Effects of vibrational motion on core-level spectra of prototype organic molecules

Janel S. Uejio, Craig P. Schwartz, Richard J. Saykally, David Prendergast

A computational approach is presented for prediction and interpretation of core-level spectra of complex molecules. Applications are presented for several isolated organic molecules, sampling a range of chemical bonding and structural motifs. Comparison with gas phase measurements indicates that spectral lineshapes are accurately reproduced both above and below the ionization potential, without resort to ad hoc broadening. Agreement with experiment is significantly improved upon inclusion of vibrations via molecular dynamics sampling. We isolate and characterize spectral features due to particular electronic transitions enabled by vibrations, noting that even zero-point motion is sufficient in some cases.

Article history:
Received 21 August 2008
In final form 27 October 2008
Available online 7 November 2008

When applied to molecular systems, core-level spectroscopies are powerful probes of both occupied and unoccupied electronic states, uniquely revealing intimate details of both intra- and inter-molecular interactions [1]. Methods involving X-ray absorption (XAS, NEXAFS, XANES) or X-ray photo-electron spectroscopy (XPS) are increasingly being applied to complex molecular systems, including nucleotides, peptides and large organic molecules [2]. However, a major limitation of this technology is the fact that extraction of molecular information from these experiments often depends explicitly on comparisons with theoretical calculations, which are extremely challenging to perform at experimental accuracy. In this Letter, we describe the extension of a recently developed method for predicting core-level spectra of condensed phases [3] to isolated organic molecules – pyrrole, s-triazine, pyrrolidin and glycine – which demonstrates qualitative improvements over existing methods [4–6] in comparison with experiment and provides new insights into the origins of particular spectral features in terms of coupling of electronic and vibrational degrees of freedom.

The challenges for simulating gas phase core-level spectra are maintaining accuracy in the following areas: (1) description of the core-hole excited state; (2) representation of both bound excitonic states below the ionization potential (IP) and resonance states in the continuum above the IP; and (3) inclusion of vibrational effects, either due to experiments being performed near room temperature, or from intrinsic zero-point motion.

Density functional theory (DFT) [7,8] has proved accurate in reproducing the excitation energies associated with core-level spectra via total energy differences (so-called ∆SCF or ∆KS) [9]. Accordingly, we model the lowest energy core-level excited state self-consistently using a full core-hole and excited electron (XCH) [3]. This is particularly important for molecular systems, where screening of the core-hole excitation is greatly enhanced by the presence of the excited electron, which can be strongly bound to the core-hole in the lowest energy excited state. In contrast, for non-molecular condensed phases, such as covalent and ionic crystalline solids, the inherent dielectric screening of the valence charge density often dominates, and so, explicit inclusion of the excited electron may prove insignificant in such cases [5]. We use the PBE form of the generalized gradient approximation to the exchange-correlation potential [10]. Transition amplitudes are estimated in the single-particle and dipole approximations and excitations to states above this first excited state are approximated using the unoccupied Kohn–Sham eigenstates computed from the XCH self-consistent potential. This is in contrast to the closely related full core-hole (FCH) approximation [5,6], which ignores the excited electron, or replaces it with a uniform background charge density, and the half core-hole (HCH) approach [4] related to Slater’s transition-state potential (TP) [11]. The HCH (or TP) approach has been applied extensively to molecular and cluster models of materials using linear combinations of atomic orbitals (LCAO) to describe the electronic structure. These have included applications to isolated molecules [9], molecules on surfaces [12], and condensed phase molecular liquids [13]. In our XCH implementation [3] we use norm-conserving pseudopotentials. Core-hole matrix elements with valence electrons are calculated by reconstructing the core region of the pseudostates within an atomic frozen core approximation. Other approaches often treat the core-excited atom at the all-electron level, while using effective core potentials for the surrounding unexcited atoms.

There are other approaches to modeling core-level excitations: the static exchange approach (STEX) of Ågren et al. [14] describes...
final state of the core-level excitation by freezing the orbitals of the molecular ion and calculates their exchange interaction with the excited electron; the multiple scattering approach of Ankudinov and Rehr [15] has been used extensively to examine the X-ray absorption of compounds at the edges of (typically) heavier elements using real-space cluster models [16]; accurate solutions to the Bethe-Salpeter equation have been used effectively by Shirley [17] in the context of core-level spectra of crystalline solids.

We apply our XCH approach within periodic boundary conditions using a plane-wave basis, enabling uniform convergence in accuracy for representing both localized and delocalized states. In contrast, LCAO approaches have difficulty in describing delocalized states, particularly those appearing above the IP in core-level spectra of isolated molecules. This has been mitigated in the past using Stieltjes imaging techniques [18] or some ad hoc numerical broadening. However, in our approach, the use of plane-waves to represent electronic states of isolated molecules engenders a significant computational cost. Spurious interactions between periodic images must be reduced by increasing the size of the supercell, which, in turn, increases the size of the basis and the density of electronic states to be determined at energies above the ionization potential. Our compromise is to use supercells large enough to represent the excited states below the IP. Then we take full advantage of the periodic boundary conditions to approximate the high energy continuous electronic density of states by numerically converging an integration over the first Brillouin zone (BZ). This will have no impact on bound states localized fully within the supercell. However, for states which span the supercell, an accurate determination of the electronic density of states can be achieved by BZ sampling [19]. Such delocalized states should be close analogues of free electron states scattered from the molecule. The weakness of this approach is in describing bound states below the IP having a spatial extent larger than the supercell, but we can mitigate this effect by increasing the size of our supercells.

For each of the molecules studied in this work we used supercell volumes of $20^3 \text{Å}^3$ and a plane-wave kinetic energy cut-off of 85 Ry. In all cases, approximately 100 Kohn–Sham eigenstates are used in constructing transition matrix elements. This is only sufficient to extend our spectra approximately 3 eV above the estimated IP. To reduce the significant computational cost of a numerically converged BZ sampling, we exploit a recently implemented interpolation scheme (based on an approach by Shirley [20]) that requires only the electronic states at the zone center as input. Furthermore, this scheme also increases the number of electronic states beyond the $\sim 3$ eV limit. The accuracy of these states is not guaranteed (see fixed-nuclei spectra in Fig. 1), but finite-temperature sampling improves the agreement with experiment. The zone-center electronic structure is calculated using the PWSCF code [21].

Typically, core-level spectra of isolated molecules are simulated within the fixed-nuclei approximation, particularly for molecules in their vibrational ground state under the experimental conditions. A candidate (lowest energy) structure is chosen and the electronic structure is calculated while modeling the atomic nuclei as fixed point charges, located at the mean of the nuclear distribution or, more commonly, at an energy minimum derived from a formalism which models the electrons as quantum particles and the nuclei as classical point charges. Usually, the core-excitation transition amplitudes are estimated without the impact of nuclear dynamics on the electronic subsystem. Often finite temperature effects are approximated by increased numerical broadening of calculated spectral peaks. More detailed approaches calculate Franck–Condon factors based on a vibrational eigenmode analysis in the ground and excited states. These factors are used to modulate a single electronic transition and help to reproduce asymmetric lineshapes associated with accompanying transitions to excited vibrational modes [9,22,23]. However, the Franck–Condon approximation ignores the impact of nuclear motion on the electronic transition amplitude. To first order, this impact is referred to as the Herzberg–Teller effect [24].

In this work, we attempt to include the impact of nuclear motion on core-level excited electronic states and transition
amplitudes. We model the nuclear degrees of freedom in these molecules using classical molecular dynamics (MD) performed at 300 K using a Langevin thermostat. We used AMBER 9 with the generalized AMBER force field and Antechamber [25,26]. The resulting distribution of nuclear coordinates is sampled at regular intervals, spaced at least 10 ps apart to reduce correlation between nuclear snapshots, for at least 100 snapshots. For the small molecules studied here, the molecular dynamics calculations represent an insignificant computational overhead with respect to the 100 or more plane-wave electronic structure calculations required to simulate the core-level spectra. For larger systems, or for first principles molecular dynamics sampling, such long trajectories may not be computationally tractable. In this case, Monte Carlo sampling may prove more efficient in sampling accessible molecular configurations, but it was not used in this work. We recognize that for each of these molecules, some (or all) of the vibrational eigenmodes (estimated using DFT calculations in good agreement with experiment), are in their quantum ground states at room temperature and will exhibit systematically different spatial distributions from the classical model. The use of a Langevin thermostat leads to distributions of nuclear coordinates which resemble quantum distributions (they are not peaked at the classical turning points), although their mean-square displacements are typically underestimated. Nevertheless, using our computationally inexpensive uncorrelated sampling approach, we still find that significant improvements in experimental agreement are possible over using only the vibrationless mean nuclear positions of the fixed-nuclei approximation. For molecules occupying their vibrational ground states at experimental temperatures, we use this improved agreement as an indicator, that features missing from the fixed-nuclei spectra, appear in experiment due to zero-point vibronic coupling effects, which might be well-reproduced by the Herzberg–Teller approximation. We see evidence for Herzberg–Teller effects in the spectra of S-triazine and glycine and future work will address these in more detail.

The approach of sampling molecular dynamics trajectories has been applied already in simulating core-level spectra of molecular clusters and liquids, particularly for photoelectron spectroscopy [27,28], X-ray absorption spectroscopy [3], and X-ray emission spectroscopy [13], where distinct changes have been observed based on configurational changes. For molecules with multiple low-energy conformations or significant anharmonicity, the molecular dynamics approach has clear advantages over vibrational eigenmode analysis. Calculation of eigenmodes is only possible about a minimum in the potential energy surface and for more complex systems there can be many such minima. We shall see evidence for this in the case of glycine below.

All calculated spectra are numerically broadened using Gaussian of 0.2 eV full width at half maximum (FWHM). Previous simulations have used larger and nonuniform numerical broadening in order to simultaneously approximate electronic and vibronic coupling. For example, in Ref. [9] a 0.3 eV FWHM broadening was used below the IP, and this was linearly increased to 4.5 eV for the next 30 eV and then held constant. (In the same work a smaller broadening of 0.15 eV is used when discussing vibronic effects.) In contrast, we use a relatively small and uniform broadening with the aim of simulating and distinguishing electronic and vibrational effects explicitly, thereby arriving at a predictive computational approach.

Fig. 1 provides comparison between measured core-level spectra from NEXAFS and inner-shell electron energy loss (ISEELS) and calculated spectra using the HCH LCAO and XCH approaches both with and without sampling of nuclear degrees of freedom. All calculated spectra are aligned by their IP estimates. We observe systematic contraction of XCH spectra along the energy axis, consistent with previous work [3,29]. The HCH LCAO calculations were performed with a commercially available package, STO-3G. [9,30] The excited nitrogen of interest was modeled using the IGLO-III basis set [31], the hydrogens were modeled with a diffuse basis set and the remaining non-hydrogen atoms were modeled with double zeta valence plus polarization basis sets included in the STO-3G package. We used the same PBE functional in both XCH and HCH calculations. We found that using larger basis sets and functionals led to only minor spectral changes in HCH calculations.

Pyrrole – found in biological systems (Fig. 1A), this molecule has an aromatic rigid ring structure and occupies its vibrational ground state at the experimental temperature. The NEXAFS experimental data were collected using total electron yield at beamline 8.0.1 at the Advanced Light Source, using prior methods [32] and the ISEELS was previously taken [33]. Peaks 1 and 2 are reproduced well by XCH without vibrations, but with an incorrect peak height ratio. HCH LCAO calculations produce a spurious peak after the first main peak and the second peak is too intense compared to experiment. Previous experiment has identified a shape resonance [34], labeled 3 here, which has some oscillator strength above the decay that occurs well above the ionization potential; some features are apparent there, but nothing definitive emerges from either calculation.

Averaging 100 MD snapshots, XCH is able to produce an improved intensity ratio between features 1 and 2 and a smoother continuum region above 410 eV. MD sampling broadens out the features in the HCH LCAO calculation, but incorrect peak height ratios and an overly structured continuum region remain. The latter is most likely due to basis set limitations.

S-triazine – a very rigid prototypical aromatic molecule; much like pyrrole, this molecule is in its vibrational ground state at the experimental temperature. The fixed-nuclei XCH spectrum (Fig. 1B) does not compare well with experiment [35,36]. Peak 1 is reproduced by XCH but with overestimated oscillator strength. The small shoulder labeled 2 corresponds to the LUMO(+1) and is attributed to vibronic coupling. HCH LCAO and XCH are unable to reproduce this peak without inclusion of vibrations, but they capture peak 3. Both methods are plagued by spurious features in the continuum region in the absence of vibrations.

It was expected that using MD snapshots would produce only small changes due to S-triazine being in the vibrational ground state at the experimental temperatures. However, inclusion of MD sampling results in a clear improvement with extra broadening induced by small displacements of the nuclei. The XCH approach more accurately captures peaks 1, 3, 5 and 6; the continuum is smooth and in better agreement with experiment. HCH LCAO is similarly broadened, with large sampling error bars around peak 1; the continuum region remains overly structured.

Peak 2 is evident as a large sampling error bar from MD sampling and is visible for individual snapshots for both HCH LCAO and XCH calculations. One such snapshot is examined in detail in Fig. 2. Here we see two electronic transitions which are forbidden in the absence of vibrations, but turn on for displaced nuclei. Breaking the in-plane mirror symmetry modifies the electronic structure at the excited nitrogen atom. For the LUMO(+1), this disrupts a nodal plane through the excited nitrogen, while for the LUMO(+2) a ring-like state in the molecular plane shifts adding to the p-character at the excited nitrogen.

We note that none of our calculations reproduce peak 4. It is believed, by comparison with a similar feature in molecular nitrogen [1,35], that this is multi-electron in origin. Our calculations do not include excitations of more than one electron. However, we have performed test calculations which include an additional electron excitation from near the top of the valence band (shake-up), indicating some transitions in the right energy range, such as HOMO(-2) or HOMO(-3) to LUMO.
Pyrrroldine – structurally similar to pyrrole but non-aromatic, pyrrroldine has an active low frequency vibrational mode centered on its nitrogen atom. The XCH fixed-nuclei approximation predicts three distinct large peaks, corresponding to the three experimental features [33], (1, 3, and 4 respectively); it also produces the small three distinct large peaks, corresponding to the three experimental on its nitrogen atom. The XCH fixed-nuclei approximation predicts pyrrolidine has an active low frequency vibrational mode centered from mean nuclear positions of the vibrational ground state, corresponding to the first two states above the lowest unoccupied molecular orbital (LUMO). Top: iso-surfaces of the LUMO+1 (left) and LUMO+2 (right) with (red) or without (dash-dot blue) nuclear displacements; green (red) indicates positive (negative) values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Glycine – in Fig. 1D, we see that the spectrum of glycine, the simplest amino acid, comprises two well resolved peaks, 1 and 2, a smaller less defined peak 3, and a broad peak 4 [37,38]. In the absence of vibrations, the XCH method reproduces all 4 peaks, as does the traditional HCH-LCAO. However neither method yields the correct general shape of the experimental spectra. Note that the MD averaged XCH spectrum has been rescaled by a factor of three.

A recent study on core-level spectra of glycine, examined only its four lowest energy conformers, and was expected to represent all individual conformers with populations >2% of the total [38]. Due to their atomistic sensitivity, core-level spectra can be strongly influenced by sparsely populated conformations [39]. By comparison with a Boltzmann-averaged spectrum of the four dominant conformers (not shown) it was apparent, in this case, that merely analyzing the four lowest energy conformers may not be adequate to accurately reproduce core-level spectra.

MD sampling has the advantage of not merely sampling the lowest energy state conformers, but also phase space the molecule would have to occupy as it changes conformations. Including MD sampling in HCH LCAO calculations produces a spurious low-energy feature with low intensity, a second peak that is much too in-
tense, and the fourth feature is not resolved. However, the overall spectral appearance is much closer to that of experiment, particularly in the low-energy range. At higher energy, the same nonphysical peaks present for the fixed-nuclei structure are also present with MD. Using the IP for alignment is beneficial, providing objective alignment when the first peak is not experimentally observed.

Using a 0.2 eV FWHM broadening scheme there appears to be an extra peak just above the onset in the XCH MD spectrum, however when the peak widths are increased to 0.4 eV FWHM (Fig. 3), this extra peak merges with the onset, and the measurement is reproduced. This indicates that we have insufficiently or incorrectly sampled the full configuration space of glycine using our classical MD approach. Perhaps more sampling might broaden this new peak correctly. This example indicates how tempting it can be to broaden theoretical spectra in order to ‘fit’ to experiment, at the sacrifice of predictability and correct interpretation. At higher energies, the XCH approach performs reasonably for glycine, where there is a reasonable decay, devoid of any unphysically sharp features.

Using a finer 0.2 eV FWHM numerical resolution, it is clear from Fig. 1D that the ‘new’ spectral feature between the measured peaks 1 and 2 is not present in the absence of vibrations, for the lowest energy structure of glycine. A survey of those molecular configurations within our molecular dynamics ensemble which contain this extra peak indicates that it results from a non-zero dihedral angle along the molecular backbone from the nitrogen atom to the carboxyl oxygen (NCC = O), which breaks the mirror symmetry of glycine. Characterization of the first four core-level electronic transitions is provided in Fig. 4. Using the XCH approach, we find the first allowed transition of the lowest energy structure has $\sigma_{\text{NH}}$ character, in agreement with previous work [37] using the STEX approach [14]. However, the agreement stops here. The next two allowed transitions, computed using XCH, have $\sigma_{\text{IC}}$ character and $2\beta$-like symmetry respectively. (We imagine the NH$_2$ group like the water molecule which also possesses a true $2\beta$ unoccupied state.) Each of these three transitions persists upon variation of the NCC = O dihedral angle, albeit with some small energy-reordering and shifting in position and oscillator strength which can be understood by analysis of expansion or contraction of the excited electron density in the presence of these structural perturbations.

However, there also exists an optically forbidden transition in the lowest energy structure, only 14 meV higher in energy than the first excited state. From the electronic wave function iso-surfaces, we see that this forbidden state is of $\pi_{\text{C=O}}$ character and is forbidden due to its lack of significant overlap with the core-excited nitrogen. Upon increasing the dihedral angle, this state leaks onto...
In conclusion, we have demonstrated an accurate computational approach for prediction and interpretation of core-level spectra, enabling new electronic transitions which are forbidden when nuclear quantum effects are ignored. Future work in which the nuclear degrees of freedom are sampled correctly with respect to their quantum ground state distribution will further illuminate this issue.

Acknowledgements

We thank Wanli Yang at Beamline 8.0.1 for his help with the pyrolysis experiment. This work was supported by the Director, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy under Contract No. DE-AC02-05CH11231 through the LBNL Chemical Sciences Division, the Molecular Foundry and the Advanced Light Source. Computational resources were provided by NERSC, a DOE Advanced Scientific Computing Research User Facility.

References