Localized Intrinsic Valence Virtual Orbitals as a Tool for the Automatic Classification of Core Excited States

"We must be clear that when it comes to atoms, language can be used only as in poetry. The poet, too, is not nearly so concerned with describing facts as with creating images and establishing mental connections" Niels H. D. Bohr Accurate assignments of the unoccupied molecular orbitals involved in electronic excited states are crucial to the interpretation of experimental spectra. Here we present an automated approach to the assignment of excited states by introducing a unique orbital basis known as localized intrinsic valence virtual orbitals (LIVVOs), which are a special case of the previously reported valence virtual orbitals. The LIVVOs are used to quantify the local contributions to particle orbitals from orthogonality-constrained density functional theory, providing an assignment with atomic-level/angular momentum shell specificity. This localized set also allows us to define the total valence character of an excited state. We highlight the utility of our approach by studying the local orbital changes in core-excited states at the sulfur K-edge of ethanethiol and benzenethiol as well as hydrogen bonding in the water dimer.

1 Introduction

Any interpretation of electronic spectroscopy depends upon assigning spectral features of the experimental spectrum to discrete electronic transitions calculated using excited state quantum chemistry methods. This requires characterization of the molecular orbitals involved in the discrete transition. Transitions are usually assigned by visual inspection of the contributing virtual molecular orbitals (MOs). Often, when these MOs do not have well defined character, orbital labels are used instead (e.g. LUMO, LUMO + 1, etc.). These approaches are somewhat unsatisfactory. The assignment of orbital character based on MO plots is potentially ambiguous. Furthermore, characterizing a transition using orbital labels is problematic because virtual orbitals are not uniquely defined and they strongly depend on the basis set and level of theory employed.

A unique branch of electronic spectroscopy involving the excitation of core electrons using X-ray photons called near-edge X-ray absorption fine structure (NEXAFS) spectroscopy has garnered particular interest due to advancements in high-resolution tunable synchrotron light sources. NEXAFS spectroscopy has become an indispensable experimental tool for probing local electronic and geometrical information in a wide variety of molecular systems.[1–4] The significant features of the NEXAFS spectrum arise from core orbital to unoccupied valence orbital transitions.[5] The atomic nature of the core orbitals involved in these transitions (1s, 2s, 2p, etc.) makes NEXAFS a spectroscopic technique for elucidating local electronic structure information. Quantum chemistry calculations serve as a vital tool for gleaning meaningful electronic structure information from NEXAFS spectra including the nature of unoccupied orbitals, orbital mixing effects, and covalent character.[2,5–8] A myriad of theoretical methods have been developed for the computation of core-excited states, ranging from rigorous methods focused on high accuracy[9–24] to computationally efficient methods focused on applications to larger systems.[25–34] The primary motivation underpinning these theoretical developments has been to obtain more accurate excitation energies and/or oscillator strengths. In contrast, the character of the wave function is analyzed in a more cursory fashion.

A recent work by the group of Dreuw[35–37] has taken a significant step toward a more rigorous description of core-excited states. Using the one-particle transition density matrix and the one-particle difference density matrix, these authors were able to quantify the effects of relaxation, estimate multireference character, and provide information on the exciton size. These methods are a significant step toward a robust characterization of excited-states. Myhre, Coriani, and Koch[38] also recently used visualization of the one-particle difference density matrix in order to track the localization of core-excited states, relying on analysis of the excitation amplitudes of coupled cluster response calculations to assign the character of core-valence transitions. Although a fair amount of attention has been put forth to add a breadth of different descriptors for excited states, little attention has been put toward trying to improve upon the method of assigning orbital character.

Canonical molecular orbitals often complicate the assignment process because valence virtual MOs are mixed with polarization and diffuse functions included in the basis set. A common approach to obtain more chemically meaningful valence orbitals is to project the virtual space onto a set of atomic valence orbitals. This approach hinges on the assumption that valence atomic orbitals are the dominant contributors to the valence molecular orbitals. This assumption was justified in the 1980s by Rudenberg and coworkers by obtaining overlaps close to unity between free-atom AOs and their projection onto the molecular orbitals of MCSCF wave functions.[39] Perhaps, the most popular method to analyze wave functions in terms of atomic valence contributions is Weinhold's natural atomic/bond orbital (NAO/NBO) analysis.[40,41] This method partitions the full density matrix into atomic sub-blocks that are diagonalized in order to produce atom-localized hybrid valence orbitals. AO projection based methods have been presented throughout the literature including polarized atomic orbitals,[42,43] enveloping localized orbitals,[44] intrinsic minimal atomic basis

orbitals,[45] molecule-adapted atomic orbitals,[46] and intrinsic atomic orbitals.[47] The approach presented in this work is based on the valence virtual orbitals (VVOs) scheme.[48] VVOs are built by bringing virtual valence MOs into maximum coincidence with a minimal atomic basis set via a unitary transformation. The VVOs have been used with success to identify the valence LUMO in a variety of molecular systems,[49] to characterize bond breaking/formation along dissociation pathways,[50] and as a set of starting orbitals for multi-configurational SCF calculations.[51] VVOs can be easily localized and provide an accurate representation of the localized valence antibonding orbitals in a molecule.

In this work we report a new method for automated classification of virtual orbitals in electronic excited states by quantifying the local contributions using a localized set of VVOs. To generate the VVOs an accurate atomic minimal basis set (AAMBS) that spans the atomic valence space is required. We propose to represent the AAMBS using a minimal basis set of quasiatomic orbitals.[48] These are generated using projection operators according to Knizia's intrinsic atomic orbitals (IAOs) procedure[47] and localized using the Pipek–Mezey localization scheme.[52] The combination of VVOs projected onto a localized intrinsic orbital basis is abbreviated as LIVVOs. To compute excitation energies we employ our recently proposed orthogonality-constrained DFT (OCDFT) approach[53] to compute a manifold of states.[54] OCDFT is a variational time-independent DFT method built upon a generalized Kohn-Sham scheme. In OCDFT excited state energies are approximated by a Kohn–Sham procedure augmented with orthogonality constraints between ground and excited state wave functions.

This new classification method is applied to the core-excited states involved in NEXAFS spectroscopy. We utilize the quantification of local contributions to analyze substituent effects in thiols (ethanethiol, benzenethiol) and intermolecular interactions in the water dimer. The thiols were chosen as an example of molecules where the relevant excited states have significant contributions from more than one valence antibonding MO. Previous studies have not fully characterized the antibonding character of orbitals involved in core-valence excitations.[55–57] We show that we are able to provide insight into the nature of these excited states using LIVVOs. Hydrogen bonding interactions in water have received plenty of attention within the X-ray absorption community.[58–64] Studies have shown that NEXAFS spectral changes upon the formation of a hydrogen bond is a result of localization of excited states along either the free OH bond or the hydrogen bond.[60,61,64–70] However, the degree of localization has not been quantified. In this work we show that LIVVOs help quantify the localization of core excited states of the water dimer and that this metric is correlated some some spectral features.

2   Theory

In this section we discuss three aspects of our approach to analyze core excitations. First, we consider the construction of the localized intrinsic valence virtual orbitals (LIVVOs) from canonical molecular orbitals (CMOs) and intrinsic atomic orbitals (IAOs). Next, we discuss how to automate the determination of the atomic character of the IAOs and the LIVVOs. Finally, we discuss the procedure used to analyze the particle orbital in OCDFT computations of core-excited states.

2.1   Construction of Localized Intrinsic Valence Virtual Orbitals

Here we summarize the construction of VVOs and discuss how to obtain the LIVVOs. The procedure outlined here is illustrated in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1 for the case of water. The construction of VVOs begins by first identifying the virtual subset of the canonical molecular orbitals ($\{\phi_a\}$) from a ground state SCF calculation. The virtual CMOs are expanded in an atom-centered basis $\{\chi_\mu\}$ and are expressed in terms of MO coefficients $C_{\mu a}$

$$|\phi_a\rangle = \sum_{\mu}^{N_{\mathrm{AO}}} |\chi_\mu\rangle C_{\mu a}, \quad a = 1, \ldots, N_{\mathrm{vir}} \tag{1}$$

where N_{AO} and N_{vir} are the number of atomic and virtual orbitals, respectively. As shown in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals

Figure 1: Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOs.

obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1A, the virtual CMOs may have contributions from diffuse functions and their valence character is difficult to interpret.

To identify the valence character of the CMOs, we perform a VVO analysis[49,71,72] using the intrinsic atomic orbitals (IAOs) of Knizia[47] as the underlying accurate atomic minimal basis set. IAOs are a set of orthonormal polarized atomic orbitals that can exactly express the occupied molecular orbitals of the ground state Kohn-Sham (KS) determinant $|\Phi\rangle$. IAOs are expressed in terms of the atomic orbital basis as

$$|\psi_\rho^{\mathrm{IAO}}\rangle = \sum_\mu^{N_{\mathrm{AO}}} |\chi_\mu\rangle \tilde{C}_{\mu\rho}, \quad \rho = 1, \ldots, N_{\mathrm{IAO}} \tag{2}$$

where $\tilde{C}_{\mu\rho}$ is the IAO coefficient matrix and N_{IAO} is the number of IAO orbitals. As shown in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecu-

3

lar orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1B, in the case of water there are 7 IAOs that correspond to the oxygen atom 1s, 2s, and 2p shells and a single 1s orbital for each hydrogen atom.

Next, the overlap matrix ($\mathbf{S}$) of the canonical virtual orbitals with the IAO functions ψ_ρ^{IAO} is evaluated:

$$(\mathbf{S})_{a\rho} = \langle \phi_a | \psi_\rho^{\mathrm{IAO}} \rangle \tag{3}$$

and as suggested in Ref. 49 a singular value decomposition (SVD) is performed on $\mathbf{S}$:

$$\mathbf{S} = \mathbf{U}\boldsymbol{\sigma}\mathbf{V}^\dagger \tag{4}$$

to yield orthogonal transformation matrices $\mathbf{U}$ and $\mathbf{V}$. These matrices are rotations of the virtual space and the IAO space, respectively, that bring the two sets of orbitals into maximum coincidence. The number of VVOs (N_{VVO}) is given by the number of IAOs minus the number of occupied molecular (N_{occ}) orbitals: $N_{\mathrm{VVO}} = N_{\mathrm{IAO}} - N_{\mathrm{occ}}$. Assuming that the singular value decomposition orders the singular values (σ_v) in descending order, then the VVOs are obtained by transforming the canonical virtual orbitals with the first N_{VVO} columns of the matrix $\mathbf{U}$:

$$|\psi_v^{\mathrm{VVO}}\rangle = \sum_a^{N_{\mathrm{vir}}} |\phi_a\rangle U_{av}, \quad v = 1, \ldots, N_{\mathrm{VVO}}. \tag{5}$$

The VVOs for water shown in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1C, have valence character and correspond to the plus and minus combination of localized $\sigma_{\mathrm{O-H}}^*$ orbitals. Following Pipek–Mezey localization,[52] we arrive at a set of localized VVOs that span the antibonding interactions of the molecular environment. We refer to this localized set as localized intrinsic valence virtual orbitals (LIVVOs). In the case of water, when the VVOs are localized we obtain the $\sigma_{\mathrm{O-H}}^*$ orbitals drawn in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1D.

2.2 Determination of the character of the IAOs and LIVVOs

The goal of our analysis is to express excited states using the LIVVO basis. In addition, we are also interested in determining the dominant atomic character of each excited state. To this end, we have automated the analysis of the atomic character of the LIVVOs. Our approach starts with the assignment of the atomic character of each IAO via a Mulliken population analysis.[73,74] For each IAO, ψ_ρ^{IAO}, we compute the population matrix:

$$P_{\mu\nu}(\rho) = \tilde{C}_{\mu\rho}S_{\mu\nu}\tilde{C}_{\nu\rho}, \quad \rho = 1, \ldots, N_{\mathrm{IAO}} \tag{6}$$

Gross populations on atom A and angular momentum shell l [$\mathrm{GP}_{A_l}(\rho)$] are obtained as partial sums of the population matrix:

$$\mathrm{GP}_{A_l}(\rho) = \sum_{\mu \in A_l} \sum_\nu^{N_{\mathrm{AO}}} P_{\mu\nu}(\rho) \tag{7}$$

where the sum over μ is restricted to all functions centered on atom A with angular momentum quantum number l.

Values of $GP_{A_l}(\rho)$ for the water example are reported in Table 4.1Atom-centered gross orbital populations [$GP_{A_l}(\rho)$] for the seven IAOs of water. Calculations were performed at the B3LYP/jun-cc-pVTZ level of theorytable.4.1. The assignment of IAOs using the gross population is straightforward since there is always a dominant contribution from a single atom/shell. For example, in the case of IAO ψ_6^{IAO}, the largest contribution to $GP_{A_l}(\rho)$ is from the H_s^2 shell (1.18), followed by smaller contributions from the O_s^1, O_s^1, and H_s^3 shells (with absolute value less or equal to 0.08). Therefore, we define the character of ψ_ρ^{IAO} as the pair atom/shell (A,l) for which $GP_{A_l}(\rho)$ is maximum:

$$\text{character}(\psi_\rho^{IAO}) = \arg\max_{(A,l)} GP_{A_l}(\rho) \tag{8}$$

In the case of water, when eq (4.8Determination of the character of the IAOs and LIVVOsequation.4.2.8) is applied to the IAOs we obtain the assignment reported in Figure 4.1Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3Construction of Localized Intrinsic Valence Virtual Orbitalsequation.4.2.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOsfigure.4.1B.

Once the atomic character of each IAO is determined, we characterize the atomic contributions to each LIVVO. For each LIVVO, ψ_l^{LIVVO}, we evaluate the overlap with all the IAOs (ψ_ρ^{IAO}):

$$S'_{l\rho} = |\langle \psi_l^{LIVVO} | \psi_\rho^{IAO} \rangle|^2 \tag{9}$$

The initial step in classifying the LIVVOs is to determine their overall orbital character $(\sigma, \pi, \dots)$. A LIVVO is classified as σ^* if the largest elements of $S'_{l\rho}$ include contributions from IAOs with s and p character. Similarly, a LIVVO is assigned π^* character if the largest elements of $S'_{l\rho}$ arise exclusively from p-type IAOs. The atoms involved in the LIVVO are identified by the character of the largest elements of $S'_{l\rho}$. For this purpose, we only include those IAOs whose overlap with a given LIVVO is greater than a threshold, which in the following examples is set to 0.1.

To illustrate this analysis we consider the case of acrolein (C_3H_4O). The following LIVVOs are characterized respectively as $\sigma^*_{C^2-C^3}$ and $\pi^*_{C^2-C^3}$:

$$= 0.20\,\psi^{IAO}(C_s^2) + 0.24\,\psi^{IAO}(C_p^2)$$
$$+ 0.20\,\psi^{IAO}(C_s^3) + 0.25\,\psi^{IAO}(C_p^3)$$

$$= 0.47\,\psi^{IAO}(C_p^2) + 0.51\,\psi^{IAO}(C_p^3)$$

Note that the in the case of the $\sigma^*_{C^2-C^3}$ LIVVO there are contributions from both s and p orbitals, while 98% of the $\pi^*_{C^2-C^3}$ LIVVO is comprised of p orbitals. The construction and analysis of the LIVVOs is fully automated and is performed only once after a ground state KS computation.

2.3 Analysis of OCDFT Particle Orbitals Using LIVVOs

The procedure used to generate the LIVVOs relies only the information contained in the reference ground state determinant and is therefore applicable to any method. The focus of this work is to apply the LIVVO to OCDFT computations of K-edge spectra. First, we briefly summarize the features of OCDFT. In OCDFT,[53] a generalized Kohn–Sham picture is assumed, where to the n-th electronic state corresponds an auxiliary system of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho^{(n)}$. The wave function for the auxiliary system is a single Slater

Table 1: Atom-centered gross orbital populations [$GP_{A_l}(\rho)$] for the seven IAOs of water. Calculations were performed at the B3LYP/jun-cc-pVTZ level of theory.

A_l	ψ_1^{IAO}	ψ_2^{IAO}	ψ_3^{IAO}	ψ_4^{IAO}	ψ_5^{IAO}	ψ_6^{IAO}	ψ_7^{IAO}
O_s^1	1.00	1.16	0.00	0.00	0.00	−0.08	−0.08
O_p^1	0.00	0.00	1.04	1.11	1.00	−0.07	−0.07
H_s^2	0.00	−0.08	−0.02	−0.05	0.00	1.18	−0.03
H_s^3	0.00	−0.08	−0.02	−0.05	0.00	−0.03	1.18

determinant, $\Phi^{(n)} = [1]\phi_1^{(n)}\phi_2^{(n)}\cdots\phi_N^{(n)}$, where the set of orbitals $\{\phi_i^{(n)}\}$ are different for each electronic state. To avoid variational collapse, OCDFT enforces orthogonality conditions between all determinants:

$$\langle\Phi^{(m)}|\Phi^{(n)}\rangle = \delta_{mn} \quad \forall m, n \tag{10}$$

In this work we employ our constrained multiple hole/particle (CMHP) algorithm[75] for the solution of multiple orthogonally constrained excited states. The CMHP scheme fully accounts for relaxation of all orbitals, but rather than solving for the exact minimum orthogonality conditions, it imposes sufficient orbital orthogonality conditions. Namely, we assume that the hole orbitals in core-excited states are localized and are orthogonal. It follows then that one-electron excited states can be characterized by fully relaxed hole $[\phi_{\rm h}^{(n)}]$ and particle $[\phi_{\rm p}^{(n)}]$ orbitals, which span the occupied and virtual spaces of the ground state wave function, respectively.

For every excited state, we analyze the character of the particle orbital $\phi_{\rm p}^{(n)}$ by evaluating its overlap with each LIVVO ($\psi_l^{\rm LIVVO}$):

$$\Omega_{{\rm p}l}^{(n)} = |\langle\phi_{\rm p}^{(n)}|\psi_l^{\rm LIVVO}\rangle|^2. \tag{11}$$

The individual overlaps $\Omega_{{\rm p}l}^{(n)}$ are used to assign the character of the particle orbital to the l-th LIVVO. In addition, we also define the total valence character $t_{\rm p}^{{\rm val},(n)}$ for any given particle orbital as the sum its overlap with all LIVVOs:

$$t_{\rm p}^{{\rm val},(n)} = \sum_{l}^{N_{\rm LIVVO}} \Omega_{{\rm p}l}^{(n)}. \tag{12}$$

Since the LIVVOs give an accurate representation of valence orbitals in the molecule, it is possible to quantify the most important local contributions to the particle orbital. While this offers a very robust description for localized valence particle orbitals, it is important to note that the particle orbital describing the excited state can also be fairly diffuse with respect to the molecular environment. Such an orbital will obviously have relatively low overlap with the LIVVOs described here. This can be a desirable feature of our method as it allows for immediate identification of spectral contributions that arise from more diffuse orbitals. However, any further classification of the diffuse nature or Rydberg character of virtual orbitals is outside the scope of the current study.

3   Computational Details

OCDFT is currently implemented as a plugin for the PSI4[76] *ab initio* quantum chemistry package. All core excitation energies and oscillator strengths are obtained using the B3LYP[77–80] functional and a triple zeta correlation-consistent basis set jun-cc-pVTZ,[81,82] which is obtained by deleting diffuse functions from the highest subshell of every atom in the aug-cc-pVTZ basis set.[83] The two-electron integrals are factorized using the density-fitting approximation[84–88] with the jun-cc-pVTZ-JKFIT auxillary fitting basis.[89,90] Basis set dependence of the LIVVO analysis is investigated using the cc-pVXZ and aug-cc-pVXZ (X=D,T,Q) family of basis sets.[82,91] Intrinsic atomic orbitals were obtained using Robert Parrish's implementation in PSI4.[92] The geometries for ethanethiol, benzenethiol, water monomer, and water dimer were all optimized at the B3LYP/aug-cc-pVTZ level of theory.

Peak intensities for each transition from the ground state to excited state (n) are based on the oscillator strength $(f_{\rm osc})$:

$$f_{\rm osc} = \frac{2}{3}|\boldsymbol{\mu}_{n0}|^2\omega_n \tag{13}$$

where ω_n is the excitation energy calculated in OCDFT and $\boldsymbol{\mu}_{n0}$ is the transition dipole moment vector. OCDFT does not provide a direct route toward computation of the transition dipole moments, however, these can be approximated using the noninteracting Kohn–Sham determinants:

$$\boldsymbol{\mu}_{n0} = \langle\Phi^{(n)}|\mathbf{r}|\Phi^{(0)}\rangle \tag{14}$$

where $\Phi^{(n)}$ is the KS determinant for the n-th excited state and $\mathbf{r}$ is the position vector. Spectroscopic broadening effects are simulated by representing each transition using a Gaussian lineshape with a FWHM of 0.4 eV.

Table 2: Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57.

| | | OCDFT | | Experiment | |
State	Energy (eV)	f_{osc}	LIVVO Assignment	Energy (eV)	Assignment
			Ethanethiol		
1	2472.9	0.0022	58% $\sigma^*_{\text{S}3-\text{H}9}$	2472.2	$\sigma^*_{\text{S}-\text{H}}$
			22% $\sigma^*_{\text{S}3-\text{C}2}$		Weak $\sigma^*_{\text{S}-\text{C}}$
2	2473.9	0.0011	57% $\sigma^*_{\text{S}3-\text{C}2}$	2473.1	$\sigma^*_{\text{S}-\text{C}}$
			18% $\sigma^*_{\text{S}3-\text{H}9}$		
3	2475.6	0.0004	9% $\sigma^*_{\text{C}2-\text{H}8}$	2474.9	Poorly Defined
			7% $\sigma^*_{\text{C}2-\text{H}7}$		
4	2476.2	0.0007	8% $\sigma^*_{\text{S}3-\text{C}2}$		
			5% $\sigma^*_{\text{C}2-\text{H}8}$		
5	2476.1	0.0005	6% $\sigma^*_{\text{C}1-\text{H}6}$		
			3% $\sigma^*_{\text{C}2-\text{H}7}$		
			Benzenethiol		
1	2472.9	0.0023	61% $\sigma^*_{\text{S}7-\text{H}8}$	2472.4	$\sigma^*_{\text{S}-\text{H}}$
			22% $\sigma^*_{\text{S}7-\text{C}3}$		Weak $\sigma^*_{\text{S}-\text{C}}$
2	2474.2	0.0011	53% $\sigma^*_{\text{S}7-\text{C}3}$	2473.5	Weak $\pi^*_{C=C}$
			16% $\sigma^*_{\text{S}7-\text{H}8}$		
3	2474.9	0.0001	68% $\pi^*_{\text{C}2,3,4}$	2475.3	σ^*_{S-C}
			15% $\pi^*_{\text{C}4,5,6}$		
4	2475.3	0.0000	48% $\pi^*_{\text{C}1,2,6}$		
			34% $\pi^*_{\text{C}4,5,6}$		
5	2476.0	0.0006	7% $\sigma^*_{\text{S}7-\text{H}8}$		

4 Results

4.1 Analysis of substituent effects in the spectra of thiols

Our first application of the LIVVO analysis for the classification of NEXAFS spectral features focuses on the effect of substituents on peak position and intensity. In this section we compare the sulfur K-edge of ethanethiol (C_2H_5SH) and benzenethiol (C_6H_5SH). The OCDFT NEXAFS spectra for these two compounds are shown in Figure 4.2Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.[93]figure.4.2. Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2 reports a comparison of experimental and theoretical excitation energies, OCDFT oscillator strengths, the LIVVO assignments, and total valence character for each state.
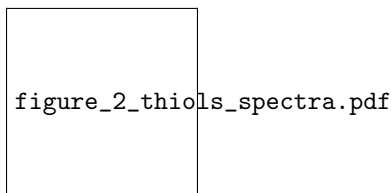

Figure 2: Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.[93]

The gas-phase experimental excitation energies shown in Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2 for ethanethiol represents the three peaks of the spectrum at 2472.17, 2473.11, and 2474.87 eV. Our calculated OCDFT spectrum is also characterized by three peaks with transition energies that are in good agreement with the experimental spectrum. The first two are a result of strong single transitions at 2472.9 and 2473.9 eV. The final peak is a combination of three weak transitions that occur at 2475.6, 2476.2, and 2476.1 eV.

The gas phase experimental spectrum of benzenethiol is characterized by three peaks at 2472.4, 2473.5, and 2475.3 eV. The theoretical spectrum has strong single transitions at 2472.9 eV and 2474.2 eV. The last peak is composed of weaker transitions at 2474.9, 2475.3, and 2476.0 eV. In total, this gives us errors that are on average less than 1.0 eV across these two molecules. This agreement with the experimental spectrum is consistent with that seen in previous studies with OCDFT[54,94] and are obtained without application of shifts to match experimental values.

By analyzing the overlap of the particle orbitals with all orbitals in the LIVVO basis, it is possible to quantify the degree of localization along each bond and discern the origin of some features of the ethanethiol and benzenethiol spectra. The lowest energy transition in the NEXAFS spectrum of ethanethiol and benzenethiol occurs at at 2472.9 eV, and was attributed mostly to a S 1s $\rightarrow \sigma^*_{\text{S-H}}$ transition with a "weak" contribution from a S $1s \rightarrow \sigma^*_{\text{S-C}}$ transition.[56] The LIVVO assignments, displayed in Figure 4.3Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) ethanethiol (C_2H_5SH) and (B) benzenethiol (C_6H_5SH). Percentages shown represent the overlap of each particle orbital with the LIVVOs. Particle orbitals are numbered according to the calculated transitions reported in Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2 while atom numberings correspond to those shown in Figure 4.2Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.[93]figure.4.2figure.4.3, show that in both cases the first state (corresponding to the particle orbital ϕ^1_{p}) can be assigned to $\sigma^*_{\text{S-H}}$ (50%) and $\sigma^*_{\text{S-C}}$ (22%) character. Going from ethanethiol to benzenethiol there is a slight increase in the degree of localization along the thiol bond. This is evidenced by the $\sigma^*_{\text{S-H}}$ character of the first excited state increasing from 58% in ethanethiol, to 61% in benzenethiol.

The second peak in the NEXAFS spectra of the thiols occurs at 2473.9 eV for ethanethiol and at 2474.2 eV for benzenethiol. Again in this case, simple inspection of the particle orbitals (ϕ^2_{p}) shows orbital character along both the S–H and S–C bonds. The LIVVO assignments show that over 50% of the character of the particle orbital can be assigned to the $\sigma^*_{\text{S-C}}$ orbital in each molecule, while less than 19% can be assigned to the $\sigma^*_{\text{S-H}}$ orbital. Comparing our results with a previous study[56] that employed the Hartree–Fock static exchange (HF-STEX) method,[95–100] our assignments agrees well for the case of ethanethiol. However, the OCDFT assignment is at odds with their assignment for benzenethiol where they attributed this peak feature to core transitions to π^* orbitals associated with the phenol ring. The discrepency in this assignment likely stems from the differences in the energy ordering of the virtual orbitals between HF and DFT.

The third peak in the spectrum for both thiols is a combination of three low-intensity transitions. Of these, the most intense transition is predicted by OCDFT to be at 2476.2 eV for ethanethiol and 2476.0 eV for benzenethiol and corresponds to the particle orbital ϕ^4_{p}. Considering the total valence characters shown in Figure 4.3Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) ethanethiol (C_2H_5SH) and (B) benzenethiol (C_6H_5SH). Percentages shown represent the overlap of each particle orbital with the LIVVOs. Particle orbitals are numbered according to the calculated transitions reported in Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2 while atom numberings correspond to those shown in Figure 4.2Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.[93]figure.4.2figure.4.3 and the oscillator strengths shown in Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2, we see a clear trend relating the valence character and the oscillator strength of the transition. In the case of ethanethiol, the low valence character of the transitions (28% or lower) suggests that the smaller oscillator strength of the third peak is due to the diffuse mixed Rydberg character of particle orbitals ϕ^3_{p}, ϕ^4_{p}, and ϕ^5_{p}.

In the case of benzenethiol, the valence character of states ϕ^3_{p} and ϕ^4_{p} is high (87% and 84%, respectively) but these

Table 3: Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-D.

State	Energy (eV)	f_{osc}	LIVVO Assignment
Monomer			
1	533.7	0.00811	42% $\sigma^*_{\mathrm{O_1-H_2}}$
			42% $\sigma^*_{\mathrm{O_1-H_3}}$
2	535.4	0.01839	38% $\sigma^*_{\mathrm{O_1-H_2}}$
			38% $\sigma^*_{\mathrm{O_1-H_3}}$
3	537.3	0.00786	Rydberg/Diffuse
4	537.8	0.00252	Rydberg/Diffuse
5	538.5	0.00167	6% $\sigma^*_{\mathrm{O_1-H_2}}$
			6% $\sigma^*_{\mathrm{O_1-H_3}}$
Dimer-A			
1	534.0	0.00737	41% $\sigma^*_{\mathrm{O_1-H_2}}$
			41% $\sigma^*_{\mathrm{O_1-H_3}}$
2	535.8	0.01862	39% $\sigma^*_{\mathrm{O_1-H_2}}$
			39% $\sigma^*_{\mathrm{O_1-H_3}}$
3	537.7	0.00552	14% $\sigma^*_{\mathrm{O_4-H_5}}$
			9% $\sigma^*_{\mathrm{O_4-H_6}}$
4	537.9	0.00529	4% $\sigma^*_{\mathrm{O_4-H_6}}$
			3% $\sigma^*_{\mathrm{O_4-H_5}}$
5	539.3	0.00282	12% $\sigma^*_{\mathrm{O_4-H_6}}$
			8% $\sigma^*_{\mathrm{O_4-H_5}}$
Dimer-D			
1	533.5	0.00605	52% $\sigma^*_{\mathrm{O_4-H_6}}$
			16% $\sigma^*_{\mathrm{O_4-H_5}}$
2	534.9	0.00878	24% $\sigma^*_{\mathrm{O_4-H_6}}$
			18% $\sigma^*_{\mathrm{O_1-H_2}}$
3	536.1	0.00574	3% $\sigma^*_{\mathrm{O_1-H_2}}$
			3% $\sigma^*_{\mathrm{O_1-H_3}}$
4	536.6	0.00284	2% $\sigma^*_{\mathrm{O_4-H_5}}$
			2% $\sigma^*_{\mathrm{O_1-H_3}}$
5	537.3	0.00623	25% $\sigma^*_{\mathrm{O_4-H_5}}$
			5% $\sigma^*_{\mathrm{O_4-H_5}}$

orbitals are mostly localized on the phenyl ring and give small contributions to the transition dipole moment. State ϕ_{p}^3 is a mixture of $\pi^*_{\mathrm{C2,3,4}}$ and $\pi^*_{\mathrm{C4,5,6}}$ at 68% and 15% respectively, while state ϕ_{p}^4 is a mixture of $\pi^*_{\mathrm{C1,2,6}}$ and $\pi^*_{\mathrm{C4,5,6}}$ at 48% and 34% respectively. This analysis suggests that the small overlap of these states with the sulfur atom is the cause of the small intensity of the corresponding transitions. State ϕ_{p}^5 of benzenethiol has a much lower total valence character (16%), however, since it has 7% overlap with the $\sigma^*_{\mathrm{S-H}}$ thiol bond orbital, the corresponding transition still has a modest intensity. We should note that thanks to the LIVVO analysis we were able to quantify the weak local contributions to the states that contribute to the third peak. Previous studies[56] were unable to assess the orbital character of this set of transitions.

4.2   Signatures of hydrogen bonding in the NEXAFS spectrum of the water dimer

Due to the critical role it plays in many processes in nature, hydrogen bonding is one of the most important noncovalent interactions in chemistry.[101–105] The water dimer $(\mathrm{H_2O})_2$ is one of the simplest yet most important examples of hydrogen bonding in chemistry and has garnered plenty of attention from both theory and experiment.[106–117] This simple dimer is the basic building block of the structures of liquid water and ice, which have both been studied extensively with NEXAFS spectroscopy to uncover their underlying hydrogen bond coordination.[118–120] A host of computational methods have been applied to study the effects of hydrogen bonding on the NEXAFS spectrum of water, including transition potential DFT (TPDFT),[61–63,121,122] coupled-cluster singles and doubles (CCSD),[9,123,124] and complex polarization propagator DFT (CPP-DFT).[125] Here, we utilize our LIVVO analysis in order to quantify these effects on the NEXAFS spectrum.

Water monomer. A host of computational methods have been applied to study the effects of hydrogen bonding on the NEXAFS spectrum of water, including transition potential DFT (TPDFT),[61–63,121,122] coupled-cluster singles and doubles (CCSD),[9,123,124] and complex polarization propagator DFT (CPP-DFT).[125]. Here, we utilize our LIVVO analysis in order to quantify these effects on the NEXAFS spectrum. Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4A shows the NEXAFS spectrum for water monomer and the lowest-energy configuration of the water dimer. To facilitate our analysis, we separate the dimer spectrum into contributions from the hole localized on the oxygen accepting the hydrogen bond (O_1, denoted as dimer-A) and a spectrum for the hole localized on the oxygen donating the hydrogen bond (O_4, denoted as dimer-D). These partial contributions are shown in panels B and C of Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4, respectively. Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 shows all of the calculated excitation energies, oscillator strengths and final LIVVO assignments for each state.

The OCDFT computed NEXAFS spectrum of water monomer at the O K-edge, shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4, is defined by three characteristic peaks at 533.7, 535.4, and 537.3 eV respectively. The positions of these peaks are in excellent agreement with the gas-phase experimental peak features which occur at 534.0, 535.9, and 537.0 eV.[126] The first peak at 533.7 eV has a relative intensity of 0.44 and is characterized by the particle orbital ϕ_{p}^1 shown in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5A, while the second peak at 535.4 is the most intense of the spectrum and is characterized by the particle orbital ϕ_{p}^2. Two Rydberg transitions at 537.3 and 537.8 eV contribute to the third peak in the spectrum and are accompanied by a higher energy shoulder feature at 538.5 eV.

The water monomer particle orbitals are shown in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5A along with the overlap with the LIVVOs and their total valence character. The first two transitions in the water monomer are heavily localized resulting in total valence character of 85% and 75%, respectively. The three higher energy transitions are to more diffuse orbitals, as evidenced by the extremely low total valence percentages (in the range 0–12 %). Rydberg transitions that contain no overlap with the LIVVO basis due to their high degree of delocalization can not be assigned a well-defined valence character and are labeled "Rydberg/Diffuse" in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3. Analyzing the contributions of each LIVVO to the respective particle orbitals reveals that for this case, both the $\sigma^*_{\text{O}_1-\text{H}_2}$ and $\sigma^*_{\text{O}_1-\text{H}_3}$ LIVVOs contribute equally to each excited state. This is a characteristic feature of particle orbitals in the spectrum of the water monomer and will be useful in the comparison with the water dimer spectrum.

Water dimer, hydrogen bond acceptor. The spectra for localized excitations in the water dimer shown in panels B

and C of Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4 allow us to discern features that arise from hydrogen bonding in the accepting and donating water molecules. Comparing panels A (monomer) and B (dimer-A) of Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4, the two spectra are very similar with three distinct peaks. The first two core excitations for the dimer-A spectrum occur at 534.0 and 535.8 eV and exhibit a 0.3 and 0.4 eV blueshift, respectively, when compared to the monomer. The oscillator strengths are also similar, with the second transition being the most intense in the spectrum, and the first transition having similar oscillator strength in the dimer-A (0.00737) and monomer (0.00811) spectra.

The similarities between these transitions becomes obvious when comparing the particle orbitals for the monomer and the dimer (shown in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5). The particle orbitals for the first two transitions in the dimer-A spectrum have no contribution from the donating water molecule and are localized exclusively on the accepting water molecule. In addition, the particle orbitals ϕ_{p}^{1} and ϕ_{p}^{2} for the dimer-A look nearly identical to those for the water monomer. These two particle orbitals are localized on the hydrogen atoms of the acceptor water and thus are weakly perturbed by the hydrogen bond. In contrast, the three higher energy transitions show some degree of delocalization onto the donating water molecule.

The LIVVOs can be used in an effort to quantify the degree of localization/delocalization in the particle orbitals involved. Water dimer contains four LIVVOs, two that are localized on the accepting water ($\sigma^{*}_{\mathrm{O_1-H_2}}$ and $\sigma^{*}_{\mathrm{O_1-H_3}}$), one that is localized on the free OH bond in the donor water ($\sigma^{*}_{\mathrm{O_4-H_6}}$), and one that represents the hydrogen bonding interaction between the two water molecules ($\sigma^{*}_{\mathrm{O_4-H_5}}$). While similar in orbital character to the other LIVVOs, $\sigma^{*}_{\mathrm{O_4-H_5}}$ has contributions from the accepting water molecule which gives us a direct way to track the overlap of an excited state orbital with O-H antibonding interaction from the hydrogen bond.

The values reported in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5B correspond to the percentage of overlap between each water dimer LIVVO and the particle orbitals contributing to the dimer-A spectrum. Considering that orbitals ϕ_{p}^{1} and ϕ_{p}^{2} are heavily localized on the accepting water, it is unsurprising that the first two states have equal contributions from the LIVVOs localized on the accepting water ($\sigma^{*}_{\mathrm{O_1-H_2}}$ and $\sigma^{*}_{\mathrm{O_1-H_3}}$). The total valence character of these states (84% and 77%) is similar to the total valence character of the first two particle orbitals in the monomer spectrum (85% and 75%), implying that the presence of the donor water has a negligible effect on ϕ_{p}^{1} and ϕ_{p}^{2}. It is encouraging to see that our LIVVO analysis parallels the uniformity seen in these states with regard to their orbital character, peak position, and peak intensity.

The three particle orbitals corresponding to higher energy transitions in the dimer-A spectrum (ϕ_{p}^{3}, ϕ_{p}^{4}, ϕ_{p}^{5}) have significant overlap with $\sigma^{*}_{\mathrm{O_4-H_5}}$ and $\sigma^{*}_{\mathrm{O_4-H_6}}$. It is in this higher energy region of the spectrum where we see the most significant distinctions between the dimer-A and monomer spectrum. The monomer spectrum shows a shoulder feature at higher energy while in the dimer-A this shoulder feature is now a distinct peak at higher energy. Our analysis based on LIVVOs shows that states 3, 4, and 5 experience a stark increase in total valence character from

0%, 0%, and 12% respectively in the monomer to 23%, 7%, and 29% in the dimer-A. This drastic increase in valence character is caused by interactions with the donor water, in all cases the final LIVVO assignments (shown in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3) can be attributed to the two LIVVOs localized on the donor water. The spectral changes seen in this region of the dimer-A spectrum can be rationalized by the presence of distinct orbital character from the donor water. This result is particularly encouraging, since simple inspection of the particle orbitals give little indication of these contributions from the donor water. In fact, especially in the cases of ϕ_{p}^3 and ϕ_{p}^4, these contributions look fairly insignificant. However our analysis reveals that these seemingly harmless contributions account for a majority, if not all of the total valence character for these states.

Water dimer, hydrogen bond donor. In contrast to dimer-A, the spectrum for dimer-D bears little resemblance to the monomer spectrum. Although there are still three distinct peak features, the oscillator strength of all transitions drop significantly. We also observe a redshift of all three peaks, the first peak exhibits a mild shift of 0.2 eV while the higher energy peaks show more dramatic shifts of 0.7 and 1.0 eV respectively. The stark differences in the dimer-D spectrum, can be rationalized using our particle orbital/LIVVO analysis. The particle orbitals for each state shown in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5C clearly show significant differences from the monomer spectrum for the dimer-D. The particle orbital for the first state in the dimer-D spectrum has similar orbital character to the first state in the monomer spectrum. However, ϕ_{p}^1 for dimer-D is clearly perturbed by the formation of the hydrogen bond and localizes along the free OH bond. This phenomenon has been observed in previous studies,[59,124] but the degree of localization has not been rigorously quantified.

The LIVVO overlaps reported in Figure 4.5Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4figure.4.5C for ϕ_{p}^1 shows that 52% of the orbital character lies along the $\sigma^*_{\mathrm{O}_4-\mathrm{H}_6}$ orbital with only 16% localized along the hydrogen bond and 10% of the orbital character coming from LIVVOs localized on the accepting water molecule. For ϕ_{p}^2, a similar localization is observed, however, a significantly higher percentage (36%) of orbital character is accounted for by LIVVOs on the accepting water. This sharing of electron density with the accepting water molecule is indicative of the effect of hydrogen bonding on the dimer spectrum and has a direct correlation to the changes in intensity that we see in the simulated NEXAFS spectrum. Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 shows a drop in f_{osc} when comparing the monomer to dimer-D, with a modest decrease for state 1 ($0.00811 \rightarrow 0.00605$) and a far more drastic decrease for state 2 ($0.01839 \rightarrow 0.00878$). This intensity drop can be attributed to delocalization of the particle orbital and correlates well with the population of LIVVOs on the accepting water molecule (10% and 36%, respectively).

In conclusion, LIVVOs help to address the origin behind the stark differences in the dimer-A and dimer-D spectra. Perhaps the most revealing signature of hydrogen bonding in the water dimer is the decrease in intensity of the lowest two peaks in the dimer-D contribution. Past studies have reasoned that the differences in these two spectra arise from the polarization of the orbitals in water.[59,124] Due to the large dipole moment the virtual orbital are expected to be polarized toward the two hydrogen atoms and the NEXAFS spectrum will show more pronounced changes for the water molecule donating a hydrogen bond than for the one accepting it. Our LIVVO analysis is consistent with this physical interpretation. In addition, it quantifies the degree of delocalization of the particle orbitals in terms of localized valence orbitals.

Table 4: Basis set dependence of the IAO assignment of LIVVOs and overlap $[\Omega_{\mathrm{p}l}^{(n)}]$ of the $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ and $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ LIVVOs with the first two particle orbitals of ethanethiol.

LIVVO	$\psi_7^{\mathrm{IAO}}(\mathrm{C}_s^2)$	$\psi_9^{\mathrm{IAO}}(\mathrm{C}_p^2)$	$\psi_{10}^{\mathrm{IAO}}(\mathrm{C}_p^2)$	$\psi_{13}^{\mathrm{IAO}}(\mathrm{S}_s^3)$	$\psi_{17}^{\mathrm{IAO}}(\mathrm{S}_p^3)$	$\psi_{18}^{\mathrm{IAO}}(\mathrm{S}_p^3)$	$\psi_{19}^{\mathrm{IAO}}(\mathrm{S}_p^3)$	$\psi_{25}^{\mathrm{IAO}}(\mathrm{H}_s^9)$	$\Omega_{\mathrm{p}l}^{(1)}$	$\Omega_{\mathrm{p}l}^{(2)}$
cc-pVDZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.08	0.23	0.00	0.15	0.54	63.7	22.3
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.09	0.33	0.09	0.07	0.00	0.32	0.09	0.00	22.6	67.1
cc-pVTZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	59.3	23.9
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	24.6	61.5
cc-pVQZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	57.8	22.8
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	24.3	59.2
aug-cc-pVDZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.16	0.54	60.8	11.1
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	15.3	63.1
aug-cc-pVTZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	57.9	14.7
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	19.2	59.1
aug-cc-pVQZ										
$\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$	0.00	0.00	0.00	0.07	0.23	0.00	0.16	0.54	56.5	15.7
$\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$	0.08	0.33	0.09	0.07	0.00	0.32	0.09	0.00	20.4	57.3

4.3 Basis Set Dependence of LIVVO Analysis

An important aspect to address concerning this method is its dependence upon the choice of basis set. To fully address this issue, there are three aspects of the method that must be tested with respect to their basis set dependence: 1) How consistent are the LIVVO assignments, i.e. is there significant fluctuation in the LIVVO/IAO overlaps (eq 4.9Determination of the character of the IAOs and LIVVOsequation.4.2.9)? 2) Are the LIVVOs assigned to the excited state consistent? and 3) How much fluctuation is there in the particle orbital/LIVVO overlaps (eq 4.11Analysis of OCDFT Particle Orbitals Using LIVVOsequation.4.2.11)? We will address these three points in this section, using the two LIVVOs used to assign the first two excited states in ethanethiol, $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ and $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$.

First we will take a look at the assignment of the atomic character of the LIVVOs, Table 4.4Basis set dependence of the IAO assignment of LIVVOs and overlap $[\Omega_{\mathrm{p}l}^{(n)}]$ of the $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ and $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ LIVVOs with the first two particle orbitals of ethanethioltable.4.4 shows the overlap of each VVO with eight of the IAOs in ethanethiol. These values are extremely consistent regardless of the choice of basis set, only deviating by a maximum of 0.01.

Next we will evaluate if the LIVVO assignments of the particle orbital are consistent for different basis sets. For ethanethiol, ϕ_{p}^1 was assigned primarily to the $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ LIVVO with a weaker contribution from $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ while ϕ_{p}^2 received a complementary assignment. Regardless of the choice of basis set, the LIVVO classification of both excited states is consistent. In all cases, ϕ_{p}^1 has greater than 57% overlap with the $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ LIVVO and 25% or less overlap with the $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ LIVVO. Similarly, ϕ_{p}^2 has 57% or greater overlap with $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ and 25% or less overlap with $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$. The characterization of the orbital character of excited states is by far the most important feature of this technique, the consistency seen in the data across different basis sets is very encouraging.

Lastly, we will address the consistency of the LIVVO/particle orbital overlap quantity presented in eq 4.11Analysis of OCDFT Particle Orbitals Using LIVVOsequation.4.2.11 with respect to choice of basis set. Table 4.4Basis set dependence of the IAO assignment of LIVVOs and overlap $[\Omega_{\mathrm{p}l}^{(n)}]$ of the $\sigma_{\mathrm{S}^3-\mathrm{H}^9}^*$ and $\sigma_{\mathrm{S}^3-\mathrm{C}^2}^*$ LIVVOs with the first two particle orbitals of ethanethioltable.4.4 shows that this value can potentially fluctuate $\pm$ 10% with the choice of basis set. However, since we have shown that the formulation of the LIVVOs has a negligible dependence on the choice of basis set, we conclude that the basis set dependence seen in these values can solely be attributed to the basis set dependence of the particle orbitals. Although these values can change fairly significantly, it is encouraging that they do not effect the overall interpretation of the excited state.

5 Conclusions

In this work, we present an automated method for the characterization of core-valence excited states. Our approach utilizes intrinsic atomic orbitals (IAOs) to derive a set of localized intrinsic valence virtual orbitals (LIVVOs) and assign their orbital character. LIVVOs are in turn used to classify particle orbitals calculated using orthogonality constrained density functional theory (OCDFT).

For molecular orbitals with dominant valence character, this analysis provides keen insight into the localized σ^* and π^* orbitals that contribute to it. For example, in the first two core excited states of ethanethiol, an electron is promoted to virtual orbitals that span both the S–C and S–H bonds. These contributions are difficult to discern by visual inspections of the MOs. For the first state, our analysis shows that the particle orbital is dominated by the $\sigma^*_{S^3-H^9}$ LIVVO (58%) with a smaller contribution from the $\sigma^*_{S^3-C^2}$ (22%) LIVVO, revealing that this transition involves primarily the thiol bond. Instead, for the second state we find that weights are reversed, with the $\sigma^*_{S^3-C^2}$ LIVVO being the dominant contribution (57%).

We have also shown how LIVVOs may be useful to quantify differences in excited states due to changes in the molecular environment. For example, in the case of the water dimer we analyze the correlation between orbital localization due to hydrogen bonding and spectral features. The NEXAFS spectrum of water dimer at the oxygen K-edge is known to have distinct contributions from the oxygens at water donating (dimer-D) or accepting (dimer-A) the hydrogen bond. Analysis of the two lowest particle orbitals for excitations on dimer-A show equal localization along intramolecular OH bonds. However, the corresponding excitations for dimer-D are significantly different. In the lowest energy state, 70% of the electron is localized on the dimer-D, and this percentage decreases to 34% in the second state. These changes in degree of localization explain changes in intensity in the oxygen K-edge spectrum predicted by theory.

We have also considered the basis set dependence of the LIVVOs and the orbital analysis. Due to the robustness of the intrinsic atomic orbitals, the LIVVOs and our analysis of the orbital character are largely insensitive to the basis set size.

While our LIVVO analysis is applied here within the context of orthogonality constrained density functional theory, the method is general and can be used to decompose any set of virtual orbitals. Overall, our LIVVO-based analysis of particle orbitals is a first step toward creating a general, robust, and automatic method to assign the character of excited states.
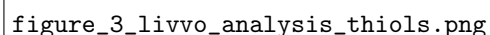
Acknowledgments

References

[1] Westre, T. E.; Kennepohl, P.; DeWitt, J. G.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1997**, *119*, 6297–6314.

[2] Löble, M. W. et al. *J. Am. Chem. Soc.* **2015**, *137*, 2506–2523.

[3] Idrees, M.; Nadeem, M.; Sung, N.-E.; Asanova, T.; Shin, T. J. *Chem. Phys. Lett.* **2014**, *612*, 262–265.

[4] Nelson, R. C.; Miller, J. T. *Catal. Sci. Technol.* **2012**, *2*, 461–470.

[5] Stöhr, J. *NEXAFS Spectroscopy*; Springer Series in Surface Sciences; Springer, 1992.

[6] Nelson, R. C.; Miller, J. T. *Catal. Sci. Technol.* **2012**, *2*, 461–470.

[7] Milne, C. J.; Penfold, T. J.; Chergui, M. *Coord. Chem. Rev.* **2014**, *277*, 44–68.

[8] Cabaret, D.; Bordage, A.; Juhin, A.; Arfaoui, M.; Gaudry, E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5619–5633.

[9] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A* **2012**, *85*, 022507.

[10] Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 6735–6756.

[11] Besley, N. A. *Chem. Phys. Lett.* **2012**, *542*, 42–46.

[12] Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.

[13] Dutta, A. K.; Gupta, J.; Vaval, N.; Pal, S. *J. Chem. Theory Comput.* **2014**, *10*, 3656–3668.

[14] Brabec, J.; Bhaskaran-Nair, K.; Govind, N.; Pittner, J.; Kowalski, K. *J. Chem. Phys.* **2012**, *137*, 171101.

[15] Sen, S.; Shee, A.; Mukherjee, D. *Mol. Phys.* **2013**, *111*, 2625–2639.

[16] Schirmer, J. *Phys. Rev. A* **1982**, *26*, 2395.

[17] Salpeter, E. E.; Bethe, H. A. *Phys. Rev.* **1951**, *84*, 1232–1242.

[18] Barth, A.; Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. *Chem. Phys.* **1980**, *46*, 149–164.

[19] Butscher, W.; Buenker, R. J.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1977**, *52*, 449–456.

[20] Roemelt, M.; Neese, F. *J. Phys. Chem. A* **2013**, *117*, 3069–3083.

[21] Asmuruf, F.; Besley, N. *Chem. Phys. Lett.* **2008**, *463*, 267–271.

[22] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.

[23] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.

[24] Fernando, R. G.; Balhoff, M. C.; Lopata, K. *J. Chem. Theory Comput.* **2015**, *11*, 646–654.

[25] Rehr, J. J.; Albers, R. C. *Rev. Mod. Phys.* **2000**, *72*, 621–654.

[26] Natoli, C. R.; Krüger, P.; Hatada, K.; Hayakawa, K.; Sébilleau, D.; Šipr, O. *J. Phys-Condens. Mat.* **2012**, *24*, 365501.

[27] Joly, Y. *Phys. Rev. B* **2001**, *63*, 125120.

[28] Slater, J. C. In *Statistical Exchange-Correlation in the Self-Consistent Field*; Löwdin, P.-O., Ed.; Adv. Quantum Chem.; Academic Press, 1972; Vol. 6; pp 1–92.

[29] Stener, M.; Fronzoni, G.; Simone, M. d. *Chem. Phys. Lett.* **2003**, *373*, 115–123.

[30] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039.

[31] Lestrange, P. J.; Nguyen, P. D.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.

[32] Kuramoto, K.; Ehara, M.; Nakatsuji, H. *J. Chem. Phys.* **2005**, *122*, 014304.

[33] Ohtsuka, Y.; Nakatsuji, H. *J. Chem. Phys.* **2006**, *124*, 054110.

[34] Kolczewski, C.; Püttner, R.; Plashkevych, O.; Ågren, H.; Staemmler, V.; Martins, M.; Snell, G.; Schlachter, A. S.; Sant'Anna, M.; Kaindl, G.; Pettersson, L. G. M. *J. Chem. Phys.* **2001**, *115*, 6426–6437.

[35] Wenzel, J.; Dreuw, A. *J. Chem. Theory Comput.* **2016**, *12*, 1314–1330.

[36] Plasser, F.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2014**, *141*, 024106.

[37] Plasser, F.; Bappler, S. A.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2014**, *141*, 024107.

[38] Myhre, R. H.; Coriani, S.; Koch, H. *J. Chem. Theory Comput.* **2016**, *12*, 2633–2643.

[39] Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51–64.

[40] Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.

[41] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.

[42] Lee, M. S.; Head-Gordon, M. *Int. J. Quantum Chem.* **2000**, *76*, 169–184.

[43] Subotnik, J. E.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2005**, *123*, 114108.

[44] Auer, A. A.; Noojien, M. *J. Chem. Phys.* **2006**, *125*, 024104.

[45] Laikov, D. N. *Int. J. Quantum Chem.* **2011**, *111*, 2851–2867.

[46] Szczepanik, D.; Mrozek, J. *J Math Chem* **2013**, *51*, 2687–2698.

[47] Knizia, G. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

[48] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.

[49] Schmidt, M. W.; Hull, E. A.; Windus, T. L. *J. Phys. Chem. A* **2015**, *119*, 10408–10427.

[50] West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. *J. Phys. Chem. A* **2015**, *119*, 10376–10389.

[51] West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2013**, *139*, 234107.

[52] Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[53] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.

[54] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, 14360–14374.

[55] Behyan, S.; Hu, Y.; Urquhart, S. G. *Chem. Phys. Lett.* **2014**, *592*, 69–74.

[56] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2013**, *138*, 214302.

[57] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2011**, *134*, 244304.

[58] Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Näslund, L. Å.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995–999.

[59] Cavalleri, M.; Ogasawara, H.; Pettersson, L. G. M.; Nilsson, A. *Chem. Phys. Lett.* **2002**, *364*, 363–370.

[60] Smith, J. D.; Cappa, C. D.; Messer, B. M.; Drisdell, W. S.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2006**, *110*, 20038–20045.

[61] Prendergast, D.; Galli, G. *Phys. Rev. Lett.* **2006**, *96*, 215502.

[62] Iannuzzi, M. *J. Chem. Phys.* **2008**, *128*, 204506.

[63] Leetmaa, M.; Ljungberg, M. P.; Lyubartsev, A.; Nilsson, A.; Pettersson, L. G. M. *J. Electron. Spectrosc. Relat. Phenom.* **2010**, *177*, 135–157.

[64] Kühne, T. D.; Khaliullin, R. Z. *J. Am. Chem. Soc.* **2014**, *136*, 3395–3399.

[65] Tse, J. S. *Phys. Rev. Lett.* **2008**, *100*.

[66] Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *Science* **2004**, *306*, 851–853.

[67] Smith, J. D.; Cappa, C. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *Science* **2005**, *308*, 793–793.

[68] Chen, W. *Phys. Rev. Lett.* **2010**, *105*.

[69] Clark, G. N.; Cappa, C. D.; Smith, J. D.; Saykally, R. J.; Head-Gordon, T. *Mol. Phys.* **2010**, *108*, 1415–1433.

[70] Vinson, J. *Phys. Rev. B* **2012**, *85*.

[71] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.

[72] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2638–2651.

[73] Mulliken, R. *J. Chem. Phys.* **1955**, *23*, 1833–1840.

[74] Mulliken, R. *J. Chem. Phys.* **1955**, *23*, 1841–1846.

[75] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A.* **2012**, *85*, 042518.

[76] Turney, J. M. et al. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 556–565.

[77] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

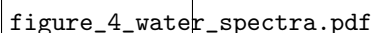[78] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[79] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[80] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[81] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[82] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[83] Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.

[84] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993–4999.

[85] Schrader, D. M.; Prager, S. *J. Chem. Phys.* **1962**, *37*, 1456–1460.

[86] Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys.* **1973**, *2*, 41 – 51.

[87] Dunlap, B. I. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2113–2116.

[88] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *Int. J. Quantum Chem., Symp.* **1977**, *12*, 81–87.

[89] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[90] Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.

[91] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[92] Parrish, R. M.; Sherrill, C. D. *J. Am. Chem. Soc.* **2014**, *136*, 17386–17389.

[93] Legault, C. Y. CYLview. 2009; http://www.cylview.org.

[94] Verma, P.; Derricotte, W. D.; Evangelista, F. A. *J. Chem. Theory Comput.* **2016**, *12*, 144–156.

[95] Otero, E.; Urquhart, S. G. *J. Phys. Chem. A* **2006**, *110*, 12121–12128.

[96] Cooney, R. R.; Urquhart, S. G. *J. Phys. Chem. B* **2004**, *108*, 18185–18191.

[97] Urquhart, S. G.; Turci, C. C.; Tyliszczak, T.; Brook, M. A.; Hitchcock, A. P. *Organometallics* **1997**, *16*, 2080–2088.

[98] Urquhart, S. G.; Hitchcock, A. P.; Lehmann, J. F.; Denk, M. *Organometallics* **1998**, *17*, 2352–2360.

[99] Urquhart, S. G.; Smith, A. P.; Ade, H. W.; Hitchcock, A. P.; Rightor, E. G.; Lidy, W. *J. Phys. Chem. B* **1999**, *103*, 4603–4610.

[100] Kosugi, N.; Ueda, K.; Shimizu, Y.; Chiba, H.; Okunishi, M.; Ohmori, K.; Sato, Y.; Shigemasa, E. *Chem. Phys. Lett.* **1995**, *246*, 475–480.

[101] Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294–300.

[102] Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1–22.

[103] Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117–4128.

[104] Stiopkin, I. V.; Weeraman, C.; Pieniazek, P. A.; Shalhout, F. Y.; Skinner, J. L.; Benderskii, A. V. *Nature* **2011**, *474*, 192–195.

[105] Pflugrath, J. W.; Quiocho, F. A. *Nature* **1985**, *314*, 257–260.

[106] Feyereisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993–2997.

[107] Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089–10094.

[108] Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066–4073.

[109] Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15–22.

[110] Szalewicz, K.; Cole, S. J.; Kołos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662–3673.

[111] Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765–1769.

[112] Wang, Y.; Carter, S.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2008**, *128*, 071101.

[113] Ch'ng, L. C.; Samanta, A. K.; Czakó, G.; Bowman, J. M.; Reisler, H. *J. Am. Chem. Soc.* **2012**, *134*, 15430–15435.

[114] Shank, A.; Wang, Y.; Kaledin, A.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2009**, *130*, 144314.

[115] Mukhopadhyay, A.; Cole, W. T. S.; Saykally, R. J. *Chem. Phys. Lett.* **2015**, *633*, 13–26.

[116] Gómez, S.; Nafziger, J.; Restrepo, A.; Wasserman, A. *J. Chem. Phys.* **2017**, *146*, 074106.

[117] Zhang, Z.; Jiang, W.; Wang, B.; Wang, Z. *Chemical Physics Letters* **2017**, *678*, 98–101.

[118] 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*, 3346–3349.

[119] Wilson, K. R.; Cavalleri, M.; Rude, B. S.; Schaller, R. D.; Nilsson, A.; Pettersson, L. G. M.; N Goldman,; Catalano, T.; Bozek, J. D.; Saykally, R. J. *J. Phys.: Condens. Matter* **2002**, *14*, L221.

[120] Parent, P.; Laffon, C.; Mangeney, C.; Bournel, F.; Tronc, M. *J. Chem. Phys.* **2002**, *117*, 10842–10851.

[121] Hetényi, B.; Angelis, F. D.; Giannozzi, P.; Car, R. *The Journal of Chemical Physics* **2004**, *120*, 8632–8637.

[122] Cavalleri, M.; Odelius, M.; Nordlund, D.; Nilsson, A.; Pettersson, L. G. M. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2854–2858.

[123] Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. *J. Chem. Theory Comput.* **2012**, *8*, 1616–1628.

[124] Fransson, T.; Zhovtobriukh, I.; Coriani, S.; Wikfeldt, K. T.; Norman, P.; Pettersson, L. G. M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 566–583.

[125] Ekström, U. *Phys. Rev. A* **2006**, *74*.

[126] Schirmer, J. *Phys. Rev. A* **1993**, *47*, 1136–1147.

Figure 3: Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) ethanethiol (C_2H_5SH) and (B) benzenethiol (C_6H_5SH). Percentages shown represent the overlap of each particle orbital with the LIVVOs. Particle orbitals are numbered according to the calculated transitions reported in Table 4.2Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57table.4.2 while atom numberings correspond to those shown in Figure 4.2Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.[93]figure.4.2.

Figure 4: Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]

Figure 5: Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-Dtable.4.3 while atom numberings correspond to those shown in Figure 4.4Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red cicle. Structures were plotted using the CYLview software.[93]figure.4.4.