Monopeptide versus Monopeptoid: Insights on Structure and Hydration of Aqueous Alanine and Sarcosine via X-ray Absorption Spectroscopy

Janel S. Uejio,[†,‡] Craig P. Schwartz,[†,‡] Andrew M. Duffin,[†,‡] Alice England,[†,‡] David Prendergast,[§] and Richard J. Saykally*[,†,‡]

Department of Chemistry, University of California, Berkeley, California 94720-1460 and Chemical Sciences Division, Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received: November 19, 2009; Revised Manuscript Received: March 3, 2010

Despite the obvious significance, the aqueous interactions of peptides remain incompletely understood. Their synthetic analogues called peptoids (poly-N-substituted glycines) have recently emerged as a promising biomimetic material, particularly due to their robust secondary structure and resistance to denaturation. We describe comparative near-edge X-ray absorption fine structure spectroscopy studies of aqueous sarcosine, the simplest peptoid, and alanine, its peptide isomer, interpreted by density functional theory calculations. The sarcosine nitrogen K-edge spectrum is blue shifted with respect to that of alanine, in agreement with our calculations; we conclude that this shift results primarily from the methyl group substitution on the nitrogen of sarcosine. Our calculations indicate that the nitrogen K-edge spectrum of alanine differs significantly between dehydrated and hydrated scenarios, while that of the sarcosine zwitterion is less affected by hydration. In contrast, the computed sarcosine spectrum is greatly impacted by conformational variations, while the alanine spectrum is not. This relates to a predicted solvent dependence for alanine, as compared to sarcosine. Additionally, we show the theoretical nitrogen K-edge spectra to be sensitive to the degree of hydration, indicating that experimental X-ray spectroscopy may be able to distinguish between bulk and partial hydration, such as found in confined environments near proteins and in reverse micelles.

Introduction

Peptoids have recently emerged as an important subject of research.[1,2] They can be readily synthesized[3] and are resistant to enzymatic, chemical, and thermal denaturation.[4,5] Peptoids are currently being evaluated for numerous applications, including cell penetration for drug delivery,[6] protein−protein interaction inhibitors,[7] lung surfactants,[8] and other biomimetics.[1,2,9] Salient characteristics of peptoids include their robust higher-order structure, folding dynamics, increased resistance to denaturation relative to proteins, and the means by which the structure and interactions can be intelligently designed by the primary sequence of the chain.[2,3,10−12] Previous methods of studying peptides and peptoids have included circular dichroism of oligomers,[5,11,13] infrared spectroscopy,[12] NMR,[11] mass spectrometry,[14,15] and computational modeling;[16,17] these methods primarily focused on design and characterization of the secondary structure.

We sought to understand the structural aspects of aqueous peptoids on a fundamental level by comparing sarcosine with its complementary peptide, alanine, using near-edge X-ray absorption fine structure (NEXAFS) spectroscopy, which is sensitive to details of charge interactions, presence of counterions, conformation, and hydration.[18−24] Comprehensive and predictive design requires an understanding of the environment and its impact on the molecule of interest, and in the case of biological applications, this necessitates addressing the aqueous systems. Previous NEXAFS studies of peptides have primarily been limited by technical requirements to experiments on the gas, solid phase, thin films, or small static samples.[23−31] To enable direct measurements of aqueous systems, we employed liquid microjet technology, where a thin stream of liquid is injected into vacuum and is intersected by a synchrotron X-ray beam.[19,20,32,33] With results from this system, we present the first direct comparison of the properties of aqueous alanine and sarcosine using NEXAFS.

To understand a solute embedded in a solvent, several factors must be initially considered, including but not limited to solute dynamics, intramolecular interactions of the solute (interactions between the carboxylate and charged nitrogen group), solute−solute, and solute−solvent interactions. As these effects are hard to separate, calculations are necessary for interpretation and predictive capability. Our first principles density functional theory (DFT) approach has been shown to be successful for addressing both isolated organic molecules and for water.[34−37] Many X-ray calculations done in the past on hydrated systems have used clusters to simulate bulk hydration;[38−42] the question arises as to the number of waters necessary to accurately represent solvation. Further, previous practice has placed waters around a moiety of interest, instead of the entire molecule but whether this is an appropriate practice has not been addressed.[18−20] Additionally, finite clusters of water possess surface electronic states due to dangling hydrogen bonds; these states are not present for calculations with periodic boundary conditions. For these reasons, we use a periodic simulation cell to model bulk hydration for X-ray calculations.

In this paper, we present the first NEXAFS spectra of aqueous alanine and sarcosine, along with accompanying DFT calculations. We untangle the spectral impact of molecular conformation versus the strength of hydration and contrast these effects between alanine and sarcosine. We also find that for NEXAFS simulations care must be exercised in how one approximates

* To whom correspondence should be addressed.
† University of California.
‡ Chemical Sciences Division, Lawrence Berkeley National Laboratory.
§ Molecular Foundry, Lawrence Berkeley National Laboratory.

10.1021/jp911007k
Published on Web 03/17/2010

the system under consideration, for example, the spectra of clusters do not represent the electronic structure of the condensed phase and vice versa.

Methods

Samples. Samples of alanine and sarcosine were purchased from Sigma Aldrich at stated purity of 99% or higher and were used without further purification; 1 M solutions were prepared using 18 MΩ (Millipore) water.

Experimental Section. The NEXAFS spectra were obtained at Beamline 8.0.1 of the Advanced Light Source (ALS) at Lawrence Berkeley National Lab. Nitrogen K-edge spectra were collected both from the liquid jet as well as from the gas phase background. A baseline intensity, I_o, was collected farther up the beamline using a gold mesh, which was used to account for intensity fluctuations. The spectra were subsequently area normalized for ease of comparison. We collected spectra for the carbon, nitrogen, and oxygen K-edge for both sarcosine and alanine, but found the largest and most interesting differences on the nitrogen K-edge. Water dominates the oxygen K-edge spectra; the carbon K-edge spectra only minimally differ between sarcosine and alanine, being relatively insensitive to the change in the location of the methyl group.

The experimental apparatus comprises a main chamber wherein the X-ray beam intersects with the microjet stream, a skimmer, a section for cryogenically trapping the microjet that reduces pressures, and a differential pumping section that allows for windowless coupling to the beamline. The measurements in this paper were obtained using a chopped fused silica capillary of 30 μm attached to an HPLC pump with an average flow rate of 1 mL/min. The temperature of the sample was approximately 300 K. The signal was collected by total electron yield (TEY) with a copper electrode biased at +1.8 kV; the use of TEY has been shown to provide bulk characterization due to the large escape depths of electrons.[43] A more detailed description of the experimental apparatus has been previously published by Wilson et al.[33]

Theoretical. As has previously been described,[34−37] accurately simulating NEXAFS of molecular systems requires numerical spectral convergence from adequate conformational space sampling. Classical molecular dynamics snapshots, in addition to fully relaxed structures, were input to first principles density functional theory (DFT) calculations for theoretical K-shell spectra.[44,45] We used AMBER 9 with generalized AMBER force field and Antechamber for the construction of the solute, box, and the MD.[46,47] The MD simulations were calculated in an approximately (12 Å)^3 orthorhombic simulation cell under periodic boundary conditions, where the solute and solvent are replicated. We note that the presence of spurious interactions between periodic images of the same molecule and its surrounding solvent does introduce some false correlation in the structural and electronic contributions to the X-ray absorption spectrum; however, by analogy with previous work on water,[36] this effect should be spectrally insignificant for cell volumes of this size, although we have not explicitly investigated this point. The system was first relaxed, then equilibrated at constant standard pressure followed by equilibration at constant temperature before recording a trajectory at 300 K using a Langevin thermostat. The simulation cell used contained 78 TIP3P waters;[48] this corresponds to a concentration of 0.7 M, slightly less concentrated than experiment. Higher concentrations necessitate either smaller simulation cells with fewer waters but incomplete solvation shells or much larger simulation cells including more waters plus an additional solute molecule, making the first principles calculations too expensive. Published studies establish that polarizable force fields for water are critical for simulations to achieve correct structure and dynamics for an ionic solute.[49,50] The effect of a first principles molecular dynamics on the spectra of ionic solutes is the subject of future work. Periodic boundary conditions also eliminate surface electronic states; surface electronic states occur in finite cluster X-ray calculations due to dangling hydrogen bonds.

X-ray absorption spectra for each MD snapshot were calculated using DFT, where we self-consistently calculate the lowest energy core-level excited state with a full core hole and explicitly including the excited electron (XCH).[35−37] Transition amplitudes are approximated within the single-particle and dipole approximations. We employed the Perdew−Burke−Ernzerhof form of the generalized gradient approximation to the exchange-correlation potential.[51] These simulations employed the same periodic boundary conditions as the MD simulations, explicitly including all 78 waters and using a plane wave basis set to represent the electronic structure. Norm-conserving pseudopotentials were used in the course of the calculation, necessitating a plane-wave kinetic energy cutoff of 85 Ry. Under periodic boundary conditions, the electron wavevector **k** is well-defined and expectation values require an integration over the first Brillouin zone in **k**-space. Upon numerical convergence, this approach has been found to accurately approximate the spectra of disordered condensed phases.[36,37] A grid of 5 × 5 × 5 k-points were used to converge the Brillouin zone integration for the X-ray absorption calculations. For more details about the theoretical approach and its development, see previous work.[34−37] Self- consistent charge densities and states were computed using PWSCF.[52] Calculations were performed on the Franklin supercomputer at NERSC, requiring over 700 CPU hours per hydrated snapshot.

Each calculated spectrum was homogeneously broadened by Gaussian convolution with a constant full width at half-maximum of 0.2 eV. By numerical experience, this broadening is considered small with respect to inhomogeneous broadening introduced by nuclear motion for molecular systems, where nondegenerate molecular orbital energies are typically well-separated at the onset of absorption but can vary within a 1 eV range with respect to molecular conformation.[35] These numerically broadened spectra were then averaged over all available snapshots. In this work, we found that using 100 snapshots is more than sufficient for spectral convergence; that is, after 60 snapshots, the addition of more snapshots did not alter the average spectrum. Unlike all-electron calculations, DFT pseudopotential calculations lack an absolute energy reference. Consequently, we employ a relative energy scale, wherein the spectra are aligned to one another using total energy differences.[34] Individual features within a given spectrum should have the correct alignment but with an underestimation of their energy spacing due to typical DFT underestimation of excited state energies.[36] Finally, for comparison with experiment the computed spectra are rigidly shifted to the experimental onset.

Discussion

Figure 1 shows the structures for both sarcosine and alanine, illustrating their chemical differences when in water. In Figure 2, we present the experimental nitrogen K-edge spectra for solvated alanine and sarcosine compared with the calculated spectra. The theoretical spectra are computed using 100 snapshots from an MD trajectory where the solute is free to move and is fully hydrated. Upon inspection, the main difference between the two spectra is the slight blue shift apparent in the

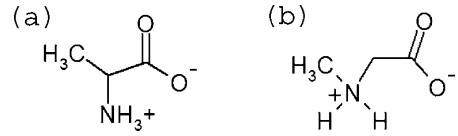

Figure 1. Structures of (a) alanine and (b) sarcosine in their zwitterionic forms; the negative charge is resonant between the two oxygens.

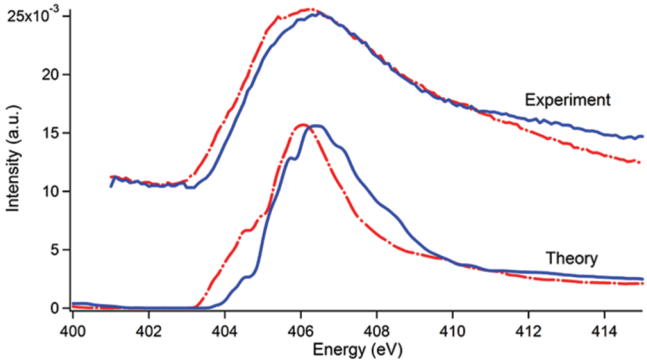

Figure 2. Measured nitrogen K-shell NEXAFS spectra of alanine and sarcosine and the computed spectra aligned to experiment are shown above. For both experiment and theory, sarcosine (blue) is shifted slightly to higher energy than alanine (red and dashed); the slight shift is reproduced well in the computed spectra and is assigned primarily to methyl group substitution on the nitrogen in sarcosine.

leading edge (the lower energy features) of the sarcosine spectrum with respect to alanine, indicating a change in the electronegativity or degree of bonding with neighboring atoms. This is analogous to what is seen for a carbonyl group; when changing its neighboring moiety to a more electronegative functional group, a blue shift occurs on the oxygen K-edge.[53] The methyl group substitution on the nitrogen in sarcosine altered the bonding around the nitrogen, shifting transitions to a lower energy. The computed sarcosine spectra reproduce the general shape as well as the slight blue-shift found in the experimental spectra. Spectral features for both simulated alanine and sarcosine appear closer together in energy than in the experiment due to the known DFT bandwidth underestimation. We deliberately chose a narrow Gaussian broadening to expose distinct spectral features underlying the entire spectrum. The good agreement of experimental trends with the calculated spectra, Figure 2, is a validation of the computational method used and of the predictions we make in the following sections.

Spectral Effects of Conformation and Hydration. The MD trajectory sampled for the spectra shown in Figure 2 comprised 78 explicit waters and an unconstrained solute; this is referred to as "hydrated-free" and includes both conformation and hydration effects. To examine and enumerate the contributions of different factors to the spectrum (e.g., hydration, conformation), MD simulations were run to create different sets of coordinates under various conditions. Previous work has shown that the inclusion of internal molecular motion is very important for computational X-ray absorption studies;[35] this was confirmed by first performing calculations of the completely relaxed gas phase structures of alanine and sarcosine for both zwitterion and neutral forms and finding these to be a poor representation of experiment (spectra not shown). To isolate the impact of solute conformation on the spectrum, the unconstrained and fully solvated MD trajectory (hydrated-free) was stripped of waters, leaving a trajectory with only the solute present in different conformations. These snapshots include all the motions of the solute that occur in the presence of water; thus, we are able to examine whether the changes in the spectrum result from the

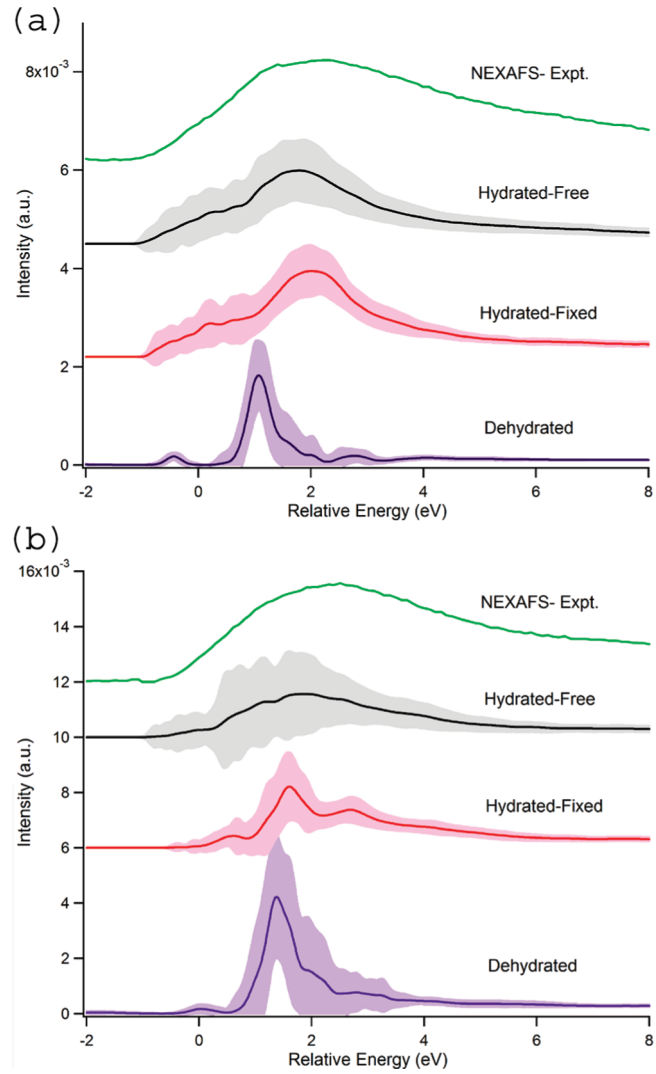

Figure 3. Computed (a) alanine (b) sarcosine nitrogen K-shell spectra with hydrated-free, hydrated-fixed, and dehydrated coordinates (see text) compared to experiment. The error (shown as shaded area) indicate one standard deviation based on MD sampling. The experimental spectra were shifted to align with the onset of the hydrated-free spectra.

conformations induced by interactions with the water or from the presence of the surrounding solvent itself. These structures were used for the DFT calculations, are shown in Figure 3, and are referred to as "dehydrated." Notably, the dehydrated calculations are not physically meaningful, as both alanine and sarcosine would be neutral, not zwitterionic in the gas phase. Nor does the dehydrated reproduce a gas phase MD trajectory based on the zwitterion, which would lack conformational changes induced by solvent interaction. Finally, to isolate the impact of hydration without the inclusion of conformation effects, another MD trajectory was generated where the solute structure was constrained, while the surrounding waters were allowed to move freely; these calculations are labeled as "hydrated-fixed." The chosen solute conformation of the hydrated fixed trajectory was taken from the original hydrated-free MD simulation and chosen for its typical conformation with regard to intermolecular angles (such as the dihedral angle), distances and being a reasonably low-energy structure. We recognize that both the hydrated-fixed and the dehydrated simulations are unphysical situations, but they serve as models to isolate the effects of conformation and hydration on the calculated spectra.

Comparing spectra generated by the three MD sampling techniques and experimental spectrum for alanine, Figure 3a, it is clear that holding the hydrated molecule rigid (hydrated-fixed) or removing solvation (dehydrated) produces inadequate representations of the measured spectrum. The most accurate calculated spectrum is, unsurprisingly, the unconstrained and fully hydrated MD-based spectrum (hydrated-free); this approach reproduces the general shape and the overall shift between sarcosine and alanine observed in experiment (Figure 2), verifying that this is the best technique for hydrated NEXAFS. However, in the case of alanine, the hydrated-fixed coordinates approximate both the experimental spectrum and the hydrated-free spectrum better than the corresponding dehydrated spectrum. This suggests that hydration has a larger impact on the electronic structure of alanine than conformational changes. The strong effect of hydration implies that given a different solvent, the electronic structure would again be altered, that is, a solvent dependence. Alanine's solvent dependence has been found in other studies; in vacuum, a polyalanine molecule assumed an α-helix and was only marginally stable once it was hydrated.[15] Additionally, a blocked alanine was found to have solvent dependence in FT-IR experiments.[54] Notably, more broadening occurs in the hydrated-free spectrum due to the additional structural changes (versus the hydrated-fixed) and the inclusion of hydrogen bonding structures (versus the dehydrated spectrum).

Figure 3b shows a similar comparison for sarcosine, between experiment and simulation for the hydrated unconstrained solute (hydrated-free), hydrated with constrained solute (hydrated-fixed), and dehydrated solute scenarios. The best approximation for the experimental spectrum is, similar to alanine, the unhindered fully solvated case. Yet, for sarcosine there is less agreement between the hydrated-free and hydrated-fixed scenarios implying conformation is a larger spectral contributor than it is for alanine. As there are numerous conformations, it is impossible to say whether this is true for each; rather, this typical (in terms of structure and energy) structure is an indicator of lesser solvent contribution to the electronic structure. This will be examined further later in the text with regard to hydrogen bonding. Intramolecular interaction is possible for both sarcosine and alanine in the gas phase for both the neutral and zwitterion.[15,55] On the basis of visual inspection of the snapshots, intramolecular interaction is reduced upon hydration; still, the degree of interaction is greater for sarcosine than alanine based on the proximity brought about by the methyl substitution. Given increased intramolecular interaction any change in angle would result in changes in the strength of these interactions and the electronic structure and would be seen spectrally.

To further examine the relative impact of conformation on spectra, individual configurations with various dihedral angles were chosen and their spectra calculated; Figure 4a,b presents the results for a variety of angles for both sarcosine and alanine, respectively; these measurements are based on the trajectories from the dehydrated calculations. These snapshots are from the dehydrated trajectory to isolate the impact of conformation alone and simplify interpretation. The angles indicated are the sum of dihedral angles illustrated in the insets of Figure 4; dihedral angles were used because they provide a better representation of possible intramolecular interaction and overall molecular strain. Previous computational work has shown that for gas phase glycine the nitrogen K-edge is sensitive to such conformational changes, therefore some angle dependence was predicted.[22,34,35] The distribution of dihedral angles for sarcosine (standard deviation approximately 90°) is much wider than for alanine (standard deviation approximately 50°). The free rotation of the NH_3^+ of alanine gives somewhat degenerate spectral features; whereas, the methyl on the nitrogen of sarcosine results in lower symmetry around the nitrogen and the sampling of more unique positions such that conformation has a greater spectral role. Examining Figure 4, it is evident that sarcosine exhibits larger spectral variation depending on the dihedral angle as opposed to alanine; this implies that sarcosine is more conformation dependent than alanine.

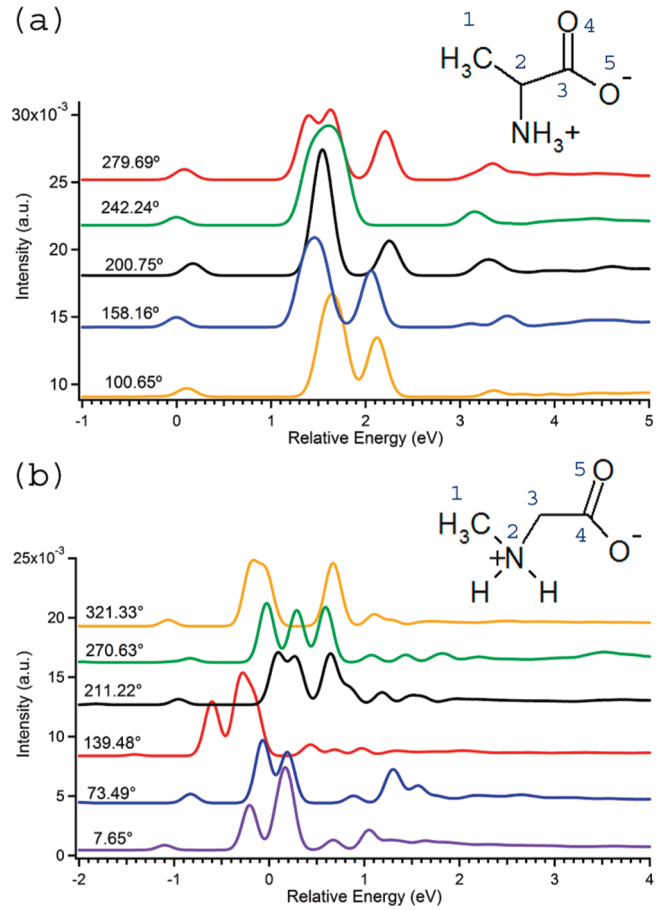

Figure 4. Calculated N K-edge spectra of (a) alanine (b) sarcosine for a range of dihedral angles; the number shown on the left is the sum of the two dihedral angles. These spectra do not include any water molecule but are from the dehydrated trajectory, and serve to isolate the impact of changing the dihedral angles. The indicated values are the sum of ∠1,2,3,4 and ∠1,2,3,5 for alanine, and ∠1,2,3,4 combined with ∠2,3,4,5 for sarcosine.

As a further test of the effect of solvent on sarcosine and alanine we compare different configurations from the hydrated-fixed calculations to examine the impact of the hydrogen bonding without conformational changes (Figure 5). For sarcosine and alanine, two snapshots each are chosen with high and low number of hydrogen bonds. A hydrogen bond was defined as having a distance less than 2.5 Å, an acceptor angle of 120° or greater, and a donor angle of 90° or more; a schematic is shown in Figure 5c. For alanine, Figure 5a, there are large spectral differences between the two snapshots, despite having the same solute coordinates. For alanine, neither of the snapshots resemble the ensemble-averaged spectrum, which requires an average of many snapshots. Sarcosine's two hydrated snapshots, Figure 5b, show little change between the two spectra other than minor shifts and a slight change in the intensity of the shoulder to the blue of the main peak. The two spectra from individual snapshots seem to be narrower versions of the average

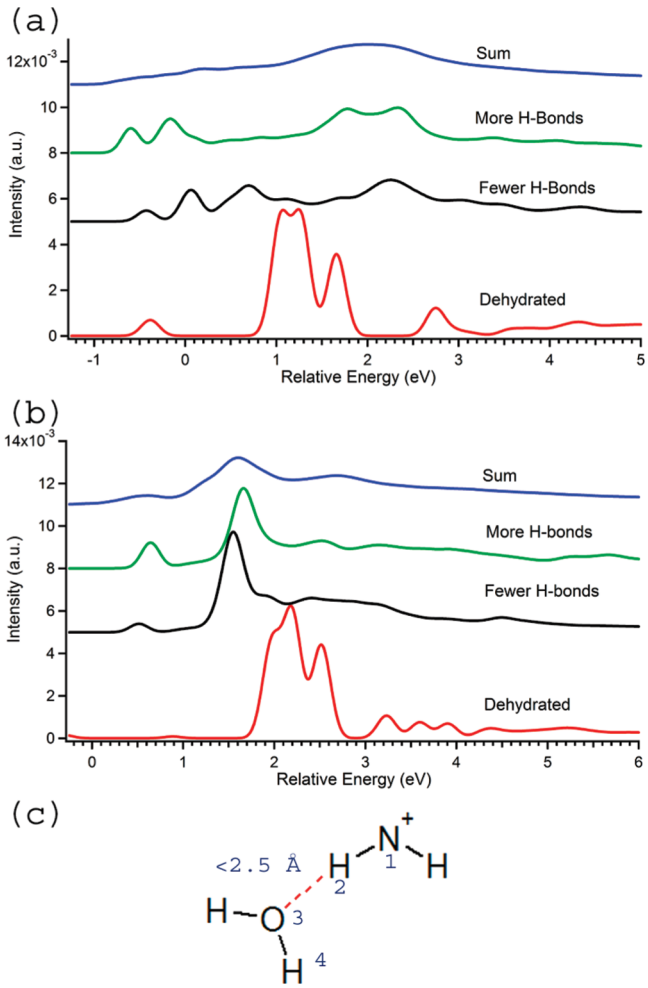

Figure 5. Using snapshots from the MD trajectory with a constrained solute, nitrogen K-edge spectra were calculated for cases where there is a higher or lower number of hydrogen bonds for (a) alanine, and (b) sarcosine. The sum is the hydrated-fixed average also shown in Figure 3. A hydrogen bond, shown in (c) was defined as having a distance between acceptor and donor of less than 2.5 Å, a donor angle greater than 120° (∠1,2,3), and an acceptor angle greater than 90° (∠2,3,4).

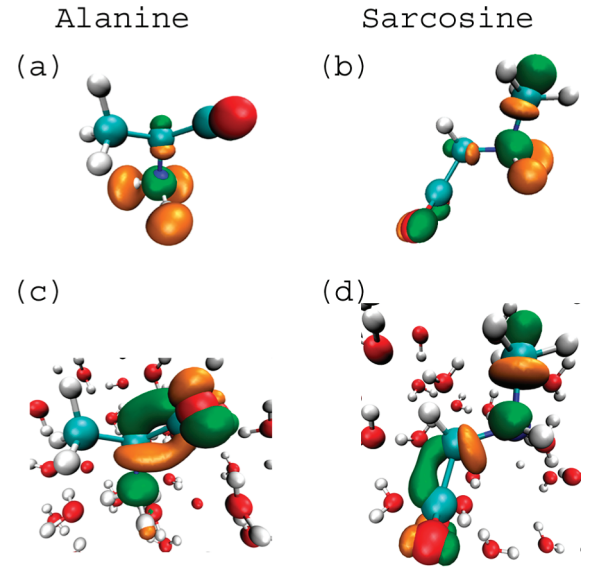

Figure 6. LUMOs from a single snapshot for (a) alanine without water, (b) sarcosine without water, (c) alanine with water, and (d) sarcosine with water. For the lower-lying unoccupied orbitals, sarcosine shows more molecular character than do the states in alanine and are therefore less affected by hydration. These states have isosurface values of 30% of the total integrated value.

sarcosine hydrated-fixed spectrum. The two snapshots for alanine and sarcosine are representative and indicative of the larger spectral impact of hydration on alanine than sarcosine and the importance of the strength or type of interaction. This suggests if sarcosine were placed in another polar solvent, the states would be less impacted than for alanine. This has been seen in previous studies of the effect of other polar solvents on the secondary structure of peptoid oligomers using circular dichrosim.[5]

Hydration has a large spectral impact, particularly for alanine. Certain excited states are more sensitive to hydration; the more sensitive states appear to be those that have more density on the hydrogens of the nitrogen. For sarcosine, most of the low-lying excited states are molecular in nature (dispersed over the entire molecule) with relatively more density along the backbone of the molecule as opposed to alanine where low-lying excited are more localized on the terminal NH_3^+ group. This can be seen in Figure 6, which contrasts the lowest unoccupied molecular orbitals (LUMOs) of alanine and sarcosine with and without water. These LUMOs are representative of the excited states as our XCH calculations explicitly place the excited electron in the LUMO in addition to the core hole. The LUMO of alanine without water, Figure 6a, is extremely localized on the amine group; whereas, sarcosine's LUMO without water, Figure 6b, has localized density on the nitrogen but also density along the methyl group and the carboxylate group. Both dehydrated LUMOs alter upon hydration; a large change occurs in the hydrogens upon hydrogen bonding, there is no density on the involved hydrogen. The hydrated LUMO of alanine, Figure 6c, shows a large migration to the backbone of the molecule with increased density near the carboxylate group. Alanine's LUMOs exhibit a density difference upon hydration particularly along the backbone; this can be attributed to the localization of the dehydrated LUMO on the dangling nitrogen group. For sarcosine, the change to the LUMO upon hydration is less than alanine as the dehydrated LUMO has character that is more molecular than alanine. Molecular states change less upon hydration, as the hydrogen bonding tends to affect specific terminal moieties in the molecule as opposed to the molecule in its entirety. Yet, a more molecular state will be impacted more by dihedral angle variations than a more localized state, such as what was found for sarcosine and alanine. Overall, the low-lying unoccupied states of sarcosine are more molecular than alanine; thus, giving sarcosine less sensitivity to the hydrogen bonding network in its entirety, but increased sensitivity to conformation.

Partial versus Bulk Hydration. Hydration can be a vague term, partial or bulk hydration? Are the waters around the atom of interest the only ones important, and how many solvation shells are important? We examine the question of using microhydration (a limited but localized number of water molecules) in simulations to approximate the bulk. Similarly, can we use bulk calculations or experiments to approximate real microhydration situations such as confined water? These concerns need to be addressed to approximate experimental conditions for NEXAFS, maintain a tractable calculation, and to move toward predictive computational work.

Clusters of various sizes were constructed for alanine and sarcosine using the nitrogen atom as the center and selecting waters by minimum distance to a specified center, the spectra of these clusters were calculated using DFT and are shown in

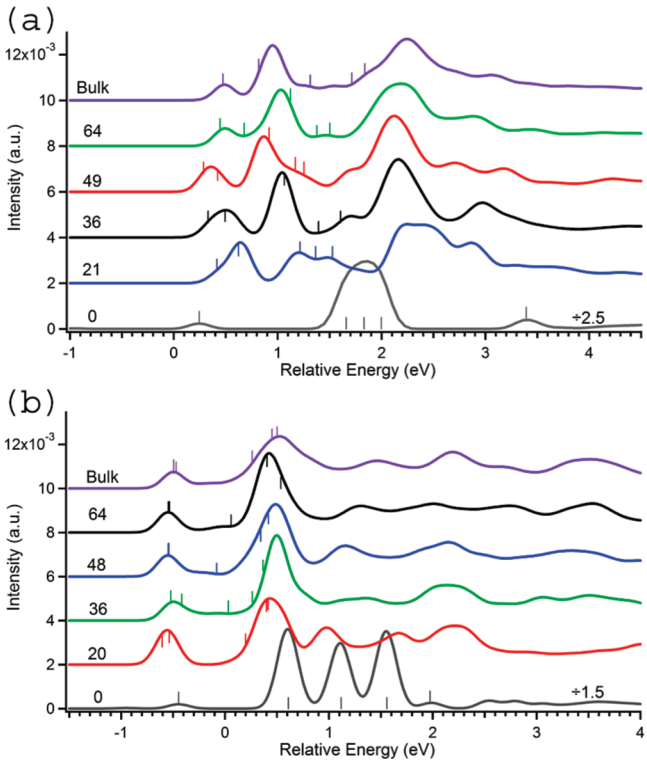

Figure 7. Computed N K-edge spectra for various aqueous cluster sizes solvating (a) alanine and (b) sarcosine with the focal point of the cluster being the nitrogen position. The number of waters for each cluster are labeled with 78 filling the periodic box (bulk). Spectra were aligned to a reference snapshot as described in the test. The vertical dashes identify the first several states.

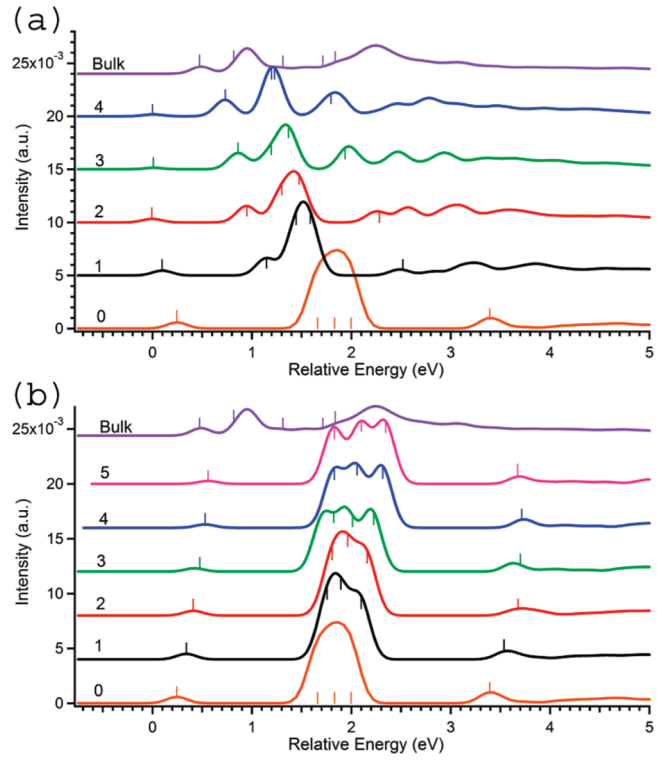

Figure 8. Nitrogen K-edge spectra of (a) NH_3^+ and (b) COO^- centered clusters of water molecules around alanine, examining the effects of microhydration. The nitrogen-centered clusters show differences in the appearance of new spectral features and large-scale shifts of features, indicated by the vertical dashes. In the case of the carboxylate-centered clusters, no new large peaks appear and there is more splitting of the main feature at approximately 2 eV. Bulk refers to 78 waters, the fully hydrated case, where the periodic box is full.

Figure 7; alanine is discussed below and sarcosine is discussed later in the text.

Inspecting the spectra of the clusters for alanine, Figure 7a, there is a marked difference between the spectra with 21 waters versus the other cluster sizes. In that configuration, there are only enough waters molecules to hydrate the amine group; thus, the carboxylate group is not hydrated until 36 waters are present. Thus, attempting to examine bulk hydration simulations with incomplete solvation may lead to incorrect interpretation.[56]

To examine the stability of our calculations and the effects of specific interactions on the nitrogen K-edge, smaller clusters of alanine were constructed, and their spectra calculated. The water clusters are centered on the NH_3^+ or around the carboxylate group in Figure 8a,b, respectively. The waters included were those closest to the functional group in question. As expected, the spectra of the NH_3^+ centered cluster exhibit definite spectral changes upon hydration. There is an increase in the number of states, changes in the intensity and energies of transitions, and in the density of states with addition of each water molecule in the NH_3^+ centered small cluster. The character (spatial extent and relative density on a particular atom) of the states change with the increased interaction; this in turn affects the oscillator strength and thus the spectrum. This can be seen in Figure 6 in the case of the LUMO; even small alterations to these final states can strongly affect their overlap with the initial electronic core state and thus alter the transition strength. The addition of waters to the COO^- centered water cluster impacts the energy spacing of the states, removing degeneracy and lowering the energy of excited states of this charged species, as opposed to the creation of new spectral features. In the dehydrated spectrum, the largest feature (at approximately 1.8 eV) comprises three smaller transitions. After hydrating the carboxylate, this approximate degeneracy is broken and the three smaller features become distinct. Hydration of the carboxylate group leads to a decrease in perturbations of the electronic structure, particularly in the case of a small zwitterion; this is notable by a slight lowering in the energy of the states. Whether this effect is important for polypeptides, where charge groups may be more spatially separated is the subject of future research. In cases of microsolvation (solvation by a limited number of waters) or clusters, this effect would be important if other nearby charged groups were unsolvated, such as in proteins or other polymers exhibiting preferential hydration. We recognize that while illustrative the small cluster studies presented here are unphysical, since the zwitterion is unstable in vacuo or with limited waters; research has suggested that at least 7−8 waters are required to stabilize the alanine zwitterion.[16,17]

Hydrated sarcosine was similarly examined using various water cluster sizes; their spectra are shown in Figure 7b. These clusters show similarities to those of alanine, Figure 7a, with regard to the spacing of spectral features. The spectral change produced when the cluster size is increased to 36 waters is less evident for sarcosine than for alanine; this is likely due to the nearest waters around sarcosine being closer to the carboxylate group than for alanine due to the difference in molecular structures. To isolate the spectral impact of water on sarcosine further, smaller clusters were constructed (Figure 9) for both nitrogen and carboxylate centered clusters. A reasonable approximation to the bulk-hydrated spectrum is reached by having 4 or 5 waters in the nitrogen-centered cluster for sarcosine, implying that the impact of hydration on the nitrogen K-edge spectrum of sarcosine converges within the first solvation shell,

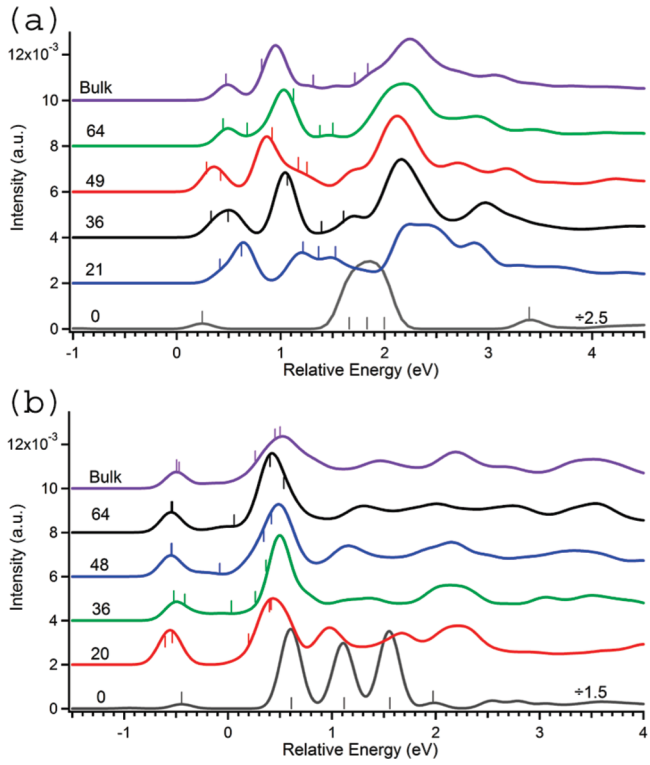

Figure 9. Nitrogen K-edge spectra of small clusters around sarcosine with (a) focused around the nitrogen and (b) centered around the carboxylate. The first few unoccupied states are labeled on each trace by small vertical lines. As found for alanine, the nitrogen-based clusters exhibit the appearance of new spectral features with increased numbers of water molecules, whereas the carboxylate-based clusters exhibit changes in the energy of the states, rather than the appearance of new transitions. Seventy-eight waters (bulk) is the fully solvated case, where the periodic box is full.

as opposed to the case of alanine. The shifting of states' energies and the appearance of new transitions occurs with the addition of each water for the small nitrogen-centered clusters' spectra, similar to what was found for alanine; this is particularly visible for the LUMO and LUMO+1. There are intensity changes of the features but the number of states within the first 1.5 eV of features remains constant after the spectrum with 3 waters for the nitrogen-based small clusters, corroborating a higher dependence on short-range interactions than was found for alanine. In the case of hydration around the carboxylate, there appears to be a contraction of the existing states, as well as small energy shifts, rather than the appearance of new states, suggesting that carboxylate-based hydration only modifies the energy levels of the states. For both sarcosine and alanine, it is important to account for the fact that there are two charged groups nearby and this will primarily affect the states and not their spectral intensities. This implies that groups that have only one such charged group would not yield good representations of those with two, and vice versa.

Conclusions

We present the first NEXAFS measurements of aqueous sarcosine and alanine and accompanying first principles DFT calculations and analysis; calculations are in good agreement with experiment, further validating our recently developed theoretical (XCH) approach. On the basis of our theoretical analysis, we show that the effect of hydration of alanine has a larger spectral effect than does conformation; the reverse is true for sarcosine. This was further supported by examining spectra of alanine and sarcosine conformers over a range of dihedral angles and with different hydrogen bonding environments. Part of alanine's higher sensitivity to hydrogen bonding is likely related to the structural difference of having a terminal nitrogen versus an embedded nitrogen, as in the case of sarcosine. It is important to note that some of the findings revealed in this study will change upon the formation of a polypeptide or polypeptoid due to the altered bonding on the nitrogens.[29] Presently, we can only predict what the spectral implications of secondary structure would be. In the case of alanine, the restrictions of motion effected by a longer peptide chain would likely affect the excited states, and thus the spectrum, less than if the solvent was changed or mixed, as has been found previously for alanine in different environments.[15,54] In general, side chains that have dangling hydrogen bonds would probably be more affected by hydration than by conformation, and this would be reflected in computed and likely experimental NEXAFS spectra. This is supported by many of the excited states of sarcosine, a molecule with a nonterminal nitrogen, resembling delocalized molecular orbitals more than moiety-specific orbitals, and therefore being less explicitly dependent on the solvent and more dependent on conformation. For polypeptoids, the existence of a longer chain may constrain the conformational movements due to the establishment of secondary structure in the absence of amino-group hydrogen-bond donors along the backbone, thus leading to more rigid higher-order structure than found in polypeptide analogues. We predict that this rigidity should exhibit a change in spectral signature given the increased population of certain dihedral angles. In addition, this work also demonstrates the sensitivity of nitrogen K-edge NEXAFS to details of the solvation environment, including degree of hydration, and solvent identity. Thus, NEXAFS spectroscopy, in combination with accurate excited state theory, appears to be a useful tool for the examination of the secondary structure of polypeptoids, polypeptides, and quite probably DNA and RNA.

Acknowledgment. We thank Wanli Yang at beamline 8.0.1 for his help with the experimental results. We also thank Ronald Zuckermann at the Molecular Foundry for suggesting candidate molecules. This work was done under the auspices of the Chemical Sciences Division, Office of Basic Energy Sciences, the Molecular Foundry, Lawrence Berkeley National Laboratory, supported by the Office of Science, and computational resources were provided by the National Energy Resource Center, Office of Science; all of which are supported by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Thanks to the Advanced Light Source (ALS) Fellowship fund and the ALS itself, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

References and Notes

(1) Bautista, A. D.; Craig, C. J.; Harker, E. A.; Schepartz, A. *Curr. Opin. Chem. Biol.* **2007**, *11* (6), 685–692.

(2) Yoo, B.; Kirshenbaum, K. *Curr. Opin. Chem. Biol.* **2008**, *12* (6), 714–721.

(3) Wu, C. W.; Kirshenbaum, K.; Sanborn, T. J.; Patch, J. A.; Huang, K.; Dill, K. A.; Zuckermann, R. N.; Barron, A. E. *J. Am. Chem. Soc.* **2003**, *125* (44), 13525–13530.

(4) Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Bioorg. Med. Chem. Lett.* **1994**, *4* (22), 2657–2662.

(5) Sanborn, T. J.; Wu, C. W.; Zuckerman, R. N.; Barron, A. E. *Biopolymers* **2002**, *63* (1), 12–20.

(6) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (24), 13003–13008.

(7) Hara, T.; Durell, S. R.; Myers, M. C.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128* (6), 1995–2004.

(8) Seurynck, S. L.; Patch, J. A.; Barron, A. E. *Chem. Biol.* **2005**, *12* (1), 77–88.

(9) Patch, J. A.; Barron, A. E. *J. Am. Chem. Soc.* **2003**, *125* (40), 12092–12093.

(10) Lee, B. C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N. *J. Am. Chem. Soc.* **2008**, *130* (27), 8847–8855.

(11) Gorske, B. C.; Blackwell, H. E. *J. Am. Chem. Soc.* **2006**, *128* (44), 14378–14387.

(12) Gomez-Zavaglia, A.; Fausto, R. *Vib. Spectrosc.* **2003**, *33* (1−2), 105–126.

(13) Shin, S. B. Y.; Kirshenbaum, K. *Org. Lett.* **2007**, *9* (24), 5003–5006.

(14) Lemoff, A. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (7), 1014–1024.

(15) Kohtani, M.; Jones, T. C.; Schneider, J. E.; Jarrold, M. F. *J. Am. Chem. Soc.* **2004**, *126* (24), 7420–7421.

(16) Aikens, C. M.; Gordon, M. S. *J. Am. Chem. Soc.* **2006**, *128* (39), 12835–12850.

(17) Ahn, D. S.; Park, S. W.; Jeon, I. S.; Lee, M. K.; Kim, N. H.; Han, Y. H.; Lee, S. *J. Phys. Chem. B* **2003**, *107* (50), 14109–14118.

(18) Uejio, J. S.; Schwartz, C. P.; Duffin, A. M.; Drisdell, W. S.; Cohen, R. C.; Saykally, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (19), 6809–6812.

(19) Messer, B. M.; Cappa, C. D.; Smith, J. D.; Drisdell, W. S.; Schwartz, C. P.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2005**, *109* (46), 21640–21646.

(20) Messer, B. M.; Cappa, C. D.; Smith, J. D.; Wilson, K. R.; Gilles, M. K.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2005**, *109* (11), 5375–5382.

(21) Ide, M.; Maeda, Y.; Kitano, H. *J. Phys. Chem. B* **1997**, *101* (35), 7022–7026.

(22) Gordon, M. L.; Cooper, G.; Morin, C.; Araki, T.; Turci, C. C.; Kaznatcheev, K.; Hitchcock, A. P. *J. Phys. Chem. A* **2003**, *107* (32), 6144–6159.

(23) Aziz, E. F.; Ottosson, N.; Bonhommeau, S.; Bergmann, N.; Eberhardt, W.; Chergui, M. *Phys. Rev. Lett.* **2009**, *102* (6), 068103.

(24) Kaznacheyev, K.; Osanna, A.; Jacobsen, C.; Plashkevych, O.; Vahtras, O.; Agren, H. *J. Phys. Chem. A* **2002**, *106* (13), 3153–3168.

(25) Boese, J.; Osanna, A.; Jacobsen, C.; Kirz, J. *J. Electron Spectrosc. Relat. Phenom.* **1997**, *85* (1−2), 9–15.

(26) Carravetta, V.; Plashkevych, O.; Agren, H. *J. Chem. Phys.* **1998**, *109* (4), 1456–1464.

(27) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Prince, K. C.; Carravetta, V. *J. Phys. Chem. A* **2008**, *112* (34), 7806–7815.

(28) Nyberg, M.; Hasselstrom, J.; Karis, O.; Wassdahl, N.; Weinelt, M.; Nilsson, A.; Pettersson, L. G. M. *J. Chem. Phys.* **2000**, *112* (12), 5420–5427.

(29) Otero, E.; Urquhart, S. G. *J. Phys. Chem. A* **2006**, *110* (44), 12121–12128.

(30) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Carravetta, V. *J. Electron Spectrosc. Relat. Phenom.* **2007**, *155* (1−3), 47–53.

(31) Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *J. Phys. Chem. B* **2008**, *112* (15), 4478–4480.

(32) Wilson, K. R.; Rude, B. S.; Catalano, T.; Schaller, R. D.; Tobin, J. G.; Co, D. T.; Saykally, R. J. *J. Phys. Chem. B* **2001**, *105* (17), 3346–3349.

(33) Wilson, K. R.; Rude, B. S.; Smith, J.; Cappa, C.; Co, D. T.; Schaller, R. D.; Larsson, M.; Catalano, T.; Saykally, R. J. *Rev. Sci. Instrum.* **2004**, *75* (3), 725–736.

(34) Schwartz, C. P.; Uejio, J. S.; Saykally, R. J.; Prendergast, D. *J. Chem. Phys.* **2009**, *130* (18), 184109.

(35) Uejio, J. S.; Schwartz, C. P.; Saykally, R. J.; Prendergast, D. *Chem. Phys. Lett.* **2008**, *467* (1−3), 195–199.

(36) Prendergast, D.; Galli, G. *Phys. Rev. Lett.* **2006**, *96* (21), 21522–21526.

(37) Prendergast, D.; Grossman, J. C.; Galli, G. *J. Chem. Phys.* **2005**, *123* (1), 14501–14505.

(38) Cappa, C. D.; Smith, J. D.; Wilson, K. R.; Messer, B. M.; Gilles, M. K.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2005**, *109* (15), 7046–7052.

(39) Cappa, C. D.; Smith, J. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2006**, *110* (11), 5301–5309.

(40) Cappa, C. D.; Smith, J. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2006**, *110* (3), 1166–1171.

(41) Cavalleri, M.; Odelius, M.; Nilsson, A.; Pettersson, L. G. M. *J. Chem. Phys.* **2004**, *121* (20), 10065–10075.

(42) Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Naslund, L. A.; Hirsch, T. K.; Ojamae, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304* (5673), 995–999.

(43) Stöhr, J. *NEXAFS Spectroscopy*; Springer: New York, 1996; pp 20−31.

(44) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136* (3B), B864.

(45) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140* (4A), 1133–1138.

(46) Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247–260.

(47) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(49) Warshel, A.; Kato, M.; Pisliakov, A. V. *J. Chem. Theory Comput.* **2007**, *3* (6), 2034–2045.

(50) Johnson, M. E.; Head-Gordon, T.; Louis, A. A. *J. Chem. Phys.* **2007**, *126*, (14), 144509−144509-10.

(51) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(52) Baroni, S.; Corso, A. D.; de Gironcoli, S.; Gianozzi, P. PWSCF. www.pwscf.org.

(53) Urquhart, S. G.; Ade, H. *J. Phys. Chem. B* **2002**, *106* (34), 8531–8538.

(54) Lee, M.-E.; Lee, S. Y.; Joo, S.-W.; Cho, K.-H. *J. Phys. Chem. B* **2009**, *113* (19), 6894–6897.

(55) Cocinero, E. J.; Villanueva, P.; Lesarri, A.; Sanz, M. E.; Blanco, S.; Mata, S.; Lopez, J. C.; Alonso, J. L. *Chem. Phys. Lett.* **2007**, *435* (4−6), 336–341.

(56) Notably, we are using plane waves in these calculations, as opposed to the localized atomic orbitals used in previous work, which means we can better describe delocalized final states and avoid unintentional localization based on basis set choice.

JP911007K