THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Scattering off molecules far from equilibrium

OPEN ACCESS

Scattering off Molecules far from Equilibrium

Haiwang Yong[1], Jennifer M. Ruddock[1], Brian Stankus[1], Lingyu Ma[1], Wenpeng Du[1], Nathan Goff[1], Yu Chang[1], Nikola Zotev[2], Darren Bellshaw[2], Adam Kirrander[2], Michael P. Minitti[3], and Peter M. Weber[1]*

[1]*Brown University, Department of Chemistry, Providence, Rhode Island 02912, USA*

[2]*School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, United Kingdom*

[3]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*

* Correspondence to: peter_weber@brown.edu (P.M.W.)

Abstract
Pump-probe gas phase X-ray scattering experiments, enabled by the development of X-ray Free Electron Lasers, have advanced to reveal scattering patterns of molecules far from their equilibrium geometry. For polyatomic molecular systems, large amplitude vibrational motions are associated with anharmonicity and shifts of interatomic distances (known as the 'shrinkage effect' in linear molecules), making analytical solutions using traditional harmonic approximations inapplicable. More generally, the interatomic distances in a polyatomic molecule are not independent and the traditional equations commonly used to interpret the data may give unphysical results. Here we introduce a novel method based on molecular dynamic trajectories and illustrate it on two examples of hot, vibrating molecules at thermal equilibrium. When excited at 200 nm, 1,3-cyclohexadiene (CHD) relaxes on a sub-picosecond time scale back to the reactant molecule, the dominant pathway, and to various forms of 1,3,5-hexatriene (HT). With internal energies of about 6 eV, the energy thermalizes quickly, leading to structure distributions that deviate significantly from their vibrationless equilibrium. The experimental and theoretical results are in excellent agreement and reveal that a significant contribution to the scattering signal arises from transition state structures near the inversion barrier of CHD. In HT, the analysis clarifies that previous inconsistent structural parameters determined from electron diffraction were artifacts resulting from the use of inapplicable analytical equations.

1

Introduction

Molecular structures at or near equilibrium have been extensively studied using static gas-phase x-ray and electron scattering[1,2,3,4]. The scattering signals are, fundamentally, the Fourier transforms of the electron or charge density distributions in the molecules, so that the measurements can be interpreted in terms of the geometrical information such as bond distances and angles. As quantum mechanical systems, molecules vibrate even at absolute zero temperature, which needs to be taken into consideration for accurate determinations of the molecular structures. For thermally stable molecules near equilibrium, a Boltzmann distribution in harmonic oscillations serves as a reasonable approximation[3,5,6]. But, as we show here, these approximations fail when the molecular systems are far from equilibrium as is the case in thermally highly excited systems. We describe a molecular dynamics approach to derive an accurate description of molecular scattering in such extreme conditions.

The imaging of molecular structures far from equilibrium is central to the study of chemical reactions[7,8,9]. The recent development of ultrafast pulsed x-ray free electron lasers[10] and MeV electron sources[11] have enabled measurements on molecular structures far from equilibrium. This resulted in numerous exciting findings, such as photo-induced ring-opening reactions[12,13,14], coherent nuclear vibrations[15,16], and photodissociation reactions[17,18]. While in many systems the nuclear dynamics starts out as a well-defined wavepacket with a structure distribution that resembles a classical system, the large amount of energy in play during the reaction typically relaxes into a hot bath as the molecule returns to a thermal state[13,19]. A complete description of the time resolved scattering patterns therefore requires not only a simulation of the transient dynamic structure, but also a model of the scattering patterns of molecules in thermally hot vibrational conditions. For polyatomic molecules with anharmonic degrees of freedom, this includes structures that are thermally populated yet far from the equilibrium geometry.

Several effects need to be considered for molecules with large-amplitude vibrations and much has been discussed in the context of static scattering signals of thermally excited molecules. Pioneering diffraction studies showed that the distances between non-bonded atoms in linear molecules may be shorter than the sum of bond distances (Figure 1)[20,21,22]. While small, this so-called Bastiansen-Morino shrinkage effect may be large enough to manifest itself in high quality experiments[23,24]. Bartell et al. have systematically explored the effect of shrinkage on scattering experiments[25,26,27,28,29,30]. The shrinkage effect exists for both harmonic and anharmonic vibrations,[30] which can be ascribed to the fact that interatomic distances are not necessarily congruous with normal mode vibrations. For nonlinear polyatomic molecules, the concept of shrinkage could be generalized to a temperature dependence of the interatomic distance distributions in vibrating molecules. This could include an enlargement instead of a shrinkage of the distances between some non-bonded atoms in non-linear molecules, as in the case of CHD that is subject of this study.
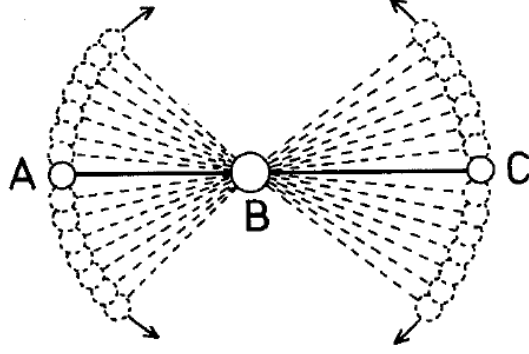
2

Figure 1. (*Something like this, will change to a new one*)For molecules with as few as three atoms, even harmonic normal mode vibrations may result in an apparent 'shrinkage' of interatomic distances compared to a rigid body at rest.

Treating the electron density of molecules as a sum of isolated-atom electron densities centered at the positions of the nuclei, the Independent Atom Model (IAM) expresses the scattering cross sections using tabulated atomic form factors in a convenient way[4]. For rotationally isotropic, non-vibrating molecules the Debye scattering equation for gas phase X-ray scattering is written as[3],

$$\frac{I_{total}(q)}{I_{Th}} = \sum_{i,j}^{N_{at}} f_i^0(q)f_j^0(q)\frac{\sin qR_{ij}}{qR_{ij}} + \sum_{i=1}^{N_{at}} S_{inel,i}^0(q), \tag{1}$$

where I_{Th} is the Thomson cross section for a free electron, N_{at} represent the total number of atoms in the molecule, R_{ij} is the interatomic distance between the ith and jth atoms, $f_i^0(q)$ and $S_{inel,i}^0$ are the elastic and inelastic atomic form factors,[4] respectively, for the ith atom. For harmonically vibrating molecules in a thermal Boltzmann distribution at temperature T, the formula is traditionally modified as[3],

$$\frac{I_{vib,har}(q)}{I_{Th}} = \sum_{i,j}^{N_{at}} f_i^0(q)f_j^0(q) \exp\left(-\frac{1}{2}l_{h,ij}^2 q^2\right)\frac{\sin q\left(R_{ij}-\frac{l_{h,ij}^2}{R_{ij}}\right)}{qR_{ij}} + \sum_{i=1}^{N_{at}} S_{inel,i}^0(q), \tag{2}$$

where $l_{h,ij}$ is the mean vibrational amplitude centered at R_{ij}, which depends on the temperature by[7]

$$l_{h,ij}^2 = \frac{h}{8\pi^2\mu\nu}\coth\left(\frac{h\nu}{2kT}\right) \quad, \tag{3}$$

where μ is the reduced mass, ν is the vibrational frequency, h is Planck's constant and k is the Boltzmann constant.

The use of the above equations requires care. The R_{ij} are interpreted as the centers of the Gaussian distributions that represent the interatomic distances during vibrations. Thus they are not always identical to the interatomic distance at equilibrium, as illustrated in Figure 1, when large-amplitude vibrations lead to shrinkage or enlargement effects. In general, an N-atomic non-linear molecule will have $\frac{N(N-1)}{2}$ interatomic distances but only 3*N*-6 vibrational degrees of freedom. This implies that for molecules with more than 4 atoms there will be redundant parameters of interatomic distances and vibrational amplitudes. The interatomic distances are not independent parameters, but rather are

3

correlated through the limited number of intrinsic vibrational modes. It may be useful to simulate a scattering pattern when all the exact R_{ij} and $l_{h,ij}$ values are known, but it is problematic to extract all the R_{ij} and $l_{h,ij}$ values by fitting a scattering pattern using Equation 2.

For highly excited systems, the anharmonicity of large-amplitude vibrations adds to the complexity of expressing interatomic distances. Analytical solutions have been derived for relatively simple polyatomic molecules such as CO_2, H_2O and SF_6, for temperatures below 1000 K[31]. Unfortunately, for more complicated molecular systems with less symmetry, and for the higher temperature that often result from optical excitation, reliable approximations and values for the asymmetry constants have not been obtained. Thus a better way of modeling scattering patterns of hot molecules is needed.

In this article we investigate a method for modeling scattering patterns of hot molecules based on molecular dynamics (MD) simulations. With the rapid development of the computational chemistry tool set, molecular dynamics simulations are widely applied to the interpretation of pump-probe scattering experiments. For example, Minitti et al. observed the real time ring-opening molecular motion of CHD within 200 fs by performing a global fit to time-resolved X-ray scattering patterns using ab initio multiconfigurational Ehrenfest trajectories[12]. A comparison of MeV electron scattering data with *ab initio* multiple spawning wavepacket simulations by Wolf et al. showed the subsequent conformational isomerism of HT following ring-opening of CHD in the sub-picosecond time regime[13]. It was even possible to follow the coherent vibrational motions of a molecule following electronic excitation using ultrafast X-ray scattering[16]. That analysis was made possible by utilizing a large number of possible molecular structures obtained from MD simulations. Thus we note that recent investigations of time-resolved pump-probe scattering have begun to take advantage of large-scale modeling of scattering patterns of systems that include far from equilibrium structures. Nevertheless, so far no systematic investigation of hot, vibrating molecules has been reported.
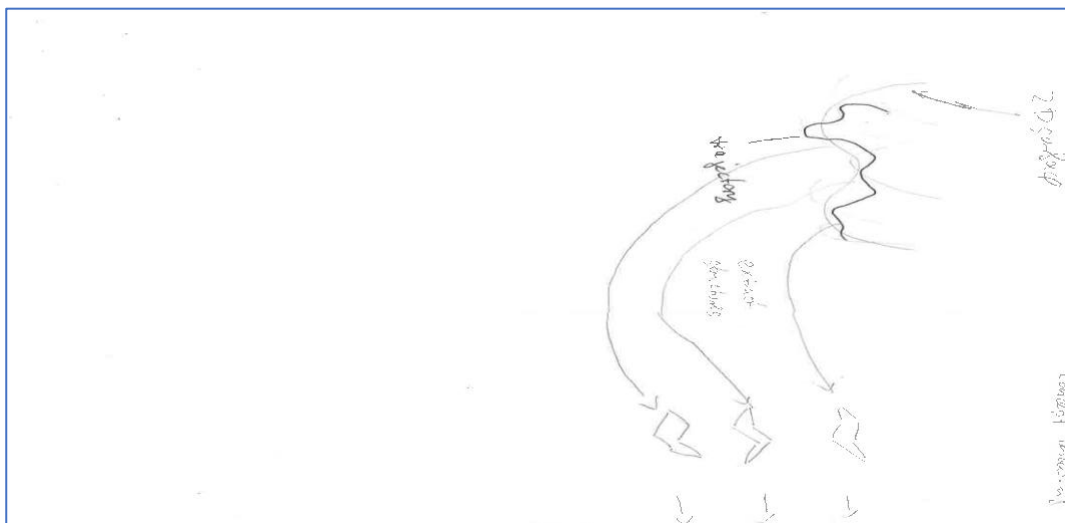

Figure 2. Concept of the method for modeling scattering patterns of hot molecules based on molecular dynamics (MD) simulations.

Figure 2 illustrates the method to calculate scattering patterns of hot molecules. An MD simulation creates a large number of molecular structures that include, depending on the initial energy used, structures that are far from equilibrium. The output of the MD calculation is inherently a time sequence of structures, but we disregard the temporal aspect and simply extract a large number of structures resembling the ensemble of vibrating molecules. For each of these structures the IAM scattering pattern is calculated following Equation 1. To simulate a scattering pattern of the hot ensemble of molecules the average over all the individual patterns is taken. By taking the entire pool of MD structures, the method inherently accounts for all previously mentioned effects such as shrinkage/enlargement, anharmonicity and correlation between interatomic distances. Moreover, the different structures are inherently weighted by their kinetic energies. Finally, the method can be applied to molecular structures of any complexity, limited only by the ability of MD methods to calculate reasonable trajectories. In principle, the method is not limited to the IAM model, as one could also use more comprehensive codes to simulate scattering patterns for each structure from ab initio wavefunction methods[32,33,34].

While the concepts of this article can be applied equally to X-ray and electron scattering, we have tested it on data from a time-resolved pump-probe X-ray experiment. Specifically, we explored gas-phase 1,3-cyclohexadiene and its isomer 1,3,5-hexatriene (HT) at thermally hot conditions. The interconversion between CHD and HT is an important model for electrocyclic reactions, and its motif is used in the synthesis of vitamin D in the skin upon exposure to sunlight[35]. The CHD/HT system is favorable for the present investigation because optical excitation at 200 nm is followed by a rapid electronic relaxation that dominantly leads to the vibrationally hot CHD in the ground electronic state[36]. The resulting molecules are very hot (~2870 K for CHD, assuming harmonic frequencies reported by Autrey et al.[37]), and undergo a thermal ring-opening reaction to HT. Because both the hot CHD reactant and the HT products are on the ground electronic surface, complications to the scattering patterns arising from the electronic excitation are avoided[38]. The thermal energy enables the electrocyclic reaction on the ground electronic surface, eventually leading to an equilibrium between hot CHD and hot HT. By introducing a forward and backward rate equations for this thermal reaction with the scattering patterns of hot CHD and hot HT as adjustable parameters, the two pure hot scattering patterns are extracted from the fits of the experimental data[14]. The ring-opening and ring-closing time constants were determined to be 174 ± 13 ps and 355 ± 45 ps, respectively[14].

Experimental and Computational Methods

Femtosecond time-resolved gas-phase X-ray scattering
The experimental apparatus has been introduced previously[39], the schematic of the experimental setup is shown in Figure 3. In short, the measurements were performed at the CXI instrument[40] of the LCLS X-ray free electron laser at the SLAC National Accelerator Laboratory. An ensemble of free CHD molecules at room-temperature is controlled by a piezoelectric needle valve to ~6 Torr of pressure at the interaction region. The molecules in the gas cell are excited by a 200 nm optical pump pulse with a ~80 fs

5

pulse duration, and probed by 9.5 keV X-ray photons with ~30 fs pulse duration, containing ~10^{12} photons/pulse. The pump and probe pulsed were focused collinearly into the scattering cell, with about 30 µm FWHM spot sizes for the X-rays and 50 µm for the laser. The time delay between the laser and X-ray pulses was controlled by a motorized delay stage and the timing jitter was monitored by a spectrally encoded cross correlator[41] to achieve a higher time resolution (~30 fs). To obtain the best signal to noise ratio, the shot by shot X-ray pulse intensity was monitored by a photodiode downstream of the scattering cell. Single shot scattering patterns are detected via a 2.3-megapixel Cornell-SLAC Pixel Array Detector (CSPAD)[42], binned by the delay time between the laser pump and X-ray probe pulses. The gas cell and the CSPAD detector are in a vacuum with about 2.6·10^{-4} Torr background pressure outside the cell, mostly comprised of CHD flowing out of the windowless scattering cell.
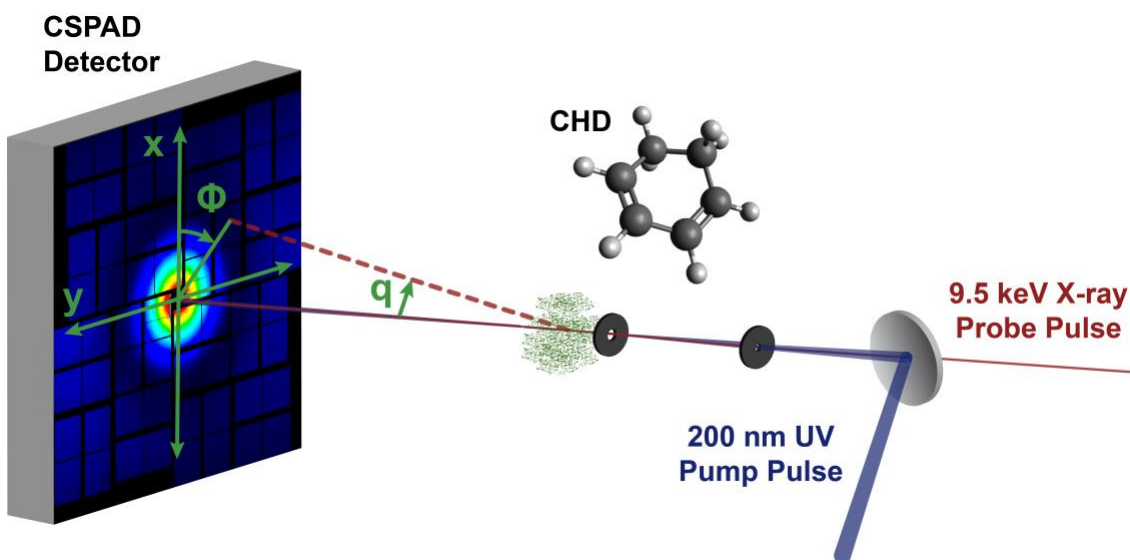

Figure 3. A schematic of the experimental setup. The reaction of CHD is initiated with a 200 nm pump pulse, and the time-evolving molecular structure is probed via scattering using X-ray probe pulses with variable time delay. The scattering signals are recorded as a function of the momentum transfer vector (q) and azimuthal angle (Φ) using a CSPAD detector.

Due to the quadrant symmetry of the $\cos^2(\theta)$ distribution resulting from a single-photon excitation process, the angular symmetry of the 2-dimensional scattering patterns can be described by second order Legendre polynomials.[43] In this case, the percent difference scattering pattern can be decomposed into an isotropic component that contains all the intrinsic molecular properties in the molecular frame and an anisotropic component that reflects both rotational and structural dynamics of the excited molecules in the laboratory frame[44]. Our analysis focuses on the isotropic component because the intramolecular vibrations of excited systems are fully described by the rotationally averaged scattering patterns. The time-dependent percent difference signal is expressed as

$$\%\Delta\mathrm{I}(q,t) = \gamma \cdot 100 \cdot \frac{\mathrm{I}_{on}(q,t) - \mathrm{I}_{off}(q)}{\mathrm{I}_{off}(q)} , \qquad (4)$$

where q is the magnitude of the momentum transfer vector, $I_{on}(q,t)$ is the scattering signal at time delay t, and $I_{off}(q)$ is the reference scattering signal when the pump laser is off. γ is the excitation fraction, which is kept below 10% to minimize undesired multi-photon processes. The advantage of expressing the pump-probe scattering signal as a percent difference is that poorly characterized experimental parameters such as background signals, gas pressure fluctuations and pixel noise are largely minimized. The decomposition into isotropic and anisotropic signals, the detector calibration, and the analysis of the measured scattering are discussed in more detail elsewhere[18].

Modeling scattering patterns of hot molecules using molecular dynamics
The ground state dynamics of hot CHD and hot HT are simulated using the SHARC program[45] interfaced with the electronic structure package MOLPRO[46]. More than 100,000 geometries calculated from MD are extracted to represent an ensemble of hot vibrating molecules for each CHD and HT at their thermal equilibrium. The internal energy of the system is adjustable by varying the energy input in the simulation and in this case was chosen to match the laser photon energy. Equation 1 is used to calculate the scattering pattern for each extracted geometry, and these individual patterns are then averaged to yield a single scattering pattern for the hot CHD or HT at the thermal equilibrium. The sequence of the averaging steps cannot be reversed: if one were to average the structures first and then calculate a scattering pattern of the averaged structure, important information about the distribution of interatomic distances during the vibrations is lost, resulting in the absence of damping that is characteristic of scattering patterns of hot molecules.

The initial conditions for the hot CHD are sampled from Wigner distributions to account for the zero point vibrations. The 6.2 eV of excess kinetic energy, corresponding to the energy of the 200 nm photons, was randomly distributed amongst all degrees of freedom of the molecules. The electronic structure calculations during the molecular dynamics are performed at the SS-CAS(2,2)/6-311+G(d) level. Similarly, the initial conditions for three HT conformers (cZc-HT, cZt-HT, and tZt-HT) were sampled from Wigner distributions. The relative equilibrium energies of cZc (1.02 eV), cZt (0.81 eV) and tZt (0.69 eV), relative to CHD, were obtained at the same level of theory and are in good agreement with previous reported literature values (0.37 eV of cZc and 0.13 eV of cZt relative to tZt[47]). For the 6.2 eV of photon energy, the total excess kinetic energies are therefore 5.18 eV for cZc, 5.39 eV for cZt and 5.51 eV for tZt. A total of 50 trajectories were calculated for CHD up to 5 ps. For HT, a total of 150 trajectories were calculated up to 3 ps, with 50 trajectories starting each from cZc-HT, cZt-HT and tZt-HT, so that the energy allows the molecules to traverse the entire ground state potential energy surface. Time steps of 0.5 fs are employed for all trajectory simulations.

In order to calculate the percent difference patterns accurately, the reference ground state signal I_{off} is very important. In the current study, Equation 2 is used to simulate the reference scattering pattern including harmonic vibrations at room temperature. The optimized structure at the SS-CAS(2,2)/6-311+G(d) level of theory is used as the equilibrium structure. A summary of several interatomic distances and structural angles are compared with previous experimental and theoretical results in Table 1. The structure

7

used here is deemed reliable as it agrees with previously reported computational and experimental results. The mean vibrational amplitudes of CHD at room temperature reported by S. J. Cyvin and O. Gebhardt are used for the calculation of the scattering patterns [48]. These values agree well with electron diffraction data measured by M. Traetteberg[49].

Table 1. Selected interatomic distances and structural angles of CHD at the ground state equilibrium. CAS: the structure calculated in this study, optimized from SS-CAS(2,2)/6-311+G(d). Ref 50: Optimized structure by Merchán et al. [50] Ref 49: Experimental structure from electron scattering measured by Traetteberg[49].

Parameter	CAS	Ref 50	Ref 49
C_1-C_2 (Å)	1.511	1.511	1.519
C_2-C_3 (Å)	1.532	1.533	1.538
C_5-C_6 (Å)	1.478	1.472	1.465
C_4-C_5 (Å)	1.343	1.343	1.348
$\angle C_5C_6C_1$ (º)	120.6	120.5	120.3
$\angle C_6C_1C_2$ (º)	120.9	120.6	120.3
$\angle C_1C_2C_3$ (º)	111.9	111.8	110.9

Convergence

To evaluate the convergence of the MD method for calculating hot scattering patterns, Figure 4a shows the average of the inversion angles for different time bins in the CHD MD simulation. All simulations start with positive values of the inversion angle sampled from zero point vibrations near 42.6º, which is the equilibrium structure inversion angle. Structures at early MD run times, before equilibrium is reached, should be dominated by positive inversion angles. Once the system has thermalized, structures with positive and negative inversion angles are expected to be equally populated. The approach to thermal equilibrium can therefore be illustrated by plotting the average inversion angles as the MD calculation progresses in time. As seen in Figure 4a, convergence is reached around 2 ps of MD simulations. Similar to the convergence of CHD, the time dependence of the convergence in HT is examined in Figure 4b. Since the MD simulations for HT are initiated from three different equilibrium structures (cZc, cZt and tZt), the difference between Φ_1 and Φ_3 is about 41°, as caused by the distinguishable labeling of the Φ_1 and Φ_3 in cZt conformers at the beginning (see labeling in Figure 8 insert). Once the system has thermalized, the assignment of these two angles are expected to be equivalent. As seen in Figure 4b, convergence is reached around 1 ps of MD run time.

To get a well-thermalized ensemble that accurately reflects the hot CHD molecules, a total of 200,000 structures extracted from 50 trajectories with MD run times between 3 and 5 ps are used. The resulting simulated scattering pattern of hot CHD is shown in Figure 6a. For hot HT molecules, a total of 300,000 structures are extracted from 2 to 3 ps time of run time, for 150 trajectories. Those structures are used to simulate the scattering pattern of the hot HT shown in Figure 9.
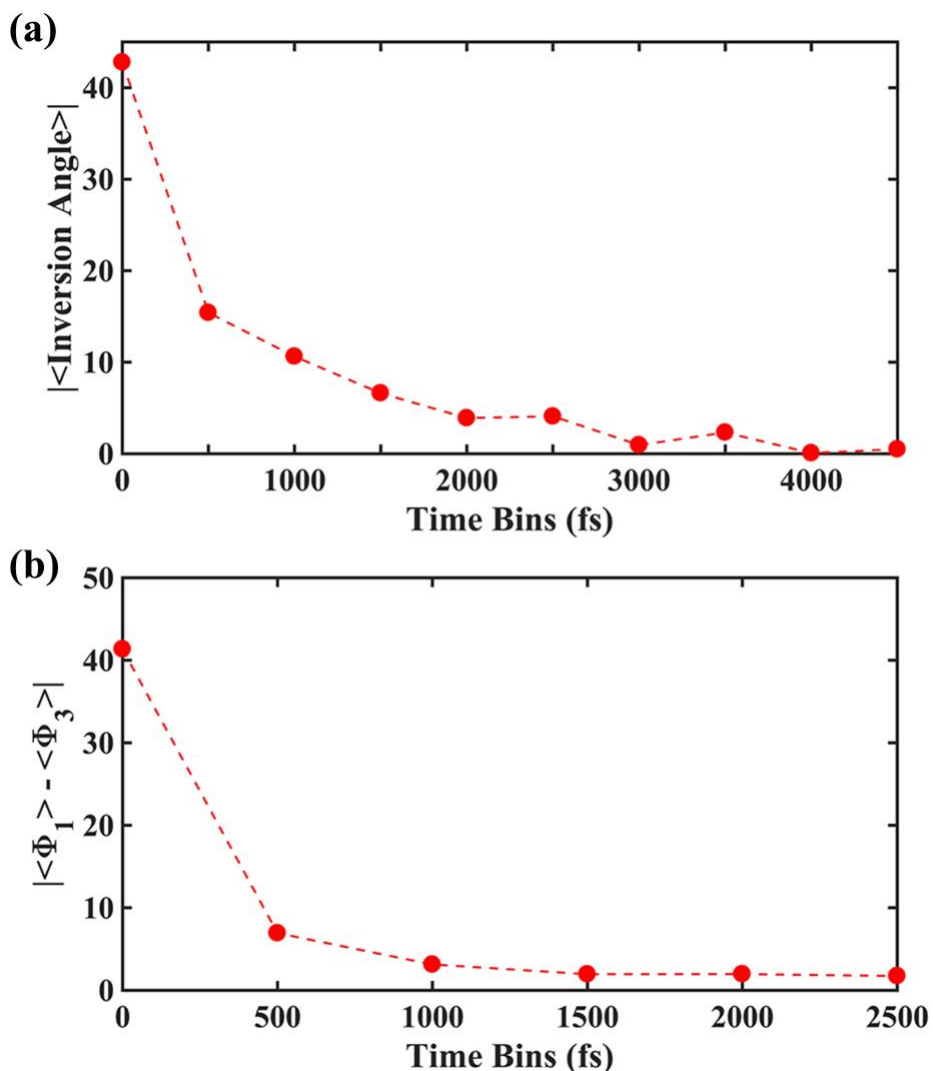
(a)

(b)

Figure 4. The convergence of the hot MD trajectories for (a) CHD and (b) HT. The inversion angle of hot CHD structures from all 50 trajectories with a one ps time bin centered at different time points are averaged to get the mean inversion angle shown as y axis values in (a). The Φ_1 and Φ_3 dihedral angles of hot HT structures from all 150 trajectories with a one ps time bin centered at different time points are averaged to calculate the mean Φ_1 and Φ_3 angles, respectively. The absolute difference between two mean angles are then shown as y axis values in (b). The center of each time bin is shown as the x axis value for each red dot in both plots.

Results and Discussions

The scattering signals of the hot CHD and the HT isomers were calculated using the MD methodology. Since detailed structural measurements exist for CHD, including vibrational amplitudes, a direct comparison of the signals calculated using the traditional approach (Equation 2), the MD approach and with experimental data is possible. For the HT system, comparable experimental data has not been reported, so that our comparison

9

is only between the MD calculated patterns and the experimental pattern of the hot isomeric system.

1,3-cyclohexadiene (CHD)
With two double bonds in a six-ring system, CHD has a non-planar C_2 structure with a double minimum in the C_1-C_2-C_3-C_4 torsional angle, Figure 3. The height of the barrier calculated at the SS-CAS(2,2)/6-311+G(d) level of theory is with 0.10 eV in close agreement with a previous experimental result of 0.13 eV[37]. A higher level calculation, using (SS-CAS(6,4)/aug-cc-pVDZ), offers confirmation, suggesting that the level of theory used for the MD simulation is sufficient for modeling the hot CHD. Equipartitioning the available energy amongst harmonic quantum oscillators indicates the inversion mode to have about 0.24 eV, which is more than twice the barrier to inversion. It stands to reason that a fair fraction of the molecules are near the barrier to inversion, i.e. in a region of the phase space that would not be easily accessible by harmonic motions, and that that this should have an effect on the calculated scattering patterns.
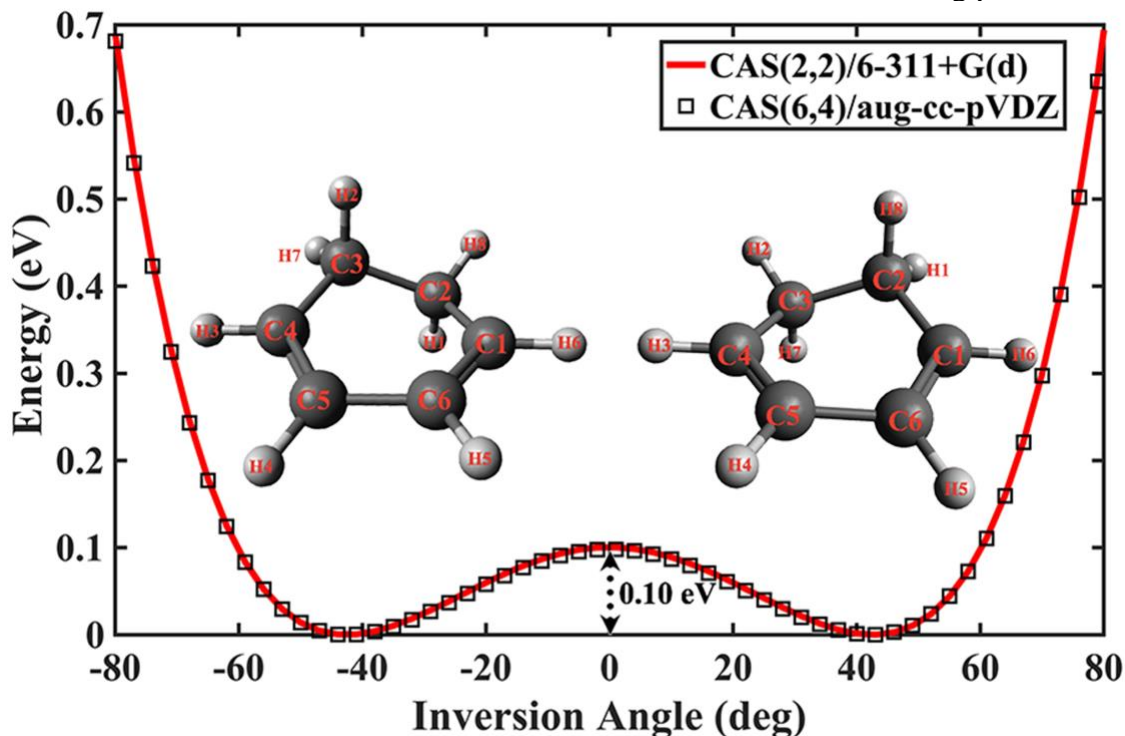

Figure 5. Potential energy of CHD as a function of the C_1-C_2-C_3-C_4 torsional angle. The level of computational method used for each curve is as shown in the legend. Insert: ground state equilibrium structures of CHD at negative and positive inversion angles at the potential energy minima.

Comparison of experimental and simulated percent difference scattering patterns
The percent difference scattering curves calculated with the MD method is compared to the experimentally measured pattern in Figure 6a. The MD scattering pattern is obtained using the method described above. The traditional scattering pattern is calculated using Equation (2), with $l_{h,ij}^2$ determined from Equation (3) based on the values of CHD at 0 K and 298 K reported by S. J. Cyvin and O. Gebhardt[48]. For comparison, a scattering

10

pattern of transition state structures calculated using the MD method but choosing only those structures near the inversion barrier in the ensemble is also shown. The noise in the experimental pattern is large because the difference between the scattering of hot and cold CHD molecules is small, approaching the detection limit (~0.05%) of the current time-resolved X-ray scattering experiments[16]. Nevertheless, it is seen that the agreement with the MD result is excellent. The model using Equation (2), on the other hand, fails to reproduce the experimental result. Importantly, the near-barrier scattering pattern shows a shape and phase that is quite similar to the experimental and MD patterns, indicating that the transition state structures make an essential contribution to the overall scattering signal of the hot molecule.

To illustrate the importance of including far-from-equilibrium structures in the simulation of scattering patterns, Figure 6b shows the distributions of the C1-C3 distance. Since Equation (2) uses the equilibrium distance amended only by a finite width distribution, the resulting scattering pattern is dominated by those structures that are near the equilibrium. However, the MD simulation shows that the maximum of the C1-C3 distance distribution shifts to a much larger value (blue bars). For the structures near the inversion barrier, red symbols, the C1-C3 distribution shifts to even larger distances. The structures that are very different from the equilibrium structures have larger percent difference patterns. It is therefore not too surprising that the experimental and MD scattering patterns show close similarity with the pattern of the transition state structures.
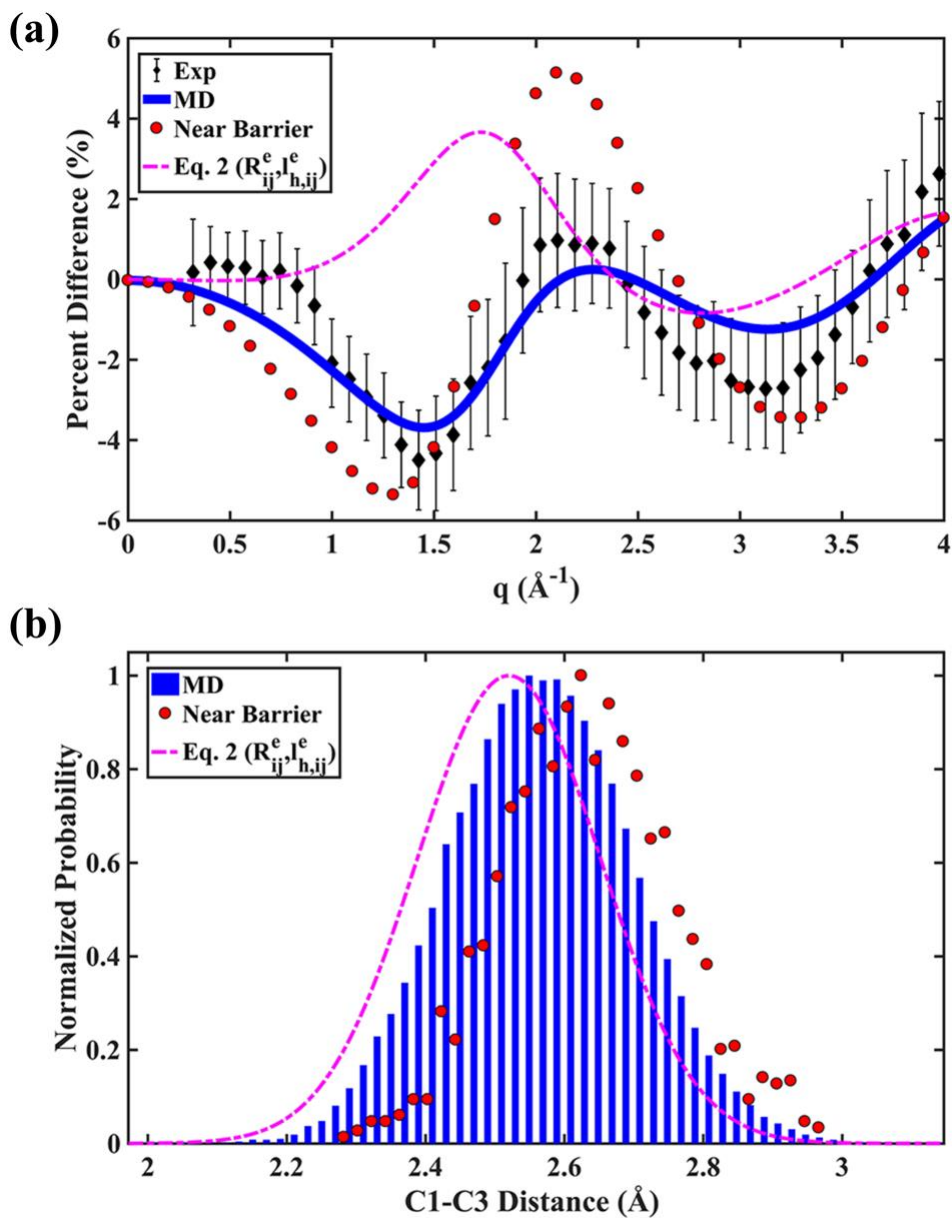
Figure 6. (a) Percent difference scattering patterns of hot CHD, assuming 100% excitation. Exp: experimental results, scaled to account for a 6.0% excitation fraction. MD: simulated hot scattering pattern from MD trajectories, averaged from 3-5 ps interval of 50 trajectories. Near Barrier: simulated scattering pattern from MD structures, averaged using structures with inversion angles between -1° and +1° from the 3-5 ps interval of 50 trajectories. Eq. 2 ($R_{ij}^e, l_{h,ij}^e$): simulated hot scattering pattern from Equation 2, with vibrational amplitudes calculated by Equation 3 using values reported by S. J. Cyvin and O. Gebhardt[48]. (b). Distributions of C1-C3 distances arising from the three models described in (a).

Discussion of various effects on simulated scattering patterns of hot molecules
From the agreement of the scattering signals computed by the MD method with the experimental results we conclude that the ensemble of molecular structures extracted

12

from MD after convergence is reached presents a satisfactory approximation of the distribution of molecular parameters. We can therefore use the results of the MD calculation as a benchmark to further discuss the impact of various effects on the simulated scattering patterns using different methods and structural parameters, Figure 7. To examine the isolated impact of each effect on the hot scattering pattern, the effects from anharmonicity, correlated distance, distance shifts and exact $l_{h,ij}$ value are filtered out as given in the legend of Figure 7a. From the curve A, which is the same as the MD curve in Figure 6a and which inherently takes all four effects into account, to the curve E, which is the same as the Eq. (2) curve in Figure 6a and which lacks any careful considerations of those four effects, the agreement becomes worse. The differences between sequential traces before and after each of those steps are presented in Figure 7b, which manifests the isolated difference from each of the individual effects.

It is apparent that during the process of filtering out the different contributions to the scattering signal from hot, vibrating molecules, the percent difference patterns increasingly deviate from the benchmarked MD result. The Eq. (2) result, which includes all the errors, has the largest deviation from the MD result (blue lines) or experimental result (black diamonds) as shown in Figure 6a. The different contributions are presented in Figure 7b. Both anharmonicity and distance shifts (shrinkage/enlargement effects) have quite large impacts on the accuracy of the hot scattering pattern, with almost similar magnitudes (~4%) when compared to the original MD scattering pattern at low q (< 3 Å^{-1}). Surprisingly, the curve labeled "Exact $l_{h,ij}$" suggests that Equation (3) can provide a fair estimate of the vibrational amplitudes: the scattering patterns calculated using the $l_{h,ij}$ values from the Equation 3 or extracted from the MD are quite similar to each other.

In addition to the anharmonicity and distance shifts, the correlation between internuclear distances also contributes to the accuracy of the scattering patterns. Although all the distributions of interatomic distances of "C: Eq.2 ($R_{ij}^H, l_{h,ij}^H$)" are exactly same as those of "B: MD harmonic ensemble", the scattering patterns calculated from them are quite different. This is because in Equation 2, each interatomic distance and its associated vibrational amplitudes is treated as independent parameters. However, these distances are correlated with each other during the vibrational motions of the molecules. In another words, the distributions of all or some of the other parameters will change when one of these distributions is changed. For example, choosing a small region of the inversion angle results in a noticeable change of the distribution (red dots) of the C1-C3 distance from its original distribution (blue bars), as shown in Figure 6b. The correlation of these distances could be ignorable for molecules near the equilibrium, as the effect from the vibrational distributions themselves is small for vibrating molecules near equilibrium. However as manifested here, this effect can no longer be ignored for molecules far from equilibrium. Since the number of interatomic distances and vibrational amplitudes, 2•91 in total, is much larger than the number of intrinsic degrees of freedom of vibrating molecules, 36 in total (overall 2•36 independent parameters if assign a width of Gaussian to each normal mode), one should be very cautious when using Equation 2 to analyze scattering patterns of molecular systems far from equilibrium. Even though the pattern resulting from Equation 2 might reproduce an experimental pattern very well, the so-

determined R_{ij} and $l_{h,ij}$ values would have no physical meaning. A specific example for will be discussed in the following section about hot HT.
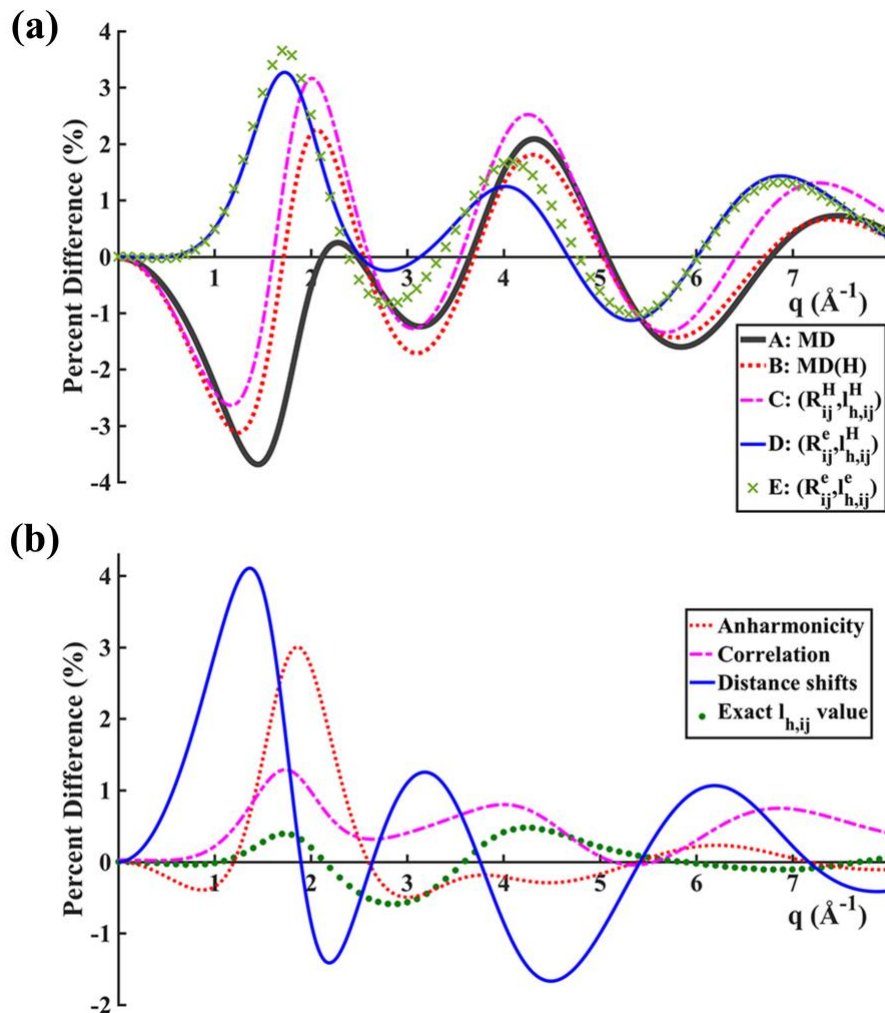

Figure 7. (a) Percent difference scattering patterns of hot CHD using different methods. A: the same MD ensemble as shown in Figure 6a. B: MD harmonic ensemble with anharmonicity filtered out from the MD ensemble by fitting each distributions of interatomic distance to a single Gaussian. C,D,E: all calculated by Equation 2, with different values of R_{ij} and $l_{h,ij}$. R_{ij}^{H}, $l_{h,ij}^{H}$: distances of maximum probability and half-widths at half maximum of the interatomic distance distributions extracted from the harmonic MD ensemble (B). R_{ij}^{e}: interatomic distances of equilibrium structure at ground state. $l_{h,ij}^{e}$: extrapolated $l_{h,ij}$ value at 2870 K using Equation 3. (b) Anharmonicity: B-A. Correlation: C-B. Distance shifts: D-C. Exact $l_{h,ij}$ value: E-D.

1,3,5-hexatriene (HT)

The 1,3,5-hexatriene system is complex because rotation about the single bonds creates structures that chemists would identify as distinct conformers. These conformers are separated by barriers that are easily surmounted by the energized molecules at the present energy. No experimental values for the individual conformer structures, or their vibrational amplitudes, have been reported. Several experimental investigations of the formation of hot HT from CHD in the ring-opening reaction initiated by 267 nm excitation have encountered the problem, and the interpretation of the data has been difficult due to the lack of understanding of the high temperature scattering patterns. In recent ultrafast time resolved X-ray and electron scattering experiments the issue was bypassed by comparing the experimental data directly with time-dependent theoretical trajectories[12,13]. This is a viable approach as long as the structure of the transient molecule is can be approximated by a dynamically evolving wavepacket with a well-defined center. In another electron scattering investigation, however, experimental data was least-square fitted to Equation (2)[7]. As pointed out in the previous sections, that method can be expected to lead to structural parameters without physical.

For the present study, the energy distributed into the vibrations of the molecules is about 6 eV, somewhat higher than in the experiments of the prior literature. As shown in Figure 7a, at that energy hot HT can transition freely among all three conformers, resulting in a broad and anharmonic distribution of structures, Figure 7b, represented by the torsional angles Φ_1 and Φ_3. The MD simulation allows the molecules to traverse the entire potential energy landscape, approaching convergence on a picosecond time scale. The nuclear probability histogram of the dihedral angles for the converged ensemble is shown in Figure 7 (b). As expected, the distributions of two of the dihedral angles (Φ_1 and Φ_3) become identical. Lacking detailed experimental data on the individual conformers, we compare the results solely from the MD simulation to the experimental data without further comparison to the Eq. (2) method of modeling the hot scattering patterns.
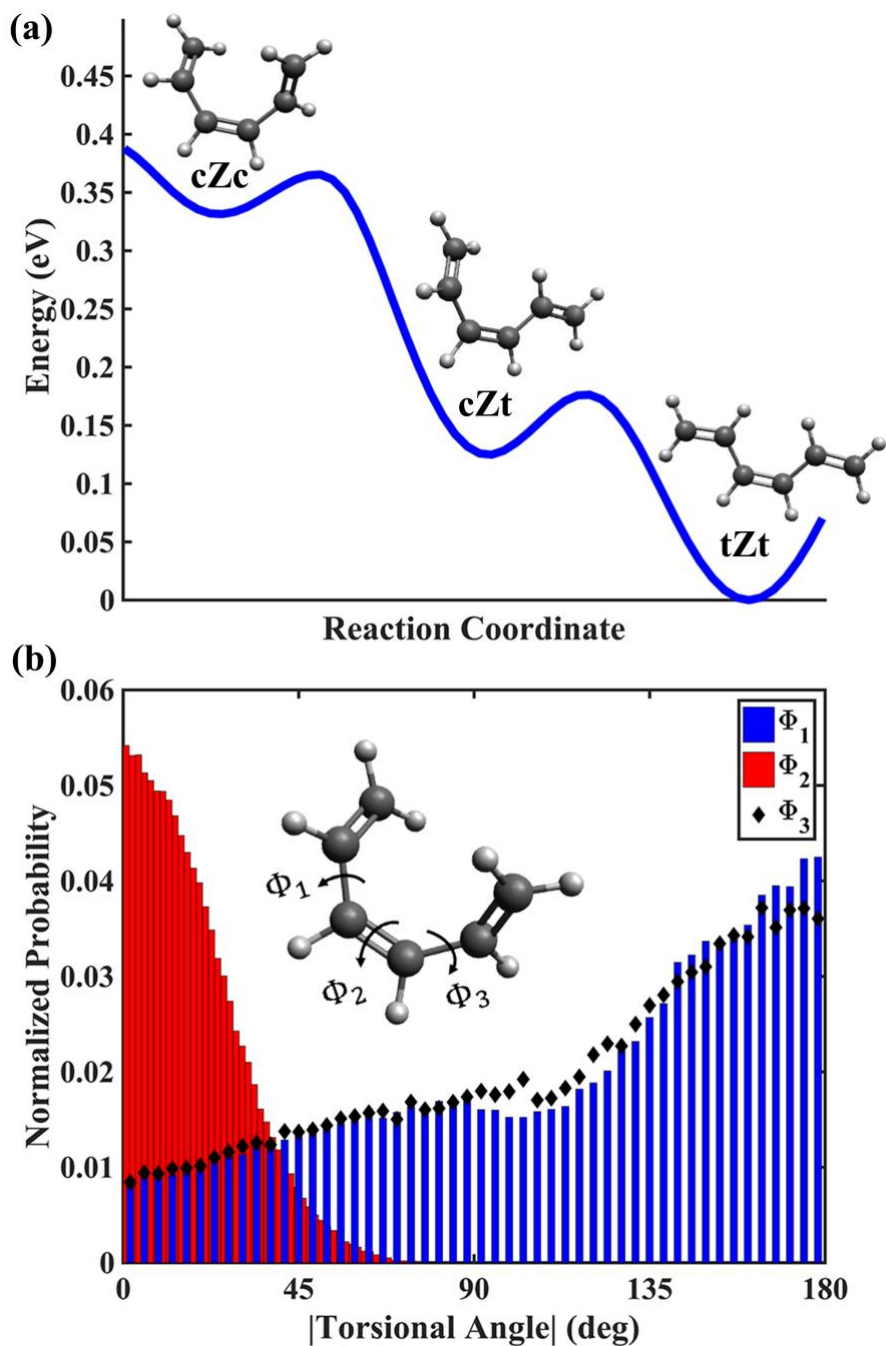
15

Figure 8. (a) Potential energy surface of HT showing the transition among three isomers. The surface is obtained by relaxed scans along the dihedral angles Φ_1 and Φ_3 at the SS-CAS(2,2)/6-311+G(d) level of theory. (b) Histogram of three torsional angles of HT in the ensemble used for modeling the MD pattern in Figure 9.

Comparison of experimental and simulated percent difference scattering patterns
The scattering pattern of hot HT calculated from the MD analysis is compared to the experimental measurement in Figure 9. The error bars of the experimental data for the hot HT are much smaller than those of the hot CHD because of the larger structural

16

difference between hot HT and the reference equilibrium CHD structure. The pattern calculated from the MD method shows an excellent agreement with the experimental result, further validating the accuracy of the method represented in this article.
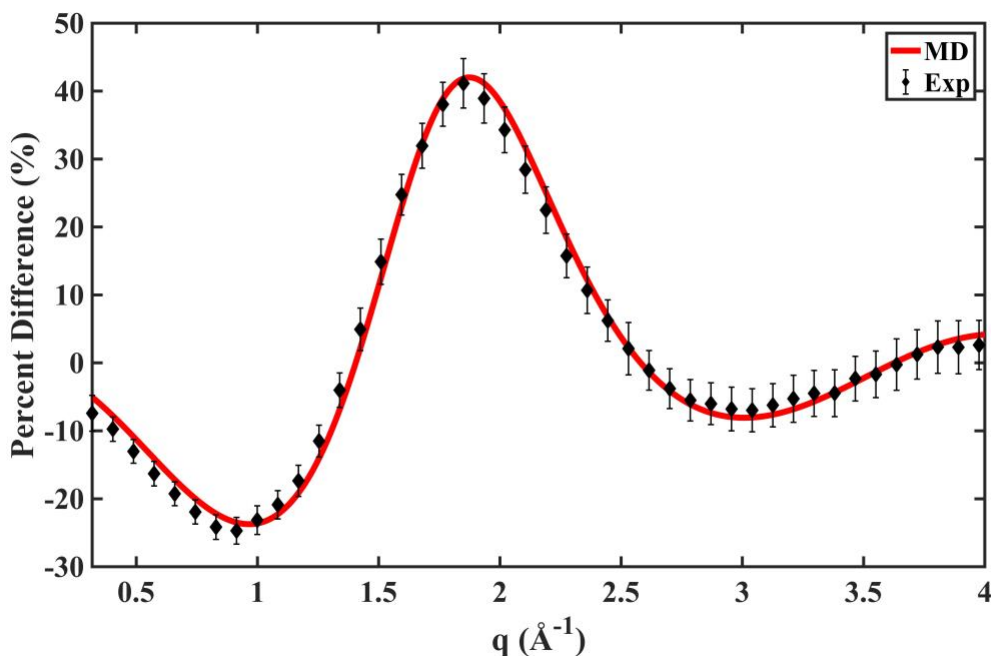

Figure 9. Percent difference scattering patterns of hot HT, assuming 100% excitation. Exp: experimental results, scaled by the 6.0% excitation fraction. MD: simulated hot scattering pattern from MD trajectories, averaged from 2-3 ps interval of 150 trajectories.

Distributions of structures of HT ensemble at high temperature

From the agreement of the scattering signals between MD method and the experimental result, we conclude that the ensemble of molecular structures extracted from MD after convergence is reached represent the hot HT ensemble very well. We can therefore use the result of the MD method to further study the distributions of structural parameters of hot HT, particularly all bonded interatomic distances and the farthest interatomic distance, Figure 10. Because of the rigidity of the chemical bonding, the distributions of bonded interatomic distances remains well characterized with a distinct maximum peak. However, for large non-bonded interatomic distances, the shape of the distributions become complicated and not well-defined, as illustrated in Figure 10b. The distribution seems to have at least three peaks (around 3.6 Å, 4.8 Å and 5.3 Å), indicating that it is indeed a hot ensemble of HT containing all three different conformers as well as various transition state structures. It is evident that it is implausible to use Equation 2 and expect meaningful bond parameters to result.

Table 2 lists the distances of maximum probability of all bond distances of hot HT extracted from MD. For comparison, a set of experimental structural parameters determined previously from electron diffraction data by Ruan et al. are also listed[7]. These latter distances were obtained by fitting the experimental scattering data using Equation 2 with R_{ij} and $l_{h,ij}$ as adjustable parameters. Although the fitted scattering curve matched well with the experimental data, the resulting fit parameters are questionable. As

17

highlighted in bold, the fits gave a 1.40 Å for the C2-C3 bond distance and a very large 1.71 Å bond distance of C4-C5. The distributions of HT structures from the MD ensemble give C2-C3 and C4-C5 values that are centered around 1.50 Å, even though here the energy distributed to the HT molecules is higher (200 nm photon). This again emphasize the importance of the careful treatment of the hot-molecule scattering pattern for the interpretation of the time-resolved X-ray and electron scattering data.
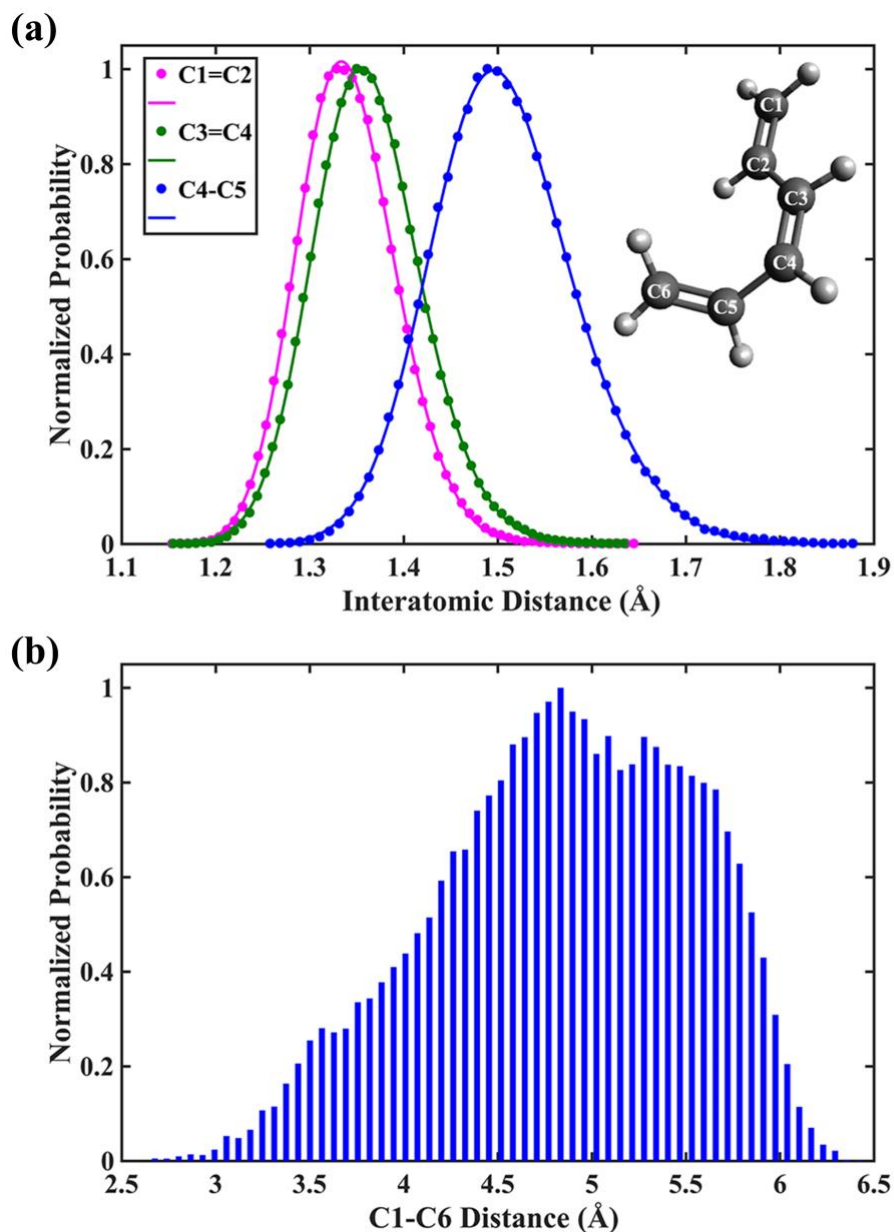

Figure 10. Distributions of selected interatomic distances of hot HT modeled by MD trajectories used in calculating the pattern in Figure 9. (a) Three bonded interatomic distances of HT as identified in the legend; dots are the original histograms, lines are the interpolated lines for each histogram. (b) Histogram of C1-C6 distance of hot HT ensemble.

18

Table 2. Bond distances and vibrational amplitudes of hot HT. Ref 7: experimental determined value by Ruan et al.[7]. Distance of maximum probability: the x values where the distributions in Figure 10a have the largest probability.

Bond Distance (Å)	Ref 7	Distance of maximum probability
C1=C2	1.29	1.334
C2-C3	**1.40**	**1.495**
C3=C4	1.41	1.355
C4-C5	**1.71**	**1.495**
C5=C6	1.32	1.333

Conclusion

In summary, a method for modeling scattering patterns of hot molecules using molecular dynamics simulations is introduced. The gas-phase 1,3-cyclohexadiene and its isomer 1,3,5-hexatriene (HT) at thermally hot conditions (~2870K) are used as examples. Excellent agreements between the hot scattering patterns calculated from the MD method and the experimentally measured X-ray scattering signals are obtained. Three important effects affect the scattering patterns: distance shifts, anharmonicity and the correlation between internuclear distances. While the change in distances and the anharmonicity contributes prominently, the correlation between internuclear distances during the vibrational motions cannot be neglected.

It is noteworthy that although the examples in this paper are based on X-ray scattering, the concept is equally applicable to electron scattering. A set of unreasonable structural parameters determined previously from electron diffraction data is traced to the inappropriate use of the conventional analytical equations. Future investigations of pump-probe scattering data need to be mindful of the limitations of conventional approaches.

The structure of the transition state, on top of the barrier to interconversion between the CHD isomers, is found to contribute significantly to the scattering signal. Because the transition state has a very different structure compared to the equilibrium structure at the bottom of the double well potential, its difference signal is strong. Consequently, even though only a small fraction of the molecular ensemble is near the barrier at any given time, the contribution to the scattering signal of the hot system cannot be neglected. It is possible that this finding could lead to novel methods to measure the structures of molecules at transition states. We anticipate that the present studies on modeling hot molecules provide an essential toolbox for exploring chemical reactions far from equilibrium and transition state structures using scattering techniques.

19

References

[1] T. S. Koritsanszky, Chem. Rev. 101, 1583 (2001).

[2] D. A. Kohl and L. S. Bartell, J. Chem. Phys. 51, 2891 (1969).

[3] I. Hargittai and M. Hargittai, Stereochemical Applications of Gas-Phase Electron Diffraction, VCH, Weinheim (1988).

[4] E. Prince, International Tables for Crystallography. Vol. C: Mathematical, Physical and Chemical Tables, 3rd Ed., Springer: Dordrecht, The Netherlands (2006).

[5] J. D. Geiser and P. M. Weber, J. Chem. Phys. 108, 8004 (1998).

[6] M. R. Ware, J. M. Glownia, A. Natan, J. P. Cryan and P. H. Bucksbaum, Phil. Trans. R. Soc. A 377, 20170477 (2018).

[7] C. Y. Ruan, V. A. Lobastov, R. Srinivasan, B. M. Goodson, H. Ihee and A. H. Zewail, PNAS 98, 7117 (2001).

[8] H. Ihee, V. A. Lobastov, U. M. Gomez, B. M. Goodson, R. Srinivasan, C. Y. Ruan and A. H. Zewail, Science 291, 458 (2001)

[9] R. C. Dudek and P. M. Weber, J. Phys. Chem. A 105, 4167 (2001).

[10] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F. -J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky and J. Galayda, Nat. Photonics 4, 641 (2010).

[11] S. P. Weathersby, G. Brown, M. Centurion, T. F. Chase, R. Coffee, J. Corbett, J. P. Eichner, J. C. Frisch, A. R. Fry , M. Gühr, N. Hartmann, C. Hast, R. Hettel, R. K. Jobe, E. N. Jongewaard, J. R. Lewandowski, R. K. Li, A. M. Lindenberg, I. Makasyuk, J. E. May, D. McCormick, M. N. Nguyen, A. H. Reid, X. Shen, K. Sokolowski-Tinten , T. Vecchione, S. L. Vetter, J. Wu, J. Yang, H. A. Dürr , and X. J. Wang, Rev. Sci. Instrum 86, 073702 (2015).

[12] M. P. Minitti, J. M. Budarz, A. Kirrander, J. S. Robinson, D. Ratner, T. J. Lane, D. Zhu, J. M. Glownia, M. Kozina, H. T. Lemke, M. Sikorski, Y. Feng, S. Nelson, K. Saita, B. Stankus, T. Northey, J. B. Hastings, and P. M. Weber, Phys. Rev. Lett. 114, 255501 (2015).

[13] T. J. A. Wolf, D. M. Sanchez, J. Yang, R. M. Parrish, J. P. F. Nunes, M. Centurion, R. Coffee, J. P. Cryan, M. Gühr, K. Hegazy, A. Kirrander, R. K. Li, J. Ruddock, X. Shen, T. Vecchione, S. P. Weathersby, P. M. Weber, K. Wilkin, H. Yong, Q. Zheng, X. J. Wang, M. P. Minitti and T. J. Martínez, Nat. Chem., just accepted (2019).

[14] J. M. Ruddock, H. Yong, B. Stankus, W. Du, N. Goff, Y. Chang, A. Odate, A. M. Carrascosa, D. Bellshaw, N. Zotev, M. Liang, S. Carbajo, J. Koglin, J. S. Robinson, S. Boutet, A. Kirrander, M. P. Minitti and P. M. Weber, Sci. Adv., submitted.

[15] Jie Yang, Markus Guehr, Xiaozhe Shen, Renkai Li, Theodore Vecchione, Ryan Coffee, Jeff Corbett, Alan Fry, Nick Hartmann, Carsten Hast, Kareem Hegazy, Keith Jobe, Igor Makasyuk, Joseph Robinson, Matthew S. Robinson, Sharon Vetter, Stephen Weathersby, Charles Yoneda, Xijie Wang, and Martin Centurion, Phys. Rev. Lett. 117, 153002 (2016).

[16] B. Stankus, H. Yong, N. Zotev, J. Ruddock, D. Bellshaw, T. J. Lane, M. Liang, S. Boutet, E. Carbajo, J. S. Robinson, W. Du, N. Goff, Y. Chang, J. E. Koglin, M. P. Minitti, A. Kirrander, P. M. Weber, Nat. Chem., submitted.

[17] J. Yang, X. Zhu, T. J. A. Wolf , Z. Li, J. P. F. Nunes, R. Coffee, J. P. Cryan, M. Gühr, K. Hegazy, T. F. Heinz, K. Jobe, R. Li, X. Shen, T. Veccione, S. Weathersby, K. J. Wilkin, C. Yoneda, Q. Zheng, T.J. Martinez, M. Centurion, X. Wang, Science 361, 64 (2018).

[18] J. M. Ruddock, N. Zotev, B. Stankus, H. Yong, D. Bellshaw, S. Boutet, T. J. Lane, M. Liang, S. Carbajo, W. Du, A. Kirrander, M. Minitti and P. M. Weber, Angew. Chem. Int. Ed. 58, (2019).

[19] A. Kirrander and P. M. Weber, Appl. Sci. 7, 534 (2017).

[20] A. Almenningen, O. Bastiansen, T. Munthe-Kaas, Acta Chem. Scand. 10, 261 (1956).

[21] A. Almenningen, O. Bastiansen, M. Traetteberg, Acta Chem. Scand. 13, 1699 (1959).

[22] A. Almenningen, O. Bastiansen, M. Traetteberg, Acta Chem. Scand. 15, 1557 (1961).

[23] Y. Morino, Acta Crystallogr. 13, 1107 (1960).

[24] S. J. Cyvin, Molecular Vibrations and Mean Square Amplitudes, Universitetsforlaget: Oslo, Elsevier, Amsterdam (1968).

[25] S. R. Goates and L. S. Bartell, J. Chem. Phys. 77, 1866 (1982).

[26] L. S. Bartell, W. Vance and S. R. Goates, J. Chem. Phys. 80, 3923 (1984).

[27] L. S. Bartell and J. F. Stanton, J. Chem. Phys. 81, 3792 (1984).

[28] L. S. Bartell, J. Chem. Phys 23, 1219 (1955).

[29] K. Kuchitsu and L. S. Bartell, J. Chem. Phys. 35, 1945 (1961).

[30] S. R. Goates and L. S. Bartell, J. Chem. Phys. 77, 1874 (1982).

[31] L. S. Bartell, J. Mol. Struct 63, 259 (1980).

[32] T. Northey, N. Zotev, and A. Kirrander, J. Chem. Theory Comput. 10, 4911 (2014).

[33] A. M. Carrascosa, H. Yong, D. L. Crittenden, P. M. Weber and A. Kirrander, J. Chem. Theory Comput., just accepted (2019).

[34] R. M. Parrish and T. J. Martínez, J. Chem. Theory Comput. 15, 1523 (2019).

[35] S. Deb and P. M. Weber, Annu. Rev. Phys. Chem., 62, 19 (2011).

[36] C. C. Bühler, M. P. Minitti, S. Deb, J. Bao, and P. M. Weber, Journal of Atomic and Molecular Physics, 2011, 637593 (2011).

[37] D. Autrey, J. Choo and J. Laane, J. Phys. Chem. A 105, 10230 (2001).

[38] H. Yong, J. M. Ruddock, N. Zotev, B. Stankus, A. M. Carrascosa, W. Du, N. Goff, Y. Chang, D. Bellshaw, M. Simmermacher, M. Liang, S. Carbajo, J. Koglin, J. S. Robinson, S. Boutet, M. P. Minitti, A. Kirrander and P. M. Weber, Nature, submitted.

[39] J. M. Budarz, M. P. Minitti, D. V. Cofer-Shabica, B. Stankus, A. Kirrander, J. B Hastings and P. M. Weber, J. Phys. B: At. Mol. Opt. Phys. 49, 034001 (2016).

[40] M. Liang, G. J. Williams, M. Messerschmidt, M. M. Seibert, P. A. Montanez, M. Hayes, D. Milathianaki, A. Aquila, M. S. Hunter, J. E. Koglin, D. W. Schafer, S. Guillet, A. Busse, R. Bergan, W. Olson, K. Fox, N. Stewart, R. Curtis, A. A. Miahnahri and S. Boutet, J. Synchrotron Radiat 22, 514 (2015).

[41] M. R. Bionta, N. Hartmann, M. Weaver, D. French, D. J. Nicholson, J. P. Cryan, J. M. Glownia, K. Baker, C. Bostedt, M. Chollet, Y. Ding, D. M. Fritz, A. R. Fry, D. J. Kane, J. Krzywinski, H. T. Lemke, M. Messerschmidt, S. Schorb, D. Zhu, W. E. White, and R. N. Coffee, Rev. Sci. Instrum. 85, 083116 (2014).

[42] H. T. Philipp, M. Hromalik, M. Tate, L. Koerner, S. M. Gruner, Nucl. Instruments Methods Phys. Res. Sect. A 649, 67 (2011).

[43] U. Lorenz, K. B. Møller and N. E. Henriksen, New J. Phys. 12, 113022 (2010).

[44] H. Yong, N. Zotev, B. Stankus, J. M. Ruddock, D. Bellshaw, S. Boutet, T. J. Lane, M. Liang, S. Carbajo, J. S. Robinson, W. Du, N. Goff, Y. Chang, J. E. Koglin, M. D. J. Waters, T. I. Sølling, M. P. Minitti, A. Kirrander, and P. M. Weber, J. Phys. Chem. Lett. 9, 6556 (2018).

[45] S. Mai, P. Marquetand, L. González, WIREs Comput. Mol Sci. 8, 1370 (2018).

[46] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, WIREs Comput. Mol. Sci. 2, 242 (2012).

[47] S. H. Pullen, N. A. Anderson, L. A. Walker and R. J. Sension, J. Chem. Phys. **107**, 4985 (1997).

[48] S. J. Cyvin and O. Gebhardt, J. Mol. Struct 27, 435 (1975).

[49] M. Traetteberg, Acta Chem. Scand. 22, 2305 (1968).

[50] M. Merchán, L. Serrano-Andrés, L. S. Slater, B. O. Roos, R. McDiarmid and X. Xing, J. Phys. Chem. A 103, 5468 (1999).