THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A XXXX, XXX, XXX−XXX

pubs.acs.org/JPCA

Water-Mediated Selenium Hydrogen-Bonding in Proteins: PDB Analysis and Gas-Phase Spectroscopy of Model Complexes

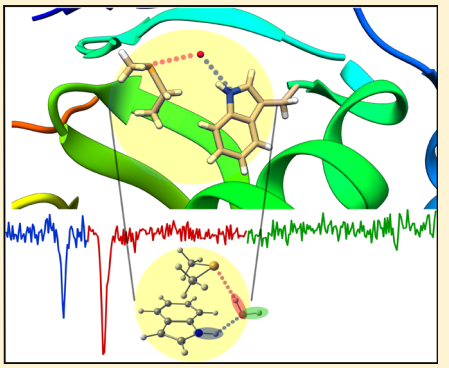
Kamal K. Mishra,[†] Santosh K. Singh,[†,∥] Satish Kumar,[†] Gulzar Singh,[‡] Biplab Sarkar,*[,§] M. S. Madhusudhan,*[,‡] and Aloke Das*[,†]

[†]Department of Chemistry and [‡]Department of Biology, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pashan, Pune 411008, India

[§]Department of Chemistry, North Eastern Hill University, Shillong, Meghalaya 793022, India

Supporting Information

ABSTRACT: High-resolution X-ray crystallography and two-dimensional NMR studies demonstrate that water-mediated conventional hydrogen-bonding interactions (N−H···N, O−H···N, etc.) bridging two or more amino acid residues contribute to the stability of proteins and protein−ligand complexes. In this work, we have investigated single water-mediated selenium hydrogen-bonding interactions (unconventional hydrogen-bonding) between amino acid residues in proteins through extensive protein data bank (PDB) analysis coupled with gas-phase spectroscopy and quantum chemical calculation of a model complex consisting of indole, dimethyl selenide, and water. Here, indole and dimethyl selenide represent the amino acid residues tryptophan and selenomethionine, respectively. The current investigation demonstrates that the most stable structure of the model complex observed in the IR spectroscopy mimics single water-mediated selenium hydrogen-bonded structural motifs present in the crystal structures of proteins. The present work establishes that water-mediated Se hydrogen-bonding interactions are ubiquitous in proteins and the number of these interactions observed in the PDB is more than that of direct Se hydrogen-bonds present there.



1. INTRODUCTION

Subtle interplay among various types of noncovalent interactions in the backbone and side chains of amino acid residues plays a significant role for protein folding, protein−protein as well as protein−ligand interactions.[1−3] Water molecules also contribute significantly to the folding of proteins and biomolecular recognition processes through the hydrophobic effect.[4−6] However, water molecules specifically present in the cavity of proteins and at the interface of proteins and ligands in protein−ligand complexes can play a crucial role in their stability.[7−16] It has been found from high-resolution X-ray crystallography and two-dimensional NMR studies that water-mediated conventional hydrogen-bonding interactions bridging two or more amino acid residues contribute to the stability of proteins and protein−ligand complexes.[9,11,12,17−19] Helms et al. performed molecular dynamics simulation of protein−ligand complexes and concluded that a water molecule in the cavity stabilizes the complexes only if it forms more than one hydrogen-bond with the ligand and neighboring residues of the protein.[12]

Conventional hydrogen-bonding interactions are defined as X−H···Y, where both X and Y are strongly electronegative atoms (i.e., O, N).[20,21] In the case of unconventional hydrogen-bonding interactions, X and Y are not conventional electronegative atoms. Although water-mediated conventional hydrogen-bonding interactions present in the cavity contributing to the stability of proteins are reported in the literature, there is surprisingly no study on water-mediated unconventional hydrogen-bonding interactions present in the protein structures.

In this work, we have investigated water-mediated selenium (Se) hydrogen-bonding interactions between amino acid residues involving selenomethionine in proteins through extensive protein data bank (PDB) analysis. To better understand the nature, strength, and binding motif of this interaction, conformation-specific IR spectroscopy and quantum chemistry calculations of a model complex mimicking water-mediated Se hydrogen-bonding has been studied further in an isolated gas phase. The model complex studied here consists of indole, dimethyl selenide, and H_2O, while indole and dimethyl selenide represent the side chains of the amino acid residues tryptophan (Trp) and selenomethionine (Mse), respectively. Mse is the selenium analogue of methionine (Met) in which sulfur (S) is replaced by selenium (Se). Mse can be incorporated into proteins without affecting their structures and functions. The presence of Mse in proteins helps in solving their crystal structures.[22] It has been reported recently that both Se- and S-centered hydrogen-bond interactions are of similar strength to oxygen- (O) and nitrogen (N)-centered conventional hydrogen-bond interac-

Received: May 2, 2019
Revised: June 21, 2019
Published: June 21, 2019

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

tions.[23−25] It has also been shown from PDB analysis that direct Se and S hydrogen-bond interactions between two amino acid residues are extensively present in proteins.[23,24] The present work establishes that water-mediated Se hydrogen-bonding interactions are ubiquitous in proteins and these interactions observed in the PDB are more abundant than the direct Se hydrogen-bonds present there.

2. EXPERIMENTAL SECTION

2.1. Experimental Methods. The experimental setup to study mass-selected conformation-specific electronic and IR spectra of molecules and complexes has been reported elsewhere,[26−28] and we have provided here only a brief description of it. Here, the Ind···H_2O···Me_2Se complex was synthesized by supersonic expansion of mixed vapor of indole (Ind), H_2O, and dimethyl selenide (Me_2Se) seeded in a He−Ne (30:70) carrier gas through a pulsed valve (General Valve, series 9, 10 Hz, 0.5 mm diameter) into a high-vacuum chamber. The carrier gas (4 bar) mixed with Me_2Se (>90%, Alfa Aesar) and H_2O and maintained at −78 and 5 °C, respectively, was passed through indole (99%, Sigma Aldrich) heated at 75 °C before the expansion through the orifice. The temperature of Me_2Se at −78 °C was maintained by taking it in a sample holder kept in a cold bath of dry ice, while H_2O was taken in another sample holder placed in an ice−water mixture to maintain the temperature at 5 °C.

Mass-selected electronic spectrum of the Ind···H_2O···Me_2Se complex was measured using the 2-color resonant 2-photon ionization (2C-R2PI) technique. In this technique, the first photon of about 100 μJ pulse energy (dye laser ND6000, Continuum, pumped by an Nd/YAG laser, ns, 10 Hz, Surelite II-10, Continuum) was scanned through the electronic spectral region, while the second photon (~300 μJ pulse energy) from another dye laser (ND6000, Continuum, pumped by an Nd:YAG laser, ns, 10 Hz, Surelite II-10, Continuum) fixed at 355 nm was spatially and temporarily overlapped with the first photon. UV−UV hole-burning spectroscopy was performed to discriminate the presence of different species or different conformers of the same species.[29,30] Conformer-specific IR spectra were measured using the resonant ion dip infrared (RIDIR) spectroscopy technique. All of the spectroscopic techniques employed in this experiment were described previously in detail.[26]

2.2. Bioinformatics. The details about downloading the PDB files and the analysis of single water-mediated Se hydrogen-bonding interactions in protein structures retrieved from the PDB are described in the Results and Discussion section.

2.3. Computational Methods. Quantum chemical calculations were performed using the Gaussian16 suite of programs[31] to obtain the optimized structures of various possible conformers of Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se complexes using density functional theory-based ωB97X-D and M06-2X functionals with the 6-311++G(d,p) basis set. Geometry optimizations were carried out using tight convergence criteria and an ultrafine grid to explore various minima having no imaginary frequencies. Initial geometries of the trimeric and tetrameric complexes were obtained from chemical intuition as well as conformational search using the genetic algorithm option in molecular mechanics with a universal force field available in Avogadro software.[32,33] Basis set superposition error (BSSE)[34] and zero point energy (ZPE) corrections were included in the binding energies of the complexes.

Theoretical harmonic vibrational frequencies of the NH and OH groups in the complexes were scaled with respect to the experimental N−H stretching frequency of the indole monomer[35] and the symmetric O−H stretching frequency of the water monomer,[36,37] respectively, reported in the literature. A scaling factor of 0.9463 was obtained for the ωB97X-D/6-311++G(d,p)-level-calculated NH frequency in the complexes by taking the ratio of the experimental (3525 cm^{-1}) and the theoretical (3725 cm^{-1}) NH stretching frequency of the indole monomer. Similarly, the scaling factor for the ωB97X-D/6-311++G(d,p)-level-calculated OH frequency in the complexes was found to be 0.9378. The scaling factors for the NH and OH groups in the complexes were also determined at the M06-2X/6-311++G(d,p) level of theory, and these values were 0.9503 and 0.9398, respectively. These scaled vibrational frequencies were used to assign the experimental IR spectra of the complexes.

Water-mediated Se hydrogen-bonding interactions between amino acid residues in a few PDB structures were characterized by quantum theory of atoms-in-molecules using AIM 2000 software.[38,39] Natural bond orbital (NBO) analysis was performed using the NBO 6.0 package.[40,41] Binding energies of the complexes were decomposed using reduced variational space energy decomposition analysis (RVS-EDA)[42] and localized molecular orbital energy decomposition analysis (LMO-EDA)[43] implemented in the GAMESS package.[44]

3. RESULTS AND DISCUSSION

3.1. Analysis of PDB Structures. We downloaded 9175 protein structures containing selenomethionine (Mse) from the PDB.[45] On filtering these proteins for X-ray structures with a resolution of ≤2.5 Å and at 30% sequence identity, using PISCES,[46] we were left with 5186 structures for further analysis. Hydrogen atoms were added to the amino acid residues in the protein structures using REDUCE.[47] In these data, there were 4334 direct Se hydrogen-bonding interactions between Mse and other amino acid residues.[23] While such direct hydrogen bonds have been reported earlier, in this study, we report single water-mediated Se hydrogen-bonding interactions from PDB for the first time. In the water-bridged interactions, Se of Mse is always the hydrogen bond acceptor with water as its donor. However, the residues that are bridged by the water could either have atoms that donate or accept hydrogen from water (Figure 1a). The distance and angle criteria considered for searching single water-mediated Se hydrogen-bonding interactions are as follows: 2.5 Å ≤ $d_{O\cdots Se}$ ≤ 4.5 Å; 2.0 Å ≤ $d_{D/A\cdots O}$ ≤ 3.5 Å; 3.8 Å ≤ $d_{D/A\cdots Se}$ ≤ 8 Å; 40° ≤ (∠D···O···Se) ≤ 180°, and 75° ≤ (∠A···O···Se) ≤ 135° (Figure 1a).

We observed 4017 and 3509 single water-mediated Se hydrogen-bonding interactions (water bridges) considering interactions of H_2O with hydrogen-bond donors (D) or acceptors (A) of the amino acid residues, respectively (Figure 1a). Thus, we have a nonredundant set of 7526 single water-mediated Se hydrogen-bonding interactions from 2978 different PDB entries. It is interesting to note that the number of single water-mediated Se hydrogen-bonding interactions observed in the PDB is almost twice than that of the direct Se hydrogen-bonding interactions.[23] This is an important result and may have implications for the relative stability of this type of association.

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

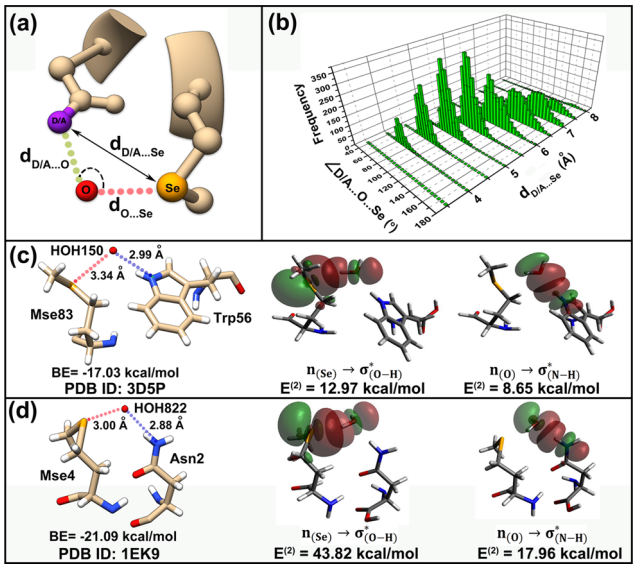


Figure 1. (a) Schematic showing the features of the Se···H_2O···amino acid hydrogen-bond network in PDB structures. The selenium, oxygen of the water, and the donor (D) or acceptor (A) atom from another amino acid are shown in orange, red, and purple ball representation, respectively, and labeled accordingly. (b) Three-dimensional (3D) plot showing the frequency of single water-mediated Se hydrogen-bonding interactions screened from PDB as a function of ∠D/A···O···Se (°) and $d_{D/A\cdots Se}$ distance (Å). Examples of single water-mediated Se hydrogen-bonding interactions between (c) selenomethionine (Mse) and tryptophan (Trp) residues in PDB ID 3d5p and (d) selenomethionine (Mse) and asparagine (Asn) residues in PDB ID 1EK9. ωB97X-D/6-311++G(d,p)-level-calculated binding energy (BE) and NBO plots showing individual hydrogen-bond interactions between the residues in the PDB structures are also shown in the figure. $E^{(2)}$ stands for NBO second-order perturbation energy between the lone pair donor and antibonding acceptor orbitals.

Figure 1b shows a 3D plot representing the frequency of single water-mediated Se hydrogen-bonding interactions or water bridges observed in the PDB as a function of ∠D/A···O···Se and $d_{D/A\cdots Se}$, where D/A is a hydrogen-bond donor/acceptor atom of any amino acid residue except Mse (see Figure 1a). It should be noted that the distances ($d_{D/A\cdots Se}$) are binned at an interval of 0.5 Å for representation purposes. The plot shows that the occurrence of single water-mediated Se hydrogen-bonding interactions is maximum at $d_{D/A\cdots Se}$ = 5−5.5 Å and ∠D/A···O···Se = 80−90°.

Two representative examples of single water-mediated Se hydrogen-bonding interactions with neighboring residues extracted from the proteins in the PDB are shown in Figure 1c,d.

Hydrogen-bond distances are shown between the heavy atoms, i.e., the Se atom of Mse and the oxygen atom of water as well as the hydrogen-bond donor/acceptor atom of any other residue and the oxygen atom of water as it is quite difficult to determine the exact position of the hydrogen atoms of free water in the protein from X-ray crystallography. The distances depicted between water oxygen and neighboring hydrogen-bond donor/acceptor atoms indicate the presence of water-mediated hydrogen-bonding interactions in the PDB. The van der Waals (vdW) radii of Se, H, O, and N are 1.9, 1.2, 1.52, and 1.55 Å, respectively.[48]

Figure 1c,d also shows BSSE-corrected binding energies (BE) and NBO second-order perturbation energies ($E^{(2)}$) of the amino acid residues interacting with water in the two selected PDB structures calculated at the ωB97X-D/6-311++G(d,p) level of theory. We have put the H atoms on the amino acid residues interacting with water in the PDB structures through REDUCE software while the H atoms on water oxygen atoms using pymol software.[49] We have performed partial optimization of the hydrogen atoms keeping all other coordinates of the selected interacting residues of the PDB structures fixed. Electron density topology of the amino acid residues interacting with water in the selected PDB structures (Figure 1c,d) obtained from the AIM calculation has been provided in Figure S1. The BE, NBO, and AIM results indicate that single water-mediated Se hydrogen-bonding interaction between amino acid residues is more stable than the direct Se hydrogen bond as the former one makes at least two hydrogen bonds.

Figure 1b as well as Figure 1c,d reveals that there is a large variation in the Se···O and N···O distances involving water-mediated Se hydrogen-bonding interactions in the PDB due to the constraint of the optimum arrangement of other surrounding residues of the proteins. However, it is obvious that water-mediated Se hydrogen-bonding interactions present in the PDB structures contribute to the stability of proteins. We have further investigated the intrinsic nature, strength, and binding motif of single water-mediated Se hydrogen-bonding interactions by studying a noncovalently bonded model complex, consisting of two amino acid residues including Mse and a water molecule, in an isolated condition.

3.2. Electronic Spectra of the Complexes. A weakly bound 1:1:1 model complex of indole, dimethyl selenide, and water has been studied in the isolated gas phase using UV and IR laser-based spectroscopic techniques. Indole and dimethyl selenide represent the amino acid residues tryptophan (Trp) and selenomethionine (Mse), respectively. Figure 2a shows the electronic spectrum measured in the mass channel of the Ind···

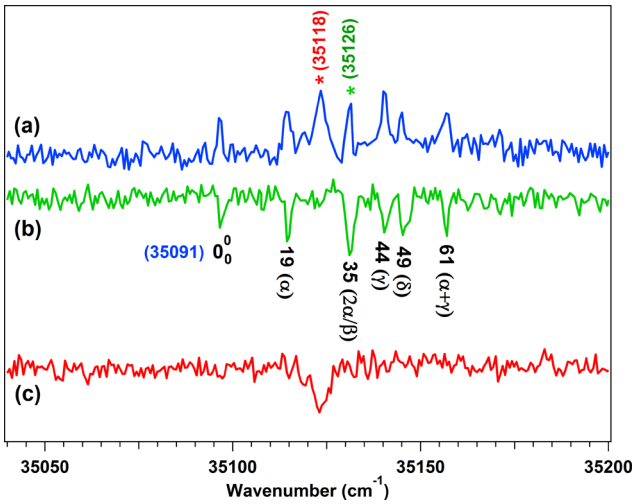


Figure 2. (a) Electronic spectrum measured in the Ind···H_2O···Me_2Se mass channel using 2C-R2PI spectroscopy. (b)-(c) UV−UV hole-burning spectra by fixing the probe UV laser at 35 126 and 35 118 cm^{-1} bands, respectively, marked by asterisks in the electronic spectrum shown in part (a). The low-frequency bands riding on the origin band of Ind···H_2O···Me_2Se in (b) are intermolecular vibrations of the complex.

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

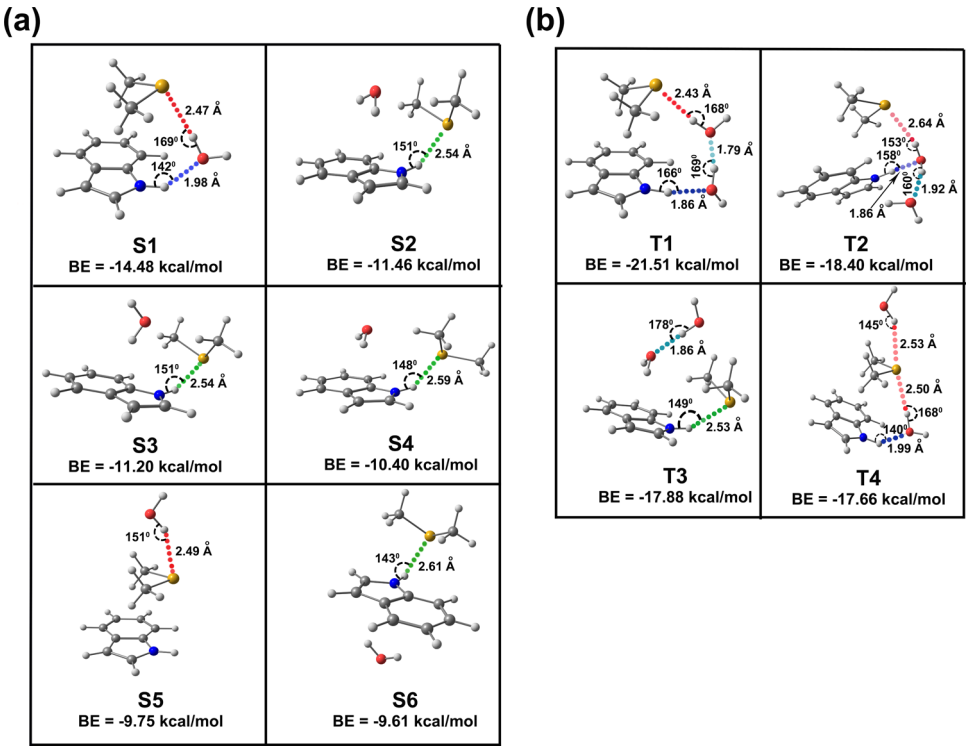


Figure 3. Various low-energy structures of (a) indole···H_2O···Me_2Se and (b) indole···$(H_2O)_2$···Me_2Se complexes optimized at the ωB97X-D/6-311++G(d,p) level of theory. Binding energy (BE) values are BSSE- and zero-point energy-corrected.

H_2O···Me_2Se complex employing 2C-R2PI spectroscopy. The electronic spectrum consists of several sharp but weak bands in the 35040−35200 cm^{-1} spectral region.

UV−UV hole-burning spectrum depicted in Figure 2b has been recorded by probing the 35 126 cm^{-1} band marked by a green asterisk in Figure 2a. The hole-burning spectrum shown in Figure 2b indicates that all of the electronic bands of Figure 2a, except for the 35 118 cm^{-1} peak marked by a red asterisk, belong to a single conformer of the Ind···H_2O···Me_2Se complex. The UV−UV hole-burning spectrum presented in Figure 2c by probing the 35 118 cm^{-1} band shows only its own depletion. Thus, the 35 118 cm^{-1} band present in the electronic spectrum of the Ind···H_2O···Me_2Se complex belongs to either its different conformer or a higher-order cluster fragmented into the trimer mass channel. Measurement of the IR spectra by probing one of the electronic bands of Ind···H_2O···Me_2Se and the 35118 cm^{-1} band in combination with quantum chemistry calculations can reveal the source of the single band observed in the hole-burning spectrum displayed in Figure 2c, and this has been discussed in the latter section.

It should be noted that the electronic origin band for the $S_1 \leftarrow S_0$ transition of the Ind···H_2O···Me_2Se trimeric complex (Figure 2b), which appears at 35 091 cm^{-1}, is red-shifted compared to that of the indole monomer and indole···H_2O complex by 132 and 150 cm^{-1}, respectively.[35] The low-frequency vibronic bands riding on the origin band of the trimeric complex (Figure 2b) can be explained tentatively in terms of fundamental, overtone, and combination modes of intermolecular vibrations denoted α, β, γ, and δ.

3.3. Conformational Landscape of Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se Complexes. Figure 3a,b shows various low-energy conformers of Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se, respectively, optimized at the ωB97X-D/6-311++G(d,p) level of theory. BSSE- and ZPE-corrected binding energies, as well as selected hydrogen bond parameters of different conformers of these two complexes, are listed with the structures. Six low-energy structures of Ind···H_2O···Me_2Se are reported in Figure 3a. Interestingly, we have found that the most stable trimeric structure S1 mimics the single water-bridge-mediated Se hydrogen-bonded structural motif observed in protein crystal structures. Our comparison here is with a subset of the 4017 amino acid donors that are the indole group of tryptophan (99 cases). In the S1 structure, water acts as a hydrogen-bond donor to Se as well as a hydrogen-bond acceptor to the NH group of indole. The S1 structure of Ind···H_2O···Me_2Se qualitatively resembles specifically the structural motif of single water-bridge-mediated hydrogen-bonding interactions between Mse (residue number 83) and Trp (residue number 56) via H_2O (residue number 150) in a protein (PDB 3d5p, Figure 1c) named the putative glucan synthesis regulator of SMI1/KNR4 family. It can be noted that Se···O and N···O distances and ∠N···O···Se in the PDB 3d5p (Figure 1c) are 3.34 Å, 2.99 Å, and 108.47°, respectively, while the same parameters in the S1 structure of Ind···H_2O···Me_2Se are 3.43 Å, 2.98 Å, and 105°, respectively. Thus, the structural motif and hydrogen bond parameters for single water-mediated Se hydrogen-bonding interactions in the trimeric model structure (S1) and the PDB 3d5p have a close resemblance. The other higher-energy structures (S2, S3, S4, S5, and S6) of Ind···H_2O···Me_2Se are not stabilized through a water bridge, and there is a direct interaction between indole and Me_2Se.

Four low-energy structures of the Ind···$(H_2O)_2$···Me_2Se complex are listed in Figure 3b. The most stable tetrameric structure is a double water-bridge-mediated Se hydrogen-bonded structure denoted T1, where indole and Me_2Se interact with each other through a double water bridge.

3.4. IR Spectra of the Complexes. Figure 4c,e shows IR spectra in the N−H and O−H stretching regions by probing

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

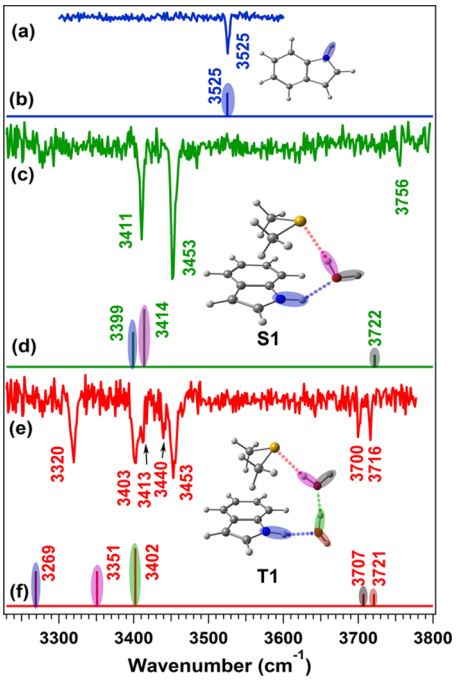


Figure 4. IR spectra of (a) indole, (c) Ind···H_2O···Me_2Se, and (e) Ind···$(H_2O)_2$···Me_2Se complexes measured using resonant ion dip infrared (RIDIR) spectroscopy. (b, d, f) Theoretical scaled IR stick spectra of the indole monomer, Ind···H_2O···Me_2Se (S1), and Ind···$(H_2O)_2$···Me_2Se (T1) complexes, respectively, calculated at the ωB97X-D/6-311++G(d,p) level of theory. Scaling factors of 0.9463 and 0.9378 have been used to correct the harmonic NH and OH stretching frequencies, respectively (see the Computational Methods section). Assignment of the IR bands is shown through color code provided in the theoretical IR spectra as well as the optimized structures.

the 35126 cm^{-1} band (0-0 + 35 cm^{-1}) of the Ind···H_2O···Me_2Se complex and the 35118 cm^{-1} band, respectively, observed in the electronic spectrum presented in Figure 2a. The two IR spectra are completely different in terms of the number of bands, and this observation indicates that the 35 118 cm^{-1} band cannot be due to a conformer of Ind···H_2O···Me_2Se trimer. The 35 118 cm^{-1} band could be rather due to a higher-order cluster. The IR spectrum of the indole monomer[35] in the N−H stretching region has been provided in Figure 4a to compare the N−H stretching frequency of the complexes with the free N−H stretching frequency.

Figure 4d,f shows theoretical harmonic scaled IR spectra of the most stable structures of the Ind···H_2O···Me_2Se (S1) and Ind···$(H_2O)_2$···Me_2Se (T1) complexes, respectively, obtained at the ωB97X-D/6-311++G(d,p) level of theory. The global minimum structures of the Ind···H_2O···Me_2Se (S1) and Ind···$(H_2O)_2$···Me_2Se (T1) complexes are provided with the theoretical IR spectra in Figure 4d,f, respectively. The N−H and O−H groups in the structures of the complexes and the corresponding theoretical IR bands are marked with specific colors for a clear understanding of the assignment of the experimental IR bands of the complexes. There is a reasonably good agreement of the experimental IR spectra presented in Figure 4c,e with the theoretical IR spectra of Ind···H_2O···Me_2Se (S1) and Ind···$(H_2O)_2$···Me_2Se (T1) complexes, respectively, shown in Figure 4d,f.

A comparison of the experimental IR spectra of Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se with the theoretical scaled IR spectra of their various structures is shown in Figures S2 and S3, respectively. It is quite clear from Figures S2 and S3 that the theoretical IR spectra of only the S1 structure of Ind···H_2O···Me_2Se and the T1 structure of Ind···$(H_2O)_2$···Me_2Se corroborate with the experimental IR spectra. IR spectra of all of the structures of the trimeric and tetrameric complexes are also calculated at the M06-2X/6-311++G(d,p) level of theory. Figures S4 and S5 show a comparison of the experimental IR spectra of the Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se complexes, respectively, with the theoretical scaled IR spectra of all of the structures calculated at the ωB97X-D/6-311++G(d,p) and M06-2X/6-311++G(d,p) levels of theory. Interestingly, there is a similar trend in the calculated frequencies obtained at both the levels of theory and thus the assignment of the observed trimeric and tetrameric complexes to the S1 and T1 structures, respectively, is unequivocal.

In Figure 4c, the 3756 cm^{-1} band is assigned to the free O−H frequency of H_2O of the Ind···H_2O···Me_2Se complex, while the 3453 and 3411 cm^{-1} bands are assigned to hydrogen-bonded O−H (O−H···Se) of H_2O and N−H (N−H···O) of indole, respectively. A red shift of 303 cm^{-1} in the O−H frequency of the O−H···Se hydrogen bond with respect to that of the free O−H indicates the presence of a very strong hydrogen-bond interaction there. The N−H stretch frequency of indole in this trimeric complex is red-shifted by 114 cm^{-1} compared to that of the indole monomer, while the red shift in the N−H stretch frequency in the indole···H_2O complex is reported to be 89 cm^{-1}.[35] In the case of Ind···$(H_2O)_2$···Me_2Se, two free O−H groups of the two water molecules appear at 3716 (free OH of indole-bound water) and 3700 cm^{-1} (free OH of Me_2Se-bound water). On the other hand, stretching frequencies of the two hydrogen-bonded bridged O−H groups (O−H···Se and O−H···O) of two water molecules are strongly coupled with each other and appear at 3403 and 3453 cm^{-1}, respectively. Assignment of these two O−H groups held together by a strongly hydrogen-bonded bridge is performed by theoretical calculation of deuterated complexes by replacing at a time one of the bridging hydrogens of the two water molecules in the T1 structure of the Ind···$(H_2O)_2$···Me_2Se complex (Figure S6, Table S1, and Figure S7). Broadening of the two bands at 3403 and 3453 cm^{-1} and appearance of additional two weak bands at 3413 and 3440 cm^{-1} could be due to strong coupling of the hydrogen-bonded bridged OH vibrations, which are close in frequencies.[50,51]

Both the experimental and the theoretical IR spectra as well as the energetics of different structures of Ind···H_2O···Me_2Se and Ind···$(H_2O)_2$···Me_2Se confirm the observation of the S1 and T1 structures, respectively, in the experiment. Our spectroscopic studies on the model complex Ind···H_2O···Me_2Se demonstrate the finding of single water-mediated Se hydrogen-bonded structural motif (S1 structure), which is abundant in the PDB. A 3D histogram showing the frequency of specific S1-type structural motif for single water-mediated Se hydrogen-bonding interactions retrieved from PDB has also been shown in Figure S8.

3.5. Natural Bond Orbital (NBO) Analysis. NBO second-order perturbation energy ($E^{(2)}$) between the lone pair orbitals of the hydrogen-bond acceptor atom (Y) and the σ^* orbital of the hydrogen-bond donor group (X−H) is a measure of the strength of an X−H···Y hydrogen bond.[40] The $E^{(2)}$ and red-shift ($\Delta\nu_{X-H}$) values in the X−H stretching frequencies of all of the hydrogen-bond interactions present in

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

the S1 trimer and T1 tetramer structures are provided with the NBO plots shown in Figure 5. Both $\Delta\nu_{X-H}$ and $E^{(2)}$ values

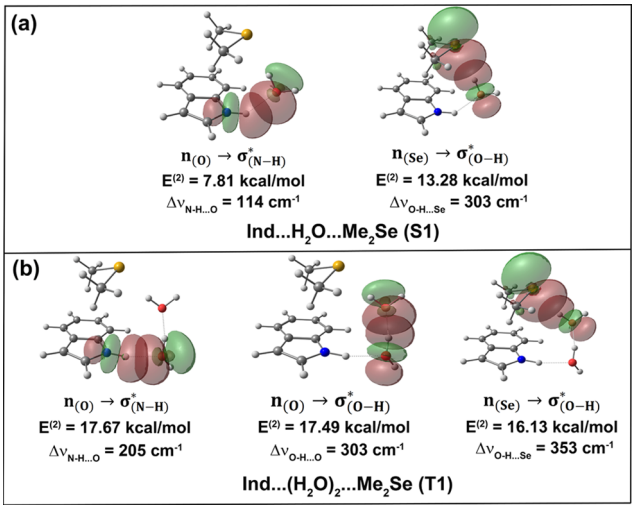


Figure 5. NBO views for the hydrogen-bonds present in the (a) S1 structure of Ind···H_2O···Me_2Se (N−H···O, O−H···Se) and (b) T1 structure of Ind···$(H_2O)_2$···Me_2Se (N−H···O, O−H···O, O−H···Se) complexes calculated at the ωB97X-D/6-311++G(d,p) level of theory. Red-shift values in the X−H stretching frequency ($\Delta\nu_{X-H}$) of the hydrogen bonds present in the trimer and the tetramer are also provided in the figure.

demonstrate that the Se-centered hydrogen-bond is of similar strength to that of the O-centered hydrogen bond. It is intriguing to note that the $E^{(2)}$ values for the O−H···Se and N−H···O hydrogen-bonds in the S1 structure of the Ind···H_2O···Me_2Se complex and the PDB structure (ID: 3d5p, Figure 1c) of a putative glucan synthesis regulator protein are qualitatively similar.

3.6. Energy Decomposition Analysis (EDA). Energy decomposition analysis (EDA) of the S1 structure of Ind···H_2O···Me_2Se and the T1 structure of Ind···(H2O)$_2$···Me_2Se is performed using the reduced variational space (RVS) method[42] to explain the observation of the unusually strong Se-centered hydrogen bond. It is expected that the contribution of the electrostatic component of the total interaction energy in the Se hydrogen-bonded complexes is relatively smaller. The form of the RVS-EDA interaction energy (ΔE_{Total}) partitioning is $\Delta E_{Total} = \Delta E_{ESX} + \Delta E_{POL} + \Delta E_{CT} + \Delta E_{MIX}$, where the ΔE_{ESX} term is equal to the sum of the electrostatic and exchange terms. ΔE_{POL} and ΔE_{CT} are the polarization and charge-transfer components, respectively. The mixing component, ΔE_{MIX}, of the RVS-EDA is the residual energy due to counterpoise (CP) correction.

Figure 6 shows the percentage components of the total interaction energies of the most stable conformer of Ind···H_2O···Me_2X as well as Ind···$(H_2O)_2$···Me_2X by varying X from O to S and Se using the RVS-EDA method. The green, red, and blue colors in the figure represent CT, POL, and ESX components, respectively. It is obvious that moving from the complexes of Me_2O to Me_2S and Me_2Se, % ESX (coulomb/exchange) decreases while % CT increases significantly. Thus, our finding demonstrates that the strength of the S- and Se-centered hydrogen bonds in terms of the IR red shift in the stretching frequency of the hydrogen-bond donor could be due to the significant contribution of the charge-transfer

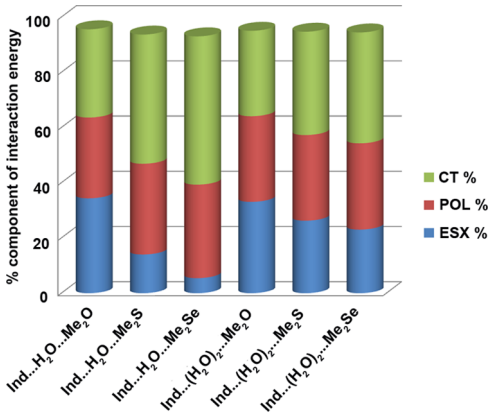


Figure 6. Percentage (%) component of the total interaction energies (kcal/mol) of the most stable structure of Ind···H_2O···Me_2X as well as Ind···$(H_2O)_2$···Me_2X (X = O, S, Se) obtained with the 6-31G basis set using the RVS-EDA method. CT and POL stand for charge transfer and polarization, respectively, while ESX is the sum of electrostatic and exchange terms.

component apart from the electrostatic and polarization interactions present there. A similar explanation has been provided recently by Das and co-workers for the observation of strong Se hydrogen-bonds.[25] We have also determined from localized molecular orbital energy decomposition analysis (LMO-EDA) method[43] that the percentage of the dispersion component (% DISP) of the total interaction energy is higher in the Me_2S and Me_2Se complexes compared to that in the Me_2O complexes (Table S2). It has been reported by Patwari and co-workers that the IR red shift in the stretching frequency of the hydrogen-bond donor is governed by electrostatic, polarization, and charge-transfer components of the interaction energy, while the overall stability of the complexes is controlled by the dispersion energy.[52]

4. CONCLUSIONS

Single water-mediated Se hydrogen-bonding interactions between amino acid residues in proteins have been investigated by a statistical analysis of the structures in the PDB. Gas-phase spectroscopy and quantum chemical calculations of model complexes consisting of indole, dimethyl selenide, and H_2O have been carried out to understand these interactions present in proteins. Mass-selected electronic and IR spectroscopy of the model complex Ind···H_2O···Me_2Se qualitatively mimics the nature, motif, and energetics of single water-mediated Se hydrogen-bonding interactions present in protein structures. Interestingly, we found that the number of Se hydrogen bonds involving water was almost twice than that of the direct Se-amino acid hydrogen-bonds in the PDB. Energy decomposition analysis of the total interaction energy of the complexes demonstrates that the charge-transfer interaction plays a significant role in the strength of the S- and Se-centered hydrogen bonds in terms of the IR red shift in the stretching frequency of the hydrogen-bond donor. This work, for the first time, reports a systematic PDB analysis of this water-mediated Se hydrogen-bonding interaction in proteins and understanding of this interaction through the spectroscopic study of a model complex.

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b04159.

Results from AIM calculations, comparison of experimental IR spectrum with theoretical IR spectra of all of the structures of the complexes, comparison of experimental IR spectrum with theoretical IR spectra of all of the structures of the complexes at different levels of theory, optimized geometries and theoretical vibrational frequencies of deuterium-substituted tetrameric complexes, 3D histogram showing frequency of specific S1-type structural motif for single water-mediated Se hydrogen-bonding interactions retrieved from PDB, Cartesian coordinates of the optimized structures of all of the structures of the trimer and tetramer, and list of PDB IDs for single water-mediated Se H-bonding interactions (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: biplabs@nehu.ac.in (B.S.).
*E-mail: madhusudhan@iiserpune.ac.in (M.S.M.).
*E-mail: a.das@iiserpune.ac.in (A.D.).

ORCID

Biplab Sarkar: 0000-0003-2640-4401
M. S. Madhusudhan: 0000-0002-2889-5884
Aloke Das: 0000-0002-2124-0631

Present Address

[‖]Department of Chemistry, University of Hawaii, Manoa, Honolulu, Hawaii 96822-2275, United States (S.K.S.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to acknowledge the financial support provided by IISER Pune, Science and Engineering Research Board (SERB), India (Grant No. EMR/2015/000486 and EMR/2015/002322) and a Wellcome-DBT India alliance senior fellowship to carry out this research. Computational support received from the PRITHVI supercomputing facility of IISER Pune as well as BRAF supercomputing facility of C-DAC, Pune, is acknowledged. We thank Dr. Himansu Biswal, NISER Bhubaneswar, and Dr. Shachi Gosavi, NCBS Bangalore, for a fruitful discussion on the results.

■ REFERENCES

(1) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer: Berlin, 1991.
(2) Anfinsen, C. B. Principles That Govern Folding of Protein Chains. *Science* **1973**, *181*, 223−230.
(3) Karshikoff, A. *Non-covalent Interactions in Proteins*; Imperial College Press, 2006.
(4) Saenger, W. Structure and dynamics of water surrounding biomolecules. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 93−114.
(5) Finney, J. L.; Eley Daniel, D.; Richards Rex, E.; Franks, F. The organization and function of water in protein crystals. *Philos. Trans. R. Soc., B* **1977**, *278*, 3−32.
(6) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. Affinities of Amino-Acid Side-Chains for Solvent Water. *Biochemistry* **1981**, *20*, 849−855.
(7) Buckle, A. M.; Cramer, P.; Fersht, A. R. Structural and energetic responses to cavity-creating mutations in hydrophobic cores: Observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry* **1996**, *35*, 4298−4305.
(8) Sleigh, S. H.; Tame, J. R. H.; Dodson, E. J.; Wilkinson, A. J. Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyllysine. *Biochemistry* **1997**, *36*, 9747−9758.
(9) Takano, K.; Funahashi, J.; Yamagata, Y.; Fujii, S.; Yutani, K. Contribution of water molecules in the interior of a protein to the conformational stability. *J. Mol. Biol.* **1997**, *274*, 132−142.
(10) Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B. S.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. Structure of Hiv-1 Protease with Kni-272, a Tight-Binding Transition-State Analog Containing Allophenylnorstatine. *Structure* **1995**, *3*, 581−590.
(11) Wang, Y. X.; Freedberg, D. I.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Bhat, T. N.; Erickson, J. W.; Torchia, D. A. Bound water molecules at the interface between the HIV-1 protease and a potent inhibitor, KNI-272, determined by NMR. *J. Am. Chem. Soc.* **1996**, *118*, 12287−12290.
(12) Helms, V.; Wade, R. C. Thermodynamics of Water Mediating Protein-Ligand Interactions in Cytochrome P450cam - a Molecular-Dynamics Study. *Biophys. J.* **1995**, *69*, 810−824.
(13) Li, Z.; Lazaridis, T. Water at biomolecular binding interfaces. *Phys. Chem. Chem. Phys.* **2007**, *9*, 573−581.
(14) Williams, M. A.; Goodfellow, J. M.; Thornton, J. M. Buried Waters and Internal Cavities in Monomeric Proteins. *Protein Sci.* **1994**, *3*, 1224−1235.
(15) Jiang, L.; Kuhlman, B.; Kortemme, T. A.; Baker, D. A "solvated rotamer" approach to modeling water-mediated hydrogen bonds at protein-protein interfaces. *Proteins* **2005**, *58*, 893−904.
(16) Fenwick, R. B.; Oyen, D.; Dyson, H. J.; Wright, P. E. Slow Dynamics of Tryptophan-Water Networks in Proteins. *J. Am. Chem. Soc.* **2018**, *140*, 675−682.
(17) Otting, G.; Liepinsh, E.; Wuthrich, K. Protein Hydration in Aqueous-Solution. *Science* **1991**, *254*, 974−980.
(18) Matthews, B. W.; Ernst, J. A.; Clubb, R. T.; Zhou, H. X.; Gronenborn, A. M.; Clore, G. M.; et al. Use of Nmr to Detect Water within Nonpolar Protein Cavities - Response. *Science* **1995**, *270*, 1847−1849.
(19) Clore, G. M.; Bax, A.; Wingfield, P. T.; Gronenborn, A. M. Identification and Localization of Bound Internal Water in the Solution Structure of Interleukin-1-Beta by Heteronuclear 3-Dimensional H-1 Rotating-Frame Overhauser N-15-H-1 Multiple Quantum Coherence Nmr-Spectroscopy. *Biochemistry* **1990**, *29*, 5671−5676.
(20) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.
(21) Scheiner, S. *Hydrogen Bonding. A Theoretical Perspective*; Oxford University Press: Oxford, 1997.
(22) Hendrickson, W. A.; Horton, J. R.; Lemaster, D. M. Selenomethionyl Proteins Produced for Analysis by Multiwavelength Anomalous Diffraction (Mad) - a Vehicle for Direct Determination of 3-Dimensional Structure. *EMBO J.* **1990**, *9*, 1665−1672.
(23) Mundlapati, V. R.; Sahoo, D. K.; Ghosh, S.; Purame, U. K.; Pandey, S.; Acharya, R.; Pal, N.; Tiwari, P.; Biswal, H. S. Spectroscopic Evidences for Strong Hydrogen Bonds with Selenomethionine in Proteins. *J. Phys. Chem. Lett.* **2017**, *8*, 794−800.
(24) Biswal, H. S.; Bhattacharyya, S.; Bhattacherjee, A.; Wategaonkar, S. Nature and strength of sulfur-centred hydrogen bonds: laser spectroscopic investigations in the gas phase and quantum-chemical calculations. *Int. Rev. Phys. Chem.* **2015**, *34*, 99−160.
(25) Mishra, K. K.; Singh, S. K.; Ghosh, P.; Ghosh, D.; Das, A. The nature of selenium hydrogen bonding: gas phase spectroscopy and quantum chemistry calculations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24179−24187.
(26) Kumar, S.; Kaul, I.; Biswas, P.; Das, A. Structure of 7-Azaindole···2-Fluoropyridine Dimer in a Supersonic Jet: Competition

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX

between N−H···N and N−H···F Interactions. *J. Phys. Chem. A* **2011**, *115*, 10299−10308.
(27) Kumar, S.; Das, A. Effect of acceptor heteroatoms on pi-hydrogen bonding interactions: A study of indole...thiophene heterodimer in a supersonic jet. *J. Chem. Phys.* **2012**, *137*, No. 094309.
(28) Kumar, S.; Singh, S. K.; Calabrese, C.; Maris, A.; Melandri, S.; Das, A. Structure of saligenin: microwave, UV and IR spectroscopy studies in a supersonic jet combined with quantum chemistry calculations. *Phys. Chem. Chem. Phys.* **2014**, *16*, 17163−17171.
(29) Page, R. H.; Shen, Y. R.; Lee, Y. T. Infrared−ultraviolet double resonance studies of benzene molecules in a supersonic beam. *J. Chem. Phys.* **1988**, *88*, 5362−5376.
(30) Das, A.; Mahato, K. K.; Chakraborty, T. Jet spectroscopy of van der Waals dimers of 1-methoxynaphthalene: A laser induced fluorescence study. *J. Chem. Phys.* **2001**, *114*, 8310−8315.
(31) Frisch, M. J. et al. *Gaussian 16*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2016.
(32) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(33) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminf.* **2012**, *4*, 17.
(34) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(35) Carney, J. R.; Hagemeister, F. C.; Zwier, T. S. The hydrogen-bonding topologies of indole-$(water)_n$ clusters from resonant ion-dip infrared spectroscopy. *J. Chem. Phys.* **1998**, *108*, 3379−3382.
(36) Watanabe, T.; Ebata, T.; Tanabe, S.; Mikami, N. Size-selected vibrational spectra of phenol-$(H_2O)_n$ (n = 1−4) clusters observed by IR-UV double resonance and stimulated Raman-UV double resonance spectroscopies. *J. Chem. Phys.* **1996**, *105*, 408−419.
(37) Benedict, W. S.; Plyler, E. K. Absorption spectra of water vapor and carbon dioxide in the region of 2.7 microns. *J. Res. Natl. Bur. Stand.* **1951**, *46*, 246−265.
(38) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Clarendon Press: Oxford, U.K., 1990.
(39) Biswal, H. S.; Chakraborty, S.; Wategaonkar, S. Experimental evidence of O−H—S hydrogen bonding in supersonic jet. *J. Chem. Phys.* **2008**, *129*, No. 184311.
(40) Weinhold, F.; Landis, C. R. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press: Cambridge, 2005.
(41) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429−1437.
(42) Stevens, W. J.; Fink, W. H. Frozen Fragment Reduced Variational Space Analysis of Hydrogen-Bonding Interactions - Application to the Water Dimer. *Chem. Phys. Lett.* **1987**, *139*, 15−22.
(43) Su, P.; Li, H. Energy decomposition analysis of covalent bonds and intermolecular interactions. *J. Chem. Phys.* **2009**, *131*, 14102−14115.
(44) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(45) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(46) Wang, G.; Dunbrack, R. L., Jr. PISCES: a protein sequence culling server. *Bioinformatics* **2003**, *19*, 1589−1591.
(47) Word, J. M.; Lovell, S. C.; LaBean, T. H.; Taylor, H. C.; Zalis, M. E.; Presley, B. K.; Richardson, J. S.; Richardson, D. C. Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* **1999**, *285*, 1711−1733.
(48) Bondi, A. van der Waals Volumes and Radii. *J. Phys. Chem. A* **1964**, *68*, 441−451.
(49) *The PyMOL Molecular Graphics System*, version 2.0; Schrödinger, LLC.
(50) Zwier, T. S. Laser Spectroscopy of Jet-Cooled Biomolecules and Their Water-Containing Clusters: Water Bridges and Molecular Conformation. *J. Phys. Chem. A* **2001**, *105*, 8827−8839.
(51) Carney, J. R.; Zwier, T. S. Infrared and Ultraviolet Spectroscopy of Water-Containing Clusters of Indole, 1-Methylindole, and 3-Methylindole. *J. Phys. Chem. A* **1999**, *103*, 9943−9957.
(52) Dey, A.; Mondal, S. I.; Sen, S.; Ghosh, D.; Patwari, G. N. Electrostatics determine vibrational frequency shifts in hydrogen bonded complexes. *Phys. Chem. Chem. Phys.* **2014**, *16*, 25247−25250.

DOI: 10.1021/acs.jpca.9b04159
J. Phys. Chem. A XXXX, XXX, XXX−XXX