# UC Berkeley
## UC Berkeley Electronic Theses and Dissertations

**Title**

X-ray Absorption Spectroscopy Study of Prototype Chemical Systems: Theory vs. Experiment

**Permalink**

https://escholarship.org/uc/item/0xd515cp

**Author**

Schwartz, Craig Philip

**Publication Date**

2010

Peer reviewed|Thesis/dissertation

X-ray Absorption Spectroscopy Study of Prototype Chemical Systems: Theory vs. Experiment

by

Craig Philip Schwartz

A dissertation in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in Charge:
Professor Richard J. Saykally, Chair
Professor Daniel M. Neumark
Professor Robert M. Glaeser

Spring 2010

X-ray Absorption Spectroscopy Study of Prototype Chemical Systems: Theory vs.

Experiment

Abstract
X-ray Absorption Spectroscopy Study of Prototype Chemical Systems: Theory vs. Experiment
by
Craig Philip Schwartz

Doctor of Philosophy in Chemistry
University of California, Berkeley

Professor Richard J. Saykally, Chair

Understanding the details of the intensities and spectral shapes of x-ray absorption spectra is a long-standing problem in chemistry and physics. Here, I present detailed studies of x-ray absorption for prototypical liquids, solids and gases with the goal of enhancing our general understanding of core-level spectroscopy via comparisons of modern theory and experiment.

In Chapter 2, I investigate the importance of quantum motions in the x-ray absorption spectra of simple gases. It is found that rare fluctuations in atomic positions can be a cause of features in the spectra of gaseous molecules.

In Chapter 3, I explore a novel quantization scheme for the excited and ground state potential surfaces for an isolated nitrogen molecule. This allows for the explicit calculation of the "correct" transition energies and peak widths (i.e. without any adjustable parameters).

In Chapter 4, the importance of nuclear motion in molecular solids is investigated for glycine. We find that the inclusion of these motions permits the spectrum to be accurately calculated without any additional adjustable parameters.

In Chapter 5, I provide a detailed study of the hydroxide ion solvated in water. There has been recent controversy as to how hydroxide is solvated, with two principal models invoked. I show that some of the computational evidence favoring one model of solvation over the other has been either previously obtained with inadequate precision or via a method that is systematically biased.

In Chapter 6, the measured and computed x-ray absorption spectra of pyrrole in both the gas phase and when solvated by water are compared. We are able to accurately predict the spectra in both cases.

In Chapter 7, the measured x-ray absorption of a series of highly charged cationic salts ($YBr_3$, $CrCl_3$, $SnCl_4$, $LaCl_3$ and $InCl_3$) solvated in water are presented and explained.

In Chapter 8, the measured x-ray absorption spectrum at the nitrogen K-edge of aqueous triglycine is presented, including effects of various salts which can alter its solubility. This is used to show that while x-ray absorption is sensitive to salt interactions with small peptides, it is unlikely to be a sensitive probe for overall protein structures, i.e. to distinguish beta sheet from an alpha helix at the nitrogen K-edge.

Finally, in Chapter 9 future directions are discussed.

1

For those who have come before me, and for those who will come after me.

# Contents

## Acknowledgements

So many people have contributed in various ways to this thesis that it will be almost impossible to thank everyone adequately.  Certainly none of this would have been possible without my advisor Rich, who always gave me the freedom to look at whatever problem I wanted to – even when it didn't coincide with what he necessarily would have liked me to be doing.  David wrote valuable tools and helped me interpret nearly all the data in this thesis.  Although I doubt he ever reads this document, his fingerprints are all over this work.  It is no small consolation to me that perhaps some of this work will help him receive the LBL equivalent of tenure.  The entire Saykally group, particularly Janel, Andrew, Alice, Walter and Jared helped with both much of the experiment and theoretical work in this thesis.  I enjoyed the chapter I wrote so much with Shervin, I put it in my thesis despite the fact that he did most of the writing.  It'd be great if we could continue this work someday in the future.   I would like to thank my friends, many of whom were previously mentioned, and family.  There is no way I could have gotten through this alone.  Finally, to you, the reader, although I personally wouldn't recommend reading this thesis as everything in here is likely to be published, and there are other things that I've published which aren't in here, I commend you on making your way through this, and I would point out that your dedication is truly impressive.

# Chapter 1 – Introduction

The absorption of x-ray photons by matter has long been known to be indicative of molecular structure. The first x-ray absorption spectrum of nitrogen gas was recorded in 1969 by photographic plate;[1] this first paper not only immediately linked the spectrum with the electronic structure of the nitrogen molecule, but also with its molecular structure. It is for this reason - the ability to link spectroscopy to both electronic and geometric structure - that near edge X-ray absorption fine structure (NEXAFS) spectroscopy has become an important technique in the characterization of matter in the gas, liquid and solid phases.[2]

The proliferation of synchrotron sources throughout the world has led to a precipitous growth in the number of high resolution NEXAFS studies of matter.[3] Unfortunately, the growth in experimental techniques has not been matched by a corresponding evolution of theoretical techniques, as accurately simulating the absorption event at energies of several hundred electron volts remains a difficult problem.[4]

The experimental study of core level transitions in gaseous molecules is a well established field.[5] A controlled amount of gas is leaked into the sample chamber and the synchrotron x-ray energy is tuned with a monochromator. Any of electrons, fluorescent photons, or direct absorption events can be used as the detection technique, depending on the desired information and experimentally accessible technology. The resolution of the experiment is usually limited by the Gaussian broadening of the X-ray source and the Lorentzian broadening of the sample due to lifetime effects.[6] With modern high resolution synchrotrons, total full widths of 130 meV at the nitrogen K-edge are common.[7]

The study of simple solids by NEXAFS has seen large advances in recent years.[8] The current state of the art involves grinding a solid powder into an indium film and collecting Auger electrons.[9] This allows for the minimizing of "space charging" effects caused by secondary electrons, as the indium film acts as an efficient electron scavenger, and by grinding the powder into a film, linear dichroism is minimized.[10] This has been used to efficiently study amino acids and many other forms of soft matter.[9] For single crystalline samples, such as water ice, the absorption measurements are less straightforward, and for this reason, much controversy still exists over the correct water ice NEXAFS spectrum.[11] A simple solution to this would be to attain high resolution X-ray Raman spectra (identical spectral information as NEXAFS) but as of this time, a high resolution X-ray Raman spectrum of water ice has not been measured.[12]

Recently, the experimental study of core level absorption spectra of liquids in the soft X-ray region has advanced dramatically, following the introduction of liquid microjet technology by Wilson *et al.* in 2001[13,14] which is the method employed in this thesis. This approach has the advantage of allowing windowless coupling to the high vacuum of a synchrotron. It also provides a convenient way to control the liquid temperature by evaporative cooling, as well as to avoid radiation damage. Recently, improvements in silicon-nitride X-ray cell technology have also advanced the sudy of liquids.[15] These allow for vaporless coupling to a synchrotron, have been purported to yield minimal sample damage problems when used as part of a flow cell (although some healthy skepticism of these claims may be appropriate) and allow for fast and convenient scanning of all relevant edges, including nitrogen.[16] Furthermore, windowed setups conveniently allow for the study of x-ray emission spectroscopy as well as detection of smaller concentrations of samples.[15,16] In contrast, liquid jets have been shown to allow for the study of

x-ray photoelectron spectroscopy.[17] In any case, soft X-ray NEXAFS spectra of volatile liquids are no longer prohibitive to obtain.

Despite the fact that all of these synchrotron measurements can be obtained relatively easily, the interpretation of these spectra remain challenging. A large cause of this is due to the difficulty in accurately simulating an x-ray absorption event for any large system. There have been a number of techniques developed over the years. Within density functional theory, it is very common for chemists to apply the half-core hole technique based on a formalism (the transition potential method) developed by Slater.[18] This technique works by removing half an electron from the core level to model the excitation process, and has shown reasonable results for isolated molecules. Physicists have often used the full-core hole model, where an entire electron is removed from the system.[19] This method has been shown to work well for metals. The method used throughout this thesis, called XCH, is based upon Fermi's Golden Rule, which treats the initial state as the electronic ground state and the excited state as the core excited state, with explicit treatment of the excited electron. This method can accurately predict the NEXAFS spectrum of liquid water within DFT without invoking strange and complex models for water, unlike the aforementioned methods.[20]

Other approaches have been explored. An efficient multiple scattering code has been developed which works well for both EXAFS and NEXAFS of heavy elements. Unfortunately, it is not accurate for atoms containing hydrogen, precluding its use from all the molecules studied herein, which all contain light elements.[21] Another approach which has garnered use is based upon the so called static exchange method, which freezes the excited molecular ion with respect to interaction with an excited electron, but allows for full accounting of exchange. This can shift the energy of excited states in a nonphysical way.[22] Techniques like the full core approximation have been applied within higher level theories such as coupled cluster and configuration interaction, with the expected increase in accuracy of using a higher level theory.[23] Finally, it should be noted that the most formally correct formalism for core-level spectra was recently developed. This method uses a complex polarization propagator to simulate the x-ray absorption event.[24] While this method is promising, it will be extremely difficult to implement on either a grid or in a complex plane due to issues related to grid spacing in order to accurately capture the core-hole. This essentially limits the method to finite Gaussian orbital methods at the current time, which is of limited utility. The Bethe-Salpeter equation, the gold standard for this type of problem, can be solved but due to poor scaling, it is infeasible to use for all but the smallest of systems.[25]

Independent of the method chosen, one must first choose molecular positions. Starting with the simplest systems, isolated molecules, the most commonly used method to simulate the molecular positions of a molecule is to minimize the nuclear forces of the Schrodinger equation.[26] Following an x-ray absorption calculation, the resultant stick spectra must then be artificially broadened. This unfortunately will not necessarily lead to an accurate lineshape and will miss transitions that are not allowed from this geometry.[27]

In order to avoid these problems, two different methods have been developed. As previously noted for liquids, motions can have a large spectral effect.[28] Applying this idea, an accurate quantum sampling was performed using path-integral molecular dynamics (PIMD) to obtain a quantum distribution of molecular positions, detailed in Chapter 2. This enables the calculation of the proper spectral shape of transitions. It leads to a noticeable improvement in

spectral quality for a few different examples that have been studies, glycine and *s*-triazine. This work was published in *J. Chem. Phys.*

The second method to account for motions, detailed in Chapter 3, is to quantize both the excited and ground state potential energy wells, and then simulate transitions. The transition intensity can be determined independently by separate quantum calculations. The inclusion of vibrational and rotational quantization allows for the explicit calculation of the "correct" stick spectrum, which can then be broadened to match experiment. This Franck-Condon simulation demonstrates the cause of spectral broadening found in NEXAFS spectra. This technique has applications in determining highly accurate excited state potential energy surfaces. This work was published in *J. Chem. Phys.*

We also note the importance of motions in molecular solids. Chapter 4 describes how motions within a glycine crystal lead to a NEXAFS spectrum without any sharp features. The underlying causes of the observed spectral broadening is detailed. It is shown that the spectrum in essence "grows in" as temperature is increased.

Chapter 5 describes a detailed investigation of the solvation of hydroxide in water. Based on NEXAFS measurements, it was recently argued that hydroxide was most probably hypercoordinated, in support of theoretical predictions to that effect.[29] The more detailed investigation of hydroxide performed herein shows that NEXAFS is actually not particularly sensitive to the solvation environment of hydroxide. There have been complimentary studies using photoelectron spectroscopy which also supported this prediction.[30] However, by carefully analyzing the projected density of states of hydroxide dissolved in water, we have determined that the photoelectron spectroscopy is not very sensitive to the details of the ion solvation. Hence, while hydroxide is likely hypercoordinated, the case is not as robust as previously declared.

In Chapter 6 the NEXAFS spectrum of pyrrole, a prototypical aromatic molecule, measured both in the gas phase and dissolved in water, is discussed. We accurately simulate these spectra, and the simulations are accurate enough that we can attribute a particular spectral feature to the auto-oligomerization of pyrrole in water. Interestingly, due to the weak interactions that are found between pyrrole and water, the solvated and gaseous spectra are almost identical. This work demonstrates the ability of the XCH method to accurately describe systems with delocalized bonding, which is problematic for competing theoretical methods. This was was published in *J. Chem. Phys.*

A study of the spectral effects that salts with highly charged cations ($YBr_3$, $CrCl_3$, $SnCl_4$, $LaCl_3$ and $InCl_3$) have on water is detailed in Chapter 7. These effects are found to be quite small, which is attributed to offsetting effects from the cations and anions. The data also help settle the debate as to the cause of some low energy features apparent in the spectra of certain salt water solutions.[29] The data presented here provide a strong indication that the cause of this feature is hydroxide formed by hydrolysis of the particular cations. This has been submitted to *Chem. Phys. Lett.*

Chapter 8 details the NEXAFS spectrum of the nitrogen K-edge of aqueous triglycine in the presence various salts which are known to either increase or decrease protein solubility.[31] It is clear that the salts directly interact with triglycine; the nature of these interactions are discussed in detail. We note that this work suggests NEXAFS will not be able to resolve peptide structures, e.g. beta sheet vs. alpha helix at the nitrogen K-edge.

Finally, in Chapter 9 future directions are discussed. A proposal of combining a x-ray water cluster experiment with detailed electronic structure calculations is detailed. If this proposal were successful it would help resolve the current controversy surrounding the nature of the hydrogen bond network of water.[32,33]

1    M. Nakamura, M. Sasanuma, S. Sato, M. Watanabe, Yamashit.H, Y. Iguchi, A. Ejiri, S. Nakai, Yamaguch.S, T. Sagawa, Y. Nakai, and T. Oshio, Physical Review **178** (1), 80 (1969).
2    J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).
3    T. H. Yoon, Applied Spectroscopy Reviews **44** (2), 91 (2009).
4    C. Witte, S. D. Findlay, M. P. Oxley, J. J. Rehr, and L. J. Allen, Physical Review B **80** (18) (2009).
5    A. P. Hitchcock, Journal of Electron Spectroscopy and Related Phenomena **112** (1-3), 9 (2000).
6    K. C. Prince, M. Vondracek, J. Karvonen, M. Coreno, R. Camilloni, L. Avaldi, and M. de Simone, Journal of Electron Spectroscopy and Related Phenomena **103**, 141 (1999).
7    B. W. Yates, Y. F. Hu, K. H. Tan, G. Retzlaff, R. G. Cavell, T. K. Sham, and G. M. Bancroft, Journal of Synchrotron Radiation **7**, 296 (2000).
8    Y. Zubavichus, A. Shaporenko, M. Grunze, and M. Zharnikov, Journal of Physical Chemistry B **112** (15), 4478 (2008).
9    Y. Zubavichus, M. Zharnikov, A. Schaporenko, and M. Grunze, Journal of Electron Spectroscopy and Related Phenomena **134** (1), 25 (2004).
10   S. G. Urquhart and R. Gillies, Journal of Chemical Physics **124** (23) (2006).
11   T. T. Fister, K. P. Nagle, F. D. Vila, G. T. Seidler, C. Hamner, J. O. Cross, and J. J. Rehr, Physical Review B **79** (17) (2009).
12   J. S. Tse, D. M. Shaw, D. D. Klug, S. Patchkovskii, G. Vanko, G. Monaco, and M. Krisch, Physical Review Letters **100** (9) (2008).
13   K. R. Wilson, B. S. Rude, T. Catalano, R. D. Schaller, J. G. Tobin, D. T. Co, and R. J. Saykally, Journal of Physical Chemistry B **105** (17), 3346 (2001).
14   K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, and R. J. Saykally, Review of Scientific Instruments **75** (3), 725 (2004).
15   O. Fuchs, M. Zharnikov, L. Weinhardt, M. Blum, M. Weigand, Y. Zubavichus, M. Bar, F. Maier, J. D. Denlinger, C. Heske, M. Grunze, and E. Umbach, Physical Review Letters **100** (2) (2008).
16   M. Blum, L. Weinhardt, O. Fuchs, M. Bar, Y. Zhang, M. Weigand, S. Krause, S. Pookpanratana, T. Hofmann, W. Yang, J. D. Denlinger, E. Umbach, and C. Heske, Review of Scientific Instruments **80** (12) (2009).
17   B. Winter, Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment **601** (1-2), 139 (2009).
18   A. Nilsson and L. G. M. Pettersson, Surface Science Reports **55** (2-5), 49 (2004).
19   R. L. C. Wang, H. J. Kreuzer, and M. Grunze, Physical Chemistry Chemical Physics **8** (41), 4744 (2006).
20   D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).
21   J. J. Rehr, J. J. Kas, M. P. Prange, A. P. Sorini, Y. Takimoto, and F. Vila, Comptes Rendus Physique **10** (6), 548 (2009).
22   V. Carravetta, O. Plashkevych, and H. Agren, Chemical Physics **263** (2-3), 231 (2001).
23   D. Duflot, J. P. Flament, A. Giuliani, J. Heinesch, M. Grogna, and M. J. Hubin-Franskin, Physical Review A **75** (5) (2007).
24   U. Ekstrom, P. Norman, V. Carravetta, and H. Agren, Physical Review Letters **97** (14) (2006).
25   J. J. Rehr, J. A. Soininen, and E. L. Shirley, Physica Scripta **T115**, 207 (2005).

26  I. Minkov, F. Gel'mukhanov, H. Agren, R. Friedlein, C. Suess, and W. R. Salaneck, Journal of Physical Chemistry A **109** (7), 1330 (2005).

27  J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).

28  J. D. Smith, C. D. Cappa, B. M. Messer, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry B **110** (40), 20038 (2006).

29  C. D. Cappa, J. D. Smith, B. M. Messer, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry A **111** (22), 4776 (2007).

30  E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, and B. Winter, Nature **455** (7209), 89 (2008).

31  Y. J. Zhang and P. S. Cremer, Current Opinion in Chemical Biology **10** (6), 658 (2006).

32  J. D. Smith, C. D. Cappa, K. R. Wilson, B. M. Messer, R. C. Cohen, and R. J. Saykally, Science **306** (5697), 851 (2004).

33  P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. A. Naslund, T. K. Hirsch, L. Ojamae, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, Science **304** (5673), 995 (2004).

# Chapter 2 -On the Importance of Nuclear Quantum Motions in Near Edge X-ray Absorption Fine Structure (NEXAFS) Spectroscopy of Molecules

## 1. Introduction

Core level spectroscopies are unique and powerful atom-specific probes of molecular interactions via both occupied and unoccupied electronic states.[1] As methods involving x-ray absorption (XAS, NEXAFS, XANES) or x-ray photo-electron spectroscopy (XPS) mature, they are increasingly being applied to complex molecular systems, including proteins, DNA, large organic molecules and polymers.[2] However, a major limitation is that deriving molecular information from these measurements usually depends explicitly on comparison with theoretical calculations, which are extremely difficult to perform at the accuracy of modern experiments. The accurate description of an absorption event of several hundred electron-volts of energy is an ongoing challenge in theoretical chemistry. Our work indicates that these calculations are extremely sensitive to the molecular geometries; therefore, in addition to an accurate theoretical formalism to describe the spectroscopy, the molecular geometries and their thermal fluctuations must be correctly sampled.[3] Herein, we describe the importance of quantum vibrational effects on core level excitations of the nitrogen K-edge of gas phase *s*-triazine and glycine. This is relevant to both near edge x-ray absorption fine structure (NEXAFS) and inner shell electron energy loss spectroscopy (ISEELS).

It has been shown previously that density functional theory (DFT) [4] can accurately reproduce excitation energies associated with core-level spectra via total energy differences ($\Delta$SCF or $\Delta$KS).[5] We use this to our advantage to model the core hole caused by the absorption of an x-ray photon; we represent the lowest energy core-level excited state self-consistently using a full electronic core hole on the excited atom and an associated screened excited electron (XCH).[6]

XCH differs from the closely related full core hole (FCH) approximation, which ignores the excited electron entirely or replaces it with a uniform background charge density (in the case of periodic systems),[7] and from the half core hole (HCH) approach, in which one removes half an electron from the system.[8,9] Another approach is an efficient cluster-based multiple scattering method, often used for heavier elements; this method is difficult to extend to lighter atoms, particularly those with a smaller atomic number than aluminum.[10] Shirley has developed accurate methods for solving the Bethe-Salpeter equation for crystalline solids, however, this high level of accuracy comes at significant computational cost.[11] Others have applied multi-electron quantum chemistry to simple systems, but generally these methods scale poorly with system size.[12] Another approach is static exchange, which freezes the orbitals of the molecular ion and calculates their exchange energy with the excited electron within a quantum calculation;[13] however this does not include the self-consistent relaxation of valence electrons due to the presence of the excited electron, which can have a significant impact on the relative energy and character of certain transitions. Recently, Ågren and coworkers have developed a polarization propagator approach, which has the significant advantage of not having to address specific atom-centered excited states explicitly.[14] While this method is extremely promising, an implementation using plane waves or a real space numerical grid is not yet available and may be

difficult due to a variety of issues related to accurately representing both localized core and extended valence states within the same computational framework.

Methods based on linear combinations of atomic orbitals (LCAOs) are highly successful in representing localized bound states.[5] However, LCAOs are limited in their ability to accurately describe delocalized scattering events, which become particularly important for electronic excitations to states above the ionization potential.[15] Thus, we have chosen to use plane waves, which are capable of approximating equally well both localized and scattering states.[3] We approximate the high energy continuous electronic density of states by exploiting the periodic boundary conditions of our supercell calculations and numerically converging an integration in k-space over the first Brillouin zone (BZ).[3]

All of these methods require a structural model for the molecules. In general, the most common approach is to calculate the lowest energy structure by holding the electrons in their quantum ground state while modeling the atomic nuclei as classical point charges, herein referred to as the fixed-nuclei approximation.[3,16] This is clearly a poor approximation if the molecule is not populated exclusively in its nuclear ground state, and for that reason, others have investigated the importance of multiple low-energy conformers.[17] Ågren and coworkers employ a more thorough approach, deriving excited-state potential energy surface gradients and using these for calculations of Franck-Condon factors.[18] This approach becomes prohibitively expensive for large molecules and impracticable for those with multiple low energy conformations or highly anharmonic motion. Furthermore, this approach neglects the impact of nuclear motion on the electronic transition amplitude; to first order, this impact is referred to as the Herzberg-Teller effect.[19] Attempts have been made to treat atomic movements directly in XAS by approximating thermal motions with Debye-Waller factors.[20] This assumes that all the molecules will move in harmonic potentials of width matched to experimental conditions; this will likely have limited validity for isolated molecules, particularly if they exhibit large conformational changes.

We have previously investigated improvements in simulating XAS using a classical trajectory, rather than using the usual fixed-nuclei approximation.[3] Large spectral changes were observed based on these classical thermal motions. Herein, we investigate the changes produced in simulated NEXAFS of molecules at the nitrogen K-edge when treating nuclear motions quantum mechanically by using path integral molecular dynamics (PIMD). These molecules are s-triazine, a relatively rigid prototypical ring structure, and glycine, the simplest amino acid. PIMD, based on the Feynman path integral formalism,[21] treats the atomic nuclei as quantum, rather than classical particles. We find that sampling quantum nuclear motions can induce large spectral changes in NEXAFS spectra, more accurately predicting certain features at a cost of some additional configurational sampling.

This is not the first work to note the importance of quantum fluctuations on electronic structure and associated spectroscopy. Earlier work simulated the gas phase valence electronic spectra of lithium clusters[22] and hydrazine.[23] In the hydrazine study, due to the efficiency of the algorithm employed and the use of localized basis set, extensive sampling was possible using an ab initio potential surface coupled with time dependent DFT.[23] Our work differs in that it explores the x-ray region of the spectrum where, as we shall see, localized basis sets are insufficient to describe high energy scattering states.

## 2. Methods

7

## X-ray absorption

We calculate the x-ray absorption cross section to first order using Fermi's golden rule:

$$\sigma(\omega) = 4\pi^2 \alpha_0 \hbar\omega \sum_f |M_{i\longrightarrow f}|^2 \delta(E_f - E_i - \hbar\omega) \qquad \textbf{Equation 1}$$

Here $\hbar\omega$ is the energy of the incident photon, which should match the difference in energy between the final and initial states $E_f - E_i$; $\alpha_0$ is the fine structure constant. $M_{i\rightarrow f}$ are the transition amplitudes between initial and final states, evaluated within the dipole approximation using a single-particle approximation

$$M_{i\longrightarrow f} = <\Psi_f | \hat{\varepsilon}\bullet R | \Psi_i > \approx S <\psi_f | \hat{\varepsilon}\bullet r | \psi_i >, \qquad \textbf{Equation 2}$$

where $\hat{\varepsilon}$ is the polarization direction of the photon electric field and $\left|\Psi_{i,f}\right\rangle$, $\mathbf{R}$ are the many-electron initial and final states and position operator, respectively, and their lower-case versions correspond to single-electron analogues. We approximate the prefactor $S \leq 1$ as constant over all single-particle transitions.[6] In this work, the single-particle eigenstates and eigenvalues are non-interacting Kohn-Sham states as computed by density functional theory within the generalized gradient approximation. We use the Perdew-Burke Ernzerhof (PBE) form of the generalized gradient approximation to the exchange-correlation potential.[24]

For x-ray core hole excitations at the nitrogen K-edge, the intial state $\left|\psi_i\right\rangle$ is fixed to the 1s atomic eigenstate of nitrogen (calculated for the ground state electronic configuration using PBE). For reasons which will become clear below, we adopt a plane-wave representation and pseudopotential approximation for valence electronic structure. The excited state was approximated to be spin-unpolarized. In all of our calculations, we use norm-conserving pseudopotentials with a numerically converged plane-wave cut-off of 85 Rydberg. These valence pseudostates are smooth, nodeless functions of position in the neighborhood of atomic nuclei, and consequently do not possess the correct overlap with particular atomic core states relevant to x-ray excitations. To fix this, we adopt a frozen-core approximation and augment the core-region of each valence state using projections of well-defined angular momentum about the atomic site of interest,[25,26] as follows:

$$\left\langle\psi_f\left|\hat{\varepsilon}\bullet\mathbf{r}\right|\psi_i\right\rangle \approx \left\langle\tilde{\psi}_f\left|\sum_\lambda\right|\beta_\lambda^{N_I}\right\rangle\left\langle\phi_\lambda^{N_I}\left|\hat{\varepsilon}\bullet\mathbf{r}\right|\phi_{1s}^{N_I}\right\rangle$$

$$\textbf{Equation 3}$$

In this expression, $\left|\tilde{\psi}_f\right\rangle$ is the pseudo (nodeless) final state and $\left|\phi_{1s}^{N_I}\right\rangle$ is the 1s atomic core state at the site of nitrogen atom $I$. We project the relevant core-region of $\left|\tilde{\psi}_f\right\rangle$ employing a sum of atomic projections $\left|\beta_\lambda^{N_I}\right\rangle$ with composite angular momentum index $\lambda = (l, m)$ and augment this core-region using associated all-electron (with nodes) atomic valence states $\left|\phi_\lambda^{N_I}\right\rangle$. In the case of 1s core excitations, this sum over projections is limited to dipole-allowed p-projections. This expression is computationally inexpensive to evaluate: the all-electron atomic dipole matrix elements are calculated on the same log-radial grid used in the pseudopotential generation and can be stored and reused for all nitrogen K-edge calculations within the given approximation to the DFT exchange-correlation potential; the angular momentum projections are just those used to calculate the non-local potential matrix elements of the Kohn-Sham DFT Hamiltonian within the

plane wave implementation. Note that for first-row elements, the pseudopotential p-channel is often chosen as local, and therefore has no associated projector within the norm-conserving Kleinman-Bylander formalism. In this case, the use of an ad hoc localized projector is necessary.[25]

### XCH Approach

To approximate the electronic final state within our Fermi's Golden Rule expression (1) we adopt the eXcited state Core Hole (XCH) approximation. The core-level excited state of the molecule is approximated by replacing the pseudopotential of the core-excited atom with one which explicitly includes a core-excitation – for nitrogen we use the electronic configuration: $1s^1 2s^2 2p^4$. We also include the important screening presence of the excited electron, by incrementing the number of ground state valence electrons by one. We then generate a self-consistent set of Kohn-Sham valence states in the presence of both of these perturbations. The atomic nuclei remain fixed in place as they will not move appreciably on the attosecond time scale of this excitation. The resulting constrained DFT ground state of this perturbed system is well-defined and approximates the first core-excited state. We further approximate higher excited states by using the unoccupied Kohn-Sham spectrum of this XCH self-consistent field. (We acknowledge that this approach may incorrectly describe the screening of excited states of different character than the first excited state, however, we welcome the convenience of an orthonormal set of final states provided by our approach.)

### Challenges for molecular systems – localized vs. continuum states, supercells.

The use of plane-wave basis sets to model the electronic structure of isolated molecules requires the use of large supercells within periodic boundary conditions. Large cells are used ($(20 \text{ Å})^3$) to reduce spurious interactions between cells and so as to be large enough to represent excited states below the ionization potential (IP).  In all cases except PIMD, approximately 100 Kohn-Sham eigenstates are used in constructing transition matrix elements, sufficient to extend the spectra approximately 3 eV above the estimated IP; for the PIMD calculations 250 Kohn-Sham eigenstates are used extending the spectra approximately 5 eV above the IP.  Due to the large box size, and hence reduced energy spacing between electronic bands, a large number of unoccupied states are needed to describe high-energy transitions.  The zone-center electronic structure is calculated using the PWSCF code.[27]

We take full advantage of the periodic boundary conditions to approximate the continuum of electronic states found at high energy by numerically converging an integration over the Brillouin Zone (BZ).  This should not affect states that are contained entirely within a given supercell.  For states which are larger than a given supercell, the electronic density of states can be reasonably determined by BZ sampling.[28]  These delocalized states are similar to the unbound electronic states scattered from the molecule.  The weakness of such a technique is in describing localized bound states lying below the ionization potential with spatial extents larger than the chosen supercell; we minimize this effect by using very large supercells.

In order to mitigate the computational cost of numerically converging the BZ integral, we utilize a technique developed by Shirley.[29]  In our implementation of this technique for molecules, we require only the electronic structure at the zone center (k=0) as input, yet we can generate eigenvalues and eigenstates to meV accuracy throughout the BZ within the energy range of the states provided at the zone center. This interpolation technique exploits the slow variation in the Bloch periodic part of the eigenfunctions $|u_{nk}\rangle = e^{-ik.r}|nk\rangle$ with respect to k.  An optimal basis is constructed from a coarse sampling of these $|u_{nk}\rangle$ across the BZ, diagonalizing

the overlap matrix of these functions and discarding those eigenvectors having very small eigenvalues, thereby removing linear dependence. The k-dependent Hamiltonian is expressed in this basis. For eigenstates with energies beyond those of the input coarse sampling of the BZ, we cannot guarantee the accuracy of this interpolated Hamiltonian. However, we notice some transferability to higher energies, and the resulting spectra at very high energies appear to improve with averaging over the thermodynamic ensemble (see results).

**Nuclear geometry considerations**

Generally, core-level spectra of isolated molecules are simulated within the fixed-nuclei approximation,[5] particularly for molecules in their vibrational ground state under experimental conditions. The lowest energy structure is computed treating the electronic structure quantum mechanically but modeling the atomic nuclei as fixed point charges. We performed this minimization on both the molecules studied here. For molecules not in their ground state, a Boltzmann weighted average of significant conformers is often used to attempt to generate the experimental spectrum.[17] We have calculated this for glycine. Unfortunately, as we will show this proves to be inadequate to properly describe the measured broadening. Structures which are not at significant conformational minima can have large spectral contributions.

In order to explicitly account for the motions caused by temperature, we have modeled the nuclear degrees of freedom in these molecules using molecular dynamics (MD) performed at 300 K with a Langevin thermostat utilizing the generalized AMBER force field and Antechamber. [30] The resulting distribution of nuclear coordinates is spaced at least 10 picoseconds apart to eliminate correlation between configurations for at least 100 configurations. The correlation was estimated at under 5 picoseconds for both molecules. Empirical, rather than *ab initio*, force fields were used in order to avoid the computational bottleneck associated with generating uncorrelated sampling. Due to the relative insensitivity of NEXAFS to bond lengths it is believed that this approximation of using empirical rather than *ab* initio should be relatively minor compared to various other approximations. Our tests indicate that the structural parameters derived from the AMBER potential agree with DFT results within ~2%.

**Quantum vs. Classical Distributions**

One problem with using classical MD is that this is based on Newton's classical equations of motion for the atomic nuclei, and should fail for quantum systems, in particular for a molecule like *s*-triazine which is largely in its vibrational ground state at the experimental temperature. This is evident from the vibrational spectrum of s-triazine shown in Table I, which was calculated using Gaussian 03.[31] Therefore, in addition to our classical MD simulations, we have performed path integral molecular dynamics (PIMD)[21] for both glycine and *s*-triazine, in order to generate a quantum distribution of states. Path integral molecular dynamics has been used extensively by groups in the past to simulate gases, liquids and solids.[22,23,32,33] Often large differences are found based upon using quantum descriptions of molecules versus those that treat molecules classically.[23,32] We note previous work has been done using PIMD to explain certain populations of glycine conformations.[34] This is not the first work to note the importance of quantum fluctuations on spectra for electronic transitions. Previous work has explored the importance of quantum motion to valence excitations.[22,23] In general, core hole excitations should be more sensitive to changes in structure or symmetry for small molecules.

Our PIMD simulations use the same classical potentials and the same Langevin thermostat used in the classical simulations, PIMD samples the nuclear quantum distribution using a chain of coupled "quantum beads;" we used 32 beads per atom at 300K, the same temperature as was used classically. Following a 1 nanosecond equilibration, the simulations were run for 10

nanoseconds.  As will be seen in the results, significant spectral deviations are seen between the calculated PIMD spectra and the calculated classical MD spectra.  These significant spectral deviations were also seen for glycine, which should not be in its nuclear ground state at the experimental temperature.  The calculated vibrational frequencies of glycine are provided in Table II.  The lowest vibrational frequency corresponds to less than 100K.

For *s*-triazine, four uncorrelated configurations were selected from a PIMD trajectory, and the electronic structure of all 32 bead sets were used as the starting geometries for electronic structure calculations.  This sampled a total of 128 nuclear geometries.  In the case of glycine, a single bead set was followed through 100 well-spaced configurations and used as the nuclear coordinates for DFT calculations.   Only one bead was used for glycine in an attempt to sample more of the conformational space; sampling all 32 bead sets (requiring 3200 spectra) would have been prohibitively expensive to compute.  Selecting 4 configurations and calculating all 32 bead sets in a manner similar to *s*-triazine would have provided much less sampling of conformational space.  The computational bottleneck of our approach is in the XCH calculations since generating 100 uncorrelated configurations from PIMD at intervals well beyond the auto-correlation time requires only on the order of 1% of the time of the 100 XCH calculations.  We note that the structural ensembles were converged by running the simulations for an extended period of time, but only ~100 configurations of those trajectories were sampled by XCH.   These XCH calculations were performed on the Franklin supercomputer at NERSC.

### Spectral alignment across configurations

These DFT/XCH calculations produce a set of Kohn-Sham eigenvalues and associated transition probabilities which may be numerically broadened to produce a smooth spectrum for a range of energies. However, the energy scale of the eigenvalues in these pseudopotential XCH calculations will be offset from the true energy scale of such core-level excitations in experiment and this must be adjusted in a systematic way. For the K edges of light elements, the energy of the measured absorption onset $E_{onset}^{\exp}$ is typically quite sharp and so, one might align the energy of the occupied LUMO state of the XCH with the experimental onset. This would correspond to shifting the energy scale as follows:

$$E \mapsto E - \varepsilon_{LUMO}^{XCH} + E_{onset}^{\exp}.$$

**Equation 4**

Such an alignment has been common practice for pseudopotential calculations. In this work we take a slightly different approach, choosing to align to the measured ionization potential ($IP_{\exp}$) instead. The purpose of such an alignment is to provide an unambiguous separation between bound and continuum states and to remove possible ambiguities arising from DFT's underestimation of bandwidths. Estimating the IP is also complicated when using pseudopotentials, and so, instead, we estimate the relative position of the absorption onset below the IP using the difference $E_{tot}^{FCH} - E_{tot}^{XCH}$. The full core hole approximation is used to model the core-ionized molecule. Now the resulting shift in the energy scale is

$$E \mapsto E - \varepsilon_{LUMO}^{XCH} + IP_{\exp} - (E_{tot}^{FCH} - E_{tot}^{XCH}).$$

**Equation 5**

All of this applies to an individual molecular configuration. However, if we wish to align the spectra of different molecular configurations (for the purpose of ensemble averaging) then we adopt the following procedure. One molecular configuration is chosen as a reference, denoted by

11

0 (in this work we chose the fixed-nuclei structure), and the shift of its eigenvalues derives from Eq. 5 above:

$$E(0) \mapsto E - \varepsilon_{LUMO}^{XCH}(0) + IP_{\exp} - [E_{tot}^{FCH}(0) - E_{tot}^{XCH}(0)] = E - \varepsilon_{LUMO}^{XCH}(0) + \Delta(0).$$

**Equation 6**

This assumes that the ionization energy of the fixed-nuclei structure sits at the same position as the peak in the experimental x-ray photoemission spectrum (nominally referred to as the IP). In general, the error associated with such an alignment (typically ~0.1eV) is less than the inherent underestimation of band widths in our DFT calculations. For particularly asymmetric XPS peaks, the alignment may require some adjustment.

For another molecular configuration (i), we must include a further shift relative to the reference (0), derived as follows: The core-excitation energy (in an all-electron formalism) is approximated as $E_{tot}^{XCH} - E_{tot}^{GS}$, the total energy difference between core-excited and ground states. The relative energy of excitations on different molecular configurations is

$$\Delta_r(i) = [E_{tot}^{XCH}(i) - E_{tot}^{GS}(i)] - [E_{tot}^{XCH}(0) - E_{tot}^{GS}(0)]$$
$$= [E_{tot}^{XCH}(i) - E_{tot}^{XCH}(0)] - [E_{tot}^{GS}(i) - E_{tot}^{GS}(0)].$$

**Equation 7**

Note that this rearrangement also enables us to compute meaningful energy differences within pseudopotential calculations and so, the ultimate shift applied to eigenvalues of configuration (i) is

$$E(i) \mapsto E - \varepsilon_{LUMO}^{XCH}(i) + \Delta_r(i) + \Delta(0).$$

**Equation 8**

Clearly, an analogous expression to Eq. 7 exists for providing the relative alignment of spectra coming from inequivalent atoms of the same species in a given molecular configuration. In this special case, the ground state energy differences are zero.

We wish to emphasize the importance of such shifts when combining spectra from a large set of molecular configurations to provide an ensemble average -- they are vital. Without them we obtain meaningless results. In particular, $\Delta_r(i)$ provides an accurate estimate of how well the environment/configuration can screen the core-excitation. This approach to energy alignment across molecular configurations has also been adopted in our previous work.[3,6] We note that the use of total energy differences in constructing relative alignment is compatible with the original ΔSCF motivation behind using the XCH approximation, which runs contrary to other approaches based on eigenvalue differences.

**Spectral impact of vibrations**

Variations in the nuclear degrees of freedom of molecules unsurprisingly have noticeable impacts on electronic structure in terms of shifting energy levels and modifying the symmetry of electronic states. In terms of the associated spectra, features may move in energy and change in intensity. We can analyze such effects within the Born-Oppenheimer approximation, assuming that electronic and nuclear states are decoupled:

$$|\Psi> = |\psi> \times |\chi> \qquad \textbf{Equation 9}$$

where $|\psi>$ is an electronic state and $|\chi>$ is a vibrational state. We explore the impact of small nuclear displacements along normal modes $Q_\nu$ on the transition amplitudes, $M_{i \to f}$ from initial state i to final state f, where

12

$$M_{i \longrightarrow f} = \mu_{n_i \longrightarrow n_f} < \chi_{v_i} \mid \chi_{v_f} > + \sum_{v=1}^{3N-6} \frac{\partial \mu_{n_i \longrightarrow n_f}}{\partial Q_v} < \chi_{v_i} \mid Q_v \mid \chi_{v_f} > + O(Q^2)$$

**Equation 10**

Here, $\mu_{n_i \longrightarrow n_f} = < \psi_{n_i} \mid r \mid \psi_{n_f} >$, corresponds to a dipole transition between an initial and final electronic electronic state; $< \chi_{v_i} \mid \chi_{v_f} >$ corresponds to the overlap of the initial and final vibrational modes. The modulation of the first term in this expansion, due to overlap of vibrational states on the ground and excited state potential energy surfaces, is the well-known Franck-Condon effect. The term linearly dependent on the normal coordinates $Q_v$ corresponds to the Herzberg-Teller effect.[19] The Franck-Condon approximation states that because the nuclei are so much more massive than electrons, an electronic transition takes place before the nuclei can respond. Therefore, the transition probability will rise with increasing spatial overlap between the initial and final state wavefunctions. The Herzberg-Teller effect expresses that the transition intensity can change with the coordinates. For a molecule that moves far from its minimum energy positions, such as glycine at room temperature, we find this effect to be spectrally significant. Clearly, for dipole forbidden transitions, the Franck-Condon term is negligible and, in the presence of symmetry-breaking nuclear displacements, the Herzberg-Teller term will dominate.

The central assumption behind the Born-Oppenheimer approximation implies that the time-scale of electronic transitions is much faster than nuclear motion, and for core-level excitations, we make the assumption that the probe can instantaneously sample the nuclear coordinates. In effect, the x-ray or electron beam takes configurations of a large ensemble of molecules, each with their nuclear degrees of freedom displaced to some degree from their mean positions.

We note that there are existing approaches which explore the vibrational degrees of freedom within a normal mode analysis, generating accurate Franck-Condon factors. However, such analysis neglects the Herzberg-Teller effect and higher order terms in Eq.10.

Our technique of generating a quantum distribution of conformations by MD sampling, and then using these displaced structures as the nuclear coordinates for electronic structure calculations, has advantages over other techniques. In particular with our approach, in the limit of infinite sampling, we expect to describe the Herzberg-Teller effect in addition to all the higher order effects.

All calculated transitions are numerically broadened using Gaussians of 0.2 eV full width at half maximum. We use this relatively small and uniform broadening with the aim of simulating and distinguishing electronic and vibrational effects explicitly. Certain features of spectra result from the fundamental energy dispersion of certain electronic states or the short lifetime of certain transitions, while other features will be caused by the motions of the molecules.[1,3] We believe that purposely using a small numerical broadening permits a predictive computational approach which can distinguish between electronic broadening and vibrational broadening of spectral features.

## 3. Results and Discussion

*S-triazine*

     *S*-triazine ($C_3N_3H_3$) is a small, naturally occurring prototypical aromatic molecule; it consists of a resonant ring, analogous to benzene, with alternate carbon atoms replaced with nitrogen atoms, giving it a high level of symmetry ($D_{3h}$). At room temperature, the molecule is localized largely in its ground vibrational state, see Table I. Figure 1 (a) shows the spectrum of *s*-triazine along with its calculated spectra at the nitrogen K-edge using various approaches to sampling the nuclear degrees of freedom, with Figure 1 (b) focusing on the energy range from 398 eV to 403 eV. The inner-shell electron energy loss spectroscopy (ISEELS) and NEXAFS spectra are taken from the literature, as is the measured experimental ionization potential.[12,16,35] The theoretical results are aligned to the experimental ionization potential, as described in section 2. The fixed nuclei and classical MD results have been shown previously.[3] The measured spectra comprise six features which have been previously assigned as follows: (1) a N(1s) to LUMO $\pi^*$ transition; (2) a shoulder to the first feature, also due to a $\pi^*$ transition, which is turned on by vibronic symmetry-breaking; (3) a feature which has previously been assigned to mixed Rydberg states; (4) a multiple electron peak, definitively assigned recently by Duflot *et al.*; (5) and (6) peaks believed to be $\sigma^*$ in character.[12,16,35] The energy positions of these 6 features are listed in Table III. We now provide an assessment of the accuracy of the various theoretical approaches employed in this work.

     The fixed nuclei spectrum approximation overestimates the intensity of feature 1, with features 2 and 4 missing, but obtains features 3 and possibly 5. Using classical MD lead to a decrease in the intensity of feature 1, in better agreement with experiment, while agreement for peaks 3 and 5 are improved. The shoulder (feature 2) is visible as an error bar, while features 4 and 6 are missing entirely. Classical molecular dynamics samples changes in atomic structure and the subsequent changes in electronic structure which, to first order, are described by the Herzberg-Teller effect. This is crucial in generating the second feature, which results from symmetry-breaking vibrational modes, as described in Ref. 4. Classical molecular dynamics does not adequately reproduce the intensity of feature 2. Furthermore, our calculations simulate single-electron transitions only, and so we would not expect to calculate feature 4 at all, which has been assigned definitively by Duflot *et al.* as a multielectron feature.[12,16,35]

     PIMD shows a substantial improvement in the ability to predict the *s*-triazine spectrum. The first peak is accurately reproduced in terms of relative intensity, and significantly, the second feature, previously assigned to vibronic effects, is clearly visible as a shoulder, in contrast to the classical MD sampled spectra where the second feature was seen only as a non-zero sampling error bar. The PIMD spectra gives the onset feature an asymmetry similar to that of experiment, including the long tail evident in the experiment at the edge of feature 2 (Fig. 1 (b)). Features 3 and 5 are in agreement with the previous computed spectra, and notably, feature 6 is also obtained. The improvement in agreement and the absorption of new features is due to the increased vibronic symmetry breaking associated with PIMD. The use of the classical AMBER potential together with PIMD reasonably samples the quantum distribution of states necessary to reproduce the spectrum, with all the single electron features accurately captured.

     Both classical MD and PIMD underestimate the energy width of all features due (we expect) to the underestimation of bandwidth typical in DFT approximations.[3] However, PIMD provides a much more general route to accurately obtaining spectra – at least for small molecules. In the particular case of the highly symmetric *s*-triazine molecule, there are clear selection rules for electronic transitions which may be broken by nuclear motion. A probe which

can induce such transitions (x-ray) on a time scale faster than the period of nuclear vibrations allows the indirect observation of these displacements via their associated symmetry breaking. In general, for systems with lower symmetry or heavier, less mobile nuclei, this effect may not be so noticeable in core level spectroscopy.

Shown in Figure 2 are the isosurfaces of some DFT Kohn-Sham eigenstates from the electron density of the fixed-nuclei geometry without and with a core hole and from a single configuration taken from the PIMD simulation without and with a core hole placed on the lowermost nitrogen atom of *s*-triazine respectively. There is a noticeable correspondence between the ground state and the core-excited electronic structure. In Ref. 3, we noted that the presence of nuclear displacements which break symmetry is sufficient to alter the electron density in the vicinity of the core excited nitrogen atom. Electronic transitions to states A and B, the LUMO and LUMO+1 respectively, clearly have $\pi^*$ symmetry and correspond to features 1 and 2 in the measured spectra. It is worth noting that the LUMO and LUMO+1 are almost degenerate in the electronic ground state; they are separated by less than 10 meV in our DFT calculations. The presence of the core excitation alone is enough to break this degeneracy by 0.7 eV.[3] The further inclusion of vibrations causes electron density to delocalize along the C-N bond, resulting in a reduced intensity of feature 1. Feature 2 appears due to the delocalization of electron density from the neighboring carbon to the core excited nitrogen; this is only enabled by symmetry breaking nuclear motion.[3] Comparing the fixed-nuclei and PIMD electron density when both have a core hole shows that the motion caused by PIMD causes the electron density to shift onto the nitrogen. Several transitions occur in the region of feature 3, one of which is shown in Figure 2C. This particular transition has some $\sigma^*$ characteristics, but other transitions in the same energy range involve the entire *s*-triazine ring;[3] no simple assignment can be made for this feature as it is due to many different transions.[16,35] One benefit of plane waves is their ability to describe scattering states, such as one of the transitions contributing to feature 5, Figure 2D. Accurately describing such a state using a localized basis set would be problematic.

It has been shown previously that symmetry-breaking can lead to significant spectral changes, in particular, the allowing of previously forbidden transitions.[3] In *s*-triazine, different forms of symmetry breaking can allow different transitions. The large symmetry change caused by the local core excitation lowers the molecular symmetry from $D_{3h}$ to $C_{2v}$. As described in Section 2, the additional symmetry-breaking nuclear motion leads to changes in transition amplitudes via the Herzberg-Teller effect.[19] We monitor two such nuclear coordinates here.

We measure changes of in-plane symmetry for *s*-triazine using the difference between the absolute value of two carbon-nitrogen-carbon bond angles (labeled $\theta_L$ and $\theta_R$ in Figure 3) for one specific nitrogen atom sampled from a long trajectory. Nonzero values correspond to symmetry-breaking. When using classical MD or PIMD, the variance of this measure grows from 5.4° to 14.5°. The higher variance corresponds to more symmetry breaking in the case of PIMD, as compared to the fixed nuclei case, which is symmetric in-plane.

Figure 4 shows another sampled symmetry breaking nuclear coordinate in *s*-triazine – out of plane deformations, measured by tracking the carbon-nitrogen-carbon-hydrogen dihedral deformation; an angle of zero degrees corresponds to a flat molecule. It is clear from the Figure 4 that the amplitude of out of plane deformations increases from classical MD to PIMD. The PIMD sampling of larger nuclear deformations results most noticeably in increased spectral intensity for feature 2 (which is entirely caused by nuclear motion), and generally improves agreement with experiment.

It is important to mention that the error bars shown on all calculated spectra correspond to a standard deviation of the sampled spectra, but this does not correspond to the error in the experiment.  The experimental uncertainty derives from a spatial average of the number of instantaneous molecular conformations captured within the x-ray beam, followed by a time average over the duration of the exposure to the beam (both intrinsic uncertainties), followed by repeated measurements to overcome instrument noise.  Therefore, all the variance associated with the molecules being in different conformations is encompassed in the huge spatial and temporal sampling per data point obtained experimentally.  This sampling is what we attempted to approximate by using molecular dynamics in this work.  What we find is that such sampling is indeed necessary, given the marked differences between spectra computed from the fixed nuclei approximation and using PIMD, and the improved agreement with experiment resulting from using PIMD. In summary, the fluctuations in molecular conformation dominate the calculated spectrum, and by inference, the experimental spectrum.

**Glycine**

Glycine, ($NH_2CH_2COOH$), the simplest amino acid, is vibrationally active at the experimental temperature, as can be determined from the vibrational spectrum provided in Table II.  Even though the four energetically most stable conformers determined by DFT fix the five heavy atoms in the same plane, glycine has a variety of "soft" modes that enable the nitrogen-containing amine group to rotate relative to the carboxyl group;[17] therefore conformations besides the four most dominant conformers will be populated.   The populations of all conformations were monitored by tracking variations in the nitrogen-carbon-carbon-oxygen dihedral angle (the double bonded oxygen in this case), shown in Figure 5.  The four most stable conformers all have dihedral angles of either 0 or 180; there is significant population found at other angles, for both classical MD and PIMD.  The classical distribution is more sharply peaked than PIMD.  Overall, the relative change in distributions between the PIMD and classical MD simulations appears to be smaller for glycine than for *s*-triazine, likely because glycine is not predominantly in its ground state at the sampled temperature.

The measured NEXAFS and ISEELS spectra of gaseous glycine evaporated at ~415°C are shown in Figure 6, along with calculated spectra at the nitrogen K-edge using the different approaches to sample nuclear degrees of freedom; the experimental data are taken from the literature.[17,36]  The theoretical results are aligned with the experimental IP as described in Section 2.   The NEXAFS spectrum of glycine consists of four features which have been assigned previously as follows: (1) a well resolved N(1s) to LUMO feature assigned as $\sigma^*_{NH}$; (2) an easily distinguished feature assigned as $\pi^*_{NC}$; (3) a feature assigned as $\sigma^*_{NC}$; (4) a slightly more intense feature than feature 3 believed to be a Rydberg feature.[17,36]  These assignments are likely an oversimplification, as it has been shown that motions in the molecule can cause these features to shift in energy and blend together.[3]

We will now discuss the accuracy of the theoretical approaches employed in this work for glycine.  The lowest energy conformation (with fixed nuclei) produces the correct number of features with a similar intensity ratio as found in experiment; it is not possible to homogeneously broaden the spectrum to match experiment.  This is almost certainly due to the lack of certain molecular conformations at the given experimental temperature.  By Boltzmann weighting the four most populated conformers of glycine, one can obtain a spectrum that generates features 2, 3 and 4.[17]  Feature 1 is split into two smaller features, once again likely due to a lack of sampling of less populated conformations.  It appears that in order to accurately reproduce the spectrum of

glycine, it is necessary to sample a wider variety of conformations. As shown previously, classical MD produces the correct number of features when a wider linewidth is used.[3] However, the relative intensities of features 3 and 4 are not well reproduced, nor is the correct spectral lineshape obtained. While this spectra is likely not fully converged, it appears close to convergence.

It was hoped that by using a quantum distribution of states generated by PIMD, a more accurate representation of the spectra would be obtained. While features 2, 3 and 4 are indeed more accurately reproduced than by any other method in terms of feature intensity and width, the first feature is not resolved. The absence of the first feature cannot be attributed to the broadening scheme: if the broadening is narrowed further, too many features are obtained in this energy range and if broadening is increased, the whole region is flattened. We can only speculate that additional sampling would produce a more accurate spectrum, including feature 1. Unfortunately, it is not possible to converge the statistics of glycine when utilizing PIMD with only 100 configurations. Due to the wider variation of molecular configurations, it will take more configurations to converge a PIMD-sampled spectra than it does to converge a classical MD sampled spectra; this is evidenced by the larger spectral standard deviation associated with PIMD compared with that of classical MD. PIMD requires at least several hundred separate uncorrelated conformations to accurately sample the molecule, which, for now, is computationally inaccessible using our technique. We note that the classical MD spectra exhibits less variance than the PIMD spectra, which means it should take less sampling to converge.

Using classical MD to sample molecular configurations does not yield predictive results. It appears imperative to include quantum motion to accurately predict spectra, implying that a careful balance of adequate sampling and computational cost is required. One possible solution is that localized states, such as those producing features 1 and 2 could be accurately calculated using localized basis sets, significantly reducing the computational cost. We note that in going to more complicated systems, such as solvated systems, the ratio of the number of configurations needed to converge PIMD relative to the number of configurations needed to converge classical MD may decrease due to a lack of sharp features in the spectra of larger or more complicated systems.

If one were to estimate the lifetime broadening for bound states of nitrogen to be ~0.1 eV,[1] it quickly becomes apparent that nuclear motions are a large, and in certain cases, dominant cause of peak widths in molecular spectra. Note that the fixed-nuclei spectra are analogous to lifetime broadened spectra while neglecting nuclear motion. For glycine this is clearly not broad enough. In the case of *s*-triazine the first peak is too narrow, but we point out that there are Frank-Condon vibrational transitions which are also significant but our fixed-nuclei spectra will miss. It appears that most of the width of many of the bound peaks are caused by motion effects rather than lifetime effects. Lifetime widths only cause a fraction of the total width of any given bound peak, although that can change from a small fraction to a large fraction even within the same molecule (see features 1 and 4 of *s*-triazine). For unbound transitions, the distinction is less clear due to a less well known lifetime width, and likely larger lifetime broadening.

The conventional use of lifetime effects as an explanation of selective broadening appears to be an oversimplification. In certain cases, it is possible to broaden the spectra to obtain the correct general shape, but in other cases it simply is not possible, and nuclear motions must be invoked. In general, to accurately predict the spectral intensity and lineshape for light atoms it appears necessary to include both the effects of motion and lifetime. Our calculations on *s*-

17

triazine which do not calculate lifetime broadening appear to do a reasonable job of predicting the spectra, leading us to speculate that motion effects are more important in general than an accurate knowledge of the lifetime broadening.

Molecular dynamics sampling is important for NEXAFS calculations because the time length of the x-ray probe is ultrashort, ca. tens of attoseconds. This is fast enough to "freeze out" the nuclear motions of the molecule. Subtle changes in molecular geometry evolve rapidly but x-rays are fast enough to be sensitive to them. Other techniques with a longer measurement time scale (e.g vibrational spectroscopy) will average over molecular positions, meaning only part of the interaction time between the molecule and a photon will occur while the atoms are maximally displaced. Therefore, those techniques may not sample the strongest deviations from the mean positions.

## 4. Conclusions

The effect of including a quantum distribution of states in simulating the NEXAFS spectra of two isolated molecules was investigated. For *s*-triazine, a small highly symmetric aromatic molecule largely localized in its ground state, the changes observed by using a quantum distribution of states help explain the observed spectrum, in particular the asymmetry in observed features which could not be adequately explained by either classical MD sampling or fixed-nuclei data. The broadening of features appears, in large extent, to be driven by these vibrational motions, and in certain cases the vibrations are responsible for the observation of these features. This demonstrates the importance of both the Herzberg-Teller effect and of having a proper quantum distribution in order to accurately predict spectra of small molecules composed of light elements. This approach finds a much greater challenge with molecules that access a larger number of vibrational states; this is due to the difficulty in sampling an adequate number of conformations, as computational cost rises linearly with the number of configurations. Regardless, this accuracy is necessary for properly accounting for certain spectral features and indicates the hazards of drawing anything other than qualitative conclusions from calculations based on a single geometry, even for molecules in their ground vibrational state. As X-ray spectroscopy experiments continue to evolve and measurements become increasingly more detailed, it is important to realize that these vibrational motions may comprise a natural limit to the resolution of NEXAFS and cannot be ignored. Thus, the ability to interpret these underlying vibrational motions will likely become increasingly important.

**Figure and Table Captions**

**Table I.** Calculated vibrational modes of *s*-triazine listed in frequency ($cm^{-1}$). The molecular structure was optimized and then the vibrational modes were calculated, using the B3LYP exchange-correlation functional and a 6-311G* basis set. We note that the lowest energy mode corresponds to ~ 450 K.

**Table II.** Calculated vibrational modes of glycine listed in both frequency ($cm^{-1}$). The molecular structure was optimized and then the vibrational modes were calculated, using the B3LYP exchange-correlation functional and a 6-311G* basis set. The lowest energy mode corresponds to 92.97 K, well below experimental temperature.

**Table III.** Summary of transition energies (eV) obtained by various methods for *s*-triazine.

**Figure 1.** Measured and simulated core-level spectra of s-triazine (indicated top right) at the nitrogen K edge, offset vertically for clarity: (a) from top top to bottom: measured NEXAFS (blue - solid) and ISEELS (yellow - dotted) spectra in comparison with the calculated spectra using fixed nuclei (red), classical MD (black with grey error bars) and PIMD (purple with error bars). The average spectra for classical MD and PIMD are shown in darker colors with a shaded width of one standard deviation. The experimental spectra are taken from the literature and the vertical line is the experimental ionization potential (see text). (b) An enlargement of the spectral energy range 398-403 eV for PIMD, classical MD and NEXAFS data. The PIMD results reproduce a shoulder around 400eV, as found experimentally, unlike the classical MD results.

**Figure 2.** Isosurfaces of several electronic states of s-triazine from the fixed nuclei structure without and with a core hole on the bottom most nitrogen and from a single PIMD configuration without and with a nitrogen 1s core excitation respectively. Positive and negative phases are indicated in red and green respectively. The excited nitrogen is always on the lowermost (blue) atom of the selected configuration. State A corresponds to spectral feature 1, state B to spectral feature 2, state C is representative of spectral feature 3 and state D is a scattering state corresponding to feature 5. Further details are given in the text.

**Figure 3.** A plot of the difference in s-triazine CNC bond angles, with the two relevant angles labeled in red ($\theta_L$ and $\theta_R$), sampled from the classical (blue, solid) and PIMD (red, dashed) distributions. The PIMD data shows a significantly broader distribution than that of classical MD. The errors are of the size of the width of the lines.

**Figure 4.** A plot of s-triazine C-N-C-H dihedral angles, with relevant atoms labeled (1-4), sampled from classical (blue, solid) and PIMD (red, dashed) distributions. The PIMD data shows a significantly broader distribution than that of classical MD. The errors are of the size of the width of the lines.

**Figure 5.** A plot of the glycine N-C-C=O dihedral angle, with the relevant atoms labeled (1-4), sampled from classical (blue, solid) and PIMD (red, dashed) distributions. The PIMD data shows

19

a slightly broader distribution than that of classical MD. The errors are of the size of the width of the lines. At -180° PIMD is more intense than classical.

**Figure 6.** Measured and simulated core-level spectra of glycine at the nitrogen K-edge (from top to bottom): Measured NEXAFS (blue solid) and ISEELS (yellow, dash-dot) spectra in comparison with the calculated spectra using fixed nuclei (red), a Boltzmann-weighted average from the four lowest energy conformers (green), classical MD (black with grey error bars) and PIMD (purple with error bars). The average spectra for classical MD and PIMD are shown in darker colors with a shaded width of one standard deviation. The experimental spectra are taken from the literature and the vertical line is the experimental ionization potential (see text).

**Figures and Tables**

Vibrational modes of *s*-triazine

| Index | Frequency (cm-1) |
|-------|------------------|
| 1 | 309.20 |
| 2 | 310.33 |
| 3 | 665.46 |
| 4 | 665.58 |
| 5 | 734.48 |
| 6 | 892.15 |
| 7 | 972.71 |
| 8 | 985.03 |
| 9 | 985.21 |
| 10 | 1115.53 |
| 11 | 1156.60 |
| 12 | 1157.04 |
| 13 | 1162.83 |
| 14 | 1349.93 |
| 15 | 1385.86 |
| 16 | 1386.23 |
| 17 | 1530.57 |
| 18 | 1531.12 |
| 19 | 3064.39 |
| 20 | 3064.69 |
| 21 | 3070.13 |

**Table I**

Vibrational modes of glycine

| Index | Frequency (cm$^{-1}$) |
|-------|----------------------|
| 1 | 64.09 |
| 2 | 313.67 |
| 3 | 338.63 |
| 4 | 483.16 |
| 5 | 514.51 |
| 6 | 628.73 |
| 7 | 760.79 |
| 8 | 8357.35 |
| 9 | 889.82 |
| 10 | 1007.97 |
| 11 | 1042.23 |
| 12 | 1127.05 |
| 13 | 1144.61 |
| 14 | 1303.78 |
| 15 | 1304.06 |
| 16 | 1385.88 |
| 17 | 1449.56 |
| 18 | 1649.02 |
| 19 | 1726.91 |
| 20 | 2799.60 |
| 21 | 2994.08 |
| 22 | 3049.39 |
| 23 | 3438.68 |
| 24 | 3550.19 |

**Table II**

*s*-triazine feature locations - energy (eV)

| Feature number | NEXAFS | Fixed Nuclei | Classical MD | PIMD |
|---|---|---|---|---|
| 1 | 398.9 | 399.2 | 399.1 | 398.9 |
| 2 | 399.9 | | ~399.4 | ~399.8 |
| 3 | 402.9 | ~403.2 | ~403.1 | ~403.1 |
| 4 | 405.2 | | | |
| 5 | 409.0 | ~408.7 | ~408.2 | ~408.0 |
| 6 | 415.3 | | ~413.5 | ~413.7 |

**Table III**

**a)**

**b)**

**Figure 1**

**Vibrationless**                      **PIMD**

No Core Hole    Core Hole    No Core Hole    Core Hole
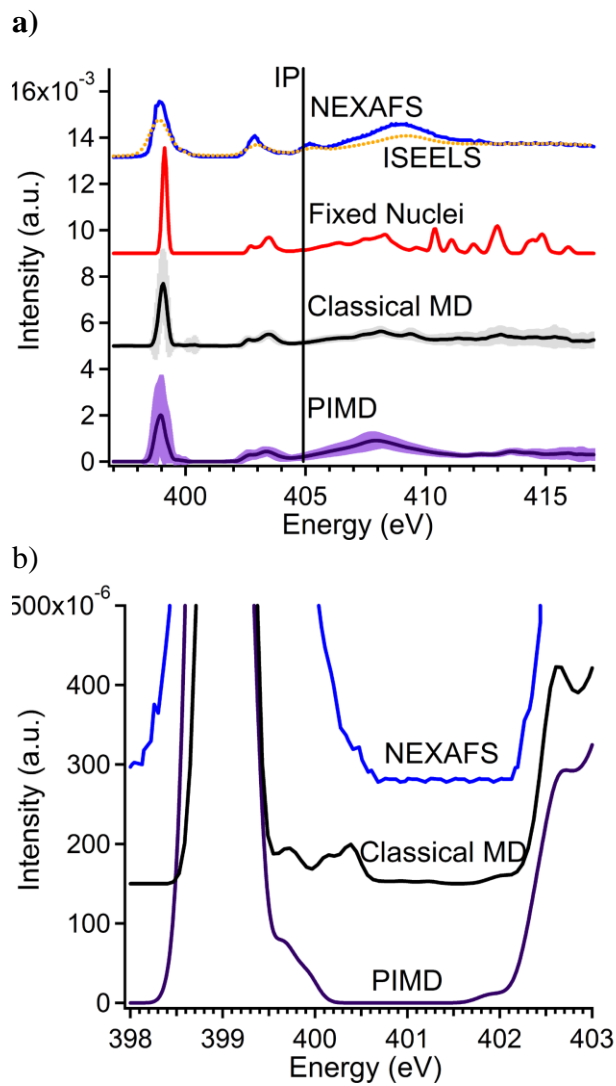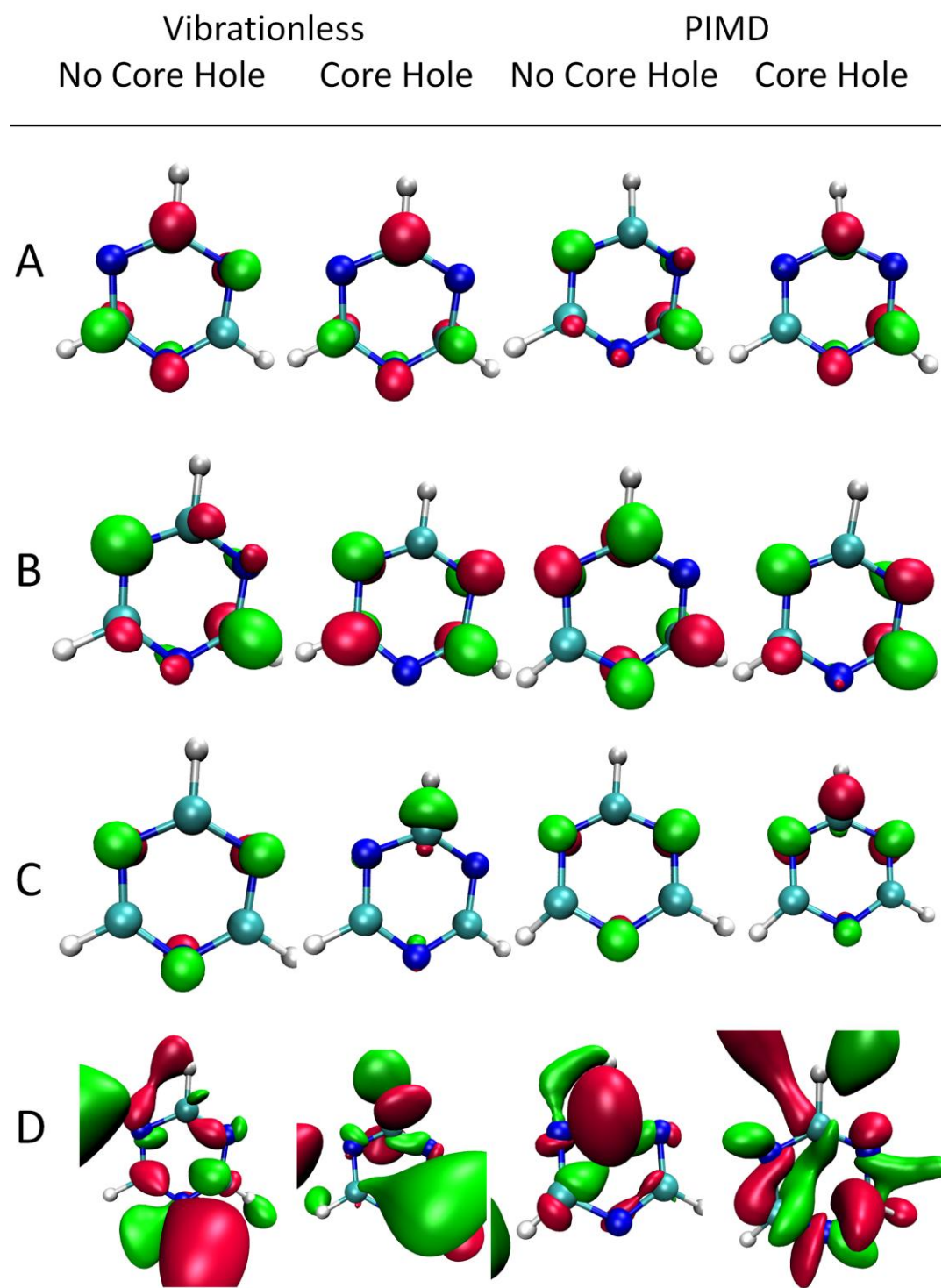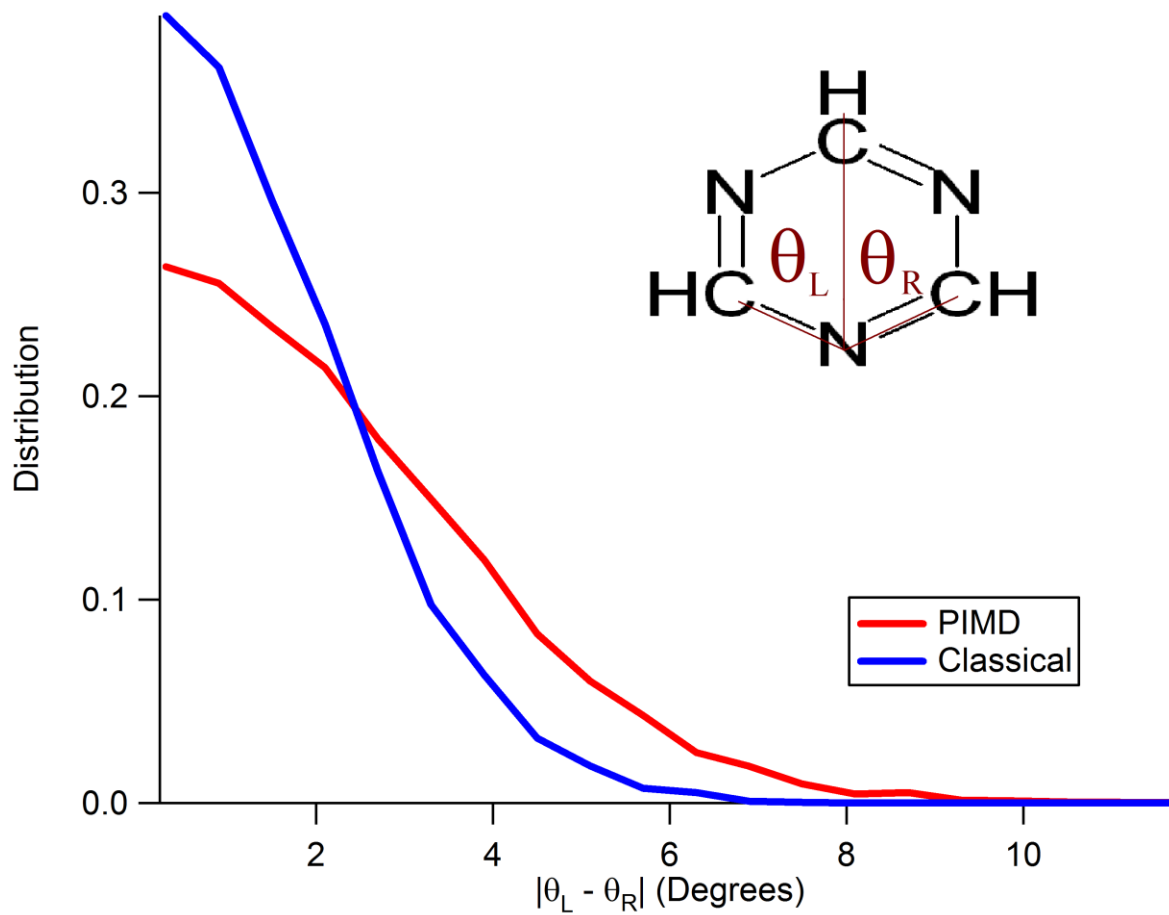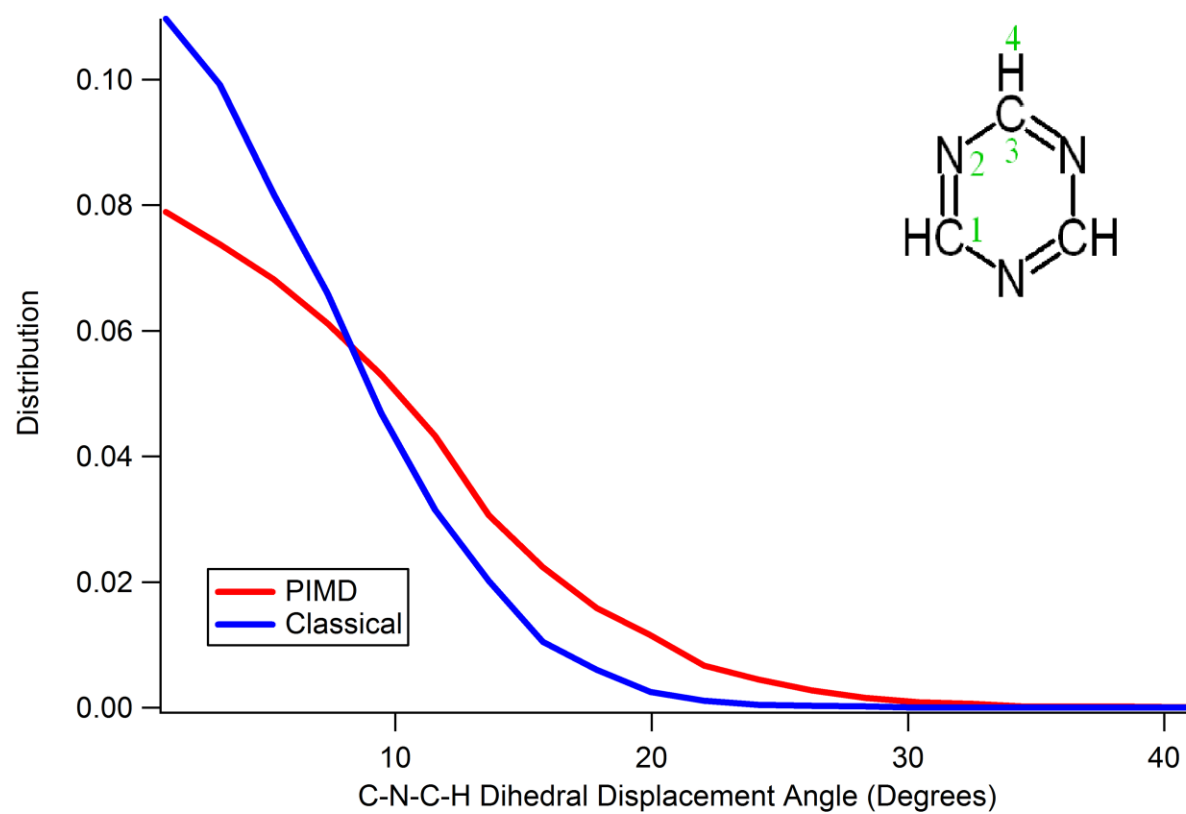
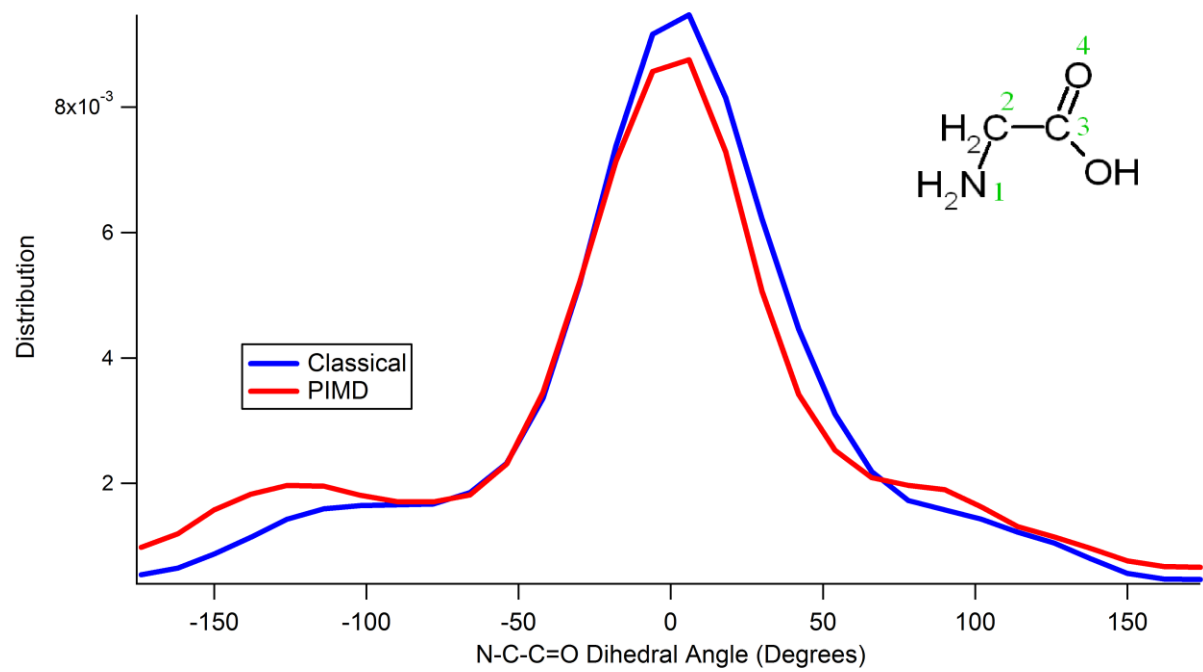**Figure 2**

**Figure 3**

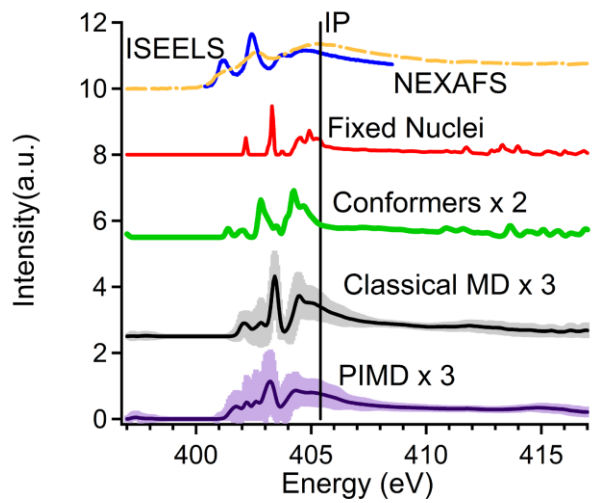**Figure 4**

**Figure 5**

**Figure 6**

1       J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).
2       J. Stewart-Ornstein, A. P. Hitchcock, D. H. Cruz, P. Henklein, J. Overhage, K. Hilpert, J. D. Hale, and R. E. W. Hancock, Journal of Physical Chemistry B **111** (26), 7691 (2007); H. Ade and A. P. Hitchcock, Polymer **49** (3), 643 (2008).
3       J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).
4       P. Hohenberg and W. Kohn, Physical Review B **136** (3B), B864 (1964); W. Kohn and L. J. Sham, Physical Review **140** (4A), 1133 (1965).
5       C. Kolczewski, R. Puttner, O. Plashkevych, H. Agren, V. Staemmler, M. Martins, G. Snell, A. S. Schlachter, M. Sant'Anna, G. Kaindl, and L. G. M. Pettersson, Journal of Chemical Physics **115** (14), 6426 (2001).
6       D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).
7       Y. Luo, H. Agren, M. Keil, R. Friedlein, and W. R. Salaneck, Chemical Physics Letters **337** (1-3), 176 (2001).
8       L. Triguero, L. G. M. Pettersson, and H. Agren, Physical Review B **58** (12), 8097 (1998).
9       A. Nilsson and L. G. M. Pettersson, Surface Science Reports **55** (2-5), 49 (2004).
10      J. J. Rehr and A. L. Ankudinov, Journal of Synchrotron Radiation **10**, 43 (2003).
11      E. L. Shirley, Physical Review Letters **80** (4), 794 (1998).
12      D. Duflot, K. Sidhoum, J. P. Flament, A. Giuliani, J. Heinesch, and M. J. Hubin-Franskin, European Physical Journal D **35** (2), 239 (2005).
13      H. Agren, V. Carravetta, O. Vahtras, and L. G. M. Pettersson, Chemical Physics Letters **222** (1-2), 75 (1994).
14      U. Ekstrom, P. Norman, V. Carravetta, and H. Agren, Physical Review Letters **97** (14) (2006).
15      A. L. Ankudinov and J. J. Rehr, Journal of Synchrotron Radiation **10**, 366 (2003).
16      G. Vall-llosera, B. Gao, A. Kivimaeki, M. Coreno, J. A. Ruiz, M. de Simone, H. Agren, and E. Rachlew, Journal of Chemical Physics **128** (4) (2008).
17      M. L. Gordon, G. Cooper, C. Morin, T. Araki, C. C. Turci, K. Kaznatcheev, and A. P. Hitchcock, Journal of Physical Chemistry A **107** (32), 6144 (2003).
18      I. Minkov, F. Gel'mukhanov, H. Agren, R. Friedlein, C. Suess, and W. R. Salaneck, Journal of Physical Chemistry A **109** (7), 1330 (2005).
19      G. Herzberg and E. Teller, Zeitschrift Fur Physikalische Chemie-Abteilung B-Chemie Der Elementarprozesse Aufbau Der Materie **21** (5/6), 410 (1933).
20      F. D. Vila, J. J. Rehr, H. H. Rossner, and H. J. Krappe, Physical Review B **76** (1) (2007); H. H. Rossner, D. Schmitz, P. Imperia, H. J. Krappe, and J. J. Rehr, Physical Review B **74** (13) (2006).
21      C. Chakravarty, International Reviews in Physical Chemistry **16** (4), 421 (1997).
22      F. Della Sala, R. Rousseau, A. Gorling, and D. Marx, Physical Review Letters **92** (18) (2004).
23      A. Kaczmarek, M. Shiga, and D. Marx, The Journal of Physical Chemistry A **113** (10), 1985 (2009).
24      J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77** (18), 3865 (1996).
25      P. E. Blochl, Physical Review B **50** (24), 17953 (1994).

26     B. Hetenyi, F. De Angelis, P. Giannozzi, and R. Car, Journal of Chemical Physics **120** (18), 8632 (2004).

27     S. Baroni, A. D. Corso, S. D. Gironcoli, and P. Giannozzi, PWSCF (2008).

28     D. Prendergast, J. C. Grossman, and G. Galli, Journal of Chemical Physics **123** (1) (2005).

29     E. L. Shirley, Physical Review B **54** (23), 16464 (1996).

30     T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., K. M. M. Luo, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S., A. R. Hayik, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H., L. Y. Gohlke, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C., and W. S. R. Schafmeister, and P.A. Kollman, (2006); J. M. Wang, W. Wang, P. A. Kollman, and D. A. Case, Journal of Molecular Graphics & Modelling **25** (2), 247 (2006).

31     M. J. T. Frisch, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.;Pople, J. A.;, Gaussian 03 (Gaussian, Inc., Wallingford, CT, 2004).

32     R. Ramirez, C. P. Herrero, and E. R. Hernandez, Physical Review B **73** (24) (2006); D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature **397** (6720), 601 (1999).

33     S. Raugei and M. L. Klein, Chemphyschem **5** (10), 1569 (2004); U. W. Schmitt and G. A. Voth, Journal of Chemical Physics **111** (20), 9361 (1999).

34     T. F. Miller and D. C. Clary, Physical Chemistry Chemical Physics **6** (10), 2563 (2004).

35     E. Apen, A. P. Hitchcock, and J. L. Gland, Journal of Physical Chemistry **97** (26), 6859 (1993).

36     O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, and V. Carravetta, Journal of Electron Spectroscopy and Related Phenomena **155** (1-3), 47 (2007).

# Chapter 3 - Nuclear quantum effects in the structure and lineshapes of the N$_2$ NEXAFS spectrum

## 1. Introduction

Core-level spectroscopies were originally utilized as structural probes of simple gases. As methods and theory have matured, they have been applied to increasingly complex systems ranging from condensed phases of single compounds to complex mixtures of polymers and biomolecules.[1,2,3] In the case of these more complicated systems, the information embedded in the resulting spectra includes both structure and signatures of molecular interactions. Even when one focuses solely on a large molecule in the gas phase – thereby eliminating the need to account for such environmental effects correctly – the computational costs associated with predicting the spectrum via many current (electronic structure) methods become prohibitive. These difficulties have provided an impetus for development of innovative methods that scale more reasonably with system size. In the process, we have shown that proper sampling of quantum-mechanically allowed geometries can significantly improve the quality of the predicted spectra.[4,5]

The present work, while built upon the same methodology, addresses a more fundamental question than whether it is possible to make such spectral predictions relatively cheaply. Instead, we aim to underscore the quality of our method by determining the upper limits of its performance when essentially all sources of error, apart from the electronic structure method itself, have been minimized. Our test case is nitrogen gas (N$_2$), a simple, relatively well-understood molecular system.[6-14] An additional advantage of this choice is that we are able, in the process of developing increasingly more realistic spectral models, to describe in exquisite detail the observable consequences of nuclear quantization and to make concrete statements about the origins of spectral structure and broadening.

We begin by outlining standard approaches to the problem. In calculating the gas-phase NEXAFS spectrum, a discrete set of transition energies and associated intensities is generated; this "stick spectrum" can then be broadened to match experiment.[5,15] Broadening is attributed to a combination of lifetime effects – energetic smearing of the molecular states due to the energy-time uncertainty principle -- and experimental sources such as a finite monochromator width.[1] There appears to be no consensus in the literature as to how much broadening to apply to a spectrum once calculated, as a variety of linewidth schemes are often used.[5,15] In the past, based on experimental fitting of the nitrogen spectrum to a Voigt line shape – a convolution of the Lorentzian and Gaussian profiles associated with lifetime and instrument broadening that is often approximated as their linear combination -- the width was determined to be approximately 120 meV FWHM (full width at half maximum) Lorentzian broadening and an experiment-dependent amount of Gaussian broadening.[6-14] Because the Voigt profile does not support a unique partition into contributions from Lorentzian and Gaussian broadening, many choices of these widths may yield a similar spectrum. Previous work by Coville and Thomas determined lifetimes for a variety of species by calculating Auger rates using semiempirical electronic structure methods; the broadening for N$_2$ was reported as 120 meV.[16] Fitting to a recent high-resolution experiment lead to an estimate of the lifetime broadening of 115 meV.[14]

We have taken a different approach to the problem of determining lifetime broadening; we sample the molecular geometries and compute transition energies and intensities essentially exactly – within the limitations of the electronic structure method used to determine the excited

state energy curve -- and then apply broadening to match the experiment. (In doing so, we account for quantum statistical effects for the nuclei – those effects due to quantization of nuclear motion on the electronic potential energy curve and, thus, the populations of the various nuclear states in the ensemble.) The choice of widths that best fits the experimental data determines the appropriate lifetime broadening. This approach mirrors the procedure typically used for calculating spectra, but as a result of the accuracy of our methods, we can determine the "exact" broadening associated with a given set of approximations. We examine a hierarchy of approximations, starting with the full analytical classical configurational distribution combined with classical transitions between the ground and excited state curves and ending with an exact treatment (within the Born-Oppenheimer approximation) of Franck-Condon transitions between rovibrational levels on the quantized electronic curves. In the literature, many have instead calculated spectra using a single molecular configuration – typically the optimized geometry from an electronic structure calculation -- or from a variety of configurations generated either by (1) optimizing conformer geometries or by (2) sampling from classical or path integral molecular dynamics simulations. Common practice has also been to treat transitions only at the classical level.

There is nevertheless a substantial body of work accounting for Franck-Condon factors for small molecules undergoing NEXAFS excitation.[17-26] The most common approach in these studies is to calculate the minimum ground state energy, the associated Hessian, and the equilibrium position of the excited state. Then it is assumed that the excited state motion is characterized by the normal modes.[23-25] A linear-coupling approximation is made, and the Franck-Condon factors may be calculated explicitly or assumed to adhere to a Poisson distribution. (Previous studies indicate that this choice yields only minor differences in the results.[20]) However, none of these studies quantizes the rotational states, as we do here.

These detailed studies have often used higher-level electronic structure methods than DFT and have involved explicit calculations of the vibrational modes of both the ground and excited state. [17-26] The molecules studied are relatively complex compared to $N_2$ and therefore provide a more challenging test for prediction of the spectrum as a whole. At the same time, because of the large number of modes present in these systems, there is typically much less spectral detail over a given range of energy. Even so, and in spite of the rigor of such a treatment, the spectra are captured only semiquantitatively, with obvious disparities (in both peak location and height) in those regions of the spectrum where fine structure is most prominent. (Agreement with less-detailed spectral features is, in any case, more difficult to judge.) We show in this paper that we can predict the spectrum comparably well with respect to these fine spectral details using a simpler approach, albeit for a far simpler system, $N_2$. The simplicity of the $N_2$ molecule is such that its entire first X-ray transition can be construed as consisting of "fine structure."

## 2. Methods

### *Density Functional Theory Calculations*

The intuitive starting point for treatment of many spectroscopic absorption phenomena is the notion of the vertical transition between electronic energy levels. Density functional theory (DFT)[27,28] can accurately reproduce the absorption energies associated with such transitions, namely, the total energy difference ($\Delta$SCF or $\Delta$KS) between the levels involved.[23] Thus, rather than using an explicit excited-state method, we solve the electronic structure problem for the ground state of the same molecule, but with constraints chosen to reflect the true quasi-bound

system. Specifically, the constraints we have determined to yield a reasonable model of the core-excited system are the inclusion of a full core hole on the atom whose core excitation we wish to examine, as well as an associated excited electron.[29]

We refer readers to our earlier work for a detailed description of the electronic structure calculations.[5] A brief summary is as follows: The functional employed is the Perdew-Burke-Ernzerhof[30] form of the generalized-gradient approximation to the exchange-correlation potential, and we use a plane-wave representation (in a large periodic box) and the pseudopotential approximation for the valence electronic structure. In the present case, the initial state of the system is chosen to be the nitrogen molecular ground state, $X\,^1\Sigma_g^+$. We approximate the excited state by replacing the pseudopotential of the core-excited nitrogen by another pseudopotential that explicitly includes the core excitation, such that its electronic configuration is $1s^1 2s^2 2p^4$. When incorporated into the $N_2$ molecule, this configuration corresponds to a degenerate pair of dipole-active molecular excited states with term symbol $^1\Pi_u$ (as well as a spin-forbidden state with term symbol $^3\Sigma_u^-$). This change in configuration is the only one relevant for treating the first transition in the $N_2$ NEXAFS spectrum, which may be characterized as $1s \rightarrow \mathrm{LUMO}$.[14]

We adopt the Born-Oppenheimer approximation and allow atomic nuclei to remain fixed. Transition amplitudes are then estimated in the single-particle and dipole approximations and used to determine the X-ray absorption cross section within a first-order Fermi's golden rule treatment. The spectra thus obtained were aligned to experiment.

### *Molecular Sampling of DFT Calculations*

Over 800 nitrogen transitions were calculated at separation distances sampled from a path-integral molecular dynamics simulation, with special attention paid to configurations lying in the ground-state potential well. Variations in the transition probability were monitored and observed to remain within 1% of the intensity at equilibrium separation. This result indicates that the transition dipole moment is insensitive to the internuclear spacing. We are therefore safely within the Franck-Condon regime (*i.e.*, the electronic contribution to the transition amplitude, $\langle \pi^* | \hat{\mathbf{\mu}} | 1s \rangle$, is essentially constant) and may neglect Herzberg-Teller effects, which are predicated on such sensitivity.[5]

### *Spectral Models for Classical-to-Classical Transitions*

Figure 1 depicts schematically the types of transition considered. The leftmost (green) arrow depicts a vertical excitation at bond length *R* from the bottom of the ground state potential $V_g(R)$ to the bottom of the excited state potential $V_x(R)$. This type of transition may be termed "classical," in the sense that we treat all such excitations as being allowed with equal (unit) probability; the structure of a spectrum based on these transitions is determined by the (canonical) classical radial density of the ground state. ("Density" and "radial density" will be used throughout this work and will always denote the nuclear radial probability distribution.) We treat the spectrum as a set of intensities continuous in the transition energy and determined solely by the relation

$$I(\Delta E_{CL}(R); \beta) \propto \rho_{CL}(R; \beta)$$

where $\Delta E_{CL}(R) = V_X(R) - V_g(R)$ is the transition energy, $\rho_{CL}(R; \beta) = \frac{1}{Q} e^{-\beta V_g(R)}$ is the density,

$Q = \int dr e^{-\beta V_g(R)}$ is the associated partition function, and the semicolon indicates that the

temperature dependence (included through the inverse thermal energy $\beta = \frac{1}{k_B T}$ ) enters only

parametrically.

In previous work,[5] another model has also been used that incorporates the quantum mechanical density but leaves the transition energy in classical form:

$$I(\Delta E_{CL}(R); \beta) \propto \rho(R; \beta)$$

Although this model is inconsistent from a physical point of view, so too is any mixture of quantum and classical energetics. We address it here for the sake of emphasizing the relation between our work and its antecedents, and also because it is readily calculable even for large systems and, therefore, of pragmatic interest.

### A Spectral Model for Quantum-to-Classical Transitions

The second class of transition, depicted by the center (blue) arrows, is distinguished from the first by (vibrational or rovibrational) quantization of the ground state; here the transition is from a state $|vJ\rangle$ on the ground state potential energy curve to a classical excited state. ($\Delta E(R)$ for this type of excitation will therefore be smaller than for the analogous classical transition.) Because the excited state remains classical, all such transitions are still allowed, and the model for the spectrum remains formally the same:

$$I(\Delta E(R); \beta) \propto \rho(R; \beta)$$

The density is now quantum mechanical: $\rho(R; \beta) = \frac{1}{Z} \sum_{vJ} g_N(J) e^{-\beta E_{vJ}} |\psi_{vJ}(R)|^2$ , where $\psi_{vJ}(R)$

is the radial wavefunction for state $|vJ\rangle$, $Z = \sum_{vJ} g_N(J) e^{-\beta E_{vJ}}$ is the quantum canonical partition function, $g_N(J)$ is the rotational degeneracy (which we will address in more detail below), and the transition energy now depends on the thermal average of the quantized energies,

$\langle E_{vJ} \rangle = \frac{1}{Z} \sum_{vJ} g_N(J) e^{-\beta E_{vJ}} E_{vJ}$ , as $\Delta E(R) = V_x(R) - <E_{vJ}>$.

The rotational degeneracy $g_N(J)$ is distinguished from the usual degeneracy factor $g(J)$ = 2J + 1 in that it also accounts for the nuclear spin statistics of the $N_2$ molecule: The $^{14}N$ nucleus is a boson with spin $I = 1$, and as a result the total wavefunction must be symmetric with respect to exchange of the nuclei. To impose this condition is equivalent to requiring that antisymmetric nuclear spin states be paired with symmetric (even $J$) rotational states -- and vice versa for symmetric nuclear spin states and antisymmetric (odd $J$) rotational states.[31] There are a total of 9 symmetrized combinations of the two nuclear spins, leading to a 2:1 ratio of symmetric to antisymmetric spin states. Thus, for every three molecules of $N_2$, two will be *ortho*nitrogen (supporting only even $J$) and one will be *para*nitrogen (supporting only odd $J$). The combined rotational/nuclear-spin degeneracy must then be

$$g_N(J) = 2(2J+1) \text{ , } J \text{ even}$$
$$g_N(J) = 2J+1 \text{ , } J \text{ odd}$$

### The Franck-Condon Spectral Model for Fully Quantum Transitions

The third class of transition, shown as the set of four (red) arrows at right in Figure 1, is simply that of Franck-Condon transitions between (ro)vibrational states on the ground potential energy curve and (ro)vibrational states on the excited state potential energy curve. The probability of transition at given $R$ is no longer uniform, and the density is therefore an inappropriate quantity for use in determining the spectrum. Since the Herzberg-Teller effect is apparently of little importance for this system -- and, therefore, the electronic transition dipole moment will be roughly constant, falling within the framework of the Franck-Condon approximation -- we now model the spectrum as deriving from all symmetry-allowed ground-to-excited-state transitions $|vJ\rangle \rightarrow |v'J'\rangle$.

We assume that both the ground and excited state fall under Hund's case (a) for coupling between the spin and orbital angular momenta, which consists of weak coupling of electronic and nuclear motions and strong coupling of the spin and orbital motion to the internuclear axis.[31] The ground state $X\ ^1\Sigma_g^+$ has angular momentum around the internuclear axis $\Omega = |\Lambda + \Sigma| = 0$, where $\Lambda$ is the maximum projection of the orbital angular momentum (corresponding to the Greek letter in the term symbol for the state) and $\Sigma$ is the appropriate projection of the spin angular momentum. The excited state $^1\Pi_u$ has $\Omega = 1$, fulfilling the $\Delta\Omega = 0, \pm 1$ selection rule for the electronic transition; further, all rotational transitions with $\Delta J = 0, \pm 1$ are allowed. (This latter rule always holds, of course, but the $\Delta J = 0$ case is forbidden for $\Delta\Omega = 0$.) Note also that, since $J$ is bounded from below by $\Omega$, there is *no* $J = 0$ rotational state on the excited electronic surface.[31]

These transitions are treated as occurring in proportions set by the relative thermal population of the state $|vJ\rangle$ with respect to the ground state $|00\rangle$, *i.e.,* with probability $P_{vJ}(\beta) = \dfrac{1}{Z} g_N(J) e^{-\beta E_{v_0 J_0}}$, where $\Delta E_{v'vJ'J} = E_{v'J} - E_{vJ}$. These considerations yield the standard spectral model

$$I(\Delta E; \beta) \propto \sum_{v'vJ'J} \delta(\Delta E - \Delta E_{v'vJ'J}) \delta(|J'-J| \le 1)\langle v'J'|vJ\rangle P_{vJ}(\beta)$$

where the first Kronecker delta restricts $\Delta E$ to those values supported by the ground and excited state energies, the second Kronecker delta enforces the rotational selection rule, and $\langle v'J'|vJ\rangle = \int dR\, \phi_{v'J'}^*(R)\psi_{vJ}(R)$ is the Franck-Condon factor (overlap integral) for the transition. Hereafter, $v'$ and $J'$ will be used to refer to quantum numbers for states on the excited potential energy curve and $v$ and $J$ to those for states on the ground potential energy curve.

### The Ground and Excited State Potentials and Fits

Our initial calculations of the nitrogen X-ray absorptions were for values of $R$ centered in the ground state potential well; this is also true for the excited-state potential, which is displaced by only $\approx 0.05\,\text{Å}$ at equilibrium. In order to capture nuclear quantum effects, however, the geometry of the potential away from the minimum may be important; for example, anharmonicity at larger values of $R$ will affect the nuclear vibrational frequencies.

We avoided sampling the electronic structure energies on an infeasibly dense grid in the radial coordinate by adding a small number of points at an even spacing of 0.1Å. These points span the entire range of physically-relevant bond separations, from well into the repulsive barrier region (0.1Å) to near the dissociation threshold (4.7Å for the ground state and 3.4Å for the excited state). The complete set of points was then used to develop a 14-Gaussian fitting function

of form $V(R) = \sum_i a_i e^{-(\frac{R-c_i}{b_i})^2} + V_e$ for each curve, where $V_e$ is the energy at the well minimum and $\{a_i, b_i, c_i\}$ are the fit parameters.

The resulting ground- and excited-state potential fits are shown in Figures 2 and 3, respectively; they are paired for the purpose of comparison with potentials obtained by fitting to experimental data.[6,32,33] Although our potential fits are in good agreement with the experimentally-derived curves in the well region, the dissociation energies are several eV too large, indicating the over-binding commonly observed for DFT.[34] Additionally, the multi-Gaussian functional form does not allow for reliable prediction of the electronic energy outside of the fitting regions, but this is not an important failing. The states of interest to us here are deep within the well, at energies no higher than 0.435 eV for the ground and 404 eV for the excited state (*vide infra* for discussion of the physical relevance of these energies).

### The Classical Density

The classical radial density for the (rotationless) ground state at 300K was determined analytically. It is very nearly Gaussian, with slight enhancement for values of $R$ greater than the most-probable bond separation and slight diminishment for smaller values of $R$. The uneven distortion with respect to a Gaussian is due to the shape of the potential, which allows for more frequent visits to the large-$R$ region than to the small-$R$ region at 300 K. To determine the classical canonical density at 0K is trivial; the particle is at a standstill, sitting at the bottom of the potential well. As such, it is a δ-function centered at the ground state equilibrium position.

### Computing the Quantum Densities: The Colbert-Miller DVR

In determining the quantum densities, it is important that the wavefunctions required can be obtained simply and straightforwardly and that they be of high quality. As such, we opted to use the DVR of Colbert and Miller, which we will briefly review here (in its radial version).[35]

As do other DVRs, the Colbert-Miller method involves setting up a grid representation of the Hamiltonian $\hat{H} = \dfrac{\hat{p}^2}{2\mu} + V(\hat{R})$ of the system, with μ the reduced mass. In this case, the grid

falls in the coordinate range $(0, \infty)$. Because the potential $V(\hat{R})$ is a function only of the position operator, its matrix representation in the basis of grid points $\{R_i\}$ is trivial, $\mathbf{V}_{ij} = \delta_{ij} V(R_i)$, and the challenge lies in determining the representation of the kinetic energy operator. Colbert and Miller show that there is a specific representation that is independent of any choice of basis functions for the grid:

$$\mathbf{T}_{ij} = \frac{\hbar^2}{2m\Delta R^2}\left[\frac{\pi^2}{3} - \frac{1}{2i^2}\right] \text{ for } i = j$$

$$\mathbf{T}_{ij} = \frac{\hbar^2}{2m\Delta R^2}(-1)^{i-j}\left[\frac{2}{(i-j)^2} - \frac{2}{(i+j)^2}\right] \text{ for } i \neq j$$

where $\Delta R$ is the grid spacing and $R_i = i\Delta R$ for $i \in \{1....., \infty\}$.

They also show that the wavefunction associated with each grid point may be written as

$$\langle R | R_i \rangle = \frac{\sin[\frac{\pi(R - R_i)}{\Delta R}]}{\pi(R - R_i)}$$

consequently, the eigenvectors obtained from the diagonalization of the Hamiltonian matrix $\mathbf{H} = \mathbf{T} + \mathbf{V}$ may be expressed as a linear combination of the grid wavefunctions, $|\psi_j\rangle = \sum_i c_{ij} |R_j\rangle$, that is evaluable for all $R$. Since we are concerned here with computing densities and Franck-Condon factors, obtaining a smooth wavefunction that may be evaluated everywhere in space is not just reason for satisfaction; it brings us nearer to the spirit of the analytical theory.

Three parameters are varied in bringing the Colbert-Miller DVR to convergence: the grid spacing $\Delta R$, the maximum value of the position $R_{max} = i_{max}\Delta R$, and a cut-off energy $V_c$. The latter allows for a simple form of adaptive grid, in which no grid points are placed where $V(R_i) > V_c$. An intuitive choice of cut-off energy is the dissociation energy for the curve in question, and $V_c$ for the excited state potential was set accordingly. The analogous choice for the ground state would be inappropriate; the potential fit was ill-behaved when points at 0.8Å and smaller were included, and so the potential would not be represented as bound with such a choice for $V_c$. The well is nevertheless sufficiently deep that $V_c \approx 10.7\,\text{eV}$ is compatible with convergence of the densities. Similarly, we wanted to allow for leakage of wavefunction amplitude via tunneling to as large a value of $R$ as physically meaningful, and so we set $R_{max}$ to values well into the classically-forbidden region for each curve; 1.5Å and 3.5 Å sufficed for the ground- and excited-state potentials, respectively. Finally, we found the point of convergence in $\Delta R$ by examining the energies of the rovibrational states and a few simple properties, such as canonical averages of the position and momentum, their uncertainties, and the kinetic and potential energies. Our working values were chosen to be roughly 3 times smaller, in order to guard against any residual distortion of the computed wavefunctions with respect to their exact counterparts, and we tested for overall convergence by varying $V_c$ and $R_{max.}$ In the end, $\Delta R = 0.003$Å and $\Delta R = 0.0125$Å were chosen for the ground and excited states. This choice is equivalent to having chosen 199 and 203 grid points for the rotationless ground and excited states, respectively; only a few points are lost to the cut-off as $J$ increases. We validated our calculations by comparison of the DVR energies for the ground state to those given by a Dunham expansion due to Le Roy *et al.* [33]; they were found to be in excellent -- although, as expected, not exact -- agreement.

### The Quantum Densities

Once the DVR for the ground state was converged, obtaining the canonical density was a simple matter. A single DVR calculation accounts for contributions to the density from all $|vJ\rangle$, with $J$ fixed and included as part of the centrifugal potential; we needed only to perform as many calculations as values of $J$ we wished to include in the average. As such, we chose an arbitrary cut-off in the rotational quantum number, $J_c$. The condition imposed is that the relative population in $|0J_c\rangle$ be of order $10^{-5}$ or smaller compared to that in $|0J_{max}\rangle$, where $J_{max}$ is that $J$ for which the population factor $P_{0J}(\beta)$ is largest. For $T = 300$K, we found the appropriate choices to be $J_{max} = 8$ and $J_c = 34$. At $T = 0$K, of course, $J_{max}$ and $J_c$ are both rigorously 0; however, nuclear spin selection rules prevent conversion of paranitrogen to orthonitrogen under typical laboratory conditions.[31] As a result, there will initially be a 2:1 ratio of molecules in $J = 0$ and $J = 1$, respectively, after a quench from room temperature.

With the canonical densities $\rho_J(R;\beta)$ in hand, we constructed the total density $\rho(R;\beta)$ by taking their weighted average with respect to the populations of the $|0J\rangle$. (This is a very minor approximation at 300K, since even the largest relative contribution from the $|1J\rangle$ states is 7 times smaller than that for $|0J_c\rangle$. At 0K, it is no approximation at all.) That is,

$$\rho(R;\beta) = \frac{Z}{Z_c} \sum_{J=0}^{J_c} P_{0J}(\beta)\rho_J(R;\beta)$$

where the partition function ratio ensures a normalized average and $Z_c = \sum_{J=0}^{J_c} g_N(J)e^{-\beta E_0 J}$ is the appropriately restricted partition function.

As was true in the classical case, the quantum density is very nearly Gaussian; the distinction is that it is significantly broader due to tunneling into the classically-forbidden regions on either side of the well. Because there is little change in the density between states with low values of *J*, the two distinct 0K scenarios are essentially indistinguishable at this level of approximation.

### Classical and Quantum-Classical Spectral Calculations

For the first class of transitions, shown in green at left in Fig. 1 (and the corresponding spectral models), densities dictate the intensity of the X-ray absorption for specified values of *R*. The associated transition energies are simply the energy differences between the fitted ground- and excited-state potentials. The resulting spectra were binned in the transition energy to the nearest meV for convenience in plotting and (more importantly) so that spectral degeneracy would be properly reflected.

Upon quantizing the ground state, transitions to the classical excited state are made almost exclusively from $\nu = 0$, with rotational levels populated commensurate with the temperature. The resulting minimum transition energy, which includes contributions from the zero-point energy of the ground state, will therefore be lower than for the analogous classical transition.

### Franck-Condon Spectral Calculations

The Franck-Condon spectral calculation was carried out for the same set of states $|\nu J\rangle$ as in the ground state density calculations. Based on the number of peaks in the experimental spectrum, we chose to include those states on the fitted excited state potential with $\nu' \in \{0,....6\}$; the range of rotational quantum number chosen was $J' \in \{1,....34\}$, determined by choosing the excited state rotational cut-off value $J_c'$ to be that for which $E_{0J_c'}$ first appreciably exceeds $E_{10}$. This criterion was used in part for convenience and in part to ensure that a reasonable spread of transition energies would be covered by our calculation. Franck-Condon factors $\langle \nu' J' | \nu J \rangle$ were then computed numerically from the DVR wavefunctions and scaled by the population factor $\rho_{\nu J}(\beta)$. The spectrum was constructed by combining the results of these calculations and applying an additional scale factor such that the height of the first peak in the computed spectrum matches that from the experiment; this latter scaling may be interpreted as incorporating the magnitude of the electronic transition dipole moment, which we have not calculated explicitly for every *R*.

### 3. Results and Discussion

Figure 4 is a direct comparison to experiment of our results for the spectra arising from purely classical transitions; we have shifted all of our spectra, irrespective of the model used to obtain them, to align energetically with the experimental spectrum of Yates *et al.*[14] In panel (A), the classical density is used, and, much as one would expect, the result is a near-Gaussian spectrum lacking any non-trivial structure. (In the 0K limit, the spectrum is essentially a $\delta$-function, though obviously not of infinite intensity.) We discuss this simplest model here in order to help establish what relevant features of the spectrum become apparent at which level of approximation.

When the quantum densities are used instead, as in panel (B) of Figure 4, the absorption profile broadens substantially, and the (now comparatively small) maximum shifts by 42 meV. Of greater interest is that the spectrum using this model is almost completely temperature invariant; because only the lowest vibrational level is appreciably populated, and because the densities of the various rotational modes are practically identical, the overall ground state density − and, thus, the spectrum -- does not change significantly between 300K and 0K. (This near-complete insensitivity to lowered temperatures holds for all of our quantized spectra; we address the question of higher temperatures in the context of our conclusions.) In other words, tunneling spreads the density over a wider range of $R$, as it must, but no model neglecting the quantum energetics will contain any additional *structure*. Unsurprisingly, one must (at the very least) quantize the ground state.

The results of the ground-state quantization and subsequent application of the second spectral model are shown in Figure 5. The spectrum calculated at 0K displays a single peak with very sharp onset and then a broader decay. The speed of the onset is due to the fact that all transitions now originate either from a pair of states spaced within $\approx 4 \, \mathrm{cm}^{-1}$ (after a quench from room temperature) or from the ground state alone (at thermal equilibrium); further, the small displacement of the electronic potential minima ($\approx 0.05$ Å) will result in many transitions to the relatively flat bottom of the excited-state well. At 300K the ground state rotational energies are no longer (essentially) completely degenerate, leading to a weaker onset and a broader decay; that is, the finite-temperature rotational statistics of the ground state directly result in broadening of the spectrum. There is also an interesting additional feature: Because the rotational progression now spans a wider range of energies and values of *J*, the nuclear spin statistics are directly observable as a series of narrowly-spaced peaks corresponding to transitions from even-*J* states. This "ragged onset" is an artifact of the questionably-physical quantum-classical spectral model, and it would be observed in some form irrespective of the shape of the excited potential curve.

As we have seen, quantization of the ground state results in a sharp Franck-Condon-like peak. It is clear, however, that an accurate representation of the Franck-Condon structure of the experimental spectrum requires that the excited state be quantized as well.

We begin by quantizing both states only vibrationally, yielding the stick spectrum shown in light blue in Figure 6(A). (There is essentially no difference between a 0K and 300K spectrum in this picture owing to the minimal change in thermal population for $\nu = 1$.) Because the pure vibrational transitions (to a good approximation) set the positions of the peaks, this model is effectively that used in fitting experimental spectra, with the peak locations and the Lorentzian and Gaussian widths contributing to the Voigt lineshape used as fitting parameters.

When our spectrum is broadened using the widths Yates *et al.* determined by such a fitting procedure − 132 meV overall broadening, with 115 meV FWHM Lorentzian and 38.4

meV FWHM Gaussian -- the purple line is obtained.[14] We note that we have adopted the Gaussian broadening reported by Yates *et al.* as a reasonable value for the instrument broadening, since it cannot be determined *a priori*. Owing to the non-uniqueness of the Lorentzian/Gaussian decomposition of the Voigt profile, this choice is merely an example, if a well-justified one.

Although the first peak is shifted such that its maximum is aligned with the experiment – and, thus, its location is immaterial in determining the quality of the predicted spectrum -- its width matches the experiment quite well. As $v'$ increases, our peaks are displaced progressively higher in transition energy from their experimental locations, and their heights do not decay sufficiently quickly for $v' > 2$. We must stress, however, that we have arrived at these results by constructing the spectrum literally "from the ground state up," while the experimental broadening figures are determined by fitting. As such, we find the agreement to be a striking confirmation of the robustness of our approach.

Although rotational transitions cannot be resolved spectroscopically, we thought it worthwhile to estimate the associated contribution to broadening of the spectrum. Results of the rovibrational Franck-Condon spectral calculation are shown at 300K in Figure 6(B) and at 0K in Figure 6(C), with stick spectra in light blue as before. When the spectra are broadened by 132 meV FWHM Voigt lineshapes, as before, the purple lines are obtained. Comparison of the stick spectra in panels 6(B) and 6(C) indicates that the addition of rotational structure leads directly to a very modest "broadening" of the spectrum, in the sense of formation of groups of closely-clustered peaks associated with a single excited vibrational state. The magnitude of this rotational effect at room temperature may be determined by adjusting the experimentally-motivated Lorentzian broadening used thus far until an optimum fit is obtained. Consequently, we estimate that quantized rotations are responsible for 3 meV of the experimentally-determined lifetime broadening for $N_2$, and we estimate the pure lifetime broadening to be 112 meV. This value is in acceptable agreement with the value of 120 meV predicted from Auger lifetime calculations by Coville and Thomas, although we note that they treated the ionized rather than the bound system.[16]

Although a discrepancy of 3 meV is very small indeed, it is nevertheless interesting to consider that $N_2$ is a common benchmark system for determining beamline resolutions. In determining a Lorentzian/Gaussian decomposition of a given Voigt lineshape, a slight overestimate of the lifetime broadening corresponds to a slight underestimate of the instrument broadening. As a result, a more careful treatment of rotations might be of interest to the X-ray spectroscopy community.

We are not aware of any experimental studies of the $N_2$ NEXAFS spectrum (or other molecules) at low temperature; given the negligible narrowing effect due to quenching of higher rovibrational states as the temperature is decreased, we think it unlikely that any will be performed soon. High-temperature experiments might be of interest, however. Guimarães *et al.* have computed the distinct vibrational progressions associated with several originating values of $v$ for the NO molecule, demonstrating for that system that high-temperature spectra will indeed display additional structure.[36] In the present case, as contributions from $v > 0$ become non-negligible, shoulders will appear that are of spacing identical to that observed in the ground state spectrum but with an additional energy shift of $\Delta E_{v0}$. A simple population analysis suggests that $v = 1$ will first make a contribution of 10% or more at ~1500K.

One possible source of the discrepancy between our peak locations and heights and those observed experimentally is that we have neglected the $\Lambda$-splitting of the + and - components of the $^1\Pi_u$ term with increasing $J$ in computing our Franck-Condon spectrum.[31] Although this effect will be on the order of fractions of a wavenumber for low values of $J$, it could be significant for the higher values included, growing as large as tens of wavenumbers (or larger). If the effect were strong enough for high $J$, contributions from orthonitrogen would disappear (as the rotational selection rule would prevent it from coupling to the symmetry-allowed $^1\Pi_u^+$ state). Even assuming that the $^1\Pi_u^-$ term were still near enough to degenerate with $^1\Pi_u^+$ that we could ignore its being symmetry-forbidden, we would have to add the $J$-dependent splitting (or combination defect) by hand, drawing values from an appropriate experiment.[31] Thus, while it is possible that $\Lambda$-splitting might contribute, it is difficult to say what the effect of including it might be.

The most straightforward explanation of the peak location and height discrepancy we observe, however, is that the DFT ground- and excited-state potential geometries are different from those of the real physical states involved in the X-ray absorption process. Specifically, as mentioned previously, we would expect that our over-binding potential curves overestimate the zero-point energy and the vibrational level spacing. The blue line in Figure 2 is the MLR$_4$(6,8) ground-state potential developed by LeRoy *et al.* from fitting to experimental data for vibrational levels up to $\nu = 19$; it compares favorably to configuration-interaction results obtained by Gdanitz.[32,33] Comparison to our ground-state fit shows that the well *is*, in fact, slightly too narrow, although the overall geometry is quite similar. An analogous comparison may be made for the excited state: The blue line in Figure 3 is the Morse potential implicit in the spectroscopic constants for nitrogen reported by Chen *et al.*[6]; once again, our fitted curve is too narrow.

The results presented in Figure 7 address the question of how much improvement will be visible in the spectrum when a better-quality ground- or excited-state potential curve is used. Panel (A) is the same spectrum as in Figure 6(B), with both states those determined by our DFT calculations and subsequent fitting. Figure 7(B) substitutes the fitted Morse potential for the excited state curve, while Figure 7(C) involves the additional replacement of the ground-state potential by MLR$_4$(6,8). [6,32,33] Perhaps surprisingly, the best fit to the experiment is obtained from the combination of the DFT ground state and the excited state given by Chen *et al.*; the peak locations are exceptionally accurate, and ratios between the heights of adjacent peaks are also in excellent agreement. (Of course, *some* improvement is to be expected, since an empirical fit must necessarily incorporate the "right answer," in some sense.) Such is not true for the (in principle) superior combination presented in Figure 7(C), which improves the predicted peak locations but actually suffers by comparison to the purely DFT-derived spectrum with respect to the intensities. This unexpected reversal might be due to the fact that the excited Morse potential was determined without consideration of rotations – and, thus, already accounts for them in some averaged sense -- or it might be the result of fortuitous cancellation of errors when the Morse fit is combined with the DFT ground state. Such inconsistent improvement in the predicted spectrum emphasizes the general need for higher-quality ground- and excited-state potentials, whether empirical or *ab initio*. In any case, our constrained DFT approach leads to a reasonably good prediction.

## 4. Concluding Remarks

In this article, we have shown that the NEXAFS spectrum can be predicted semiquantitatively using a combination of constrained density functional theory and exact methods for quantization of the nuclear motions. In the process, we have explored a hierarchy of models for the spectrum and established which of the salient features of the spectrum appear at each successively more accurate level of approximation. Specifically, we find that a sharply-peaked onset will appear once the ground state has been vibrationally quantized; full Franck-Condon structure is recovered on quantization of the excited state, unsurprisingly; and the addition of rotational quantization shows that there are associated very modest – but real – contributions to the intrinsic broadening of the spectrum. We have also addressed the possibility of non-trivial temperature dependence in the spectrum. Finally, we have shown that it is possible to improve the predicted spectrum by using constrained DFT in combination with more accurate potentials.

**Figure Captions**

**Figure 1.** A schematic depiction of the three classes of transitions associated with our spectral models. The green arrow at left depicts purely classical transitions; the blue arrows in the center depict transitions from a quantized ground to a classical excited state; and the red arrows at right depict transitions from a quantized ground to a quantized excited state. See text for details.

**Figure 2.** The electronic structure energies for the ground state in eV (red points) plotted with the 14-Gaussian fit (black) to the full region depicted. The inset between 0.9 and 1.6Å shows the quality of the fit for the bottom of the ground state well; rovibrational state $|0,34\rangle$ is deep within the well (0.435 eV).

**Figure 3.** The electronic structure energies for the excited state in eV (red points) plotted with the 14-Gaussian fit (black) to the full region depicted. The inset between 1 and 1.5 Å shows the quality of the fit for the bottom of the excited state well; rovibrational state $|6,34\rangle$ lies at an energy of 403.8 eV.

**Figure 4.** Part (A) shows the experimental nitrogen K-edge NEXAFS (red – see text) and the spectra generated by transitions from classical curves using the classical density at 300K (black) and 0K (grey, dashed). Note that the spectra are too narrow. Part (B) shows the experimental nitrogen K-edge NEXAFS and the spectra generated by transitions from classical curves using the quantum densities at 300K (brown). The 0K spectrum is completely hidden by the 300K spectrum.

**Figure 5.** Comparison of the experimental nitrogen K-edge NEXAFS (red – see text) and the spectra generated by transitions from a quantized ground state to a classical excited state at 300K (purple) and 0K (blue, dashed) with broadening of 1 meV.

**Figure 6.** Part (A) shows the experimental nitrogen K-edge NEXAFS (red – see text) and the spectrum generated by purely vibrational transitions from the ground to the excited state (blue sticks) broadened with a Voigt lineshape (purple). The 0K vibrational spectrum is indistinguishable from that for 300K. Part (B) shows the experimental nitrogen K-edge NEXAFS and the spectrum generated by rovibrational transitions from the ground to the excited state (blue sticks) broadened with a Voigt lineshape (purple) at 300 K . Part (C) is the experimental nitrogen K-edge NEXAFS and the spectrum generated by rovibrational transitions from the ground to the excited state (blue sticks) broadened with a Voigt lineshape (purple) at 0 K. In all cases the Voigt lineshape is 132 meV FWHM, and the Gaussian contribution is 38.4 meV FWHM.

**Figure 7.** Part (A) shows the experimental nitrogen K-edge NEXAFS (red – see text) and the spectrum generated by rovibrational transitions from the ground to the excited state (blue sticks), as defined by our Gaussian fits, and broadened with a Voigt lineshape (purple) at 300K. Part (B)

shows the experimental nitrogen K-edge NEXAFS and the spectrum generated by rovibrational transitions from the DFT ground state potential to the Morse potential due to Chen *et al.* (blue sticks) and broadened with a Voigt lineshape (purple). Part (C) shows the experimental nitrogen K-edge NEXAFS and the spectrum generated by rovibrational transitions from the LeRoy *et al.* ground state to the Morse potential due to Chen *et al.* (blue sticks) and broadened with a Voigt lineshape (purple). In all cases the Voigt lineshape is 132 meV FWHM, and the Gaussian contribution is 38.4 meV FWHM.
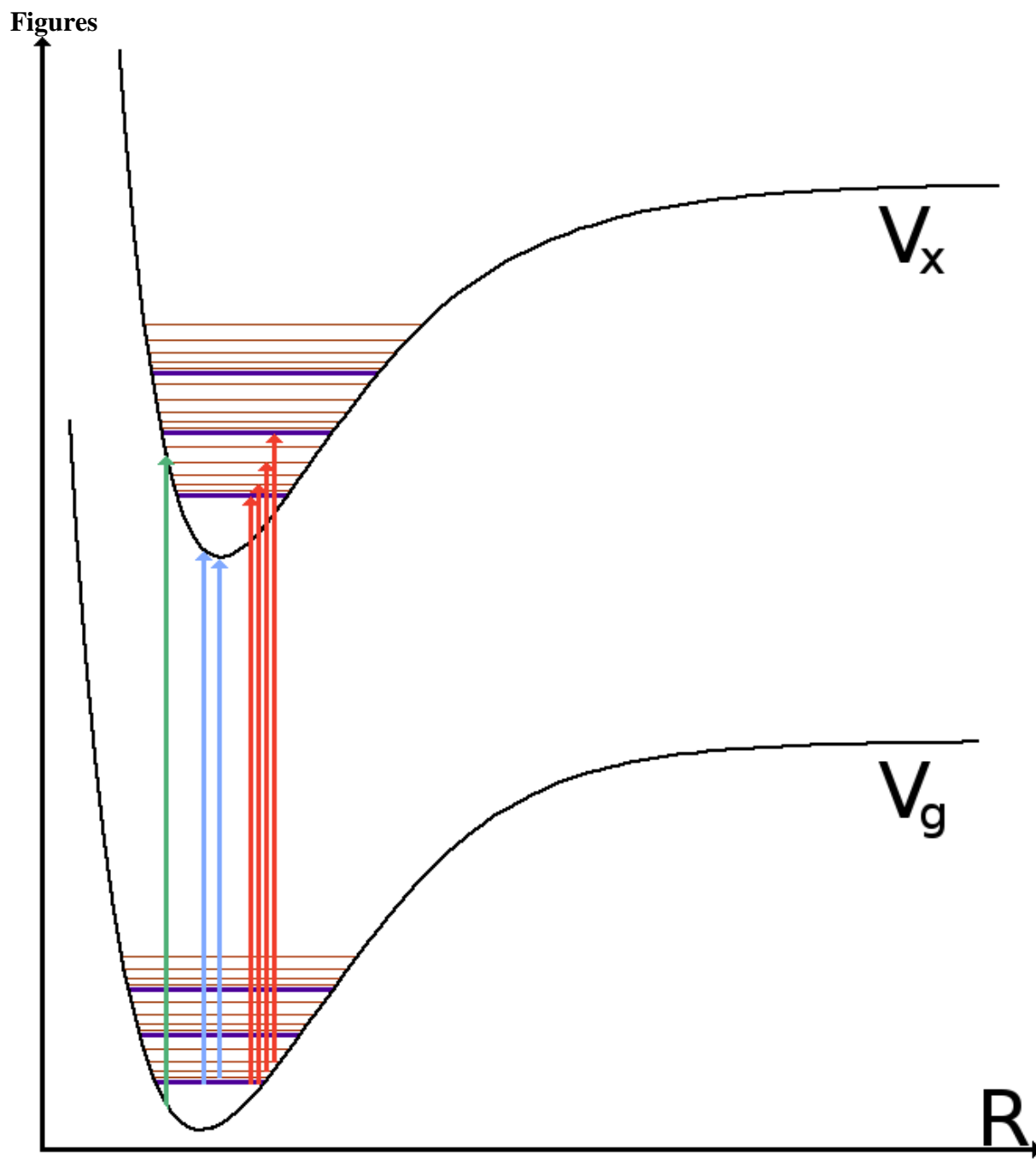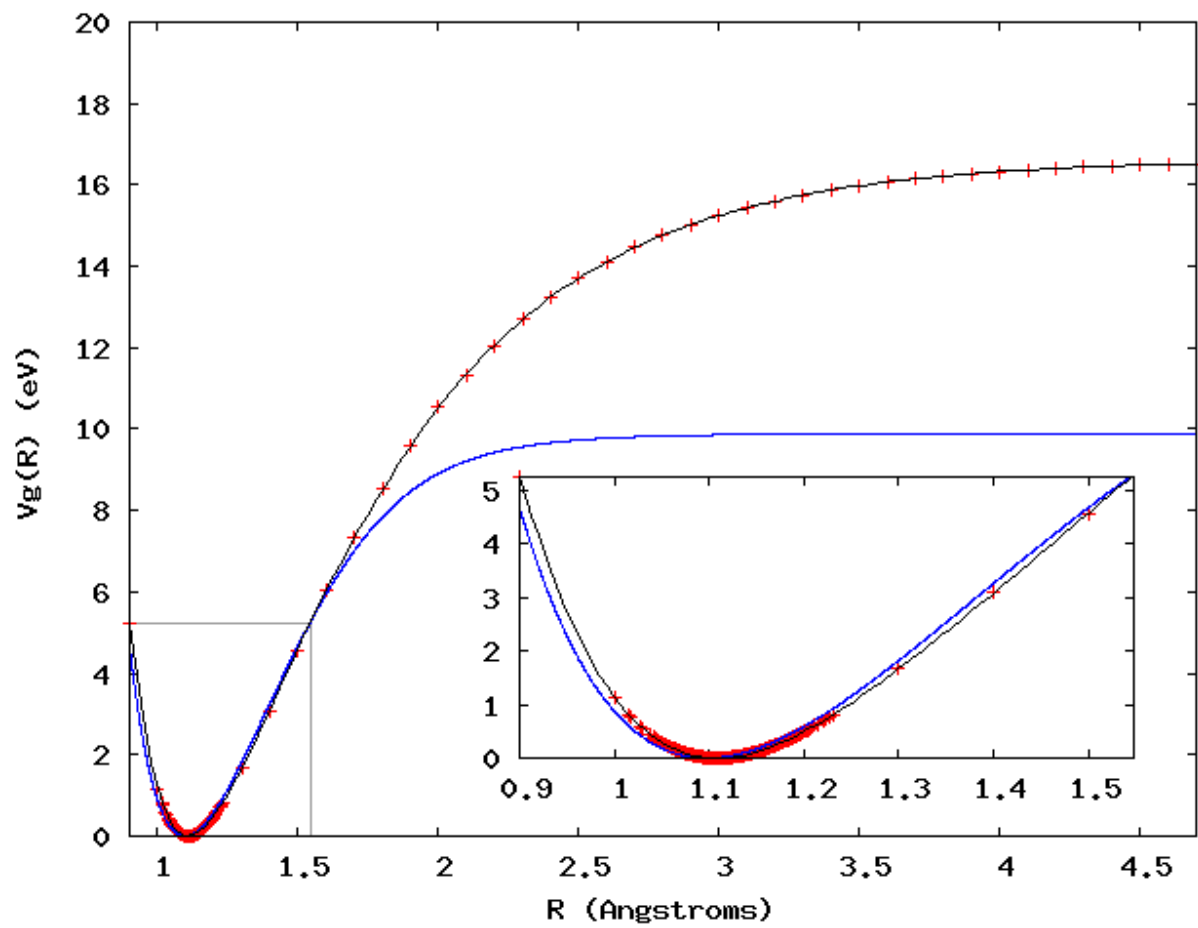
**Figures**
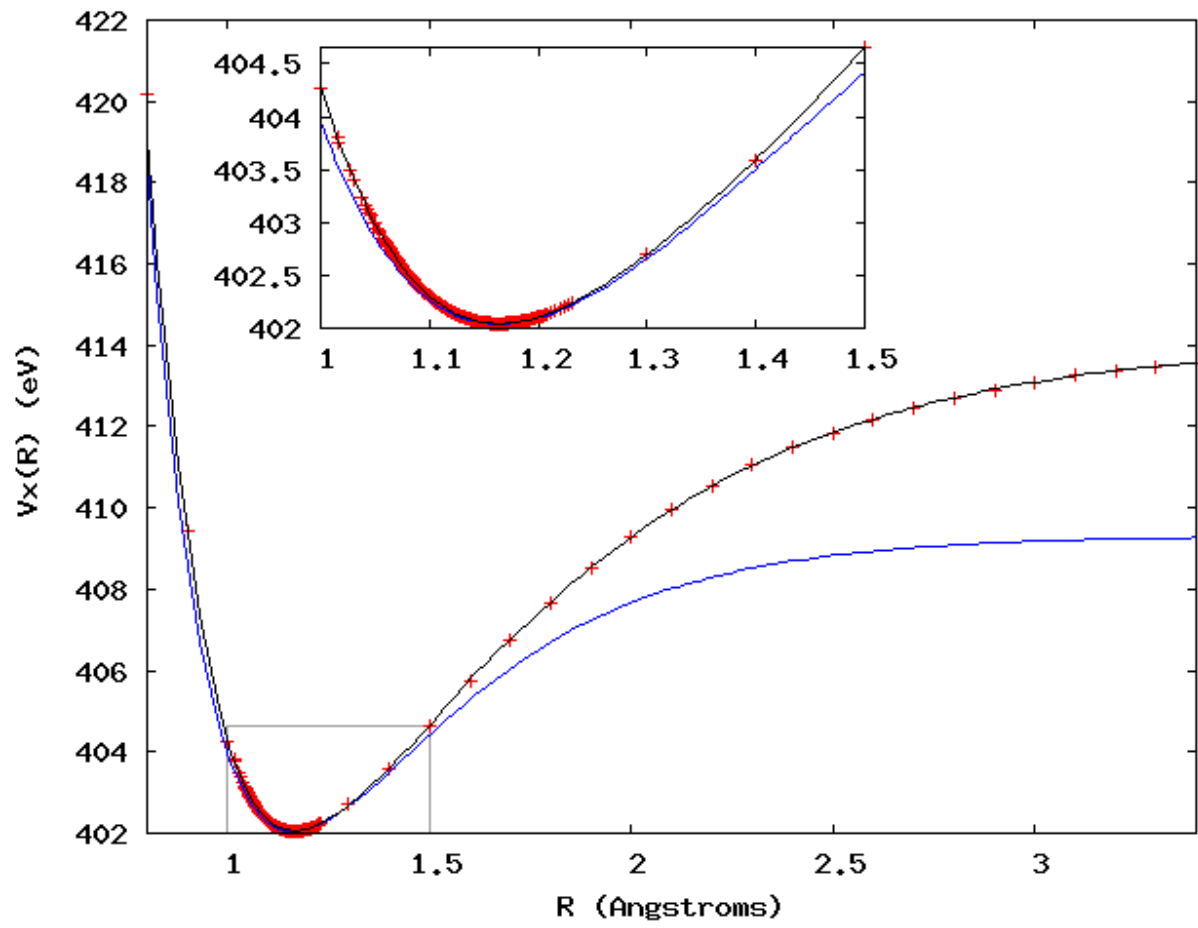


**Figure 1**
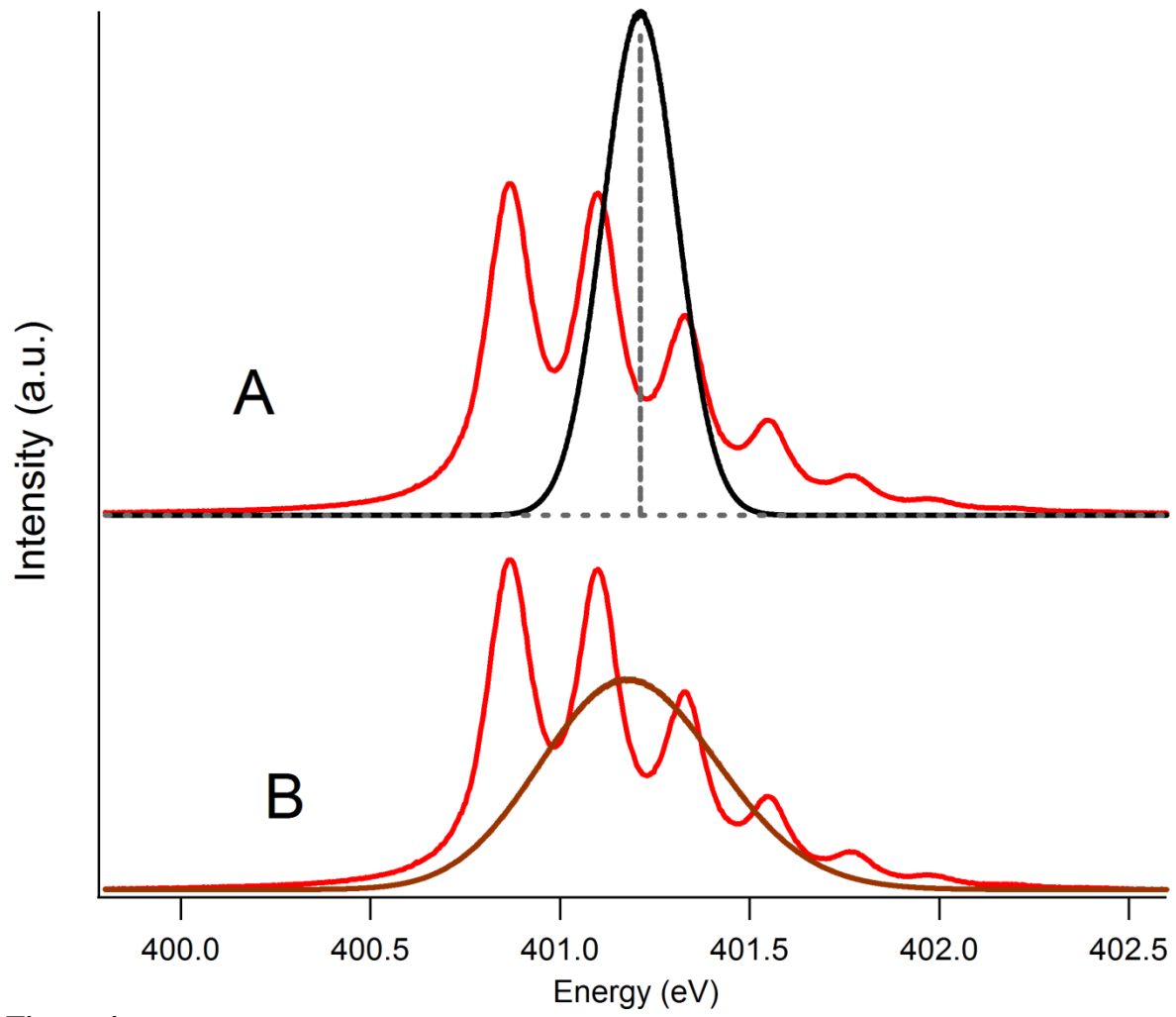
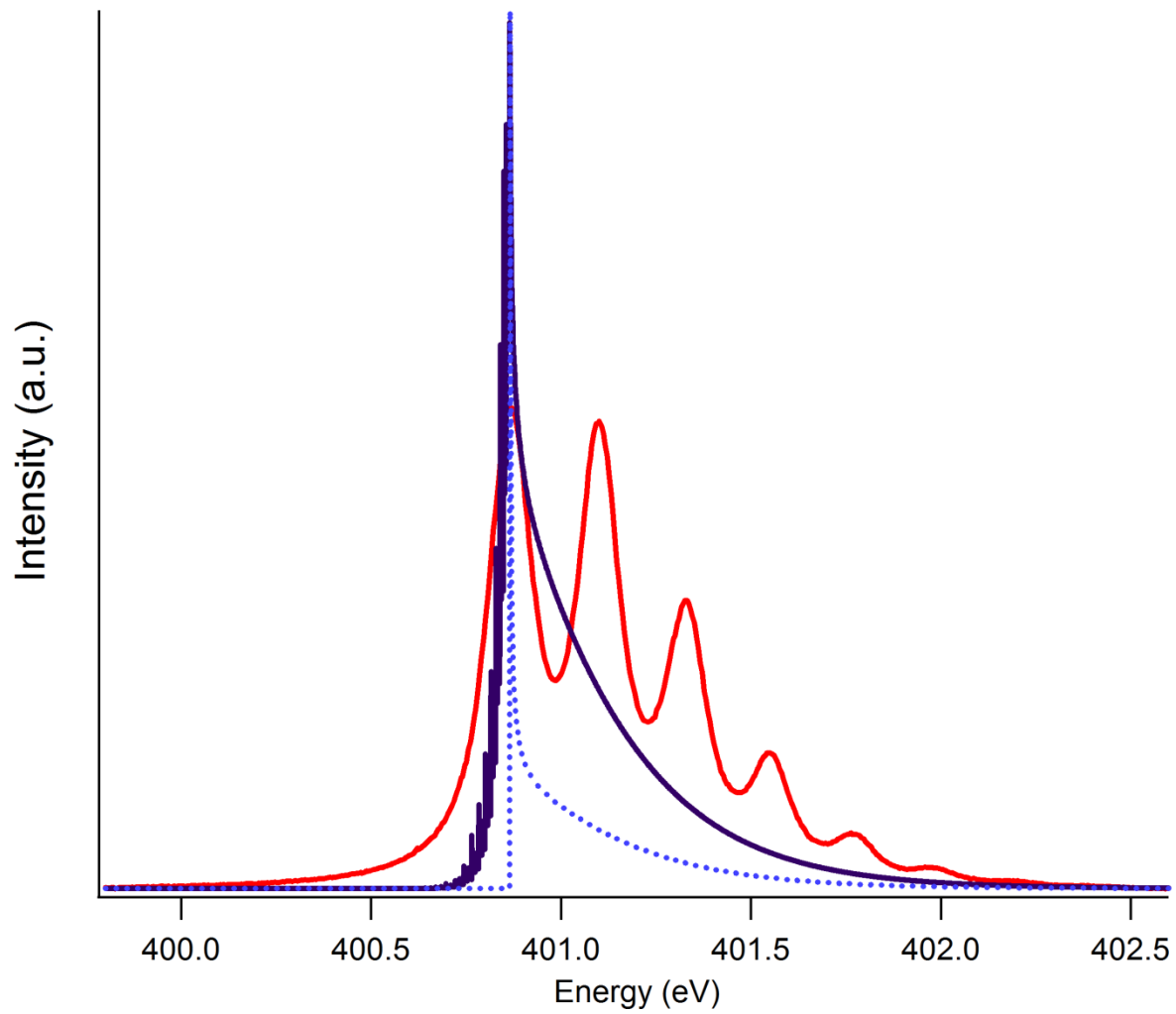**Figure 2**
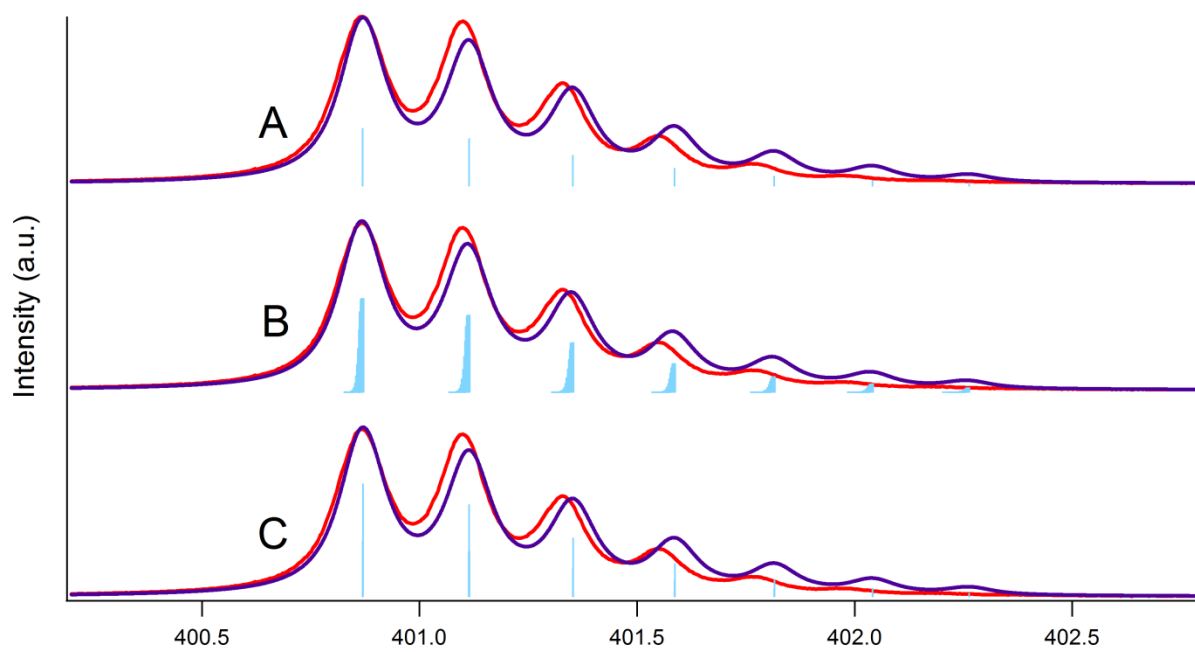
**Figure 3**
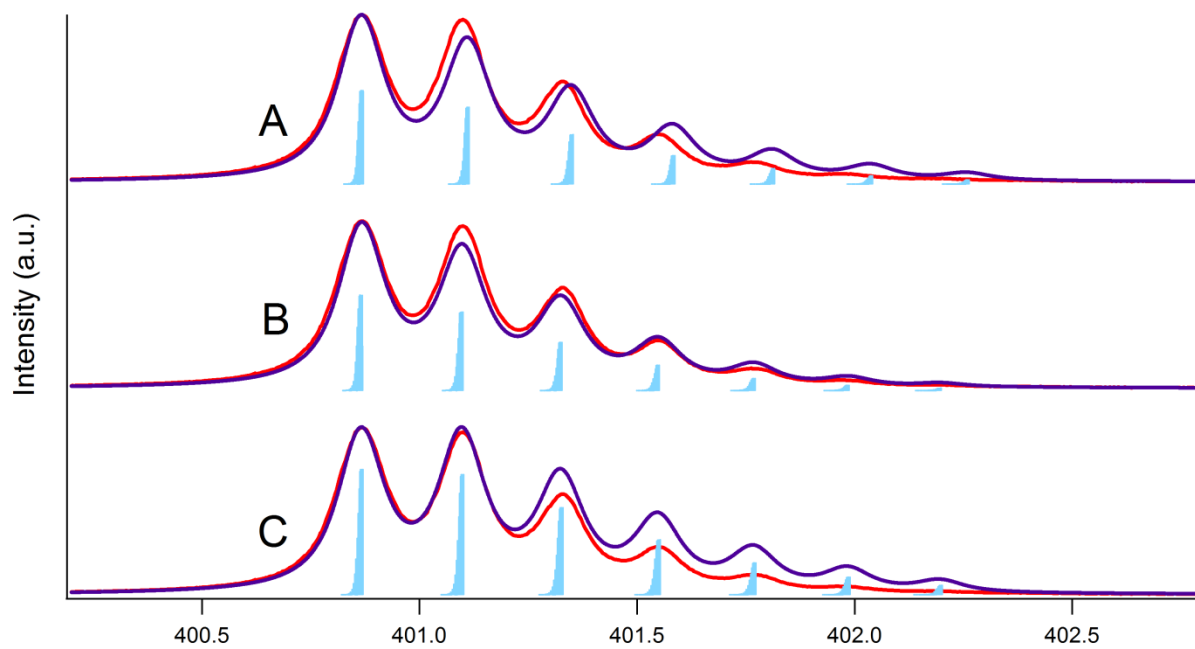
**Figure 4**

**Figure 5**

**Figure 6**

**Figure 7**

1.  J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).
2.  J. Stewart-Ornstein, A. P. Hitchcock, D. H. Cruz, P. Henklein, J. Overhage, K. Hilpert, J. D. Hale, and R. E. W. Hancock, Journal of Physical Chemistry B **111** (26), 7691 (2007).
3.  H. Ade and A. P. Hitchcock, Polymer **49** (3), 643 (2008).
4.  J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).
5.  C. P. Schwartz, J. S. Uejio, R. J. Saykally, and D. Prendergast, Journal of Chemical Physics **130** (18) (2009).
6.  C. T. Chen, Y. Ma, and F. Sette, Physical Review A **40** (11), 6737 (1989).
7.  L. Floreano, G. Naletto, D. Cvetko, R. Gotter, M. Malvezzi, L. Marassi, A. Morgante, A. Santaniello, A. Verdini, F. Tommasini, and G. Tondello, Review of Scientific Instruments **70** (10), 3855 (1999).
8.  A. P. Hitchcock and C. E. Brion, Journal of Electron Spectroscopy and Related Phenomena **18** (1-2), 1 (1980).
9.  M. Kato, Y. Morishita, M. Oura, H. Yamaoka, Y. Tamenori, K. Okada, T. Matsudo, T. Gejo, I. H. Suzuki, and N. Saito, Journal of Electron Spectroscopy and Related Phenomena **160** (1-3), 39 (2007).
10. G. C. King, F. H. Read, and M. Tronc, Chemical Physics Letters **52** (1), 50 (1977).
11. A. Kivimaki, K. Maier, U. Hergenhahn, M. N. Piancastelli, B. Kempgens, A. Rudel, and A. M. Bradshaw, Physical Review Letters **81** (2), 301 (1998).
12. H. Ohashi, E. Ishiguro, Y. Tamenori, H. Okumura, A. Hiraya, H. Yoshida, Y. Senba, K. Okada, N. Saito, I. H. Suzuki, K. Ueda, T. Ibuki, S. Nagaoka, I. Koyano, and T. Ishikawa, Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment **467**, 533 (2001).
13. K. C. Prince, M. Vondracek, J. Karvonen, M. Coreno, R. Camilloni, L. Avaldi, and M. de Simone, Journal of Electron Spectroscopy and Related Phenomena **103**, 141 (1999).
14. B. W. Yates, Y. F. Hu, K. H. Tan, G. Retzlaff, R. G. Cavell, T. K. Sham, and G. M. Bancroft, Journal of Synchrotron Radiation **7**, 296 (2000).
15. A. Nilsson and L. G. M. Pettersson, Surface Science Reports **55** (2-5), 49 (2004).
16. M. Coville and T. D. Thomas, Physical Review A **43** (11), 6053 (1991).
17. D. Duflot, S. Zeggari, and J. P. Flament, Chemical Physics **327** (2-3), 518 (2006).
18. D. Duflot, J. P. Flament, A. Giuliani, J. Heinesch, and M. J. Hubin-Franskin, International Journal of Mass Spectrometry **277** (1-3), 70 (2008).
19. A. B. Trofimov, T. E. Moskovskaya, E. V. Gromov, H. Koppel, and J. Schirmer, Physical Review A **64** (2) (2001).
20. A. B. Trofimov, E. V. Gromov, H. Koppel, J. Schirmer, K. C. Prince, R. Richter, M. De Simone, and M. Coreno, Journal of Physics B-Atomic Molecular and Optical Physics **36** (18), 3805 (2003).
21. S. Carniato, R. Taieb, E. Kukk, Y. Luo, and B. Brena, Journal of Chemical Physics **123** (21) (2005).
22. H. Koppel, F. X. Gadea, G. Klatt, J. Schirmer, and L. S. Cederbaum, Journal of Chemical Physics **106** (11), 4415 (1997).

23      C. Kolczewski, R. Puttner, O. Plashkevych, H. Agren, V. Staemmler, M. Martins, G. Snell, A. S. Schlachter, M. Sant'Anna, G. Kaindl, and L. G. M. Pettersson, Journal of Chemical Physics **115** (14), 6426 (2001).

24      I. Minkov, F. Gel'mukhanov, R. Friedlein, W. Osikowicz, C. Suess, G. Ohrwall, S. L. Sorensen, S. Braun, R. Murdey, W. R. Salaneck, and H. Agren, Journal of Chemical Physics **121** (12), 5733 (2004).

25      I. Minkov, F. Gel'mukhanov, H. Agren, R. Friedlein, C. Suess, and W. R. Salaneck, Journal of Physical Chemistry A **109** (7), 1330 (2005).

26      V. Ilakovac, S. Carniato, J. J. Gallet, E. Kukk, D. Horvatic, and A. Ilakovac, Physical Review A **77** (1) (2008).

27      P. Hohenberg and W. Kohn, Physical Review B **136** (3B), B864 (1964).

28      W. Kohn and L. J. Sham, Physical Review **140** (4A), 1133 (1965).

29      D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).

30      J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77** (18), 3865 (1996).

31      G. Herzberg, *Spectra of Diatomic Molecules*, Second ed. (Van Nostrand Reinhold Company, New York, 1950).

32      R. J. Gdanitz, Chemical Physics Letters **283** (5-6), 253 (1998).

33      R. J. Le Roy, Y. Huang, and C. Jary, Journal of Chemical Physics **125** (16) (2006).

34      T. Bally and G. N. Sastry, Journal of Physical Chemistry A **101** (43), 7923 (1997).

35      D. T. Colbert and W. H. Miller, Journal of Chemical Physics **96** (3), 1982 (1992).

36      F. F. Guimaraes, V. Kimberg, V. C. Felicissimo, F. Gel'mukhanov, A. Cesar, and H. Agren, Physical Review A **72** (1) (2005).

# Chapter 4 - An Analysis of the NEXAFS Spectra of Solid Glycine

1. **Introduction**

   Seeking to understand electronic structure and geometrical structures of biological molecules, much research has addressed amino acids by Near Edge X-ray Absorption Fine Structure (NEXAFS) spectroscopy.[1-18] These studies began with single amino acids, and progressed to dipeptides, wherein it was noted that the formation of a peptide bond causes significant spectral changes.[4,7,15,18] The belief that one could use a "building block" approach to untangle the spectra of proteins was a large driving force behind this approach, viz. that should be able to predict a protein spectrum given the amount and type of each amino acid with an appropriate spectral library and properly accounting for the peptide bond.[19] Unfortunately, while this approach has shown modest success, polypeptides and proteins are not amenable as much as was hoped.[20-22] In this paper, we address the origins of the spectral features of solid glycine, and how this relates to protein NEXAFS spectra.

   Although nitrogen K-edge spectra of glycine have been obtained with widely differing appearances, probably differing due to sample preparation issues, the literature has converged on a single "correct" solid glycine spectrum.[4,11,15-18] This spectrum is dominated by a single peak centered at 406.8 eV, assigned to a N 1s $\rightarrow$ $\sigma^*$ (N-C), and N 1s $\rightarrow$ $\sigma^*$ (N-H).[4,11,15-18] There have been assignments made for low energy shoulders and a feature well above the main peak in energy, however, the latest data indicates a single strong spectral peak, although it is not Gaussian and is almost certainly due to a superposition of transitions.[17]

   Glycine ($NH_2CH_2COOH$) is zwitterionic in its solid form, i.e. the $NH_2$ group becomes an $NH_3^+$ and the COOH group becomes $COO^-$.[4,11,15-18] This is similar to the form found when solvated in water, but different from that in the gas phase.[8,12] The N K-edge spectra are largely unchanged between glycine and glycine chloride, a protonated form of glycine where the $NH_3^+$ group remains, but changes drastically upon being switched to sodium glycinate, where the $NH_3^+$ group becomes an $NH_2$ group.[4,11,15,17,18] The $NH_2$ group leads to several prominent features before the major peak. This has also been observed for glycine in its aqueous anionic form and in several related glycine structures wherein the structure is terminated in an $NH_2$ group. Spectra of these $NH_2$ containing structures are different from those of the dipeptide, which exhibits a low energy peak resolved from the main peak.[4,11,15,17,18]

   As noted, a variety of different spectra of nominally identical solid glycine have been published. However, even when made from nominally identical powders, spectra were shown to vary greatly, likely due to glycine crystallizing extremely readily as well as linear dichroism.[11] Some changes were tentatively assigned to defects in the samples leading to neutral glycine, and to an inability to sample truly random orientations. Other causes could be radiation damage and interactions with adsorbed water.[17]

   Simulating a solid is especially challenging because of the number of atoms involved. However, several groups have made serious efforts toward accurately predicting the spectrum of solid glycine. One study explicitly calculated the spectrum of methylamine and methylamine hydrochloride in the presence of a full core hole (FCH) for both the isolated molecules and for clusters using the Hartree-Fock Hamiltonian, in an attempt to model the nitrogen edges of

glycine with an $NH_2$ and $NH_3^+$ respectively.[11]  The cluster structures were modeled after those of the solid structures.  While this work is certainly informative and explains some of the underlying causes of the observed spectra, the authors clearly elaborate the weaknesses of their approach.  First, although they are modeling a system designed to explain the glycine spectra, they aren't actually modeling glycine, so it is hard to know the accuracy of the approach.  They also use a fairly wide broad empirical linewidth, and although such an approach is common, it often obscures the spectral features, and its use is not adequately justified in the literature.  The only thorough attempt to directly simulate the spectra of solid glycine was undertaken using a multiple scattering code.[17]  In that work, a cluster of 15 glycine molecules was simulated with a Debye-Waller factor of $0.01Å^2$ to account for thermal vibrations.  While the correct general spectral form is reproduced, it comprises unrealistically sharp features, similar to those found in sodium glycinate at low energy.   This could be due to the difficulties in the method used when simulating low-Z elements.[23]

Previous work has shown that density functional theory (DFT) [24,25] can accurately reproduce the excitation energies associated with core-level spectra via total energy differences ($\Delta$SCF or $\Delta$KS).[26]  We model the lowest energy core-level state self-consistently with a full core hole and an associated screened excited electron (XCH).[27]  This differs from the previously described FCH approximation, which ignores the excited electron entirely, or replaces it with a uniform background charge density (in the case of periodic systems).[28] Others have applied multielectron quantum chemistry to simple systems, but generally these methods scale poorly and would be far too expensive for the sampling performed herein, particularly on periodic systems.[29]  Because the antibonding orbitals and Rydberg states probed by NEXAFS can be quite spatially extended,[19] they are sensitive to interactions over a considerable distance from the probed atom.  We have chosen to use plane waves, which are capable of approximating both localized and scattering states.[30]

Past research has attempted to simulate solid glycine NEXAFS spectra using the crystal structure, and perhaps including the effects of the Debye-Waller factors.  This approach has often proved to be inadequate in the gas and liquid state,[30-32] as it neglects the impact of nuclear motion on the electronic transition amplitude; to first order, this is referred to as the Herzberg-Teller effect.[30]  Furthermore, any structures which are significantly different from the ground state will be severely undersampled.  Therefore, we attempted to simulate solid glycine as a crystal at various experimental temperatures by using molecular dynamics (MD) sampling to accurately represent molecular positions.  Herein, we investigate the changes in the NEXAFS nitrogen K-edge spectra of solid glycine, using MD sampling as a function of temperature, and make an explicit attempt to account for the vibrations of the molecules.[33]  We find that the effect of temperature on the spectrum of glycine is quite large.

## 2.  Computational Methods

Our method for computing core-level spectra has been detailed previously.[27,30-32] Briefly, we calculate the absorption cross section to first order using Fermi's golden rule, using the Perdew-Burke-Ernzenhof (PBE) form of the generalized gradient approximation to the exchange-correlation potential.[34] We adopt a plane-wave representation and pseudopotential approximation for valence electronic structure, with norm-conserving pseudopotentials and a numerically converged plane-wave cut-off of 85 Rydberg.

We  adopt the eXcited state Core Hole (XCH) approximation. For x-ray core hole excitations, the intial state is fixed on the 1s atomic eigenstate of interest and we include the important screening presence of the excited electron; the excited nitrogen has the electronic configuration $1s^1 2s^2 2p^4$. Because the atomic nuclei will not move appreciably on the attosecond time scale, the atomic nuclei remain fixed in place during the excitation.  We used approximately 1000 Kohn-Sham eigenstates in constructing transition matrix elements, which is sufficient to extend the transitions approximately 25 eV above the onset.  The electronic structure is calculated using the PWSCF code.[35]  We use the periodic boundary conditions to approximate the continuum of electronic states found at high energy by numerically converging an integration over the Brillouin Zone (BZ).  These delocalized states are similar to the unbound electronic scattering states of the molecule.  Spectra were shifted to match experimental energies, and aligned to each other as previously described.[31]  A total of 125 k-points were used per spectra.[31]

Generally, core-level spectra of solids are simulated based on the crystal structure, usually determined by x-ray or neutron diffraction.  These are well known for glycine.[36]  The crystal unit cell was repeated 4 times to produce a 16 molecule crystal, thus minimizing interactions between individual glycine molecules.  This structure was then minimized both in terms of intermolecular forces and stress on the crystal by modeling the atomic nuclei as fixed point charges, located at an energy minimum derived from a formalism which models the electrons as quantum particles and the nuclei as classical point charges.  We performed this calculation on solid glycine and this result is referred to herein as 0K calculations.

In order to account for the thermal motions, we have modeled the nuclear degrees of freedom in these molecules using molecular dynamics (MD) performed at the specified temperature, ranging from 77K to 450K using a Langevin thermostat and the generalized AMBER force field.[37,38]  The resulting distribution of nuclear coordinates is spaced at least 1 nanosecond apart to eliminate correlation between snapshots for 10 snapshots.[37,38]  All nitrogen atoms in a given snapshot were sampled.

All calculated core-level transitions are numerically broadened using Gaussians of 0.2 eV full width at half maximum (FWHM).  We use this relatively small and uniform broadening scheme with the aim of simulating and distinguishing electronic and vibrational effects explicitly.  Certain features of spectra result from the fundamental energy dispersion of electronic states or the short lifetime of particular transitions, while other features will be caused by molecular motions.[19,30]  We believe that purposely using a small broadening permits a predictive computational approach which can distinguish between electronic broadening and vibrational broadening of spectral features.

## 3.  Results and Discussion

Solid glycine in its most stable form is a zwitterionic ($NH_3^+CH_2COO^-$) monoclinic crystal of space group $P2_1/n$.[44] It's structure is relatively invariant with temperature and there are several hydrogen bonds within the crystal.  An image of 4 units of the crystal structure is shown in Figure 1, as is the structure of zwitterionic glycine.  This is the cell size that was used to simulate a glycine crystal in our calculations. In this figure, several hydrogen bonds are clearly visible between the nitrogen-containing amine group and the carboxylate group.

In order to ascertain the effect of temperature on the crystal, the crystal was simulated at several different temperatures. As the crystal temperature is increased, the glycine molecules rotate more freely. This is evidenced by following the carbon-carbon-nitrogen-hydrogen dihedral angle, as is shown in Figure 2. While there is always a significant barrier to rotation, at higher temperatures the spread of angles which are accessible is significantly larger than that at 77K. There are three major peaks at every temperature because the amine group, $NH_3^+$, is three fold symmetric and there is a minimum energy location for the amine group.

The calculated NEXAFS spectrum as a function of temperature is shown in Figure 3. The results are compared to spectra from the literature at room temperature.[17,22] The calculated nitrogen K-edge spectra of solid glycine at 0K, 77K, 198K, 300K and 450K are shown. The spectral changes are due to conformational changes of glycine, as all the molecules are simulated under the same periodic conditions. It is interesting to watch the evolution of the structure as a function of temperature. At 0K, the structure of all the glycine molecules are identical, and this causes a spectrum with very well resolved features, unlike what is found experimentally at room temperature with most solids. There are two well-resolved features below 405 eV, a large one at 406 eV and other smaller higher energy features. Upon heating the sample to 77K, the features below 405 eV merge together and the features above 405 eV merge together, giving a two-feature structure with a large intensity gap at approximately 405 eV. The spectrum is not smooth however, as both feature have substructures.

When glycine is simulated at 198K, the spectrum is similar to that of the 77K spectra. The features blend together slightly more and the spectrum generally is smoothed compared to the 77K spectrum. By 300K the spectrum comprises only a single large feature. This is the spectrum which can be directly compared to the literature, and although it is slightly contracted compared to the literature results (we believe due to DFT underestimating the bandwidth)[39] the spectra are quite similar. This is more surprising considering that the transitions were only broadened by 0.2 eV FWHM Gaussians, meaning that there are many transitions which comprise the structure. The continuum is slightly too weak relative to experiment, but this is not surprising, given that we do not account for multiple scattering effects. The spectrum appears to acquire its form due to the thermal motions of the molecule, not from an inherent peak width. The spectrum at 450K is similar to that of the 300K spectrum, with a noticeable redshift of approximately 1 eV. Glycine decomposes starting at ~500K, meaning that measurements well above 450K will not be experimentally feasible.

Unfortunately, there is not a simple relationship between the dihedral angles and the calculated spectra. This is shown in Figure 4, which compares the nitrogen K-edge spectra versus the carbon-carbon-nitrogen-hydrogen dihedral angle (for the angle found between 0˚ and 120˚) at 300K. As can be seen, there is quite a bit of variation in the spectral intensity and it does not necessarily correlate simply with the dihedral angle. Similar phenomena are observed with other angles, distances and temperatures, i.e. there is no simple relationship between any simple variable and the spectra. It is important to realize that these changes associated with the different locations of the atom appear to be the major cause of the spectral structure. Even with very narrow transition widths, it is possible to generate the correct spectra because of the large thermal motions of the molecule. It would be a worthwhile future goal for x-ray absorption theorists to correlate the position of atoms (possibly via a normal mode analysis) with spectral peak positions and intensities.

It has occasionally been difficult to reproduce spectra of solid glycine, even under nominally identical conditions.[4,11,15-18] While early spectral inconsistencies could have been due

to sample preparation techniques, previous experiments supported the notion that linear dichroism could be a source of irreproducibility; however, our calculations on the nitrogen K-edge indicate that this effect is limited.[11]  In particular, new techniques avoid this problem by sampling a variety of crystalline domains.[17] In our calculations, it only takes on the order of 100 glycine spectra to converge the solid glycine nitrogen K-edge spectra, implying that avoiding insufficient crystal sampling should be readily achievable. However, other effects such as crystalline domain shifts and neutral glycine impurities could produce significant effects on the spectra.   Furthermore, it is known that these crystals may contain water impurities, which could affect the spectra.[8]

## 4. Conclusion

Nitrogen K-edge NEXAFS spectra of solid glycine have been calculated as a function of temperature, ranging from 0 K to approaching decomposition temperature.  It was found that the thermal motions can cause large spectral variations, and that these motions actually drive the spectrum that is observed at room temperature.  The room temperature spectrum, which is just one large peak, is not due to an inherently broad peak structure, but rather, to the thermal motions of the molecule.  These motions cannot be described by single simple variable, but rather are due to collective motions throughout the crystal.

**Figure Captions**

**Figure 1.** Part a) is an image of four units of the most stable crystal structure of glycine, with oxygen in red, carbon in teal, nitrogen in blue and hydrogen in white. Several hydrogen bonds are visible between the amine group and the carboxylate group. Part b) is the structure of zwitterionic glycine.

**Figure 2.** Histogram of carbon-carbon-nitrogen-hydrogen dihedral angle for glycine as a function of temperature. As temperature increases, the spread in values increases. There are three major peaks at every temperature because the amine group, $NH_3^+$, is 3-fold symmetric and there is a lowest energy location for each of the hydrogen atoms.

**Figure 3.** The calculated spectra of solid glycine at various temperatures ranging from 0K to 450K at the nitrogen K-edge. For all temperatures besides 0K, the shaded region represents 1 standard deviation. The spectra are compared with a room temperature spectrum from the literature (see text). Spectra are offset for clarity

**Figure 4.** Individual calculated nitrogen K-edge NEXAFS spectra plotted as a function of carbon-carbon-nitrogen-hydrogen dihedral angle, for the angle found between 0° and 120° at 300K.
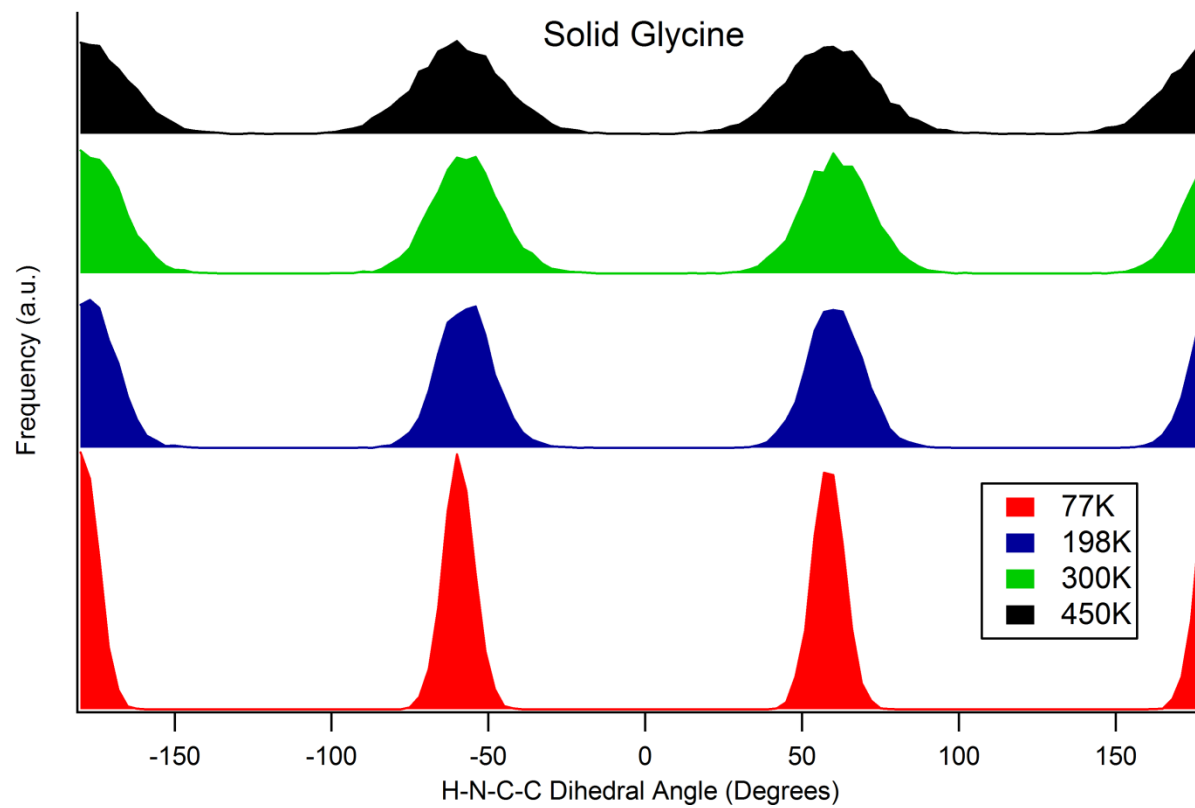
**Figures and Tables**
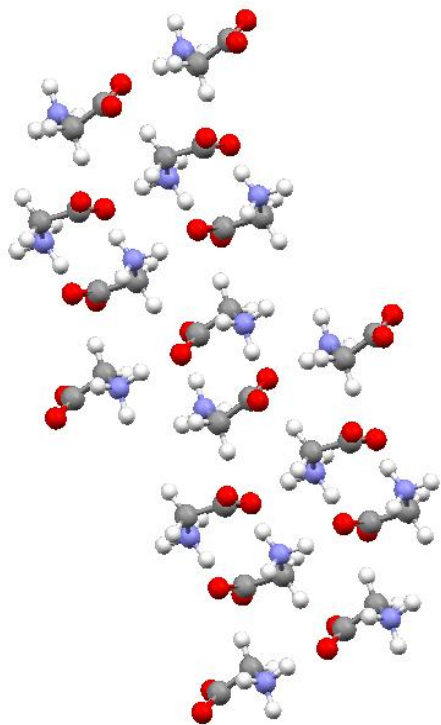
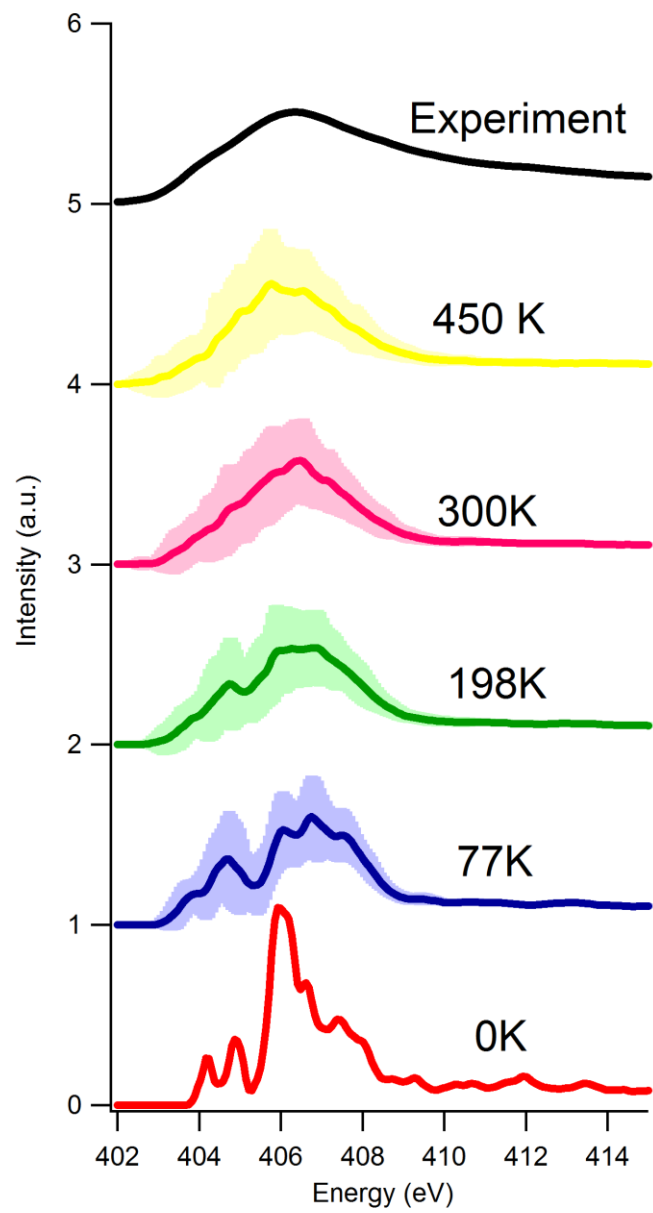

**a)**
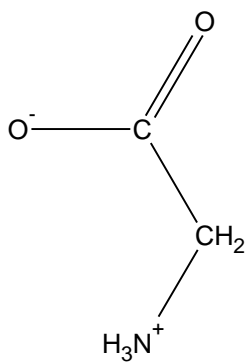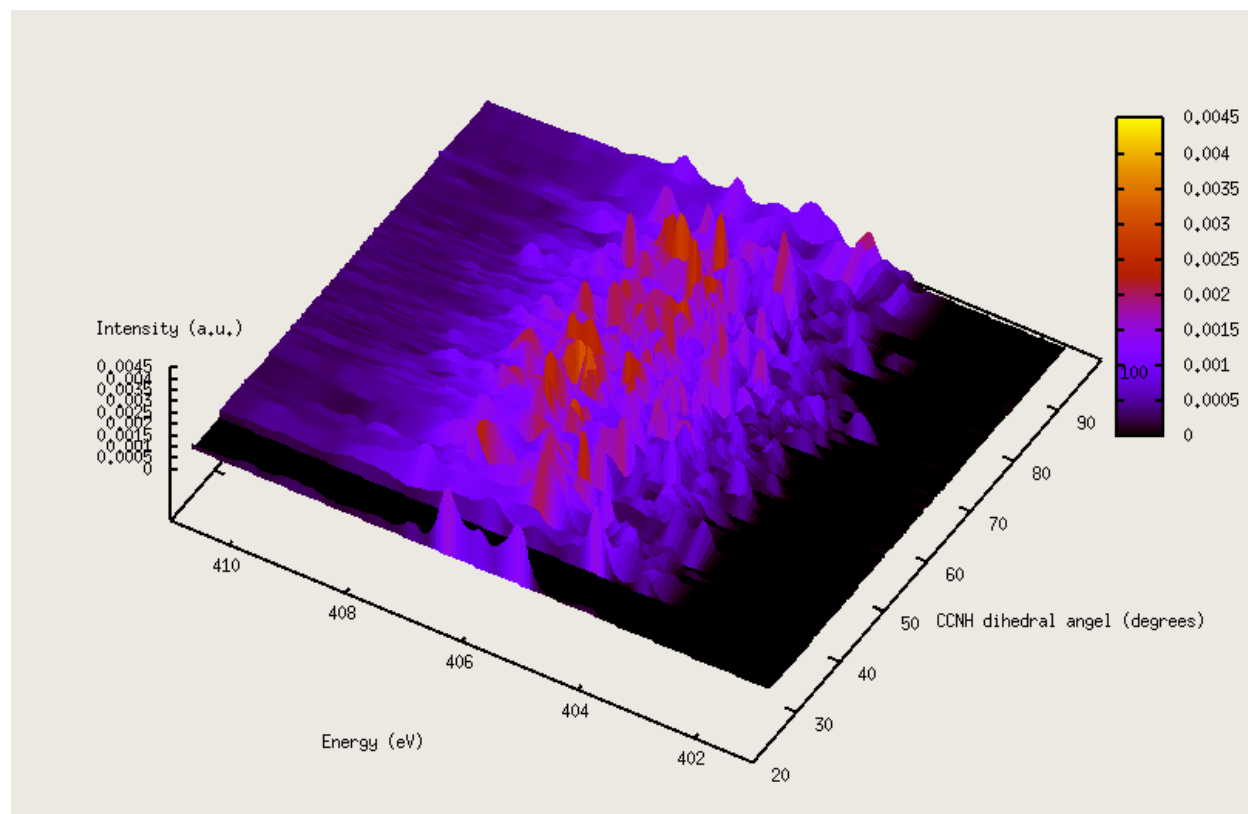


**b)**

**Figure 1**

**Figure 2**

**Figure 3**

63

**Figure 4**

(1)	Carravetta, V.; Plashkevych, O.; Agren, H. *Journal of Chemical Physics* **1998**, *109*, 1456.

(2)	Feyer, V.; Plekan, O.; Skala, T.; Chab, V.; Matolin, V.; Prince, K. C. *Journal of Physical Chemistry B* **2008**, *112*, 13655.

(3)	Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Prince, K. C.; Carravetta, V. *Journal of Physical Chemistry A* **2008**, *112*, 7806.

(4)	Gordon, M. L.; Cooper, G.; Morin, C.; Araki, T.; Turci, C. C.; Kaznatcheev, K.; Hitchcock, A. P. *Journal of Physical Chemistry A* **2003**, *107*, 6144.

(5)	Hasselstrom, J.; Karis, O.; Weinelt, M.; Wassdahl, N.; Nilsson, A.; Nyberg, M.; Pettersson, L. G. M.; Samant, M. G.; Stohr, J. *Surface Science* **1998**, *407*, 221.

(6)	Kaznacheyev, K.; Osanna, A.; Jacobsen, C.; Plashkevych, O.; Vahtras, O.; Agren, H. *Journal of Physical Chemistry A* **2002**, *106*, 3153.

(7)	Messer, B. M.; Cappa, C. D.; Smith, J. D.; Drisdell, W. S.; Schwartz, C. P.; Cohen, R. C.; Saykally, R. J. *Journal of Physical Chemistry B* **2005**, *109*, 21640.

(8)	Messer, B. M.; Cappa, C. D.; Smith, J. D.; Wilson, K. R.; Gilles, M. K.; Cohen, R. C.; Saykally, R. J. *Journal of Physical Chemistry B* **2005**, *109*, 5375.

(9)	Nyberg, M.; Hasselstrom, J.; Karis, O.; Wassdahl, N.; Weinelt, M.; Nilsson, A.; Pettersson, L. G. M. *Journal of Chemical Physics* **2000**, *112*, 5420.

(10)	Nyberg, M.; Odelius, M.; Nilsson, A.; Pettersson, L. G. M. *Journal of Chemical Physics* **2003**, *119*, 12577.

(11)	Otero, E.; Urquhart, S. G. *Journal of Physical Chemistry A* **2006**, *110*, 12121.

(12)	Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Carravetta, V. *Journal of Electron Spectroscopy and Related Phenomena* **2007**, *155*, 47.

(13)	Tanaka, M.; Nakagawa, K.; Koketsu, T.; Agui, A.; Yokoya, A. *Journal of Synchrotron Radiation* **2001**, *8*, 1009.

(14)	Yang, L.; Plashkevych, O.; Vahtras, O.; Carravetta, V.; Agren, H. *Journal of Synchrotron Radiation* **1999**, *6*, 708.

(15)	Zubavichus, Y.; Zharnikov, M.; Schaporenko, A.; Grunze, M. *Journal of Electron Spectroscopy and Related Phenomena* **2004**, *134*, 25.

(16)	Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *Journal of Physical Chemistry A* **2005**, *109*, 6998.

(17)	Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *Journal of Physical Chemistry B* **2006**, *110*, 3420.

(18)	Cooper, G.; Gordon, M.; Tulumello, D.; Turci, C.; Kaznatcheev, K.; Hitchcock, A. R. *Journal of Electron Spectroscopy and Related Phenomena* **2004**, *137-40*, 795.

(19)	Stöhr, J. *NEXAFS Spectroscopy*; Springer: Berlin, 1992; Vol. 25.

(20)	Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *Journal of Physical Chemistry B* **2008**, *112*, 4478.

(21)	Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *Journal of Physical Chemistry B* **2007**, *111*, 11866.

(22)	Stewart-Ornstein, J.; Hitchcock, A. P.; Cruz, D. H.; Henklein, P.; Overhage, J.; Hilpert, K.; Hale, J. D.; Hancock, R. E. W. *Journal of Physical Chemistry B* **2007**, *111*, 7691.

(23)	Rehr, J. J.; Ankudinov, A. L. *Coordination Chemistry Reviews* **2005**, *249*, 131.

(24)	Hohenberg, P.; Kohn, W. *Physical Review B* **1964**, *136*, B864.

(25)     Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, 1133.

(26)     Kolczewski, C.; Puttner, R.; Plashkevych, O.; Agren, H.; Staemmler, V.; Martins, M.; Snell, G.; Schlachter, A. S.; Sant'Anna, M.; Kaindl, G.; Pettersson, L. G. M. *Journal of Chemical Physics* **2001**, *115*, 6426.

(27)     Prendergast, D.; Galli, G. *Physical Review Letters* **2006**, *96*.

(28)     Luo, Y.; Agren, H.; Keil, M.; Friedlein, R.; Salaneck, W. R. *Chemical Physics Letters* **2001**, *337*, 176.

(29)     Duflot, D.; Sidhoum, K.; Flament, J. P.; Giuliani, A.; Heinesch, J.; Hubin-Franskin, M. J. *European Physical Journal D* **2005**, *35*, 239.

(30)     Uejio, J. S.; Schwartz, C. P.; Saykally, R. J.; Prendergast, D. *Chemical Physics Letters* **2008**, *467*, 195.

(31)     Schwartz, C. P.; Uejio, J. S.; Saykally, R. J.; Prendergast, D. *Journal of Chemical Physics* **2009**, *130*.

(32)     Schwartz, C. P.; Uejio, J. S.; Duffin, A. M.; England, A. H.; Prendergast, D.; Saykally, R. J. *Journal of Chemical Physics* **2009**, *131*.

(33)     Chakravarty, C. *International Reviews in Physical Chemistry* **1997**, *16*, 421.

(34)     Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865.

(35)     Baroni, S.; Corso, A. D.; Gironcoli, S. D.; Giannozzi, P. PWSCF, 2008.

(36)     Langan, P.; Mason, S. A.; Myles, D.; Schoenborn, B. P. *Acta Crystallographica Section B-Structural Science* **2002**, *58*, 728.

(37)     D.A. Case, T. A. D., T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R.; Luo, K. M. M., D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S.; Hayik, A. R., G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H.; Gohlke, L. Y., C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C.; Schafmeister, W. S. R., and P.A. Kollman. **2006**.

(38)     Wang, R. L. C.; Kreuzer, H. J.; Grunze, M. *Physical Chemistry Chemical Physics* **2006**, *8*, 4744.

(39)     Prendergast, D.; Grossman, J. C.; Galli, G. *Journal of Chemical Physics* **2005**, *123*.

# Chapter 5 - X-ray Absorption and Photoemission Spectra Do Not Prove Hydroxide to be Hypercoordinated

1. **Introduction**

   Hydroxide is uniquely solvated by water, in that it exhibits unusually fast transport.[1] There have been two competing explanations for this high transport rate. One treats the ion as a water molecule with a missing proton("proton hole"), with a diffusion rate that could be inferred from that of $H^+$(aq). This model also assumes that hydroxide will be unable to donate a hydrogen bond, and will be predominantly solvated as the triple hydrogen bond acceptor, three coordinate $OH^-(H_2O)_3$.[1] The other model is based on $OH^-$ primarily being solvated primarily as the quadruple acceptor, four coordinate (hypercoordinated) $OH^-(H_2O)_4$ which can interchange by proton transfer to $OH^-(H_2O)_3$. This complex can form a transient donor hydrogen bond between $OH^-$ and the oxygen of water.[2-4]

   This latter interpretation has recently been supported by experimental evidence. Based on x-ray scattering data coupled with CPMD, it was shown that an increase in hydroxide concentration (concentrations from ~4.6M – ~12M) leads to a general loss of tetrahedrality.[5] Detailed neutron diffraction studies have determined that the experimental radial distribution function at high concentrations are incompatible with a model that treats $H^+$ and $OH^-$ symmetrically.[6,7] Recent FTIR-ATR studies of dilute HOD in $D_2O$ have also supported the notion that $OH^-(H_2O)_4$ is the most prevalent species in water (concentrations up to ~1M).[8] There is also evidence based on electrochemistry, diffusion, and 2D-IR (concentrations up to ~11M).[9,10]

   There have also been studies of hydroxide in aqueous solutions by x-ray absorption (XAS) and x-ray photoelectron spectroscopy (PES). In XAS, tunable radiation is used to excite the core level electrons to high energy nonbonding and antibonding states; absorption is measured as a function of energy. Potassium hydroxide solutions (concentrations up to 6M) were found to have a new feature at low energy which was attributed to the hydroxide ion itself. By comparing the observed absorption with theoretical calculations of the hypercoordinated and three coordinate hydroxide, it was determined that the absorption data agreed best with the hypercoordinated form.[11]

   Photoelectron spectroscopy, in contrast to XAS, probes the occupied states of molecules.[12,13] The kinetic energy of the escaping electrons is analyzed and the binding energy of the electron is determined. Thus, despite probing at similar energies, XAS and PES provide different, but complimentary information. Photoelectron spectroscopy of hydroxide solutions (~4M) were performed with probe energies that were less than, corresponding to, and greater than the hydroxide feature in energy as observed in XAS. Sharp features were observed,[14] which could be explained only by intermolecular Coulombic decay (ICD). This phenomenon involves the transfer of energy from an atom or molecule to a neighboring atom or molecule via a Coulomb interaction, leading to the ionization of the accepting molecule,[15] which then decays as if it had been excited with that amount of energy. It was proposed that the $OH^-$ transferred energy to its hydrogen bonded neighboring water molecules, which then emitted electrons. This proposed mechanism would thus favor the hypercoordinated model.[14]

   Herein, we perform detailed quantum calculations of one hydroxide solvated by 31 waters in periodic box conditions. These comprise a significant advance over previous results in that we

use a technique which is able to properly simulate liquid water, and therefore can reliably simulate the XAS spectra of a hydroxide solution.[16] Furthermore, we examine the occupied states of molecules which have been core excited into their LUMO, identical conceptually to the excitation occurring during the PES experiment. We determine that the reason ICD was seen for hydroxide and not for the isoelectronic fluoride molecule is unrelated to hydrogen bonding, but rather due to the energy of the electronic states of the core excited molecule.

We utilize a formalism which works well for molecular liquids, employing a plane wave basis set which allows for the accurate description of scattering states. It has been shown previously that density functional theory (DFT)[17,18] can accurately reproduce excitation energies associated with core-level spectra via total energy differences ($\Delta$SCF or $\Delta$KS).[19] We use this to model the core hole caused by the absorption of an x-ray photon, representing the lowest energy core-level excited state self-consistently using a full electronic core hole on the excited atom and an associated screened excited electron (XCH).[16,20,21]

## 2. Methods

Our XAS method has been detailed previously.[16,20,21] Briefly, we calculate the x-ray absorption cross section to first order using Fermi's golden rule. We use the Perdew-Burke-Ernzenhof (PBE) form of the generalized gradient approximation to the exchange-correlation potential.[22] For x-ray core hole excitations at the oxygen K-edge, the intial state is fixed to the 1s atomic eigenstate of oxygen (calculated for the ground state electronic configuration using PBE). We adopt a plane-wave representation and use a pseudopotential approximation for valence electronic structure. The excited state was approximated to be spin-unpolarized. In all of our calculations, we use norm-conserving pseudopotentials with a numerically converged plane-wave cut-off of 85 Rydberg.

To approximate the electronic final state within our Fermi's Golden Rule expression, we adopt the eXcited state Core Hole (XCH) approximation. The core-level excited state of the molecule is approximated by replacing the ground state pseudopotential of the core-excited atom with one which explicitly includes a core-excitation – for oxygen we use the electronic configuration: $1s^1 2s^2 2p^5$. We also include the important screening presence of the excited electron, by incrementing the number of ground state valence electrons by one. The atomic nuclei remain fixed, as they will not move appreciably on the attosecond time scale of this excitation. We analyze the population of the bound states within the XCH approximation. Spectra were shifted to match experimental energies, and aligned to each other as previously described. A total of 125 k-points were used per spectra.[21]

Snapshots of 31 waters and a single OH- were provided by Prof. Mark Tuckerman. The box size was [(9.8615 Å)$^3$], giving a concentration ~1.73 M and the simulations were performed without counterions. Eleven snapshots were used for both the four- and three-coordinate hydroxide solutions.[2] These snapshots were from both plane wave and discrete variable representation (DVR) simulations.[23] These produced similar results, with six of each group of eleven snapshots being taken from DVR simulations for the XAS calculations. For water, six snapshots of water were used from a plane wave simulation of 32 water molecules, meaning a total of 192 calculations were performed for comparison.

For the occupied states analysis, ten snapshots, 5 each of PW and DVR calculations, were analyzed of both three and 4 coordinate hydroxide. The three coordinate snapshots donated a

hydrogen bond whereas the four coordinate snapshots did not donate a hydrogen bond. A core-hole and extra electron was placed on the molecule of interest, in this case hydroxide. The total density of states (DOS) of the system and the projected density of states (pDOS) of the atom of interest were then plotted. This process was repeated for six snapshots of 32 water molecules generating 192 sets of DOS and pDOS. For flouride, 10 snapshots were generated by classical molecular dynamics, with approximately 70 water molecules in periodic box conditions using AMBER and GAFF.[24] In these cases where there were more water molecules, the pDOS was rescaled to be proportional to the pDOS of the cases with fewer water molecules.

All calculated transitions are numerically broadened using Gaussians of 0.2 eV full width at half maximum (FWHM). We use this relatively small and uniform broadening with the aim of simulating and distinguishing electronic and vibrational effects explicitly. Certain features of spectra result from the fundamental energy dispersion of certain electronic states or the short lifetime of certain transitions, while other features will be caused by the motions of the molecules.[12,20] We believe that purposely using a small numerical broadening yields a more predictive computational approach which can distinguish between electronic broadening and vibrational broadening of spectral features.

## 3. Results and Discussion

### Absorption

Shown in Figure 1 are experimental and theoretical calculations of core level spectra of water and hydroxide solutions at the oxygen K-edge. The experimental oxygen K-edge data is from the literature,[11] showing pure water, 4M KOH solution and 6M KOH solution. As the $OH^-$ concentration increases an increase in intensity is found at low energy (~532 eV), a minor increase in concentration is seen in the pre-edge (~535 eV) while the main-edge decreases in intensity (~538 eV) and the post-edge remains relatively unchanged (~542 eV). These results are compared with a variety of theoretical calculations. The calculated spectra of a 1.7M solution of 3-coordinated hydroxide are shown for both the bulk solution and the hydrated hydroxide itself. These spectra are similar to those of 4-coordinate hydroxide, which are also shown for both entire 1.7M solution and the hydrated hydroxide itself. The calculated absorption spectrum of pure liquid water is shown for reference.

First, in agreement with previous work, we can ascribe the low energy feature before the pre-edge principally to $OH^-$. This is supported by noting the strong absorption due to the $OH^-$ itself. However, in contrast to previous work our calculations do not support the contention that NEXAFS is able to differentiate between three and four coordinate $OH^-$ forms, noting that even at 6M concentrations there are still approximately nine times more water molecules than hydroxide molecules, and based on a sum rule each oxygen atom must give identical integrated intensities.[12] While hydroxide certainly perturbs adjacent water molecules, the effect is not large enough to produce observable differences between 3-coordinated and 4-coordinated solvation shells. In fact, the main-edge and post-edge of both $OH^-(H_2O)_4$ and $OH^-(H_2O)_3$ are nearly identical. The pre-edges of these forms are slightly different, but not sufficiently so to differentiate between $OH^-(H_2O)_4$ and $OH^-(H_2O)_3$. Even if one could clearly isolate only the spectra of hydroxide, which is not currently experimentally possible, one still would obtain similar spectra between $OH^-(H_2O)_4$ and $OH^-(H_2O)_3$. Comparing the calculated water spectrum to that calculated for either form of hydroxide, one notes primarily a later onset in absorption and that the pre-edge is more sharply peaked for water.

The new computational approach employed here reproduced the experimental water and hydroxide spectra quite well without tunable variables. For this reason, we believe our spectral calculations are accurate, and since they show only minor differences between three and four coordinate hydroxide, we conclude that NEXAFS is insensitive to hydroxide coordination.

**Bound States**

The dramatic spectral change observed in the photoelectron spectra (PES) of hydroxide as compared to that of water, cannot be attributed to differences in the excited states, as shown previously. Instead, the cause of the sharp PES features is in the nature of the bound states. Specifically, the sharp spectral features were attributed by Aziz *et. al.* to intermolecular Coulombic decay, a phenomenon wherein energy is transferred from an excited state (in this case, the p-state of the excited hydroxide) to a neighboring molecule via a Coulomb interaction.[14] The neighboring molecule (water, in this case) then decays as if being excited with that amount of energy. This was observed previously for clusters of valence excited neon.[15] We note that in the case of hydroxide, the excited state of the initial molecule is core-hole excited.

In the Aziz *et. al.* study,[14] it was noted that the signature of ICD was present only for hydroxide, and not for fluoride or water, attributed to that hydroxide could potentially donate a hydrogen bond, unlike fluoride. Why this phenomena was not seen for water, which will also hydrogen bond, was not explained by the authors. Figure 2 shows the overall density of states (DOS) of a solution of pure water, fluoride anion in water and both three and four coordinate hydroxide in water. The total DOS of the solution is shown in red, the s-type projected density of states (pDOS) of the excited atom is shown in blue and the p-type pDOS of the excited atom is shown in black. From this figure, it is apparent that the p-type pDOS of the excited molecule does not overlap strongly with the water DOS for both fluoride and water. This evidences the absence of ICD for water and fluoride, not hydrogen bonding as previously proposed.[14] Water and fluoride, when core-excited do not have states that mix with the solvent to enable the transfer of energy to a neighboring water molecule. It is necessary for the states of the excited species (in this case the excited water or fluoride) and the acceptor species (water) to mix to enable an efficient transfer of energy. This does not occur in these cases.

In contrast, both 3- and 4- coordinate hydroxide molecules have states that strongly overlap with those of neighboring water molecules, enabling an efficient transfer of energy. Interestingly, both the three and four coordinate hydroxide species have states with a strong overlap with those of neighboring water molecules in the energetic region with the p-type DOS at approximately the same energy. Therefore, PES is insensitive to hydrogen bond donation by hydroxide. It is interesting to note that the p-state of hydroxide that overlaps with water is interacting primarily with the accepting waters, not the donating waters. Finally, we note that the reason the spectra were apparently assignable to ICD is that there is only a single state from which the energy transfer could happen from in the case of hydroxide. If this band were more spread out in energy, the resulting PES from ICD would be correspondingly "blurred out." This is why sharp peaks were obtained; the energy transfer can only occur at a single energy value.

We hypothesize that the reason that the states of water and fluoride don't mix with those of neighboring water whereas, those of hydroxide do is because of the more electronegative nature of both of those, as compared to hydroxide. Our general conclusion from calculations with our new methodology is that these calculations imply that hydroxide has a strong ICD spectra due, not to hydrogen bonding, but rather due to the inherent nature of hydroxide. Therefore, no conclusions as to the details of hydration of hydroxide can actually be deduced from the photoelectron data.

We also note that preliminary calculations indicate that this same phenomenon may occur in CN[-] as well as in amino acids and DNA base pairs, which may be important for radiation ooncology, although the resultant energy transfer will likely occur over a wider energy range, due to the variety of states overlapping with the solvent.[25]

**Other Data**

Having analyzed the relevant x-ray absorption and PES data, we will now turn our attention to the other forms of data that have been used to determine the solvation structure of hydroxide. First, we will focus on the results of scattering experiments. There have been numerous studies of hydroxide solutions by both neutron and x-ray scattering.[6,7,26-28] The data indicate a gradual evolution from four- to three-coordinate hydration as the concentration of hydroxide is increased from 2M to near the solubility limit (>12M). To some extent, it should be true that the average coordination of hydroxide lowers at very high concentration, as there are simply fewer waters in the proximity of hydroxide.

The problem with using scattering data to distinguish between these forms of hydroxide is that the scattering data is overfit, and this is exacerbated when the molecule of interest does not exhibit a strong contrast in its scattering properties from those of the background solvent. In this case, hydroxide will scatter very similarly to water. It has recently been shown that the scattering data for pure water can be fit with anywhere from 2.0 -3.5 hydrogen bonded neighboring waters.[29] For this reason, one could almost certainly find examples where the coordination number of hydroxide is either three or four in the low concentration regions.

There have also been recent studies with Fourier transform infrared attenuated total reflection (ATR) spectra of dilute HDO in $D_2O$.[8,30] In these experiments, a careful subtraction scheme is chosen to help determine the spectral contribution due to hydroxide. It was determined that to recreate the spectra via a computational model, hydroxide must donate a hydrogen bond. However, their model which involved calculating the IR spectra of several small clusters of hydroxide and water (<8 water molecules) and attempting to determine the cause of spectral features is not likely to be adequate, as a liquid is not likely to be well-represented by merely the low energy cluster structures, and assigning spectral features to specific stretches in a liquid is difficult and may not be justified.[31] Furthermore, it is unlikely this small cluster can adequately represent the system. For water, 32 water molecules was recently needed to converge a cluster.[32] The most rigorous way to simulate the infrared spectra would be to perform an MD simulation with both models of hydroxide and calculate the associated infrared spectra.[33]

Recent evidence favoring hydroxide being hypercoordinated has been obtained from 2-dimensional infrared spectroscopy.[9] While the results indicate adequate agreement between theory and experiment, a large fraction of the signal is actually due to bulk water, and information about hydroxide was determined from classical simulations. Furthermore, only a relatively small spectral window was probed.

Finally, it has also been noticed that diffusion occurs much too rapidly for hydroxide when the proton hole model is used, whereas the hypercoordinated model has an accurate diffusion rate.[10] This result is fairly robust to corrections such as tunneling. This is perhaps the single most convincing piece of evidence in favor of the hypercoordinated model of hydroxide.

## 4. Conclusions

Recent studies of the x-ray absorption spectra of hydroxide solutions concluded that the experimental data support the theoretical prediction that hydroxide is primarily hypercoordinated

in the form of $OH^-(H_2O)_4$. Calculations performed with a new method and much more accurate methodology show that x-ray absorption spectroscopy actually provides no support for $OH^-(H_2O)_4$ over the competing proposed $OH^-(H_2O)_3$ structure. Furthermore, recent PES experiments demonstrated the importance of ICD in the spectra of hydroxide, and were interpreted as indicative of hydroxide donating a hydrogen bond. Our calculations show that the origin of the ICD spectra is not hydrogen bond donation, but rather an intrinsic property of hydroxide in water. We conclude that the evidence for hydroxide being hypercoordinated has been overstated, although the proposition is still most likely correct.

**Figure Captions**

**Figure 1.**  Measured and simulated core-level spectra of hydroxide solutions at the oxygen K-edge.  A is the experimental water (pink), 4M KOH solution (blue) and 6M KOH solution (yellow) taken from the literature (see text).  B is the calculated absorption spectrum of a 1.7M solution of OH$^-$ when the OH$^-$ is 3 coordinated.  The overall solution spectra is shown (red) as is the absorption due to just OH$^-$ itself (green), and water (black).  C is the calculated absorption spectrum of a 1.7M solution of OH$^-$ when the OH$^-$ is 4 coordinated.  The overall solution spectra is shown (blue) as is the absorption due to just OH$^-$ itself (purple) and water (black).

**Figure 2.**  The average calculated total DOS (red) of a solution as well as the s-type pDOS (blue) and p-type pDOS (black) of core-excited molecules for solutions of pure water, three coordinate, four coordinate hydroxide and fluoride, averaged from 10 snapshots each.  The excited atoms s-type and p-type DOS intensity is multipled by a larger factor for fluoride to account for its lower concentration in the simulations.  Note how the DOS of the solvating water molecules do not strongly overlap with the excited water or fluoride, whereas hydroxide's DOS strongly overlaps the solvating water molecules DOS whether it is either three or four coordinated.  This is visible as a feature at approximately -8 eV binding energy.  The features from -10 eV to -19 eV tend to be due to the excited atom itself as do the features below -23 eV binding energy.
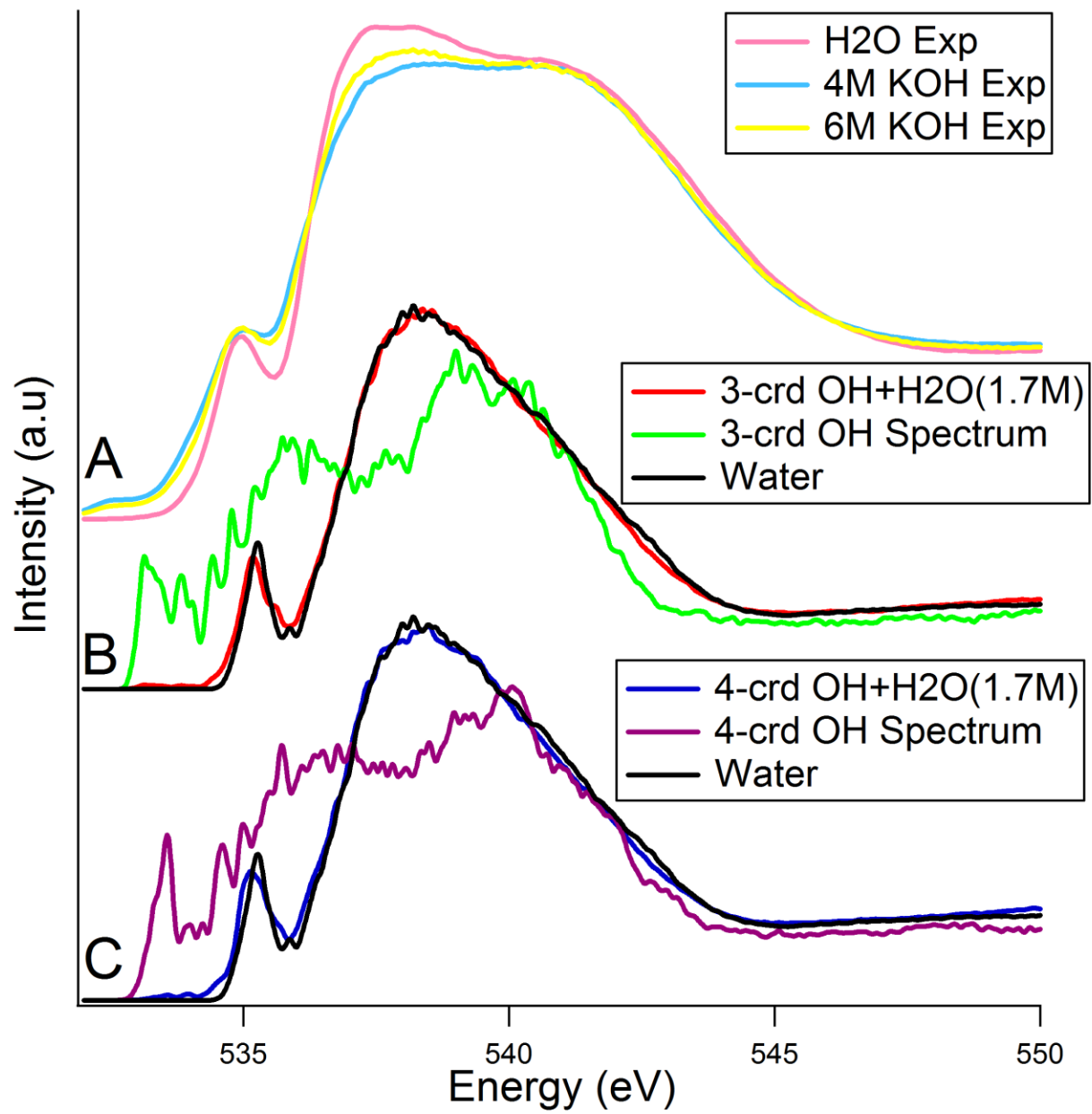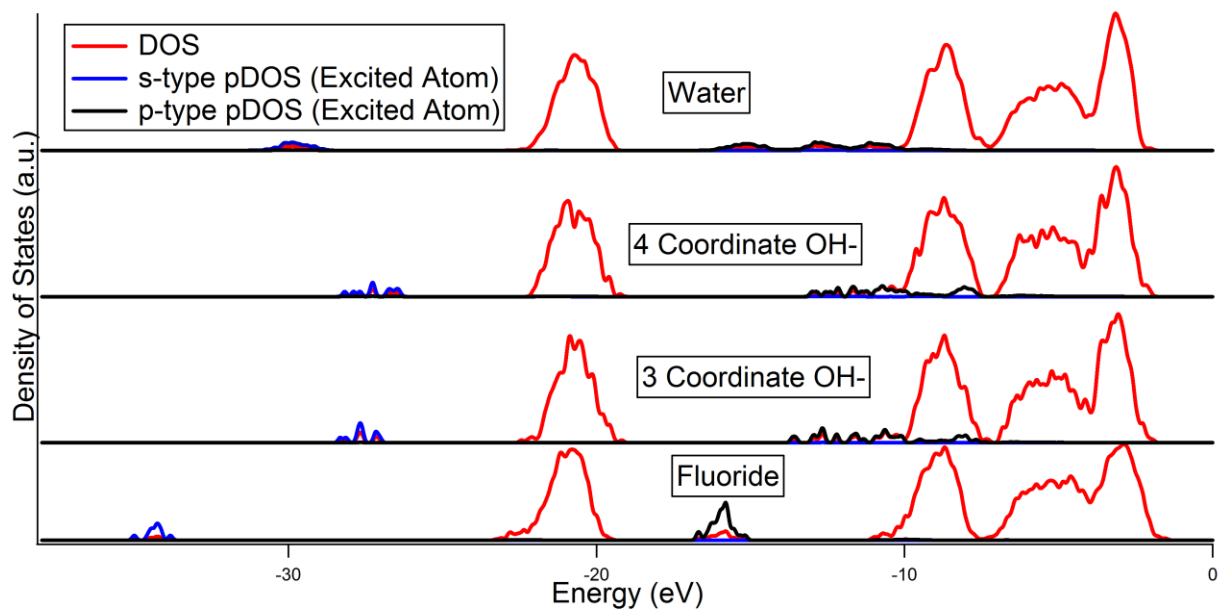
**Figure 1**

**Figure 2**

1    D. Asthagiri, L. R. Pratt, J. D. Kress, and M. A. Gomez, Proceedings of the National Academy of Sciences of the United States of America **101** (19), 7229 (2004).
2    A. Chandra, M. E. Tuckerman, and D. Marx, Physical Review Letters **99** (14) (2007).
3    M. E. Tuckerman, D. Marx, and M. Parrinello, Nature **417** (6892), 925 (2002).
4    M. E. Tuckerman, A. Chandra, and D. Marx, Accounts of Chemical Research **39** (2), 151 (2006).
5    T. Megyes, S. Balint, T. Grosz, T. Radnai, I. Bako, and P. Sipos, Journal of Chemical Physics **128** (4) (2008).
6    S. Imberti, A. Botti, F. Bruni, G. Cappa, M. A. Ricci, and A. K. Soper, Journal of Chemical Physics **122** (19) (2005).
7    A. Botti, F. Bruni, S. Imberti, M. A. Ricci, and A. K. Soper, Journal of Chemical Physics **120** (21), 10154 (2004).
8    M. Smiechowski and J. Stangret, Journal of Physical Chemistry A **111** (15), 2889 (2007).
9    S. T. Roberts, P. B. Petersen, K. Ramasesha, A. Tokmakoff, I. S. Ufimtsev, and T. J. Martinez, Proceedings of the National Academy of Sciences of the United States of America **106** (36), 15154 (2009).
10   A. Chandra, M. E. Tuckerman, and D. Marx, Physical Review Letters **99** (14), 145901 (2007).
11   C. D. Cappa, J. D. Smith, B. M. Messer, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry A **111** (22), 4776 (2007).
12   J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).
13   B. Winter, Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment **601** (1-2), 139 (2009).
14   E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, and B. Winter, Nature **455** (7209), 89 (2008).
15   L. S. Cederbaum, J. Zobeley, and F. Tarantelli, Physical Review Letters **79** (24), 4778 (1997).
16   D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).
17   P. Hohenberg and W. Kohn, Physical Review B **136** (3B), B864 (1964).
18   W. Kohn and L. J. Sham, Physical Review **140** (4A), 1133 (1965).
19   C. Kolczewski, R. Puttner, O. Plashkevych, H. Agren, V. Staemmler, M. Martins, G. Snell, A. S. Schlachter, M. Sant'Anna, G. Kaindl, and L. G. M. Pettersson, Journal of Chemical Physics **115** (14), 6426 (2001).
20   J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).
21   C. P. Schwartz, J. S. Uejio, R. J. Saykally, and D. Prendergast, Journal of Chemical Physics **Accepted** (2009).
22   J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77** (18), 3865 (1996).
23   H. S. Lee and M. E. Tuckerman, Journal of Physical Chemistry A **110** (16), 5549 (2006).
24   T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., K. M. M. Luo, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S., A. R. Hayik, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H., L. Y. Gohlke, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C., and W. S. R. Schafmeister, and P.A. Kollman,  (2006).
25   R. W. Howell, International Journal of Radiation Biology **84** (12), 959 (2008).

26    S. E. McLain, S. Imberti, A. K. Soper, A. Botti, F. Bruni, and M. A. Ricci, Physical Review B **74** (9) (2006).

27    T. Megyes, S. Balint, E. Peter, T. Grosz, I. Bako, H. Krienke, and M. C. Bellissent-Funel, Journal of Physical Chemistry B **113** (13), 4054 (2009).

28    R. Vacha, T. Megyes, I. Bako, L. Pusztai, and P. Jungwirth, Journal of Physical Chemistry A **113** (16), 4022 (2009).

29    K. T. Wikfeldt, M. Leetmaa, M. P. Ljungberg, A. Nilsson, and L. G. M. Pettersson, Journal of Physical Chemistry B **113** (18), 6246 (2009).

30    M. Smiechowski and J. Stangret, Journal of Molecular Structure **834**, 239 (2007).

31    J. D. Smith, C. D. Cappa, K. R. Wilson, R. C. Cohen, P. L. Geissler, and R. J. Saykally, Proceedings of the National Academy of Sciences of the United States of America **102** (40), 14171 (2005).

32    M. P. Ljungberg, A. P. Lyubartsev, A. Nilsson, and L. G. M. Pettersson, Journal of Chemical Physics **131** (3) (2009).

33    H. S. Lee and M. E. Tuckerman, Journal of Chemical Physics **126** (16) (2007).

# Chapter 6 - Auto-Oligomerization and Hydration of Pyrrole Revealed by X-ray Absorption Spectroscopy

## 1. Introduction

Only slightly soluble (47 g/L) in water,[1] pyrrole is a very weak base due to the partial delocalization of its nitrogen lone pair electrons. It reacts via electrophilic substitution and oligomerizes readily upon contact with light or air.[2] As a heterocyclic aromatic component of important macrocycles, including hemes and chlorophyll, it is an important prototype for detailed study. Here, we investigate pyrrole using X-ray absorption spectroscopy (XAS) experiments and a promising new theoretical formalism, seeking to explore the agreement between theory and experiment, anticipating future studies of similar, but more complex molecules.

The development of near edge X-ray absorption fine structure spectroscopy (NEXAFS) of liquid microjets has provided a useful new tool for characterizing the details of solvation.[3] NEXAFS probes the unoccupied molecular orbitals, which are highly sensitive to intermolecular interactions.[4,5] This new approach to the study of liquids is yielding important insights into the behavior of aqueous systems, but the chemical information that can be extracted from the measurements is currently limited by the available theoretical methods for computing core-level spectra. Here, we use the newly-developed excited core hole (XCH) method, which allows for accurate calculations of XAS spectra. This study provides a good test of the ability of XCH to describe highly delocalized bonding systems, which are problematic for the methods now commonly used.

NEXAFS spectra of water differ greatly upon changing from solid to liquid to gas,[5,6] as do those of glycine when comparing the solid and gaseous forms to the form solvated in water.[3,7] However, these molecules exhibit strong intermolecular interactions either via hydrogen bonding or by modifying of their charge state, so large spectral changes are expected. In both microwave and theoretical studies of clusters comprising one pyrrole and one water molecule, the N-H group appears to interact exclusively with the water, preventing water from interacting with the ring.[8,9] It has been known for some time that the vibrational spectrum of the N-H bond of pyrrole is red-shifted when dissolved in liquid water versus when in neat solution.[10] Both these results imply that the N-H group has strong intermolecular interactions with liquid water. This makes pyrrole an interesting case, because the ring carbons of pyrrole should interact with water in a manner similar to those of benzene, whereas the N-H group can interact strongly. It was previously thought that hydrophobic solvation involved generating "icebergs" in water because the hydrophobic solute would structure the water.[11] This is no longer believed to be the case, as deduced largely from classical molecular dynamics simulations.[11,12] However, such simulations are incapable of describing detailed quantum effects, such as the extent of orbital mixing between pyrrole and water, for which NEXAFS is a sensitive probe.[13] Previous study of pyrrole has shown that its solid form has very similar NEXAFS spectra to the inner shell electron energy loss spectroscopy (ISEELS) measurements of pyrrole in the gas phase.[14-17] This implies that in the solid, the orbitals of pyrrole do not exhibit significant hybridization.

Pyrrole is known to be very reactive, and previous studies have observed particles of micron size in liquid pyrrole.[18] It polymerizes readily in the presence of a platinum electrode[19] and has important autopolymerization properties for controlling the growth and properties of

gold nanoparticles.[20]  In fact, neat pyrrole is known to oligomerize readily when exposed to either air or light.[2]  However, it will not react if exposed to neither.  Unfortunately, under realistic conditions, pyrrole suffers exposure to stray light and  air, leading to impure samples over time.[21]  This process is accelerated in water at neutral pH.  It is believed that in pH-neutral water, the oligomerization proceeds by dimerizing, which, if then repeated many times,  would produce the polymer.[22]  This reactivity of pyrrole is important for X-ray absorption studies because, unlike in vibrational spectroscopy, spectral features resulting from  reaction products will overlap with the pure pyrrole spectrum.

There are several methods of obtaining high-resolution X-ray absorption spectra of liquids; in our case small jets of water solutions (~30 microns) were windowlessly coupled to a synchrotron beamline.  An advantage of this method is that the liquid is not  affected by interactions with silicon nitride windows; this is useful, as we want to probe the nitrogen K-edge. Additionally, the solution is constantly renewed,  thereby avoiding sample damage.[23]  One of the biggest advantages is that the temperature can be easily controlled by  allowing the liquid to evaporatively cool.[24]  Furthermore, one can easily obtain a sample of the vapor, enabling a direct comparison of liquid and gas spectra.

In order to model the pyrrole spectra, density functional theory (DFT) is used within the generalized gradient approimxation (GGA).[25]  DFT has exhibited reasonable accuracy in reproducing core level spectra via total energy differences ($\Delta$SCF or $\Delta$KS).[26]  There are several methods of calculating NEXAFS spectra,[27,28] but we choose to model our system  using a full core hole and the associated lowest energy excited electron (XCH).[29,30]  We represent the electronic structure using plane-waves, in order to accurately describe localized bound states as well as scattering states that are often found above the ionization and spatially extended Ryberg-like states.  Linear combinations of atomic orbitals are often inadequate for describing such extended states.[28]  We approximate the high energy continuous electron density of states by using periodic boundary conditions and numerically converging an integration over the first Brillouin zone (BZ).  For extended states that span the supercell, an accurate determination of the electronic density of states can be achieved by such BZ sampling.[31]  We simulated the atomic configurations of pyrrole both in the gas phase and aqueous solution with classical molecular dynamics.  This method has the advantage of capturing the impact of nuclear motion on the electronic transition amplitude.[32]  We note that the same techniques have been applied successfully to interpreting the spectra of solvated amino acids.[33]

## 2.  Methods

**Samples:**  Pyrrole was obtained commercially from Sigma-Aldrich, with stated purities of at least 98%.  All water used had a resitivity of 18 M$\Omega$/cm.  Samples were used without further purification.  The pH of the solution was approximately 7 and the molecule is expected to be  predominantly  neutral  at  this  pH.[34]   The  p$K_{aH}$  of  pyrrole  is  approximately  -4.   The concentration of the solution used was ~0.5 M.  We note that the sample was not  fresh and may have undergone oligomerization prior to our use.

**NEXAFS Spectroscopy of Solvated and Gaseous Pyrrole:**   Total electron yield (TEY) X-ray absorption spectra (XAS) were recorded at the carbon K-edge (~300 eV) and nitrogen K-edge (~400 eV).  These measurements were performed at Beamline 8.0.1 of the Advanced Light Source (ALS) at Lawerence Berkeley National Laboratory in Berkeley, CA.  A detailed description of the experimental system has been published previously.[23]  Briefly, an

intense ($> 10^{11}$ photons/sec), high resolution ($E/\Delta E > 4000$) tunable X-ray beam is generated from an undulator at the ALS.  The synchrotron light is then focused (~50 μm spot size) onto a small liquid jet (~30 μm diameter) and the TEY is collected to obtain spectra of the bulk liquid.[5] The jet is produced by using a syringe pump (Teledyne-Isco) to pressurize the liquid behind a fused silica capillary tip.  The jet travels parallel to the polarization of the incident radiation. Almost immediately (~0.5 mm and ~15 μs) after leaving the tip orifice, the liquid is intersected by the X-ray beam, wherein the sample is close to room temperature (~ 21.5°C ± 1.5°C).[24] Additional measurements were taken downstream (2 cm), allowing the jet to evaporatively cool before being exposed to the X-ray beam.  The energy step size used was 0.1 eV in all scans.  The liquid jet is then condensed in a separate cryogenically cooled section of the chamber to maintain low pressures (~$10^{-4}$ torr) in the interaction region.  In order to collect a vapor spectrum, the jet is moved out of the X-ray beam and a spectrum of the vapor present near the jet is taken.

This approach provides several advantages over other emergent methods for studying X-ray spectra of liquids.  Typical pressures in the main experimental chamber are ~ $10^{-4}$ torr, but two differential pumping sections allow the beamline to maintain UHV conditions without using windows, which would limit the flux.  Furthermore, sample damage is minimized by employing this constantly renewing source.  The low operating pressures permit the use of sensitive charged particle detection as a function of X-ray energy (i.e. action spectra).  The TEY is detected by locating a small (~1 $mm^2$) positively biased copper electrode close (<5 mm) to the jet.  The detected electron current is amplified and converted to a voltage before being read out by the beamline computer.  The signal is normalized to that from a gold mesh located ~3 meters up-beam of the chamber.  The energy is calibrated based on specific metal impurities located in the beamline optics and the spectra are area normalized.

**Calculations -**

**Core-level spectra:** Our theoretical approach has been described previously.[29,32] We employ periodic boundary conditions using a plane-wave basis that is accurate for both localized and delocalized states.  The zone-center electronic structure is calculated using the PWSCF code.[35]  Core-hole matrix elements with valence electrons were calculated by reconstructing the core region of the excited atom of the pseudo states within the atomic frozen core approximation. Transition amplitudes are estimated in the single-particle and dipole approximations; excitations to states above this first excited state are approximated using the unoccupied Kohn-Sham eigenstates computed from the excited core hole (XCH) self-consistent potential.  We use the Perdew-Burke-Erzenhoff (PBE) form of the generalized gradient approximation to the exchange-correlation potential.[36]  These calculations were performed on the Franklin supercomputer at NERSC.

**Isolated Pyrrole:**  Large simulation cells are used (($20$ Å$)^3$) to reduce spurious interactions between cells for the isolated molecule and so as to be large enough to accurately represent excited states below the ionization potential (IP).  Approximately 100 Kohn-Sham eigenstates are used in constructing transition matrix elements.  Due to the large box size, a large number of unoccupied states are needed in order to describe high-energy transitions.

**Solvated Pyrrole:**  Significantly smaller cells ((~$13$ Å$)^3$) were sufficient to reduce the spurious interactions between hydrated molecules.  A single pyrrole molecule was solvated by 81 water molecules, under periodic boundary conditions.  Approximately 1000 Kohn-Sham eigenstates are used to extend the spectra approximately 20 eV above the LUMO energy (absorption onset).

**Molecular Geometry:** Generally, core-level spectra of isolated molecules are simulated within the fixed-nuclei approximation,[26] particularly for molecules in their vibrational ground state. The lowest energy geometrical structure is determined by minimizing the forces derived from the Born-Oppenheimer Hamiltonian, a formalism which models the electrons as quantum particles and the nuclei as classical point charges. We generate spectra from this single molecular configuration on isolated pyrrole molecules and these are referred to herein as fixed-nuclei calculations.

Unfortunately this will be inadequate to properly describe the intrinsic spectral broadening, as a major source of broadening can be caused by thermally excited vibrations and zero-point motion of the molecule. These motions can lead to new transitions and large shifting of transition energies. Within the Born-Oppenheimer approximation, the effect of such motions on the transition strength is known to first order as the Herzberg-Teller effect.[37] In order to account for these effects, we have modeled the nuclear degrees of freedom in these molecules using molecular dynamics (MD) performed at 300 K using a Langevin thermostat, with the generalized AMBER force field. The resulting distribution of nuclear coordinates was sampled at 20 picosecond intervals to eliminate correlation between snapshots. This was done for 100 snapshots for both the isolated and solvated case of pyrrole.[38] These coordinates we subsequently used in DFT calculations which were by far the most significant part of the computational cost.

All calculated spectra are numerically broadened using Gaussians of 0.2 eV full width at half maximum. Others often resort to ad-hoc energy dependent broadening schemes.[39] We use this relatively small and uniform broadening with the aim of simulating and distinguishing electronic and vibrational effects explicitly, thereby arriving at a predictive computational approach.

## 3. Results and Discussion

### Carbon K-edge

Figure 1 shows the carbon K-edge spectrum of pyrrole in the gas phase measured using both ISEELS and NEXAFS.[14,17] Five major features have been identified in the spectra (labeled 1-5). Another feature has been reported previously in ISEELS studies at approximately 300 eV; however, the evidence for this feature, given the signal to noise ratio, is not compelling. Furthermore, this feature is not evident in the NEXAFS data, indicating that it may be due to a non-dipole transition in ISEELS. Calculations for both fixed-nuclei and MD sampled spectra are shown for comparison. Carbon 1 (C1) refers to the two identical carbon atoms bound only to other carbon atoms, whereas carbon 2 (C2) refers to the two carbon atoms bound directly to the nitrogen atom of pyrrole. The fixed-nuclei spectra are dotted lines, while the average of 100 MD snapshots is shown as the solid line, with a shading indicating one standard deviation. Calculated spectra were aligned to the experimental IP of C2, 290.8 eV.[14,17] The relative alignment of C1 and C2 is obtained by total energy differences.[32] The first feature was originally believed to comprise two transitions, however, higher resolution spectra[17] and our high resolution NEXAFS clearly show this feature to comprise at least three transitions. The first two transitions within Feature 1, centered at 285.7 and 286.2 eV, respectively, have been assigned as $1sC1 \rightarrow 1\pi^*(3b_1)$ and $1sC2 \rightarrow 1\pi^*(3b_1)$. This is confirmed by our most recent theoretical calculations. The difference in peak positions corresponds to the difference in XPS binding energy of 1.0 eV between the two carbons. The molecular motions of the molecule substantially

81

broaden these features in theoretical calculations. In the case of carbon 1, the feature is also substantially red shifted by motions of the molecule. The underestimation of the bandwidth in DFT leads to spectra which appear contracted with respect to experiment.[29] The third segment of Feature 1, visible as a shoulder at approximately 287.1 eV, has previously been assigned to a combination of a $\pi^*/\sigma^*$ orbital transition from carbon 1;[17] however, our DFT calculations did not reproduce this feature with adequate intensity. The shift is too large to be caused by vibrations; our hypothesis is that this feature arises from chemical impurities. We did witness the same transitions observed in previous calculations, but as in previous calculations the intensity is too weak to explain the observed intensity.[17] This will be discussed in more detail later. Feature 2 was previously assigned to a feature from C2, primarily involving a $\pi^*$ transition, in agreement with our calculations. Feature 3 was never observed experimentally before, but was predicted by earlier calculations.[17] Our calculations indicate that it is due to a superposition of transitions, some of which involve $\sigma^*$ orbitals. It is primarily C2 that causes this feature; however, much of the associated background intensity appears to come from C1. Features 4 and 5 are due to both carbons and have been assigned as due to $\sigma^*(C-N)$ and $\sigma^*(C-C)$ respectively.[14,17] They are clearly scattering features, and one of the advantages of a plane wave basis set is the ability to capture such scattering phenomena. Gaussian orbital calculations would have difficulty reproducing either feature.

All of the principal observed spectral features are captured, except for a shoulder of Feature 1. Our DFT calculations indicate that there are two types of transitions in this energy range. Their assignment is in agreement with previous configuration interaction quantum chemical calculations.[17] However, the spectral intensity of these features is insufficient to reproduce the observed shoulder of Feature 1. Higher level GW-corrected DFT calculations (not shown) were unable to reproduce the shoulder, and clearly, even with the inclusion of molecular motions, we were unable to reproduce this feature within DFT (Figure 1). This indicates that pyrrole is somehow different from isoelectronic furan, whose shoulder region can be more accurately reproduced by electronic structure calculations.[40]

Thus there must be some other cause for the peak intensity. We ruled out the possibility of pyrrole-pyrrole self interactions by examining the nonbonded dimer, consisting of two weakly interacting pyrroles.[41] Calculations indicate that, if present, this complex does not provide sufficiently intense signal in the shoulder region (~287.1 eV), as is shown in Figure 2. There is literature precedent indicating the dimerization of pyrrole in aqueous solutions, so this possibility was investigated.[2,22] It should be noted that this process can occur merely upon exposure to light or air.[21] We believe that the pyrrole used in this study reacted while stored in a bottle. If our interpretation is correct, this likely has happened as well in previous pyrrole studies.[17] Our pyrrole sample likely further reacted when placed in water. For this reason, much of the dimer may actually have been present before being added to water. When two pyrrole molecules are chemically bonded, as is shown at the top of Figure 2, a strong shoulder develops on Feature 1. In its lowest energy conformation, the dimer is planar, but thermal excitations at room temperature should enable dihedral angles between the rings of up to ~15˚. The shoulder increases in intensity as the molecule is twisted. The shoulder is due primarily to the carbons which bond upon dimerizing, labeled $\alpha$ in Figure 2. The spectra of the trimer, as well as polypyrrole, were also calculated in their vibrational ground states (within the fixed-nuclei approximation). The trimer consists of another pyrrole ring binding to the dimer via the $\alpha$ carbons, which could form due to subsequent reactions of the dimer in water. If the process were continually repeated, the polymer would be formed. As the oligomer chain grows longer, the

shoulder region grows more intense, primarily due to the carbons bonding to other pyrrole rings. Due to this strong feature emerging in oligomerized pyrroles, we can reliably ascribe the shoulder found in the NEXAFS and ISEELS spectra to chemically bonded pyrrole. We note that this feature may not be caused solely by water. In earlier work in which pyrrole was not dissolved in water, a similar, albeit slightly less intense, shoulder was seen.[17] It is likely that the chemical species which make up the shoulder can form in pure liquid pyrrole but will react even more rapidly when exposed to liquid water.

For illustrative purposes, we show in Figure 3 that when the planar dimer is combined with the monomer in a 1:1 ratio, all the observed low energy features are obtained. In reality, mixture of chain lengths of pyrrole oligomers would probably be present, as would more extensive molecular motions, which were not sampled here. Therefore, we are unable to determine the fraction of pyrrole that actually oligomerized. Our spectra should not be considered as quantitative, but merely as an explanation of the origin of the shoulder to Feature 1. We note that this result indicates a significant presence of oligomers. An oligomer free sample will require maintaining an air- and light-free environment after purifying the pyrrole. However, for studies under aqueous solvation oligomerization is likely unavoidable due to reactions with water.

Figure 4 displays two experimental carbon K-edge spectra of fully solvated pyrrole taken at different temperatures, and theoretical calculations for the associated carbons, with error bars corresponding to one standard deviation. The spectra look quite similar to that of the gas phase molecule, and there is only a minor temperature dependence. (Similar to the gas phase spectra, the shoulder at 287.1 eV is found here as well). In particular, the most intense part of the first feature grows slightly with an increase in temperature. This would indicate that there is less motion of the molecule at higher temperature. We believe this should be unlikely, and it is likely due at least partially to normalizing over a temperature range that is too small. It is likely that there is extra intensity outside of the temperature range relative to the low temperature spectra due to shifting of EXAFS peaks. This would cause the corresponding low energy peaks to shrink in intensity. Detector drift may also be a factor. The first feature is red-shifted by ~0.1 eV at lower temperature, but the measurement step size was also 0.1 eV so this shift cannot be quantitatively determined. This shift in energy would likely correspond to a slight change in solvation, as the change in oligomerization should be minor between these two temperatures. At higher temperature the water will likely solvate pyrrole slightly differently than at lower temperature. Feature 3 disappears in the spectrum taken at ~295 K, but reappears to a small extent upon cooling to 275 K. This may be due to increasing gas phase pyrrole background as one moves farther down the jet. The overall spectrum remains largely invariant over the ~20˚ C temperature range.

Comparing the liquid spectra measured at 295K to the calculated liquid spectra, one reproduces the correct number and location of features, and generally the proper form of the spectra. Feature 1 still lacks the observed shoulder, since auto-oligomerization was not included in the solvated simulations, and Feature 3 disappears in the simulated spectra, as found experimentally. The calculations correctly predict the blue-shift of the large 4th feature relative to that of the gas phase spectrum. By including a large number of unoccupied states in the solvated calculations, the fifth feature could also be accurately reproduced. This was not done in the gas phase calculations and therefore the fifth feature was not as well resolved for that case. When the calculated solution phase spectra are aligned with the calculated gas phase spectra, there is remarkably little difference between the two. Feature 3 disappears, Feature 1 broadens

slightly, and Feature 4 blue shifts upon condensation. We speculate that the changes to Feature 4 arise from the fact that the scattering states of the isolated molecule must be orthogonal to surrounding water states when solvated, thereby increasing their kinetic energy.

We note that we have not calculated the spectra of the oligomerization products of pyrrole in water due to computational expense. However, given the large similarity between the calculated spectra of pyrrole in the liquid and the gas we feel we can assume that the spectral change upon the products of auto-oligomerization are roughly identical both spectrally and in ratio as compared to the gas. This is not surprising in that the gas is sampled directly off the evaporating liquid, and thus should be similar in composition. The reaction should occur in the liquid, although due to the large time the liquid must rest in the pump before it can be sampled, it is extremely difficult to estimate time scale of this reaction.

The carbon K-edge spectrum of solid pyrrole has previously been measured.[15] Unfortunately, the resolution of that experiment was quite low (0.8 eV), and for this reason, it is difficult to discern changes due to solidification of the pyrrole and those due simply to the low resolution of the experiment. However, the spectrum is extremely similar to the gas phase ISEELS data (Feature 1 can be described by two Gaussians, unlike the higher resolution NEXAFS data); Feature 3 would not be visible due to the low resolution. Features 2, 4 and 5 are similar to the ISEELS data. The full width at half maximum of the first feature in the solid is similar to that of the gas phase ISEELS data. We note that the solid may have contained chemical impurities as well.

The effects of solvation on the carbon K-edge NEXAFS spectra of pyrrole are relatively minor, comprising a small shift in location of a $\sigma^*$ feature (feature 1) and the disappearance of a small feature (Feature 3, which is a also $\sigma^*$ feature). The transition underlying Feature 3 is to a large spatially extended state around the molecule in the gas phase; however, in the liquid, there are neighboring water molecules, interacting with the aromatic ring. Therefore, the disappearance of this feature is due directly to the ability of nearby water orbitals to mix with the relevant high-energy molecular orbitals, causing Feature 3 to disappear. This is likely Rydberg quenching or a similar phenomena.[13] In summary, our theoretical spectra correctly predict the subtle shift in feature locations and the disappearance of Feature 3 upon solvation, providing an encouraging benchmark for XAS theory.

**Nitrogen K-edge**

The experimental and theoretical calculations for the nitrogen K-edge of gas phase pyrrole are presented in Figure 5. This data was shown in a different format in previous work.[29] The calculated theoretical spectra are aligned to the experimental IP at 406.1 eV. The ISEELS data were taken previously, and are similar to the experimental NEXAFS vapor spectra.[14,16,17] The large first feature, Feature 1, has been assigned to a nitrogen 1s $\rightarrow$ 1$\pi$*(3b$_1$) feature. The second feature, Feature 2, has been variously assigned to a combination of transitions including those to the nitrogen 4p state and a variety of $\sigma$* features, including $\sigma$*(C-N) and $\sigma$*(C-H). The high energy resonance, Feature 3, at 412 eV has been assigned to a $\sigma$*(C-C) resonance.[14] Our calculated spectra are shown on the same figure, the dotted line depicting the spectra calculated with the fixed-nuclei approximation and the solid line depicting the spectra calculated with the average of 100 MD snapshots, with one standard deviation shaded. The first two features broaden and weaken in intensity by sampling a variety of snapshots. The first two features of the experimental spectrum are well reproduced in the gas phase calculations. We note that the second major feature is caused by a variety of resonance states, in agreement with previous

assignments, which have large spatial extent. Unlike the carbon K-edge, the nitrogen K-edge spectrum is largely unaffected by the oligomerization of pyrrole. The flat dimer, trimer and polypyrrole spectra are similar to the monomer spectrum (not shown).

The spectra for pyrrole solvated in water are shown in Figure 6. The experimental spectra are taken at two different temperatures, ~295K and ~275K. It is worth noting that the spectra exhibit almost no change with temperature. The spectrum of pyrrole solvated in water was calculated, with the shaded error bar corresponding to one standard deviation over the 100 sampled snapshots. This calculation agrees well with experiment, accurately reproducing all three features.

The experimental liquid and gas phase spectra of the nitrogen K-edge exhibit features that remain largely unaltered upon being solvated by water. It is surprising that the orbitals appear to mix so little, particularly in light of the previous work on clusters, which predicted an effect similar to charge–transfer-to-solvent for the HOMO to LUMO transition.[9] A large spectral change is not observed in solvated pyrrole in the X-ray region, despite the nitrogen interacting with water. Upon solvation, the first feature red-shifts ~.3 eV and the FWHM broadens from 0.9 eV to 1.0 eV. The second major feature is largely unaffected by solvation, which is somewhat surprising, as this is a scattering feature and should be sensitive to molecular interactions. However, the geometry of the molecule could be largely unaffected by solvation and the energy of a scattering feature is largely dependent upon intramolecular bond lengths. It is possible the lack of feature movement could be a consequence, despite the large spatial extent of the scattering.[42] The third feature is also largely unaffected by solvation. A comparison to solid pyrrole is quite difficult due to low resolution at the nitrogen K-edge beyond mentioning that the three major features are still intact, meaning that in the solid the underlying states remain largely unchanged.[15] The alignment used in the case of the solid was merely to align the solid spectra with the gas spectra from the literature, so discussing relative feature shifts is impossible.[15] The overall background at high energy is larger in the solid, and the ratio of Feature 1 to Feature 2 is similar to what was calculated for the solution.

The theoretical calculations show changes between the condensed and gaseous spectra similar to what is found experimentally. If Feature 2 is used as the alignment between the computed and experimental spectra for solvated pyrrole, the first feature red shifts by .3 eV upon condensing, and Feature 1 widens by .1 eV at FWHM, both identical to experiment. All the features lack the proper width, however, due to DFT's systematic underestimation of the bandwidth.[32]

While the carbons in pyrrole are somewhat hydrophobically solvated, the hydrogen bonded to the nitrogen should hydrogen bond with water. In our simulation, the mean distance between that hydrogen and the nearest oxygen of water is 2.34Å with a standard deviation of 0.36 Å. Although no new electronic transitions are introduced in the region of the first spectral feature at 402.4 eV, the interactions with water can cause this feature to shift in energy, explaining why the feature is so broad when pyrrole is solvated. This same effect also leads to the apparent red-shift. This antibonding orbital delocalizes to a small extent and hybridizes with the oribtals of nearby water molecules, which lowers the transition energy and due to conformational variations, broadens the feature. The second feature exhibits no change upon solvation, as predicted theoretically. Scattering features such as Feature 2 are believed to be highly sensitive to intramolecular distances, which should not change much upon solvation or temperature.[42]

Finally, we describe the orbital mixing between water and pyrrole. This leads in part to the decrease in intensity of Feature 1 observed upon solvation. The two transitions that contribute to Feature 1, the LUMO and LUMO+1, are separated by ~.02 eV in the gas phase. Shown in Figure 7 are the LUMO and LUMO+1 of pyrrole with a core hole on the nitrogen atom. These are dramatically changed upon solvation. It appears in fact, that the bare LUMO + 1 corresponds to the solvated LUMO, and the solvated LUMO + 1 corresponds to the bare LUMO. The LUMO + 1 shows a particularly large amount of mixing with the waters. This lowers the intensity of this transition and therefore the intensity of the first feature relative to that of the isolated molecule. The state exists on the nitrogen perpendicular to the nitrogen-hydrogen bond, as it must remain orthogonal to other states. In contrast, for the isolated molecule,, the LUMO extends out from the N-H into the vacuum. Hence, solvation does not alter the states enough to cause new spectral features to appear, but it does cause state reordering and the changing of spectral intensities.

## 4. Conclusions

The solvation of pyrrole by water was shown, perhaps surprisingly, to have only minor effects on the NEXAFS spectra, and hence on the electronic structure of this prototypical aromatic molecule. The nitrogen K-edge spectrum of pyrrole exhibits only slight changes with solvation in water when compared to the gas phase spectrum. There is a noticeable red-shifting of several features, as predicted by theoretical calculations, and attributed to interactions between the N-H group and the neighboring water molecules. The spectral features themselves remain largely intact, as found previously for the solid.

The carbon K-edge spectrum is complicated, resulting from two inequivalent carbon atoms and also being obscured by the oligomerization of pyrrole. The calculated gas phase spectrum generally has the correct form, although it underestimates the width of the prominent first feature. A shoulder on this feature is attributed to the oligomerization of pyrrole, which our calculations (as well as previous calculations) have been unable to reproduce, when considering only monomeric pyrrole. We investigate several candidate chemical species that contribute to this spectral feature and conclude that the fraction of carbons that have reacted is significant.

Generally excellent agreeement between our observed spectra and those calculated with the XCH method was found. This is an encouraging step in the search for a truly predictive theory of core-level spectra of complex molecules in solution.

## Figure Captions

**Figure 1.** Experimental carbon K-edge NEXAFS (green) and ISEELS (red dotted) spectra of gaseous pyrrole with features labeled and a sketch of the molecular structure. The carbons are labeled 1 and 2 to differentiate between the calculated carbon spectra. The calculated carbon 1 fixed nuclei spectrum is shown (blue dots) with the classical MD overlaid (red with shaded error bars) and for carbon 2 the fixed nuclei spectra is shown (purple dots) with the classical MD overlaid (orange with error bars). The experimental ISEELS are from Ref. 19 and the calculated spectra are aligned to the experimental ionization potential of carbon 2 which is shown, as well as that of carbon 1. The calculated spectra are summed to give the total combined carbon K-edge spectra shown as summed fixed nuclei (orange dotted) and the summed MD spectra (purple). The arrows indicate corresponding features between theory and experiment.

**Figure 2.** Calculated carbon K-edge spectra of possible aggregation products are shown. The calculated nonbonded dimer spectrum is shown in blue. The structure of the chemically bonded dimer is shown at the top, and its calculated spectrum is shown in both the flat conformation (orange) and when twisted about the carbon-carbon bond formed upon dimerizing (red). The shoulder grows more intense for the trimer (green) and polypyrrole (yellow). The carbons labeled $\alpha$ give rise to the large subfeature also labeled $\alpha$.

**Figure 3.** Experimental carbon K-edge NEXAFS (light green) and ISEELS (red dotted – from ref. 19) spectra of gaseous pyrrole. A 1:1 mixture of the flat pyrrole dimer and the pyrrole monomer (black) captures all the low energy features observed experimentally.

**Figure 4.** Experimental carbon K-edge NEXAFS of solvated pyrrole at ~295K (light green) and ~275 K (dark green) with features labeled and a drawing of the molecule. The carbons are labeled 1 and 2 to differentiate between the calculated carbon edges. The calculated carbon spectra are obtained from classical MD at 300K. The carbon 1 spectra are shown (blue with error bars) as are the carbon 2 spectra (red with error bars). The spectra are summed to give the total combined carbon K-edge spectrum (purple).

**Figure 5.** Experimental nitrogen K-edge NEXAFS (green) and ISEELS (red dotted) spectra of gaseous pyrrole and calculated fixed nuclei (purple dashed) and classical MD (blue with error bars) spectra of pyrrole with features labeled and the molecule studied pictured. The average spectrum for classical MD are shown in darker colors, and one standard deviation is shown as a lightly shaded line for the classical MD spectra. The experimental ionization potential is shown as a vertical line. The experimental ISEELS is from ref 19.

**Figure 6.** Experimental nitrogen K-edge NEXAFS spectra of pyrrole solvated in water at ~295 K (orange) and ~275 K (yellow) and a calculated spectrum at 300K of the solvated molecule from classical MD (green with error bars) with features labeled. The average spectrum for classical MD are shown in darker colors, with one standard deviation lightly shaded.

**Figure 7.** Isosurfaces are shown for the LUMO and LUMO+1 of the excited state of pyrrole with and without water, corresponding to 30% of the total integrated value. The nitrogen atom is

87

at the top of the pyrrole molecule in each of these images.  These are the transitions that cause the intense first feature of the nitrogen K-edge of pyrrole.  Part of the reason this feature decreases in intensity upon solvation us due to the mixing of the solvated LUMO+1 with the surrounding water molecules. Such mixing is not as strong in the case of the solvated LUMO.
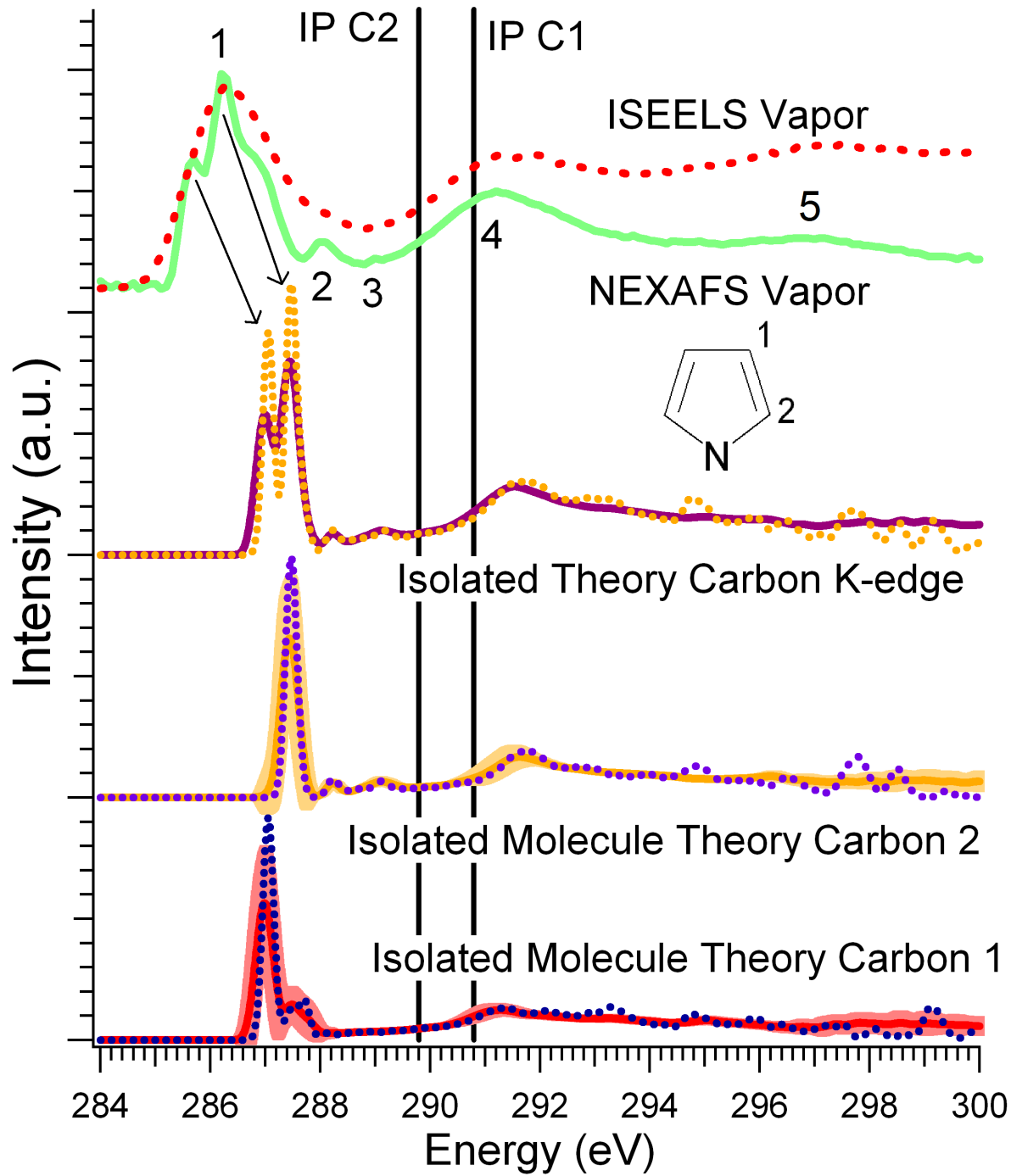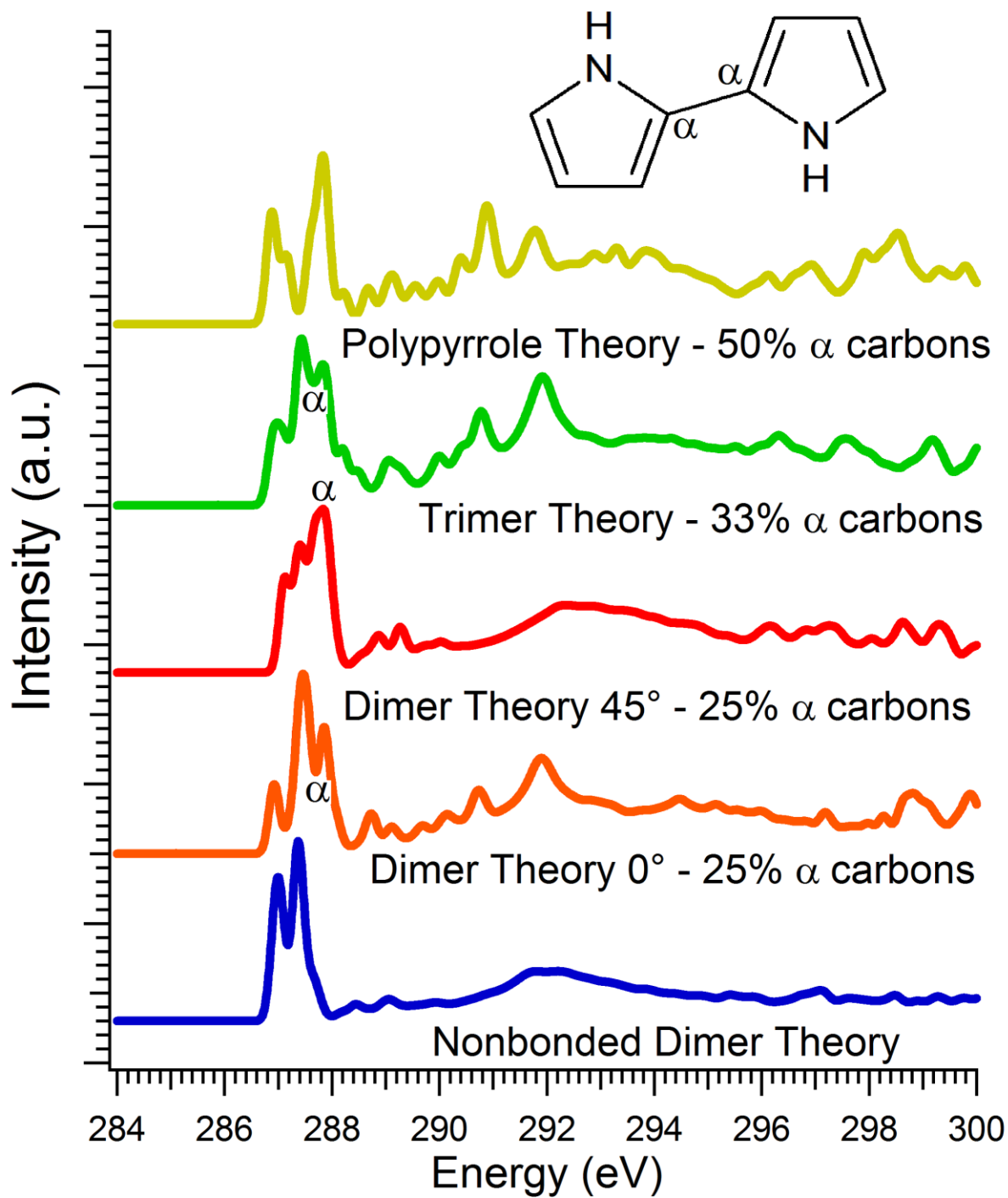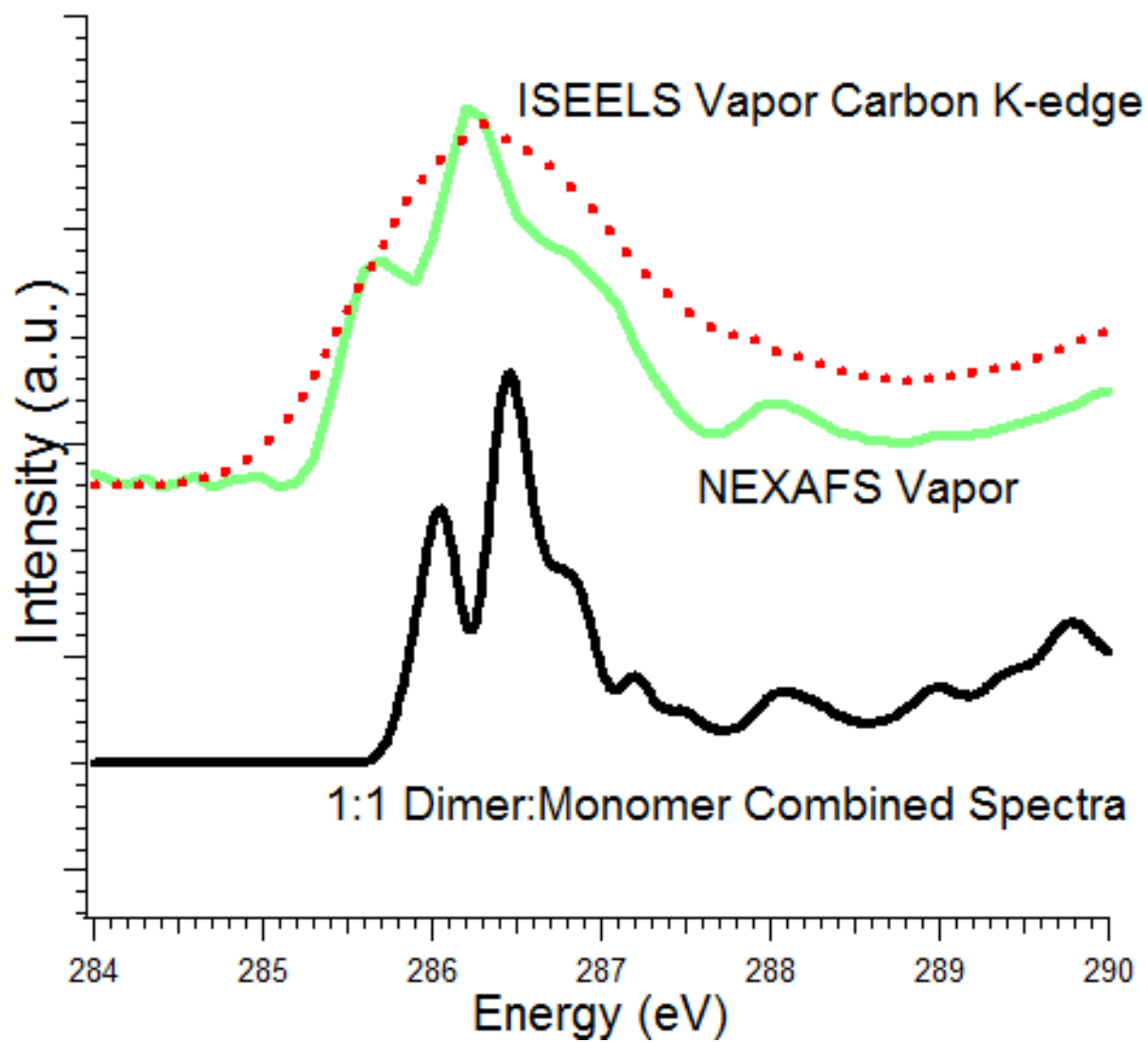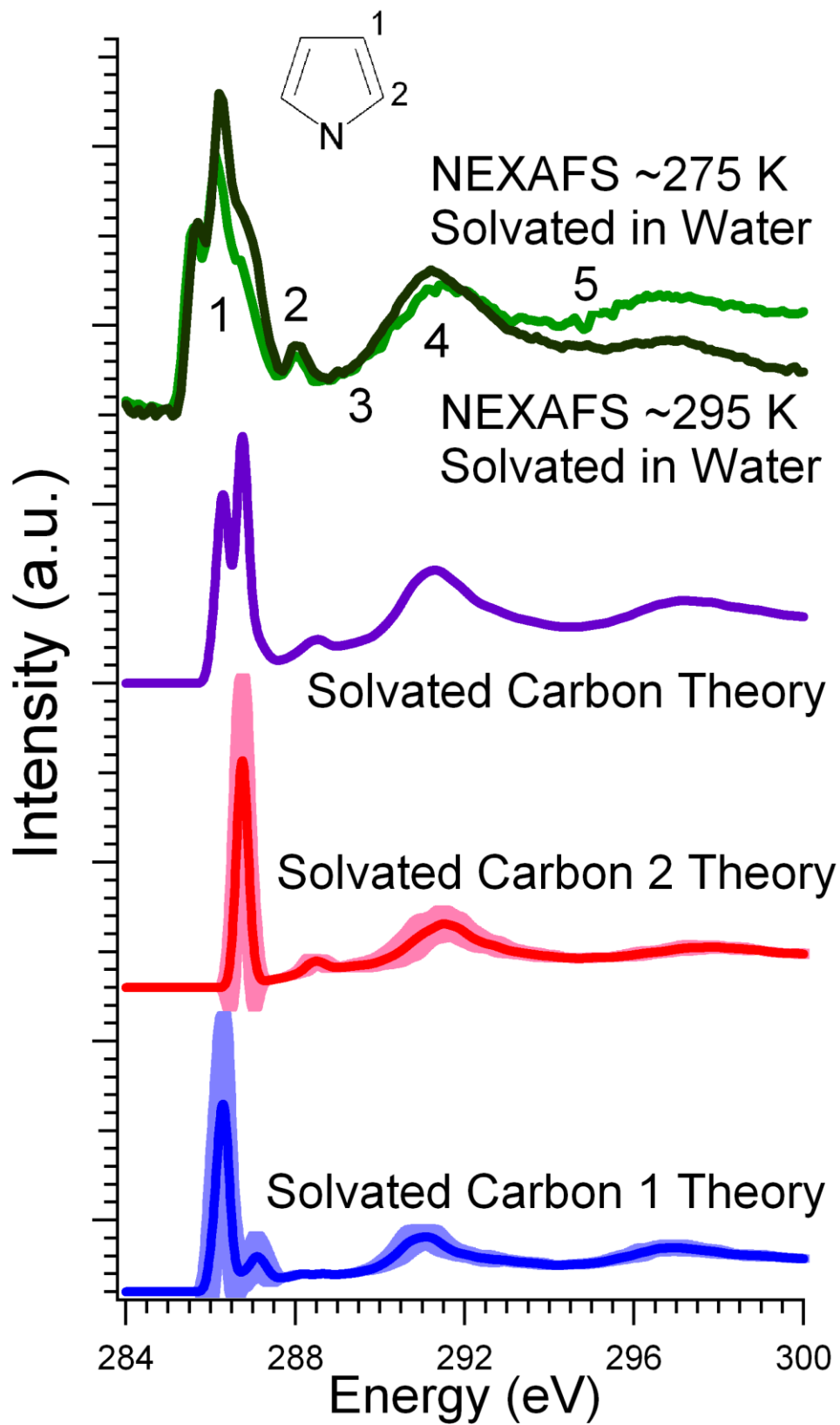
**Figures**



**Figure 1**

Figure 2

**Figure 3**

**Figure 4**

**Figure 5**

**Figure 6**

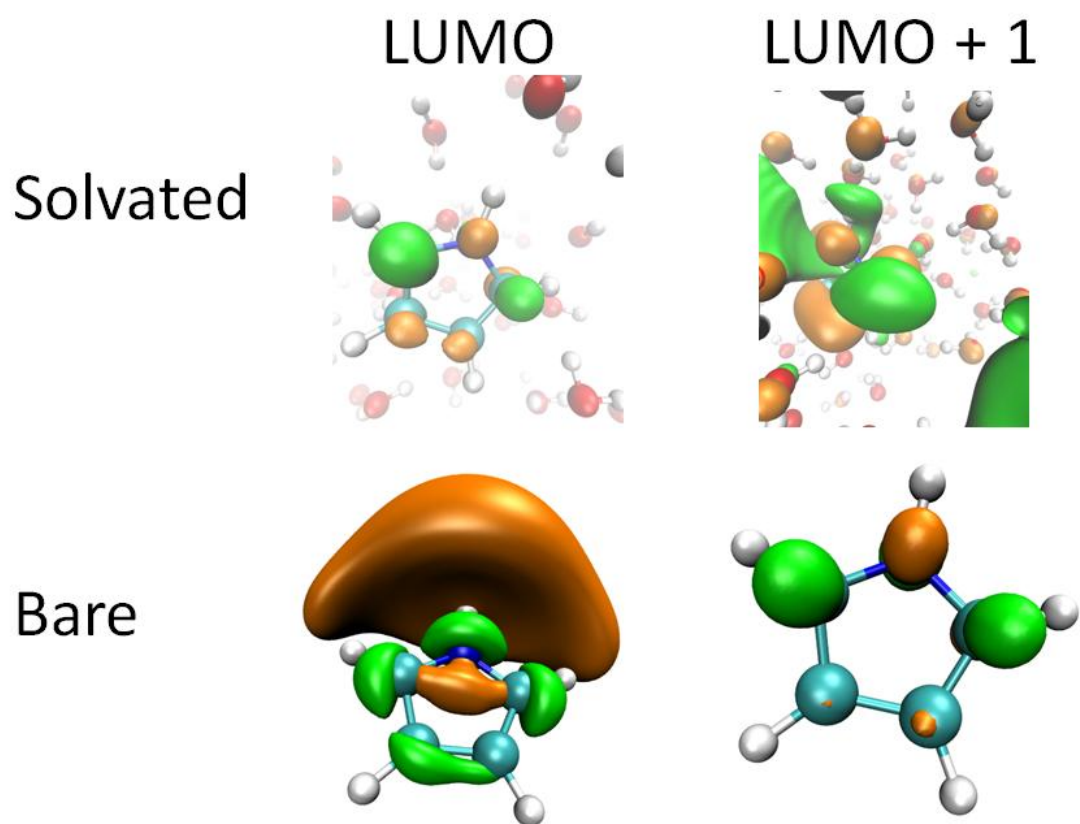LUMO    LUMO + 1

Solvated

Bare
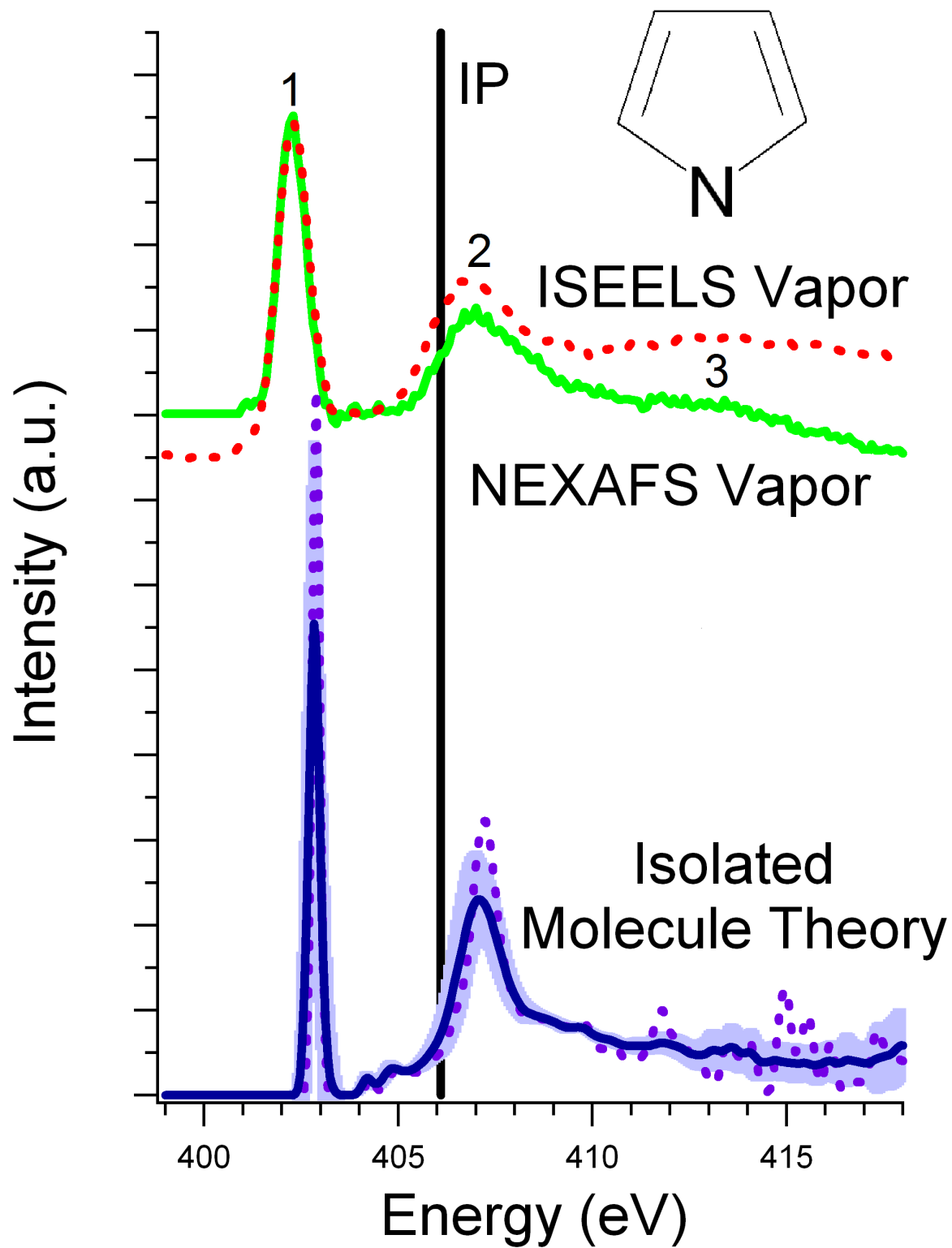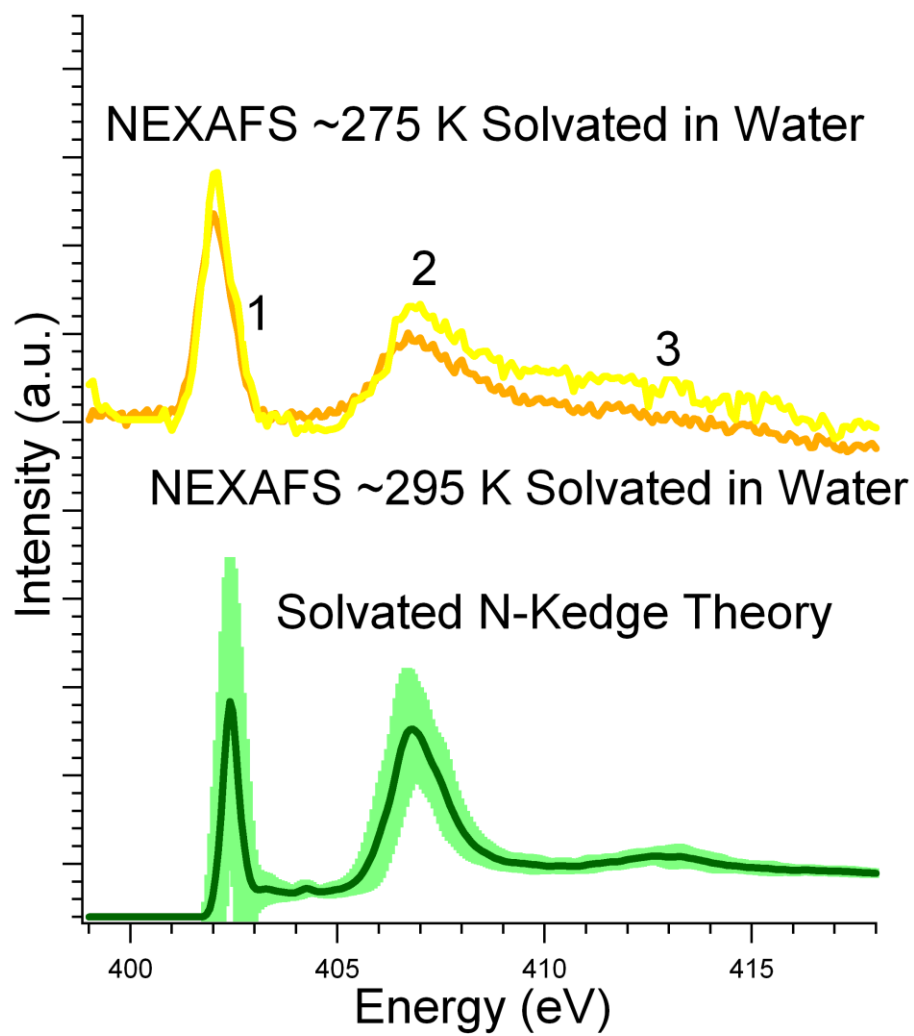
**Figure 7**

1    D. R. Lide, *CRC Handbook of Chemistry and Physics, 88th Edition*. (CRC, 2007).

2    R. H. Linnell and S. Umar, Archives of Biochemistry and Biophysics **57** (1), 264 (1955); A. Stolarz and J. Szydlowski, Journal of Radioanalytical and Nuclear Chemistry-Articles **185** (2), 219 (1994).

3    B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry B **109** (11), 5375 (2005).

4    E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, and B. Winter, Nature **455** (7209), 89 (2008); J. S. Uejio, C. P. Schwartz, A. M. Duffin, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, Proceedings of the National Academy of Sciences of the United States of America **105** (19), 6809 (2008).

5    K. R. Wilson, B. S. Rude, T. Catalano, R. D. Schaller, J. G. Tobin, D. T. Co, and R. J. Saykally, Journal of Physical Chemistry B **105** (17), 3346 (2001).

6    P. Parent, C. Laffon, C. Mangeney, F. Bournel, and M. Tronc, Journal of Chemical Physics **117** (23), 10842 (2002).

7    G. Vall-llosera, B. Gao, A. Kivimaeki, M. Coreno, J. A. Ruiz, M. de Simone, H. Agren, and E. Rachlew, Journal of Chemical Physics **128** (4) (2008); Y. Zubavichus, M. Zharnikov, A. Schaporenko, and M. Grunze, Journal of Electron Spectroscopy and Related Phenomena **134** (1), 25 (2004).

8    M. J. Tubergen, A. M. Andrews, and R. L. Kuczkowski, Journal of Physical Chemistry **97** (29), 7451 (1993); M. A. Martoprawiro and G. B. Bacskay, Molecular Physics **85** (3), 573 (1995).

9    A. L. Sobolewski and W. Domcke, Chemical Physics Letters **321** (5-6), 479 (2000).

10   P. Tuomikoski, Journal of Chemical Physics **20** (6), 1054 (1952).

11   D. Chandler, Nature **437** (7059), 640 (2005); L. R. Pratt, Annual Review of Physical Chemistry **53**, 409 (2002).

12   H. S. Ashbaugh and L. R. Pratt, Reviews of Modern Physics **78** (1), 159 (2006).

13   E. Otero and S. G. Urquhart, Journal of Physical Chemistry A **110** (44), 12121 (2006).

14   D. C. Newbury, I. Ishii, and A. P. Hitchcock, Canadian Journal of Chemistry-Revue Canadienne De Chimie **64** (6), 1145 (1986).

15   M. Mauerer, P. Zebisch, M. Weinelt, and H. P. Steinruck, Journal of Chemical Physics **99** (5), 3343 (1993).

16   C. Hennig, K. H. Hallmeier, A. Bach, S. Bender, R. Franke, J. Hormes, and R. Szargan, Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy **52** (9), 1079 (1996).

17   D. Duflot, C. Hannay, J. P. Flament, and M. J. Hubin-Franskin, Journal of Chemical Physics **109** (13), 5308 (1998).

18   I. Dauster, C. A. Rice, P. Zielke, and M. A. Suhm, Physical Chemistry Chemical Physics **10** (19), 2827 (2008).

19   M. F. Suarez-Herrera and J. M. Feliu, Physical Chemistry Chemical Physics **10** (46), 7022 (2008).

20   Y. C. Liu, S. J. Yang, T. C. Chuang, and C. C. Wang, Journal of Electroanalytical Chemistry **570** (1), 1 (2004).

21   F. Beck, M. Oberst, and R. Jansen, Electrochimica Acta **35** (11-12), 1841 (1990).

22   M. Aoki, A. Ishihara, and T. Kabe, Nippon Kagaku Kaishi (1), 57 (1991).

23   K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, and R. J. Saykally, Review of Scientific Instruments **75** (3), 725 (2004).

24   J. D. Smith, C. D. Cappa, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, Journal of the American Chemical Society **128** (39), 12892 (2006).

25   P. Hohenberg and W. Kohn, Physical Review B **136** (3B), B864 (1964); W. Kohn and L. J. Sham, Physical Review **140** (4A), 1133 (1965); J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Physical Review B **46** (11), 6671 (1992).

26   C. Kolczewski, R. Puttner, O. Plashkevych, H. Agren, V. Staemmler, M. Martins, G. Snell, A. S. Schlachter, M. Sant'Anna, G. Kaindl, and L. G. M. Pettersson, Journal of Chemical Physics **115** (14), 6426 (2001).

27   L. G. M. Pettersson, T. Hatsui, and N. Kosugi, Chemical Physics Letters **311** (3-4), 299 (1999); J. J. Rehr, Radiation Physics and Chemistry **75** (11), 1547 (2006); A. Nilsson and L. G. M. Pettersson, Surface Science Reports **55** (2-5), 49 (2004).

28   U. Ekstrom and P. Norman, Physical Review A **74** (4) (2006).

29   J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).

30   D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).

31   D. Prendergast, J. C. Grossman, and G. Galli, Journal of Chemical Physics **123** (1) (2005).

32   C. P. Schwartz, J. S. Uejio, R. J. Saykally, and D. Prendergast, The Journal of Chemical Physics **130** (18), 184109 (2009).

33   J. S. Uejio, C. P. Schwartz, A. M. Duffin, A. H. England, R. J. Saykally, and D. Prendergast, Journal of the American Chemical Society **Submitted** (2009).

34   D. K. Singh, S. K. Srivastava, A. K. Ojha, and B. P. Asthana, Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy **71** (3), 823 (2008).

35   S. Baroni, A. D. Corso, S. D. Gironcoli, and P. Giannozzi, PWSCF (2008).

36   J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77** (18), 3865 (1996).

37   G. Herzberg and E. Teller, Zeitschrift Fur Physikalische Chemie-Abteilung B-Chemie Der Elementarprozesse Aufbau Der Materie **21** (5/6), 410 (1933).

38   T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., K. M. M. Luo, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S., A. R. Hayik, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H., L. Y. Gohlke, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C., and W. S. R. Schafmeister, and P.A. Kollman,  (2006); J. M. Wang, W. Wang, P. A. Kollman, and D. A. Case, Journal of Molecular Graphics & Modelling **25** (2), 247 (2006).

39   L. A. Naslund, D. C. Edwards, P. Wernet, U. Bergmann, H. Ogasawara, L. G. M. Pettersson, S. Myneni, and A. Nilsson, Journal of Physical Chemistry A **109** (27), 5995 (2005).

40   D. Duflot, J. P. Flament, A. Giuliani, J. Heinesch, and M. J. Hubin-Franskin, Journal of Chemical Physics **119** (17), 8946 (2003).

41   G. Columberg and A. Bauder, Journal of Chemical Physics **106** (2), 504 (1997).

42   J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).

# Chapter 7 - Soft X-Ray Absorption Spectra of Aqueous Salt Solutions with Highly Charged Cations

## 1. Introduction

Highly charged cations can have a large restructuring effect on the hydrogen bond network of water.[1]   Here we study this effect via the perturbations on the near edge X-ray absorption fine structure (NEXAFS) spectrum of water.  NEXAFS is a spectroscopic technique in which the core electrons of selected atoms in a molecule are excited to high energy anti-bonding and continuum states.[2]  These unoccupied orbitals are highly sensitive to their local environments due to their large spatial extent, which make them excellent probes of changes in the local water structure.[3]

Until recently, it was prohibitive to apply the inherently high vacuum technique of soft X-ray spectroscopy to high vapor pressure liquids (such as water, octane, methanol etc.). Following the approach Faubel and coworkers' developed for UV photoelectron spectra of such volatile species,[4] in 2001 we incorporated liquid microjet technology into a synchrotron X-ray experiment, thereby opening many new systems to study by this powerful technology.[3] Dramatic changes were observed in NEXAFS spectra for strongly interacting molecules, e.g. methanol and water, as a function of state (gas, solid, liquid).[3,5]   Liquid water exhibits a spectrum with none of the well-resolved peaks that are prevalent in the gas phase spectrum and which is also quite different from that of the solid; these prominent spectral changes have been interpreted with the aid of theoretical calculations to help elucidate the liquid state of water and have engendered much recent controversy.[6-8]

Previous NEXAFS studies of aqueous salt solutions have shown that halide ions perturb the vacant electronic orbitals of water in a manner specifically dependent upon the identity of the halide.[9]  This was in contrast to monovalent cations, which exhibited relatively small effect on the oxygen K-edge spectra of water, independent of their identity.[10,11]  Interestingly, doubly charged cations produce spectral changes that were specific to the identity of the cation.  Highly charged cations have previously been investigated in the soft X-ray region, but only in the spectral region preceding the onset of the water pre-edge (energy $< 534$ eV), where an increase in intensity was observed for transitions of $Cr^{3+}$ and $Fe^{3+}$ but not $Al^{3+}$; this was attributed to d-shell mixing between metals and water,[12] although it was later argued that these features could instead be due to the formation of hydroxide ions, which also absorb in that energy range.[13]  We note that there have been many previous studies of aqueous salt solutions (reviewed previously)[14] and that structural and coordination information can be gathered by diffraction methods (X-ray and neutron) as well as nuclear magnetic resonance (NMR) and extended X-ray absorption fine structure (EXAFS) experiments in the hard X-ray region.[14]  Simulations are becoming increasingly reliable for determining coordination numbers for solvated ions.[15]  Raman spectrometry can determine the symmetry and stability of certain ion complexes.[16]   Mass spectroscopy studies of small ion-containing water clusters have been performed, but are not definitively related to the liquid because these small clusters lack counterions.[17]  Herein, we investigate the possible restructuring of the aqueous hydrogen bond network by highly charged cations via the effects observed on the oxygen K-edge NEXAFS spectra of water upon the addition of halide (chloride, except for one case) salts which contain highly charged cations.

## 2. Experimental

**Samples -** Samples purchased from Sigma-Aldrich (>97% purity, in certain cases the hydrate was purchased) and used without further purification were dissolved in 18 MΩ water. Concentrations of up to 1M were used for all salts. $CrCl_3$ was dissolved in water, and allowed to age for approximately 6 hours before measurements were taken. Under these conditions, the dominant species would be $[CrCl(H_2O)_5]^{2+}$ with smaller amounts of $[Cr(H_2O)_6]^{3+}$ and $[CrCl_2(H_2O)_4]^+$ also present.[18-20] This was previously determined by a combination of EXAFS, X-ray scattering and NMR with a high degree of certainty. In the case of $YBr_3$, yttrium will largely be in the +3 charge state coordinated by 8 waters, but there will also be significant direct ion pairing between yttrium and the counter ion, as determined recently by EXAFS, neutron scattering and Raman spectroscopy.[21-23] $InCl_3$ will predominantly be found in water as $[InCl_{2+}(H_2O)_5]^+$, as determined by EXAFS, Raman spectroscopy and X-ray scattering.[24] For $LaCl_3$, lanthanum will be in the +3 state at all concentrations studied here and solvated by 9 waters (although eightfold hydration has been invoked in the past).[25] Lanthanum salts have been studied extensively by techniques including EXAFS, X-ray diffraction and Raman spectroscopy. Based on NMR and Raman data, the dominant tin species at 1.0 molar will be $[SnCl_3(H_2O)_3]^{+3}$ and $Sn^{+4}$ but there will still be significant (>5%) amounts of $[SnCl_3(H_2O)_3]^+$ $[SnCl_4(H_2O)_2]$ $[SnCl_2(H_2O)_4]^{+2}$ $[SnCl_3(H_2O)_3]^{+3}$ $[SnCl_5(H_2O)_1]^{-1}$, $[SnCl_3(H_2O)_3]^+$ and $[SnCl_4(H_2O)_2]$ .[26]

**Absorption Measurements -** Total electron yield (TEY) X-ray absorption spectra were recorded at the oxygen K-edge (~550 eV) for each sample, as detailed previously,[3] using Beamline 8.0.1 at the Advanced Light Source (ALS) at Lawerence Berkeley National Laboratory in Berkeley, CA. A detailed description of the experimental system has been previously published.[27,28] Briefly, an intense (< $10^{11}$ photons/sec), high resolution (E/ΔE > 4000), and tunable X-ray beam is focused (~50 μm spot size) onto a small liquid jet (~30 μm diameter). The jet is produced by using a syringe pump (Teledyne-Isco) to pressurize the liquid behind a fused silica capillary tip. Almost immediately (~1mm and ~30 μs) after leaving the tip orifice, the liquid is intersected by the intense X-ray beam, wherein the sample is close to room temperature (> 15˚C).[29] The signal is normalized to that from a gold mesh located ~3 meters up beam of the chamber. The energy is calibrated to the water vapor NEXAFS spectrum. To allow for comparison between spectra, all spectra have been area normalized.

## 3. Results and Discussion

Area-normalized TEY spectra of water and 1M aqueous solutions are shown in Figure 1A for halide solutions containing several highly charged cations. When lower concentrations were investigated, similar effects were observed, but the changes from the pure water spectrum were less pronounced. The water spectrum is typically in terms of the pre-edge (~535 eV), the main-edge (~538 eV) and the post-edge (~542 eV).[6,7] The solution spectra show small but salt-specific shifts in spectral intensity compared to pure water, or compared to a comparable concentration of chloride from sodium chloride. This is in contrast to what was found when monovalent cation salts were used, wherein spectral changes were shown to be independent of the cation identity. Cation-specific changes were observed for divalent cations, however, and ascribed to charge exchange interactions.[10,11] Due to the small magnitude of the spectral change apparent in the present data, a theoretical analysis of sufficient accuracy (i.e. using the recently

developed XCH approach)[8] to be reliable was deemed not possible at this current time. This problem is especially acute due to the difficulty in accurately simulating both the dynamics of the system and describing the electronic structure of the excitation within experimental accuracy.[8,30,31] The combined error on both of these calculations coupled with the small magnitude of observed experimental changes makes accurately simulating the experimental spectra beyond the current scope of theoretical calculations.

Interestingly, and as previously found for $CrCl_3$,[12] dissolution of certain salts ($CrCl_3$ and $SnCl_4$) leads to the observation of relatively strong features at an energy below the water pre-edge (Figure 1B). The observation of peaks before the pre-edge was previously attributed to d-orbital mixing of the cation with the neighboring water molecules.[12] Later, it was noted that these features could actually be due to hydroxide and hydroxide-ion complexes, as hydroxide also absorbs strongly in this region.[13] The data presented here support the conclusion that it is likely due to formation of hydroxide, as other species which contain d-orbitals, viz. $InCl_3$, $LaCl_3$ and $YBr_3$, do not exhibit this feature in the present data set. Therefore, d-orbital mixing appears unlikely to be the cause of this spectral feature; hydroxide-containing species in $CrCl_3$ and $SnCl_4$ appears a more likely explanation. We note that both $SnCl_4$ and $CrCl_3$ are very reactive species.[18-20,26]

In the pre-edge region, all the salt solutions show a decrease in intensity relative to pure water. Previously, a decrease in pre-edge intensity has been interpreted as indicating the existence of fewer broken hydrogen bonds,[7] but this type of qualitative analysis can no longer be justified.[32] Furthermore, this analysis is not appropriate when the spectral features are due to several chemical species, such as hydroxide in addition to the water. Where the pre-edge evolves into the main-edge, the addition of salts leads to an increase in intensity relative to water at ~536 eV for all cases except $LaCl_3$. We note that some other species other than bulk water must cause this feature, so as the number of ion-paired species that are present in water increase, the likelihood there is some species that will strongly absorb at ~536 eV should increase. We note that $LaCl_3$ will be predominantly solvated without chloride in its first shell, minimizing the variety of ion pairs compared to the other salts considered in this study. Chromium, indium, yttrium and tin will form a large variety of complexes, as noted in the samples section, which have previously been characterized by techniques other than NEXAFS.[18-24,26]

The main-edge and post-edge regions exhibit salt-specific perturbations to the water spectrum. $InCl_3$ produces a small increase in intensity at the main-edge, with a post-edge similar to that of water, whereas $YBr_3$ produces an increase in intensity at the main-edge and decrease at the post-edge relative to water. These spectra are noticeably different from those of $LaCl_3$ solutions, which show a similar intensity to pure water at the main-edge, but greater intensity at the post-edge. Both $SnCl_4$ and $CrCl_3$ produce slight decreases in intensity at the main-edge and post-edge relative to that of pure water. This may be a function of normalization, as for these species there is increased intensity below the pre-edge in there must correspondingly be less intensity elsewhere in the spectrum. This region of the spectrum is especially difficult to interpret as it will have competing effects induced by both the cation and the anion which may to some extent offset, explaining the relatively small degree of spectral changes observed. It is known that anions can effect large perturbations,[9] and given the relatively high concentration of anions in this study (3M or greater) and the small degree of spectral change, it is likely that the perturbations to the water electronic structure by the highly charged cations are offsetting much of the spectral changes induced by the anions.[9] Overall, the spectral changes induced by these salts are surprisingly small.

## 4. Conclusions

X-ray absorption spectra of aqueous solutions of $InCl_3$, $YBr_3$, $LaCl_3$, $SnCl_4$ and $CrCl_3$ have been measured at the oxygen K-edge, using liquid microjet technology. The spectra of the aqueous solutions differ only slightly from that of pure water, although the observed differences do depend on the identity of the salt. Due to the relatively minor effects apparent in the main-edge and post-edge regions, as compared to similar solutions of concentrated alkali halide salts, it is probable that the spectral perturbations effected by a highly charged cation offset much of the spectral effect of the anion. More detailed analysis must await further developments in theoretical methods for calculating XAS spectra of ionic solutions.

**Figure Captions**

**Figure 1.** Oxygen NEXAFS K-edge spectra of water and several salt solutions with a concentration of 1M. Part A - section i shows the spectrum of bulk water (black, solid) and 1M solutions of $InCl_3$ (red, - - -), $YBr_3$ (blue, $\cdot \cdot \cdot$) and $LaCl_3$ (green, -- --); the difference spectra between the salt containing spectra and the spectra of bulk water are shown above the main spectra. Section ii shows the spectrum of bulk water (black, solid) and 1M solutions of $SnCl_4$ (purple, - $\cdot$ - $\cdot$) and $CrCl_3$ (yellow, $\cdot \cdot \cdot$); the difference spectra between the salt containing spectra and the spectra of bulk water are shown above the main spectra. The approximate locations of the pre-edge, main-edge and post-edge are labeled. Part B - an expansion of Figure 1 section ii in the pre-edge region. The spectrum of bulk water (black, solid) and 1M solutions of $SnCl_4$ (purple, - $\cdot$ - $\cdot$) and $CrCl_3$ (yellow, $\cdot \cdot \cdot$).

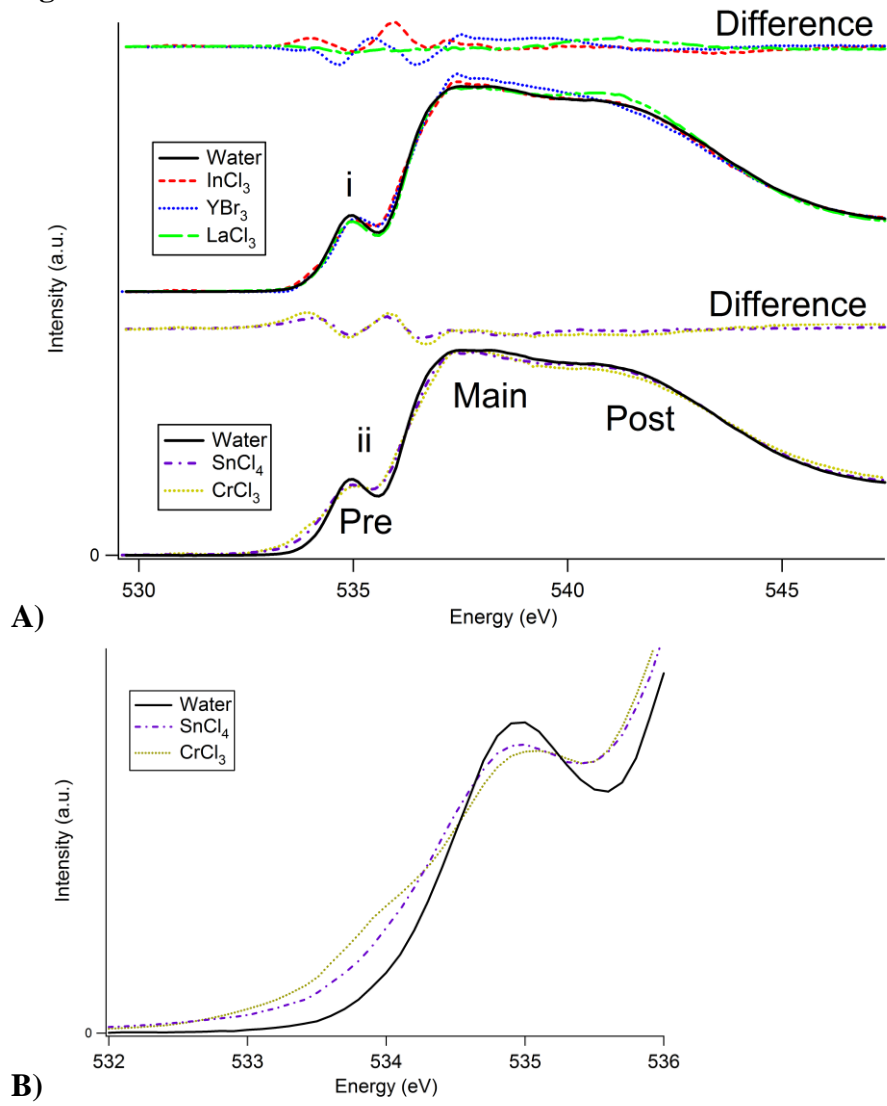**Figures**



**A)**



**B)**

**Figure 1**

1    U. Cosentino, A. Villa, D. Pitea, G. Moro, and V. Barone, Journal of Physical Chemistry B **104** (33), 8001 (2000).

2    J. Stohr, *NEXAFS Spectroscopy*. (Springer-Verlag, Berlin, 1992).

3    K. R. Wilson, B. S. Rude, T. Catalano, R. D. Schaller, J. G. Tobin, D. T. Co, and R. J. Saykally, Journal of Physical Chemistry B **105** (17), 3346 (2001).

4    M. Faubel, B. Steiner, and J. P. Toennies, Journal of Chemical Physics **106** (22), 9013 (1997).

5    K. R. Wilson, M. Cavalleri, B. S. Rude, R. D. Schaller, T. Catalano, A. Nilsson, R. J. Saykally, and L. G. M. Pettersson, Journal of Physical Chemistry B **109** (20), 10194 (2005).

6    J. D. Smith, C. D. Cappa, K. R. Wilson, B. M. Messer, R. C. Cohen, and R. J. Saykally, Science **306** (5697), 851 (2004).

7    P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. A. Naslund, T. K. Hirsch, L. Ojamae, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, Science **304** (5673), 995 (2004).

8    D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).

9    C. D. Cappa, J. D. Smith, K. R. Wilson, B. M. Messer, M. K. Gilles, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry B **109** (15), 7046 (2005).

10   C. D. Cappa, J. D. Smith, B. M. Messer, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry B **110** (11), 5301 (2006).

11   L. A. Naslund, D. C. Edwards, P. Wernet, U. Bergmann, H. Ogasawara, L. G. M. Pettersson, S. Myneni, and A. Nilsson, Journal of Physical Chemistry A **109** (27), 5995 (2005).

12   L. A. Naslund, M. Cavalleri, H. Ogasawara, A. Nilsson, L. G. M. Pettersson, P. Wernet, D. C. Edwards, M. Sandstrom, and S. Myneni, Journal of Physical Chemistry A **107** (35), 6869 (2003).

13   C. D. Cappa, J. D. Smith, B. M. Messer, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry A **111** (2007).

14   E. V. Vinogradov, P. R. Smirnov, and V. N. Trostin, Russian Chemical Bulletin **52** (6), 1253 (2003).

15   V. Buzko, I. Sukhno, and M. Buzko, International Journal of Quantum Chemistry **107** (13), 2353 (2007).

16   M. I. Cabaco, M. Marques, A. M. Gaspar, M. A. Marques, and M. M. Costa, Journal of Molecular Liquids **136** (3), 323 (2007).

17   M. F. Bush, R. J. Saykally, and E. R. Williams, Journal of the American Chemical Society **130** (28), 9122 (2008).

18   S. DiazMoreno, A. MunozPaez, J. M. Martinez, R. R. Pappalardo, and E. S. Marcos, Journal of the American Chemical Society **118** (50), 12654 (1996).

19   M. Magini, Journal of Chemical Physics **73** (5), 2499 (1980).

20   L. Helm and A. E. Merbach, Coordination Chemistry Reviews **187**, 151 (1999).

21   M. I. D. Marques, M. A. Marques, and J. R. Rodrigues, Journal of Physics-Condensed Matter **4** (38), 7679 (1992).

22   S. Ramos, G. W. Neilson, A. C. Barnes, and A. Mazuelas, Journal of Physical Chemistry B **105** (14), 2694 (2001).

23   D. T. Bowron and S. Diaz-Moreno, Journal of Physical Chemistry B **111** (39), 11393 (2007).

24   M. I. Cabaco, A. M. Gaspar, C. M. de Morais, and M. A. Marques, Journal of Physics-Condensed Matter **12** (12), 2623 (2000).

25   M. A. Marques, M. I. Cabaco, M. I. D. Marques, A. M. Gaspar, and C. M. de Morais, Journal of Physics-Condensed Matter **13** (20), 4367 (2001).

26   M. J. Taylor and J. M. Coddington, Polyhedron **11** (12), 1531 (1992).

27   K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, and R. J. Saykally, Review of Scientific Instruments **75** (3), 725 (2004).

28   J. D. Smith, C. D. Cappa, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, Journal of the American Chemical Society **128** (39), 12892 (2006).

29   B. M. Messer, C. D. Cappa, J. D. Smith, K. R. Wilson, M. K. Gilles, R. C. Cohen, and R. J. Saykally, Journal of Physical Chemistry B **109** (11), 5375 (2005).

30   M. Odelius, M. Cavalleri, A. Nilsson, and L. G. M. Pettersson, Physical Review B **73** (2) (2006).

31   C. Clavaguera, R. Pollet, J. M. Soudan, V. Brenner, and J. P. Dognon, Journal of Physical Chemistry B **109** (16), 7614 (2005).

32   J. S. Tse, D. M. Shaw, D. D. Klug, S. Patchkovskii, G. Vanko, G. Monaco, and M. Krisch, Physical Review Letters **100** (9) (2008).

# Chapter 8 - Probing Protein Interactions with NEXAFS: Sensitivity of Polypeptide Spectra to Conformation and Added Salts

## 1. Introduction

Over 100 years ago  Hofmeister discovered that adding salts to egg white protein could alter its solubility.[1]  Certain salts would cause the protein to become more soluble, while others had the opposite effect.  This ordering of salt-protein interactions has since become known as the "Hofmeister series,"[2] and the associated "Hofmeister effects" now extend to many different phenomena, including protein denaturing,  optical rotation of sugars and bacterial growth rates.[3]  It is known that anions have a stronger  effect on protein solubility than do cations and the general anion ordering is:

$$CO_3^{2-} > SO_4^{2-} > SO_3^{2-} > Cl^- > Br^- > I^- > SCN^-$$
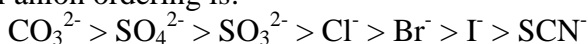
The species on the left are generally referred to as "kosmotropes", or "structure makers" and the species on the right are generally referred to as "chaotropes", or "structure breakers."[3,4]  Originally, these terms derived from the entropies of hydration and referred to making and breaking the water structure, which was believed to be the cause of the Hofmeister effects.  However,  this model of ascribing Hofmeister effects to bulk water changes has come under criticism, given newer experimental findings.[3]  It is now believed that interactions either directly between a protein and the ions or interactions mediated by a single solvation shell underlie the effects.[3,5]  Some groups have recently presented measurements to support the original contention that Hofmeister effects resulted from  changes in the bulk water structure, but their models treated water as a two state system, which has come under intense recent criticism.[6,7]

It has also been proposed that absolute free energies of hydration also underlie Hofmeister effects,[8,9] contending that if absolute hydration free energies of anionic and cationic groups match, than ion pairing will occur,[10] signifying strong ion-protein interactions .[11]  Cremer and coworkers have performed a series of studies wherein protein analogues were combined with various amounts of salt, causing the lower critical solution temperature (LCST) – the critical temperature below which the mixture is miscible in all portions – to change as a function of salt concentration and identity.[3,12-14]  They observed that kosmotropes lower the LCST more than do chaotropes.  Based on fitting of the curves combined with thermodynamic arguments, they argued that the way the salts affect the LCST is dependent upon the Hofmeister activity of the anion.  Chaotropes (such as Br⁻) were believed to directly interact with the protein analogue near the nitrogen sites, as well as to interact with the exposed hydrophobic surfaces.  Kosmotropes, on the other hand, were believed to polarize the water interacting with the nitrogen- and oxygen-containing groups as well as to interact with hydrophobic groups.  However, these conclusions were based on light scattering measurements and thermodynamic data, so the interactions had to bet be inferred, rather than being observed directly.

Here, we investigate triglycine as a model peptide via near edge x-ray absorption fine structure (NEXAFS) experiments and theoretical methods, seeking to further characterize the interactions of proteins with  selected salts.  Nitrogen K-edge (excitation from the N 1s orbital) spectra of triglycine in aqueous solution were measured, without cosolutes, with cosmotropic co-solutes $Na_2SO_3$ and $K_2SO_4$   and with co-solutes NaBr, NaCl, and NaF,  seeking to quantify observed differences in the spectra of triglycine due to the presence of the salts.  NEXAFS

probes unoccupied anti-bonding and Rydberg states, which are highly sensitive to intermolecular interactions.[15-17] This work follows similar recent studies using X-ray spectroscopy to characterize interactions between ions and small biomolecules.[17,18]

There are now several methods available for obtaining core-level spectra of liquids; we use small diameter jets of water solutions (~30 microns) with windowless coupling to a synchrotron beamline.. This approach is particularly useful for studying complex molecules, since the high velocity of the liquid jet minimizes sample damage.[19] The chemical information that can be extracted from such measurements is limited in many respects by the accuracy of theoretical methods for computing core-level spectra. Here, we use the recently-developed excited core hole (XCH) method, which accurately calculates NEXAFS spectra.[20,21] We simulated the atomic configurations of triglycine with and without additional salts in aqueous solution using classical molecular dynamics. Sampling a trajectory in our DFT simulation captures the impact of nuclear motion on the calculated spectrum.[22] This has been used previously to interpret the spectra of solvated amino acids and the prototypical aromatic molecule, pyrrole.[23,24] We also used rigid structural models of polyalanine in both an anti-parallel β-sheet and an α-helix from the literature[25] to calculate the respective N-edge X-ray absorption spectra.

There have been many studies of solid proteins and amino acids by NEXAFS; every naturally-occurring individual amino acid has had its core-level spectra characterized at the C, N and O K-edges[26], and a variety of polypeptides and protein spectra have been measured under similar conditions.[27-29] These data have been used to predict the corresponding spectrum of a protein based on its amino acid composition with reasonable success.[29] Based on such studies, it has been suggested that the difference in $\pi^*$-resonances between a model dipeptide and a protein could be due to differences between an α-helix and a planar structure.[27,28] In this work, we show preliminary evidence that the N K-edge spectrum will be relatively unchanged in these conformations.

## 2. Methods

**2.1 Samples:** Gly-Gly-Gly (Triglycine), NaBr, Na$_2$SO$_3$, NaCl, NaF, and K$_2$SO$_4$, were obtained commercially from Sigma-Aldrich, with stated purities of at least 98%. All water used had a resistivity of 18 MΩ/cm. Samples were used without further purification. The aqueous samples used in the liquid jets contained concentrations of triglycine that were 33.3 (~0.13M) grams per liter of water. One sample comprised solely triglycine, another contained triglycine and 33.3g/L (~0.33 M) of NaBr, the third contained triglycine and 33.3 g/L of Na$_2$SO$_3$ (~0.26 M), a fourth contained triglycine and 33.3 g/L of NaCl (~0.57 M), a fifth contained triglycine and 33.3 g/L NaF (~0.81 M) and the sixth contained triglycine and 50.0 g/L K$_2$SO$_4$ (~0.29 M). These correspond to approximately a 2.5:1 NaBr:triglycine molar ratio, a 2:1 Na$_2$SO$_3$:triglycine molar ratio, a 4:1 NaF:triglycine molar ratio, a 6:1 NaF:triglycine molar ratio and a 2:1 K2SO4:triglycine molar ratio. We note that these concentrations are well below those wherein precipitation would occur in the liquid jets, as higher concentrations would lead to jet clogging.

**2.2 NEXAFS Spectroscopy of Solvated Triglycine:** Total electron yield (TEY) NEXAFS were recorded at the nitrogen K-edge (~400 eV). These measurements were performed at Beamline 8.0.1 of the Advanced Light Source (ALS) at Lawerence Berkeley National Laboratory in Berkeley, CA. A detailed description of the experimental system has been published previously.[19] Briefly, an intense (> 10$^{11}$ photons/sec), high resolution (E/ΔE >

4000) tunable X-ray beam is generated from an undulator. The synchrotron light is then focused (~50 μm spot size) onto a small liquid jet (~30 μm diameter) and the TEY is collected to obtain spectra of the bulk liquid.[16] The jet is produced by using a syringe pump (Teledyne-Isco) to pressurize the liquid behind a fused silica capillary tip, and travels parallel to the polarization of the incident radiation. Almost immediately (<0.5 mm and <15 μs) after leaving the tip orifice, the liquid is intersected by the X-ray beam, wherein the sample is close to room temperature (~ 23˚C).[23] The liquid jet is then condensed in a separate cryogenically cooled section of the chamber to maintain low pressures (~$10^{-4}$ torr) in the interaction region. The signal is normalized to that from a gold mesh located up-beam of the chamber and the energy is calibrated based on specific metal impurities located in the beamline optics. The spectra are area-normalized and smoothed.

### 2.3 Calculations

**Core-level spectra:** Our theoretical approach has been described previously.[20,22] We use periodic boundary conditions and a plane-wave basis. PWSCF is used to calculate the zone-center electronic structure.[30] We use the Perdew-Burke-Erzenhoff (PBE) form of the generalized gradient approximation to the exchange-correlation potential within density functional theory.[31-34] Core-hole matrix elements with valence electrons were calculated by reconstructing the core region of the excited atom of the pseudo states within the atomic frozen core approximation. Transition amplitudes are estimated in the single-particle and dipole approximations. The lowest energy core-hole excited state is modeled by explicit inclusion of the core-hole and excited electron (XCH). Excitations to states above this first excited state are approximated using the unoccupied Kohn-Sham eigenstates computed from the XCH self-consistent potential.[20,22] The spectra were aligned to experiment. These calculations were performed on the Franklin supercomputer at NERSC. All spectra were stretched along the energy axis by 25% to improve agreement with measurements. This is a typical underestimation in DFT.[23,24]

**Broadening:** All calculated spectra are numerically broadened using Gaussians of 0.2 eV full width at half maximum. Others often resort to ad-hoc energy dependent broadening schemes.[35] We use this relatively small and uniform broadening with the aim of simulating and distinguishing electronic and vibrational effects explicitly, thereby arriving at a predictive computational approach.

**Solvated Triglycine MD Simulations:** We have modeled the nuclear degrees of freedom in these systems using molecular dynamics (MD) performed at 300 K using a Langevin thermostat, with the generalized AMBER force field.[36] The resulting distribution of nuclear coordinates was sampled at 40 picosecond intervals to minimize correlation between snapshots. This was done for 100 configurations on each nitrogen (therefore, three calculations per configuration).[36,37] These coordinates are subsequently used in DFT calculations which were by far the most significant part of the computational cost. The systems sizes were approximately 4000 $Å^3$ and contained approximately 120 water molecules. In one case the triglycine was simulated with water, in another case 6 Na and Br ions were added to the triglycine water simulation, and in the final case 4 Na ions, and 2 $SO_3^{2-}$ ions were added to the triglycine water simulation. We note that this is higher than the experimental concentration, but this was done to emphasize any effects which may exist. We explicitly simulated spectra for only sulfite and bromide salt addition because the other salts studied here (sulfate and other halides, respectively) produced spectra that were essentially identical to these, and because of the significant computational demands of these calculations. For statistical purposes, these MD simulations

were continued for a total of 20 nanoseconds and sampled every 50 femtoseconds. These simulations were used for the radial distribution functions (RDF) and the root mean square (RMS) positions shown later. The RMS positions of the atoms were determined by RMS fitting the backbone of the triglycine structure to the average structure, with the average structure first RMS fitted to the first coordinate set to remove rotational effects. The results were similar if rotational effects were not removed initially.

**Spectrum of polyalanine:** The coordinates of polyalanine in both an ideal α–helix and an ideal anti-parallel β–sheet were taken from the literature.[25] Because all of the nitrogen atoms are effectively degenerate and motions were not treated, only a single NEXAFS spectrum was calculated for both the α–helix and anti-parallel β–sheet at the nitrogen K-edge.

## 3. Results and Discussion

**Molecular Structures -** The structure of triglycine (A) is shown in Figure 1, along with that of the sulfite anion (B) and polyalanine (C), along with the structure used in calculations of polyalanine in an α-helix (D) and β-sheet (E). At neutral pH, triglycine will exist primarily in the zwitterionic structure shown. The nitrogen atoms are labeled 1-3; this scheme will be used throughout this work to reference specific nitrogen atoms within triglycine, with N1 referring to the nitrogen closest to the carboxylic group, N2 the middle nitrogen and N3 the terminal nitrogen. It should be noted that the leftmost two carboxylate oxygen atoms in triglycine are identical, as are all the oxygen atoms of the sulfite anion. In the case of polyalanine in both the anti-parallel β-sheet and the α-helix, every monomer of alanine is conformationally identical due to the coordinates used.

**Experimental Results and DFT Calculations -** The experimental and theoretical results for the nitrogen K-edge spectra of solvated triglycine are shown in Figure 2. The experimental spectrum consists of a large feature at approximately 401 eV and another large broad feature centered at approximately 406 eV, followed by a large decay with another feature possibly located at 412 eV. Our simulations provide excellent agreement, capturing the essential features of the experimental spectra. Furthermore, the character of transitions can easily be assigned by examining the spectra of the individual nitrogen atoms. The first peak is due entirely to N1 and N2. That N1 and N2 have similar spectra is not surprising, given that the local environment surrounding each amide nitrogen in the triglycine backbone should be quite similar. That N1 and N2 are similar is also expected based on the "building block" approach which is used throughout much of the NEXAFS literature.[28,38] N3 largely gives rise to the broad feature around 406 eV. It is not surprising that there are not many sharp features, as this is a large and quite flexible molecule with freedom to sample molecular phase space, and the spectral features shift considerably in both energy and intensity with peptide motions.

When sodium bromide is added to the solution of triglycine, the spectrum does not change significantly, as shown in Figure 3. The spectrum again consists of a strong feature at approximately 401 eV, followed by another strong feature centered at 406 eV and a possible weak feature at 412 eV. The experimental result again is closely modeled by the simulations. We note that the spectra of N1 and N2 at 401 eV remain similar despite the interactions with bromide. Sodium is generally repelled from the protonated nitrogens and therefore is not of much spectral importance, based on classical MD.[39] This spectrum is generally very similar to that of triglycine without sodium bromide, in that no new features are present. Further NEXAFS measurements show a similar result, viz. no new features or effects with all the other sodium

halide salts, shown in Figure 4. There are differences between the spectra but no large new features are found.

The addition of sodium sulfite to the solution of triglycine has a larger spectral effect than does sodium bromide. The spectrum shown in Figure 5 still exhibits the same features, centered at 401 eV, 406 eV, a peak at 412 eV that is quite weak if it is present and a new feature at approximately 403 eV. While some absorption intensity is contributed from all nitrogen atoms at 406 eV, the new feature is due primarily to the addition of sodium sulfite and its affect on the amino group N3. We note that a new feature was also found (at slightly different energy due to the different ion) at approximately 403 eV in our measurements (Figure 6) in the case of potassium sulfate. It would be interesting to investigate potassium sulfate by simulations method and see if the peak is also shifted theoretically, but these simulations have not yet been performed. Perhaps this feature shifts as a function of charge concentration, but currently this is a completely untested hypothesis.

**Molecular Dynamics and Spectroscopic Analysis** – In order to extract details of solvation structure, RDFs comparing the nitrogen atoms and ion proximity were calculated. These are shown for sodium bromide and triglycine with bromide – nitrogen distances in part A of Figure 7 and for sodium sulfite and triglycine with oxygen (sulfite) – nitrogen distances in part B of Figure 7. Neither for sodium bromide nor sodium sulfite does the sodium atom have a propensity for directly interacting with the triglycine nitrogen atoms, generally being repelled from the first solvation shell. In contrast, sodium strongly interacts with the sulfite, which, in turn, strongly interacts with the protonated nitrogen terminus, leading to a large density of sodium as a second nearest neighbor.

We will first discuss part A of Figure 7. The only nitrogen which is strongly associated with bromide ions is N3, the protonated terminus of the zwitterion. However, this interaction produces only a modest spectral consequence, comprising a small increase in intensity at approximately 403 eV and a small decrease in intensity at approximately 406 eV, as compared to trigylcine without the sodium bromide present.

The RDF of sodium sulfite (oxygen) with the nitrogen atoms of triglycine (nitrogen) is shown in Figure 7 B. It is important to note that the sulfite ion is almost an order of magnitude more likely to be located directly adjacent to hydrogen atoms bonded to any of the nitrogen atoms than is bromide, likely due to a higher charge. Furthermore, sodium sulfite often exhibits a high degree of ion-pairing, particularly for the sulfite near N3 and to a lesser extent that near N1. These large differences in local environments around the nitrogen atoms, in particular with regards to N1 and N2, lead to the overall spectral transition being slightly broader, as the 1s → LUMO transition is no longer energetically degenerate for N1 and N2. The N2 1s → LUMO transition shifts in energy as the local environment shifts around the N2 atom; this is particularly related to the position of the sulfite and the N-H bond length. We note that there are other factors which contribute, but a single cause cannot be isolated at the present time. The new peak at 403 eV appears to be caused primarily by the N3 group interacting with the sulfite ion. A large intensity at that energy is not found in all of the calculated spectra, but only for a fraction of the calculated snapshots. However, many snapshots have similar 1s → LUMO transition energies (i.e. the absolute energy of many transitions is approximately 403 eV). This implies a similar local environment for a variety of snapshots, indicating that sulfite induces a certain amount of local structure in or around triglycine. If the environment was shifting constantly, one would not expect this peak to appear.

While the RDFs characterize the distance between triglycine and ions, it provides no dynamical information. In order to determine the structure, the RMS fitted backbone of triglycine from the simulations were plotted and are shown in Figure 8, with the standard deviation in the RMS fitting corresponding to the spheroid size. It is important to note that for all of these structures, the backbones lie roughly in a plane, i.e. they are not twisted. However, the sodium sulfite solution exhibits a different conformation of triglycine than when sodium bromide is present. Furthermore, the RMS deviation increases when triglycine is simulated with either no salt present or with sodium bromide, as compared to sodium sulfite; this implies the existence of less conformational variations in the presence of sodium sulfite. This is in agreement with the traditional interpretation of Hofmeister effects, viz. that cosomtropic solvent molecules such as $SO_3^{2-}$ will make it harder for a protein to form a cavity (the protein becomes less conformationally flexible).[3] The chaotropic sodium bromide induces a very slight increase in the conformational flexibility of triglycine, as compared to the peptide without co-solute, in agreement with the Hofmeister series for solvent cavity formation.[3] While this overall protein conformational flexibility may be important for phenomena related to the Hofmeister series, NEXAFS appears to have limited sensitivity to this effect, based on our calculations. Conformational flexibility can still be important with regards to peak width, however.

The reason for this limited spectroscopic sensitivity can be elucidated by examining the states probed in the experiment. The LUMO+1 is shown in Figure 9. From this image, it can be seen that the state is fairly well localized on a given segment of the polypeptide chain. When N1 is excited, the excitation is not delocalized very far along the backbone, although it does reach the carbon of the carboxylic group. When N2 is excited, the case which would be most representative of the middle of a large protein, the state is almost entirely localized on N2 and atoms at most two bonds away. There is no significant off-chain mixing. The state is extremely localized. N3 mixes more than any other nitrogen, with the LUMO+1 delocalized all the way back up the chain of trigylcine to N2. It also mixes with neighboring water molecules, and for that reason, N3 is unlikely to be any more sensitive to its surroundings than are the other nitrogen molecules.

The applicability of our findings to the general elucidation of Hofmeister effects is only indirect, as the restriction of conformational flexibility due to the cosmotropic sulfite ion is largely due to the protonated (zwitterionic) nitrogen group interacting with the sodium sulfite, wherein the anion pairs with sodium ions, causing the carboxylic acid group to loop around and interact with the sodium. In other words, the present findings are probably specific to short oligopeptides; however, one could envision such salts affecting proteins when multiple charged groups are in close spatial proximity, as will often be the case in actual proteins. We note that it was recently predicted that iodide and bromide ions would not directly interact with the protein backbone, but rather would instead interact with hydrophobic regions;[39] we have here experimentally confirmed that bromine does not strongly interact with a hydrophilic polypeptide. The calculations in the previous work indicate that the destabilizing effect of bromide in the Hofmeister effect is due to attraction to hydrophobic regions of the protein.

Previous studies have addressed the thermodynamics of the Hofmeister series, but molecular level insights have been much harder to gain. Cremer and coworkers have identified systematic thermodynamic dependencies of the LCST based on specific anion identity.[3,13,14] Their conclusion − that the more kosmotropic anions would interact with the protein via mediating waters and that chaotropic anions would interact with a protein directly − is not supported by the results presented here. However, it is important to note that their measurements

were taken a salt concentrations that were significantly higher than those studied here, and it is possible that different effects occur as concentration approaches the solubility limit. Therefore, while the present results do not prove that the thermodynamic arguments used by Cremer and coworkers are incorrect, it should emphasize the perils of trying to infer molecular phenomena from measurements that are inherently macroscopic.

**Polyalanine** – Finally, because of the noted minimal effects of motion on triglycine spectra, we have investigated the case of "extreme" conformational differences in polyalanine. As shown in in Figure 10, we have compared the calculated nitrogen K-edge spectrum of polyalanine in an anti-parallel β-sheet and in an α-helix, (as shown in Figure 1), for the solid . In these frozen conformations, spectral differences will be magnified, as there will be no motions to blur spectral features – in essence, this represents a maximum possible conformational effect on the core-level spectrum.

Nevertheless, the spectra are quite similar between the anti-parallel β-sheet and the α-helix. Both feature a strong absorption band ~401 eV and peaks roughly corresponding to experimental scattering peaks at ~408 eV and ~413 eV as well as a strong transition at ~404 eV. The antiparallel β-sheet features an absorption at ~403 eV that is not present in the α-helix. This can be compared with several experimental spectra from the literature, polyisoleucine (a polypeptide), lysozyme (a protein), and 2,5-diketopiperazine (a simple cyclic glycine dimer).[27,28] The lack of absorption at ~404 eV for 2,5-diketopiperazine was thought to be due to 2,5-diketopiperazine not being in an α-helix, compared to proteins and peptides which should largely be in α-helices. Our calculations show the opposite result. This indicates that the cause of this absorption feature is probably not due solely to protein secondary structure , and other effects, such as bond lengths and angles may play a larger role. We note that this spectral insensitivity should be expected given the lack of mixing present when N2 is excited, the nitrogen most representative of bulk protein structure. In summary, NEXAFS is largely a local probe of chemical environments.

Finally, if one were to consider that vibrational effects will likely broaden the observed spectral features , it is reasonable to assume that the differences between the α-helix and the β-sheet will be further suppressed; even the minor peak at ~ 403 eV not present in the α-helix, will likely be obscured as a consequence of motions. It strongly indicates that while subtle changes in NEXAFS spectra can result from differences in conformation, the overall spectral sensitivity to conformation will be minor. However, if one could orient the proteins (for example, relative to a surface), the NEXAFS spectrum should then exhibit a large linear dichroism effect in the case of the α-helix which would not be found with a β-sheet (data not shown) and perhaps this effect could be used to identify the protein conformational state. Nevertheless, protein conformational information will be quite difficult to extract from bulk nitrogen K-edge spectra.

## 4. Conclusions

The nitrogen K-edge NEXAFS spectra of aqueous triglycine was measured, and the effects of salts near the extremes of the Hofmeister series on the spectra were observed. We simulated these spectra using the novel XCH method and were able to capture all the prominent experimental features, including those caused by salts. Specifically, the cosmotropic sodium sulfite was shown to cause a new feature in the nitrogen K-edge absorption spectrum. This feature is ascribed to a change in electronic structure as well as a reduction in conformational flexibility of triglycine, showing that for small peptides, direct ion-peptide interaction can be

detected.  However, we also point out that such effects may not be observable for a large protein. In fact, solid polyalanine will exhibit a very similar NEXAFS spectrum in either its β-sheet or α-helix forms at the nitrogen K-edge.

**Figure Captions**

**Figure 1.** Schematic drawings of triglycine (A), the sulfite anion (B) and polyalanine (C). The structure of the α-helix and anti-parallel β-sheet are shown in (D) and (E) respectively, with carbon in yellow, hydrogen in blue, oxygen in red, nitrogen in bluish grey. The unit cell used in the calculations is shown. Triglycine is in its zwitterionic form at the pH of this study. The nitrogen atoms of triglycine are labeled; this numbering scheme will be used throughout this paper. Note the two leftmost oxygen atoms on triglycine are degenerate, as are all the oxygens of the sulfite anion and every alanine monomer.

**Figure 2.** Experimental nitrogen K-edge NEXAFS spectrum of aqueous triglycine (black, solid) and the overall calculated spectrum for triglycine (purple, ••••). The calculated spectrum due to individual nitrogen atoms N1 (blue, — ••• —), N2 (red, – – –) and N3 (green, — • —) are shown.

**Figure 3.** Experimental nitrogen K-edge NEXAFS spectrum of aqueous triglycine with sodium bromide (black, solid) and the overall calculated spectrum for triglycine with sodium bromide (purple, ••••). The calculated spectrum due to individual nitrogen atoms N1 (blue, — ••• —), N2 (red, – – –) and N3 (green, — • —) are shown.

**Figure 4.** Experimental nitrogen K-edge NEXAFS spectrum of aqueous triglycine with sodium flouride (black, solid), triglycine with sodium chloride (green, solid) and triglycine with sodium bromide (red, solid). We note there are no new spectral features due to the salt interaction.

**Figure 5.** Experimental nitrogen K-edge NEXAFS spectrum of aqueous triglycine with sodium sulfite (black, solid) and the overall calculated spectrum for triglycine with sodium sulfite (purple, ••••). The calculated spectrum due to individual nitrogen atoms N1 (blue, — ••• —), N2 (red, – – –) and N3 (green, — • —) are shown.

**Figure 6.** Experimental nitrogen K-edge NEXAFS spectrum of aqueous triglycine with sodium sulfite (black, solid) and triglycine with potassium sulfate (red, solid) . We note both exhibit a new feature at approximately 403 eV.

**Figure 7.** Radial distribution functions from simulations of (A) triglycine with sodium bromide and (B) sodium sulfite with triglycine. The radial distribution functions shown are between the nitrogen atoms of triglycine and (A) Br⁻ as well as the radial distribution function between the nitrogen atoms of triglycine and (B) oxygen (from $SO_3^{2-}$). N1, N2 and N3 are shown in blue (solid), red (••••) and green (– – –) respectively for both A and B.

**Figure 8.** Result of RMS fitting the backbone of triglycine to the average structure for a given MD simulation. The carbon atoms are in blue, and the nitrogen atoms are in purple. The results correspond to the simulations of (A) triglycine, (B) triglycine and sodium bromide and (C) triglycine and sodium sulfite. The standard deviation of the fit for each atom corresponds to the

size of each atom.  Note how similar (A) and (B) are as compared to (C) and the often larger size of the spheroids in (A) and (B).

**Figure 9.**  Isosurfaces (orange and green) are shown for the LUMO+1 of the excited state of triglycine solvated with sodium sulfite and water, corresponding to 10% of the total integrated value.  Many of the waters have been removed for clarity.  The images show an excitation on N1, N2 and N3 from left to right.  Despite being a delocalized state, the state does not extend throughout the entire molecule.

**Figure 10.** Experimental nitrogen K-edge NEXAFS spectrum of solid lysozyme (purple, solid), 2,5-diketopiperazine (black, ••••) and polyisoleucine (green, — ••• —) from the literature (see text).  These are contrasted with calculated spectra at the nitrogen K-edge of solid polyalanine in an antiparallel β-sheet(red, solid) and an α-helix(blue, ••••).

**Figures**
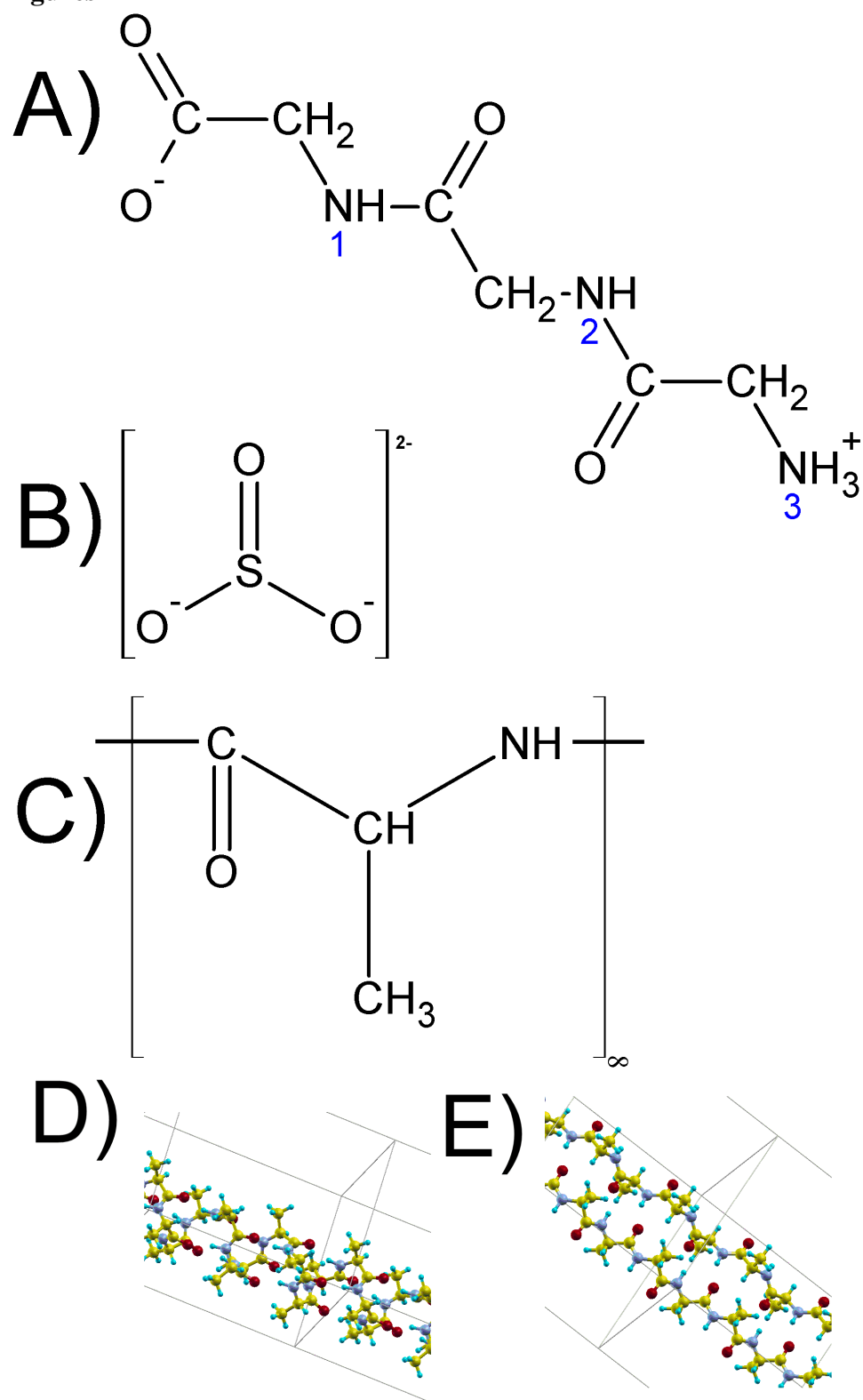


Figure 1
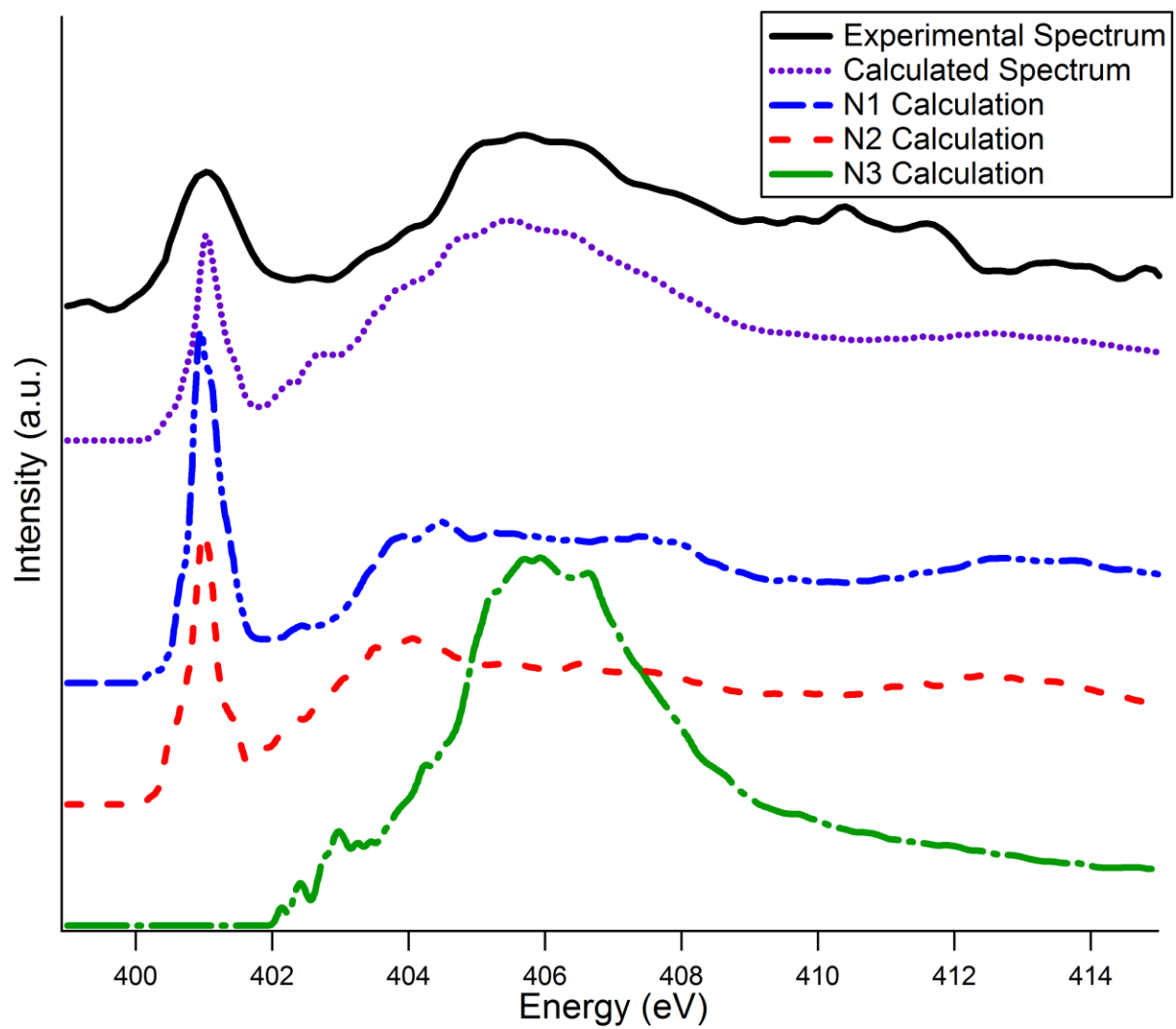
**Figure 2**
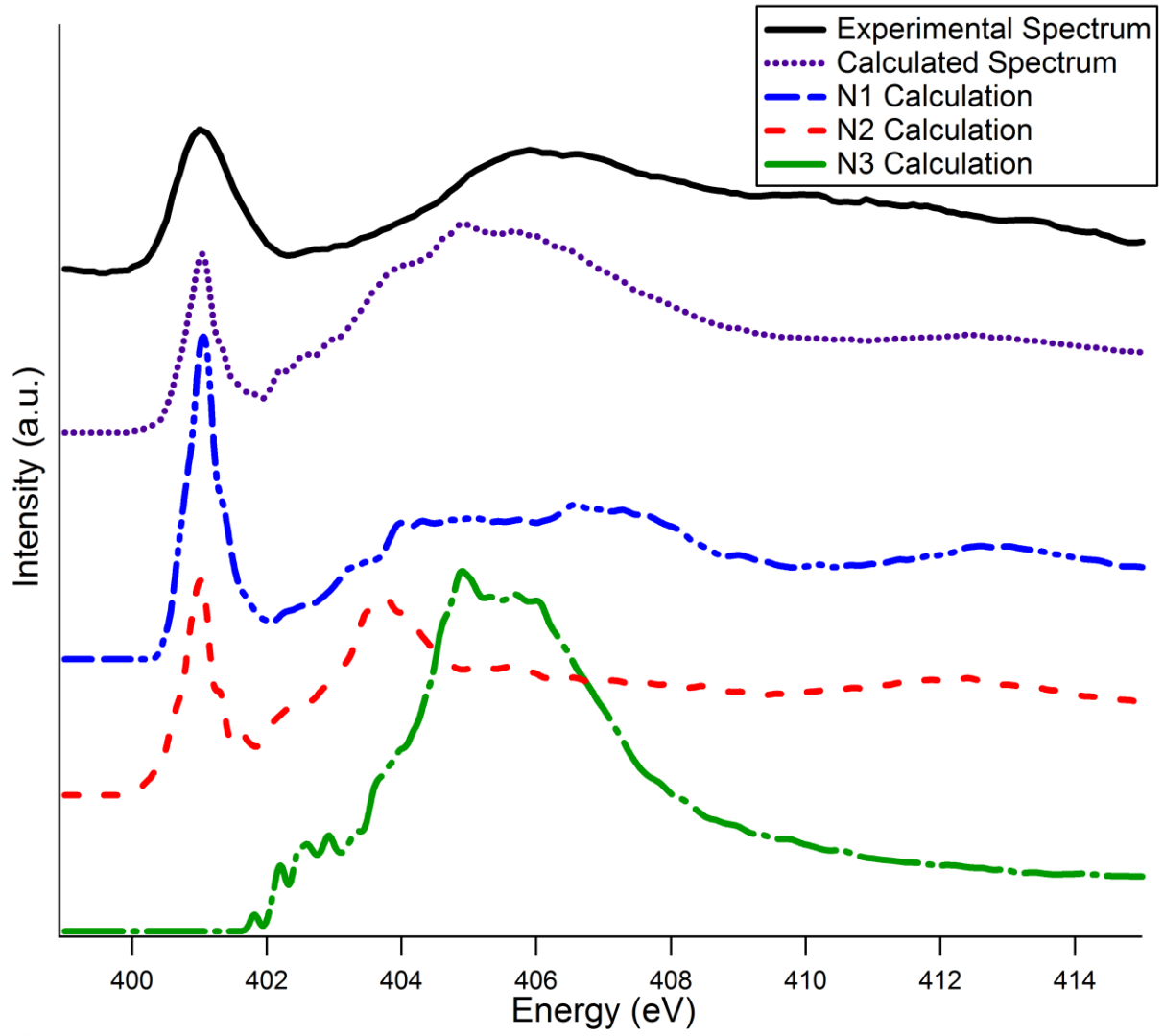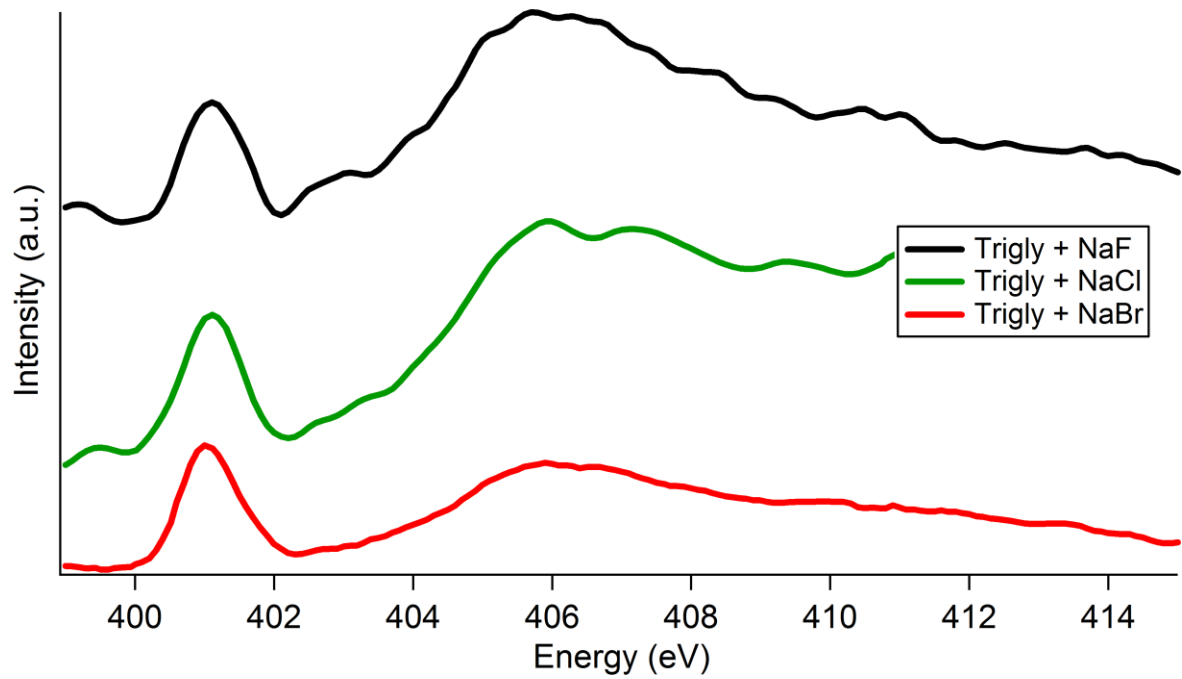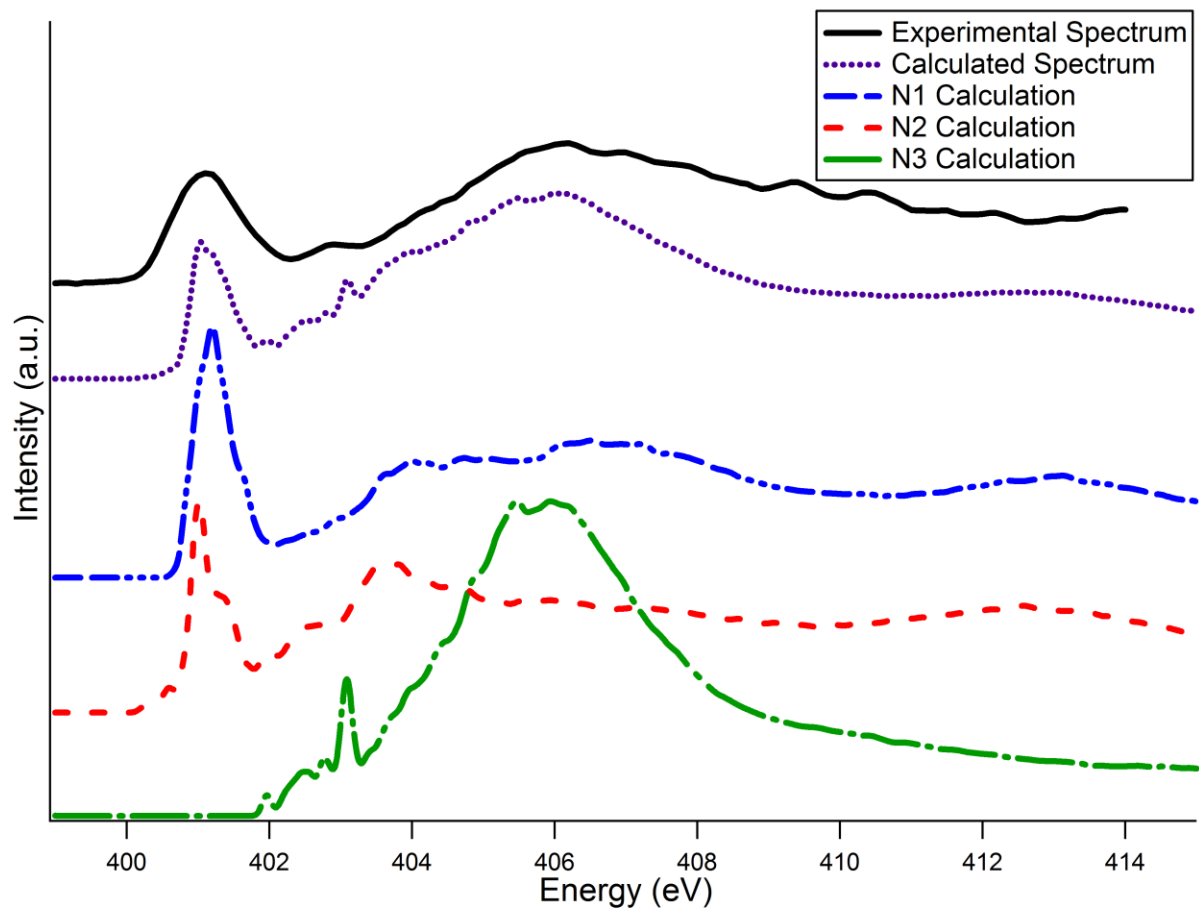
**Figure 3**
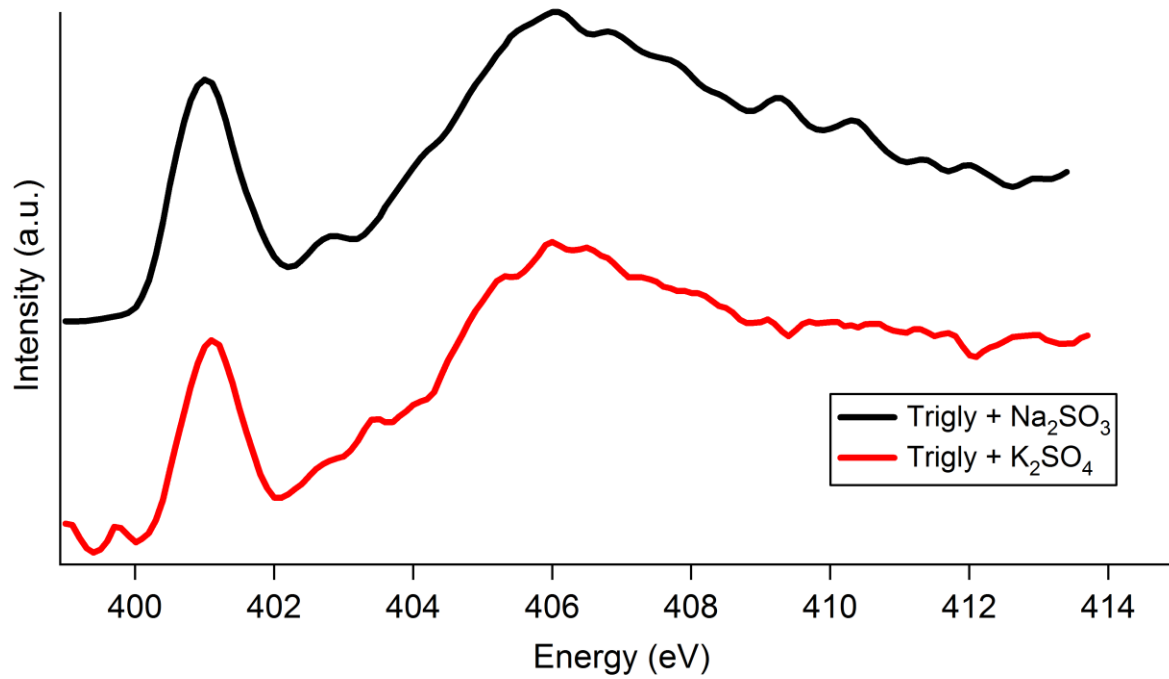
**Figure 4**

**Figure 5**

**Figure 6**

**Figure 7**

**Figure 8**

**N1**  **N2**  **N3**
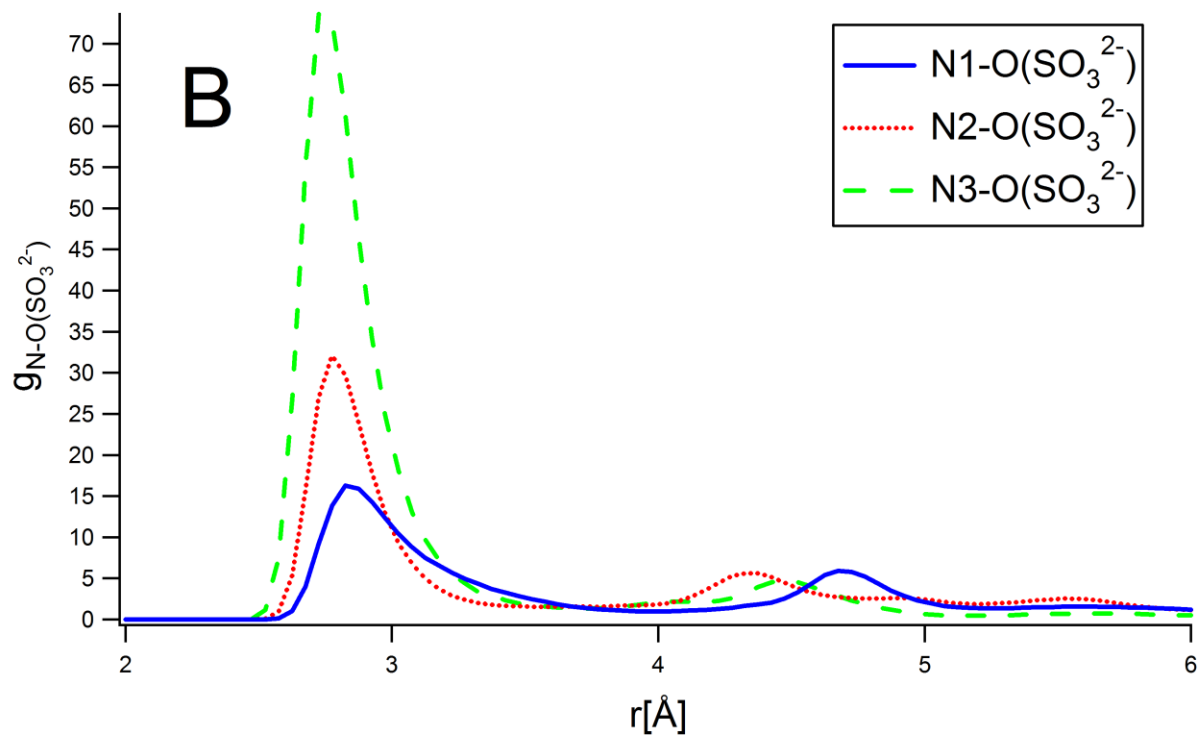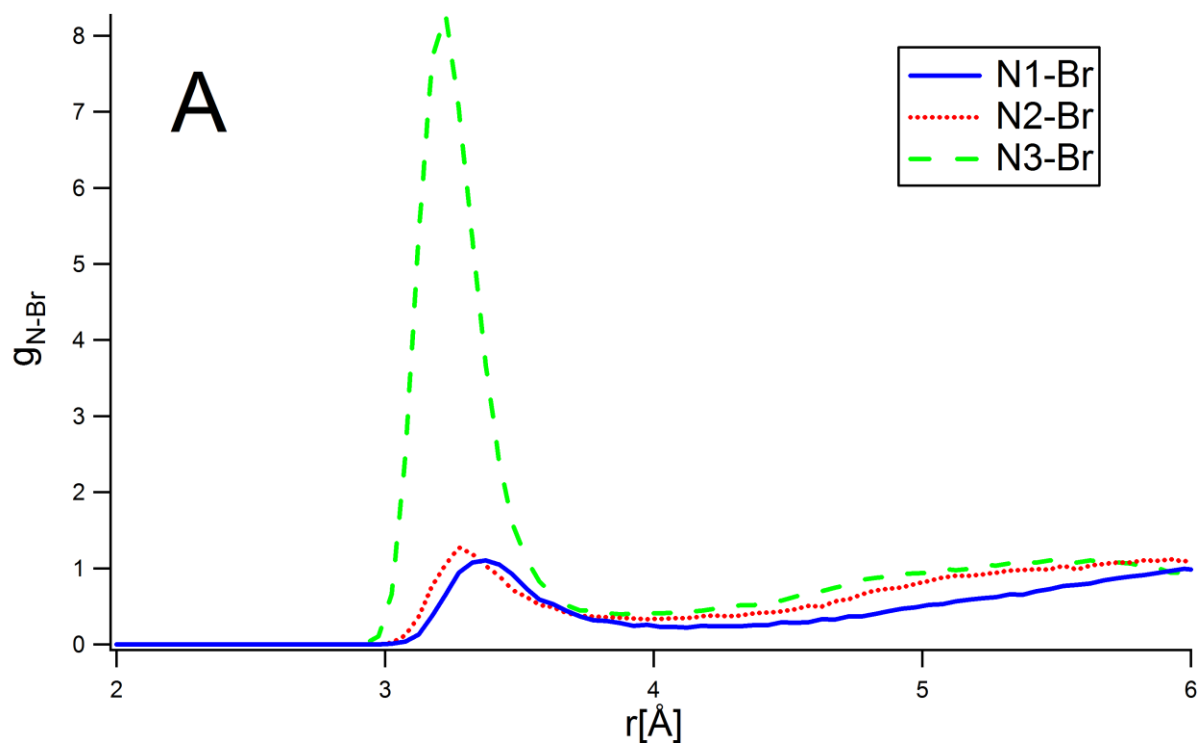
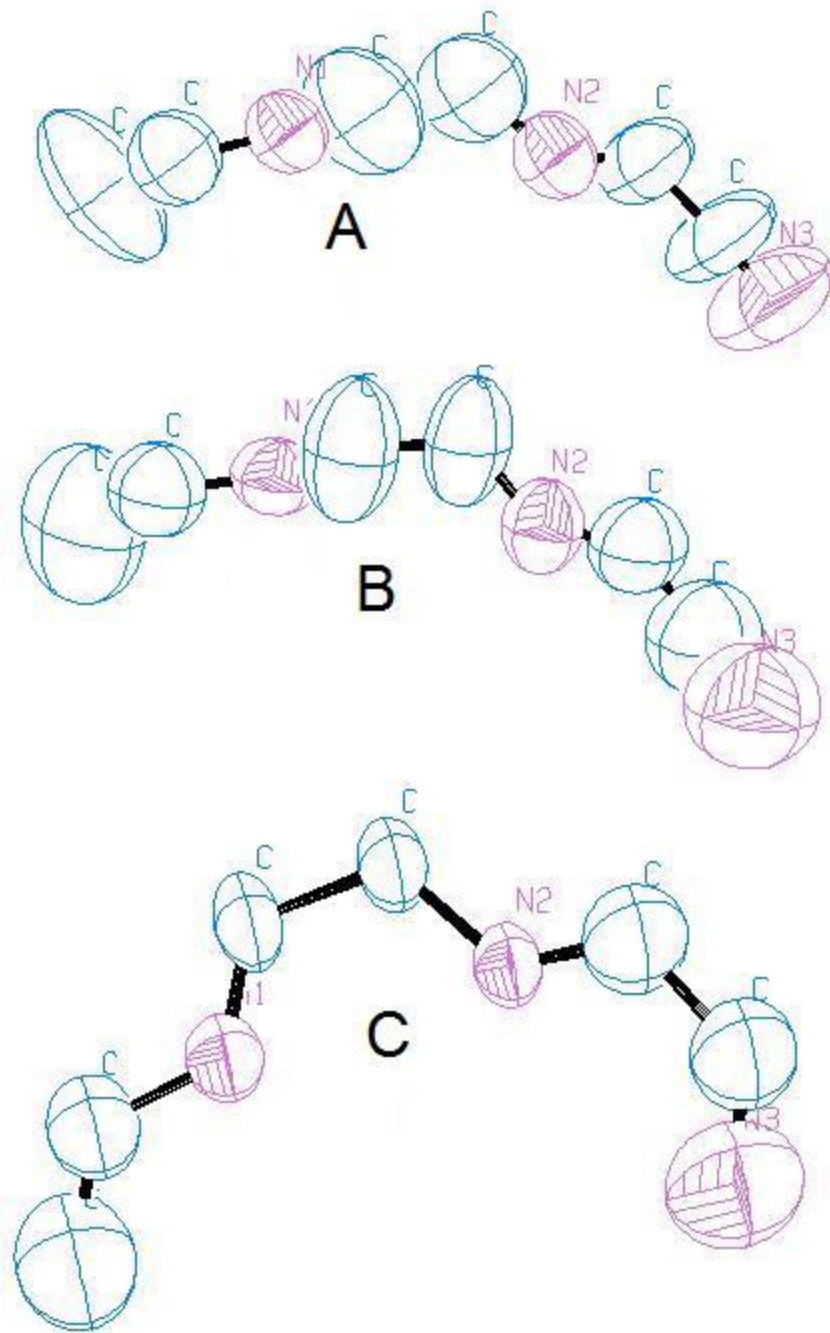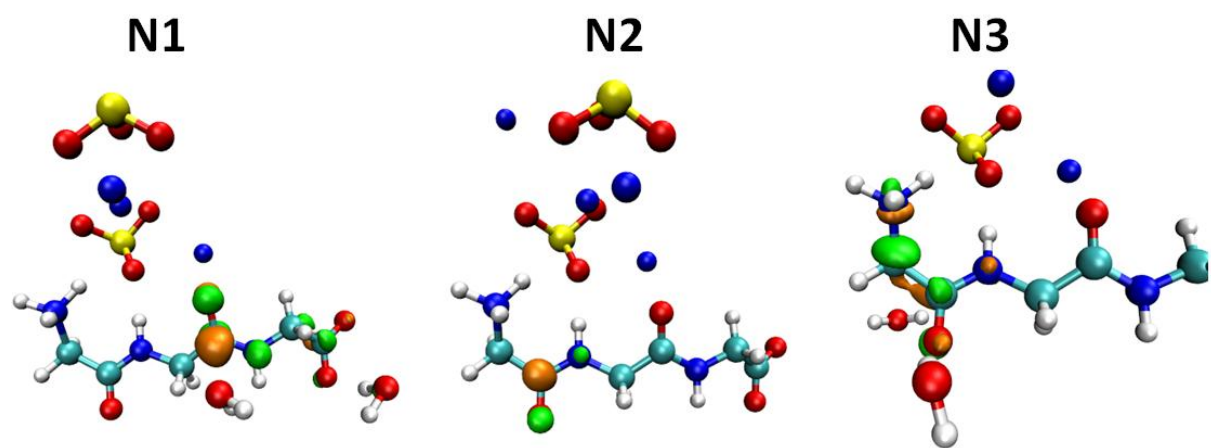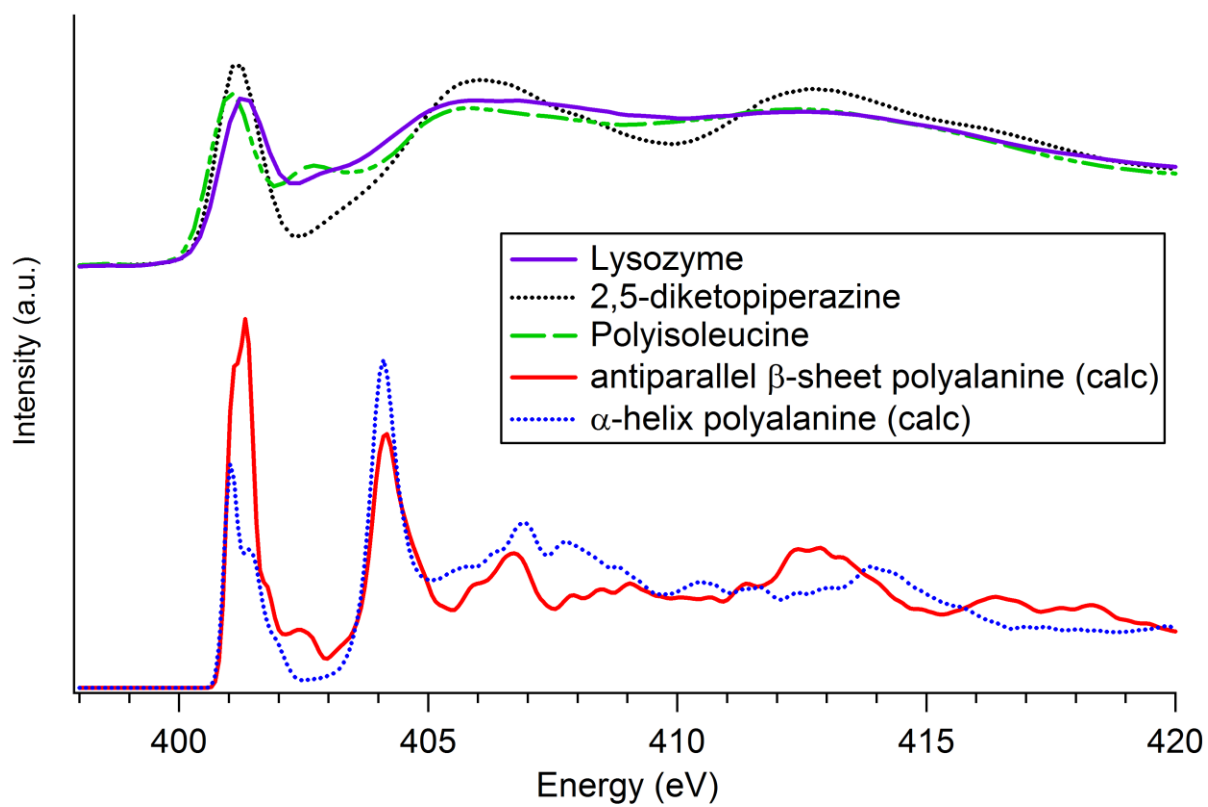**Figure 9**

**Figure 10**

1       F. Hofmeister, Arch. Exp. Pathol. Pharmakol. **24**, 247 (1888).
2       W. Kunz, J. Henle, and B. W. Ninham, Current Opinion in Colloid & Interface Science **9** (1-2), 19 (2004).
3       Y. J. Zhang and P. S. Cremer, Current Opinion in Chemical Biology **10** (6), 658 (2006).
4       E. P. Luo, Y. Q. Chai, and R. Yuan, Analytical Letters **40** (2), 369 (2007).
5       J. D. Batchelor, A. Olteanu, A. Tripathy, and G. J. Pielak, Journal of the American Chemical Society **126** (7), 1958 (2004).
6       J. D. Smith, C. D. Cappa, K. R. Wilson, R. C. Cohen, P. L. Geissler, and R. J. Saykally, Proceedings of the National Academy of Sciences of the United States of America **102** (40), 14171 (2005).
7       N. V. Nucci and J. M. Vanderkooi, Journal of Molecular Liquids **143** (2-3), 160 (2008).
8       B. Hess and N. F. A. van der Vegt, Proceedings of the National Academy of Sciences of the United States of America **106** (32), 13296 (2009).
9       K. D. Collins, Biophysical Chemistry **119** (3), 271 (2006).
10      N. Vlachy, B. Jagoda-Cwiklik, R. Vacha, D. Touraud, P. Jungwirth, and W. Kunz, Advances in Colloid and Interface Science **146** (1-2), 42 (2009).
11      P. E. Mason, C. E. Dempsey, L. Vrbka, J. Heyda, J. W. Brady, and P. Jungwirth, Journal of Physical Chemistry B **113** (10), 3227 (2009).
12      Y. H. Cho, Y. J. Zhang, T. Christensen, L. B. Sagle, A. Chilkoti, and P. S. Cremer, Journal of Physical Chemistry B **112** (44), 13765 (2008).
13      Y. Zhang, S. Furyk, L. B. Sagle, Y. Cho, D. E. Bergbreiter, and P. S. Cremer, Journal of Physical Chemistry C **111** (25), 8916 (2007).
14      Y. J. Zhang, S. Furyk, D. E. Bergbreiter, and P. S. Cremer, Journal of the American Chemical Society **127** (41), 14505 (2005).
15      E. F. Aziz, N. Ottosson, M. Faubel, I. V. Hertel, and B. Winter, Nature **455** (7209), 89 (2008).
16      K. R. Wilson, B. S. Rude, T. Catalano, R. D. Schaller, J. G. Tobin, D. T. Co, and R. J. Saykally, Journal of Physical Chemistry B **105** (17), 3346 (2001).
17      J. S. Uejio, C. P. Schwartz, A. M. Duffin, W. S. Drisdell, R. C. Cohen, and R. J. Saykally, Proceedings of the National Academy of Sciences of the United States of America **105** (19), 6809 (2008).
18      E. F. Aziz, N. Ottosson, S. Eisebitt, W. Eberhardt, B. Jagoda-Cwiklik, R. Vacha, P. Jungwirth, and B. Winter, Journal of Physical Chemistry B **112** (40), 12567 (2008).
19      K. R. Wilson, B. S. Rude, J. Smith, C. Cappa, D. T. Co, R. D. Schaller, M. Larsson, T. Catalano, and R. J. Saykally, Review of Scientific Instruments **75** (3), 725 (2004).
20      J. S. Uejio, C. P. Schwartz, R. J. Saykally, and D. Prendergast, Chemical Physics Letters **467** (1-3), 195 (2008).
21      D. Prendergast and G. Galli, Physical Review Letters **96** (21) (2006).
22      C. P. Schwartz, J. S. Uejio, R. J. Saykally, and D. Prendergast, The Journal of Chemical Physics **130** (18), 184109 (2009).
23      C. P. Schwartz, J. S. Uejio, A. M. Duffin, A. H. England, D. Prendergast, and R. J. Saykally, Journal of Chemical Physics **131** (11) (2009).
24      J. S. Uejio, C. P. Schwartz, A. M. Duffin, A. H. England, D. Prendergast, and R. J. Saykally, Journal of Physical Chemistry B **Accepted** (2010).

126

[25]   M. V. Vener, A. N. Egorova, D. P. Fomin, and V. G. Tsirelson, Journal of Physical Organic Chemistry **22** (3), 177 (2009).

[26]   Y. Zubavichus, A. Shaporenko, M. Grunze, and M. Zharnikov, Journal of Physical Chemistry A **109** (32), 6998 (2005).

[27]   Y. Zubavichus, A. Shaporenko, M. Grunze, and M. Zharnikov, Journal of Physical Chemistry B **111** (33), 9803 (2007).

[28]   Y. Zubavichus, A. Shaporenko, M. Grunze, and M. Zharnikov, Journal of Physical Chemistry B **112** (15), 4478 (2008).

[29]   J. Stewart-Ornstein, A. P. Hitchcock, D. H. Cruz, P. Henklein, J. Overhage, K. Hilpert, J. D. Hale, and R. E. W. Hancock, Journal of Physical Chemistry B **111** (26), 7691 (2007).

[30]   S. Baroni, A. D. Corso, S. D. Gironcoli, and P. Giannozzi, PWSCF (2008).

[31]   P. Hohenberg and W. Kohn, Physical Review B **136** (3B), B864 (1964).

[32]   W. Kohn and L. J. Sham, Physical Review **140** (4A), 1133 (1965).

[33]   J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Physical Review B **46** (11), 6671 (1992).

[34]   J. P. Perdew, K. Burke, and M. Ernzerhof, Physical Review Letters **77** (18), 3865 (1996).

[35]   L. A. Naslund, D. C. Edwards, P. Wernet, U. Bergmann, H. Ogasawara, L. G. M. Pettersson, S. Myneni, and A. Nilsson, Journal of Physical Chemistry A **109** (27), 5995 (2005).

[36]   T. A. D. D.A. Case, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., K. M. M. Luo, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S., A. R. Hayik, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H., L. Y. Gohlke, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C., and W. S. R. Schafmeister, and P.A. Kollman, (2006).

[37]   R. L. C. Wang, H. J. Kreuzer, and M. Grunze, Physical Chemistry Chemical Physics **8** (41), 4744 (2006).

[38]   J. Stöhr, *NEXAFS Spectroscopy*. (Springer, Berlin, 1992).

[39]   J. Heyda, J. C. Vincent, D. J. Tobias, J. Dzubiella, and P. Jungwirth, Journal of Physical Chemistry B **114** (2), 1213.

# Chapter 9 – Future Work

In the future, two things would greatly aid in the ability to understand complicated spectra.  The first is the coupling of a cluster source with the high flux X-ray free electron lasers currently under construction.[1] The combination of high flux and a cluster source would allow for the determination of the spectrum of water clusters.  This could be done on charged clusters by mass selecting the cluster of the correct size and subtracting the neutral background.  For uncharged clusters, lasers could be used to systematically destroy the relevant cluster, performing a form of loss spectroscopy.  By tuning to an energy that only, for example the water dimer absorbed, the x-ray absorption of the water dimer could be determined.[2]

In order to interpret the results generated from these experiments would require a careful mapping of the potential energy surface.  The best way to do this would be to carefully map out the potential energy surface of the system, and attempt to apply a localized mode analysis.[3] This would allow for the constantly discussed but rarely achieved building up of a liquid a single molecule at a time, and in the not that terribly distant future would allow for the analytical solution to the liquid water spectrum to be built up a molecule at a time.[4]

[1]    K. A. Nugent, Advances in Physics **59** (1), 1.
[2]    A. Moudens, R. Georges, M. Goubet, J. Makarewicz, S. E. Lokshtanov, and A. A. Vigasin, Journal of Chemical Physics **131** (20) (2009).
[3]    B. Njegic and M. S. Gordon, Journal of Chemical Physics **129** (16) (2008).
[4]    A. J. Huneycutt and R. J. Saykally, Science **299** (5611), 1329 (2003).