behavior of the dihedral γ during all trajectories (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: c.ochsenfeld@fkf.mpg.de.

ORCID

Christian Ochsenfeld: 0000-0002-4189-6558

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Gökçen Savaşçı for providing the optimized structures of the covalent organic framework pores (VI and VII) and Thomas Schnappinger and Prof. Regina de Vivie-Riedle (LMU Munich) for helpful discussions. Financial support was provided by the Innovative Training Network “Computational Spectroscopy in Natural Sciences and Engineering” (ITN-COSINE) and the DFG cluster of excellence (EXC 2089) “e-conversion”. C.O. acknowledges further support as a Max-Planck Fellow at the MPI-FKF Stuttgart.

■ REFERENCES

- (1) Turro, N. J. *Modern Molecular Photochemistry*; University Science Books: Sausalito, CA, 1991.
- (2) Avouris, P.; Gelbart, W. M.; El-Sayed, M. A. Nonradiative electronic relaxation under collision-free conditions. *Chem. Rev.* **1977**, *77*, 793–833.
- (3) May, V.; Kühn, O. *Charge and Energy Transfer Dynamics in Molecular Science*; Wiley-VCH: Weinheim, 2005.
- (4) Lozier, R. H.; Bogomolni, R. A.; Stoerkenius, W. Bacteriorhodopsin: a light-driven proton pump in *Halobacterium Halobium*. *Biophys. J.* **1975**, *15*, 955–962.
- (5) Koumura, N.; Zijlstra, R. W. J.; van Delden, R. A.; Harada, N.; Feringa, B. L. Light-driven monodirectional molecular rotor. *Nature* **1999**, *401*, 152–155.
- (6) Guentner, M.; Schildhauer, M.; Thumser, S.; Mayer, P.; Stephenson, D.; Mayer, P. J.; Dube, H. Sunlight-powered kHz rotation of a hemithioindigo-based molecular motor. *Nat. Commun.* **2015**, *6*, 8406.
- (7) Astumian, R. D. How molecular motors work—insights from the molecular machinist’s toolbox: the Nobel prize in Chemistry 2016. *Chem. Sci.* **2017**, *8*, 840–845.
- (8) Dreuw, A. Quantum chemical methods for the investigation of photoinitiated processes in biological systems: Theory and applications. *ChemPhysChem* **2006**, *7*, 2259–2274.
- (9) Roos, B. A new method for large-scale CI calculations. *Chem. Phys. Lett.* **1972**, *15*, 153–159.
- (10) Trofimov, A. B.; Krivdina, I. L.; Weller, J.; Schirmer, J. Algebraic-diagrammatic construction propagator approach to molecular response properties. *Chem. Phys.* **2006**, *329*, 1–10.
- (11) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.
- (12) Casida, M. E. Time-Dependent Density Functional Response Theory for Molecules. In *Recent Advances in Density Functional Methods*; Chong, D. P., Ed.; Recent Advances in Computational Chemistry; World Scientific, 1995; Part 1, pp 155–192.
- (13) Casida, M. E. Time-dependent density-functional theory for molecules and molecular solids. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 3–18.
- (14) Dreuw, A.; Head-Gordon, M. Failure of time-dependent density functional theory for long-range charge-transfer excited states: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.
- (15) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.
- (16) Laurent, A. D.; Jacquemin, D. TD-DFT benchmarks: A review. *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.
- (17) Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **2002**, *117*, 7433–7447.
- (18) Chernyak, V.; Mukamel, S. Density-matrix representation of nonadiabatic couplings in time-dependent density functional (TDDFT) theories. *J. Chem. Phys.* **2000**, *112*, 3572.
- (19) Baer, R. Non-adiabatic couplings by time-dependent density functional theory. *Chem. Phys. Lett.* **2002**, *364*, 75–79.
- (20) Hu, C.; Hirai, H.; Sugino, O. Nonadiabatic couplings from time-dependent density functional theory: Formulation in the Casida formalism and practical scheme within modified linear response. *J. Chem. Phys.* **2007**, *127*, No. 064103.
- (21) Tavernelli, I.; Tapavicza, E.; Rothlisberger, U. Non-adiabatic dynamics using time-dependent density functional theory: Assessing the coupling strengths. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 22–29.
- (22) Tavernelli, I.; Tapavicza, E.; Rothlisberger, U. Nonadiabatic coupling vectors within linear response time-dependent density functional theory. *J. Chem. Phys.* **2009**, *130*, 124107.
- (23) Send, R.; Furche, F. First-order nonadiabatic couplings from time-dependent hybrid density functional response theory: Consistent formalism, implementation, and performance. *J. Chem. Phys.* **2010**, *132*, No. 044107.
- (24) Tavernelli, I.; Curchod, B. F. E.; Laktionov, A.; Rothlisberger, U. Nonadiabatic coupling vectors for excited states within time-dependent density functional theory in the Tamm-Dancoff approximation and beyond. *J. Chem. Phys.* **2010**, *133*, 194104.
- (25) Fatehi, S.; Alguire, E.; Shao, Y.; Subotnik, J. E. Analytic derivative couplings between configuration-interaction-singles states with built-in electron-translation factors for translational invariance. *J. Chem. Phys.* **2011**, *135*, 234105.
- (26) Ou, Q.; Fatehi, S.; Alguire, E.; Shao, Y.; Subotnik, J. E. Derivative couplings between TDDFT excited states obtained by direct differentiation in the Tamm-Dancoff approximation. *J. Chem. Phys.* **2014**, *141*, No. 024114.
- (27) Alguire, E. C.; Ou, Q.; Subotnik, J. E. Calculating derivative couplings between time-dependent Hartree–Fock excited states with pseudo-wavefunctions. *J. Phys. Chem. B* **2015**, *119*, 7140–7149.
- (28) Subotnik, J. E.; Alguire, E. C.; Ou, Q.; Landry, B. R.; Fatehi, S. The requisite electronic structure theory to describe photoexcited nonadiabatic dynamics: Nonadiabatic derivative couplings and diabatic electronic couplings. *Acc. Chem. Res.* **2015**, *48*, 1340–1350.
- (29) Ou, Q.; Alguire, E. C.; Subotnik, J. E. Derivative couplings between time-dependent density functional theory excited states in the random-phase approximation based on pseudo-wavefunctions: Behavior around conical intersections. *J. Phys. Chem. B* **2015**, *119*, 7150–7161.

- (30) Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061–1071.
- (31) Hammes-Schiffer, S.; Tully, J. C. Proton transfer in solution: Molecular dynamics with quantum transitions. *J. Chem. Phys.* **1994**, *101*, 4657–4667.
- (32) Wang, L.; Akimov, A.; Prezhdo, O. V. Recent Progress in Surface Hopping: 2011–2015. *J. Phys. Chem. Lett.* **2016**, *7*, 2100.
- (33) Crespo-Otero, R.; Barbatti, M. Recent Advances and Perspectives on Nonadiabatic Mixed Quantum-Classical Dynamics. *Chem. Rev.* **2018**, *118*, 7026.
- (34) Mai, S.; Marquetand, P.; González, L. Nonadiabatic Dynamics: The SHARC Approach. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1370.
- (35) Tapavicza, E.; Tavernelli, L.; Rothlisberger, U. Trajectory surface hopping within linear response time-dependent density-functional theory. *Phys. Rev. Lett.* **2007**, *98*, No. 023001.
- (36) Barbatti, M.; Granucci, G.; Persico, M.; Ruckebauer, M.; Vazdar, M.; Eckert-Maksić, M.; Lischka, H. The on-the-fly surface-hopping program system Newton-X: Application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *J. Photochem. Photobiol., A* **2007**, *190*, 228–240.
- (37) Barbatti, M. Nonadiabatic dynamics with trajectory surface hopping method. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 620–633.
- (38) Tully, J. C. Perspective: Nonadiabatic dynamics theory. *J. Chem. Phys.* **2012**, *137*, 22A301.
- (39) Tapavicza, E.; Bellchambers, G. D.; Vincent, J. C.; Furche, F. Ab initio non-adiabatic molecular dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18336–18348.
- (40) Plasser, F.; Crespo-Otero, R.; Pederzoli, M.; Pittner, J.; Lischka, H.; Barbatti, M. Surface hopping dynamics with correlated single-reference methods: 9H-adenine as a case study. *J. Chem. Theory Comput.* **2014**, *10*, 1395–1405.
- (41) Nikiforov, A.; Gamez, J. A.; Thiel, W.; Filatov, M. Computational Design of a Family of Light-Driven Rotary Molecular Motors with Improved Quantum Efficiency. *J. Phys. Chem. Lett.* **2016**, *7*, 105–110.
- (42) Kazaryan, A.; Lan, Z.; Schäfer, L. V.; Thiel, W.; Filatov, M. Surface hopping excited-state dynamics study of the photoisomerization of a light-driven fluorene molecular rotary motor. *J. Chem. Theory Comput.* **2011**, *7*, 2189–2199.
- (43) Oruganti, B.; Wang, J.; Durbeej, B. Quantum chemical design of rotary molecular motors. *Int. J. Quantum Chem.* **2018**, *118*, No. e25405.
- (44) Wang, J.; Oruganti, B.; Durbeej, B. Light-driven rotary molecular motors without point chirality: A minimal design. *Phys. Chem. Chem. Phys.* **2017**, *19*, 6952–6956.
- (45) Sisto, A.; Glowacki, D. R.; Martínez, T. J. Ab initio nonadiabatic dynamics of multichromophore complexes: A scalable graphical-processing-unit-accelerated exciton framework. *Acc. Chem. Res.* **2014**, *47*, 2857–2866.
- (46) Yasuda, K. Two-electron integral evaluation on the graphics processor unit. *J. Comput. Chem.* **2008**, *29*, 334–342.
- (47) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 1. strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222–231.
- (48) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 2. direct self-consistent-field implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015.
- (49) Ufimtsev, I. S.; Martínez, T. J. Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.
- (50) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.
- (51) Kussmann, J.; Ochsenfeld, C. Preselective screening for linear-scaling exact exchange-gradient calculations for graphics processing units and general strong-scaling massively parallel calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
- (52) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU integral engine for strong-scaling ab initio methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
- (53) Peters, L. D. M.; Kussmann, J.; Ochsenfeld, C. Efficient and accurate Born-Oppenheimer molecular dynamics for large molecular systems. *J. Chem. Theory Comput.* **2017**, *13*, 5479–5485.
- (54) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. Excited-state electronic structure with configuration interaction singles and Tamm-Dancoff time-dependent density functional theory on graphical processing units. *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.
- (55) Hohenstein, E. G.; Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units. *J. Chem. Phys.* **2015**, *142*, 224103.
- (56) Snyder, J. W.; Hohenstein, E. G.; Luehr, N.; Martínez, T. J. An atomic orbital-based formulation of analytical gradients and non-adiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units. *J. Chem. Phys.* **2015**, *143*, 154107.
- (57) Curchod, B. F. E.; Sisto, A.; Martínez, T. J. Ab Initio Multiple Spawning Photochemical Dynamics of DMABN Using GPUs. *J. Phys. Chem. A* **2017**, *121*, 265.
- (58) Snyder, J. W.; Curchod, B. F. E.; Martínez, T. J. GPU-Accelerated State-Averaged Complete Active Space Self-Consistent Field Interfaced with Ab Initio Multiple Spawning Unravels the Photodynamics of Provitamin D₃. *J. Phys. Chem. Lett.* **2016**, *7*, 2444.
- (59) Hollas, D.; Šišťák, L.; Hohenstein, E. G.; Martínez, T. J.; Slaviček, P. Nonadiabatic Ab Initio Molecular Dynamics with the Floating Occupation Molecular Orbital-Complete Active Space Configuration Interaction Method. *J. Chem. Theory Comput.* **2018**, *14*, 339.
- (60) Filatov, M.; Min, S. K.; Kim, K. W. Direct Nonadiabatic Dynamics by Mixed Quantum-Classical Formalism Connected with Ensemble Density Functional Theory Method: Application to *trans*-Penta-2,4-dieniminium Cation. *J. Chem. Theory Comput.* **2018**, *14*, 4499.
- (61) Bauernschmitt, R.; Ahlrichs, R. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **1996**, *256*, 454–464.
- (62) Bauernschmitt, R.; Häser, M.; Treutler, O.; Ahlrichs, R. Calculation of excitation energies within time-dependent density functional theory using auxiliary basis set expansions. *Chem. Phys. Lett.* **1997**, *264*, 573–578.
- (63) Maurice, D.; Head-Gordon, M. Configuration interaction with single substitutions for excited states of open-shell molecules. *Int. J. Quantum Chem.* **1995**, *56*, 361–370.
- (64) Cordova, F.; Doriol, L. J.; Casida, M. E. Troubleshooting time-dependent density-functional theory for photochemical applications: Oxirane. *J. Chem. Phys.* **2007**, *127*, 164111.
- (65) Chai, J. D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, No. 084106.
- (66) Li, Z.; Liu, W. First-order nonadiabatic coupling matrix elements between excited states: A Lagrangian formulation at the CIS, RPA, TD-HF, and TD-DFT levels. *J. Chem. Phys.* **2014**, *141*, No. 014110.
- (67) Li, Z.; Suo, B.; Liu, W. First order nonadiabatic coupling matrix elements between excited states: Implementation and application at the TD-DFT and pp-TDA levels. *J. Chem. Phys.* **2014**, *141*, 244105.
- (68) Ou, Q.; Bellchambers, G. D.; Furche, F.; Subotnik, J. E. First-order derivative couplings between excited states from adiabatic TDDFT response theory. *J. Chem. Phys.* **2015**, *142*, No. 064114.
- (69) Zhang, X.; Herbert, J. M. Analytic derivative couplings in time-dependent density functional theory: Quadratic response theory versus pseudo-wavefunction approach. *J. Chem. Phys.* **2015**, *142*, No. 064109.
- (70) Parker, S. M.; Roy, S.; Furche, F. Unphysical divergences in response theory. *J. Chem. Phys.* **2016**, *145*, 134105.

- (71) Zhang, X.; Herbert, J. M. Analytic derivative couplings for spin-flip configuration interaction singles and spin-flip time-dependent density functional theory. *J. Chem. Phys.* **2014**, *141*, No. 064104.
- (72) Liu, J.; Thiel, W. An efficient implementation of semiempirical quantum-chemical orthogonalization-corrected methods for excited-state dynamics. *J. Chem. Phys.* **2018**, *148*, 154103.
- (73) Reza Ahmadi, G.; Almlöf, J. The Coulomb operator in a Gaussian product basis. *Chem. Phys. Lett.* **1995**, *246*, 364–370.
- (74) White, C.; Head-Gordon, M. A J matrix engine for density functional theory calculations. *J. Chem. Phys.* **1996**, *104*, 2620–2629.
- (75) intel c/c++ compiler v19.0, see <https://software.intel.com>.
- (76) Marques, M. A. L.; Oliveira, M. J. T.; Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Comput. Phys. Commun.* **2012**, *183*, 2272–2281.
- (77) Lehtola, S.; Steigemann, C.; Oliveira, M. J.; Marques, M. A. Recent developments in LIBXC A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.
- (78) Laqua, H.; Kussmann, J.; Ochsenfeld, C. An improved molecular partitioning scheme for numerical quadratures in density functional theory. *J. Chem. Phys.* **2018**, *149*, 204111.
- (79) Morin, J.-F.; Shirai, Y.; Tour, J. M. En Route to a motorized nanocar. *Org. Lett.* **2006**, *8*, 1713–1716.
- (80) Vyas, V. S.; Haase, F.; Stegbauer, L.; Savasci, G.; Podjaski, F.; Ochsenfeld, C.; Lotsch, B. V. A tunable azine covalent organic framework platform for visible light-induced hydrogen generation. *Nat. Commun.* **2015**, *6*, 8508.
- (81) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (82) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation made simple. *Phys. Rev. Lett.* **1997**, *78*, 1396.
- (83) Ernzerhof, M.; Scuseria, G. E. Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110*, 5029.
- (84) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158.
- (85) Weigend, F.; Ahlrichs, A. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (86) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.
- (87) White, C. A.; Head-Gordon, M. Linear and sublinear scaling formation of Hartree–Fock-type exchange matrices. *Chem. Phys. Lett.* **1994**, *230*, 8.
- (88) White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M. Linear and sublinear scaling formation of Hartree–Fock-type exchange matrices. *Chem. Phys. Lett.* **1996**, *253*, 268.
- (89) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. Linear and sublinear scaling formation of Hartree–Fock-type exchange matrices. *J. Chem. Phys.* **1998**, *109*, 1663–1669.
- (90) Ochsenfeld, C. Linear scaling exchange gradients for Hartree–Fock and hybrid density functional theory. *Chem. Phys. Lett.* **2000**, *327*, 216–223.
- (91) Verlet, L. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **1967**, *159*, 98–103.
- (92) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.
- (93) Niklasson, A. M. N.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C. J.; Holmström, E.; Zheng, G.; Weber, V. Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009**, *130*, 214109.
- (94) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
- (95) Laqua, H.; Kussmann, J.; Ochsenfeld, C. Submitted for publication, 2019.
- (96) Jaeger, H. M.; Fischer, S.; Prezhdo, O. V. Decoherence-induced surface hopping. *J. Chem. Phys.* **2012**, *137*, 22A545.
- (97) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González. SHARC: *ab Initio* Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings. *J. Chem. Theory Comput.* **2011**, *7*, 1253.

Non-adiabatic molecular dynamics on graphics processing units: performance and application to rotary molecular motors

Laurens D. M. Peters,[†] Jörg Kussmann,[†] and Christian Ochsenfeld^{*,†,‡}

[†]*Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU),
Butenandtstr. 7, D-81377 München, Germany*

[‡]*Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart,
Germany*

E-mail: c.ochsenfeld@fkf.mpg.de

Contents

1	Structures	2
2	Performance	2
3	Illustrative Examples	5
3.1	Validation	5
3.2	Relaxed Difference Densities	8
3.3	Excited State Gradients and Non-adiabatic Coupling Vectors	10
3.4	Dihedrals	12

1 Structures

All optimized structures and initial conditions (structure and velocity) are available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>.

2 Performance

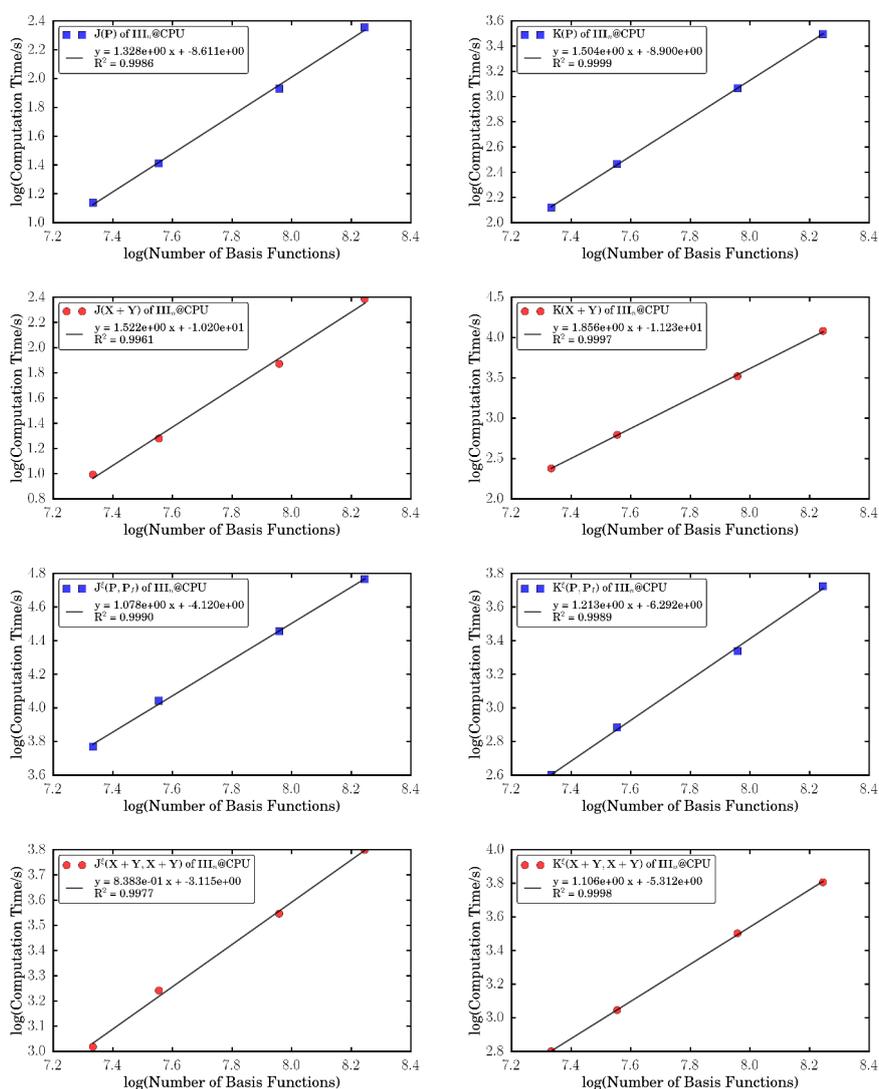


Figure S1: Log-log plot of the timings of (left) Coulomb and (right) exchange integral evaluations and their derivatives with respect to the nuclear coordinates of polyethylene (III_n) with $n = 40, 50, 75, 100$ calculated at PBE0/def2-SVP level of theory on CPUs. The slope of the linear fit is equal to the effective scaling behavior of the routine.

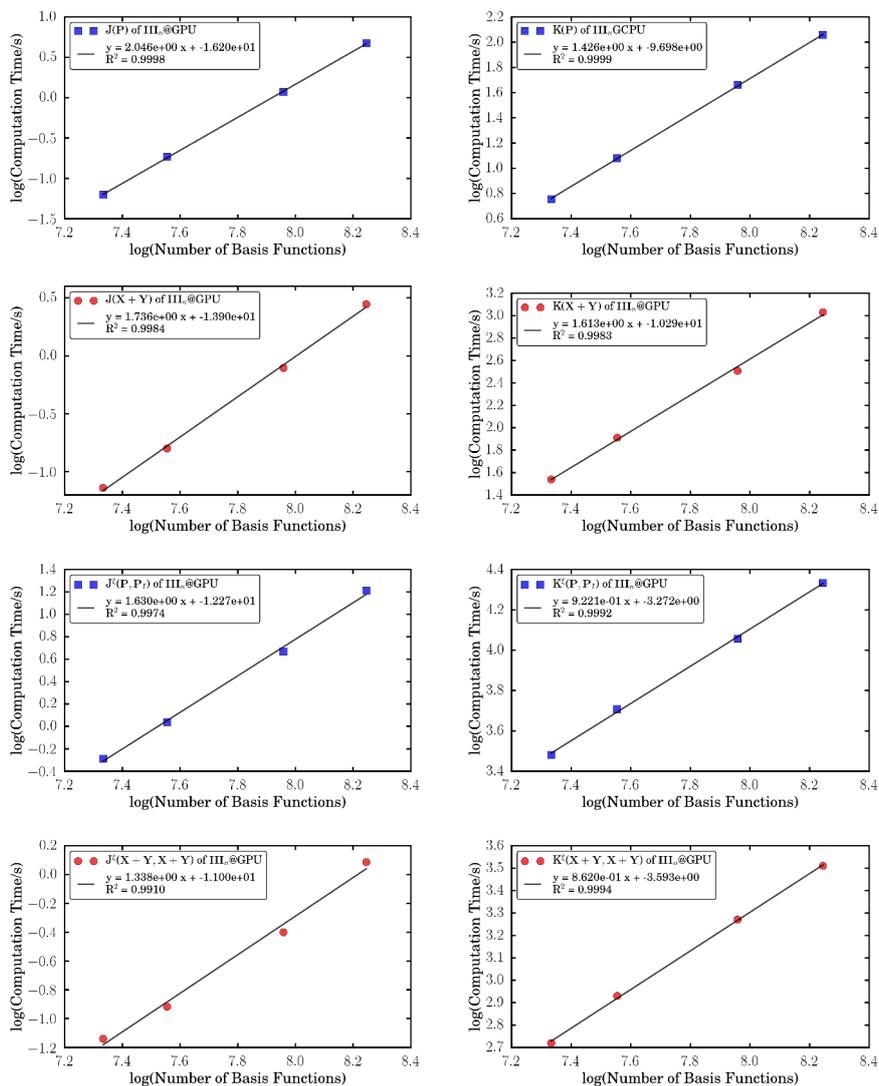


Figure S2: Log-log plot of the timings of (left) Coulomb and (right) exchange integral evaluations and their derivatives with respect to the nuclear coordinates of polyethylene (III_n) with $n = 40, 50, 75, 100$ calculated at PBE0/def2-SVP level of theory on GPUs. The slope of the linear fit is equal to the effective scaling behavior of the routine.

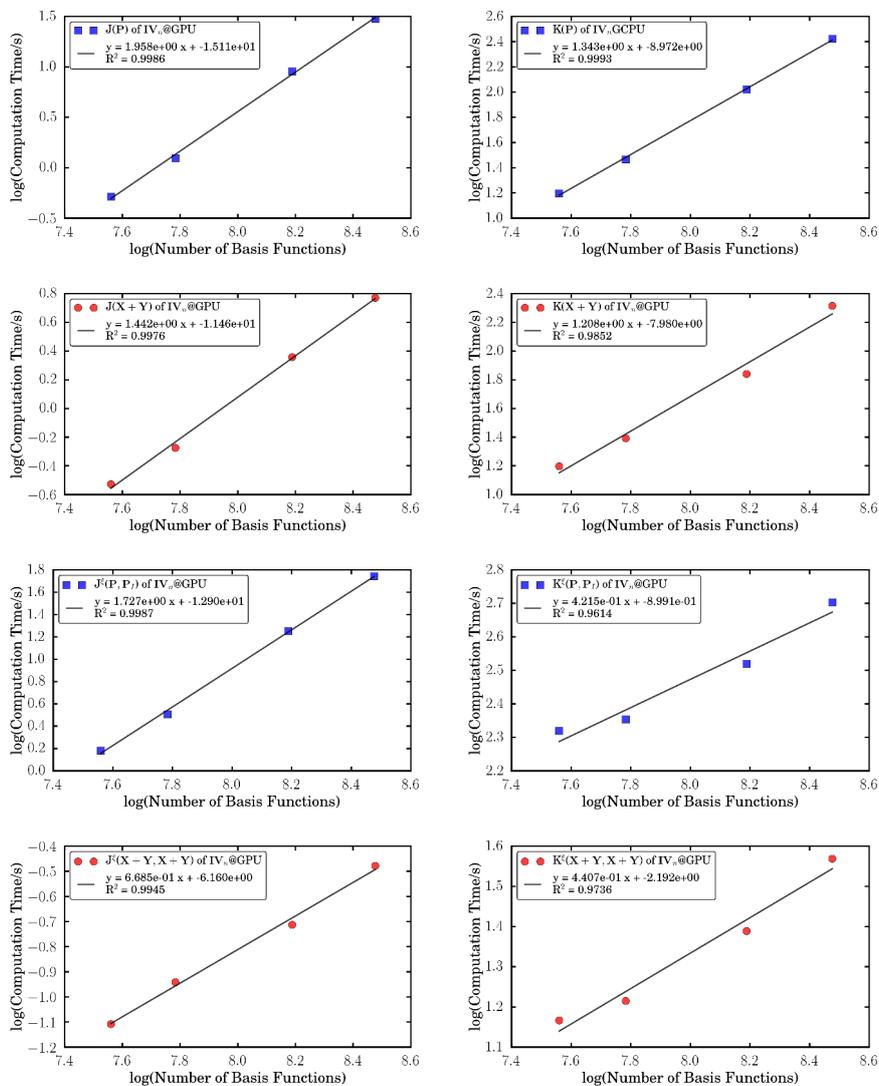


Figure S3: Log-log plot of the timings of (left) Coulomb and (right) exchange integral evaluations and their derivatives with respect to the nuclear coordinates of dialkylethene (IV_n) with $n = 40, 50, 75, 100$ calculated at PBE0/def2-SVP level of theory on GPUs. The slope of the linear fit is equal to the effective scaling behavior of the routine.

3 Illustrative Examples

3.1 Validation

Table 1: Mean absolute errors (MAE; in atomic units) of excited state energies (ω_I), gradients (ω_I^ξ), and non-adiabatic coupling vectors ($\tau_{I \rightarrow J}^\xi$) of the four rotary molecular machines (**C**, **N**, **S**, and **O**) calculated at RPA and TDA (ω B97/def2-SVP) level of theory on GPUs, comparing two different thresholds for preLink (ϑ_{pre}), the preLink gradient ($\vartheta_{\text{pre}}^\nabla$), and the TDDFT convergence (ϑ_{TDDFT}).

Screening Thresholds and Convergence Criteria		
ϑ_{pre}	10^{-3} vs 10^{-4}	
$\vartheta_{\text{pre}}^\nabla$	10^{-10} vs 10^{-11}	
ϑ_{TDDFT}	10^{-5} vs 10^{-6}	
C		
	RPA	TDA
MAE(ω_1)	3.19×10^{-4}	6.55×10^{-7}
MAE(ω_1^ξ)	4.03×10^{-5}	4.34×10^{-6}
MAE($\tau_{0 \rightarrow 1}^\xi$)	9.89×10^{-4}	2.71×10^{-5}
N		
	RPA	TDA
MAE(ω_1)	1.98×10^{-4}	2.90×10^{-6}
MAE(ω_1^ξ)	2.42×10^{-5}	1.22×10^{-5}
MAE($\tau_{0 \rightarrow 1}^\xi$)	9.43×10^{-4}	6.34×10^{-5}
S		
	RPA	TDA
MAE(ω_1)	3.91×10^{-4}	1.09×10^{-6}
MAE(ω_1^ξ)	2.84×10^{-5}	8.33×10^{-6}
MAE($\tau_{0 \rightarrow 1}^\xi$)	7.85×10^{-4}	4.32×10^{-5}
O		
	RPA	TDA
MAE(ω_1)	1.94×10^{-4}	9.07×10^{-7}
MAE(ω_1^ξ)	2.82×10^{-5}	7.30×10^{-6}
MAE($\tau_{0 \rightarrow 1}^\xi$)	6.35×10^{-4}	3.59×10^{-5}

Table 2: Mean absolute errors (MAE; in atomic units) of excitation energies (ω_I), gradients (ω_I^ξ), and non-adiabatic coupling vectors ($\tau_{I \rightarrow J}^\xi$) of the four rotary molecular machines (**C**, **N**, **S**, and **O**) calculated at RPA and TDA (ω B97/def2-SVP) level of theory on GPUs, comparing RPA and TDA.

RPA vs. TDA				
	C	N	S	O
MAE(ω_1)	6.49×10^{-3}	8.17×10^{-3}	5.41×10^{-3}	6.70×10^{-3}
MAE(ω_1^ξ)	6.43×10^{-4}	5.86×10^{-4}	5.28×10^{-4}	4.60×10^{-4}
MAE($\tau_{0 \rightarrow 1}^\xi$)	8.73×10^{-3}	8.14×10^{-3}	8.81×10^{-3}	8.86×10^{-3}

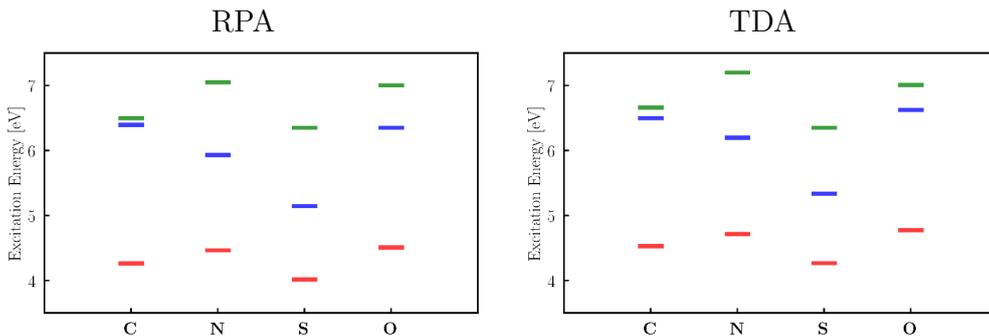


Figure S4: Excitation energies of the four rotary molecular machines (**C**, **N**, **S**, and **O**) calculated at (left) RPA and (right) TDA (ω B97/def2-SVP) level of theory.

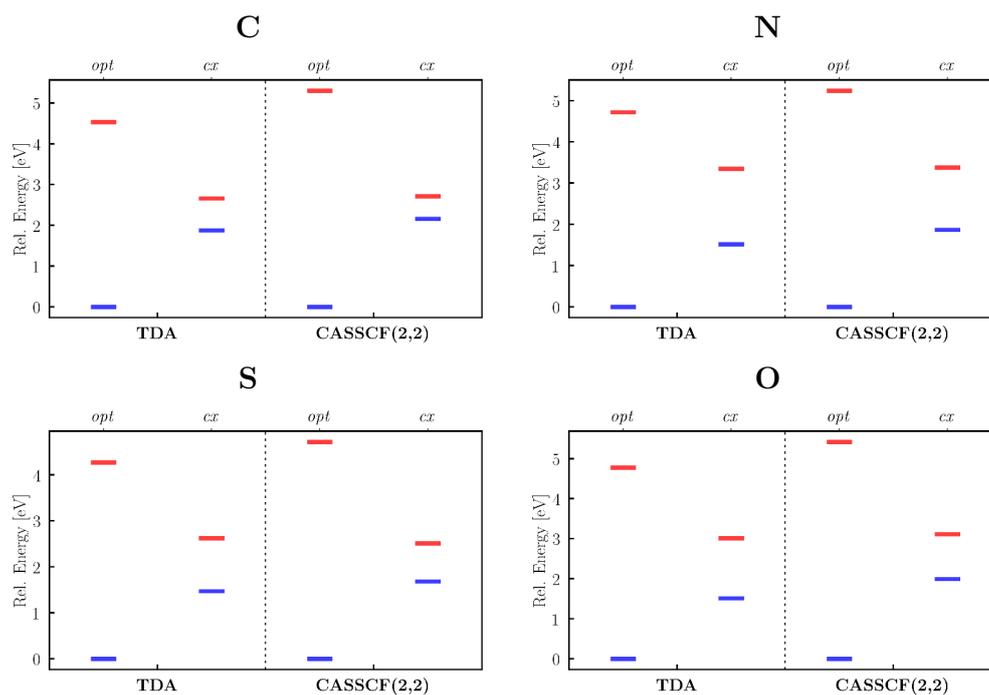


Figure S5: Comparison of relative S_0 and S_1 energies of the four rotary molecular machines (**C**, **N**, **S**, and **O**) calculated at TDA (ω B97/def2-SVP) and CASSCF(2,2)/def2-SVP level of theory. The two geometries were obtained from geometry optimizations of the ground state (*opt*) and the S_1 state (close to the conical intersection, *cx*) at TDA (ω B97/def2-SVP) level of theory. CASSCF calculations were performed with ORCA v4.0 (Neese, F. The ORCA program system, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73-78). TDA describes both states at *cx* and trends between the different molecular machines remarkably well.

3.2 Relaxed Difference Densities

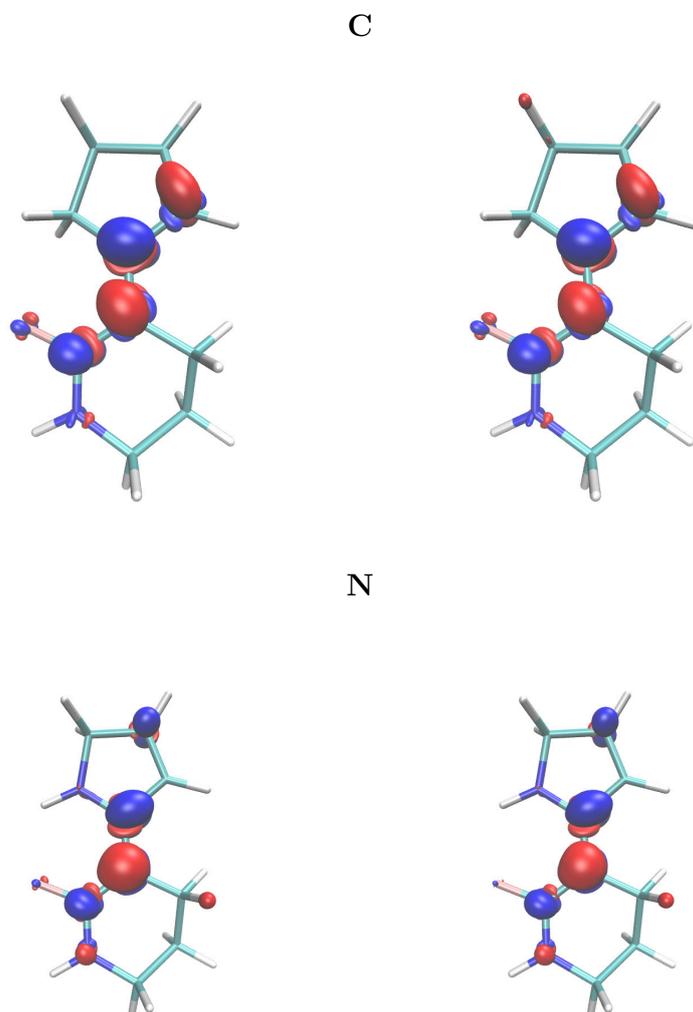


Figure S6: Plots of the relaxed difference densities of the first excited state (\mathbf{P}_1) of (up) **C** and (down) **N** calculated at (left) RPA and (right) TDA (ω B97/def2-SVP) level of theory.

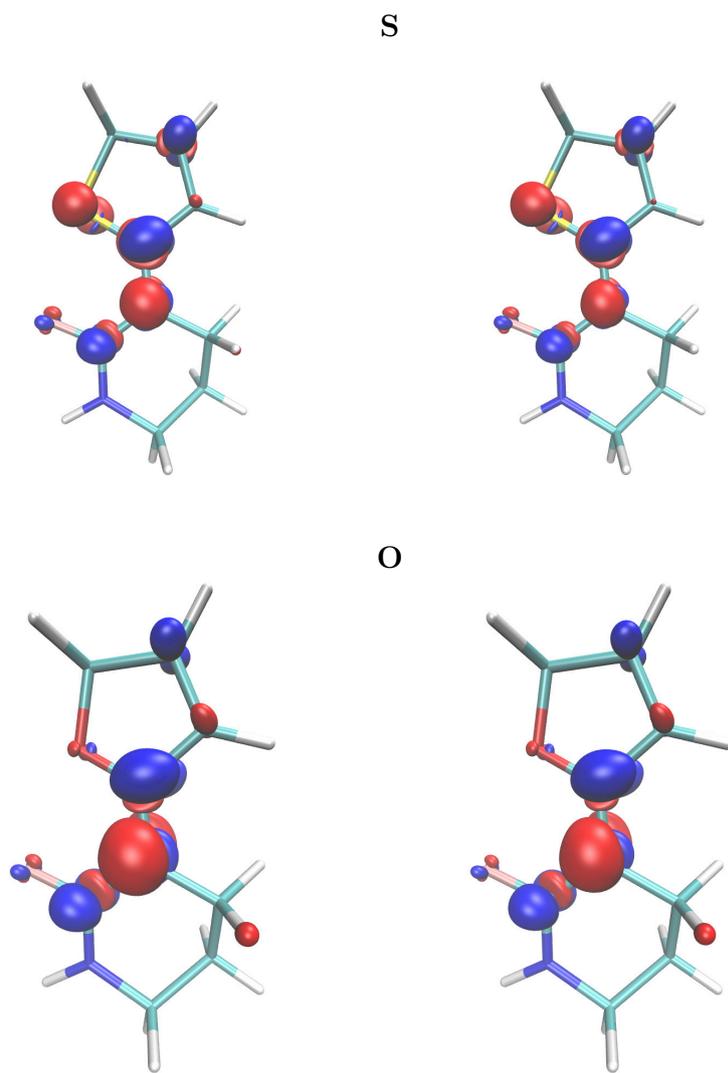


Figure S7: Plots of the relaxed difference densities of the first excited state (\mathbf{P}_1) of (up) **S** and (down) **O** calculated at (left) RPA and (right) TDA (ω B97/def2-SVP) level of theory.

3.3 Excited State Gradients and Non-adiabatic Coupling Vectors

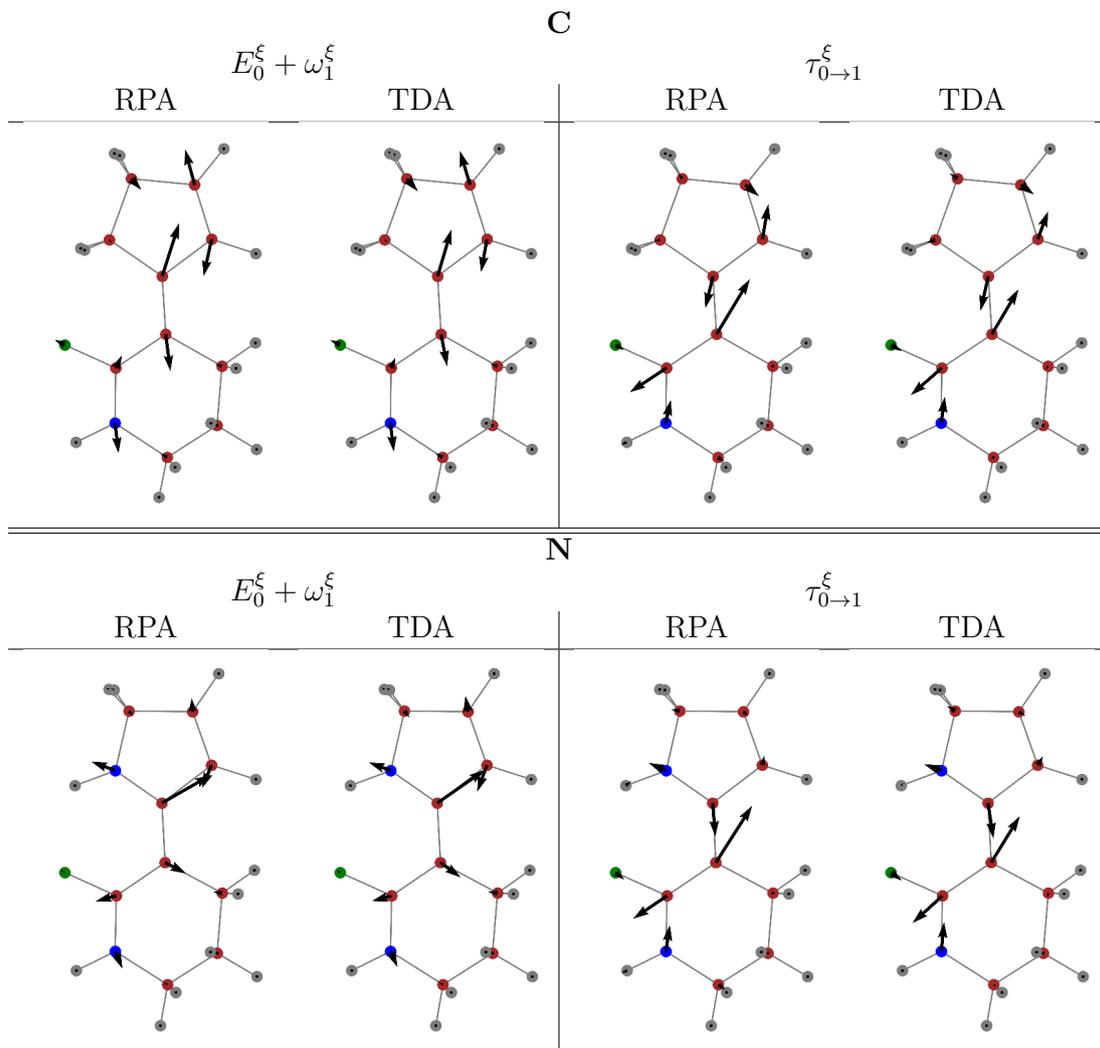


Figure S8: Excited state gradients of the first excited state ($E_0^\xi + \omega_1^\xi$) and non-adiabatic coupling vectors between the ground and the first excited state ($\tau_{0 \rightarrow 1}^\xi$) of (up) **C** and (down) **N** calculated at RPA and TDA (ω B97/def2-SVP) level of theory.

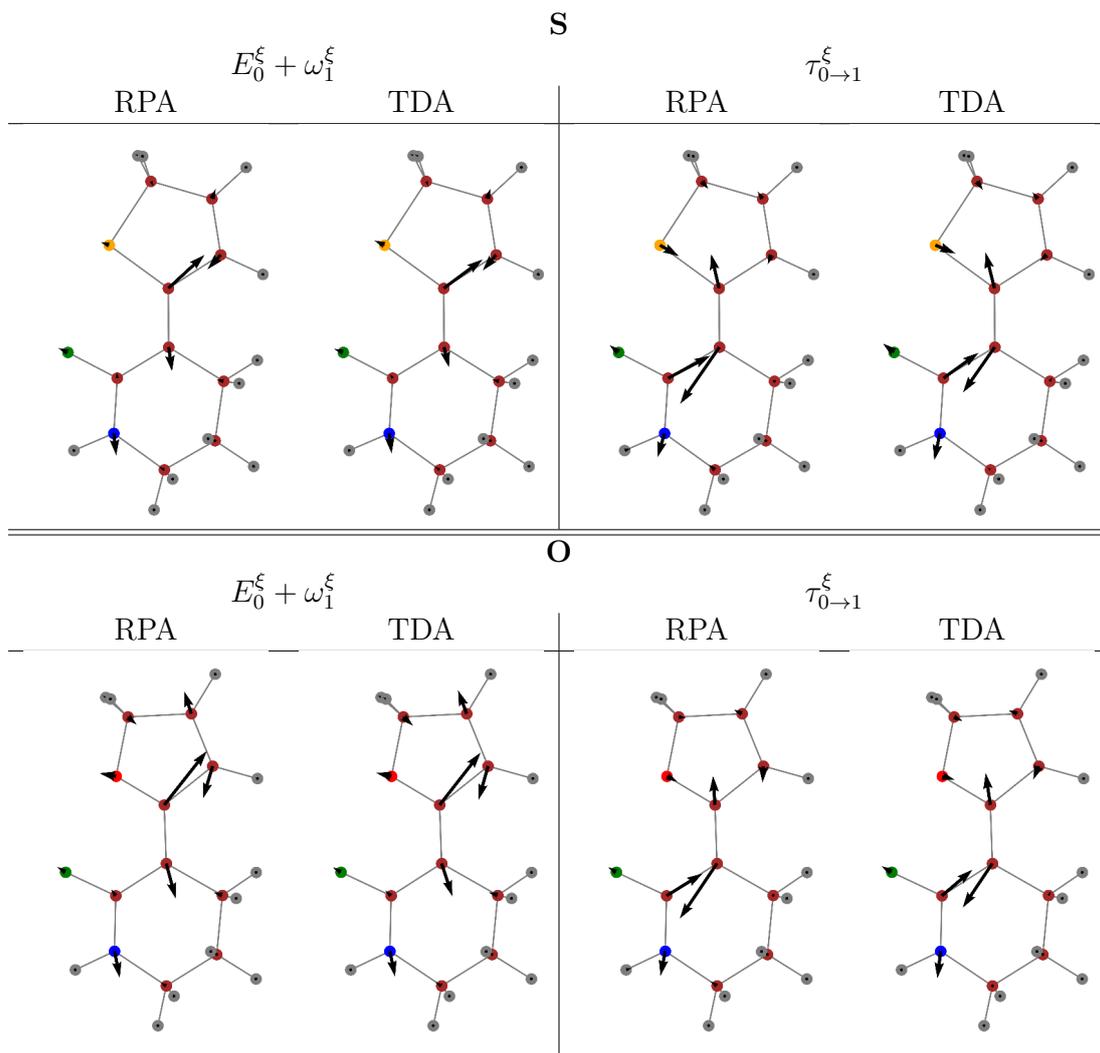


Figure S9: Excited state gradients of the first excited state ($E_0^\xi + \omega_1^\xi$) and non-adiabatic coupling vectors between the ground and the first excited state ($\tau_{0 \rightarrow 1}^\xi$) of (up) **S** and (down) **O** calculated at RPA and TDA (ω B97/def2-SVP) level of theory.

3.4 Dihedrals

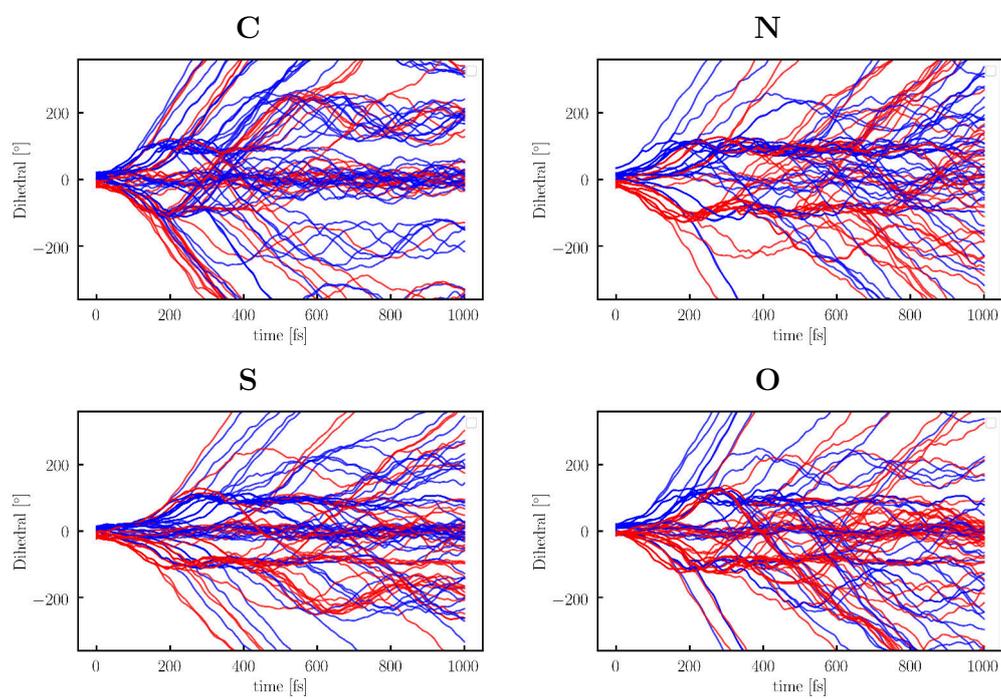


Figure S10: Dihedrals during the non-adiabatic molecular dynamics simulations of **C**, **N**, **S**, and **O**. Blue lines indicate simulations starting from a dihedral > 0 and red a dihedral < 0 . All rotors show no clear preference towards clockwise (< 0) or counterclockwise (> 0) rotations.

3.3 Manuscript III: Combining Graphics Processing Units, Simplified Time-Dependent Density Functional Theory, and Finite-Difference Couplings to Accelerate Non-adiabatic Molecular Dynamics

L. D. M. Peters, J. Kussmann, C. Ochsenfeld,

"Combining Graphics Processing Units, Simplified Time-Dependent Density Functional Theory, and Finite-Difference Couplings to Accelerate Non-adiabatic Molecular Dynamics",

in preparation

Abstract: Starting from our recently published implementation of non-adiabatic molecular dynamics (NAMD) on graphics processing units (GPUs) [*J. Chem. Theory Comput.* **15**, 6647 (2019)], we explore further approaches to accelerate *ab initio* NAMD calculations at the time-dependent density functional theory (TDDFT) level of theory. We employ (1) the simplified TDDFT schemes of Grimme *et al.* [*J. Chem. Phys.* **138**, 244104 (2013), *Comput. Theor. Chem.* **1040-1041**, 45 (2014)] and (2) the Hammes-Schiffer Tully approach [*J. Chem. Phys.* **101**, 4657 (1994)] to obtain non-adiabatic couplings from finite-difference calculations. The resulting scheme delivers an accurate physical picture while virtually eliminating the two computationally most demanding steps of the algorithm. Combined with our GPU-based integral routines for SCF, TDDFT, and TDDFT derivative calculations, NAMD simulations of systems of a few hundreds of atoms at a reasonable time scale become accessible on a single compute node. To demonstrate this and to present a first, illustrative example, we perform TDDFT/MM-NAMD simulations of the rhodopsin protein.

Combining Graphics Processing Units, Simplified Time-Dependent Density Functional Theory, and Finite-Difference Couplings to Accelerate Non-adiabatic Molecular Dynamics

Laurens D. M. Peters,¹ Jörg Kussmann,¹ and Christian Ochsenfeld^{1,2, a)}

¹⁾*Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany*

²⁾*Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart, Germany*

Starting from our recently published implementation of non-adiabatic molecular dynamics (NAMD) on graphics processing units (GPUs) [*J. Chem. Theory Comput.* **15**, 6647 (2019)], we explore further approaches to accelerate *ab initio* NAMD calculations at the time-dependent density functional theory (TDDFT) level of theory. We employ (1) the simplified TDDFT schemes of Grimme *et al.* [*J. Chem. Phys.* **138**, 244104 (2013), *Comput. Theor. Chem.* **1040-1041**, 45 (2014)] and (2) the Hammes-Schiffer Tully approach [*J. Chem. Phys.* **101**, 4657 (1994)] to obtain non-adiabatic couplings from finite-difference calculations. The resulting scheme delivers an accurate physical picture while virtually eliminating the two computationally most demanding steps of the algorithm. Combined with our GPU-based integral routines for SCF, TDDFT, and TDDFT derivative calculations, NAMD simulations of systems of a few hundreds of atoms at a reasonable time scale become accessible on a single compute node. To demonstrate this and to present a first, illustrative example, we perform TDDFT/MM-NAMD simulations of the rhodopsin protein.

^{a)}Electronic mail: c.ochsenfeld@fkf.mpg.de

Non-adiabatic molecular dynamics (NAMMD) simulations using trajectory surface hopping (TSH)¹⁻⁴ have become a powerful tool to describe dynamics of molecular systems involving multiple electronic states. Their field of application ranges from the description of rather small molecular machines⁵⁻⁹ over medium-sized photoswitches^{10,11} to the dynamics of entire photoactive proteins^{12,13}. They can be used with a variety of excited-state methods, e.g., the complete active space self-consistent field (CASSCF) method¹⁴, the algebraic-diagrammatic construction (ADC(2))¹⁵, several coupled cluster methods (e.g., CC2)¹⁶, as well as time-dependent density functional theory (TDDFT)^{17,18}. Also triplet states¹⁹ can be included.

However, the greatest challenge remains the large computational cost of NAMMD simulations, which is a result of the expensive excited-state methods mentioned above and the fact that TSH requires not only one but a series trajectories to determine observables as ensemble averages. This problem can be tackled by using semi-empirical methods^{5,6}, employing exciton models²⁰, or using graphics processing units (GPUs) to accelerate the calculations of ground- and excited-state energies and properties^{9,20-31}. Based on our recent work on the latter, we explore in our present work the use of simplified TDDFT schemes^{32,33} and the Hammes-Schiffer Tully (HST)² model in addition to GPU-based integral routines. They tackle the two major bottlenecks of NAMMD: The calculation of the state energies and the couplings between the states. After a brief summary of the corresponding theory and their validation for the investigated problems, we show timings and use our approach to simulate the photo-induced rotation of the retinal chromophore in the rhodopsin protein at TDDFT/MM level of theory. Details on the methods (thresholds, convergence criteria etc.) and the computational setup can be found in the Supporting Information.

In TSH^{1,2} a system is allowed to switch the potential energy surface (PES) within one trajectory. The occurrence of such a surface hop depends on the hopping probability: If it exceeds a randomly drawn number between zero and one, the trajectory continues on a different PES with a rescaled nuclear velocity. The average of multiple trajectories with a different series of random numbers describes the behavior of the system. The hopping probability itself is calculated from the change of the state energies and the non-adiabatic couplings (\mathbf{Q}), which can also be obtained as the product of the non-adiabatic coupling vectors ($\boldsymbol{\tau}$) and the nuclear velocity ($\dot{\mathbf{R}}$)

$$Q_{IJ} = \langle \Phi_I | \frac{\partial}{\partial t} \Phi_J \rangle = \langle \Phi_I | \frac{\partial}{\partial \mathbf{R}} \Phi_J \rangle \cdot \dot{\mathbf{R}} = \boldsymbol{\tau}_{IJ} \cdot \dot{\mathbf{R}} . \quad (1)$$

Φ_I is the wavefunction of the electronic state I . While $\boldsymbol{\tau}$ can be calculated using response theory, \mathbf{Q} cannot be determined analytically.

Energies of all considered states, gradients, and couplings are thus the main ingredients of a TSH algorithm. The first can be obtained from TDDFT by solving the TDDFT or random phase approximation (RPA) equations^{17,34,35}

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix} = \omega_I \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} \mathbf{X}_I \\ \mathbf{Y}_I \end{pmatrix}, \quad (2)$$

with ω_I being the excitation energy of state I . \mathbf{A} and \mathbf{B} are the orbital rotation Hessians and \mathbf{X}_I and \mathbf{Y}_I are the transition densities for excitation and de-excitation, respectively. Neglecting \mathbf{B} in eq. (2) is known as the Tamm-Dancoff approximation (TDA)³⁶, leading to

$$\mathbf{A}\mathbf{X}_I = \omega_I\mathbf{X}_I. \quad (3)$$

The calculation of excitation energies with eqs. (2) and (3) is time-consuming, mainly because of the evaluation of the two-electron integrals in \mathbf{A} and \mathbf{B} . To accelerate these calculations, we apply the simplified RPA and TDA methods by Grimme and coworkers^{32,33,37}. Here, Coulomb and exchange kernels are approximated (\mathbf{J}' for Coulomb and \mathbf{K}' for exchange) using the Mataga-Nishimoto-Ohno-Klopman³⁸⁻⁴⁰ damped Coulomb operators together with the transition/charge density monopoles q obtained from a Löwdin population analysis⁴¹

$$\begin{aligned} J'_{pqrs} &= \sum_{NM} q_{pq}^N \left(\frac{1}{r_{NM}^\beta + (c_x \eta_{NM})^{-\beta}} \right)^{\frac{1}{\beta}} q_{rs}^M, \\ K'_{pqrs} &= \sum_{NM} q_{pq}^N \left(\frac{1}{r_{NM}^\alpha + \eta_{NM}^{-\alpha}} \right)^{\frac{1}{\alpha}} q_{rs}^M. \end{aligned} \quad (4)$$

p, q, \dots are arbitrary molecular orbitals. r is the interatomic distance, η is the mean of the chemical hardness of the atoms N and M . α and β are global fit parameters, while c_x is the amount of exact exchange. This leads to the following approximate orbital rotation Hessians:

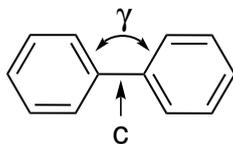
$$\begin{aligned} A'_{iajb} &= \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + s_k K'_{iajb} - J'_{ijab}, \\ B'_{iajb} &= s_k K'_{iabj} - c_x K'_{ibaj}. \end{aligned} \quad (5)$$

i, j, \dots denote occupied and a, b, \dots virtual molecular orbitals. s_k is 2 or 0 for singlet-singlet or singlet-triplet excitations and ϵ_p is the orbital energy of p .

The excitation energies (ω') are then obtained by diagonalizing \mathbf{A}' in case of sTDA or $(\mathbf{A}' - \mathbf{B}')^{\frac{1}{2}}(\mathbf{A}' + \mathbf{B}')(\mathbf{A}' - \mathbf{B}')^{\frac{1}{2}}$ in case of sRPA. To avoid the diagonalization of the entire matrix, the number of included configuration state functions (CSFs) is truncated using the thresholds ϑ_{pCSF} , ϑ_{sCSF} , and ϑ_{CSF} .³² Only primary (with an energy below ϑ_{pCSF}) and secondary CSFs (with an energy between ϑ_{pCSF} and ϑ_{CSF} and a significant coupling to the primary CSFs $> \vartheta_{\text{sCSF}}$) are considered. The sTDA scheme greatly reduces the cost of excited state calculations, giving access to absorption spectra of large molecular systems.³² The sRPA scheme yields better transition densities³³, leading to better transition dipole moments and even enabling the calculation of higher order dynamical response properties.⁴²

Excited-state gradients (ω_I^x) and $\boldsymbol{\tau}$'s can then be derived from eqs. (2) and (3) using linear response theory.⁴³⁻⁵³ The resulting equations depend on the excitation energies and transition densities. In case of sTDA/sRPA, we use the calculated ω'_I , \mathbf{X}'_I , and \mathbf{Y}'_I in the standard algorithms, as the derivatives of \mathbf{J}' and \mathbf{K}' with respect to the nuclear coordinates are, so far, neither implemented nor tested. To validate this approach, we compare optimized structures of biphenyl (**I**) at the TDDFT and sTDDFT level of theory in tab. (1) and show ω^x 's and $\boldsymbol{\tau}$'s of protonated formalimine (**II**) and the Schiff base of retinal (**III**) in fig. (1). Additional plots and a screening of the thresholds (ϑ_{pCSF} , ϑ_{CSF} , and ϑ_{sCSF}) are shown in the Supporting Information.

TABLE 1. Comparison of optimized S_1 structures of biphenyl (**I**) calculated at RPA, TDA, sRPA, and sTDA (PBE0/def2-SVP) level of theory listing the central C-C distance (c) and the dihedral γ .



	RPA	TDA	sRPA	sTDA
c [\AA]	1.42	1.44	1.44	1.44
$ \gamma $ [$^\circ$]	0.02	0.03	0.02	0.01

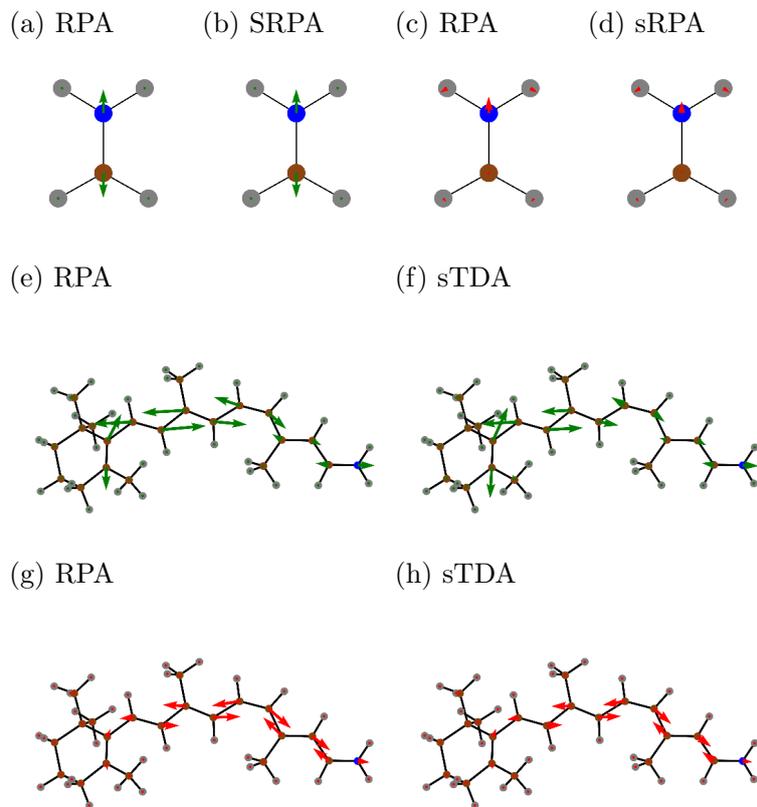


FIG. 1. (a - d) RPA and sRPA excited-state gradients of the second excited state (a + b, green) and non-adiabatic coupling vectors between the ground and the second excited state (c+d, red) of **II** at the PBE0/def2-SVP level of theory. (e - h) RPA and sTDA excited-state gradients of the first excited state (e + f, green) and non-adiabatic coupling vectors between the ground and the first excited state (g + h, red) of **III** at the ω B97/def2-SVP level of theory.

All four optimized structures of **I** in tab. (1) are nearly planar and feature a similar central C-C distance. The gradients and coupling vectors based on sTDDFT results in fig. (1) are also in good agreement with the RPA and TDA properties. The only differences are the weaker couplings of **II** and **III** and the fact that ω_1^x of **III** has a larger contribution in the six-membered ring and a smaller in the conjugated system. Both are the result of the slightly different excitation energies (**II**: 9.76 eV (RPA), 9.87 eV (sRPA); **III**: 2.74 eV (RPA), 3.17 eV (sTDA)) and transition densities. In case of **III**, two other observations can be made (see Supporting Information): sRPA performs worse than sTDA and PBE0⁵⁴⁻⁵⁷/def2-SVP^{58,59} is, in contrast to ω B97⁶⁰/def2-SVP, not able to capture the charge transfer character of the excitation.

Tab. (1) and fig. (1) indicate that simplified TDDFT might also be used for NAMD simulations which require, however, time-consuming calculations of the couplings between the states. To circumvent the analytical calculation of \mathbf{Q} via $\boldsymbol{\tau}$, we apply the numerical HST model:²

$$Q_{IJ}(t + \frac{\Delta t}{2}) \approx \frac{1}{2\Delta t}[O_{IJ}(t, t + \Delta t) - O_{IJ}(t + \Delta t, t)] = \frac{1}{2\Delta t}[\langle \Phi_I(t) | \Phi_J(t + \Delta t) \rangle - \langle \Phi_I(t + \Delta t) | \Phi_J(t) \rangle] . \quad (6)$$

Assuming that the excited-state wave functions can be obtained from the ground-state Kohn-Sham orbitals (ϕ_p) and the transition densities, O_{0I} and O_{IJ} take the following form:

$$O_{0I}(t_1, t_2) = \sum_{ia} \gamma_{ia}^{0I}(t_1) S_{ia}(t_1, t_2) , \quad (7)$$

$$O_{IJ}(t_1, t_2) = \sum_{pq} \gamma_{pq}^{IJ}(t_1) S_{pq}(t_1, t_2) + \sum_{ia} X_{ia}^I(t_1) X_{ia}^J(t_2) - \sum_{ia} Y_{ia}^I(t_1) Y_{ia}^J(t_2) , \quad (8)$$

with

$$S_{pq}(t_1, t_2) = \langle \phi_p(t_1) | \phi_q(t_2) \rangle , \quad (9)$$

$$\gamma_{0I} = \mathbf{L}_I , \quad (10)$$

$$\gamma_{IJ} = \frac{1}{2} \begin{pmatrix} -(\mathbf{R}_I^T \mathbf{L}_J + \mathbf{L}_I^T \mathbf{R}_J)_{ij} & 0 \\ 0 & (\mathbf{R}_I \mathbf{L}_J^T + \mathbf{L}_I \mathbf{R}_J^T)_{ab} \end{pmatrix} . \quad (11)$$

\mathbf{R}_I and \mathbf{L}_I are $(\mathbf{X}_I + \mathbf{Y}_I)$ and $(\mathbf{X}_I - \mathbf{Y}_I)$, respectively, in case of RPA and \mathbf{X}_I in case of TDA. The HST model is nowadays widely used in NAMD simulations⁶¹⁻⁶³ because it reduces the computational time (as shown below) and leads to more stable trajectories in the vicinity of conical intersections.⁶⁴

A validation of the presented numerical scheme for couplings used in the HST model is shown in the Supporting Information, where we have calculated numerical and analytical τ 's for formaldehyde. Additionally, we have performed NAMD simulations of **II**, using the HST model and analytically calculated τ 's as well as RPA, TDA, sRPA, and sTDA. After excitation to the S_2 state, the molecule shows a fast conversion to the S_1 state, which goes along with the elongation of the C-N bond. Further relaxation to the S_0 state is achieved via a rotation around this bond. In fig. (2) the increase ($S_2 \rightarrow S_1$) and decrease ($S_1 \rightarrow S_0$) of the S_1 occupation is shown. The change of occupation for all states as well as energies and couplings of selected trajectories are listed in the Supporting Information.

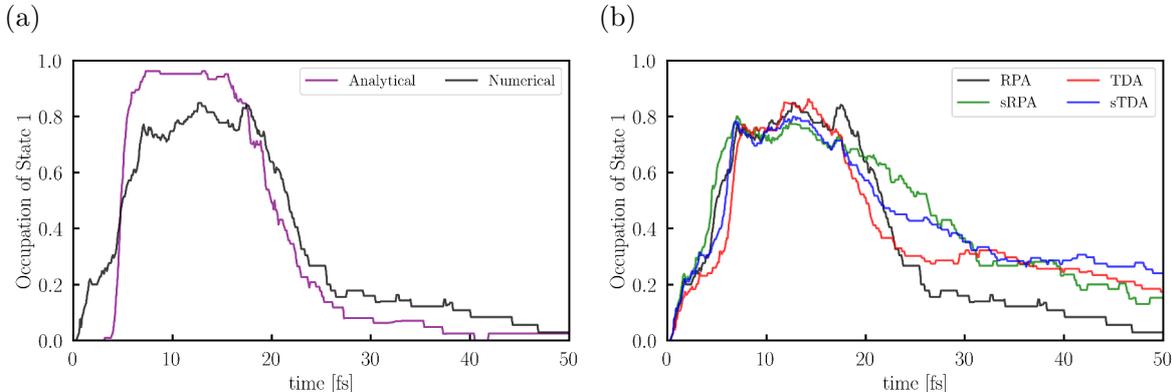


FIG. 2. Change of S_1 state occupations of protonated formalimine (**II**) calculated as an average of all NAMD simulations at (a) RPA (PBE0/def2-SVP) level of theory using analytical and numerical couplings and (b) RPA, TDA, sRPA, and sTDA (PBE0/def2-SVP) results using numerical couplings.

In all sets of trajectories, we observe a similar behavior of **II**, indicating that the HST model and the simplified TDDFT are valid approximations for this example. This is reflected in the S_1 occupations, which are very similar for all cases. The only differences are a slightly different S_2 - S_1 coupling when applying the HST model, and a slower decay of the S_1 occupation in case of the simplified TDDFT methods which is, however, also visible in case of TDA. All of the observed trends are also reflected in the couplings and τ 's. Furthermore, the results agree well with previously published data obtained from TDDFT and CASSCF simulations.⁶³ Please note that the simulations using sTDA or sRPA are not NVE simulations. Rescaling the nuclear velocities to enforce an NVE ensemble has, however, no effect on the average properties of the system.

Tab. (2) presents the impact of the approximations on the performance of a NAMD simulation of **III**. It shows that HST and sRPA virtually eliminate the computational cost of the calculations of excited-state energies and couplings. In combination with the GPU-based integral evaluations, the total speed-up in case of $N_{\text{roots}} = 2$ is ~ 4 with respect to the CPU-based RPA implementation using analytical τ 's. A NAMD simulation of **III** involving 5000 steps (e.g., 1 ps simulation using 0.2 fs time steps) can thus be conducted within ~ 5 instead of ~ 21 days. The acceleration becomes even larger when more excited states and couplings are considered (e.g., a factor of more than 20 for **III** in case of $N_{\text{roots}} = 7$). This and the fact that the performance of GPUs is better for large molecular systems (see, ref. 9) makes the presented approach interesting for the investigation of systems involving hundreds of atoms and plenty of electronic states.

TABLE 2. Computation times of ground-state energy (E_0) and gradient (E_0^x), excited-state energies (ω_I) and gradient (ω_1^x), and couplings (\mathbf{Q}) calculations of the Schiff base of retinal (**III**) at RPA and sRPA (PBE0/def2-SVP) level of theory, using a different number of roots (N_{roots}) and couplings ($\frac{(N_{\text{roots}}+1) \times N_{\text{roots}}}{2}$) as well as analytical and numerical \mathbf{Q} 's. Asterisks mark calculations that have been performed entirely on CPUs. All calculations were conducted on two Intel Xeon CPU E5 2640 v4 @ 2.20 GHz (20 threads) CPUs and four AMD FirePro 3D W8100 GPUs.

N_{roots}	\mathbf{Q}	RPA/sRPA	$t(E_0) + t(E_0^x)$	$t(\omega_I)$	$t(\omega_1^x)$	$t(\mathbf{Q})$	t (Total)
2	analytical*	RPA*	34 s	28 s	101 s	201 s	~ 6 min
2	analytical	RPA	26 s	19 s	62 s	128 s	~ 4 min
2	numerical	RPA	26 s	19 s	62 s	< 1 s	~ 2 min
2	numerical	sRPA	26 s	< 1 s	62 s	< 1 s	~ 1.5 min
3	analytical	RPA	26 s	36 s	62 s	309 s	~ 7 min
3	numerical	RPA	26 s	36 s	62 s	< 1 s	~ 2 min
3	numerical	sRPA	26 s	< 1 s	62 s	< 1 s	~ 1.5 min
7	analytical	RPA	26 s	51 s	62 s	1822 s	~ 32.5 min
7	numerical	RPA	26 s	51 s	62 s	< 1 s	~ 2.5 min
7	numerical	sRPA	26 s	< 1 s	62 s	< 1 s	~ 1.5 min

As a first application of our proposed scheme, we have calculated 42 NAMD simulations of the rhodopsin protein (**IV**) at the sTDA/MM (ω B97/def2-SVP) level of theory. The chromophore of **IV** undergoes a *cis-trans* isomerization, when exposed to light (see fig. (3)). For a review of calculations on this system, the reader is referred to ref. 65. The change of the dihedral γ_1 (defined in fig. (3b)) and the state occupations are shown in fig. (4).

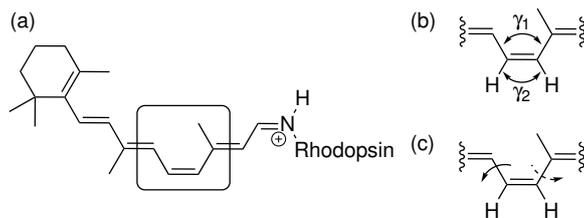


FIG. 3. (a) Structure of the chromophore of rhodopsin (**IV**). (b) Important dihedrals (γ_1 and γ_2) in and (c) the bicycle pedal isomerization mechanism of **IV**. The part of the molecule shown in (b) and (c) is located by the rectangle in (a).

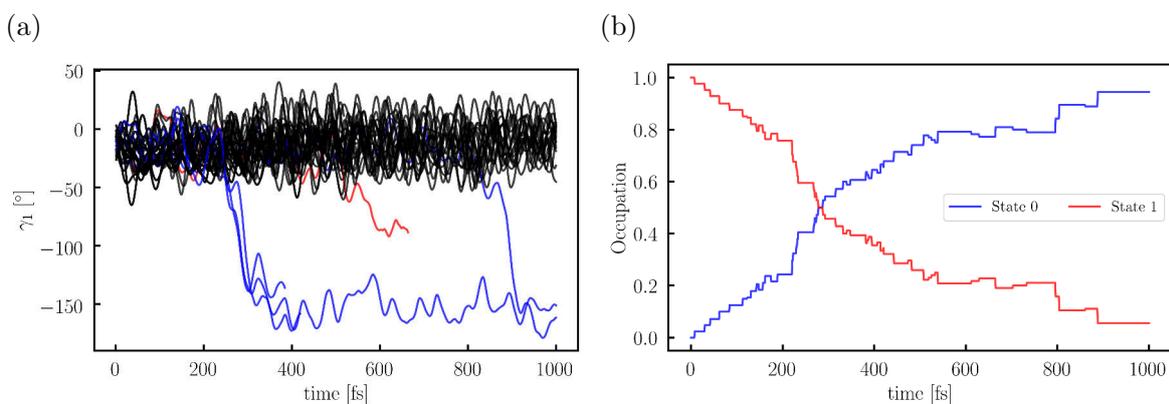


FIG. 4. NAMD simulations of **IV** at sTDA (PBE0/def2-SVP) level of theory using the HST model: (a) Change of the dihedral γ_1 indicating trajectories with no rotation (black), half rotation (red), and full rotation (blue). (b) Change of the state occupations of **IV** calculated as an average of all trajectories.

Five out of 42 calculated trajectories feature a *cis-trans* isomerization (see fig. (4a), a movie of the isomerization is available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>). Three of them reach $\gamma_1 = -90^\circ$ at ~ 280 fs, which coincides with the crossing point of the state occupations (see fig. (4b)). This hop time (t) is significantly higher and the yield (y) of 12% significantly lower than the experimental ($t = 147.7 \pm 1.0$ fs; $y = 0.63 \pm 0.01$) results reported in ref. (13). Our analysis of **III** (see fig. (1)) indicates that the weaker gradients and couplings of sTDA or the general problems of TDDFT with the system (see also ref. 65–67) may be the reason for this. However, our approach describes the direction and mechanism (see fig. (3b)) of the isomerization correctly. In contrast to previous work^{12,13,65}, our method requires, besides α , β , and the QM/MM ansatz, no further parameterizations and/or reductions of the system. One trajectory of **IV** takes ~ 5 -7 days on two Intel Xeon CPU E5 2640 v4 @ 2.20 GHz (20 threads) CPUs and four AMD FirePro 3D W8100 GPUs.

We have introduced a combination of GPU-based integral routines, simplified TDDFT schemes, and numerical couplings for efficient NAMD simulations. For all investigated systems ranging from small organic molecules (**II**) to proteins (**IV**), excited-state properties and dynamics are described qualitatively correct with a significantly reduced computational cost. The latter is due to the vanishing computational times for TDDFT energies and couplings calculations. The present approach may be used to qualitatively explore relaxation pathways and predict trends (e.g., effects of mutations or different isotopes) within these reactions.

ACKNOWLEDGMENTS

Financial support was provided by the Innovative Training Network “Computational Spectroscopy In Natural Sciences And Engineering” (ITN-COSINE) and the DFG cluster of excellence (EXC 2089) “e-conversion”. C.O. acknowledges further support as Max-Planck-Fellow at the MPI-FKF Stuttgart.

REFERENCES

- ¹J. C. Tully, *J. Chem. Phys.* **93**, 1061 (1990).
- ²S. Hammes-Schiffer and J. C. Tully, *J. Chem. Phys.* **101**, 4657 (1994).
- ³L. Wang, A. Akimov, and O. V. Prezhdo, *J. Phys. Chem. Lett.* **7**, 2100 (2016).
- ⁴R. Crespo-Otero and M. Barbatti, *Chem. Rev.* **118**, 7026 (2018).
- ⁵A. Kazaryan, Z. Lan, L. V. Schäfer, W. Thiel, and M. Filatov, *J. Chem. Theory Comput.* **7**, 2189 (2011).
- ⁶A. Nikiforov, J. A. Gamez, W. Thiel, and M. Filatov, *J. Phys. Chem. Lett.* **7**, 105 (2016).
- ⁷B. Oruganti, J. Wang, and B. Durbeej, *Int. J. Quantum Chem.* **118**, e25405 (2018).
- ⁸J. Wang, B. Oruganti, and B. Durbeej, *Phys. Chem. Chem. Phys.* **19**, 6952 (2017).
- ⁹L. D. M. Peters, J. Kussmann, and C. Ochsenfeld, *J. Chem. Theory Comput.* **15**, 6647 (2019).
- ¹⁰M. Pederzoli, J. Pittner, M. Barbatti, and H. Lischka, *J. Phys. Chem. A* **115**, 11136 (2011).
- ¹¹J. Jankowska, M. Barbatti, J. Sadlej, and A. L. Sobolewski, *Phys. Chem. Chem. Phys.* **19**, 5318 (2017).
- ¹²M. Manathunga, X. Yang, H. L. Luk, S. Gozem, L. M. Frutos, A. Valentini, N. Ferr, and M. Olivucci, *J. Chem. Theory Comput.* **12**, 839 (2016).
- ¹³C. Schnedermann, X. Yang, M. Liebel, K. M. Spillane, J. Lugtenburg, I. Fernandez, A. Valentini, I. Schapiro, M. Olivucci, P. Kukura, and R. A. Mathies, *Nat. Chem.* **10**, 449 (2018).
- ¹⁴B. Roos, *Chem. Phys. Lett.* **15**, 153 (1972).
- ¹⁵A. B. Trofimov, I. L. Krivdina, J. Weller, and J. Schirmer, *Chem. Phys.* **329**, 1 (2006).
- ¹⁶O. Christiansen, H. Koch, and P. Jörgensen, *Chem. Phys. Lett.* **243**, 409 (1995).
- ¹⁷M. E. Casida, in *Recent advances in density functional methods, part 1.*, edited by D. P. Chong (World Scientific, 1995) pp. 155–192.
- ¹⁸M. E. Casida, *J. Mol. Struct.: THEOCHEM* **914**, 3 (2009).
- ¹⁹S. Mai, P. Marquetand, and L. González, *WIREs Comput. Mol. Sci.* **8**, e1370 (2018).
- ²⁰A. Sisto, D. R. Glowacki, and T. J. Martínez, *Acc. Chem. Res.* **47**, 2857 (2014).
- ²¹K. Yasuda, *J. Comput. Chem.* **29**, 334 (2008).
- ²²I. S. Ufimtsev and T. J. Martínez, *J. Chem. Theory Comput.* **4**, 222 (2008).

- ²³I. S. Ufimtsev and T. J. Martínez, *J. Chem. Theory Comput.* **5**, 1004 (2009).
- ²⁴I. S. Ufimtsev and T. J. Martínez, *J. Chem. Theory Comput.* **5**, 2619 (2009).
- ²⁵J. Kussmann and C. Ochsenfeld, *J. Chem. Phys.* **138**, 134114 (2013).
- ²⁶J. Kussmann and C. Ochsenfeld, *J. Chem. Theory Comput.* **11**, 918 (2015).
- ²⁷J. Kussmann and C. Ochsenfeld, *J. Chem. Theory Comput.* **13**, 3153 (2017).
- ²⁸L. D. M. Peters, J. Kussmann, and C. Ochsenfeld, *J. Chem. Theory Comput.* **13**, 5479 (2017).
- ²⁹C. M. Isborn, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, *J. Chem. Theory Comput.* **7**, 1814 (2011).
- ³⁰E. G. Hohenstein, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, *J. Chem. Phys.* **142**, 224103 (2015).
- ³¹J. W. Snyder, E. G. Hohenstein, N. Luehr, and T. J. Martínez, *J. Chem. Phys.* **143**, 154107 (2015).
- ³²S. Grimme, *J. Chem. Phys.* **138**, 244104 (2013).
- ³³C. Bannwarth and S. Grimme, *Comput. Theor. Chem.* **1040-1041**, 45 (2014).
- ³⁴R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Lett.* **256**, 454 (1996).
- ³⁵R. Bauernschmitt, M. Häser, O. Treutler, and R. Ahlrichs, *Chem. Phys. Lett.* **264**, 573 (1997).
- ³⁶D. Maurice and M. Head-Gordon, *Int. J. Quantum Chem.* **56**, 361 (1995).
- ³⁷T. Risthaus, A. Hansen, and S. Grimme, *Phys. Chem. Chem. Phys.* **16**, 14408 (2014).
- ³⁸K. Nishimoto and N. Mataga, *Z. Phys. Chem.* **12**, 335 (1957).
- ³⁹K. Ohno, *Theor. Chim. Acta* **2**, 219 (1964).
- ⁴⁰G. Klopman, *J. Am. Chem. Soc.* **86**, 4450 (1964).
- ⁴¹P.-O. Löwdin, *J. Chem. Phys.* **18**, 365 (1950).
- ⁴²M. de Wergifosse and S. Grimme, *J. Chem. Phys.* **149**, 24108 (2018).
- ⁴³F. Furche and R. Ahlrichs, *J. Chem. Phys.* **117**, 7433 (2002).
- ⁴⁴R. Send and F. Furche, *J. Chem. Phys.* **132**, 44107 (2010).
- ⁴⁵S. Fatehi, E. Alguire, Y. Shao, and J. E. Subotnik, *J. Chem. Phys.* **135**, 234105 (2011).
- ⁴⁶Q. Ou, S. Fatehi, E. Alguire, Y. Shao, and J. E. Subotnik, *J. Chem. Phys.* **141**, 24114 (2014).
- ⁴⁷Z. Li and W. Liu, *J. Chem. Phys.* **141**, 14110 (2014).
- ⁴⁸Z. Li, B. Suo, and W. Liu, *J. Chem. Phys.* **141**, 244105 (2014).

- ⁴⁹Q. Ou, G. D. Bellchambers, F. Furche, and J. E. Subotnik, *J. Chem. Phys.* **142**, 64114 (2015).
- ⁵⁰Q. Ou, E. C. Alguire, and J. E. Subotnik, *J. Phys. Chem. B* **119**, 7150 (2015).
- ⁵¹X. Zhang and J. M. Herbert, *J. Chem. Phys.* **142**, 64109 (2015).
- ⁵²S. M. Parker, S. Roy, and F. Furche, *J. Chem. Phys.* **145**, 134105 (2016).
- ⁵³X. Zhang and J. M. Herbert, *J. Chem. Phys.* **141**, 64104 (2014).
- ⁵⁴J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- ⁵⁵J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **78**, 1396(E) (1997).
- ⁵⁶M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).
- ⁵⁷C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- ⁵⁸F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- ⁵⁹F. Weigend, *Phys. Chem. Chem. Phys.* **8**, 1057 (2006).
- ⁶⁰J. D. Chai and M. Head-Gordon, *J. Chem. Phys.* **128**, 84106 (2008).
- ⁶¹E. Tapavicza, I. Tavernelli, and U. Rothlisberger, *Phys. Rev. Lett.* **98**, 23001 (2007).
- ⁶²J. Pittner, H. Lischka, and M. Barbatti, *Chem. Phys.* **356**, 147 (2009).
- ⁶³I. Tavernelli, E. Tapavicza, and U. Rothlisberger, *J. Mol. Struct.: THEOCHEM* **914**, 22 (2009).
- ⁶⁴F. Plasser, R. Crespo-Otero, M. Pederzoli, J. Pittner, H. Lischka, and M. Barbatti, *J. Chem. Theory Comput.* **10**, 1395 (2014).
- ⁶⁵S. Gozem, H. L. Luk, I. Schapiro, and M. Olivucci, *Chem. Rev.* **117**, 13502 (2017).
- ⁶⁶M. Huix-Rotllant, M. Filatov, S. Gozem, I. Schapiro, M. Olivucci, and N. Ferr, *J. Chem. Theory Comput.* **9**, 3917 (2013).
- ⁶⁷O. Valsson, C. Filippi, and M. E. Casida, *J. Chem. Phys.* **142**, 144104 (2015).

Combining Graphics Processing Units, Simplified Time-Dependent Density Functional Theory, and Finite-Difference Couplings to Accelerate Non-adiabatic Molecular Dynamics - Supporting Information

Laurens D. M. Peters,[†] Jörg Kussmann,[†] and Christian Ochsenfeld^{*,†,‡}

[†]*Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU),
Butenandtstr. 7, D-81377 München, Germany*

[‡]*Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart,
Germany*

E-mail: c.ochsenfeld@fkf.mpg.de

Contents

1	Computational Details	3
1.1	General Remarks	3
1.2	Excited-State Calculations	3
1.3	Timings	3
1.4	Excited-State Optimizations	4
1.5	QM/MM Calculations of Rhodopsin (IV)	4
1.6	Initial Conditions	4
1.7	NAMD Simulations	6
2	Biphenyl (I)	7
3	Protonated Formaldimine (II)	9
4	Schiff Base of Retinal (III)	14
4.1	PBE0/def2-SVP	14
4.2	ω B97/def2-SVP	16
5	Validation of the Finite Differences Calculation of Couplings	18
	References	20

1 Computational Details

1.1 General Remarks

All calculations employ the FermiONs++ program package¹⁻³. It was compiled using the GCC compiler v7.1 and the AMD APPSDK compiler for the GPU routines. In addition, the LibXC library v4.0.1^{4,5} was used. Tight thresholds were applied for the SCF convergence (10^{-7} using the FP-commutator), the integrals (10^{-10}), the preLinK method (10^{-3}), the preLinK gradient (10^{-10}), the TDDFT convergence (10^{-5}), and the Z-vector equation convergence (10^{-5}). We did not exploit the symmetry of the molecules.

1.2 Excited-State Calculations

In all calculations (single point, timings, NAMD simulations), we solely calculated singlet excitations using the gm5 grid⁶, either the PBE0⁷⁻¹⁰ or the ω B97¹¹ method, and the def2-SVP^{12,13} basis set. For **I**, **II**, **III**, **IV**, and formaldehyde, three, three, two, two, and seven excited states were taken into account, respectively, if not stated otherwise. In case of sTDA/sRPA, the global fit parameters α , β , and c_x were set according to the literature.¹⁴⁻¹⁶ The choice of the thresholds for the truncation of the CSFs are shown in tabs. (2), (4), (6), and (8). In case of **IV**, the same thresholds as **III** were used with ϑ_{pCSF} [eV] increased to 7. Optimized ground-state geometries (available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>) were used for the single point excited state calculations.

1.3 Timings

Timings were determined as a mean of 350 NAMD steps of **III** at the PBE0/def2-SVP level of theory on two Intel Xeon CPU E5 2640 v4 @ 2.20 GHz (20 threads) CPUs and four AMD FirePro 3D W8100 GPUs. To determine $t(\omega_I)$ and $t(\mathbf{Q})$ in case of $N_{\text{roots}} = 3, 7, 20$ NAMD steps were performed. The simulations started from the optimized ground-state geometry in the first excited state. At every step, all $\frac{(N_{\text{roots}}+1) \times N_{\text{roots}}}{2}$ couplings were calculated.

1.4 Excited-State Optimizations

Excited-state optimizations of **I** at the RPA, TDA, sRPA, and sTDA (PBE0/def2-SVP) levels of theory were started from the optimized ground-state geometry following the gradient of the S_1 state. The final structures are available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>.

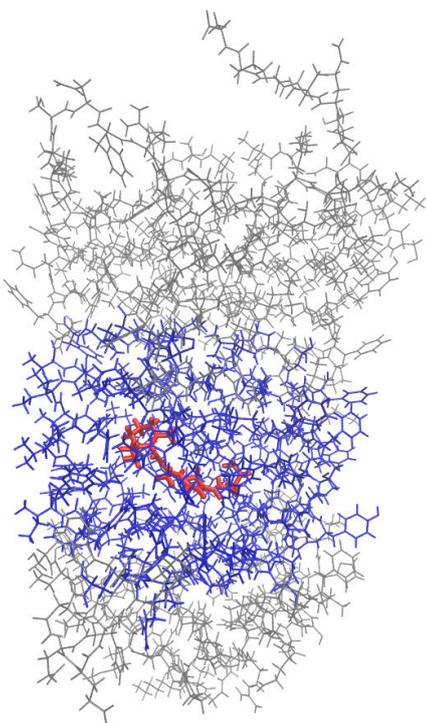
1.5 QM/MM Calculations of Rhodopsin (IV)

The crystal structure of *bovine rhodopsin* (PDB 1F88)¹⁷ was used as a starting geometry. The chromophore was parametrized with the antechamber v17.3 suite of amber¹⁸ and the GAFF2 force field was applied. Before the NVE simulation discussed in the next subsection, the entire structure was first relaxed at the MM level of theory using the NAMD suite¹⁹ (1000 steps) and then at the QM/MM level (a few hundred steps). In all QM/MM calculations, we used the same QM region (see fig. (S1b)) and kept all atoms outside a radius of 10 Å (from the chromophore) frozen (see fig. (S1a)).

1.6 Initial Conditions

Initial geometries and momenta for the NAMD simulations of **II** and **IV** (available at <https://www.cup.uni-muenchen.de/pc/ochsenfeld/download/>) were obtained from ground-state NVT simulations at the PBE0/def2-SVP level of theory using the Velocity Verlet propagator^{20,21} and a velocity rescaling thermostat²² set to 298.18 K. The time step was 0.1 fs in case of **II** and 0.2 fs in case of **IV**. 21 geometries and momenta were taken from a 10 ps production run (every 500 fs) after an equilibration of 1 ps. In case of **II**, the equilibration was performed at a constant temperature of 298.18 K, while **IV** was slowly heated from 10 K to 298.18 K in the first 900 fs. The overall rotation and translation was removed at every time step of the simulation.

(a)



(b)

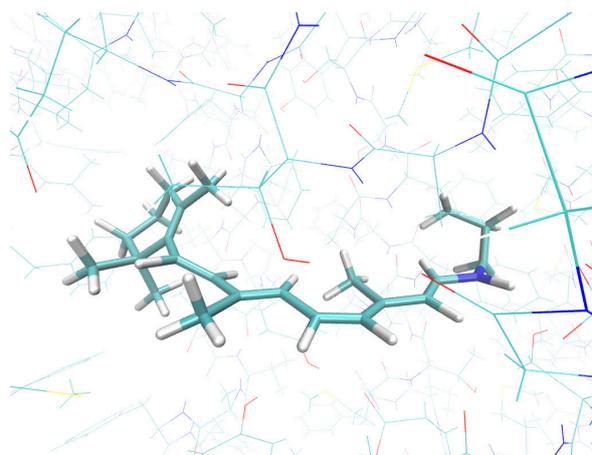


Figure S1: (a) Structure of rhodopsin (**IV**) showing the QM region (red), the active MM atoms (blue), and the frozen MM atoms (grey). (b) Expanded view of the QM region.

1.7 NAMD Simulations

NAMD simulations of **II** and **IV** were conducted at the PBE0/def2-SVP and ω B97 level of theory, respectively, using the fewest switches surface hopping algorithm^{23,24}. Again, the Velocity Verlet propagator was applied and the overall rotation and translation were removed. In case of **II**, five independent trajectories of 100 fs were started from each of the 21 initial conditions at the second excited state using a time step of 0.1 fs. Three couplings (Q_{01} , Q_{02} , and Q_{12}) were calculated either analytically or numerically (HST model) and RPA, TDA, sRPA, and sTDA were used to calculate the excited-state energies and transition densities. In case of **IV**, two independent trajectories of 1 ps (time step of 0.2 fs) were started from each of the 21 initial conditions. Here, the first excited state was the initial state and only one coupling (Q_{01}) was calculated using the HST model. The excitation energies and gradients were obtained at the sTDA level of theory.

2 Biphenyl (I)

Table 1: Comparison of RPA, TDA, sRPA, and sTDA excitation energies of **I** at the PBE0/def2-SVP level of theory.

	RPA	sRPA	TDA	sTDA
ω_1 [eV]	5.09	5.04	5.10	5.05
ω_2 [eV]	5.10	5.11	5.27	5.11

Table 2: Influence of the thresholds ϑ_{pCSF} , ϑ_{CSF} , and ϑ_{sCSF} on the first sTDA/sRPA excitation energy of **I** calculated at the PBE0/def2-SVP level of theory. The bold numbers have been applied to calculations of **I** throughout this work. Only primary configuration state functions (CSFs) with $< \vartheta_{\text{pCSF}}$ and secondary CSFs with $< \vartheta_{\text{CSF}}$ and a significant coupling to the primary CSFs ($> \vartheta_{\text{sCSF}}$) are considered.

ϑ_{pCSF} [eV]	7	7	7	7	7	12
ϑ_{CSF} [eV]	7	20	50	7	20	20
ϑ_{sCSF}	1.e-4	1.e-4	1.e-4	1.e-5	1.e-5	1.e-4
sTDA [eV]	5.07	5.05	5.05	5.07	5.05	5.05
sRPA [eV]	5.14	5.04	5.04	5.07	5.04	5.04

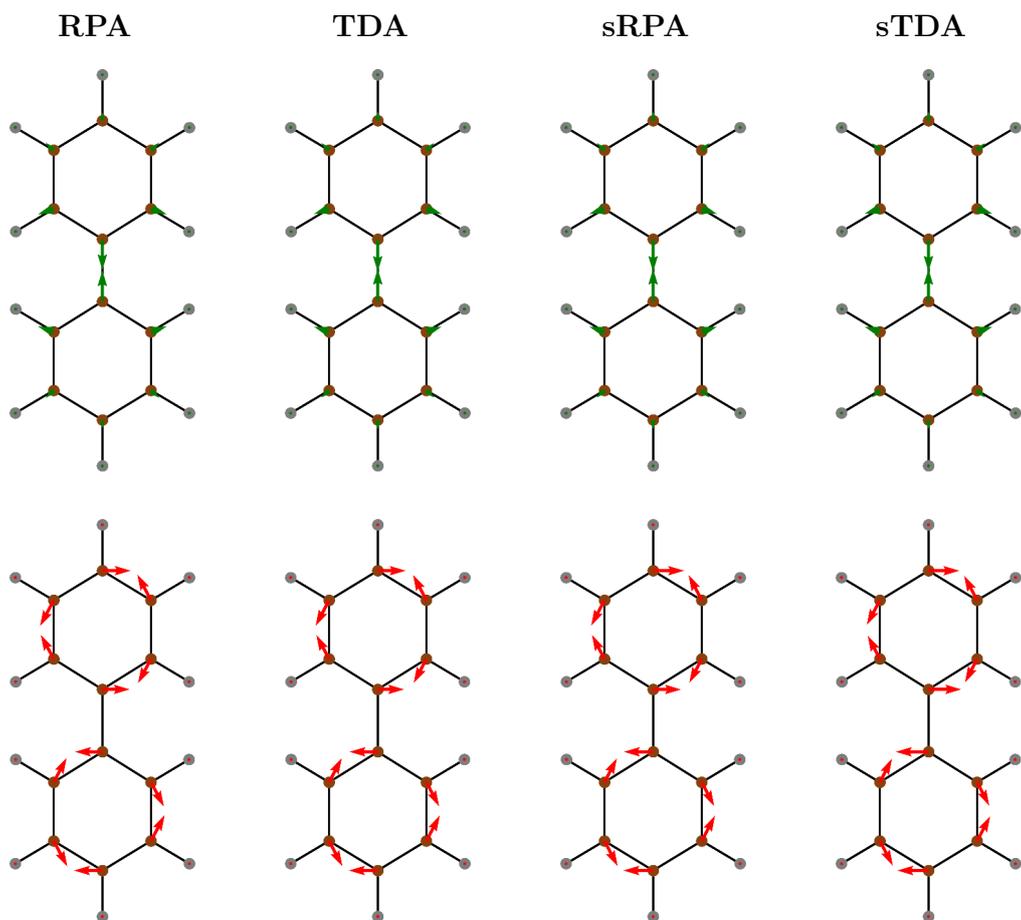


Figure S2: RPA, TDA, sRPA, and sTDA excited-state gradients of the first excited state (top, green) and non-adiabatic coupling vectors between the ground and the first excited state (bottom, red) of **I** at the PBE0/def2-SVP level of theory.

3 Protonated Formaldimine (II)

Table 3: Comparison of RPA, TDA, sRPA, and sTDA excitation energies of **II** at the PBE0/def2-SVP level of theory.

	RPA	sRPA	TDA	sTDA
ω_1 [eV]	7.86	8.07	7.87	8.07
ω_2 [eV]	9.76	9.87	10.487	10.40
ω_3 [eV]	10.87	11.06	10.92	11.06

Table 4: Influence of the thresholds ϑ_{pCSF} , ϑ_{CSF} , and ϑ_{sCSF} on the second sTDA/sRPA excitation energy of **II** calculated at the PBE0/def2-SVP level of theory. The bold numbers have been applied to calculations of **II** throughout this work. Only primary configuration state functions (CSFs) with $< \vartheta_{\text{pCSF}}$ and secondary CSFs with $< \vartheta_{\text{CSF}}$ and a significant coupling to the primary CSFs ($> \vartheta_{\text{sCSF}}$) are considered.

ϑ_{pCSF} [eV]	12	12	12	12	12	12
ϑ_{CSF} [eV]	12	20	30	100	1000	10000
ϑ_{sCSF}	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
sTDA [eV]	11.06	11.06	10.81	10.43	10.40	10.40
sRPA [eV]	10.66	10.61	10.32	9.90	9.87	9.87
ϑ_{pCSF} [eV]	12	12	12	20	20	20
ϑ_{CSF} [eV]	20	100	1000	50	100	1000
ϑ_{sCSF}	1.e-5	1.e-5	1.e-5	1.e-4	1.e-4	1.e-4
sTDA [eV]	11.06	10.41	10.39	10.52	10.42	10.39
sRPA [eV]	10.61	9.89	9.85	10.01	9.89	9.86

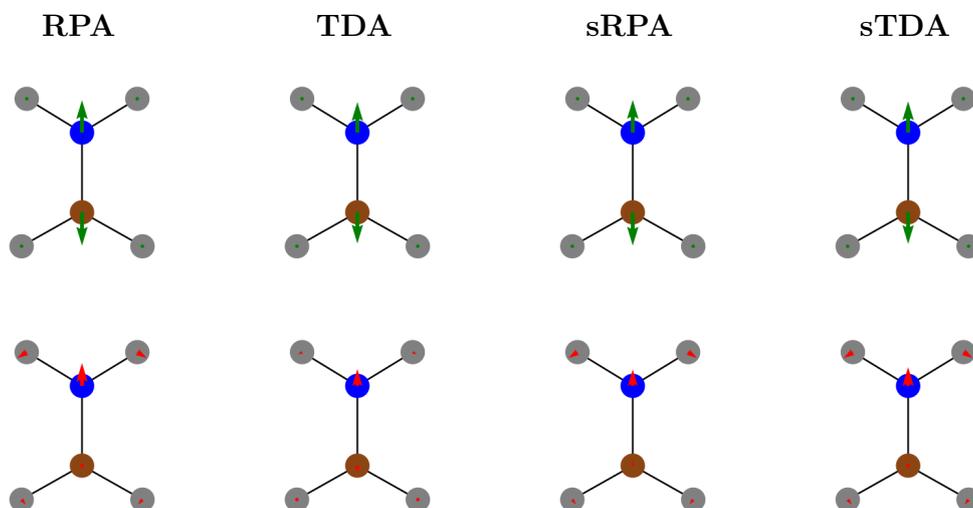


Figure S3: RPA, TDA, sRPA, and sTDA excited-state gradients of the second excited state (top, green) and non-adiabatic coupling vectors between the ground and the second excited state (bottom, red) of **II** at the PBE0/def2-SVP level of theory.

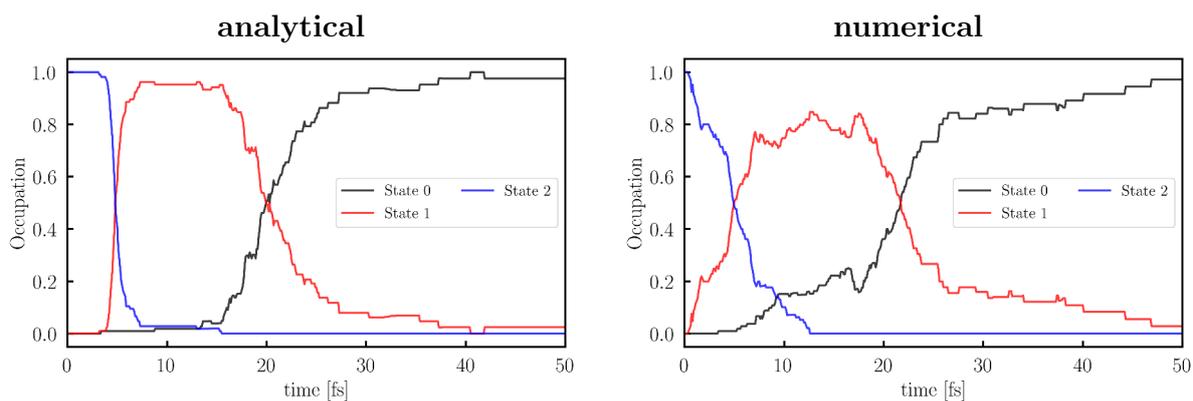


Figure S4: Change of the state occupations of **II** calculated as an average of all NAMD simulations at RPA (PBE0/def2-SVP) level of theory using (left) analytical NACVs and (right) the HST model.

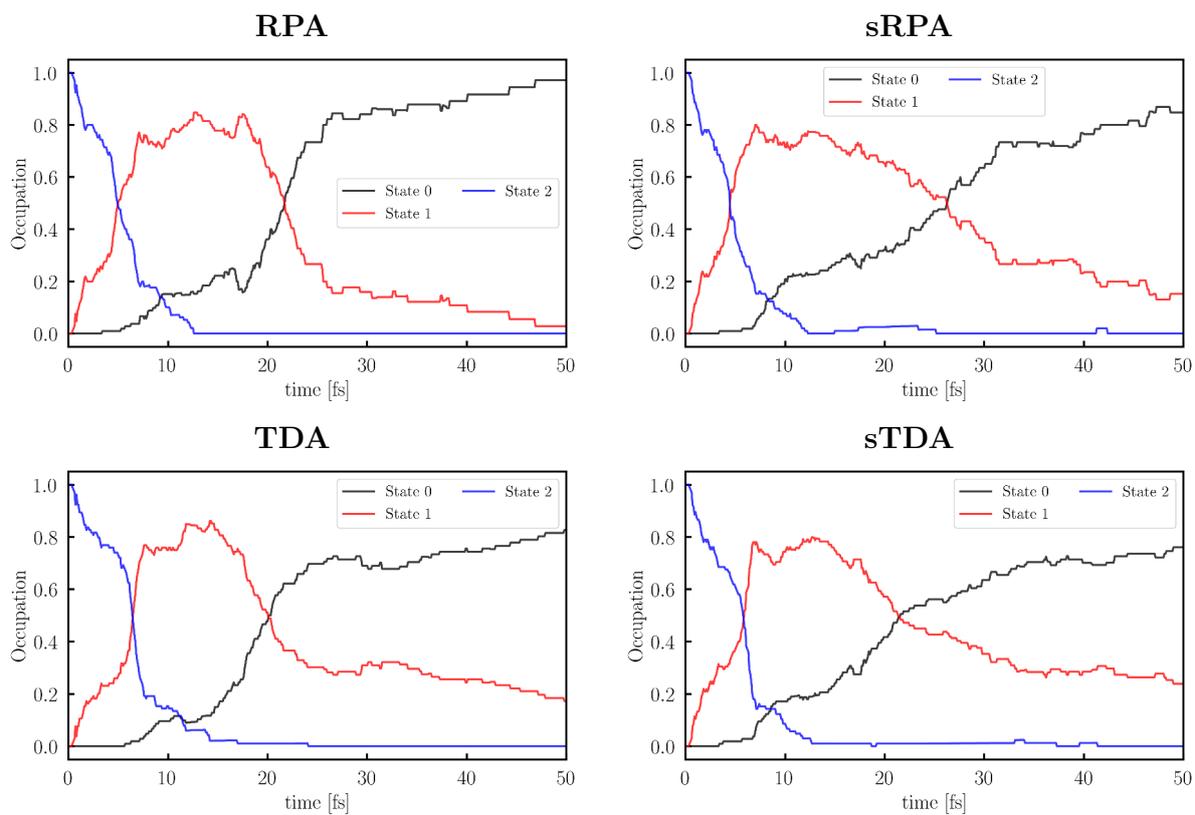


Figure S5: Change of the state occupations of \mathbf{II} calculated as an average of all NAMM simulations at RPA, TDA, sRPA, and sTDA (PBE0/def2-SVP) level of theory using the HST model.

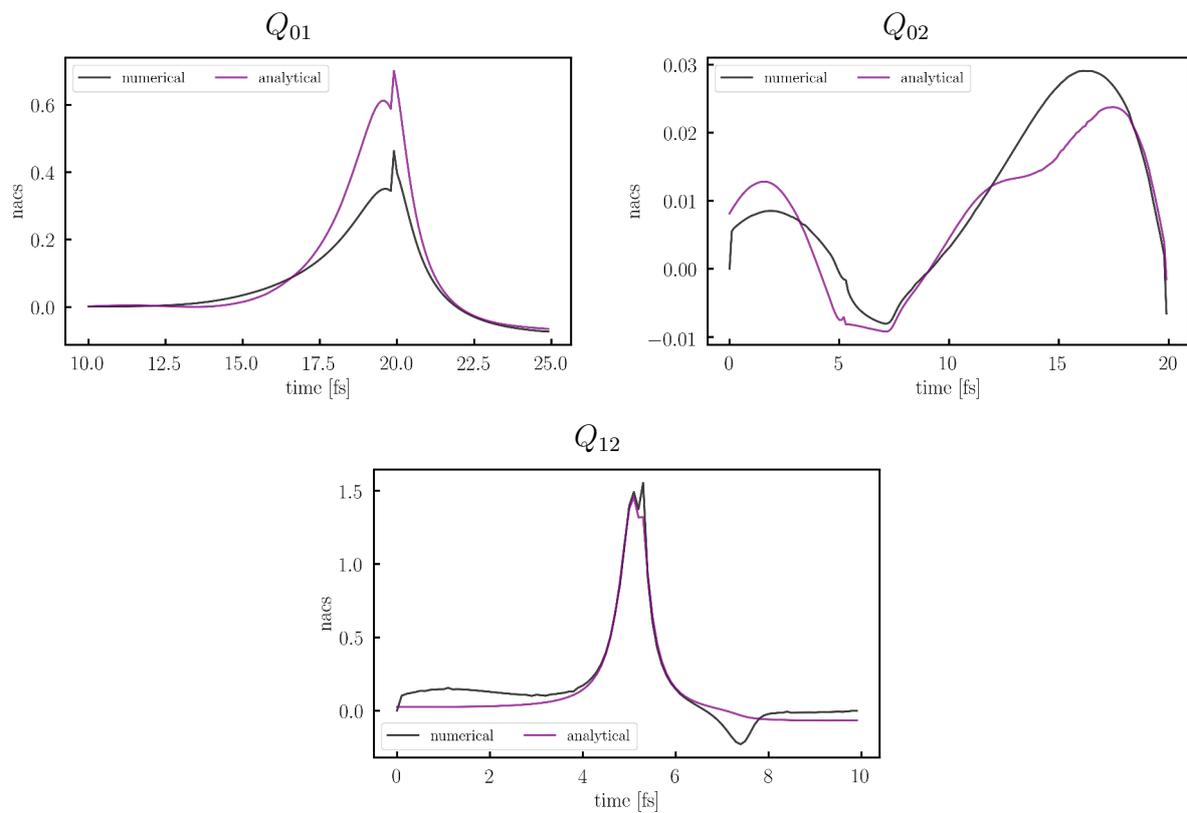


Figure S6: Couplings (Q_{01} , Q_{02} , Q_{12}) during a selected trajectory of \mathbf{II} calculated RPA (PBE0/def2-SVP) level of theory using analytical non-adiabatic coupling vectors and the HST model.

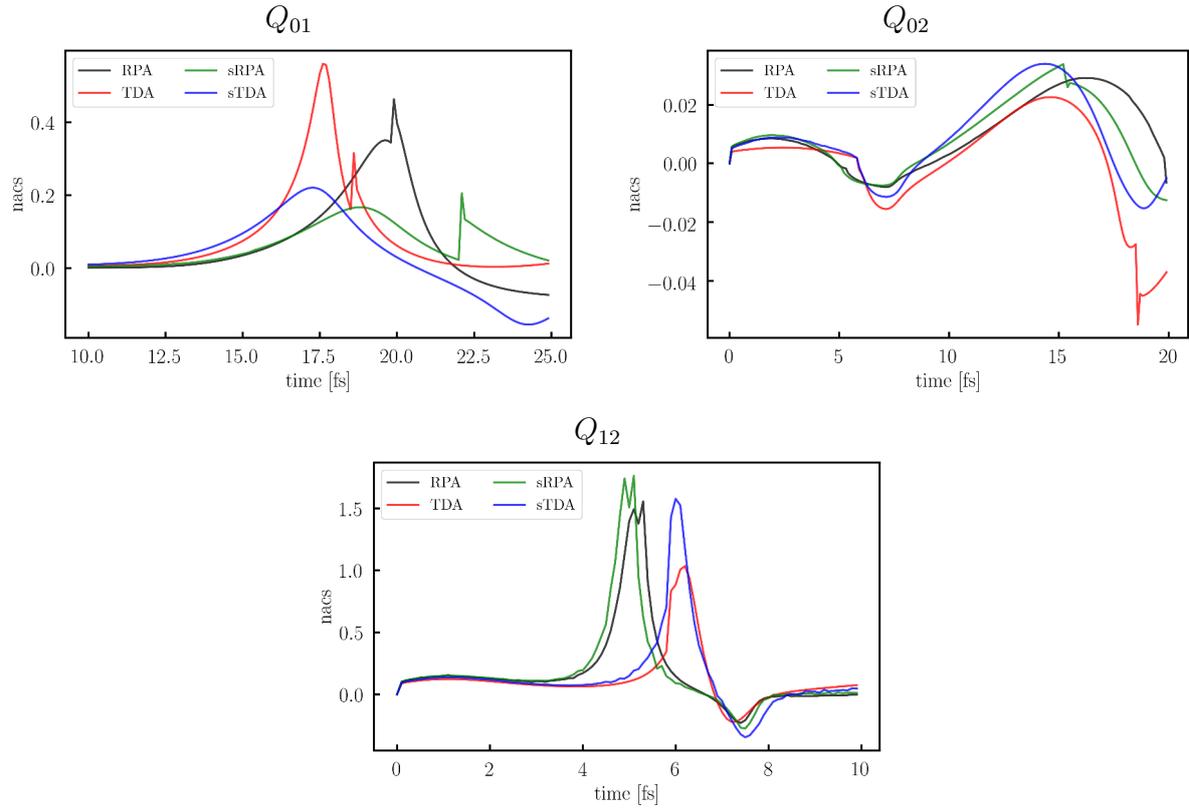


Figure S7: Couplings (Q_{01} , Q_{02} , Q_{12}) during a selected trajectory of \mathbf{II} calculated RPA, TDA, sRPA, and sTDA (PBE0/def2-SVP) level of theory using the HST model.

4 Schiff Base of Retinal (III)

4.1 PBE0/def2-SVP

Table 5: Comparison of RPA, TDA, sRPA, and sTDA excitation energies of **III** at the PBE0/def2-SVP level of theory.

	RPA	sRPA	TDA	sTDA
ω_1 [eV]	2.38	2.33	2.69	2.56
ω_2 [eV]	3.28	3.21	3.48	3.38

Table 6: Influence of the thresholds ϑ_{pCSF} , ϑ_{CSF} , and ϑ_{sCSF} on the first sTDA/sRPA excitation energy of **III** calculated at the PBE0/def2-SVP level of theory. The bold numbers have been applied to calculations of **III** throughout this work. Only primary configuration state functions (CSFs) with $< \vartheta_{\text{pCSF}}$ and secondary CSFs with $< \vartheta_{\text{CSF}}$ and a significant coupling to the primary CSFs ($> \vartheta_{\text{sCSF}}$) are considered.

ϑ_{pCSF} [eV]	5	5	5	5	5	5	5	5	5	10	10
ϑ_{CSF} [eV]	5	10	20	50	100	5	10	20	50	20	50
ϑ_{sCSF}	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4	1.e-5	1.e-5	1.e-5	1.e-5	1.e-4	1.e-4
sTDA [eV]	2.58	2.56	2.56	2.55	2.55	2.58	2.55	2.50	2.49	2.52	2.51
sRPA [eV]	2.35	2.33	2.33	2.33	2.33	2.35	2.32	2.27	2.25	2.29	2.28

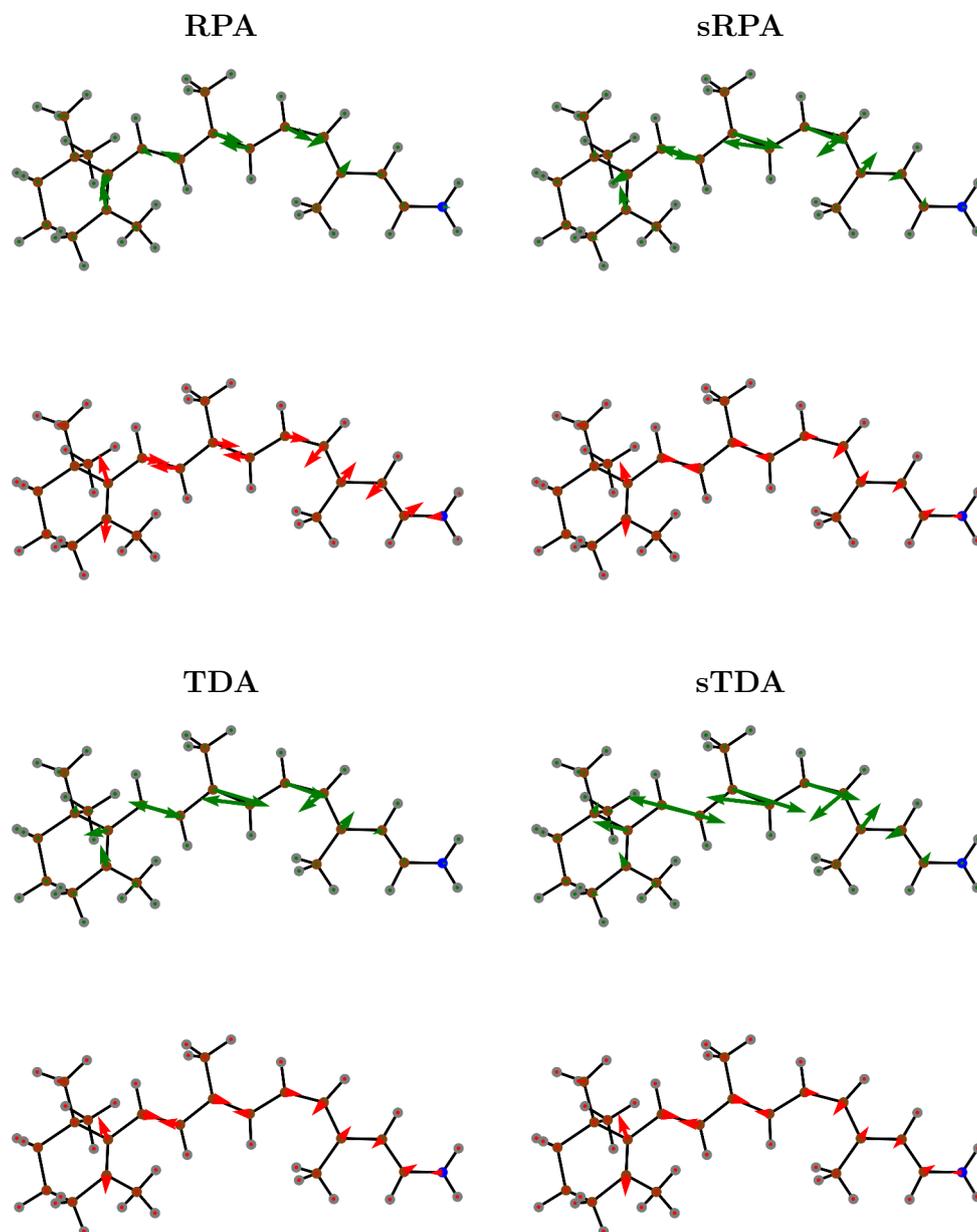


Figure S8: RPA, TDA, sRPA, and sTDA excited-state gradients of the first excited state (top, green) and non-adiabatic coupling vectors between the ground and the first excited state (bottom, red) of **III** at the PBE0/def2-SVP level of theory.

4.2 ω B97/def2-SVP

Table 7: Comparison of RPA, TDA, sRPA, and sTDA excitation energies of **III** at the ω B97/def2-SVP level of theory.

	RPA	sRPA	TDA	sTDA
ω_1 [eV]	2.74	3.08	2.90	3.17
ω_2 [eV]	4.11	4.46	4.29	4.57

Table 8: Influence of the thresholds ϑ_{pCSF} , ϑ_{CSF} , and ϑ_{sCSF} on the first sTDA/sRPA excitation energy of **III** calculated at ω B97/def2-SVP level of theory. The bold numbers have been applied to calculations of **III** throughout this work. Only primary configuration state functions (CSFs) with $< \vartheta_{\text{pCSF}}$ and secondary CSFs with $< \vartheta_{\text{CSF}}$ and a significant coupling to the primary CSFs ($> \vartheta_{\text{sCSF}}$) are considered.

ϑ_{pCSF} [eV]	5	5	5	5	5	5	5	5	5	10	10
ϑ_{CSF} [eV]	5	10	20	50	100	5	10	20	100	20	50
ϑ_{sCSF}	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4	1.e-5	1.e-5	1.e-5	1.e-5	1.e-4	1.e-4
sTDA [eV]	3.51	3.22	3.17	3.17	3.17	3.51	3.20	3.12	3.11	3.11	3.11
sRPA [eV]	3.42	3.13	3.08	3.08	3.08	3.42	3.12	3.02	3.00	3.02	3.00

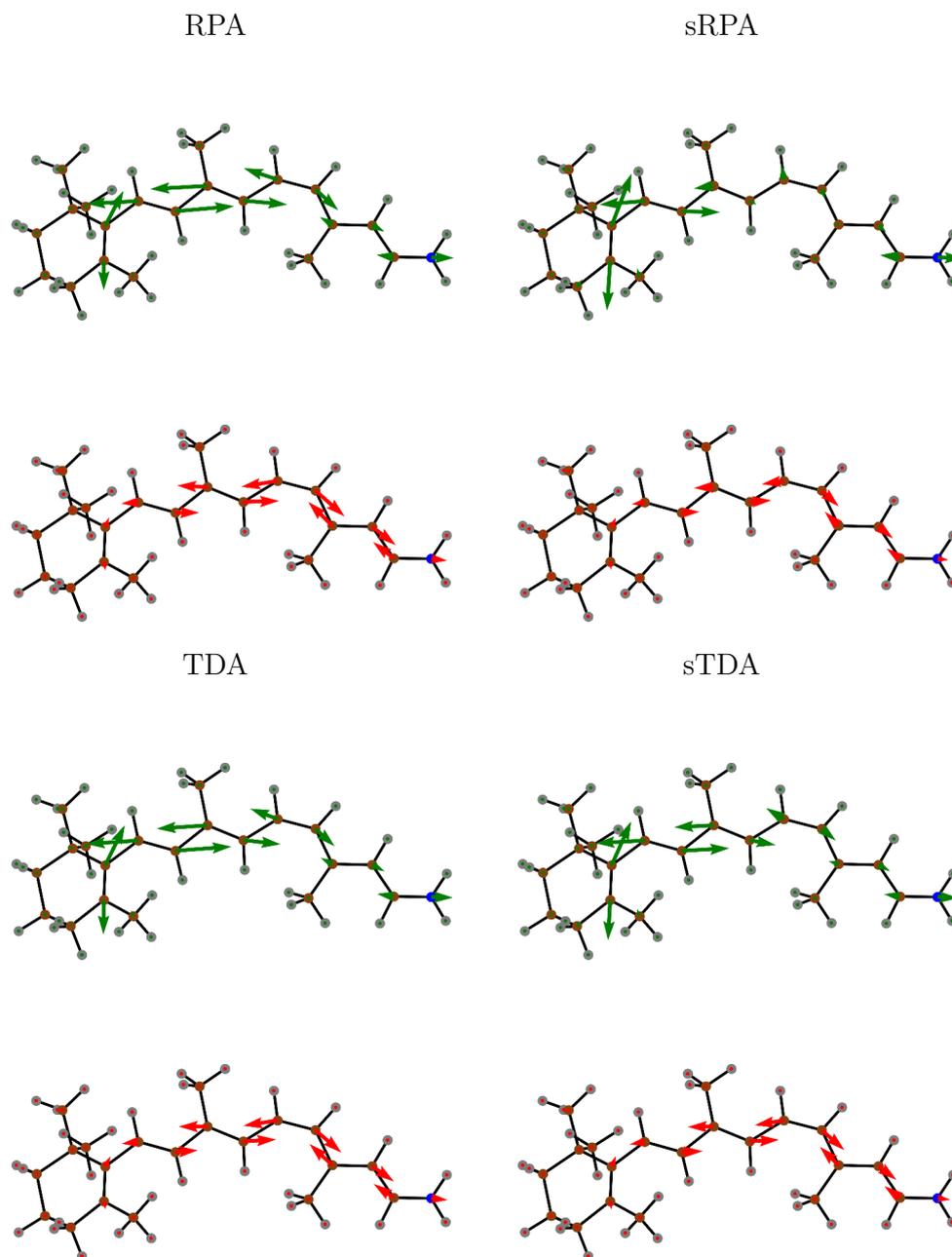


Figure S9: **RPA**, **TDA**, **sRPA**, and **sTDA** excited-state gradients of the first excited state (top, green) and non-adiabatic coupling vectors between the ground and the first excited state (bottom, red) of **III** at the ω B97/def2-SVP level of theory.

5 Validation of the Finite Differences Calculation of Couplings

In order to validate the finite differences calculation of couplings, we compare analytical and numerical non-adiabatic coupling vectors ($\boldsymbol{\tau}$). The latter are calculated as follows:

$$\boldsymbol{\tau}_{0J}(\mathbf{R}) \approx \sum_{ia} (X_{ia}^J(\mathbf{R}) - Y_{ia}^J(\mathbf{R})) \frac{O_{ia}(\mathbf{R}, \mathbf{R} + \Delta\mathbf{R}) - O_{ia}(\mathbf{R}, \mathbf{R} - \Delta\mathbf{R})}{2\Delta\mathbf{R}} \quad (1)$$

$$\begin{aligned} \boldsymbol{\tau}_{IJ}(\mathbf{R}) \approx & \sum_{ia} X_{ia}^I(\mathbf{R}) \frac{X_{ia}^J(\mathbf{R} + \Delta\mathbf{R}) - X_{ia}^J(\mathbf{R} - \Delta\mathbf{R})}{2\Delta\mathbf{R}} \\ & - \sum_{ia} Y_{ia}^I(\mathbf{R}) \frac{Y_{ia}^J(\mathbf{R} + \Delta\mathbf{R}) - Y_{ia}^J(\mathbf{R} - \Delta\mathbf{R})}{2\Delta\mathbf{R}} \\ & + \sum_{iab} (X_{ia}^I(\mathbf{R})X_{ib}^J(\mathbf{R}) - Y_{ia}^I(\mathbf{R})Y_{ib}^J(\mathbf{R})) \frac{O_{ab}(\mathbf{R}, \mathbf{R} + \Delta\mathbf{R}) - O_{ab}(\mathbf{R}, \mathbf{R} - \Delta\mathbf{R})}{2\Delta\mathbf{R}} \\ & - \sum_{ija} (X_{ia}^I(\mathbf{R})X_{ja}^J(\mathbf{R}) - Y_{ia}^I(\mathbf{R})Y_{ja}^J(\mathbf{R})) \frac{O_{ij}(\mathbf{R}, \mathbf{R} + \Delta\mathbf{R}) - O_{ij}(\mathbf{R}, \mathbf{R} - \Delta\mathbf{R})}{2\Delta\mathbf{R}} \end{aligned} \quad (2)$$

$$O_{pq}(\mathbf{R}_1, \mathbf{R}_2) = \langle \phi(\mathbf{R}_1) | \phi(\mathbf{R}_2) \rangle \quad (3)$$

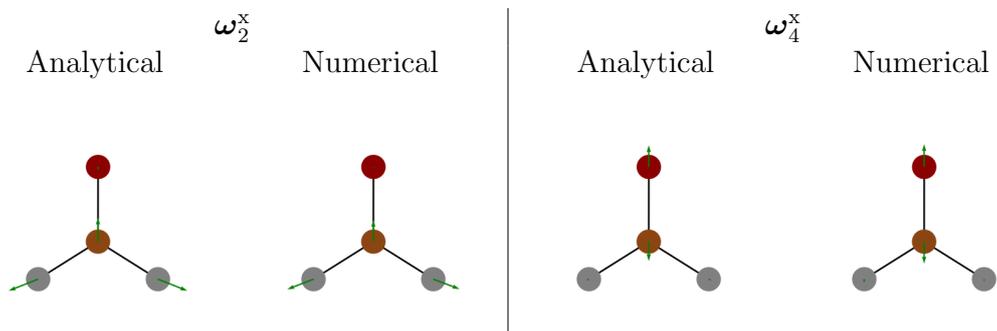


Figure S10: Numerical and analytical excited-state gradients (ω_2^x and ω_4^x) of formaldehyde calculated at the PBE0/def2-SVP level of theory.

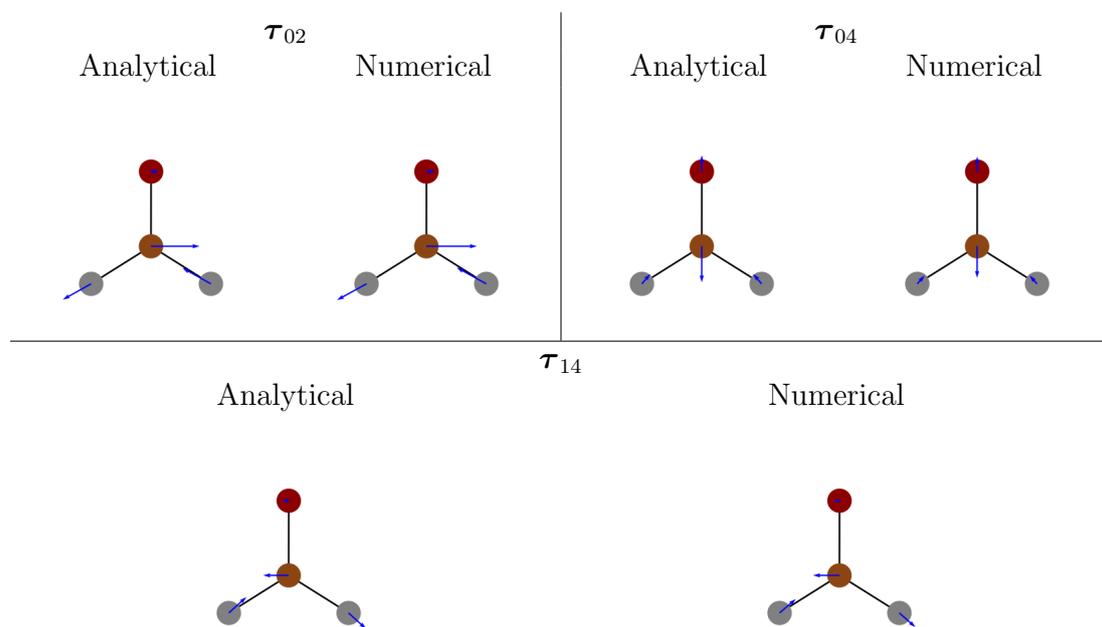


Figure S11: Numerical and analytical non-adiabatic coupling vectors (τ_{02} , τ_{04} , and τ_{14}) of formaldehyde calculated at the PBE0/def2-SVP level of theory.

References

- (1) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 134114.
- (2) Kussmann, J.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
- (3) Kussmann, J.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
- (4) Marques, M. A. L.; Oliveira, M. J. T.; Burnus, T. *Comput. Phys. Commun.* **2012**, *183*, 2272.
- (5) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. *SoftwareX* **2018**, *7*, 1–5.
- (6) Laqua, H.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2018**, *149*, 204111.
- (7) Perdew, J. P.; Burke, K.; Erzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (8) Perdew, J. P.; Burke, K.; Erzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396(E).
- (9) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.
- (10) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
- (11) Chai, J. D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 84106.
- (12) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (13) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.
- (14) Grimme, S. *J. Chem. Phys.* **2013**, *138*, 244104.
- (15) Bannwarth, C.; Grimme, S. *Comput. Theor. Chem.* **2014**, *1040-1041*, 45–53.
- (16) Risthaus, T.; Hansen, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14408–14419.

- (17) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.; Motoshima, H.; Fox, B.; Le Trong, I.; Teller, D.; Okada, T.; Stenkamp, R.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739–745.
- (18) Case, D. A.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017. 2017.
- (19) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (20) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
- (21) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637–649.
- (22) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 14101.
- (23) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061–1071.
- (24) Hammes-Schiffer, S.; Tully, J. C. *J. Chem. Phys.* **1994**, *101*, 4657–4667.

3.4 Publication IV: Calculating free energies from the vibrational density of states function: validation and critical assessment

L. D. M. Peters, J. C. B. Dietschreit, J. Kussmann, C. Ochsenfeld,
"Calculating free energies from the vibrational density of states function:
validation and critical assessment",
J. Chem. Phys., **150**, 194111 (2019)

Abstract: We explore and show the usefulness of the density of states function for computing vibrational free energies and free energy differences between small systems. Therefore, we compare this density of states integration method (DSI) to more established schemes such as Bennett's Acceptance Ratio method (BAR), the Normal Mode Analysis (NMA), and the Quasiharmonic Analysis (QHA). The strengths and shortcomings of all methods are highlighted with three numerical examples. Furthermore, the free energy of the ionization of ammonia and the mutation from serine to cysteine are computed using extensive *ab initio* molecular dynamics simulations. We conclude that DSI improves upon the other frequency-based methods (NMA and QHA) regarding the treatment of anharmonicity and yielding results comparable to BAR in all cases without the need for alchemical transformations. Low-frequency modes lead to larger errors indicating that long simulation times might be required for larger systems. In addition, we introduce the use of DSI for the localization of the vibrational free energy to specific atoms or residues, leading to insights into the underlying process, a unique feature that is only offered by this method.

Reprinted from:

L. D. M. Peters, J. C. B. Dietschreit, J. Kussmann, C. Ochsenfeld,
"Calculating free energies from the vibrational density of states function:
validation and critical assessment",
J. Chem. Phys., **150**, 194111 (2019)

with the permission of AIP Publishing.

<https://aip.scitation.org/doi/abs/10.1063/1.5079643>

Calculating free energies from the vibrational density of states function: Validation and critical assessment

Cite as: J. Chem. Phys. 150, 194111 (2019); doi: 10.1063/1.5079643

Submitted: 1 November 2018 • Accepted: 21 April 2019 •

Published Online: 20 May 2019



View Online



Export Citation



CrossMark

Laurens D. M. Peters,^{1,2,a)}  Johannes C. B. Dietschreit,^{1,2,a)}  Jörg Kussmann,^{1,2} 
and Christian Ochsenfeld^{1,2,b)} 

AFFILIATIONS

¹Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany

²Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstr. 5-13, D-81377 München, Germany

^{a)}Contributions: L. D. M. Peters and J. C. B. Dietschreit contributed equally to this work.

^{b)}Electronic mail: christian.ochsenfeld@uni-muenchen.de

ABSTRACT

We explore and show the usefulness of the density of states function for computing vibrational free energies and free energy differences between small systems. Therefore, we compare this density of states integration method (DSI) to more established schemes such as Bennett's Acceptance Ratio method (BAR), the Normal Mode Analysis (NMA), and the Quasiharmonic Analysis (QHA). The strengths and shortcomings of all methods are highlighted with three numerical examples. Furthermore, the free energy of the ionization of ammonia and the mutation from serine to cysteine are computed using extensive *ab initio* molecular dynamics simulations. We conclude that DSI improves upon the other frequency-based methods (NMA and QHA) regarding the treatment of anharmonicity and yielding results comparable to BAR in all cases without the need for alchemical transformations. Low-frequency modes lead to larger errors indicating that long simulation times might be required for larger systems. In addition, we introduce the use of DSI for the localization of the vibrational free energy to specific atoms or residues, leading to insights into the underlying process, a unique feature that is only offered by this method.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5079643>

I. INTRODUCTION

Free energy differences are closely connected to experimental thermodynamic data (e.g., binding affinities, reaction energies, and activation barriers of molecular transformations) as they incorporate contributions from the internal energy as well as from entropy.¹⁻⁷ Methods to calculate free energies or their differences are, therefore, of great interest in computational chemistry. Roughly, they can be divided into two groups: (1) frequency- and (2) energy-based methods.

In the first group, the free energy is calculated from frequencies of molecular vibrations (or rotations).⁸ These frequencies can be obtained from the second derivative of the energy with respect to the nuclear coordinates at the minimum energy geometry or from the covariance matrix taken from a molecular dynamics⁹⁻¹¹ (MD)

or Monte Carlo^{12,13} (MC) simulation referring to the Normal Mode Analysis^{14,15} (NMA) and the Quasiharmonic Analysis^{16,17} (QHA), respectively. Energy-based methods calculate free energy differences from sampled energies along MC or MD simulations, applying exponential averaging theory¹⁸ (EXP), thermodynamic integration¹⁹ (TI), or Bennett's acceptance ratio method²⁰ (BAR).

All mentioned methods have, despite their great success and broad fields of application,²¹⁻²⁸ well-known shortcomings. The use of NMA, for example, requires the search for the minimum energy geometry (or geometries). It is, therefore, usually applied to small- or medium-sized molecules, using quantum-mechanical (QM) methods. Free energy methods, using data from simulations (QHA, EXP, TI, BAR), usually require a large number of steps to converge, which is connected to the universal problem of sampling the phase space sufficiently to estimate the ratio of partition functions.

This challenge has been tackled in many publications,^{1–3,29} applying, e.g., alchemical transformations or enhanced sampling techniques. As this may still require long simulation times, the levels of theory for these calculations range from molecular-mechanical (MM) over semiempirical to combined quantum-mechanical/molecular-mechanical (QM/MM) methods, depending on the size of the simulated system and the problem at hand. Additional shortcomings are the harmonic approximation in NMA and QHA and the neglect of vibrational quantum effects³⁰ in EXP, TI, and BAR.

An alternative approach has been proposed by Berens *et al.*³¹ by calculating free energies and free energy differences as a weighted integral over the density of states function, which is determined from sampled nuclear velocities along MD simulations. Although already developed in 1983, this method [named integration of the density of states method (DSI) in the following] has only been used occasionally for absolute entropy calculations^{32,33} or the calculation of solvation entropies.^{34–36} Therefore, its convergence with respect to the simulation time and the number of independent trajectories has not been investigated in detail. Nevertheless, we expect the sampling problem to be as crucial in DSI as in the other simulation-based methods as the vibrational partition function can again only be approximated (in this case via the density of states function).

In this work, we compare DSI to the established methods mentioned above and validate its use for (1) free energy calculations of molecular transformations and (2) localization of the free energy to specific atoms or residues. First, we briefly recapitulate the theory of EXP, BAR, NMA, QHA, and DSI (Sec. II A) and introduce our novel ansatz to obtain atomic contributions of vibrational free energy changes using DSI (Sec. II B). Having listed the computational details in Sec. III, we compare the free energy methods in Sec. IV. For this purpose, we (1) investigate three numerical examples, (2) determine the ionization potential of ammonia, and (3) calculate the free transformation energy from serine to cysteine *in vacuo* from *ab initio* MD simulations^{37,38} using HF-3c.³⁹ Conclusions are drawn in Sec. V.

II. THEORY

The free energy (A) and the free energy difference between two systems 0 and 1 ($\Delta A_{0\rightarrow 1}$) can be calculated from the partition function (Q) as

$$A = -\beta^{-1} \ln Q, \quad (1)$$

$$\Delta A_{0\rightarrow 1} = -\beta^{-1} \ln \frac{Q_1}{Q_0}, \quad (2)$$

where β is equal to $1/(k_B T)$ with k_B being the Boltzmann constant and T being the absolute temperature. It is generally assumed that A can be separated into contributions of translation (A_{trans}), rotation (A_{rot}), and vibration (A_{vib}) as well as the energy of the electronic ground state (E)

$$\begin{aligned} A &= E + A_{\text{trans}} + A_{\text{rot}} + A_{\text{vib}} \\ &= E - \beta^{-1} \ln\{Q_{\text{trans}} \times Q_{\text{rot}} \times Q_{\text{vib}}\}, \end{aligned} \quad (3)$$

where Q_{trans} , Q_{rot} , and Q_{vib} are the corresponding partition functions. In this article, we restrict ourselves to the calculation of vibrational free energies. In the simulations, this is realized by removing

the center of mass translation and the overall rotation of the system in each step of the molecular dynamics simulation.

A. Review of free energy methods

1. Energy-based methods

The partition function of a canonical ensemble (NVT) is defined as

$$Q \propto \int d\mathbf{x} \exp\{-\beta U(\mathbf{x})\}. \quad (4)$$

$U(\mathbf{x})$ is the potential energy at a given nuclear structure \mathbf{x} , whereas the kinetic energy terms are part of the proportionality constant. Equation (2) can thus be transformed into

$$\Delta A_{0\rightarrow 1} = -\beta^{-1} \ln \frac{\int d\mathbf{x} \exp\{-\beta U_1(\mathbf{x})\}}{\int d\mathbf{x} \exp\{-\beta U_0(\mathbf{x})\}}. \quad (5)$$

$U_0(\mathbf{x})$ and $U_1(\mathbf{x})$ are the potential energy functions of systems 0 and 1, respectively. In exponential averaging theory (EXP), the difference between the potential energies of the two systems $\Delta U = U_1 - U_0$ is calculated so that¹⁸

$$\Delta A_{0\rightarrow 1} = -\beta^{-1} \ln \langle \exp\{-\beta \Delta U(\mathbf{x})\} \rangle_0, \quad (6)$$

where $\langle B(\mathbf{x}) \rangle_0$ denotes the ensemble average of $B(\mathbf{x})$ over configurations sampled from the reference system 0. In many cases, the underlying distribution of ΔU is too wide for an efficient calculation of ΔA so that a coupling parameter $\lambda \in [0; 1]$ is introduced, which gradually transforms system 0 into system 1 and thus creates a better overlap of the distributions^{2,18,40}

$$U_\lambda(\mathbf{x}) = (1 - \lambda)U_0(\mathbf{x}) + \lambda U_1(\mathbf{x}). \quad (7)$$

This transformation, which can be, for example, a chemical reaction or an artificial (so-called alchemical) transformation, is then separated into M sufficiently small steps of size $\Delta\lambda_i$, and the free energy difference of each step is calculated individually leading to

$$\Delta A_{0\rightarrow 1} = -\beta^{-1} \sum_{i=0}^{M-1} \ln \langle \exp\{-\beta \Delta U_i(\mathbf{x})\} \rangle_{\lambda_i}, \quad (8)$$

with

$$\Delta U_i(\mathbf{x}) = U_{\lambda_{i+1}}(\mathbf{x}) - U_{\lambda_i}(\mathbf{x}) = (\lambda_{i+1} - \lambda_i) \Delta U(\mathbf{x}) = \Delta\lambda_i \Delta U(\mathbf{x}) \quad (9)$$

and

$$\sum_{i=0}^{M-1} \Delta\lambda_i = 1. \quad (10)$$

In the additive scheme of Eq. (8), forward ($\Delta A_{0\rightarrow 1}$) and backward ($-\Delta A_{1\rightarrow 0}$) calculations of the free energy differ in almost all cases, again due to the different distributions.⁴¹ This error can be reduced by increasing the sampling of the system or by applying the double-wide sampling scheme.⁴²

A more sophisticated approach to obtain the “best” free energy from forward and backward calculations has been derived by Bennett in 1976 [see Eq. (11)].²⁰ It minimizes the variance of ΔA and is equivalent to its maximum likelihood estimator, as shown by Shirts and Pande in 2003,⁴³

$$0 = \ln \left[\frac{\sum_F^{N_{0\rightarrow 1}} f(M + \beta \Delta U_{0\rightarrow 1}^F - \beta \Delta A_{0\rightarrow 1})}{\sum_B^{N_{1\rightarrow 0}} f(-M + \beta \Delta U_{1\rightarrow 0}^B + \beta \Delta A_{0\rightarrow 1})} \right], \quad (11)$$

$$M = \ln \frac{N_{0 \rightarrow 1}}{N_{1 \rightarrow 0}}. \quad (12)$$

f is the Fermi function $f(x) = \frac{1}{1 + \exp(x)}$; $\Delta U_{0 \rightarrow 1}^F$ and $\Delta U_{1 \rightarrow 0}^B$ are independent forward and backward perturbations, respectively; and $N_{0 \rightarrow 1}$ and $N_{1 \rightarrow 0}$ are the corresponding numbers of frames. The resulting Bennett's Acceptance Ratio method (BAR) is known to be more robust than EXP or thermodynamic integration (TI) schemes.⁴⁴⁻⁴⁹

2. Frequency-based methods

Frequency-based methods assume that the potential energy function can be approximated by a sum of $N_F - 6$ harmonic oscillators (harmonic approximation)

$$U(\mathbf{x}) = \sum_{ij}^{N_F-6} k_{ij}(x_i - x_i^0)(x_j - x_j^0), \quad (13)$$

where N_F is the number of degrees of freedom of the system and k_{ij} are the force constants. In the Normal Mode Analysis (NMA),^{14,15} the Hessian matrix (the second derivative of the energy with respect to the nuclear coordinates) at the minimum energy configuration (\mathbf{x}^0),

$$H_{ij} = \frac{\partial^2 E}{\partial x_i \partial x_j}, \quad (14)$$

is diagonalized, yielding the normal modes ν_i , which are then used to calculate the vibrational free energy either classically (CL) or quantum-mechanically (QM)

$$A_{\text{vib}}^{\text{CL}} = \beta^{-1} \sum_i \ln[\beta h \nu_i], \quad (15)$$

$$A_{\text{vib}}^{\text{QM}} = \beta^{-1} \sum_i \ln \left[\frac{1 - \exp(-\beta h \nu_i)}{\exp(-\frac{1}{2} \beta h \nu_i)} \right], \quad (16)$$

where h is the Planck constant. While NMA works well for small systems, great care has to be taken in the case of large systems. Here, NMAs have to be performed at all (relevant) local minima and the results have to be weighted by the Boltzmann factor of the respective minimum. This task becomes harder with increasing system size.

The search for minima is not required in the Quasiharmonic Analysis (QHA)^{16,17} as it is performed after a molecular dynamics or Monte Carlo simulation. Assuming ergodicity and that all x_i are Boltzmann distributed, ν_i can be obtained by diagonalizing the mass weighted covariance matrix

$$\left[\mathbf{M}^{\frac{1}{2}} \boldsymbol{\sigma} \mathbf{M}^{\frac{1}{2}} - \beta^{-1} \mathbf{v} \right] \mathbf{M}^{\frac{1}{2}} \Delta \mathbf{x} = 0, \quad (17)$$

where \mathbf{M} is the kinetic energy matrix and $\boldsymbol{\sigma}$ is the covariance matrix,

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle. \quad (18)$$

$\langle \rangle$ symbolizes the average over all trajectory frames. The frequencies obtained this way can then again be transformed into vibrational free energies with Eqs. (15) and (16).

Frequencies obtained with QHA are always equal or lower to those of the NMA since it approximates all possible minima along one coordinate as well as its anharmonicity with one single harmonic oscillator. This leads to a lower curvature of the potential energy surface than actually present.

3. Integration of the density of states method

The vibrational partition function (Q_{vib}) can also be calculated as a product of the partition functions $q(\nu)$ of the single vibrational modes with frequency ν . It is assumed that these partition functions can be written as classic (q^{CL}) or quantum (q^{QM}) harmonic oscillators

$$q^{\text{CL}}(\nu) = \frac{1}{\beta h \nu}, \quad (19)$$

$$q^{\text{QM}}(\nu) = \frac{\exp(-\frac{1}{2} \beta h \nu)}{1 - \exp(-\beta h \nu)}. \quad (20)$$

The logarithm of the vibrational partition function can thus be calculated from the following integral:²¹

$$\ln Q_{\text{vib}} = \int_0^\infty d\nu D(\nu) \ln\{q(\nu)\}. \quad (21)$$

D is the density of states function, which singles out the specific frequencies of the investigated system, while an infinite number of harmonic oscillators (classical or quantum) is considered. D itself is determined as the mass-weighted Fourier transform of the nuclear velocity autocorrelation

$$D(\nu) = 2\beta \sum_{j=1}^{N_A} m_j \int dt \langle \mathbf{v}_j(\tau) \mathbf{v}_j(\tau + t) \rangle_\tau e^{-i2\pi\nu t}. \quad (22)$$

m_j and \mathbf{v}_j denote the mass and velocity vector of the nucleus j , respectively, while N_A is the number of atoms. Integration over the entire density of states functions yields the number of degrees of freedom

$$N_F = \int_0^\infty d\nu D(\nu). \quad (23)$$

Insertion of Eq. (21) into Eq. (1) yields the following expression for the free energy:

$$A = E + \beta^{-1} \int_0^\infty d\nu D(\nu) W_A(\nu). \quad (24)$$

W_A is, depending on the inserted q , the classical (W_A^{CL}) or quantum (W_A^{QM}) weighting function

$$W_A^{\text{CL}}(\nu) = \ln[\beta h \nu], \quad (25)$$

$$W_A^{\text{QM}}(\nu) = \ln \left[\frac{1 - \exp(-\beta h \nu)}{\exp(-\frac{1}{2} \beta h \nu)} \right]. \quad (26)$$

If one assumes that D consists of delta functions at the frequencies of the normal modes (ν_i) of a system, Eq. (24) with W_A^{CL} and W_A^{QM} is equal to Eqs. (15) and (16), respectively. The difference between the integration of the density of states method (DSI) and the other frequency-based methods is, thus, that instead of $N_F - 6$ harmonic oscillators

$$Q_{\text{vib}} = \prod_i^{N_F-6} q(\nu_i), \quad (27)$$

an arbitrary large number of harmonic oscillators (depending on the simulation time) weighted by the density of states function

$$Q_{\text{vib}} = \prod_i q(\nu_i)^{D(\nu_i)} \quad (28)$$

are assumed to describe the system (harmonic approximation).

From Eq. (24), one can easily obtain the expression for free energy differences

$$\Delta A_{0 \rightarrow 1} = \Delta E + \beta^{-1} \int_0^\infty dv \Delta D(v) W_A(v), \quad (29)$$

where $\Delta D = D_1 - D_0$ and $\Delta E = E_1 - E_0$.

B. Atomic contributions to the density of states functions

Our ansatz to determine atomic contributions to the vibrational free energy uses the fact that the density of states function from Eq. (22) can be rewritten as a superposition of atomic functions

$$\begin{aligned} D(v) &= \sum_{j=1}^{N_A} 2\beta m_j \int dt \langle \mathbf{v}_j(\tau) \mathbf{v}_j(\tau + t) \rangle_\tau e^{-i2\pi vt} \\ &= \sum_{j=1}^{N_A} D_j(v). \end{aligned} \quad (30)$$

This essentially means that the vibrational partition function can be written, without any further loss of generality, as

$$\begin{aligned} Q_{\text{vib}} &= \prod_j \prod_i q(v_i)^{D_j(v_i)} \\ &= \prod_j Q_j^{\text{vib}} \end{aligned} \quad (31)$$

because

$$\prod_j q(v_i)^{D_j(v_i)} = q(v_i)^{D(v_i)}. \quad (32)$$

Thus, the vibrational free energy can be written as a sum over atomic contributions

$$\begin{aligned} A_{\text{vib}} &= -\beta^{-1} \ln Q_{\text{vib}} \\ &= -\beta^{-1} \ln \left[\prod_j^{N_A} Q_j^{\text{vib}} \right] \\ &= -\beta^{-1} \sum_j^{N_A} \ln Q_j^{\text{vib}} \\ &= \sum_j^{N_A} A_j^{\text{vib}}. \end{aligned} \quad (33)$$

The above partitioning of the vibrational free energy is not restricted to atoms. It can without any further assumption be grouped into any meaningful collection of atoms such as residues or functional groups. This ansatz can be used as an aide to interpret and localize the changes occurring in the system, as shown in our previous work.⁵⁰ Please note that it uses the vibrational free energy only and not the total free energy. For the partitioning of the latter, approximate schemes exist²¹ but have been discussed to lead to unreliable results.^{51,52} Our only assumption along with the harmonic approximation is that the regions with the most prominent changes in the vibrational free energy are also those which contribute the most to the change in E and thus the total free energy change.

III. COMPUTATIONAL DETAILS

A. Classical and molecular dynamics simulations

The free energies of the numerical examples have been obtained from classical NVT simulations of a particle (of mass 1 u) in a one-dimensional harmonic (V_H), Morse (V_A), and double-well (V_D) potentials, respectively,

$$V_H(x) = \frac{1}{2} kx^2, \quad (34)$$

$$V_A(x) = D_E \times (1 - \exp\{-ax\})^2, \quad (35)$$

$$V_D(x) = \frac{1}{2} bx^2 (x - 0.5)^2. \quad (36)$$

The exact values for k , D_E , a , and b are listed in the [supplementary material](#). To obtain an exact reference, we have integrated Eq. (5) numerically on a grid using $\approx 10^6$ points and a step width of 0.01 Bohr. For the harmonic oscillator, this procedure leads to an error below 10^{-4} kJ/mol.

The free energies of the molecular systems have been obtained from Born-Oppenheimer molecular dynamics simulations at the HF-3c³⁹ level of theory using the FermiONS++ program package⁵³⁻⁵⁵ with DFTD3 v3.1^{56,57} and gCP v2.02.⁵⁸ The center of mass translation and the overall rotation of the system have been removed at every step of the simulation.

All simulations use the Velocity Verlet^{59,60} propagator and the random rescaling thermostat by Bussi, Donadio, and Parrinello,⁶¹ keeping the average temperature at 298.15 K. A different thermostat, like a Langevin-thermostat, would in general have been better suited to sample our small systems.⁴⁷ However, the random changes of the nuclear forces would severely impact the velocity autocorrelation function and render our analysis impossible. Initial velocities have been drawn from a Maxwell-Boltzmann distribution at 298.15 K. Energies, velocities, and coordinates were written to files every 1 fs. The numerical examples are simulated for 110 ps (10 ps equilibration time, 0.1 fs time step) or, in some cases, 1010 ps (10 ps equilibration time, 0.1 fs time step). The simulation times of the molecular systems are 310 ps (10 ps equilibration time, 0.1 fs time step) in the case of ammonia and 202 ps (2 ps equilibration time, 0.2 fs time step) in the case of serine and cysteine. For every λ window, we have calculated five independent trajectories and an equidistant $\Delta\lambda$ of 0.1 has been applied. To show the convergence behavior of DSI in the ammonia example, we have additionally conducted 10 independent trajectories of 910 ps (10 ps equilibration time, 0.1 fs time step) for NH_3 and NH_3^+ .

B. *Ab initio* alchemical transformations

Alchemical transformations are normally used in a molecular mechanics (MM) context, where transforming one system (0) into another system (1) is equal to gradually turning on (or off) contributions to the potential energy.² Here, we want to use this concept with *ab initio* calculations, which do not allow for such a fragmentation of the energy. To circumvent this problem, we use an ansatz developed by Reddy *et al.*⁶² We perform two energy and forces calculations (for systems 0 and 1) at every step of the simulation and continue the trajectory along a weighted force \mathbf{F}_λ ,

$$\mathbf{F}_\lambda = (1 - \lambda)\mathbf{F}_0 + \lambda\mathbf{F}_1, \quad (37)$$

where \mathbf{F}_0 and \mathbf{F}_1 denote the determined forces of system 0 and system 1, respectively. Consequently, we use the weighted mass, temperature, center of mass velocity, inertia tensor, and total angular momenta in our thermostat and when removing the overall translation and rotation. In this work, we follow the single-topology ansatz.⁶³ In the case of the serine-cysteine transformation, this means that both systems (0 and 1) share the same structure, with the oxygen in system 0 being replaced by a sulfur in system 1. The use of a dual-topology ansatz (the OH-group in system 0 and the SH-group in system 1 have different structures, while the rest of the molecule is shared) is also possible.^{62,64,65} However, when it is applied to simulations without explicit solvent molecules, an MM region, or geometrical constraints, the dual-topology ansatz leads to unstable trajectories. The reason for this is that, when $\lambda \approx 0$ or $\lambda \approx 1$, the OH- or SH-group is not “seen” by the shared part of the molecule, leading to unphysical geometries (large C–O or C–S bonds) and convergence problems of the self-consistent field algorithm. Constraints or surroundings will prevent this.

C. Free energy calculations

The density of states function (D) was calculated from the sampled velocities using Eq. (22) and subsequently rescaled so that Eq. (23) yields the $3N_A - 6$ vibrational degrees of freedom for the complete system. To allow for an easy comparison especially between chemically identical atoms, single atom spectra were rescaled so that Eq. (23) yields three. ΔA^{DSI} was calculated following the integration in Eq. (24) or (29) and for the molecular examples adding E . E is determined as the potential energy at the minimum geometry, which was for the intermediate systems ($0 < \lambda < 1$) obtained by performing a geometry optimization with the weighted forces [see Eq. (37)] until $\mathbf{F}_\lambda \approx \mathbf{0}$. ΔA^{BAR} is determined from the sampled potential energies by solving Eq. (11). In the case of the serine-cysteine transformation, the free energy change due to the mass change of the atom (oxygen to sulfur) was corrected by an analytically derived constant for each window. For the derivation, see the Appendix. ΔA^{NMA} and ΔA^{QHA} were obtained using Eqs. (15) and (16). The frequencies (ν_i) for the NMA were determined using

the numerically calculated Hessian at the minimum energy geometry [Eq. (14)], while ν_i for the QHA were calculated as presented in Eq. (17).

The vibrational parts $\Delta A_{\text{vib}}^{\text{DSI}}$ and $\Delta A_{\text{vib}}^{\text{BAR}}$ are calculated as

$$\Delta A_{\text{vib}}^{\text{DSI}} = \Delta A^{\text{DSI}} - \Delta E, \quad (38)$$

$$\Delta A_{\text{vib}}^{\text{BAR}} = \Delta A^{\text{BAR}} - \Delta E. \quad (39)$$

As we have simulated several replicas for each λ window, we conduct a statistical analysis calculating the average free energy difference ($\langle \Delta A \rangle$) as

$$\langle \Delta A \rangle = \frac{1}{N_E N_P} \sum_i^{N_E} \sum_j^{N_P} \Delta A_{ij}, \quad (40)$$

where ΔA_{ij} is the free energy difference between the replicas i and j of the educt and product, respectively, and N_E and N_P are their total numbers. We, additionally, calculate the standard deviation of the different ΔA_{ij} . We do not list the inherent statistical error calculated by Shirts and Pande⁴⁵ since there is no analog for DSI.

IV. RESULTS AND DISCUSSION

A. Numerical examples

In order to prove that DSI yields the same results as other free energy methods and to investigate the effect of the shape and curvature of the potential on its accuracy, we carried out classical simulations in an one-dimensional harmonic [Eq. (34)], a Morse [Eq. (36)], and a double-well [Eq. (36)] potential and calculated three free energy changes for each potential (for details, see Sec. III). The free energy changes consist of changes in the curvature of the potential caused by variation of parameters k , a , D_E , and b , resembling changes in molecular angles, bonds, and dihedral angles, respectively (the exact values are listed in the [supplementary material](#)). The resulting $\langle \Delta A \rangle$ s calculated using DSI, BAR, QHA, and NMA as well as the exact results are shown in [Tables I and II](#). For the simulation-based methods, we also provide the standard deviation of ΔA from multiple trajectories. $\langle \Delta A \rangle$ s of the individual λ -windows, potential plots, and density of states plots can be found in the

TABLE I. Calculated free energy changes (average and standard deviation of ΔA_{vib} in kJ/mol) of the harmonic and anharmonic potential (three transformations each) using NMA, QHA, BAR, and DSI. The exact result obtained from numerical integration is given as a reference. The wavenumber (in cm^{-1}) refers to the curvature of the potential at $x = 0$.

Potential	Wavenumbers	NMA	QHA	BAR	DSI	Exact
Harmonic	1000 → 2000	1.718	1.758 ± 0.041	1.709 ± 0.013	1.720 ± 0.002	1.718
	500 → 1000	1.718	1.780 ± 0.068	1.755 ± 0.013	1.732 ± 0.003	1.718
	100 → 500	3.990	4.488 ± 0.056	4.614 ± 0.042	4.166 ± 0.015	3.990
	100 → 500 ^a		3.996 ± 0.048	3.858 ± 0.024	4.007 ± 0.002	
Anharmonic	1000 → 2000	1.718	1.796 ± 0.041	1.735 ± 0.015	1.757 ± 0.002	1.766
	500 → 1000	1.718	1.875 ± 0.076	1.787 ± 0.015	1.766 ± 0.004	1.764
	100 → 500	3.990	4.635 ± 0.043	4.806 ± 0.032	4.260 ± 0.011	4.033
	100 → 500 ^a		4.104 ± 0.070	3.885 ± 0.030	4.047 ± 0.002	

^aTrajectories with longer simulation times.

TABLE II. Calculated free energy changes (average and standard deviation of ΔA_{vib} in kJ/mol) of the double well potential (three transformations) using NMA, QHA, BAR, and DSI. The exact result obtained from numerical integration is given as a reference. The wavenumber (in cm^{-1}) refers to the curvature of the potential at $x = 0$. Double well potentials can be seen as “worst-case” examples for the frequency-based methods.

Potential	Wavenumbers	NMA	QHA	BAR	DSI	Exact
Double well	2000 \rightarrow 2500	0.553	0.592 ± 0.062	0.581 ± 0.004	0.598 ± 0.005	0.608
	1500 \rightarrow 2000	0.713	4.243 ± 0.076	0.810 ± 0.008	0.910 ± 0.014	0.825
	1000 \rightarrow 1500	1.005	-0.159 ± 0.147	1.269 ± 0.008	1.682 ± 0.036	1.068
	1000 \rightarrow 1500 ^a		-0.151 ± 0.042	1.224 ± 0.008	1.611 ± 0.029	

^aTrajectories with longer simulation times.

supplementary material. Please note that ΔE is for all numerical cases zero so that $\Delta A = \Delta A_{\text{vib}}$.

Comparing the methods that use sampled data along classical simulations (QHA, BAR, DSI) in the harmonic and anharmonic cases (see Table I), DSI performs best, showing smaller errors and standard deviations. In the latter case, it even outperforms NMA, which is only exact, when harmonic potentials are investigated. The error of all three simulation-based methods increases with decreasing curvature (wavenumbers) of the investigated potentials. The standard deviations of BAR and DSI also increase, while it remains constant in the case of QHA. The first reason for the larger errors (and standard deviations) is the sampling error as slower vibrations require longer simulation times to be sampled accurately. The second reason is the choice of the thermostat. As already discussed in Sec. III A, we are bound to velocity rescaling thermostats when applying DSI. These thermostats, however, introduce errors in the free energy calculations, especially for systems with only a few degrees of freedom and for slow modes. Consequently, we observe a decrease in the errors when the relaxation time of the thermostat is increased (resulting in a weaker thermostat). Longer simulation times tackle both problems discussed above and, therefore, improve the errors (and standard deviations of BAR and DSI) for the low-frequency harmonic and anharmonic cases significantly (see footnote a in Table I), which are then in good agreement with numerical test potentials found elsewhere.^{45,66}

For the frequency-based methods, the double well potential is a “worst-case” example as it cannot be described exactly within the harmonic approximation. It features two types of movements: the movement within one well and the slower movement over the barrier (the inversion). With decreasing curvature (wavenumber), the barrier height shrinks, increasing the probability of the inversion. This set up is not problematic for BAR as it relies on energy averages and distributions and uses intermediate systems ($0 < \lambda < 1$) to enhance the sampling efficiency. This explains the small error and standard deviation in Table II, which is comparable to the values of the other examples.

For NMA, QHA, and DSI, it serves as a good showcase to illustrate the conceptual differences between these methods and to show how well the actual potential can be approximated. In Fig. 1(a), we have plotted the double well potential (1000 cm^{-1}) and the corresponding harmonic potentials, which are used by NMA and

QHA to approximate the potential and to calculate the free energy. Figure 1(b) presents the same for DSI. The difference is that not a single but a series of harmonic potentials (illustrated by the red and blue areas) weighted by the density of states functions [also plotted in Fig. 1(b)] are assumed to describe the system.

The NMA harmonic oscillator, derived from a finite difference calculation around $x = 0$, mimics the fast vibration, while the QHA harmonic oscillator, derived from the distribution of x ,

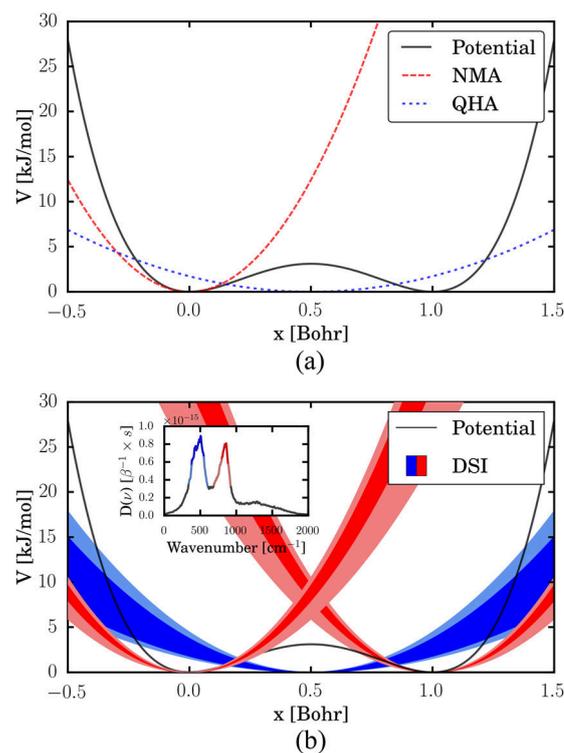


FIG. 1. (a) Double well potential (1000 cm^{-1}) and corresponding harmonic oscillators of NMA and QHA from which the free energy is calculated. (b) Double well potential (1000 cm^{-1}) and the series of weighted harmonic oscillators used in DSI to obtain the free energy. The frequencies have been extracted from the density of states function (see the subplot). Please note that the density of states function ($D(\nu)$) is given in $\beta^{-1} \times s$; we have omitted the factor β in Eq. (22).

is dominated by the slow vibration. In this example, the first is a good approximation (most likely due to error compensation), while the latter leads to erratic results (Table II). The density of states function shows that both vibrations are considered in DSI as two peaks with the same intensity appear in the spectrum. In this case, both vibrations have the same probability due to the low barrier height of around 2.5 kJ/mol $\approx RT$. When larger barrier heights are applied, the intensity of the fast vibration becomes significantly larger than the intensity of the slow vibration (see the [supplementary material](#)). Note that the parameter b changes not only the barrier height but also the curvature of the potential at the same time.

In this example, longer simulation times lead to better DSI results, but the difference to the exact result is still significantly larger than in BAR. This indicates that even an arbitrary large number of harmonic oscillators are incapable of describing the system correctly. However, the description of the system in DSI is physically more correct than in the case of NMA and QHA, and the resulting free energy estimation is significantly better than in the case of QHA.

B. Ionization energy of ammonia

As a first molecular example, we have chosen the ionization of ammonia ($\text{NH}_3 \rightarrow \text{NH}_3^+$). (Alchemical) *ab initio* molecular dynamics simulations have been performed at the HF-3c level of theory (see Sec. III for details). The results for the overall reaction and for the individual λ -windows are listed in Table III.

Table III proves that the BAR result can be improved by taking into account intermediate λ -windows as the cumulative result (using all intermediate windows) differs from the direct result (using only the two end points) while featuring a 10 times smaller spread. This is not the case for DSI. Since all contributions of the intermediate λ -simulations for E and D cancel out, the results are

TABLE III. Calculated vibrational free energy changes (average and standard deviation of ΔA_{vib} in kJ/mol) of the ionization of ammonia (direct from the first to the last λ -value and cumulative over all windows) and the intermediate λ -windows using BAR, DSI, and NMA.

λ -window	$\Delta A_{\text{vib}}^{\text{DSI}}$	$\Delta A_{\text{vib}}^{\text{BAR}}$	$\Delta A_{\text{vib}}^{\text{NMA}}$
0.0 \rightarrow 0.1	-0.26 ± 1.05	-1.08 ± 0.03	...
0.1 \rightarrow 0.2	-0.45 ± 0.82	-0.94 ± 0.03	...
0.2 \rightarrow 0.3	-0.31 ± 0.65	-0.44 ± 0.04	...
0.3 \rightarrow 0.4	-0.96 ± 0.59	0.22 ± 0.04	...
0.4 \rightarrow 0.5	0.90 ± 0.79	0.35 ± 0.06	...
0.5 \rightarrow 0.6	-0.48 ± 1.06	0.08 ± 0.05	...
0.6 \rightarrow 0.7	-0.24 ± 0.86	-0.03 ± 0.03	...
0.7 \rightarrow 0.8	-0.39 ± 0.55	-0.11 ± 0.02	...
0.8 \rightarrow 0.9	0.36 ± 0.42	-0.17 ± 0.03	...
0.9 \rightarrow 1.0	-0.41 ± 0.30	-0.22 ± 0.03	...
Cumulative	-2.24 ± 2.70	-2.36 ± 0.12	...
Direct	-2.24 ± 0.77	-2.50 ± 1.56	-1.97
Quantum corrected	-13.76 ± 5.06	...	-13.26

identical. Only the standard deviation is larger in the cumulative case due to the noise between the intermediate states. The convergence of the direct DSI result and the cumulative BAR result are shown in Fig. 2.

The standard deviation on both curves decreases with increasing simulation time. After 250 ps simulation time (per trajectory), both methods yield the same result within one standard deviation of the BAR curve, and at ≈ 300 ps, the mean results are nearly identical (see also Table III). Even longer simulations (up to 900 ps) do not substantially affect the average DSI result, while the standard deviation is reduced (see Fig. S7 of the [supplementary material](#)). However, even when 2×10 trajectories of 900 ps are used in the DSI calculation, the standard deviation is still approximately two times larger than the one observed in BAR featuring an (almost) equivalent amount of data points. If one takes into account that no alchemical simulations (one energy and force calculation per step instead of two) are required for DSI, we could say that (for this example) the standard deviations of DSI and BAR behave similarly with respect to the computation time, while the average free energy change seems to converge faster in the case of DSI. At this point, we also want to mention that there are two factors which can decrease the accuracy of DSI: Too short simulations and too long intervals between the sampling of the nuclear velocities are applied (see Figs. S7 and S8 of the [supplementary material](#)).

The results of DSI and BAR for the intermediate λ -windows differ usually by about one standard deviation, except for 0.3 \rightarrow 0.4. The histograms of the improper dihedral of ammonia (Fig. S4 of the [supplementary material](#)) reveal that for these cases the system is similar to the double well system we have discussed in Sec. IV A, which explains the larger error. In the other windows, the barrier is either too high for a frequent inversion of the molecule or vanishes entirely. The figure also shows that for $\lambda = 0.0$ the simulation has not spent equal amounts of time in the two minima of ammonia, which should bias the BAR results. In general, we observe a relatively high standard deviation for the DSI free energies of the intermediate λ -windows. The reason for this could be that mixed potential energy surfaces tend to be more anharmonic or even nonharmonic (e.g., the λ -window 0.3 \rightarrow 0.4), showing larger errors and slower convergence.

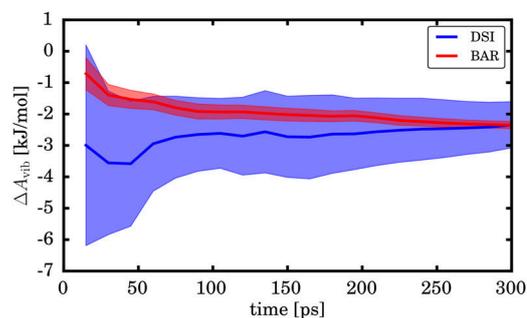


FIG. 2. Convergence of the total free energy change ($\langle \Delta A \rangle$, solid line) of the ionization of ammonia and the standard deviation (lighter area) with respect to the length of the used trajectories using BAR and DSI. BAR contains information from 11×5 trajectories (five replicas for all 11 λ -windows), whereas the DSI result is only based on 2×5 trajectories (five replicas for NH_3 and NH_3^+ , respectively).

Additionally, the DSI method offers two features that are not accessible in energy-based methods. One can easily calculate the quantum corrected free energy change (see the last line in Table III), and one can map the change in $\langle \Delta A_{\text{vib}}^{\text{DSI}} \rangle$ to each atom or when dealing with larger problems, groups of atoms or molecules. In the case of the ionization of ammonia, the hydrogen atoms and the nitrogen atom gain vibrational free energy (0.59 kJ/mol and 0.46 kJ/mol, respectively) since the bonds in NH_3^+ are weaker than in NH_3 . This is also reflected in the power spectrum (D , see Fig. 3), where nearly all modes of NH_3 are red-shifted in NH_3^+ .

Figure 3 also shows the results of NMA and QHA as vertical dashed lines. As one can see, the frequencies estimated with NMA are in good agreement with D and are always positioned at the upper bound of the peaks in D . This is due to the fact that NMA does not consider any anharmonicity in the bond vibrations, which causes the slight decrease in the vibrational frequency and the vibrational free energy change (see Table III). QHA results clearly underestimate all frequencies and suggest unreasonably slow motions, especially for the inversion motion of NH_3 .

C. Mutation from serine to cysteine

The second example consists of the mutation from serine to cysteine in vacuum. Mutations are a widely used tool in free

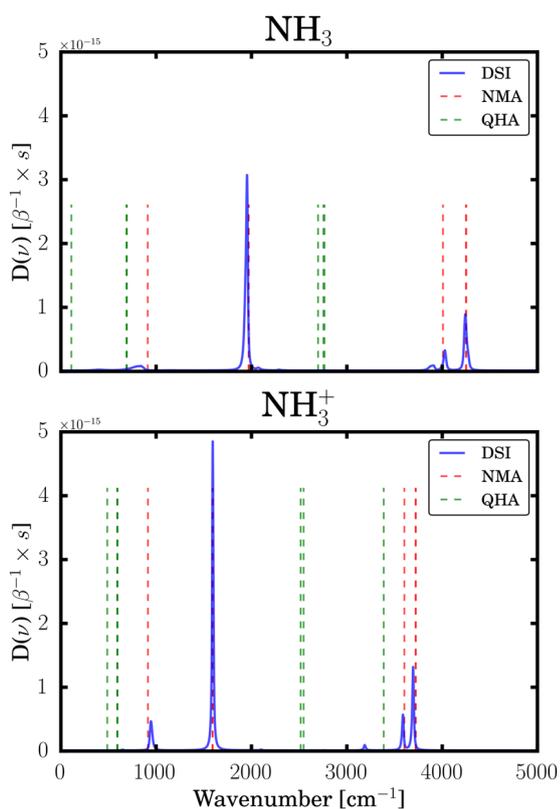


FIG. 3. Velocity density of states spectrum of NH_3 and NH_3^+ . The frequencies obtained from NMA and QHA are shown in red and green, respectively. Please note that $D(\nu)$ is given in $\beta^{-1} \times \text{s}$; we have omitted the factor β in Eq. (22).

TABLE IV. Calculated vibrational free energy changes (average and standard deviation of ΔA_{vib} in kJ/mol) of the mutation from serine to cysteine (direct from the first to the last λ -value and cumulative over all windows) and the intermediate λ -windows using BAR, DSI, and NMA.

λ -Window	$\Delta A_{\text{vib}}^{\text{DSI}}$	$\Delta A_{\text{vib}}^{\text{BAR}}$	$\Delta A_{\text{vib}}^{\text{NMA}}$
0.0 \rightarrow 0.1	-0.79 ± 3.03	0.56 ± 0.40	...
0.1 \rightarrow 0.2	0.06 ± 2.94	-0.06 ± 0.51	...
0.2 \rightarrow 0.3	-2.10 ± 3.70	-0.74 ± 0.26	...
0.3 \rightarrow 0.4	0.40 ± 4.03	-0.43 ± 0.31	...
0.4 \rightarrow 0.5	-3.37 ± 3.80	-0.35 ± 0.32	...
0.5 \rightarrow 0.6	0.59 ± 3.47	-0.27 ± 0.10	...
0.6 \rightarrow 0.7	-1.66 ± 3.57	-0.48 ± 0.10	...
0.7 \rightarrow 0.8	-1.26 ± 4.38	-0.35 ± 0.10	...
0.8 \rightarrow 0.9	0.66 ± 4.31	-0.35 ± 0.05	...
0.9 \rightarrow 1.0	-1.47 ± 3.16	-0.31 ± 0.05	...
Cumulative	-8.94 ± 12.46	-2.76 ± 0.85	...
Direct	-8.94 ± 2.76	290.93 ± 177.98	-3.69
Quantum corrected	-23.55 ± 7.24	...	-15.76

energy calculations as they give access to, e.g., binding free energies. We conducted extensive (alchemical) *ab initio* molecular dynamics simulations of serine, cysteine, and intermediate structures (see Sec. III for details) and calculated the free energy for the overall reaction and for the individual λ -windows. The results are presented in Table IV.

The one-step application of BAR in Table IV shows the wrong sign and is about two orders of magnitude too large. Significantly better results for similar one-step mutations have been reported for MM simulations.⁶⁷ However, the underlying data consisted of four 168 ns trajectories, which contain nearly 1000 times more conformations than our *ab initio* simulations.

The results of direct DSI are, in comparison with direct BAR, significantly better. The final results of DSI and BAR are within two

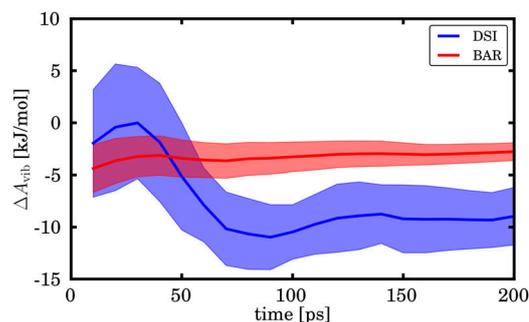


FIG. 4. Convergence of the total free energy change ($\langle \Delta A \rangle$, solid line) of the serine-cysteine transformation and its standard deviation (lighter area) with respect to the length of the used trajectories using BAR and DSI. BAR contains information from 11×5 trajectories (five replicas for all 11 λ -windows), whereas the DSI result is only based on 2×5 trajectories (five replicas for serine and cysteine, respectively).

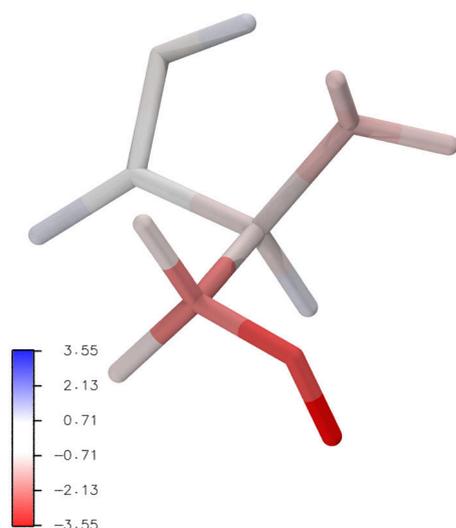


FIG. 5. Vibrational free energy change ($\langle \Delta A_{\text{vib}}^{\text{DSI}} \rangle$ in kJ/mol) of each atom during the mutation from serine to cysteine. The major changes occur exactly at the atoms that are connected or close to the O/S-mutation.

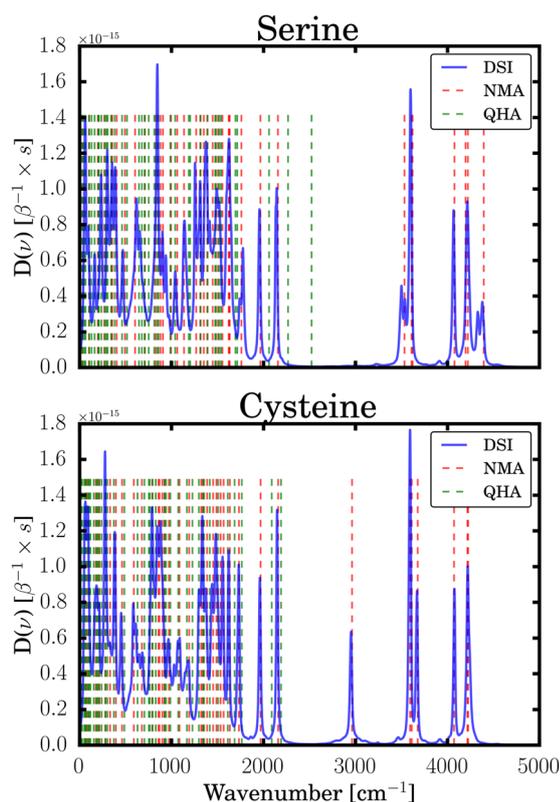


FIG. 6. Velocity density of states spectrum of serine and cysteine. The frequencies obtained from NMA and QHA are shown in red and green, respectively. Please note that $D(\nu)$ is given in $\beta^{-1} \times \text{s}$; we have omitted the factor β in Eq. (22).

standard deviations of one another and both are also more than one standard deviation off from the NMA results which lies between the two results. Nevertheless, the standard deviations of DSI are one order of magnitude larger than those of BAR. This and the convergence plot of the free energies (see Fig. 4) indicate that further sampling is required, mainly due to the larger amount of low-frequency modes (see Fig. 6) in this example, which is beyond the focus of the present manuscript.

Despite the numerical noise, the trend given by DSI is correct, and localizing the vibrational free energy (Fig. 5) offers an explanation where and why the free energy changes during the mutation. It clearly shows that main contributors are the C_{β} -atom as well as the connected hydroxyl- or thiol-group. Slight contributions stem also from the C_{α} -atom and the amino-group as the strength of the intermittently formed hydrogen bond between hydroxy/thiol-group and amino-group changes. This also has a small effect on the carboxyl-group. These results agree well with chemical intuition.

Figure 6 shows the power spectra of serine and cysteine as well as the NMA and QHA results. The position of the NMA frequencies and the main peaks in the power spectrum agree well. Both show a large amount of low-frequency modes, so-called “breathing modes.” Nevertheless, NMA neglects the anharmonicity of the vibrations and the different minima of the system, leading to a different free energy (see Table IV). QHA significantly overestimates the existence of these modes and fails to find the high-frequency bond vibrations.

V. CONCLUSION

In this work, we have tested and compared the density of states method (DSI) to the more established free energy methods such as Normal Mode Analysis (NMA), Quasiharmonic Analysis (QHA), and Bennett’s Acceptance Ratio method (BAR), calculating several numerical and two chemical examples. We show that DSI works similar to NMA and QHA, but features the ability to correctly include anharmonicities, as the partition function is approximated by an arbitrary large number of harmonic oscillators weighted by the density of states function.

DSI delivers the same result as BAR for the numerical examples and the ionization of ammonia. Regarding the mutation from serine to cysteine, DSI correctly reflects the trend of the free energy, but features larger standard deviations, mainly due to the large number of low-frequency modes in the systems. This indicates that long simulation times will be required for larger systems. Additional downsides of the method regarding free energy calculations are as follows: (1) larger memory requirements ($3 \times N_A$ velocities have to be stored in short intervals instead of one energy at arbitrary long intervals), (2) Monte Carlo simulations and enhanced sampling methods cannot be combined with DSI, and (3) ΔE has to be determined, which will become tedious for large systems with many degrees of freedom.

There are, however, also important advantages of the method, when comparing to BAR:

1. For DSI, only the end points (no intermediates) are required. This gives access to free energies of nearly all molecular

transformations even at the *ab initio* level of theory, circumventing endpoint-catastrophes and alchemical transformations. Additionally, this can lead to a reduction of computation time when DSI is applied to small molecular systems.

- Quantum-corrected vibrational free energies are directly accessible.
- A straightforward pattern to determine atomwise or residue-wise contributions to the vibrational free energies exists.

We, therefore, think that DSI can be a good alternative to standard free energy methods, especially when expensive *ab initio* methods are applied to transformations of small to medium-sized molecules. Furthermore, its ability to localize free energy changes at atoms or residues is a valuable tool to gain insights into the underlying process(es), which can always be combined with energy-based methods such as BAR.

SUPPLEMENTARY MATERIAL

The [supplementary material](#) comprises details on our numerical examples where we give details on the used potentials and show the corresponding density of states plots as well as ΔA 's for the intermediate λ -windows. For the ionization of ammonia, we present dihedral distributions and ΔA convergence studies of the intermediate λ -windows. Additionally, we show the mapping of $\Delta A_{\text{vib}}^{\text{DSI}}$ on the individual atoms and the convergence of $A_{\text{vib}}^{\text{DSI}}$ with respect to simulation time and the sampling frequency. For the serine-cysteine transformation, we present the distributions of the C-O/S bond lengths for all λ -windows.

ACKNOWLEDGMENTS

Financial support was provided by the SFB 1309 "Chemical Biology of Epigenetic Modifications" (DFG), SFB 749 "Dynamics and Intermediates of Molecular Transformations" (DFG), and the DFG cluster of excellence (EXC 114) "Center for Integrative Protein Science Munich" (CIPSM). C.O. acknowledges further support as Max-Planck-Fellow at the MPI-FKF Stuttgart.

APPENDIX: INFLUENCE OF THE MASS CHANGE IN THE FREE ENERGY OF THE TRANSFORMATION FROM SERINE TO CYSTEINE

Changing the mass of a particle has an impact on the free energy as the kinetic energy distribution of the particle changes. The canonical partition function of a system consisting of N_A distinguishable particles has the form

$$Q = \frac{1}{h^{3N_A}} \int dx^{3N_A} \int dp^{3N_A} \exp\{-\beta H(x^{3N_A}, p^{3N_A})\}. \quad (\text{A1})$$

The Hamiltonian (H) is usually split into the potential (U) and kinetic energy (T) which are functions of the generalized coordinates (\mathbf{x}) and generalized impulses (\mathbf{p}), respectively. Hence, the above integral can be split into the product of kinetic and potential energy contributions

$$\begin{aligned} Q &= \frac{1}{h^{3N_A}} \int dx^{3N_A} \exp\{-\beta U(x^{3N_A})\} \int dp^{3N_A} \exp\{-\beta T(p^{3N_A})\} \\ &= \frac{1}{h^{3N_A}} \int dx^{3N_A} \exp\{-\beta U(x^{3N_A})\} \int_{-\infty}^{\infty} dp^{3N_A} \exp\{-\beta \sum_i^{3N_A} \frac{p_i^2}{2m_i}\}. \end{aligned} \quad (\text{A2})$$

The integration over the kinetic part can be carried out analytically and yields

$$Q = \frac{1}{h^{3N_A}} \prod_i^{3N_A} \sqrt{\frac{2\pi m_i}{\beta}} \int dx^{3N_A} \exp\{-\beta U(x^{3N_A})\}. \quad (\text{A3})$$

If we consider now the free energy difference between two systems, where not only the potential energy function changes but also the mass of one particle, we can write

$$\begin{aligned} \Delta A_{0 \rightarrow 1} &= -\beta^{-1} \ln \left[\frac{Q_1}{Q_0} \right] \\ &= -\beta^{-1} \ln \left[\frac{\prod_i^{3N_A} \sqrt{\frac{2\pi m_i}{\beta}} \int dx^{3N_A} \exp\{-\beta U_1(x^{3N_A})\}}{\prod_i^{3N_A} \sqrt{\frac{2\pi m_i}{\beta}} \int dx^{3N_A} \exp\{-\beta U_0(x^{3N_A})\}} \right] \\ &= -\beta^{-1} \ln \left[\sqrt{\frac{m_1}{m_0}} \langle \exp\{-\beta \Delta U\} \rangle_0 \right]. \end{aligned} \quad (\text{A4})$$

In our case, m_1 is m_{O} , the atomic mass of oxygen, and m_1 is m_{S} , the atomic mass of sulfur

$$\Delta A_{0 \rightarrow 1} = -\frac{3}{2} \beta^{-1} \ln \frac{m_{\text{S}}}{m_{\text{O}}} - \beta^{-1} \ln \langle \exp\{-\beta \Delta U\} \rangle_0. \quad (\text{A5})$$

Thus, the results of BAR have to be corrected by

$$\Delta A_{0 \rightarrow 1}^{\text{mass}} = -\frac{3}{2} \beta^{-1} \ln \frac{m_{\text{S}}}{m_{\text{O}}} = -2.58 \text{ kJ/mol}. \quad (\text{A6})$$

For each individual λ -window, the correction reads

$$\Delta A_{\lambda \rightarrow \lambda + \Delta \lambda}^{\text{mass}} = -\frac{3}{2} \beta^{-1} \ln \left[1 + \Delta \lambda \frac{m_{\text{S}} - m_{\text{O}}}{(1 - \lambda)m_{\text{O}} + \lambda m_{\text{S}}} \right]. \quad (\text{A7})$$

REFERENCES

- Free Energy Calculations*, edited by C. Chipot and A. Pohorille (Springer-Verlag, Berlin, Heidelberg, 2007).
- P. Kollman, *Chem. Rev.* **93**, 2395 (1993).
- C. D. Christ, A. E. Mark, and W. F. van Gunsteren, *J. Comput. Chem.* **31**, 1569 (2010).
- M. A. Olsson and U. Ryde, *J. Chem. Theory Comput.* **13**, 2245 (2017).
- U. Ryde and P. Söderhjelm, *Chem. Rev.* **116**, 5520 (2016).
- O. K. Dudko, G. Hummer, and A. Szabo, *Phys. Rev. Lett.* **96**, 108101 (2006).
- E. Neria, S. Fischer, and M. Karplus, *J. Chem. Phys.* **105**, 1902 (1996).
- D. A. McQuarrie and S. D. Simon, *Molecular Thermodynamics* (University Science Books, Sausalito, CA, 1999).
- E. Fermi, J. Pasta, and S. Ulam, Los Alamos Report No. LA-1940, 1955.
- B. J. Alder and T. E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959).
- A. Rahman, *Phys. Rev. A* **136**, A405 (1964).
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).

- ¹³W. K. Hastings, *Biometrika* **57**, 97 (1970).
- ¹⁴B. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. U. S. A.* **80**, 6571 (1983).
- ¹⁵B. Tidor and M. Karplus, *J. Mol. Biol.* **238**, 405 (1994).
- ¹⁶M. Karplus and J. N. Kushick, *Macromolecules* **14**, 325 (1981).
- ¹⁷R. M. Levy, A. R. Srinivasan, and W. K. Olson, *Biopolymers* **23**, 1099 (1984).
- ¹⁸R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
- ¹⁹J. G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).
- ²⁰C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
- ²¹J. Gao, K. Kuczera, B. Tidor, and M. Karplus, *Science* **244**, 1069 (1989).
- ²²M. Karplus and G. A. Petsko, *Nature* **347**, 631 (1990).
- ²³J. M. Rickman and R. LeSar, *Annu. Rev. Mater. Res.* **32**, 195 (2002).
- ²⁴P. A. Bash, U. C. Singh, R. Langridge, and P. A. Kollman, *Science* **236**, 564 (1987).
- ²⁵P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge, and P. A. Kollman, *Science* **235**, 574 (1987).
- ²⁶R. Rathore, M. Sumakanth, M. S. Reddy, P. Reddanna, A. A. Rao, M. D. Erion, and M. Reddy, *Curr. Pharm. Des.* **19**, 4674 (2013).
- ²⁷C. Chipot, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **4**, 71 (2014).
- ²⁸D. L. Mobley and M. K. Gilson, *Annu. Rev. Biophys.* **46**, 531 (2017).
- ²⁹A. P. Bhati, S. Wan, D. W. Wright, and P. V. Coveney, *J. Chem. Theory Comput.* **13**, 210 (2017).
- ³⁰M. Cecchini, *J. Chem. Theory Comput.* **11**, 4011 (2015).
- ³¹P. H. Berens, D. H. J. Mackay, G. M. White, and K. R. Wilson, *J. Chem. Phys.* **79**, 2375 (1983).
- ³²S.-T. Lin, M. Blanco, and W. A. Goddard, *J. Chem. Phys.* **119**, 11792 (2003).
- ³³S.-T. Lin, P. K. Maiti, and W. A. Goddard, *J. Phys. Chem. B* **114**, 8191 (2010).
- ³⁴R. A. X. Persson, V. Pattni, A. Singh, S. M. Last, and M. Heyden, *J. Chem. Theory Comput.* **13**, 4467 (2017).
- ³⁵S. Belsare, V. Pattni, M. Heyden, and T. Head-Gordon, *J. Phys. Chem. B* **122**, 5300 (2018).
- ³⁶M. Heyden, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **9**, e1390 (2019).
- ³⁷D. Marx and J. Hutter, in *Modern Methods and Algorithms of Quantum Chemistry – Proceedings*, 2nd ed., NIC Series, Vol. 3, edited by J. Grotendorst (NIC-Directors, Juelich, 2000), pp. 329–477.
- ³⁸B. Kirchner, P. J. di Dio, and J. Hutter, *Top. Curr. Chem.* **307**, 109 (2012).
- ³⁹R. Sure and S. Grimme, *J. Comput. Chem.* **34**, 1672 (2013).
- ⁴⁰D. M. Zuckerman and T. B. Woolf, *Phys. Rev. Lett.* **89**, 180602 (2002).
- ⁴¹N. Lu and D. A. Kofke, *J. Chem. Phys.* **114**, 7303 (2001).
- ⁴²W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.* **83**, 3050 (1985).
- ⁴³M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, *Phys. Rev. Lett.* **91**, 140601 (2003).
- ⁴⁴S. Bruckner and S. Boresch, *J. Comput. Chem.* **32**, 1303 (2011).
- ⁴⁵M. R. Shirts and V. S. Pande, *J. Chem. Phys.* **122**, 144107 (2005).
- ⁴⁶A. de Ruiter, S. Boresch, and C. Oostenbrink, *J. Comput. Chem.* **34**, 1024 (2013).
- ⁴⁷G. König and B. R. Brooks, *J. Comput.-Aided Mol. Des.* **26**, 543 (2012).
- ⁴⁸G. König, B. R. Brooks, W. Thiel, and D. M. York, *Mol. Simul.* **44**, 1062 (2018).
- ⁴⁹F. M. Ytreberg, R. H. Swendsen, and D. M. Zuckerman, *J. Chem. Phys.* **125**, 184114 (2006).
- ⁵⁰J. C. B. Dietschreit, L. D. M. Peters, J. Kussmann, and C. Ochsenfeld, *J. Phys. Chem. A* **123**, 2163 (2019).
- ⁵¹P. E. Smith and W. F. van Gunsteren, *J. Phys. Chem.* **98**, 13735 (1994).
- ⁵²A. E. Mark and W. F. van Gunsteren, *J. Mol. Biol.* **240**, 167 (1994).
- ⁵³J. Kussmann and C. Ochsenfeld, *J. Chem. Phys.* **138**, 134114 (2013).
- ⁵⁴J. Kussmann and C. Ochsenfeld, *J. Chem. Theory Comput.* **11**, 918 (2015).
- ⁵⁵J. Kussmann and C. Ochsenfeld, *J. Chem. Theory Comput.* **13**, 3153 (2017).
- ⁵⁶S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- ⁵⁷S. Grimme, S. Ehrlich, and L. Goerigk, *J. Comput. Chem.* **32**, 1456 (2011).
- ⁵⁸H. Kruse and S. Grimme, *J. Chem. Phys.* **136**, 154101 (2012).
- ⁵⁹L. Verlet, *Phys. Rev.* **159**, 98 (1967).
- ⁶⁰W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).
- ⁶¹G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
- ⁶²M. R. Reddy, U. C. Singh, and M. D. Erion, *J. Am. Chem. Soc.* **126**, 6224 (2004).
- ⁶³S. Boresch and M. Karplus, *J. Phys. Chem. A* **103**, 103 (1999).
- ⁶⁴H. Hu and W. Yang, *J. Chem. Phys.* **123**, 041102 (2005).
- ⁶⁵G. König, P. S. Hudson, S. Boresch, and H. L. Woodcock, *J. Chem. Theory Comput.* **10**, 1406 (2014).
- ⁶⁶G. König, B. T. Miller, S. Boresch, X. Wu, and B. R. Brooks, *J. Chem. Theory Comput.* **8**, 3650 (2012).
- ⁶⁷G. König, S. Bruckner, and S. Boresch, *J. Comput. Chem.* **30**, 1712 (2009).

**Calculating Free Energies from the Vibrational Density of States Function:
Validation and Critical Assessment - Supporting Information**

Laurens D. M. Peters,^{1,2, a)} Johannes C. B. Dietschreit,^{1,2, a)} Jörg Kussmann,^{1,2} and
Christian Ochsenfeld^{1,2, b)}

¹⁾*Chair of Theoretical Chemistry, Department of Chemistry,
University of Munich (LMU), Butenandtstr. 7, D-81377 München,
Germany*

²⁾*Center for Integrated Protein Science (CIPSM) at the Department of Chemistry,
University of Munich (LMU), Butenandtstr. 5-13, D-81377 München,
Germany*

(Dated: 12 April 2019)

^{a)}These two authors contributed equally to this work

^{b)}Electronic mail: christian.ochsenfeld@uni-muenchen.de

I. NUMERICAL EXAMPLES

A. Harmonic Potential

TABLE S1. Used values (in atomic units) for k in the simulations of the initial (index 0) and final (index 1) state.

Wavenumbers	k_0	k_1
100 \rightarrow 500	3.78e-04	9.46e-03
500 \rightarrow 1000	9.46e-03	3.78e-02
1000 \rightarrow 2000	3.78e-02	1.51e-01

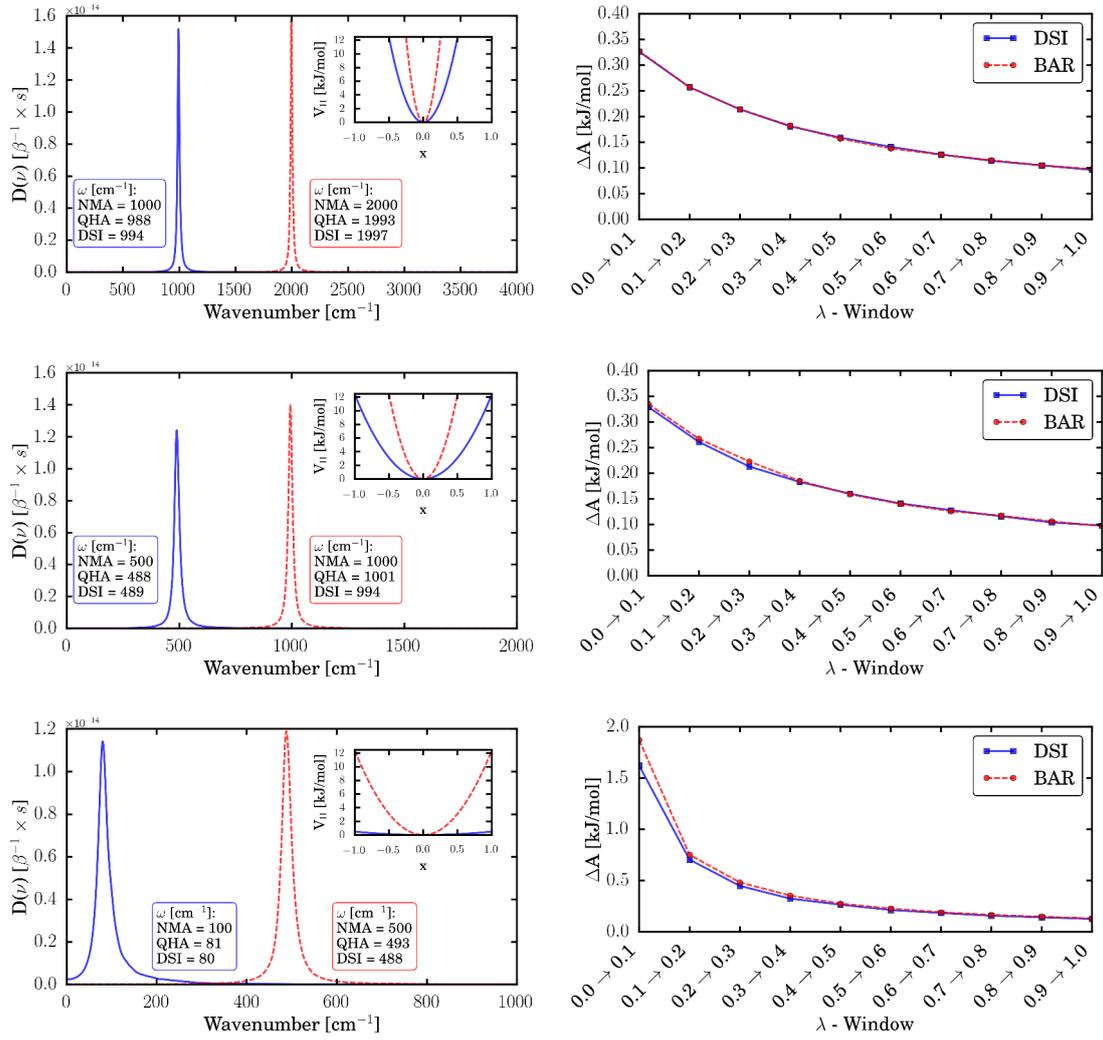


FIG. S1. (left) Harmonic potentials and density of states plots of the simulations. In addition, their maxima (DSI) and the frequencies calculated with NMA and QHA are listed. (right) $\langle \Delta A \rangle$ s of the individual λ -windows of the corresponding simulations. Please note that the density of states function ($D(\nu)$) is given in $\beta^{-1} \times s$; We have omitted the factor β in eq. (22).

B. Anharmonic Potential

TABLE S2. Used values (in atomic units) for D_E and a in the simulations of the initial (index 0) and final (index 1) state.

Wavenumbers	D_{E0}	D_{E1}	a_0	a_1
100 \rightarrow 500	1.55e-02	3.10e-02	1.11e-01	3.91e-01
500 \rightarrow 1000	1.55e-02	3.10e-02	5.53e-01	7.82e-01
1000 \rightarrow 2000	1.55e-02	3.10e-02	1.11e+00	1.56e+00

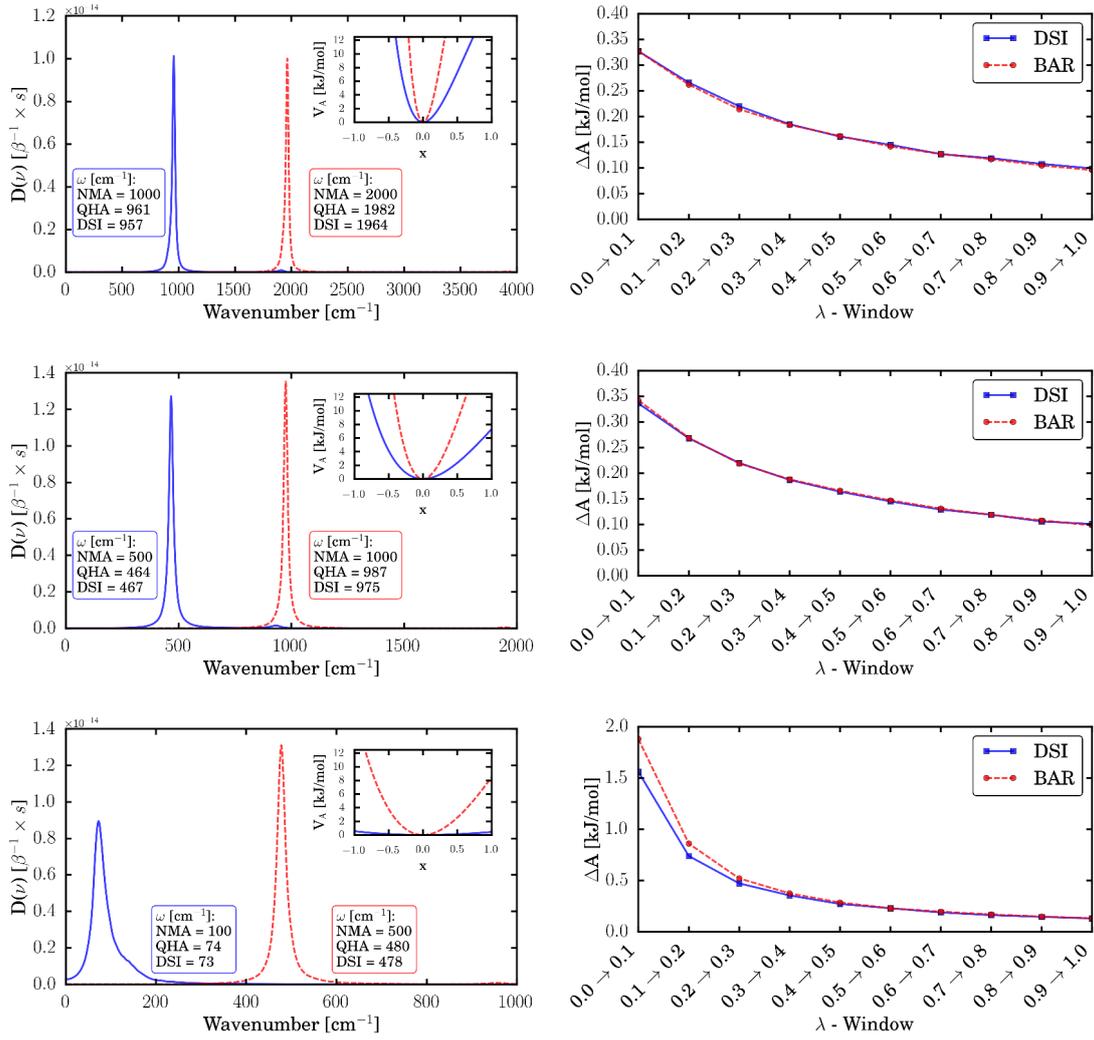


FIG. S2. (left) Anharmonic potentials and density of states plots of the simulations. In addition, their maxima (DSI) and the frequencies calculated with NMA and QHA are listed. (right) $\langle \Delta A \rangle$ s of the individual λ -windows of the corresponding simulations. Please note that the density of states function ($D(\nu)$) is given in $\beta^{-1} \times s$; We have omitted the factor β in eq. (22).

C. Double Well Potential

TABLE S3. Used values (in atomic units) for b in the simulations of the initial (index 0) and final (index 1) state.

Wavenumbers	b_0	b_1
100 \rightarrow 1500	3.78e-02	8.51e-02
1500 \rightarrow 2000	8.51e-02	1.51e-01
2000 \rightarrow 2500	1.51e-01	2.37e-01

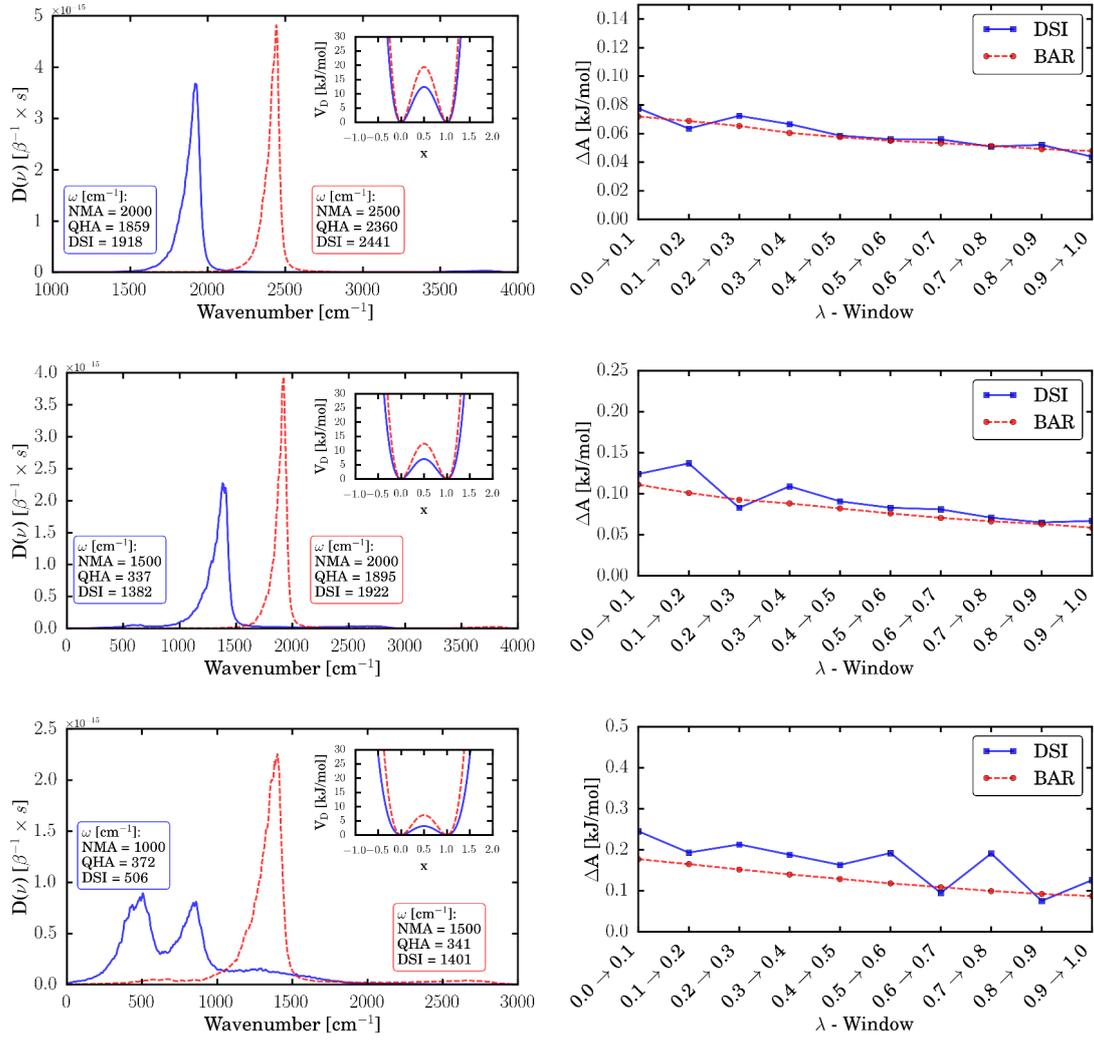


FIG. S3. (left) Double well potentials and density of states plots of the simulations. In addition, their maxima (DSI) and the frequencies calculated with NMA and QHA are listed. (right) $\langle \Delta A \rangle$ s of the individual λ -windows of the corresponding simulations. Please note that the density of states function ($D(\nu)$) is given in $\beta^{-1} \times s$; We have omitted the factor β in eq. (22).

II. IONIZATION ENERGY OF AMMONIA

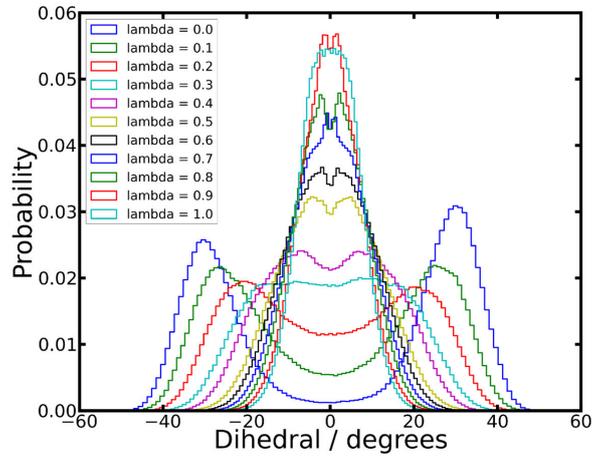


FIG. S4. Normalized distribution of the improper dihedral of ammonia for each λ -value over 100 bins. The histograms include the data of all five simulations per λ -value.

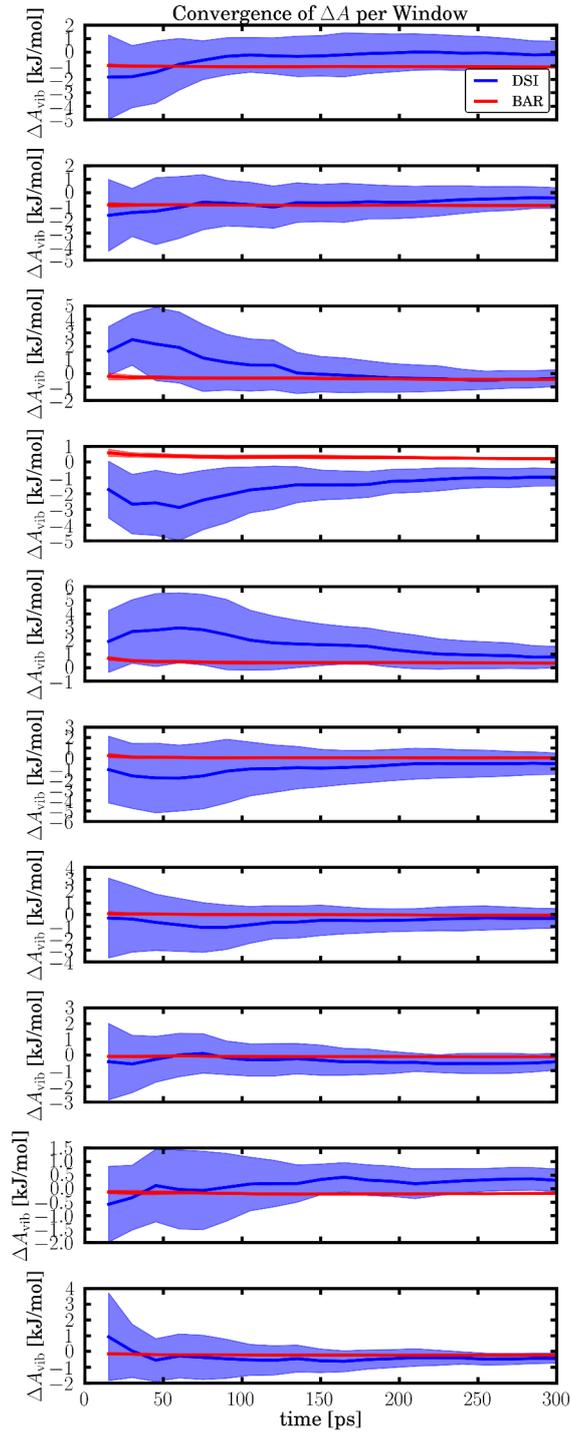


FIG. S5. Convergence of the mean vibrational free energy change ($\langle\Delta A\rangle$, solid line) and the standard deviation of different trajectories (lighter area) using BAR and DSI for each λ -window. The top panel shows the changes for $0.0 \rightarrow 0.1$, the second panel from the top pertains to $0.1 \rightarrow 0.2$, and so forth.

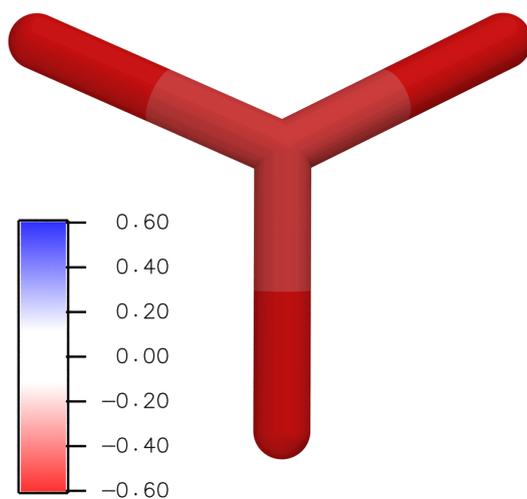


FIG. S6. Vibrational free energy change ($\langle \Delta A_{\text{vib}}^{\text{DSI}} \rangle$ in kJ/mol) of each atom during the ionization of ammonia. The hydrogen atoms lose 0.59 kJ/mol and the nitrogen atom 0.46 kJ/mol.

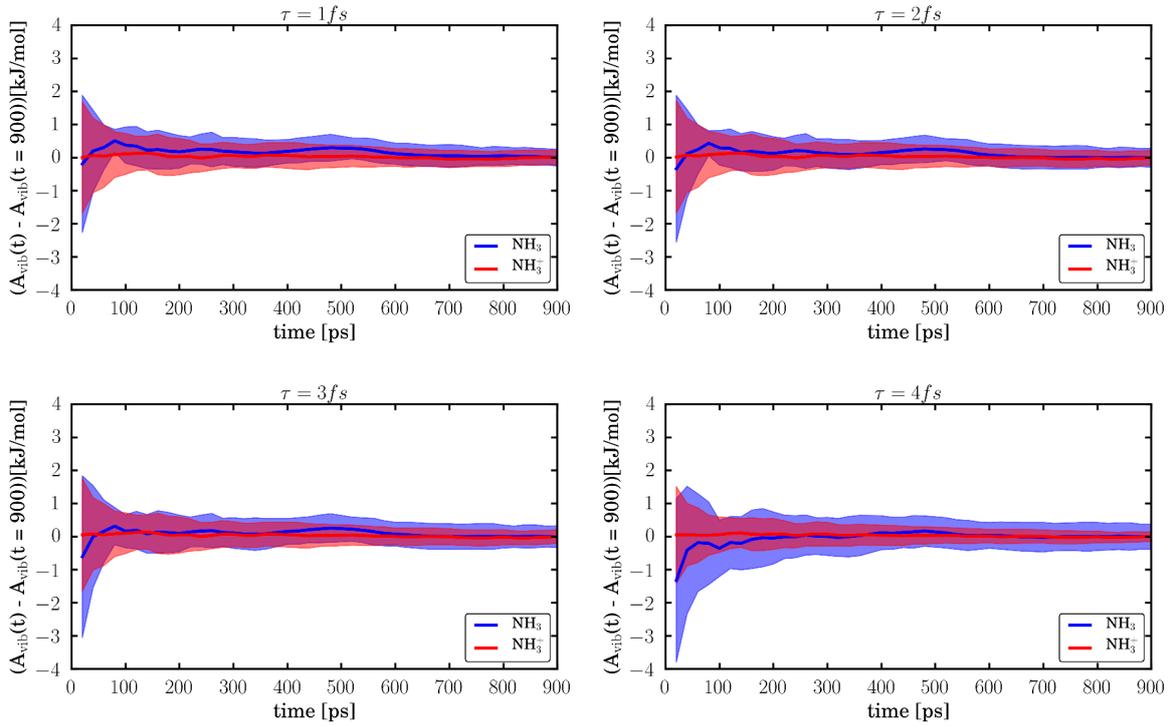


FIG. S7. Convergence of the mean vibrational free energy ($\langle A_{\text{vib}} \rangle$, solid line) and the standard deviation of ten different trajectories (lighter area) of NH_3 and NH_3^+ using DSI and different intervals (τ) between the sampling of the nuclear velocities. The mean value at $t = 900$ ps is set to zero. The mean vibrational free energy converges after a sampling of ≈ 200 ps. Its standard deviation decreases constantly with increasing simulation time. The values of NH_3 show a slower convergence and a larger standard deviation, due to the higher amount of low-frequency modes.

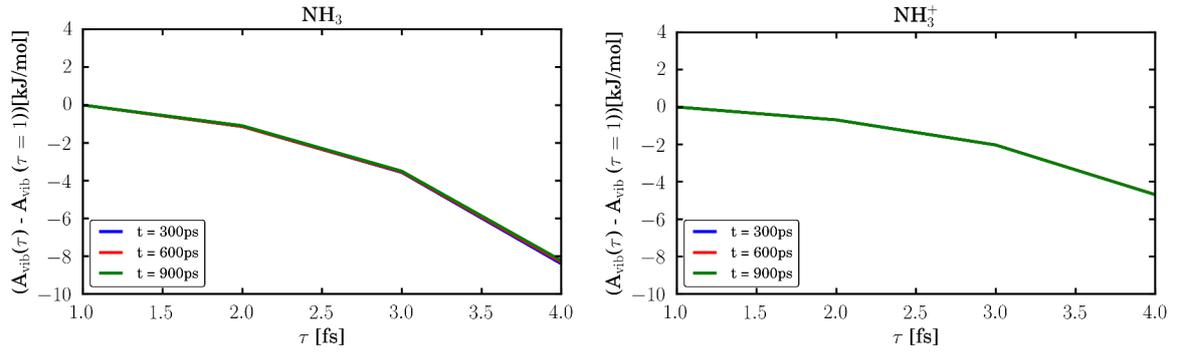


FIG. S8. Change of the mean vibrational free energies with the length of the interval between the sampling of the nuclear velocities (τ) using DSI. Ten independent trajectories of NH_3 and NH_3^+ and three different simulation times (t) are considered. The mean values at $\tau = 1$ fs are set to zero. The sampling rate has a larger impact on the result than the simulation length. An increasing τ decreases the intensity of the N-H bond-stretching modes, leading to changes in the vibrational free energy. The reason for this is that the sampling of these high-frequency modes becomes worse. Applying $\tau > 4$ fs the bond vibrations do not appear in the density of states spectrum. The effect of τ on NH_3 is larger, as the frequency of its bond vibrations is higher.

III. MUTATION FROM SERINE TO CYSTEINE

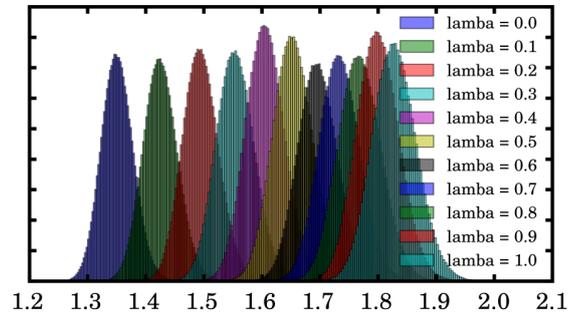


FIG. S9. Histograms of the C-O/S bond length for each λ -window. The bond length histograms are closely connected to the distribution of ΔU in each window.

3.5 Publication V: Identifying free energy hot-spots in molecular transformations

J. C. B. Dietschreit, L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Identifying free energy hot-spots in molecular transformations",
J. Phys. Chem. A, **123**, 2163 (2019)

Abstract: The free energy is one of the central quantities in material and natural sciences. While being well-established, e.g., in drug design or catalyst optimization, computational methods lack a straightforward way to gain deeper insights into the calculated free energy, and thus the underlying chemical or physical processes. Here, we present a generally applicable, spectrum-based ansatz that tackles this shortcoming by identifying contributions from specific atoms or groups to the vibrational free energy. We illustrate this in studies of the bromodomain-inhibitor binding and the anomeric effect in glucose providing quantitative evidence in line with chemical intuition in both cases. For the latter example we also report an experimental infrared spectrum and find excellent agreement with our simulated spectra.

Reprinted with permission from:

J. C. B. Dietschreit, L. D. M. Peters, J. Kussmann, C. Ochsenfeld,
"Identifying free energy hot-spots in molecular transformations",
J. Phys. Chem. A, **123**, 2163 (2019)

Copyright 2019 American Chemical Society.

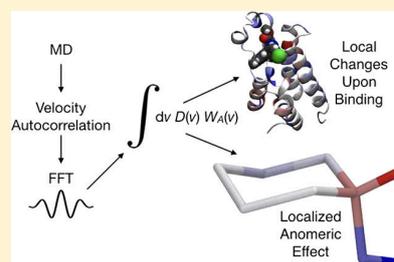
<https://pubs.acs.org/doi/10.1021/acs.jpca.8b12309>

Identifying Free Energy Hot-Spots in Molecular Transformations

Johannes C. B. Dietschreit,^{†,‡,§} Laurens D. M. Peters,^{†,‡,§} Jörg Kussmann,^{†,‡}
and Christian Ochsenfeld^{*,†,‡,§}[†]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany[‡]Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstr. 5–13, D-81377 München, Germany

Supporting Information

ABSTRACT: The free energy is one of the central quantities in material and natural sciences. While being well-established, e.g., in drug design or catalyst optimization, computational methods lack a straightforward way to gain deeper insights into the calculated free energy, and thus the underlying chemical or physical processes. Here, we present a generally applicable, spectrum-based ansatz that tackles this shortcoming by identifying contributions from specific atoms or groups to the vibrational free energy. We illustrate this in studies of the bromodomain-inhibitor binding and the anomeric effect in glucose providing quantitative evidence in line with chemical intuition in both cases. For the latter example we also report an experimental infrared spectrum and find excellent agreement with our simulated spectra.



INTRODUCTION

The free energy is the driving force behind every chemical reaction. It determines, for example, the rate of an enzymatic reaction or the scope of products formed during a catalytic reaction. The prediction of free energies is, therefore, a key challenge in modern quantum chemistry.^{1,2} For small, unimolecular systems this is usually done via a frequency analysis of the molecule using quantum mechanics (QM) calculations. It is assumed that, in the vicinity of the minimum energy geometry, all vibrations can be described as independent harmonic oscillators (harmonic approximation). For larger or multimolecular systems, this approach is not feasible, as the harmonic approximation is not valid anymore (due to the increasing number of anharmonic modes) and the potential energy surface features an enormous number of local minima, which have to be considered. In these cases, one focuses on free energy differences, which can be computed without having to know the absolute energy of both states, and determines the free energy from sampled energies along Monte Carlo^{3,4} or molecular dynamics (MD)^{5–7} simulations applying, e.g., exponential averaging theory⁸ or Bennett's acceptance ratio method.⁹

While the mentioned methods have been used extensively in different fields,^{10–13} the interpretation of their results is in most cases not straightforward. The reason for this is that it is not possible to separate the total free energy change into contributions from different atoms or residues¹⁴ and, therefore, to understand the underlying effects (e.g., bond weakening, steric clashes, new noncovalent interactions) causing the free energy to change. Applying the conventional energy-based methods,^{8,9} some approximate fragmentation is possible for

simple force fields,^{12,15} however, this is not possible when nonadditive force fields (like the emerging polarizable force fields^{16–18}), QM calculations, or combined quantum mechanics/molecular mechanics (QM/MM) are used.

In this work, we use and present a method that calculates the vibrational part of the free energy from the vibrational density of states function, which itself was the topic of experimental¹⁹ and theoretical studies.^{20,21} This approach has originally been introduced by Berens et al.²² to estimate quantum corrections to thermodynamic properties. It has been used occasionally to compute absolute entropies,²³ solvation effects such as entropy,²⁴ or helped identifying different water species around a protein in solution²⁵ by employing the additional assumptions of the two-phase model.²⁶ The calculation of free energy changes in discrete volume units (so-called “voxels”) by Heyden is also based on this approach.²⁷

The applicability of the method of Berens et al.²² to free energy calculations has been determined in a different study of ours,²⁸ where a more detailed derivation, validation, and analysis of the method can be found. Here, we will focus entirely on its capability of calculating atom- or residue-wise contributions to the vibrational free energy and how these free energy hot-spots can help to understand and interpret free energy changes during molecular transformations. We start with a brief summary of the density of states integration method (DSI)^{22,28} in Section 2. There, we will also discuss shortcomings of the method and how they affect the

Received: December 21, 2018

Revised: February 14, 2019

Published: March 1, 2019

applicability and the interpretation of the results of our approach. In Section 3 we list computational (and experimental) details. In Section 4 we apply our method to two prototypical, illustrative examples: (1) The binding of an inhibitor to a bromodomain-containing protein and (2) the visualization of the anomeric effect in glucose. An outlook is given in Section 5.

THEORY

Density of States Integration Method. We extract the free energy from the velocities (\mathbf{v}_j) of each atom j during a molecular dynamics simulation. This is done by calculating the density of states function ($D(\nu)$) as the Fourier transform of the velocity autocorrelation function

$$D(\nu) = 2\beta \sum_{j=1}^N m_j \int dt \langle \mathbf{v}_j(\tau) \mathbf{v}_j(t + \tau) \rangle_{\tau} e^{-i2\pi\nu t} \quad (1)$$

β is equal to $1/(k_B T)$ with k_B being the Boltzmann constant and T the absolute temperature. m_j is the mass of atom j , and N is the total number of nuclei. When neglecting contributions from translation and rotation, the free energy (A) can be calculated from a weighted integral over the frequency (ν)^{22,28}

$$A = E + A_{\text{vib}} = E + \beta^{-1} \int_0^{\infty} d\nu D(\nu) \ln[\beta h \nu] \quad (2)$$

with E being the potential energy at the global minimum energy geometry and h the Planck constant. The vibrational part of the free energy (A_{vib}) is thus calculated as a sum of an arbitrarily large number of harmonic oscillators weighted by $D(\nu)$.

Equations 1 and 2 indicate that A_{vib} can be split into contributions from the individual nuclei or residues, as $D(\nu)$ is calculated as a sum over all atoms. In order to obtain atom- or residue-resolved free energies, we recast eq 1 to

$$\begin{aligned} D(\nu) &= 2\beta \sum_{j=1}^N m_j \int dt \langle \mathbf{v}_j(\tau) \mathbf{v}_j(t + \tau) \rangle_{\tau} e^{-i2\pi\nu t} \\ &= \sum_{j=1}^N 2\beta m_j \int dt \langle \mathbf{v}_j(\tau) \mathbf{v}_j(t + \tau) \rangle_{\tau} e^{-i2\pi\nu t} \\ &= \sum_{j=1}^N D_j(\nu) \\ &= \sum_i^{N_{\text{regions}}} \sum_j^{\{N\}_i} D_j(\nu) \end{aligned} \quad (3)$$

N_{regions} is the number of regions in which we split the total system and $\{N\}_i$ is the set of atoms that belong to the region i . The regions can be chosen completely freely ranging from the entire system over residues to individual atoms. This helps us rewrite eq 2 to

$$A = E + \sum_i^{N_{\text{regions}}} A_{\text{vib}}(i) \quad (4)$$

$A_{\text{vib}}(i)$ is the vibrational free energy localized in region i . If we consider free energy changes

$$\Delta A = \Delta E + \sum_i^{N_{\text{regions}}} \Delta A_{\text{vib}}(i) \quad (5)$$

$\Delta A_{\text{vib}}(i)$ indicates a change in the potential energy surface in region i . Please note that the fragmentation of ΔA_{vib} directly evolves from its calculation from $D(\nu)$ and requires no additional approximations or assumptions.

Interpreting Free Energy Hot-Spots. Two shortcomings have to be considered when interpreting $\Delta A_{\text{vib}}(i)$ for different atoms or residues i . The first one is that we are investigating only the vibrational part of the free energy ($\Delta A_{\text{vib}}(i)$ instead of $\Delta A(i)$) without $\Delta E(i)$ and contributions from rotations and translations. The latter are neglected, as the overall rotation and translation of the system are removed, because keeping them can lead to unwanted artifacts. Therefore, the sum over all hot-spots yields ΔA_{vib} and not the full free energy difference (ΔA). A comparison to other free energy methods is, thus, only meaningful when adding ΔE , which has been done for small systems in ref 28 and is not possible for larger systems (as the exact determination of ΔE becomes impossible). However, A_{vib} identifies regions where the potential energy surface (and thus A) is changing and is, therefore, an excellent tool to find and quantify free energy hot-spots, as shown below.

The second shortcoming is that the Density of States Integration (DSI) uses the harmonic approximation. It was shown in ref 28 that DSI can (in contrast to other free energy methods based on vibrational frequencies) describe the anharmonic behavior of vibrations, as it considers the system as a linear combination of a (nearly) infinite number of harmonic oscillators. Free energy changes arising from vibrations involving movements over local maxima in the potential energy surface and very slow modes (such as rotations of entire protein domains) can only be described qualitatively but not quantitatively, leading to errors in the calculation of A_{vib} .²⁸ However, these errors partly cancel out, because we investigate free energy differences. Moreover, slow modes are normally delocalized over large parts of the systems and thus do not substantially affect the free energy hot-spots and their interpretation, as we focus mainly on local changes.

METHODS

Classical Mechanical Simulations. The crystal structures of the apo-bromodomain (PDB 5O38)²⁹ and the inhibitor-domain complex (PDB 5O3B)²⁹ were used as starting structures. All molecules that were not protein, inhibitor, or water were removed. Antechamber, part of the AmberTools 16,³⁰ was used to parametrize the inhibitor. The force field ff14SB³¹ was used for the simulations. The proteins were solvated in a rectangular box with 10 Å of TIP3P³² water, and neutralized with 2 chlorine ions. The simulation engine NAMD³³ was used. The energy of the system was minimized: For the first 10,000 steps only the water molecules and for the next 10,000 steps the full system. The system was heated over 30 ps to 300 K. In the following it was equilibrated for 200 ps, and then a production run of 1 ns was carried out. The time step was 0.5 fs, as no constraints such as SHAKE³⁴ or RATTLE³⁵ were imposed on the system during the production runs. Nonbonded interactions were evaluated at every step. Periodic electrostatic interactions were computed with the particle mesh Ewald summation method, with a sixth order interpolation. We used a cutoff radius of 12 Å and a switching

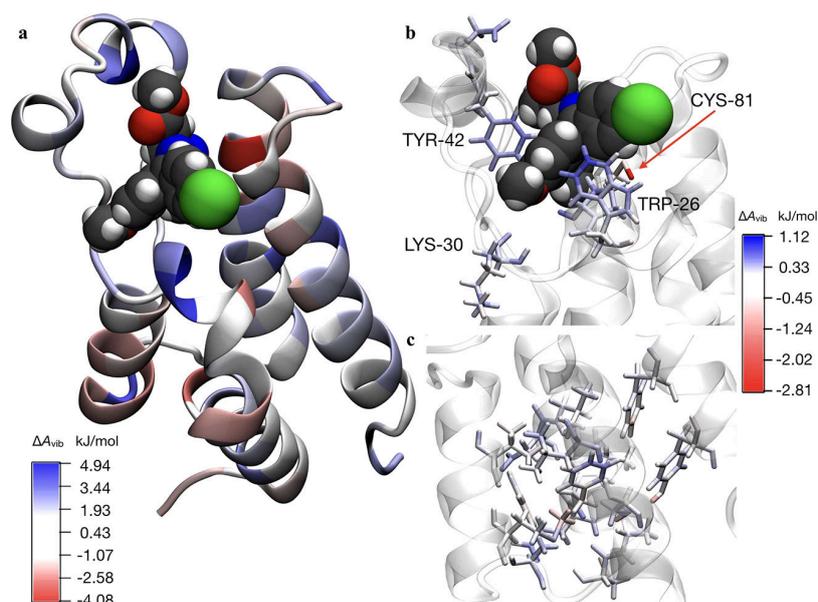


Figure 1. (a) Changes of the vibrational free energy (ΔA_{vib}) within the bromodomain upon binding the inhibitor. The residues are colored according to the changes in A_{vib} going from the apo- to the complexed-form per residue. Blue residues indicate a gain ($\Delta A_{\text{vib}} > 0$, less movement), and red residues indicate a loss in vibrational free energy ($\Delta A_{\text{vib}} < 0$, more freedom). The inhibitor is shown with van-der-Waals-spheres colored according to atom types. (b) Interaction between the residues and the inhibitor. TRP-26 (van-der-Waals interaction), TYR-42 (hydrogen bond), and CYS-81 (two hydrogen bonds) are highlighted. (c) Interactions within the helical part of the domain. Note that the color scales differ between part a and parts b and c of the figure.

function that smoothly switches off interaction between 10 and 12 Å. A Verlet nearest neighbor list with a radius of 13.5 Å was used. The temperature was controlled with the Berendsen rescaling algorithm.³⁶ Translation and rotation of the protein were removed from the velocities after the simulation. *MDAnalysis*^{37,38} was used to extract and process the velocities. The convergence of the free energy difference is shown for residue TRP-26 in Figure S6.

Quantum Mechanical Simulations. The quantum chemistry package *FermiONs++*^{39–41} developed in our group was used for the *ab initio* Born–Oppenheimer molecular dynamics simulations. We used the HF-3c⁴² method that includes dispersion (DFTD3 v3.1)^{43,44} and counterpoise corrections (gCP v2.02).⁴⁵ The Velocity Verlet algorithm^{46,47} and a stochastic rescaling thermostat⁴⁸ were applied. The structures of α -Glu, β -Glu, α -HCT, and β -HCT were minimized at the same level of theory before the calculations. The initial atom velocities were drawn at random from a Maxwell–Boltzmann distribution at 298.15 K. The time step is 0.5 fs, and a ninth order extended Lagrangian scheme⁴⁹ was used to improve the SCF convergence. The system was equilibrated for 5 ps. The production runs were 200 ps long, and 20 independent simulations (different starting velocity vectors and pseudorandom numbers in the thermostat) were conducted for each molecule. Translation and rotation of the molecule were removed at every step of the simulation. As starting points, we used two different minima, both obtained by energy minimizations. Additional conformers or a subsequent weighting of the single simulations were not required as the thermal energy of the molecule and the simulation time were sufficient to explore the conformational space. The sampling was monitored by the convergence of the mean free energy difference; see Figure S5.

Free Energy Calculations. Vibrational free energies (A_{vib}) are calculated from the sampled nuclear velocities applying eqs 1 and 2. All atomic spectra were rescaled such that every atom receives the same fraction of the total amount of degrees of freedom.

Infrared Spectra. The experimental spectrum of D(+)-glucose 1-hydrate (ITW Reagents, > 99%) has been measured in this work as an average of 16 scans with 1 cm^{-1} resolution using a Thermo Fischer Nicolet 6700 FT-IR apparatus.

RESULTS AND DISCUSSION

Inhibitor Bound to the Bromodomain—An MM Application. As a first demonstration of the presented approach, we investigate the change of A_{vib} during the binding of a bromodomain-containing protein to an inhibitor. Proteins of the bromo- and extra-terminal domain (BET) family are involved in the recognition of acetylated lysine residues and play an important role in epigenetic communication.⁵⁰ Very recently, potent mutant-selective inhibitors for BET have been developed,²⁹ which are meant as a tool for future *in vivo* studies. Upon binding to the inhibitor, the potential energy surface of BET is modified leading to conformational changes in the protein which one would generally assume causes the binding site to become tighter. Here, we want to stress that we are focusing on calculating the *changes* of A_{vib} upon binding and do not attempt to compute the binding free energy, for which energy-based methods such as the Bennett’s Acceptance Ratio method⁹ are more suitable. We expect though that the atoms highlighted by our method are those which are the main contributors to the free energy of binding.

We used the cocrystal structure of 9-ME-1 and BET as well as the apo-crystal structure (PDB 5O3C and 5O38²⁹) as a starting point to investigate the effect of inhibitor binding to the bromodomain motif. We conducted two independent

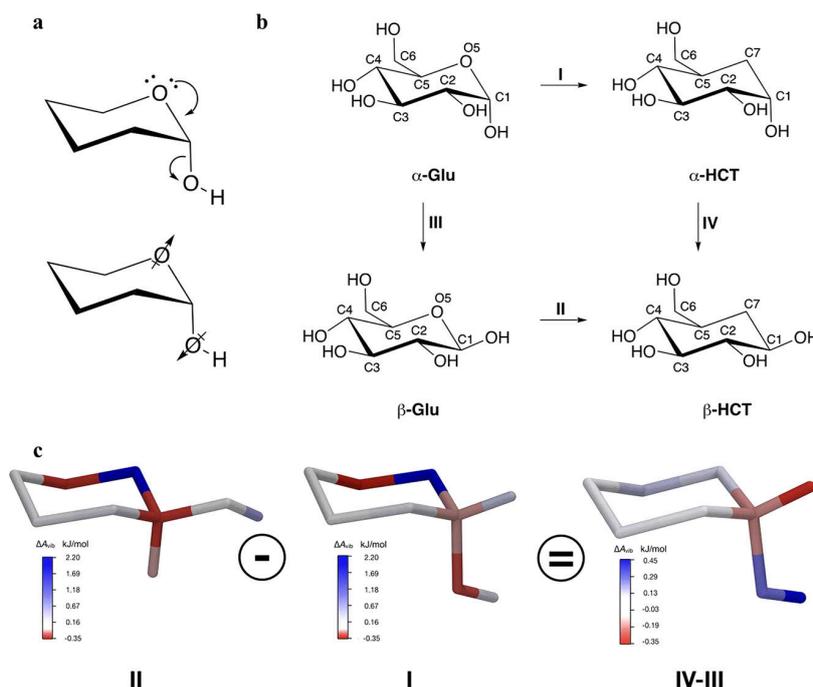


Figure 2. (a) (top) Hyperconjugation and (bottom) dipole interaction discussed as origins of the anomeric effect in glucose.⁵² (b) Structures, abbreviations, and atom labels of the investigated molecules. The possible transformations I to IV are arranged in a thermodynamic cycle. (c) Change in the vibrational free energy per atom from β -Glu \rightarrow β -HCT (transformation II), α -Glu \rightarrow α -HCT (transformation I), and their difference (II-I = IV-III). The latter is equivalent to the change of the free energy during the appearance of the anomeric effect reflecting the bond strengthening of the O5–C1 bond and the bond weakening of the C1–O1 and C5–O5 bond.

classical molecular dynamics simulations in a water box and computed the difference in vibrational free energy (ΔA_{vib}) per residue and per atom, where the residue-wise vibrational free energy is just the sum over the corresponding atoms. The changes in the vibrational free energy are shown in Figure 1. Blue indicates a stiffening of the residue ($\Delta A_{vib} > 0$, shift toward vibrations of higher frequency) upon the binding of the inhibitor, whereas red ($\Delta A_{vib} < 0$, shift toward vibrations of lower frequency) means that the residue can move more freely.

The overall change in ΔA_{vib} is positive, already indicating a stiffening of the motif due to the interaction with the inhibitor. Many of those stiffer residues can be found in the binding pocket (see Figure 1a+b), where their side chains show distinct interactions with the inhibitor. Examples are TRP-26, which interacts with the inhibitor via van-der-Waals interactions, and TYR-42, which forms a hydrogen bond. Interactions between the inhibitor and the backbone are also occurring. The peptide group of LYS-30, for example, communicates with 9-ME-1 via a water molecule. The only residue in the binding pocket, which becomes more flexible, is CYS-81; its SH-group switches between two hydrogen bond acceptors (one at BET and one at the inhibitor). Both hydrogen bonds are not formed in the apoprotein, where CYS-81 is constantly bound to a residue which is inaccessible in the presence of the inhibitor. The analysis, however, also shows that A_{vib} does not only change at the binding site, as parts of the α -helices are colored as well (see Figure 1c). Reasons for these significantly smaller contributions are subtle changes in the arrangement of the helices.

In order to discuss if the values of the free energy hot-spots ($\Delta A_{vib}(i)$) can be interpreted quantitatively (not only qualitatively), one has to consider the individual vibrations

contributing to the free energy hot-spots. The changes of the free energy in the binding pocket and at the helices are dominated by local changes of stretching or bending vibrations, arising from the interactions with the inhibitor or the resulting changes in the arrangement of the helices. These (rather) high-frequency modes should be described quantitatively correct with our method, as already discussed in Section 2 and ref 28. CYS-81 might be an exception as the switching between two hydrogen bond acceptors features a low-frequency mode over a local maximum. It is, however, an excellent example how local effects change ΔA_{vib} dramatically and how this is detected by our method. Of course, we cannot exclude that our method neglects possible low-frequency modes, which affect large parts of the system. Yet, we were not able to find such vibrations in RMSD-plots.

In previous experimental and theoretical studies,^{19,51} a general mode softening (increase of the density of states function for very small wave numbers) was reported for similar protein-inhibitor systems. These new low-frequency modes were identified as linear combinations of the rotational or translational modes of the inhibitor and modes of the protein. In our present study, this effect is not visible. The main reason for this is that all our simulations were carried out including explicit solvent molecules, while previous studies involved gas-phase simulations and measurements of dried samples. As a consequence, we describe the more realistic system and the replacement of solvent by the inhibitor in the binding pocket and not the mere binding event. Since the solvent molecules located in the binding pocket also couple to the protein, no mode softening during its replacement by the inhibitor is observed.

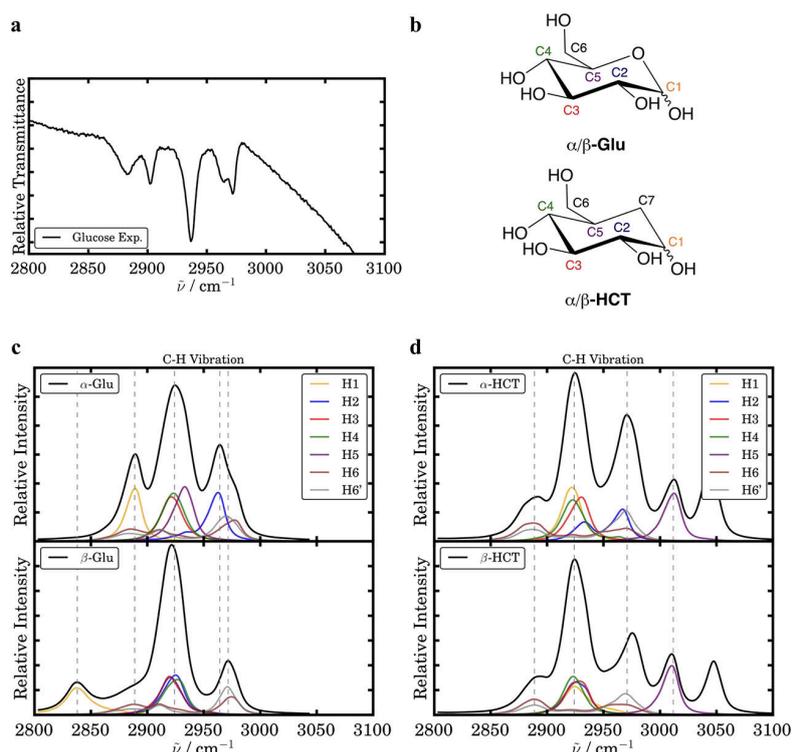


Figure 3. (a) Excerpt of the experimental IR-spectrum of crystalline glucose (monohydrate and mixture of the α - and β -anomer) showing the C–H stretching vibrations for comparison to the simulated spectra below. (b) Labels of carbon atoms in α -Glu, β -Glu, α -HCT, and β -HCT. (c) Calculated spectra ($D(\tilde{\nu})$) of (top) α -Glu and (bottom) β -Glu showing the C–H stretching vibrations of the entire molecule (black) and the contributions from the different C–H bonds (color). The splitting of molecular peaks enables a detailed inspection of the surroundings of the individual atoms. For comparison to the experiment, the frequencies of the simulated spectra have been scaled by a factor of 0.82 (similar to the reported 0.81 in our previous work⁵⁵). (d) Same analysis for (top) α -HCT and (bottom) β -HCT. The peaks of H7 and H7' (around 3050 cm^{-1}) are not shown as they cannot be compared to glucose.

In summary, our approach has revealed the formation of a tight inhibitor–protein complex, which also affects parts of the helices of the protein. All interactions reported here were found simply through the application of the presented method, and no complex analysis of bond distances or dihedral angle distributions and, thus, no *a priori* knowledge about the binding process was necessary to identify any of them. In addition, the power spectra ($D(\nu)$) of the residues or atoms were used to interpret the free energy hot-spots.

Anomeric Effect—A QM Application. In the first example we focused on intermolecular interactions, but our method can also visualize changes of covalent bonds. Therefore, we use, as a second example, *ab initio* molecular dynamics instead of force field calculations to investigate the anomeric effect. The anomeric effect appears in heterocycles based on cyclohexane and leads to a stabilization of the axial position of heteroatomic substituents adjacent to the heteroatom within the six-membered ring.⁵² We investigate this effect by looking at one famous representative, namely glucose. Here, one encounters an unexpected stabilization of α -glucose (α -Glu) with respect to β -glucose (β -Glu). The origin of the anomeric effect has been under discussion for a long time including experimental⁵³ as well as theoretical⁵⁴ contributions (see Figure 2a). The two debated causes are hyperconjugation and dipole interactions, both stereoelectronic effects. In this work, we restrict ourselves to the visualization of the effect and highlight the involved atoms, as

our method does not allow for a distinction between the two models.

The sole comparison of α -Glu and β -Glu would not only incorporate the anomeric but also other effects, for example, changes of the hydrogen bonds or the 1,3-diaxial interactions. To isolate and visualize the anomeric effect, we have simulated α -Glu, β -Glu, and their two analogues with the ring oxygen being replaced by a CH_2 moiety (α -HCT and β -HCT, the abbreviation HCT stands for the IUPAC name 5-Hydroxymethyl-cyclohexane-(1,2,3,4)-tetrole) at the HF-3c⁴² level of theory. Please note that the anomeric effect is only present in α -Glu. It can, therefore, be investigated by analyzing the difference between the transformations α -Glu \rightarrow α -HCT (I) and β -Glu \rightarrow β -HCT (II) or between the transformations α -Glu \rightarrow β -Glu (III) and α -HCT \rightarrow β -HCT (IV), as the other effects cancel out. For structures and atomic labels see Figure 2b.

In order to visualize the anomeric effect, we show the vibrational free energy differences (ΔA_{vib}) of the transformations I and II for selected atoms in Figure 2c (for the corresponding vibrational spectra see Figure S1). Their difference, which can be interpreted as the appearance of the anomeric effect, is also shown. In both transformations (I and II), the centers near the mutation site C7/O5 and (in the case of I) O1 contribute to ΔA_{vib} , whereas C2–4 are not affected by the mutation. Their difference (II – I) reveals that ΔA_{vib} of the anomeric effect is mainly localized at C1, O1, and their hydrogen atoms as well as at C5 and O5. It is in good

agreement with the bond strengthening of the O5–C1 bond (C1 is red) and the bond weakening of the C1–O1 bond (O1 is blue), when the anomeric effect appears. Additionally, the C5–O5 bond is slightly weakened (C5 is blue). The comparably small contribution at O5 is the result of two counteractive effects, the simultaneous strengthening and weakening of its bonds to C1 and C5, respectively. All effects are also visible in the distributions of bond lengths during the simulations (see Figure S2) and can be interpreted quantitatively, as they feature solely changes in (rather) high-frequency modes.

The C–H stretching vibrations of the systems (Figure 3) can also be used to prove the anomeric effect, as they are very good sensors for changes in the surrounding chemical environment. The superposition spectra of α -Glu and β -Glu (black lines in Figure 3c) are in good agreement with the experimental infrared spectrum measured in the present work (Figure 3a) featuring both six peaks. Comparing α -Glu and β -Glu (Figure 3c), we can identify two red-shifts (C1–H1 and C5–H5), which do not appear in the HCT-spectra (Figure 3d). They can, therefore, be assigned to the anomeric effect corroborating the previous result that the anomeric effect affects the vibrations of C1, C5, and O1 as well as the connected hydrogen atoms and not C2–C4 and C6.

Again, our method has discovered all atoms involved in the anomeric effect, verifying the common picture of this complex stereoelectronic effect (see Figure 2a) without any prior knowledge or assumption. A detailed inspection of the spectra ($D(\nu)$) offers even more insights in the vibrational behavior of the investigated molecules, as shown for this specific case.

CONCLUSION

Overall, the use of vibrational spectra calculated from nuclear velocities, can lead to new and valuable insights into molecular transformations. As the two examples have shown, we are able to localize and, therefore, explain free energy changes. The calculation is straightforward and does not require any *a priori* knowledge about the system before the actual evaluation. We have also shown that it is applicable to any level of theory for the molecular dynamics simulations, ranging from force-field to full quantum-mechanical calculations. Our results are in absolute agreement with chemical intuition for which our method provides a solid and generally valid fundament. Although the central quantity of the approach is the vibrational free energy (ΔA_{vib}) and not the total free energy (ΔA), our method allows for a quantification of effects, especially when rather high-frequency vibrations are involved and when one compares different ΔA_{vib} 's of, e.g., different inhibitors or different substituents. We suggest this ansatz to be used, e.g. in drug or catalyst design, in addition to the calculation of energies and the investigation of structural parameters for gaining the complete picture of the problem at hand.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.8b12309.

The Supporting Information includes details about our image processing, atomic VDOS spectra of all heavy atoms in glucose and HCT, distributions of important bond lengths and dihedral angles in glucose and HCT, comparison of experimental and simulated IR-spectra, a

selected region from the VDOS spectrum featuring the anomeric effect, convergence plots for glucose and HCT, and a convergence plot for a selected residue of the bromodomain (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: christian.ochsenfeld@uni-muenchen.de.

ORCID

Christian Ochsenfeld: 0000-0002-4189-6558

Author Contributions

§J.C.B.D. and L.D.M.P. contributed equally to this work

Notes

The authors declare no competing financial interest.

An interactive tutorial with analysis scripts is available at <http://www.cup.lmu.de/pc/ochsenfeld/download/>.

ACKNOWLEDGMENTS

The authors acknowledge Sophia Schwarz and Professor Oliver Trapp (LMU Munich) for their help in measuring the experimental infrared spectrum, and Professor Fritz Schaefer (University of Georgia, Athens, USA) for useful comments on our manuscript. Financial support was provided by the SFB 749 "Dynamics and Intermediates of Molecular Transformations" (DFG), the SFB 1309 "Chemical Biology of Epigenetic Modifications" (DFG), and the DFG cluster of excellence (EXC 114) "Center for Integrative Protein Science Munich" (CIPSM). C.O. acknowledges further support as Max-Planck-Fellow at the MPI-FKF Stuttgart.

REFERENCES

- (1) Chipot, C.; Pohorille, A., Eds. *Free Energy Calculations*; Springer-Verlag: Berlin Heidelberg, 2007.
- (2) Hansen, N.; Van Gunsteren, W. F. Practical aspects of free-energy calculations: A review. *J. Chem. Theory Comput.* **2014**, *10*, 2632–2647.
- (3) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
- (4) Hastings, W. K. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* **1970**, *57*, 97–109.
- (5) Fermi, E.; Pasta, J.; Ulam, S. Studies of nonlinear problems. *Los Alamos Rep. LA-1940*; **1955**, DOI: 10.2172/4376203.
- (6) Alder, B. J.; Wainwright, T. E. Studies in molecular dynamics. I. General method. *J. Chem. Phys.* **1959**, *31*, 459.
- (7) Rahman, A. Correlations in the motion of atoms in liquid argon. *Phys. Rev.* **1964**, *136*, 405–411.
- (8) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- (9) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (10) Rickman, J. M.; LeSar, R. Free-Energy calculations in materials research. *Annu. Rev. Mater. Res.* **2002**, *32*, 195–217.
- (11) Karplus, M.; Petsko, G. A. Molecular dynamics simulations in biology. *Nature* **1990**, *347*, 631–639.
- (12) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science* **1989**, *244*, 1069–1072.
- (13) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. Free energy calculations by computer simulation. *Science* **1987**, *236*, 564–568.
- (14) Smith, P. E.; van Gunsteren, W. When are free energy components meaningful? *J. Phys. Chem.* **1994**, *98*, 13735–13740.

- (15) Irwin, B. W. J.; Huggins, D. J. Estimating atomic contributions to hydration and binding using free energy perturbation. *J. Chem. Theory Comput.* **2018**, *14*, 3218–3227.
- (16) Warshel, A.; Kato, M.; Pliaskov, A. V. Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.
- (17) Lipparini, F.; Barone, V. Polarizable Force Fields and Polarizable Continuum Model: A Fluctuating Charges/PCM Approach. 1. Theory and Implementation. *J. Chem. Theory Comput.* **2011**, *7*, 3711–3724.
- (18) Baker, C. M. Polarizable force fields for molecular dynamics simulations of biomolecules. *WIREs Comput. Mol. Sci.* **2015**, *5*, 241–254.
- (19) Balog, E.; Becker, T.; Oettl, M.; Lechner, R.; Daniel, R.; Finney, J.; Smith, J. C. Direct determination of vibrational density of states change on ligand binding to a protein. *Phys. Rev. Lett.* **2004**, *93*, 028103.
- (20) Zhang, C.; Guidoni, L.; Kühne, T. D. Competing factors on the frequency separation between the OH stretching modes in water. *J. Mol. Liq.* **2015**, *205*, 42–45.
- (21) Martinez, M.; Gaigeot, M.; Borgis, D.; Vuilleumier, R. Extracting effective normal modes from equilibrium dynamics at finite temperature. *J. Chem. Phys.* **2006**, *125*, 144106.
- (22) Berens, P. H.; Mackay, D. H. J.; White, G. M.; Wilson, K. R. Thermodynamics and quantum corrections from molecular dynamics for liquid water. *J. Chem. Phys.* **1983**, *79*, 2375–2389.
- (23) Lin, S.-T.; Maiti, P. K.; Goddard, W. A., III Two-Phase thermodynamic model for efficient and accurate absolute entropy of water from molecular dynamics simulations. *J. Phys. Chem. B* **2010**, *114*, 8191–8198.
- (24) Persson, R. A.; Pattni, V.; Singh, A.; Kast, S. M.; Heyden, M. Signatures of solvation thermodynamics in spectra of intermolecular vibrations. *J. Chem. Theory Comput.* **2017**, *13*, 4467–4481.
- (25) Pattni, V.; Vasilevskaya, T.; Thiel, W.; Heyden, M. Distinct protein hydration water species defined by spatially resolved spectra of intermolecular vibrations. *J. Phys. Chem. B* **2017**, *121*, 7431–7442.
- (26) Lin, S. T.; Blanco, M.; Goddard, W. A., III The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of Lennard-Jones fluids. *J. Chem. Phys.* **2003**, *119*, 11792–11805.
- (27) Heyden, M. Disassembling solvation free energies into local contributions - Toward a microscopic understanding of solvation processes. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, No. e1390.
- (28) Peters, L. D. M.; Dietschreit, J. C. B.; Kussmann, J.; Ochsenfeld, C. Calculating free energies from the vibrational density of states function: Validation and critical assessment. *J. Chem. Phys.* **2018**, submitted.
- (29) Runcie, A. C.; Zengerle, M.; Chan, K.-H.; Testa, A.; van Beurden, L.; Baud, M. G. J.; Epemolu, O.; Ellis, L. C. J.; Read, K. D.; Coulthard, V.; et al. Optimization of a 'bump-and-hole' approach to allele-selective BET bromodomain inhibition. *Chem. Sci.* **2018**, *9*, 2452–2468.
- (30) Case, D.; Cerutti, D.; Cheatham, T.; Darden, T.; Duke, R.; Giese, T.; Gohlke, H.; Goetz, A.; Greene, D.; Homeyer, N.; et al. *AMBER 2017*; 2017.
- (31) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (35) Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.
- (36) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (37) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Dotson, D. L.; Domanski, J.; Buchoux, S.; Kenney, I. M.; et al. In *Proceedings of the 15th Python in Science Conference*; Benthall, S., Rostrup, S., Eds.; 2016; *Chapter MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations*, pp 102–109.
- (38) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (39) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.
- (40) Kussmann, J.; Ochsenfeld, C. Preselective screening for linear-scaling exact exchange-gradient calculations for graphics processing units and general strong-scaling massively parallel calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
- (41) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU Integral Engine for Strong-Scaling Ab Initio Methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
- (42) Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- (43) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (44) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (45) Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **2012**, *136*, 154101.
- (46) Verlet, L. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **1967**, *159*, 98–103.
- (47) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.
- (48) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (49) Niklasson, A. M. N.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C. J.; Holmström, E.; Zheng, G.; Weber, V. Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009**, *130*, 214109.
- (50) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J.-P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **2012**, *149*, 214–231.
- (51) Moritsugu, K.; Njunda, B. M.; Smith, J. C. Theory and Normal-Mode analysis of change in protein vibrational dynamics on ligand binding. *J. Phys. Chem. B* **2010**, *114*, 1479–1485.
- (52) Filloux, C. M. The problem of origins and origins of the problem: Influence of language on studies concerning the anomeric effect. *Angew. Chem., Int. Ed.* **2015**, *54*, 8880–8894.
- (53) Cocinero, E. J.; Carcabal, P.; Vaden, T. D.; Simons, J. P.; Davis, B. G. Sensing the anomeric effect in a solvent-free environment. *Nature* **2011**, *469*, 76–80.
- (54) Mo, Y. Computational evidence that hyperconjugative interactions are not responsible for the anomeric effect. *Nat. Chem.* **2010**, *2*, 666–671.

(55) Peters, L. D. M.; Kussmann, J.; Ochsenfeld, C. Efficient and accurate Born-Oppenheimer molecular dynamics for large molecular systems. *J. Chem. Theory Comput.* **2017**, *13*, 5479–5485.

Supporting Information: Identifying Free Energy Hot-Spots in Molecular Transformations

Johannes C. B. Dietschreit,^{1,2} Laurens D. M. Peters,^{1,2}
Jörg Kussmann^{1,2}, Christian Ochsenfeld^{1,2}

¹Center for Integrated Protein Science (CIPSM) at the Department of Chemistry,
University of Munich (LMU), Butenandtstr. 5–13, D-81377 München, Germany

²Chair of Theoretical Chemistry, Department of Chemistry,

University of Munich (LMU), Butenandtstr. 7, D-81377 München, Germany

Contents

1 Image Processing	S2
2 Data and Materials Availability	S2
3 Figures	S3
References	S8

1 Image Processing

All images containing molecular geometries which are coloured according to changes in vibrational free energy were produced using *VMD*.^[1] All plots showing spectra and distributions were produced using the python-package *matplotlib*.^[2] The chemical structures were drawn with *ChemDraw*.

2 Data and Materials Availability

All inputs and trajectories are available upon request. PDB files (with the free energy colouring), analysis scripts, and an interactive tutorial are available at <http://www.cup.lmu.de/pc/ochsenfeld/download/>. *NAMD* is freely available for non-commercial users, while *FermiONS++* is not yet available.

3 Figures

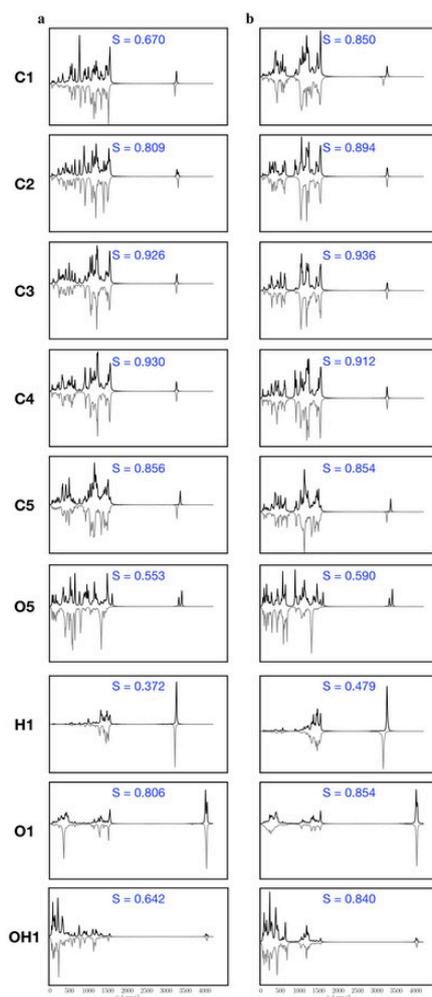


Figure S 1: a+b, Vibrational spectra per atom for (a) α -HCT (black) and α -Glu (inverted grey) and (b) β -HCT (black) and β -Glu (inverted grey). The overlap (S) between the two spectra is calculated as $S = \int_0^\infty I_1(\tilde{\nu})I_2(\tilde{\nu})d\tilde{\nu} / \sqrt{\int_0^\infty I_1^2(\tilde{\nu})d\tilde{\nu} \int_0^\infty I_2^2(\tilde{\nu})d\tilde{\nu}}$, with I_1 and I_2 being the intensity of the two spectra at a wavenumber ($\tilde{\nu}$). The overlap between α -Glu and α -HCT is generally smaller than the overlap between β -Glu and β -HCT. The wave number ($\tilde{\nu}$) increases from left to right.

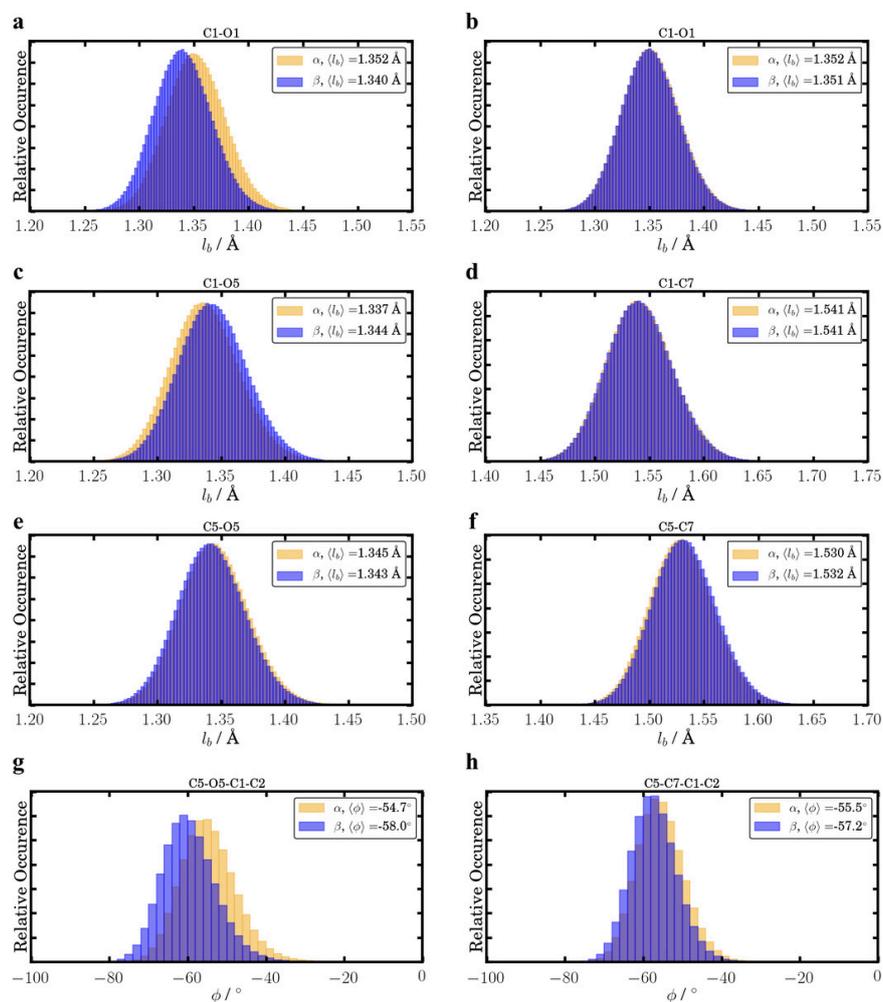


Figure S 2: Distribution of bond lengths and angles of the simulations of (left column) α -Glu and β -Glu and (right column) α -HCT and β -HCT, that serve as an indicator for the anomeric effect. α -Glu exhibits, in comparison to β -Glu, a shorter C1-O5 bond, a longer C1-O1 and C5-O5 bond, and a smaller C5-O5-C1-C2 dihedral angle. Similar observations cannot be made, when comparing the simulations of α -HCT and β -HCT.

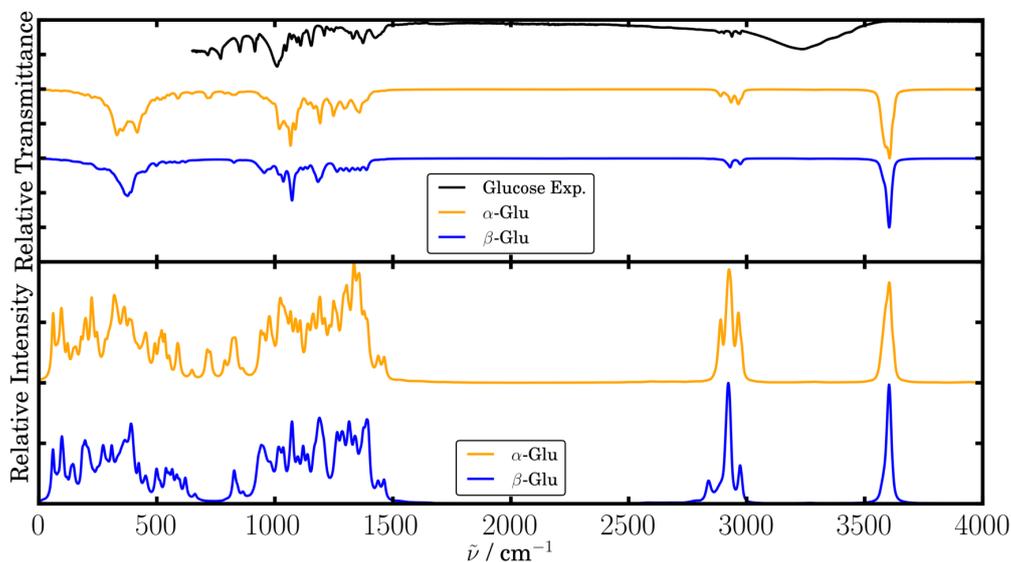


Figure S 3: (Top) Experimental IR spectrum of crystalline glucose-mono-hydrate (black) and simulated IR spectra (calculated as presented in ref. [3]) of α -Glu (orange) and β -Glu (blue). (Bottom) Computed vibrational power spectra of α -Glu (orange) and β -Glu (blue). The power spectrum has different intensities and also shows not IR-active vibrations exhibiting a small or no change in the dipole moment. For comparison to the experiment, the frequencies of the simulated spectra have been scaled by a factor of 0.82 (similar to the reported 0.81 in ref. [3]).

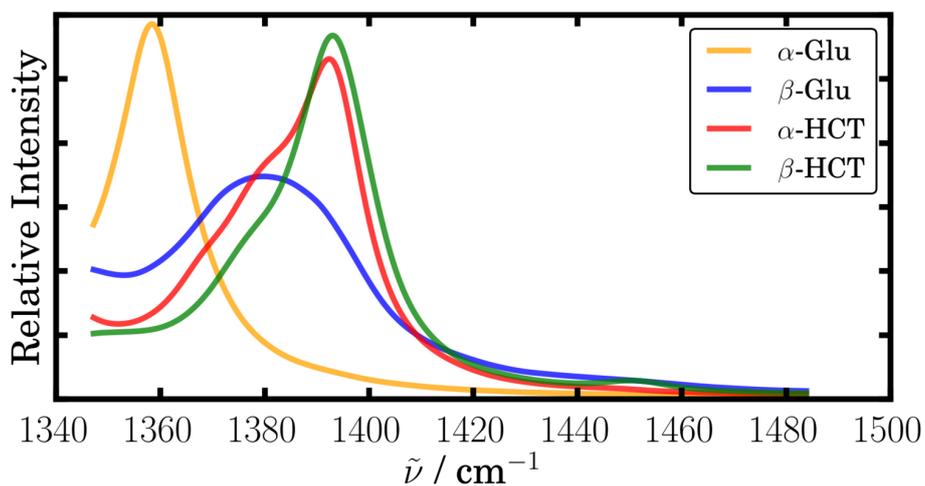


Figure S 4: Other vibrations than the C-H stretching bond are affected by the anomeric effect, e.g., the C-H deformation mode, which is clearly shifted in the case of α -Glu. For comparison to the experiment, the frequencies of the simulated spectra have been scaled by a factor of 0.82 (similar to the reported 0.81 in ref. [3]).

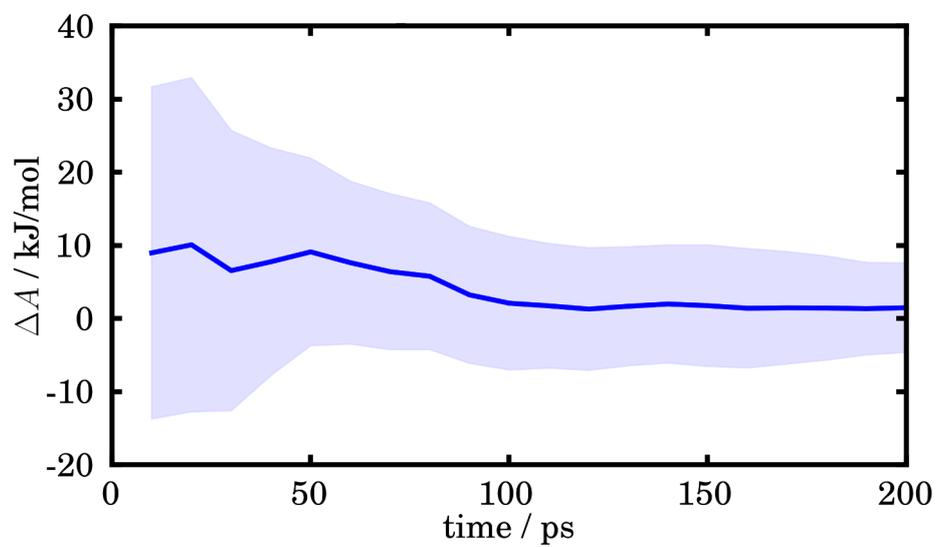
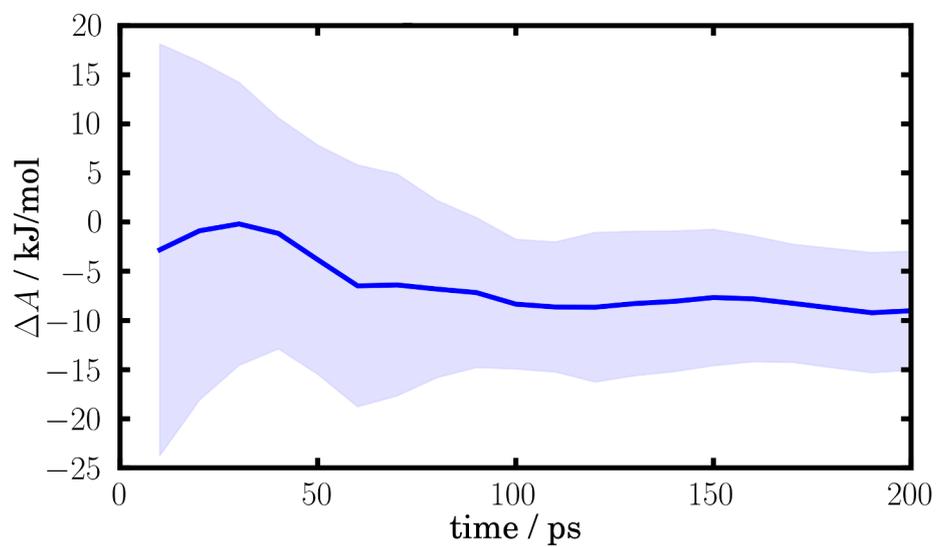


Figure S 5: Convergence of ΔA_{vib} for $\alpha \rightarrow \beta$ glucose (top) and HCT (bottom).

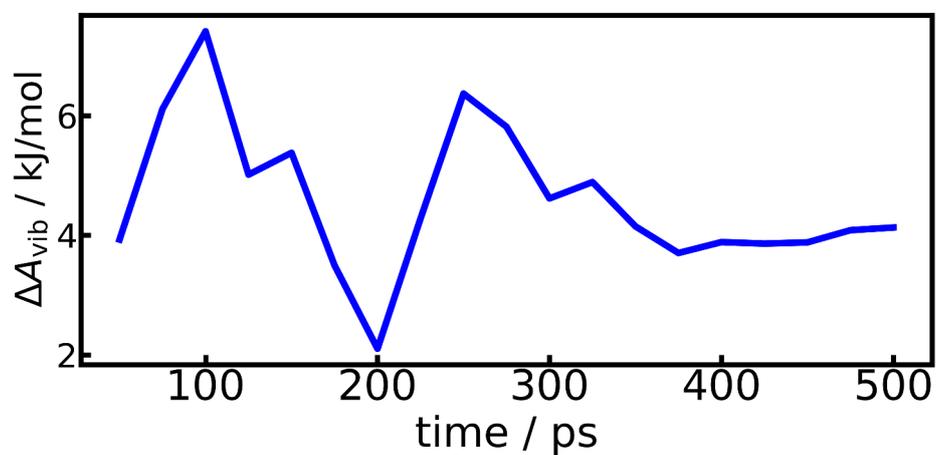


Figure S 6: Convergence of ΔA_{vib} for the Bromodomain shown for residue TRP-26, as it showed a large change in the free energy.

References

- [1] W. Humphrey, A. Dalke, K. Schulten, *Journal of Molecular Graphics* **1996**, *14*, 33–38.
- [2] J. D. Hunter, *Computing In Science & Engineering* **2007**, *9*, 90–95.
- [3] L. D. M. Peters, J. Kussmann, C. Ochsenfeld, *J. Chem. Theory Comput.* **2017**, *13*, 5479–5485.

Chapter 4

Conclusions and Outlook

This thesis has investigated ways to accelerate AIMD calculations, overcoming different bottlenecks of these time-consuming calculations, and pushing their applicability to larger system sizes. The efficiency and accuracy of the presented routines have been demonstrated in ground-state BOMD and excited-state NAMD simulations. The resulting trajectories offer novel insights into a variety of interesting properties, such as free energies, vibrational spectra, and relaxation pathways.

The main bottleneck of AIMD is the evaluation of the two-electron integrals, which needs to be done at each time step. Their extremely efficient calculation on GPUs is, therefore, the first and crucial step towards fast MD simulations. Consequently, this thesis expands the scope of FermiONS++ and its GPU-based integral routines, giving access to ground-state MD, excited-state properties, and consequently NAMD. The resulting speed-ups of the calculations are remarkable, albeit not for the moment enough to address chemically relevant systems and problems. For this purpose, fast, yet still accurate electronic structure methods are required. HF-3c and simplified TDDFT have been tested and proven to be reliable for ground-state and excited-state calculations in this thesis. Additionally, the extended Lagrangian method was used to accelerate the convergence (or circumvent) the SCF cycle of the electronic structure method, while the Hammes–Schiffer–Tully model served to efficiently calculate the couplings between electronic states.

With efficient BOMD and NAMD methods at hand, the thesis presents and discusses several state-of-the-art applications. Accurate IR spectra of large systems (e.g., β -carotene) or even liquid water can be computed within several days (instead of weeks), thereby showing good agreement with experimental data. Dynamics of excited rotary molecular motors and the chromophore of the rhodopsin protein reveal the expected rotation around a double bond accompanying the S_1 - S_0 transition. Here, the speed-up of (at least) a factor of two allows for more and/or longer trajectories, and thus better sampling of the investigated process. The efficient BOMD method is also tested for free energy calculations, using the standard energy-based methods and a reintroduced method based on the vibrational density of states function. The latter gives comparable results to the established methods, while additionally allowing for a partitioning of the free energy. This approach is suitable to identify so-called free energy hot-spots, which could be firmly demonstrated through the investigation of the anomeric effect and protein-inhibitor binding.

Many future projects can be thought of, but the incorporation of the extended tight-binding method GFN(2)-xTB [143, 144], machine learning algorithms [145], or the expansion of the NAMD algorithm towards decoherence corrections [146] or triplet states [147] would be of particular interest. Exploring the introduced approaches together with increasing computing capacities will open up interesting and novel ways for investigating, e.g., biomolecules, photo-switches, and prebiotic chemistry. Along with free energy predictions and analysis (using the free energy hot-spots), simulations of spectra, and investigations of excited-processes, nanoreactors [37, 38] are expected to provide new and more in-depth insights into the fascinating chemistry of these and many other fields.

Bibliography

- [1] M. Born, R. Oppenheimer, *Annalen der Physik* **1927**, *84*, 457.
- [2] D. A. Case, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, AMBER 2017, University of California, San Francisco, **2017**.
- [3] K. A. Beauchamp, Y. S. Lin, R. Das, V. S. Pande, *J. Chem. Theory Comput.* **2012**, *8*, 1409–1414.
- [4] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, A. C. T. van Duin, *npj Comput. Mater.* **2016**, *2*, 15011.
- [5] D. Marx, J. Hutter in *Modern Methods and Algorithms of Quantum Chemistry*, (Ed.: J. Grotendorst), NIC Series, Jülich, **2000**, pp. 301–449.
- [6] B. Kirchner, P. J. di Dio, J. Hutter, *Top. Curr. Chem.* **2012**, *307*, 109–154.
- [7] D. R. Hartree, *Math. Proc. Camb. Philos. Soc.* **1928**, *24*, 89–110.
- [8] V. Fock, *Z. Phys.* **1930**, *61*, 126–148.
- [9] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- [10] G. G. Hall, J. E. Lennard-Jones, *Proc. Royal Soc. A* **1951**, *205*, 541–552.
- [11] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–871.
- [12] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *137*, A1697–1705.
- [13] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, Oxford, **1989**.
- [14] E. Schrödinger, *Phys. Rev.* **1926**, *28*, 1049–1070.
- [15] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618–622.
- [16] J. Čížek, *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- [17] J. Čížek, J. Paldus, Šroubková, *Int. J. Quantum Chem.* **1969**, *3*, 149–167.
- [18] N. J. Turro, *Modern Molecular Photochemistry*, University Science Books, Sausalito, CA, **1991**.
- [19] L. S. Cederbaum in *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*, (Eds.: W. Domcke, D. R. Yarkony, H. Köppel), World Scientific, Singapore, **2004**, pp. 3–40.

- [20] P. Avouris, W. M. Gelbart, M. A. El-Sayed, *Chem. Rev.* **1977**, *77*, 793–833.
- [21] X. Li, J. C. Tully, H. B. Schlegel, M. J. Frisch, *J. Chem. Phys.* **2005**, *123*, 84106.
- [22] B. F. E. Curchod, A. Sisto, T. J. Martínez, *J. Phys. Chem. A* **2017**, *121*, 265.
- [23] J. W. Snyder, B. F. E. Curchod, T. J. Martínez, *J. Phys. Chem. Lett.* **2016**, *7*, 2444.
- [24] J. C. Tully, *J. Chem. Phys.* **1990**, *93*, 1061–1071.
- [25] S. Hammes-Schiffer, J. C. Tully, *J. Chem. Phys.* **1994**, *101*, 4657–4667.
- [26] L. Wang, A. Akimov, O. V. Prezhdo, *J. Phys. Chem. Lett.* **2016**, *7*, 2100.
- [27] R. Crespo-Otero, M. Barbatti, *Chem. Rev.* **2018**, *118*, 7026.
- [28] M. E. Casida in *Recent advances in density functional methods, part 1*. (Ed.: D. P. Chong), World Scientific, **1995**, pp. 155–192.
- [29] M. E. Casida, *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 3–18.
- [30] B. Roos, *Chem. Phys. Lett.* **1972**, *15*, 153–159.
- [31] *Free Energy Calculations*, (Eds.: C. Chipot, A. Pohorille), Springer-Verlag, Berlin Heidelberg, **2007**.
- [32] R. P. Futrelle, D. J. McGinty, *Chem. Phys. Lett.* **1971**, *12*, 285–287.
- [33] M.-P. Gaigeot, M. Martinez, R. Vuilleumier, *Mol. Phys.* **2007**, *105*, 2857.
- [34] M. Thomas, M. Brehm, R. Fligg, P. Vöhringer, B. Kirchner, *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608–6622.
- [35] M. Dračinský, P. Bouř, *J. Comput. Chem.* **2012**, *33*, 1080–1089.
- [36] A. Massolle, T. Dresselhaus, S. Eusterwiemann, C. Doerenkamp, H. Eckert, A. Studer, J. Neugebauer, *Phys. Chem. Chem. Phys.* **2018**, *20*, 7661–7675.
- [37] L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande, T. J. Martínez, *Nat. Chem.* **2014**, *6*, 1044–1048.
- [38] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847–2862.
- [39] E. Tapavicza, I. Tavernelli, U. Rothlisberger, *Phys. Rev. Lett.* **2007**, *98*, 23001.
- [40] M. Barbatti, G. Granucci, M. Persico, M. Ruckebauer, M. Vazdar, M. Eckert-Maksić, H. Lischka, *J. Photochem. Photobiol. A* **2007**, *190*, 228–240.
- [41] J. Pittner, H. Lischka, M. Barbatti, *Chem. Phys.* **2009**, *356*, 147.
- [42] M. Barbatti, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 620–633.
- [43] J. C. Tully, *J. Chem. Phys.* **2012**, *137*, 22A301.
- [44] E. Tapavicza, G. D. Bellchambers, J. C. Vincent, F. Furche, *Phys. Chem. Chem. Phys.* **2013**, *15*, 18336–18348.
- [45] F. Plasser, R. Crespo-Otero, M. Pederzoli, J. Pittner, H. Lischka, M. Barbatti, *J. Chem. Theory Comput.* **2014**, *10*, 1395–1405.
- [46] A. Nikiforov, J. A. Gamez, W. Thiel, M. Filatov, *J. Phys. Chem. Lett.* **2016**, *7*, 105–110.

- [47] A. Kazaryan, Z. Lan, L. V. Schäfer, W. Thiel, M. Filatov, *J. Chem. Theory Comput.* **2011**, *7*, 2189–2199.
- [48] B. Oruganti, J. Wang, B. Durbeej, *Int. J. Quantum Chem.* **2018**, *118*, e25405.
- [49] J. Wang, B. Oruganti, B. Durbeej, *Phys. Chem. Chem. Phys.* **2017**, *19*, 6952–6956.
- [50] R. Sure, S. Grimme, *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- [51] S. Grimme, *J. Chem. Phys.* **2013**, *138*, 244104.
- [52] C. Bannwarth, S. Grimme, *Comput. Theor. Chem.* **2014**, *1040-1041*, 45–53.
- [53] A. M. N. Niklasson, P. Steneteg, A. Odell, N. Bock, M. Challacombe, C. J. Tymczak, E. Holmström, G. Zheng, V. Weber, *J. Chem. Phys.* **2009**, *130*, 214109.
- [54] J. Kussmann, C. Ochsenfeld, *J. Chem. Phys.* **2013**, *138*, 134114.
- [55] J. Kussmann, C. Ochsenfeld, *J. Chem. Theory Comput.* **2015**, *11*, 918–922.
- [56] J. Kussmann, C. Ochsenfeld, *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.
- [57] K. Yasuda, *J. Comput. Chem.* **2008**, *29*, 334–342.
- [58] I. S. Ufimtsev, T. J. Martínez, *J. Chem. Theory Comput.* **2008**, *4*, 222–231.
- [59] I. S. Ufimtsev, T. J. Martínez, *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015.
- [60] I. S. Ufimtsev, T. J. Martínez, *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.
- [61] C. M. Isborn, N. Luehr, I. S. Ufimtsev, T. J. Martínez, *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.
- [62] A. Sisto, D. R. Glowacki, T. J. Martínez, *Acc. Chem. Res.* **2014**, *47*, 2857–2866.
- [63] E. G. Hohenstein, N. Luehr, I. S. Ufimtsev, T. J. Martínez, *J. Chem. Phys.* **2015**, *142*, 224103.
- [64] J. W. Snyder, E. G. Hohenstein, N. Luehr, T. J. Martínez, *J. Chem. Phys.* **2015**, *143*, 154107.
- [65] P. H. Berens, D. H. J. Mackay, G. M. White, K. R. Wilson, *J. Chem. Phys.* **1983**, *79*, 2375–2389.
- [66] R. W. Zwanzig, *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- [67] C. H. Bennett, *J. Comput. Phys.* **1976**, *22*, 245–268.
- [68] M. R. Shirts, E. Bair, G. Hooker, V. S. Pande, *Phys. Rev. Lett.* **2003**, *91*, 140601.
- [69] D. J. Tannor, *Introductions to Quantum Mechanics A Time-Dependent Perspective*, University Science Books, Sausalito, CA, **2007**.
- [70] P. Ehrenfest, *Z. Phys.* **1927**, *45*, 455–457.
- [71] L. Verlet, *Phys. Rev.* **1967**, *159*, 98–103.
- [72] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *J. Chem. Phys.* **1982**, *76*, 637–649.
- [73] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.

- [74] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 14101.
- [75] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- [76] J. C. Slater, *Phys. Rev.* **1930**, *35*, 210.
- [77] A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications Inc., New York, **1996**.
- [78] D. Sherrill, H. F. Schaefer, *Advances in Quantum Chemistry* **1999**, *34*, 143–269.
- [79] J. P. Perdew, K. Schmidt, *AIP Conf. Proc.* **2001**, *577*, 1–29.
- [80] M. J. Stott, E. Zaremba, *Phys. Rev. A* **1980**, *21*, 2293.
- [81] G. D. Mahan, *Phys. Rev. A* **1980**, *22*, 1790.
- [82] E. Runge, E. K. U. Gross, *Phys. Chem. Lett.* **1984**, *52*, 997.
- [83] A. Zangwill, P. Soven, *Phys. Rev. A* **1980**, *21*, 1561.
- [84] R. Bauernschmitt, R. Ahlrichs, *Chem. Phys. Lett.* **1996**, *256*, 454–464.
- [85] R. Bauernschmitt, M. Häser, O. Treutler, R. Ahlrichs, *Chem. Phys. Lett.* **1997**, *264*, 573–578.
- [86] E. R. Davidson, *J. Comp. Phys.* **1975**, *17*, 87–94.
- [87] D. Maurice, M. Head-Gordon, *Int. J. Quantum Chem.* **1995**, *56*, 361–370.
- [88] A. Dreuw, M. Head-Gordon, *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.
- [89] A. Dreuw, M. Head-Gordon, *Chem. Rev.* **2005**, *105*, 4009–4037.
- [90] A. D. Laurent, D. Jacquemin, *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.
- [91] J. D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 84106.
- [92] F. Cordova, L. J. Doriol, M. E. Casida, *J. Chem. Phys.* **2007**, *127*, 164111.
- [93] F. Furche, R. Ahlrichs, *J. Chem. Phys.* **2002**, *117*, 7433–7447.
- [94] R. Send, F. Furche, *J. Chem. Phys.* **2010**, *132*, 44107.
- [95] Z. Li, W. Liu, *J. Chem. Phys.* **2014**, *141*, 14110.
- [96] Z. Li, B. Suo, W. Liu, *J. Chem. Phys.* **2014**, *141*, 244105.
- [97] S. Fatehi, E. Alguire, Y. Shao, J. E. Subotnik, *J. Chem. Phys.* **2011**, *135*, 234105.
- [98] Q. Ou, S. Fatehi, E. Alguire, Y. Shao, J. E. Subotnik, *J. Chem. Phys.* **2014**, *141*, 24114.
- [99] Q. Ou, E. C. Alguire, J. E. Subotnik, *J. Phys. Chem. B* **2015**, *119*, 7150–7161.
- [100] E. C. Alguire, Q. Ou, J. E. Subotnik, *J. Phys. Chem. B* **2015**, *119*, 7140–7149.
- [101] X. Zhang, J. M. Herbert, *J. Chem. Phys.* **2014**, *141*, 64104.
- [102] Q. Ou, G. D. Bellchambers, F. Furche, J. E. Subotnik, *J. Chem. Phys.* **2015**, *142*, 64114.
- [103] X. Zhang, J. M. Herbert, *J. Chem. Phys.* **2015**, *142*, 64109.

- [104] S. M. Parker, S. Roy, F. Furche, *J. Chem. Phys.* **2016**, *145*, 134105.
- [105] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
- [106] *Computer Simulations in Condensed Matter: From Materials to Chemical Biology. Volume 1*, (Eds.: M. Ferrario, G. Ciccotti, K. Binder), Springer-Verlag, Berlin Heidelberg, **2006**.
- [107] A. M. N. Niklasson, C. J. Tymczak, M. Challacombe, *J. Chem. Phys.* **2007**, *126*, 144103.
- [108] A. M. N. Niklasson, C. J. Tymczak, M. Challacombe, *Phys. Rev. Lett.* **2006**, *97*, 123001.
- [109] A. M. N. Niklasson, *Phys. Rev. Lett.* **2008**, *100*, 123004.
- [110] M. J. Cawkwell, A. M. N. Niklasson, *J. Chem. Phys.* **2012**, *137*, 134105.
- [111] P. Souvatzis, A. M. N. Niklasson, *J. Chem. Phys.* **2013**, *139*, 214102.
- [112] A. Odell, A. Delin, B. Johansson, N. Bock, M. Challacombe, A. M. N. Niklasson, *J. Chem. Phys.* **2009**, *131*, 244106.
- [113] E. R. Davidson, D. Feller, *Chem. Rev.* **1986**, *86*, 681.
- [114] W. Kolos, *Theor. Chim. Acta* **1979**, *51*, 219.
- [115] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [116] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [117] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.
- [118] H. Kruse, S. Grimme, *J. Chem. Phys.* **2012**, *136*, 154101.
- [119] P.-O. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365.
- [120] T. Risthaus, A. Hansen, S. Grimme, *Phys. Chem. Chem. Phys.* **2014**, *16*, 14408–14419.
- [121] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2018**, *149*, 24108.
- [122] G. R. Ahmadi, J. Almlöf, *Chem. Phys. Lett.* **1995**, *246*, 364–370.
- [123] C. White, M. Head-Gordon, *J. Chem. Phys.* **1996**, *104*, 2620–2629.
- [124] E. B. Wilson, *Molecular vibrations : the theory of infrared and Raman vibrational spectra*, McGraw-Hill, New York, **1955**.
- [125] K. Wendler, M. Brehm, F. Malberg, B. Kirchner, L. Delle Site, *J. Chem. Theory Comput.* **2012**, *8*, 1570–1579.
- [126] T. Hrenar, H.-J. Werner, G. Rauhut, *J. Chem. Phys.* **2007**, *126*, 134108.
- [127] J. Neugebauer, B. A. Hess, *J. Chem. Phys.* **2003**, *118*, 7215–7225.
- [128] M. Kumarasiri, C. Swalina, S. Hammes-Schiffer, *J. Phys. Chem. B* **2007**, *111*, 4653–4658.
- [129] M. Thomas, M. Brehm, B. Kirchner, *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207–3213.
- [130] P. Kollman, *Chem. Rev.* **1993**, *93*, 2395–2417.
- [131] C. D. Christ, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2010**, *31*, 1569–1582.

- [132] M. A. Olsson, U. Ryde, *J. Chem. Theory Comput.* **2017**, *13*, 2245–2253.
- [133] U. Ryde, P. Söderhjelm, *Chem. Rev.* **2016**, *116*, 5520–5566.
- [134] O. K. Dudko, G. Hummer, A. Szabo, *Phys. Rev. Lett.* **2006**, *96*, 1–4.
- [135] E. Neria, S. Fischer, M. Karplus, *J. Chem. Phys.* **1996**, *105*, 1902–1921.
- [136] D. A. McQuarrie, S. D. Simon, *Molecular Thermodynamics*, University Science Books, Sausalito, CA, **1999**.
- [137] B. Brooks, M. Karplus, *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*, 6571–6575.
- [138] B. Tidor, M. Karplus, *J. Mol. Biol.* **1994**, *238*, 405–414.
- [139] D. M. Zuckerman, T. B. Woolf, *Phys. Rev. Lett.* **2002**, *89*, 180602.
- [140] J. Gao, K. Kuczera, B. Tidor, M. Karplus, *Science* **1989**, *244*, 1069–1072.
- [141] P. E. Smith, W. van Gunsteren, *J. Phys. Chem.* **1994**, *98*, 13735–13740.
- [142] B. W. J. Irwin, D. J. Huggins, *J. Chem. Theory Comput.* **2018**, *14*, 3218–3227.
- [143] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
- [144] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- [145] S. Chmiela, H. E. Saucedo, K.-R. Müller, A. Tkatchenko, *Nat. Commun.* **2018**, *9*, 3997.
- [146] H. M. Jaeger, S. Fischer, O. V. Prezhdo, *J. Chem. Phys.* **2012**, *137*, 22A545.
- [147] M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, González, *J. Chem. Theory Comput.* **2011**, *7*, 1253.