THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

# Combination of Resonance Raman Spectroscopy and Docking Simulations to Study the Nonspecific Binding of a Free-Base Porphyrin to a Globular Protein

James E. Parker,[†,‡] Robert J. Thomas,[§] Dayla Morisson,[∥] and Lorenzo Brancaleon*[,‡]

[†]General Dynamics Information Technology, San Antonio, Texas 78234, United States

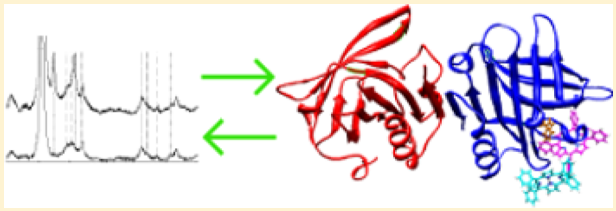
[‡]Department of Physics and Astronomy, The University of Texas at San Antonio, San Antonio, Texas 78249, United States

[§]Optical Radiation Bioeffects Branch, Bioeffects Division, Air Force Research Laboratory, Fort Sam Houston, Texas 78234, United States

[∥]Department of Physics, The University of Texas at Arlington, Arlington, Texas 76019, United States

Supporting Information

**ABSTRACT:** Understanding the conformational changes induced by small ligands noncovalently bound to proteins is a central problem in biophysics. We focus on the binding location of the water-soluble porphyrin, meso-tetrakis (*p*-sulfonatophenyl) porphyrin, to a globular protein, *β*-lactoglobulin, which has been observed to partially unfold when irradiated by laser light. Identifying the binding location is necessary to determine the mechanism of action as well as the atoms and residues involved in the photoinduced partial unfolding. Such atomic details are typically investigated by nuclear magnetic resonance or X-ray crystallography. However, for biomolecules in solution at the low concentrations (*μ*M) required to deliver uniform laser irradiation, these traditional techniques do not currently provide sufficient information, and one must rely upon less direct spectroscopic methods. We describe a method that uses resonance Raman spectroscopy and density functional theory (DFT) to select the most likely binding configuration among a set of solutions yielded by computational docking algorithms. This methodology may be generalized to use with other ligand−protein complexes where the ligand structure is amenable to DFT simulations.



## INTRODUCTION

Lasers and other light sources have been used to trigger protein conformational changes via diffusion-limited mechanisms, where the typical diffusion length is several orders of magnitude larger than the diameter of the proteins.[1−3] Conversely, in the past few years, other studies have focused on localized photoinduced protein effects led by photoisomerization of a covalently linked ligand.[4,5] Our group has demonstrated that partial protein unfolding can be prompted through the use of photoactive, noncovalently attached ligands via the mediation of a different type of photophysical/photochemical mechanism.[6] This approach could lead to future protein−dye molecule complexes designed to optimize photoinduced manipulation of proteins in order to (1) induce non-native properties of polypeptides, (2) engineer hybrid biomaterials, and (3) study the effects of laser irradiation in tissues at the molecular level. The model used to demonstrate the effects of dye irradiation on globular proteins was provided by *β*-lactoglobulin (BLG) bound to meso-tetrakis (*p*-sulfonatophenyl) porphyrin (TSPP)[7] or protoporphyrin IX (PPIX).[8] BLG is a small (162 amino acids) globular protein found mostly in a dimeric form and undergoes a conformational change, the Tanford transition, which involves a rotation of the EF loop and the deprotonation of a glutamic acid (Glu) residue in position 89 located underneath the EF loop.[9] We showed that both TSPP and PPIX trigger laser-induced conformational changes at pH values above the Tanford transition, whereas, under identical light exposure and equivalent ligand binding characteristics, there is no conformational change at lower pH.[7,8] The increased solvent accessibility of BLG and the weakening of some hydrogen bonds at alkaline pH[9] would favor the observed photoinduced unfolding.

Despite the success in triggering and measuring the porphryin-induced partial unfolding, several aspects of the model remain to be characterized; among them is the location of the bound porphyrin within the protein. This manuscript will focus on a method to establish the binding location of the water-soluble TSPP on BLG. This information is necessary to determine the mechanism of action as well as the atoms and residues involved in the photoinduced partial unfolding. Such atomic details are typically investigated by nuclear magnetic resonance (NMR) or X-ray crystallography. However, for biomolecules in solution at the low concentrations (*μ*M) required to deliver uniform laser irradiation, these traditional

**Received:** May 4, 2012
**Revised:** August 10, 2012
**Published:** August 13, 2012





dx.doi.org/10.1021/jp304310z | J. Phys. Chem. B 2012, 116, 11032−11040

techniques do not currently provide sufficient information, and one must rely upon less direct spectroscopic methods. We have, therefore, adopted the multifaceted approach described in this manuscript that uses a combination of experimental techniques (i.e., resonance Raman (RR) spectroscopy) and computational methods to reconstruct the protein−ligand complex at atomic-level detail. The use of RR spectroscopy, where the laser wavelength is within the absorption band of the ligand, allows us to focus on the vibrational modes of the ligand even with the protein present.[10] The differences in frequency and intensity of the Raman modes can be used to probe in-plane (*ip*) and out-of-plane (*oop*) distortions of the bonds in the ligand induced by the proximity of the protein. To quantitatively relate these distortions to the observed vibrational spectra, a computational model, which incorporates DFT to develop a complete set of normal modes[11] from the Hessian matrix, is coupled with molecular mechanics docking simulations. Typically, Raman spectroscopy only detects a subset of the possible lines due to selection rules.[12] This produces an underdetermined system of possible solutions for the ligand−protein complex. Nevertheless, other authors have shown that only a few lines are required to generate the majority of the distortions observed in real molecules.[13] Why are these distortions so important? Their relevance relies on the assumption that unique ligand distortions are induced by binding to a site; therefore, they could be "markers" for a specific binding location. The current manuscript shows a method that will use RR spectra and DFT to select the most likely binding configuration among a set of solutions yielded by computational docking. The importance of this methodology is that it could be generalized to other ligand−protein complexes where the ligand structure is amenable to DFT simulations.

## EXPERIMENTAL METHODS

**Sample Preparation.** TSPP (Frontier Scientific Inc., Logan, UT) and BLG (Sigma Aldrich, St. Louis, MO) were used without further purification. Solutions were prepared in 10 mM phosphate buffer at pH 9.0. The pH was adjusted by adding small amounts of 0.1 M NaOH. This pH level was chosen to replicate the conditions in which conformational changes of BLG were observed.[7] TSPP and BLG were dissolved directly into the buffer, and the final concentrations were obtained spectroscopically using $\varepsilon_{413} = 5.1 \times 10^5\ M^{-1}\ cm^{-1}$ for TSPP and $\varepsilon_{280} = 1.7 \times 10^{-4}\ M^{-1}\ cm^{-1}$ for BLG.[7]

**UV−vis Absorption Spectroscopy.** TSPP samples were placed in 1 cm path length, 1.4 mL, four-clear-sided quartz cells (Starna Cells, Atascadero, CA) for spectra collection using a dual-beam Cary 5 UV spectrophotometer (Agilent, Santa Clara, CA) at a scan rate of 600 nm/min, 0.1 s averaging time, and 1 nm bandwidth. Spectra were recorded from 350 to 700 nm and corrected with baseline subtraction.

**Resonant Raman Spectroscopy.** Laser excitation of 15.0 mW, measured at the sample, was provided by a Coherent Innova Krypton ion laser (Coherent, Santa Clara, CA) at 413.1 nm corresponding to the peak absorption of the Soret band in TSPP.[14] The scattered light was collected and dispersed by a Spex 1404 double monochromator (Horiba Jobin Yvon, Edison, NJ) equipped with an Andor iDus 420A-BV back-illuminated 1024 × 256 pixel CCD camera (Andor, Belfast, United Kingdom). Spectra were digitized and stored using LabSpec (Horiba Jobin Yvon, Edison, NJ). RR spectra were obtained using a 90° offset scattering geometry from an aqueous solution of 2.9−3.1 $\mu$M TSPP and 9.3 $\mu$M BLG in 1 cm path length, 1.4 mL, four-clear-sided quartz flow cells (Starna Cells, Atascadero, CA). The spectra are collected in nine different spectral windows each 250 $cm^{-1}$ wide, manually overlapped to provide a complete Raman spectra from 30 to 1600 $cm^{-1}$. The fluorescence background was subtracted utilizing an automated method developed by Lieber and Mahadevan-Jansen.[15] The spectra were corrected for the instrument response through a multiple line calibration process outlined by McCreery.[16] The minimal SNR for the Raman peaks was 40:1. Each peak was fit using a Levenberg−Marquardt algorithm to a Voigt line shape as implemented in the program fityk-0.8.9[17] to determine the spectral peak and width. The nearly 1:3 ratio of TSPP to BLG was chosen to maximize the probability that only one TSPP molecule was bound with a given BLG molecule.[6] To ensure the BLG−TSPP complex did not undergo photobleaching during laser exposure, 11 mL of TSPP−BLG solution was prepared at the desired concentration and circulated in a closed loop through a flow cell at 5 mL/min. Absorption spectra were collected in a separate flow cell every 20 min using a USB2000+UV−vis miniature spectrometer (Ocean Optics, Dunedin, FL). The consistency of the optical density and position of the absorption peak at 413 nm were taken as proof that photobleaching did not occur during the collection of the RR spectra.

## THEORETICAL AND COMPUTATIONAL MODELS

**Atomic Structure.** The initial configuration of the atomic structure of TSPP (Figure 1) was generated using the molecular editing tools Avogadro[18] and Maestro.[19] The phenyl groups have been rotated 90° into the plane of the page for ease of display (Figure 1), but in simulations, the phenyls are oriented perpendicular to the plane of the macrocycle as dictated by the lowest energy configuration of the molecule.[20] Each carbon atom along the periphery of the molecule is bonded to a hydrogen atom (not shown in the image). The molecule was initially built without the interior hydrogens or sulfonic acid groups and constrained to a $D_{4h}$ symmetry. Then, hydrogen atoms were added to nitrogens in pyrrole units 1 and 3 (opposing sides of the macrocycle), reducing the symmetry to a $D_{2h}$ point group. Finally, during the geometry optimization, the molecule was distorted from its planar symmetry to a $D_2$

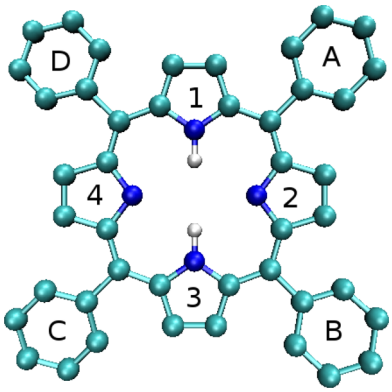



**Figure 1.** Structure of TSPP. The hydrogen atoms along the periphery of the molecule have been removed to avoid clutter. Carbon atoms are depicted in light blue, nitrogen atoms are dark blue, and hydrogen is white. The numbers 1−4 are used to identify specific pyrrole groups, and the letters A−D are used to identify phenyl groups.


dx.doi.org/10.1021/jp304310z | J. Phys. Chem. B 2012, 116, 11032−11040

symmetry in order to remove negative vibrational modes and find an approximate ground state. Partial charges were assigned using the AM1-BCC model[21] and an additional 0.5 $e^-$ was added to the terminus of each phenyl group to make the net charge −2. A dimeric structure for BLG was obtained from the Research Collaboratory for Structural Bioinformatics (RSCB) Protein Data Bank (PDB) as PDB id 1B8E,[22] which is a monomer; however, the published file provides translation and rotation tensors to generate a dimeric form of BLG from the monomer unit. A second dimer structure, PDB id 1BEB, valid at pH 6.5[23] was also used for comparisons but not in the docking calculations.

**Docking.** Two computational docking algorithms, Dock 6.4[24] and AutoDock 4.2,[25] were used to find the bound configuration of ligand−protein complexes based upon molecular mechanics (MM) computations of free energies. For Dock, we utilized the simplex search algorithm, and for Autodock, we used the Genetic search algorithm. Copies of our input files which provide specifics on the search parameters can be found in the Supporting Information. The chosen scoring function was the grid score plus the internal energy. By incorporating the internal energy, the score accounts for the bending of the macrocycle which is rigid during the docking. Larger distortions will lead to increasingly large internal energy values that are unfavorable and smoothly terminate the search of possible configurations, while the grid score accounts for the nonbonded interaction between the ligand and protein. To search the possible conformation space in reasonable time, the docking algorithms make the simplifying assumptions that the entire protein and the aromatic bonds (if any) in the ligand are rigid. This restriction, in porphyrins, prevents any distortions of the atoms in the macrocycle from the initial planar configuration, which are known to occur and are key to understanding their biological and chemical functions.[26−28] As the deformation of the ring is expected to be a critical parameter in understanding binding and the subsequent photophysics of the bound porphyrin,[29] the constraint imposed by the docking algorithm may not adequately sample our desired configuration space. Various methods have been proposed to allow systematic breaking of the ring into pieces that the computational engine can manipulate to allow a more thorough search of the conformational space.[30,31] However, these methods generally fail for porphyrins to return the bond distances in the final structure to be within 10% of the expected values required to close the macrocycle again. In this model, we propose a conceptually simpler but computationally expensive algorithm that provides multiple conformations of the macrocycle as the input of the docking program and compare the resulting binding scores to determine the favored ligand conformation and its location. Such a model can be implemented by using the experimental and computational data present in eq 11 and varying the distortion scaling factor $a$ to generate the associated porphyrin structures. Each structure corresponding to a value of $a$ is then docked separately to BLG in order to find the minimum docking energy. Ideally, a single value of $a$ should be found to scale the distortions of the ligand that represent the best docked configuration of the ligand consistent with the RR spectra.

A possible complication arises from the effects pH may have on the protein. The experimental TSPP−BLG complex is in a pH 9.0 solution, and the PDB file, 1B8E, for BLG characterizes the protein at pH 7.9. To assign appropriate protonation states to the residues in BLG at pH 9.0, the programs PDB2PQR[32] and PropKa[33] were used with the CHARMM force-field. The resulting protein structure was identical to the simulated protein at pH 7.9. In fact, the simulated structures generated by the PDB2PQR program were insensitive to any changes in pH between 7 and 9. This was because the only amino acid with a p$K_a$ in the range 7−9 is cysteine, and all 10 cysteine residues in BLG are buried or bound in the interior of the protein. Because the residues are not solvent accessible, the program does not change the protonation state and generates identical structures for pH 7 and 9. However, the glutamic acid residue Glu89 is known to deprotonate during the Tanford transition. By comparing the position of the EF loop between the published structures of BLG in the PDB database at pH 6.5 (1BEB) and pH 7.9 (1B8E) and the change in the solvent accessible areas at each pH,[9] we estimated that the difference in position of the EF loop at pH 9.0 would be minimally different than the pH 7.9 structure, and thus, the 1B8E structure was used for docking. The simulated BLG structure has a net charge of 16 $e^-$, and the charged surface is shown in Figure 2.

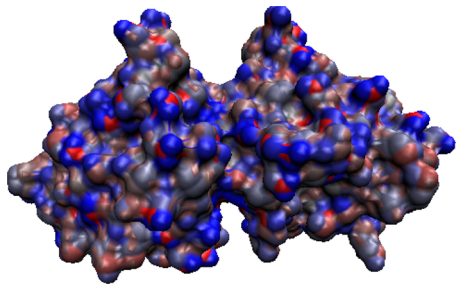


**Figure 2.** Charged surface of BLG in a pH 9 solution as computed by PDB2PQR. Negatively charged areas are blue, positive areas are red, and neutral are gray.

Another challenge represented by our system is that BLG is not known to possess a unique binding site for TSPP. Since the binding is nonspecific, the entire protein must be checked to find the best docking solution. As the protein fills a volume of approximately 96 000 Å$^3$ and the recommended resolution for docking simulations is 0.027 Å$^3$, the required problem space was larger than would fit in the memory of the available computers. To resolve this conflict, the problem space was divided into four rectangular regions, each composed of approximately 2 million voxels, which required 4 GB of memory to store. The computational regions overlapped by 10 Å (60% of the length of TSPP) in any direction along the surface of BLG to ensure that the surface of the BLG near the boundaries of a computational region were adequately sampled. To improve the sampling of possible conformations, 500 docking simulations (each with a unique random seed) were computed for each region. A search for the lowest energy conformation among all solutions yielded the orientation and location that was chosen to represent the best conformer for the TSPP−BLG complex.

**Density Functional Theory.** A significant simplication of the molecular model was made by removing the sulfonic acid groups ($SO_3^-$) at the terminus of each phenyl in TSPP and replacing them with a hydrogen atom; i.e., the simulation was reduced to a molecule of tetraphenylporphyrin (TPP) involving only hydrogen, carbon, and nitrogen atoms. Such simplification is justified by the fact that the laser wavelength (413 nm) primarily excites the porphyrin macrocycle; therefore, the

vibrational modes of the macrocycle will dominate the RR spectra, which will minimize the effects of the groups appended to the phenyl *meso* substituents. This is consistent with other reports on the substituent effects on porphyrins.[34,35] This simplification reduces the ligand molecule from 90 to 78 atoms, from 482 to 322 electrons, and removes the sulfur atoms, which would require a larger basis set to simulate. The combined effect is to reduce the computational time required for the DFT calculations without affecting the more relevant vibrational modes for our study (i.e., the ones associated with the macrocycle). The TPP molecule was then structurally manipulated by inducing saddling and rippling of the skeletal ring and twisting of the meso-phenyl rings by 0, 45, 90, and 100° from the plane of the skeletal ring. Each structure was geometrically optimized, except for a restriction of the point group symmetry, in a solution of water at neutral pH with an induced charge varying from 0 to −4 using the Becke−Lee−Young−Parr (B3LYP) hybrid-DFT level of theory with the 6-31+G basis set available with the Gaussian 09 suite.[36] A normal-mode analysis and Raman spectra calculation was conducted based upon the geometry optimized structure and a read-in wavelength of 413.1 nm to match that used for the experimental Raman spectra.

The resulting $3N - 6 = 228$ modes are classified according to the irreducible representations of the $D_2$ point group of the distorted TSPP molecule as

$$\Gamma = 57A + 58B_1 + 56B_2 + 57B_3 \tag{1}$$

For each irreducible representation, the modes are arranged in order of increasing frequency to enable the definition of basis sets at different orders that span subspaces of the full 228-dimensional orthonormal space in the internal coordinate system. Additionally, using correlation matrices between the $D_{2h}$ and $D_2$ symmetry groups, the modes can be further classified as *ip* or *oop* vibrations which helps refine the correlations with the Raman spectra. The first-order basis set ($m = 1$) contains only one mode, i.e., the lowest frequency mode from each of the irreducible representations. The $m$th-order basis set contains $m$ modes from each irreducible representation. Jarzecki and Spiro[13] have shown that the first-order basis can be sufficient to describe most of the distortions present in the molecule. Here, we use the sixth-order basis which provides 24 modes to provide a sufficient description of the distortions. Additionally, although Raman activity and eigenfrequencies may be perturbed due to the exclusion of the sulfonic acid groups, the main purpose of the DFT computation is to generate the Cartesian displacements of each atom for each vibrational mode, which result from diagonalization of the eigenvalue problem. However, as Cremer and Kraka have shown,[37] one can study the behavior of the full molecule from the sum of the parts, allowing the removal of the sulfonic groups with minimal perturbations of the computed displacements of the normal modes. Lastly, during the geometry optimization, the molecule had 1 mode out of 228 that exhibited a slight negative frequency, indicating that a true ground state had not been found. The negative mode was close enough to zero that Gaussian associated it with a translation or rotation mode, and thus, the geometry was considered acceptable. Thus, for each normal mode of the chosen configuration of the TSPP molecule, the DFT solution yields an eigenfrequency, symmetry group representation, Raman activity, and the normalized Cartesian displacements of each atom.

**Relating Normal Modes to Distortions.** While focusing on the energy gradients in the Franck−Condon region, Jarzecki and Spiro[13] outlined the relation of the Raman spectral intensity to structural changes in the molecule. According to this model, the RR intensity of the $k$th mode, $I_k$, is proportional to the square of the displacement along the $k$th normal mode of the minimum in the excited state potential surface from the ground state minimum, $\Delta_k$, and the modulus squared of the Kramers−Kronig transform of the absorption band, $\Phi(\omega)$[38,39]

$$I_k \propto \omega_L(\omega_L - \omega_k)^3 \Delta_k^2 \, |\Phi(\omega_L) - \Phi(\omega_L - \omega_k)|^2 \tag{2}$$

where $\omega_L$ is the frequency of the laser excitation and $\omega_k$ is the frequency of the vibrational mode. The Kramers−Kronig transform of the absorption band is defined as

$$\Phi(\omega) = P \int_{-\infty}^{\infty} d\omega' I(\omega')(\omega' - \omega)^{-1} + i\pi I(\omega) \tag{3}$$

where $P$ denotes the principal part and $I(\omega)$ is a normalized line shape function for the absorption spectra denoted by

$$I(\omega) = \frac{\alpha(\omega)}{\omega}\left[\int_{-\infty}^{\infty} d\omega' \alpha(\omega')/\omega'\right]^{-1} \tag{4}$$

where $\alpha(\omega)$ is the optical absorption. Lee and Heller[40] and Heller et al.[41] related the displacements, $\Delta_k$, to distortions in the internal atomic coordinates, $\Delta X_j$, by

$$\Delta_k \propto \omega_k^{1/2} \sum_j^{3N-6} L_{kj}^{-1} \Delta X_j \tag{5}$$

where $L_{kj}^{-1}$ are the elements of the inverse of $L$ which is a matrix determined by the ground-state eigenvectors of the normal mode problem. Substituting eq 5 into eq 2 yields

$$I_k \propto \omega_L(\omega_L - \omega_k)^3 \omega_k \left(\sum_j^{3N-6} L_{kj}^{-1} \Delta X_j\right)^2 |\Phi(\omega_L) - \Phi(\omega_L - \omega_k)|^2 \tag{6}$$

which can be used to find the relative intensity between two modes $k$ and $k'$ in a Raman spectrum obtained from the same laser source using the following ratio

$$\frac{I_{k'}}{I_k} = \frac{(\omega_L - \omega_{k'})^3}{(\omega_L - \omega_k)^3}\frac{\omega_{k'}}{\omega_k}\frac{\left(\sum_j^{3N-6} L_{k'j}^{-1} \Delta X_j\right)^2}{\left(\sum_j^{3N-6} L_{kj}^{-1} \Delta X_j\right)^2} \frac{|\Phi(\omega_L) - \Phi(\omega_L - \omega_{k'})|^2}{|\Phi(\omega_L) - \Phi(\omega_L - \omega_k)|^2} \tag{7}$$

The term $\sum_j^{3N-6} L_{kj}^{-1} \Delta X_j$ can be related to a sum of distortions over all symmetry modes present in the molecule[13]

$$\sum_j^{3N-6} L_{kj}^{-1} \Delta X_j = \sum_{\Gamma,m}^{N_\Gamma, n_m} d_m^\Gamma = d_k \tag{8}$$

where $\Gamma$ is a irreducible representation, $N_\Gamma$ are the number of different irreducible representations, $n_m$ are the number of modes with a given symmetry $\Gamma$, and $d_m^\Gamma$ is the distortion along the $m$th normal mode with symmetry $\Gamma$. Using eqs 7 and 8, one can solve for the relative distortions in terms of the Raman intensities and find that

Table 1. Mode Assignments and Fitted Raman Line Parameters for the RR Spectra of the TSPP−BLG Complex[a]

| mode | freq ($cm^{-1}$) | fwhm ($cm^{-1}$) | symmetry | DFT mode | reference |
|---|---|---|---|---|---|
| $\gamma_{19}$ | 123.1 | 12.4 | $B_2/B_3$ | 16, 17 | A |
| $\nu_{35}$ | 167.6 | 5.7 | $B_1$ | 23, 24 | A |
| $\phi''_{10}$ | 215.5 | 12.0 | $B_2/B_3$ | 25, 26, 28 | A |
| $\nu_8$ | 312.3 | 5.5 | $A$ | 34, 35 | B |
| $\nu_8$ | 348.4 | 6.1 | $A$ | 38 | B |
| phenyl | 393.8 | 17.3 | | 40, 43 | A |
| $\sigma_{3/3'}$ | 412.8 | 6.9 | $A$ | 42 | A |
| $\nu_{49}$ | 439.1 | 11.1 | $B_2/B_3$ | 41 | A |
| phenyl | 627.6 | 10.8 | $A$ | 60, 61 | A |
| $\phi_9$ | 646.8 | 11.3 | $A$ | 59 | A |
| $\phi_{9'/9''}$ | 675.4 | 2.3 | $B_1/B_2/B_3$ | 66−69, 71−73 | A |
| $\gamma(C_b–H)$ | 717.4 | 9.4 | | 78−80 | B |
| $\gamma_{20}$ | 734.7 | 7.1 | $B_2/B_3$ | 78−80 | A |
| $\gamma_5$ | 808.9 | 7.6 | $B_1$ | 84, 87 | A |
| $\nu_{32}$ | 863.6 | 6.4 | $B_1$ | 94 | A |
| $\nu_7$ or $\nu_{48}$ | 886.8 | 3.7 | $A/B_2/B_3$ | 101, 102 | A and B |
| $\nu_6$ | 966.3 | 3.7 | $A$ | 109 | A |
| $\nu_{15}$ or $\phi_8$ | 1004.4 | 3.1 | $A$ | 119, 124 | A |
| $\nu_9$ or $\nu_{17}$ | 1085.6 | 7.2 | $A$ | 130 | A and B |
| N−H bend | 1145.8 | 5.5 | | 146, 147 | A |
| $\nu_{51}$ | 1215.9 | 3.8 | $B_2/B_3$ | 151, 152 | A |
| $\nu_1$ | 1235.0 | 5.5 | $A$ | 149 | A |
| $\nu_{26}$ | 1239.9 | 6.7 | $B_1$ | 153 | A |
| C−H or C−N vib | 1247.6 | 6.7 | | not used | C |
| $\nu_{36}$ | 1276.2 | 5.0 | $B_2$ | 166, 169 | A |
| $\nu_{12}$ | 1295.3 | 9.3 | $A$ | 168 | A |
| $\nu_4$ | 1366.5 | 5.2 | $A$ | 170 | A |
| $\nu_{29}$ | 1384.1 | 4.8 | $A$ | 178 | A |
| $\psi_4$ or $\nu(C_\beta–C_\beta)$ | 1440.3 | 3.8 | $A$ | 185, 187 | A and B |
| $\nu_3$ | 1462.1 | 13.3 | $A$ | 186 | A |
| $\nu_{11}$ | 1498.9 | 8.4 | $A$ | 193 | A |
| $\nu_2$ | 1550.3 | 5.7 | $A$ | 194 | A |
| $\nu(C_a–C_m)$ | 1564.1 | 6.1 | | 195, 197, 198 | B |

[a]Mode labels and symmetry are based upon the notation in references A = Saini,[43] B = Zhang et al.,[14] and C = Bour et al.[44] In cases where the mode assignment was ambiguous between two references, both assignments are listed.

$$\frac{d_{k'}^{\,2}}{d_k^{\,2}} = \frac{I_{k'}}{I_k}\frac{(\omega_L - \omega_k)^3}{(\omega_L - \omega_{k'})^3}\frac{\omega_k}{\omega_{k'}}\frac{|\Phi(\omega_L) - \Phi(\omega_L - \omega_k)|^2}{|\Phi(\omega_L) - \Phi(\omega_L - \omega_{k'})|^2} \quad (9)$$

Since the terms involving the absorption spectra transforms $\Phi(\omega)$ may be computed numerically from a digitized absorption spectra,[39] the relative distortions along the normal coordinates, $d_k$, are determined by readily available spectroscopic quantities: the relative Raman intensity, the mode frequency, $\omega_k$, and the transform of the absorption spectra, $\Phi(\omega)$. Relating the observed Raman modes to a vector composed of $3N$ Cartesian coordinates of $N$ atomic positions, $\vec{D}_{obs}$, requires a relation of the form

$$\mathbf{L}\cdot a\vec{d} = \mathbf{D}^{\mathbf{T}}\cdot\vec{D}_{obs} \quad (10)$$

where $a$ is the scaling factor to convert the relative distortions $d$ to absolute distortions, $\mathbf{D}$ is a matrix composed of the eigenvectors associated with the observed modes in the RR spectra written as columns, and the superscript $\mathbf{T}$ indicates the matrix transpose. Since $\mathbf{D}^{\mathbf{T}}$ is not a square matrix, a unique inverse is not defined, which is expected, since the number of observable modes $k$ in Raman spectra is less than the $3N - 6$ vibrational modes possible in the molecule, and moreover is less than the $3N$ atomic coordinates in $\vec{D}_{obs}$. However, using singular value decomposition (SVD), a unique inverse can be defined to allow the determination of $\vec{D}_{obs}$

$$\vec{D}_{obs} = a\mathbf{V}^{\dagger}\cdot\mathbf{S}\cdot\mathbf{U}^{\mathbf{T}}\cdot\mathbf{L}\cdot\vec{d} \quad (11)$$

where the matrices $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices defined by the SVD of $\mathbf{D}^{\mathbf{T}}$, † represents the conjugate transpose, and $\mathbf{S}$ is a $3N - 6 \times 3N$ rectangular matrix that has the singular values along the diagonal and zeroes elsewhere.[42] In addition to this tranformation providing a resultant structure for the observed molecule, the singular values indicate the relative importance of each mode in determining the final structure. Thus, the distortions of the $3N$ atomic positions $\vec{D}_{obs}$ can be computed by applying eq 9 to each mode and through the application of eq 11 to the resulting system of equations. The constant $a$ will be determined by the docking algorithm as described in the section Docking.

**Assignment of Spectral Lines.** The computational spectrum, obtained using DFT, generates $3N - 6 = 228$ vibrational modes. The observable, experimental spectra had 33 lines, each of which had a typical fwhm of 7 $cm^{-1}$, which often span multiple lines in the DFT spectrum. The challenge is to unambiguously correlate the experimental and computed spectra. Referring to published line assignments,[14,43,44] the experimental lines could be associated with the motions of


dx.doi.org/10.1021/jp304310z | J. Phys. Chem. B 2012, 116, 11032−11040

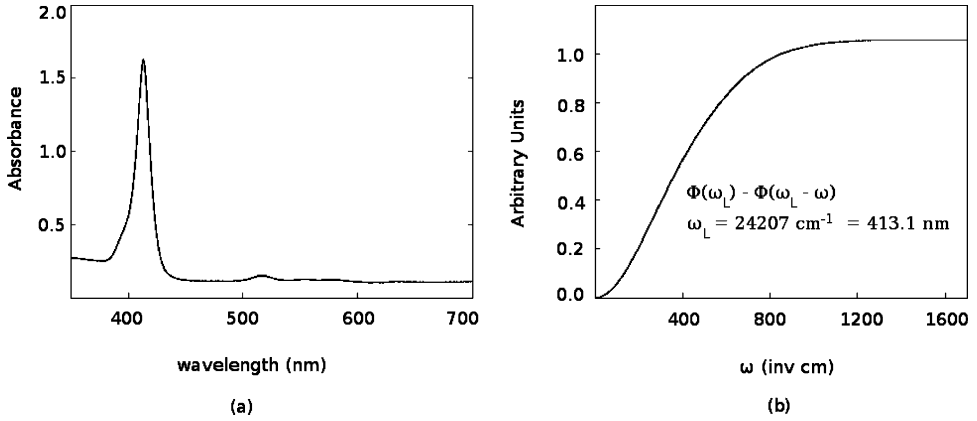



**Figure 3.** (a) Absorption spectra of TSPP at pH 9.0. (b) Transform of the absorption spectra in part a relative to the laser excitation at 413 nm.

certain porphyrin atoms. The use of relative ordering of the line assignments combined with atomic motions and scaling the DFT computed frequencies according to the scaled quantum mechanical factors provided by Kozlowski et al.[45] generally reduced the possible choices to 2−5 lines. If in addition symmetry assignments for the experimental spectra were available, a given experimental line could be correlated with just one or two lines in the computed spectra. For the cases where a unique pairing of experimental spectral lines to computed modes did not exist, the method assigned the distortion based upon the root-mean-square of the individual distortions. For example, if a line $k'$ was associated with two modes $\alpha$ and $\beta$, then the distortion for $d_{k'}$ is found from

$$d_{k'} = \sqrt{d_\alpha{}^2 + d_\beta{}^2} \tag{12}$$

where $d_{\alpha,\beta}$ is computed with eq 9, $I_{k'}$ is the same for both $\alpha$ and $\beta$, and $\omega_{k'}$ is the associated frequency for modes $\alpha$ or $\beta$. Since only 33 modes were observed in the spectra (Figure 4 and Table 1), only the lowest six vibrational orders were used to develop the distortions. As the lowest frequency modes correspond to the largest distortions in the molecule, these modes had the largest effect on determining the ligand conformation. As a final consistency check of the assignments, one examines the singular values resulting from the SVD of eq 11 to ensure the modes most sensitive to the fitting are accurately correlated. If sensitive modes are questionable, then one can iteratively analyze the fitting to estimate the uncertainty in the methodology due to the mode assignments.

**Error Analysis.** To estimate the uncertainty in the agreement between the Raman spectra and the predicted bound structures, one must quantify the measurement errors in the Raman spectral line amplitudes, the ambiguity in line assignments, and the precision of the docking algorithm. The standard deviation of the amplitude for the Raman spectra was computed for each spectral line using the difference between successive scans. These deviations were then propagated into the distortions using a perturbation form of eq 11,

$$\overrightarrow{\Delta D}_{\text{obs}} = a\mathbf{M}\cdot\overrightarrow{\Delta d} \tag{13}$$

where the matrix product of $\mathbf{V}^{\dagger}\mathbf{S}\mathbf{U}^{\mathrm{T}}\mathbf{L}$ has been simplified into a single matrix $\mathbf{M}$ for notational convenience. The net distortion due to line $k$, which due to the ambiguity of Raman spectral line assignments is associated with $m_k$ normal modes, is described by eq 12. The error associated by the combined effect of the $m_k$ modes can be estimated from combining eqs 9 and 12

$$\Delta d_k = \frac{1}{d_k}\sum_{i=1}^{m_k} d_i{}^2\frac{\Delta I_i}{2I_i} \tag{14}$$

where $\Delta I_i$ is the uncertainty in the Raman intensity $I_i$ of line $i$. To determine the effect of these errors on the docked conformations, we generated distortions of the bound TSPP molecule using eq 13 for each component of $\vec{d}$ in a positive and negative direction along the related normal mode at the amplitude $a$ determined for that conformation. The distorted molecule was then rescored at the binding site to determine an associated binding energy variation. The rms value of the change in binding energy due to the variations along each spectral line is taken as the uncertainty in the binding energy. However, since these perturbations are from an approximately optimal docked configuration, the computed energy variations almost always indicate an increase in energy (reduction in binding affinity). Thus, the perturbed energy values were split into two groups: those that were higher energy $E_{\text{high}}$ and those lower $E_{\text{low}}$. The rms value of each group was found separately, and the energy error is reported as the range $E_{\text{eq}} - E_{\text{low}}$ to $E_{\text{eq}} + E_{\text{high}}$, where $E_{\text{eq}}$ is the energy of the docked conformation at a given amplitude $a$.

## RESULTS AND DISCUSSION

**Experimental Spectra.** The characteristic visible absorption spectrum of a 3.2 $\mu$M TSPP solution is shown in Figure 3a. It exhibits a strong absorption peak at 413 nm (Soret band) and four lesser peaks from 500 to 650 nm (Q-bands). Figure 3b shows the Kramers−Kronig (K−K) transform $\Phi(\omega)$ of the absorption spectrum calculated according to eq 3. The transform shows only the region of interest to the RR spectra discussed in this manuscript (30−1600 cm$^{-1}$) and exhibits the initial quadratic dependence on wavenumber described in Jarzecki and Spiro,[13] but flattens out above 800 cm$^{-1}$. Because the change in the K−K transformed spectrum is significant, $\Phi(\omega)$ cannot be approximated by a quadratic term,[13] and thus, the full transform is needed for substitution in eq 9. The Raman spectra of free-base TSPP in a phosphate buffer at pH 9.0 are shown in Figure 4 and can be compared to the spectra of the bound TSPP−BLG complex in a similar solution, also in Figure 4. Most RR lines in the bound TSPP spectra are shifted higher by 1−2 cm$^{-1}$ than in the free molecule. From an energy standpoint, they cannot be taken as a clear representation of porphyrin distortions unless they showed large changes in intensity. Conversely, other lines exhibited distinct changes, and they are the ones that were more significant in generating the


dx.doi.org/10.1021/jp304310z | J. Phys. Chem. B 2012, 116, 11032−11040

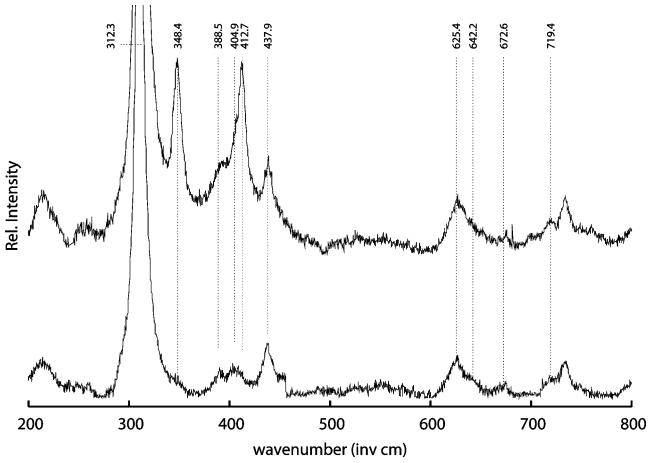



**Figure 4.** RR spectra of free-base TSPP and TSPP–BLG complex in a pH 9.0 phosphate solution using 413.1 nm excitation. The spectra of the bound TSPP–BLG complex have been displaced vertically.

distorted structure of the porphyrin. For example, the intensity of the modes in the region 348–439 $cm^{-1}$ is strongly enhanced upon binding. In particular, the lines at 348 and 412 $cm^{-1}$ only appear in the bound spectra. Additionally, the lines at 388, 404, and 749 all upshift by more than 6 $cm^{-1}$ in the bound spectra. Only the 717 $cm^{-1}$ line is shifted to the red. If the core size indicator lines in the region 1350–1600 $cm^{-1}$ follow the metalloporphyrin empirical relationships provided by Parthasarathi et al.,[46] then a shift $\nu_{bound} - \nu_{free}$ of 1 $cm^{-1}$ of the mode $\nu_4$ would indicate a core size increase of 0.007 Å in the bound form. The other modes indicate similar small changes in core size from +0.002 to +0.007 Å for the bound form. The Raman lines of the bound TSPP–BLG complex are labeled with the associated mode assignments from Table 1, which also lists the fitted center frequencies and widths of each line along with mode assignments that were determined from literature values. In some cases, e.g., the line at 886 $cm^{-1}$, the observed value was near two different modes listed in the literature and it was not possible a priori to decide which was the correct assignment. For those lines, both modes are listed together with the associated references. Some of the references provide the symmetry groups of a mode relative to the $D_{4h}$ point group of metalloporphyrins, but they are listed in Table 1 with the correlated $D_2$ symmetry representation of the nonplanar distortions of the simulated TSPP molecule. Where multiple representations are possible, both are listed.

**Computed Structures.** As mentioned earlier, for each normal mode of the equilibrium configuration of the TSPP molecule, the DFT solution yields the eigenfrequency, symmetry group representation, Raman activity, and normalized Cartesian displacements of each atom. An example of the results for one mode may be found in Table S1 of the Supporting Information (requests for a complete table of all 228 modes will gladly be entertained by the corresponding author). By combining the transformed absorption spectra, the RR spectra, the DFT computed normal modes, and the correlations between the computed and experimental spectra, one can now determine the relative distortions along each normal coordinate in the TSPP structure that are reflected in the RR spectra of the bound configuration of TSPP and BLG from eq 9. To resolve the actual magnitude of the distortion in the molecule, the constant $a$ is assigned a set of appropriate values, (0.1, 0.2, 0.3, ..., 0.8). Each value produces a proportionally distorted ligand, an example of which is demonstrated in Figure 5, that portrays a relative distortion of $a$ = 0.4 from two different perspectives to show the out-of-plane character of the atomic displacements. For each value of $a$ (i.e., for each distortion), the resulting TSPP molecule was processed with the computational docking algorithm to find the best protein–ligand configuration. Figure 6 displays the dependence of the best docking conformer on the distortion coordinate $a$. The figure shows a global minimum at 0.4 with possible local minimum at 0.8. Additionally, error bars are plotted for each data point to indicate the variation in the computed dock energy for each configuration that results from the uncertainty in the Raman amplitudes and mode assignments. Although not clear from Figure 6, the docking algorithm converged to two probable binding sites. This was expected because TSPP does not have a known specific binding site for BLG.[7]

The lowest energy solution in cyan and its location relative to the protein are shown in Figure 7. The second possible binding

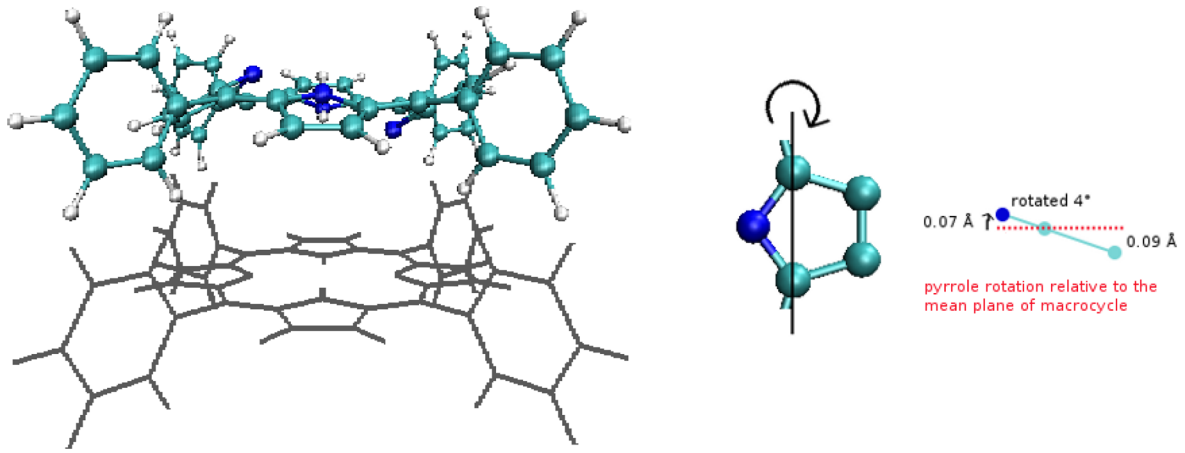



**Figure 5.** Relative distortion of TSPP bound to BLG along normal mode coordinates predicted by the Raman and absorption spectra of TSPP. The base porphyrin structure is depicted in the gray wire diagram. The distorted molecule is drawn in ball–stick mode and translated 5 Å along the $z$-coordinate to separate the molecules. The pyrrole groups 1 and 2 are rotated about 4° out of the plane of the macrocycle as shown; groups 3 and 4 are rotated in the opposite direction.


dx.doi.org/10.1021/jp304310z | J. Phys. Chem. B 2012, 116, 11032–11040

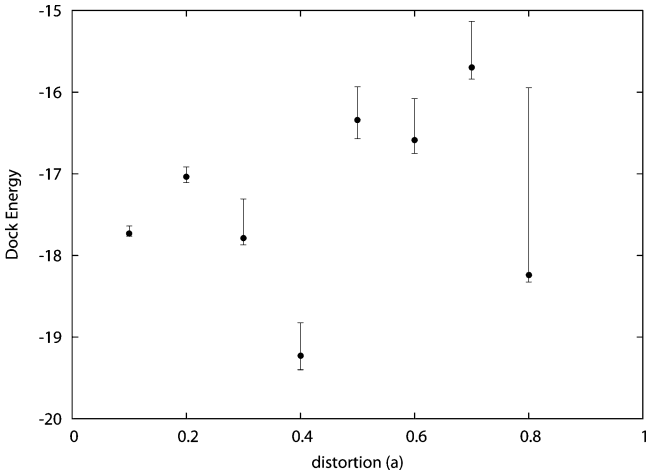



**Figure 6.** Dependence of relative binding energy of the ligand on the distortion scaling factor *a*. Energy error bars are computed according to the section Error Analysis.

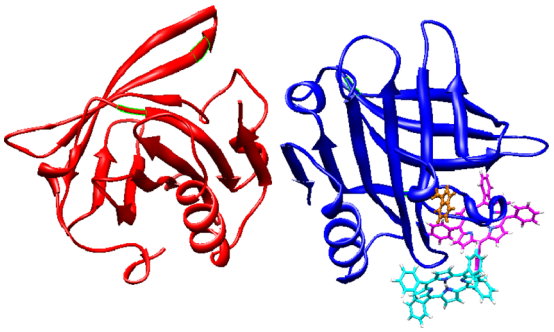


**Figure 7.** The best two solutions for the bound shape and location of TSPP on the BLG surface. The lowest energy solution is colored in cyan, the second best solution is colored purple. The Trp19 residue is depicted in orange. The beta sheets for one monomer unit are labeled in accordance with Brownlow et al.[23] (The diagram was produced using Chimera.[47])

site is indicated by the purple TPP molecule in Figure 7 and corresponds to the docking solutions associated with the scaling factor $a$ = 0.1, 0.2, 0.3, 0.5, or 0.7. The Trp19 residue is drawn in ball−stick representations to highlight the relative orientations of the aromatic rings in tryptophan and the porphyrin macrocycles between the two conformations. The plane of the macrocycle for the lowest energy solution relative to the indole group of the Trp19 is 10 Å away and oriented perpendicular to the group, whereas the alternative site is 7 Å away and oriented parallel to the group.

The methodology employed in this paper combines the results of several types of analyses—absorption spectroscopy, resonance Raman spectroscopy, DFT, and normal-mode analysis—to deduce a probable structure for a protein−ligand complex when the ligand does not have a specific binding site and NMR and X-ray crystallography are not viable techniques. Specifically, in our previous research using fluorescence spectroscopy and lifetime, we have observed that TSPP binds to BLG and that only one TSPP molecule appears to bind each TSPP dimer.[6] It remained unknown the likely location of the binding site. From Figure 4, one observes that, even at the relatively low concentrations used in our samples, RR spectroscopy is able to clearly detect the Raman lines for the porphyrin and to identify a subset of lines (e.g., 348 and 412 $cm^{-1}$) that can be representative of the bound state of the ligand. The 348 $cm^{-1}$ mode is correlated with a $B_{1u}$ mode which is an asymmetric in-plane vibration.[14] The 412 $cm^{-1}$ mode is correlated with bending out of the planes of the phenyl groups but in the plane of the macrocycle.[43] As listed in Table 1, it is possible to assign the most binding-affected modes to the vibrational modes of specific atoms in the porphyrin; however, additional analysis was required to use RR spectra as guidance to select the most likely bound configuration. This can be achieved using eq 11 whereby the connection between vibrational modes and the distortions of the atomic positions from equilibrium can be obtained, based upon the normal-mode analysis provided by DFT. The vibrational modes provided by DFT are matched to the experimental RR spectral intensity as described in the theoretical introduction to yield the porphyrin distorted structure[13,38−42] (Figure 5) that best matches the experimental spectra. In other words, the docked structure of the porphyrin has to be similar to the one of Figure 5 aside from the scaling factor $a$ which determines the extent of the distortion. As shown in Figure 6, we have determined that the value of $a$ that produces the lowest docking energy is 0.4. The search for the best computationally docked configuration corresponding to the porphyrin structure whose distortion best matched the RR spectra returned the conformation shown in Figure 7. The binding configuration of Figure 7 is not only self-consistent with the RR spectra and the DFT normal-mode analysis but contains features that are consistent with both our previous results[7] and what is expected of the binding site of a water-soluble ligand to a protein. First, the computationally determined binding site of Figure 7 places the four negative charges of TSPP adjacent (3−4 Å) to the positively charged residues Lys14, Lys47, Lys100, and Lys101, and Arg124 on the surface of BLG. In addition, the computed site is near a protein feature that is likely to be involved in the observed photoinduced protein unfolding: the base of the $\beta$-barrel that is uncovered by the EF-loop,[6] which is an area of increased instability of the protein structure.[9] Finally, the binding configuration places the macrocyle 7−10 Å from the indole ring of Trp19 (shown in Figure 7) which has been shown to be quenched by TSPP binding.[7] The value of this analysis is that the distortions and location of the molecular structure are consistent with the observed Raman spectra, the molecular mechanics simulations, and previous results.[7]

## CONCLUSION

The methodology outlined here used well-established experimental and computational techniques—absorption and Raman spectroscopy, docking, and DFT—to estimate the bound configuration of TSPP to BLG. When neither NMR nor X-ray crystallography offer information at the low concentrations required for these experiments, the method that we adopted may help provide a coherent picture of ligand binding and aid future studies to establish the mechanism that leads to the photoinduced conformational changes prompted by TSPP on BLG.[7]

## ASSOCIATED CONTENT

### Supporting Information

Plots of the Raman spectra for the free-base TSPP molecule and the TSPP−BLG complex at pH 9 over the full range collected 30−1700 $cm^{-1}$. Table listing one of the normal modes computed for the undistorted TSPP molecule in the $D_2$ point group. The table lists the number, irreducible

representation, frequency ($cm^{-1}$) for the mode, and normalized Cartesian displacements for each atom. The example mode is an *oop* mode, and hence, the *z*-coordinate displacements predominate. Additionally, listings of typical input files for Dock 6.4 and Autodock 4.2 are provided to document the search and scoring algorithms employed. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION


### Corresponding Author

*E-mail: lorenzo.brancaleon@utsa.edu.


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors acknowledge the assistance of Mr. Gary Noojin and Mr. David Stolarski in the developing, aligning, and calibrating of the laser and optical setup for the Raman spectroscopy apparatus and Mr. Larry Estlack for his help with protein sample preparation. We also acknowledge the use of the facilities at the Air Force Research Laboratory in conducting the Raman spectroscopy studies. Funding was provided by the Air Force Research Laboratory, 711th Human Performance Wing, Human Effectiveness Directorate, Bioeffects Division, through contract number FA8650-07-D-6800 (to J.E.P.) and contract number FA8650-08-D-6930 (to L.B.). This work was supported in part by a grant of computer time from the DoD High Performance Computing Modernization Program at the Army Research Laboratory, DoD Supercomputing Resource Center. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).

## REFERENCES


(1) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76−88.

(2) Abbruzzetti, S.; Sottini, S.; Viappiani, C.; Corrie, J. E. T. *Photochem. Photobiol. Sci.* **2006**, *5*, 621−628.

(3) Jacobson, K.; Rajfur, Z.; Vitriol, E.; Hahn, K. *Trends Cell Biol.* **2008**, *18*, 443−450.

(4) Löweneck, M.; Milbrandt, A. G.; Root, C.; Satzger, H.; Zinth, W.; Moroder, L.; Renner, C. *Biophys. J.* **2006**, *90*, 2099−2108.

(5) C. Ted Lee, J.; Smith, K. A.; Hatton, T. A. *Biochemistry* **2005**, *44*, 524−536.

(6) Silva, I.; Sansone, S.; Brancaleon, L. *Protein J.* **2009**, *28*, 1−13.

(7) Belcher, J.; Sansone, S.; Fernandez, N. F.; Haskins, W. E.; Brancaleon, L. *J. Phys. Chem. B* **2009**, *113*, 6020−6030.

(8) Fernandez, N. F.; Sansone, S.; Mazzini, A.; Brancaleon, L. *J. Phys. Chem. B* **2008**, *112*, 7592−7600.

(9) Qin, B. Y.; Bewley, M. C.; Creamer, L. K.; Baker, H. M.; Baker, E. N.; Jameson, G. B. *Biochemistry* **1998**, *37*, 14014−14023.

(10) Spiro, T. G.; Smulevich, G.; Su, C. *Biochemistry* **1990**, *29*, 4497−4508.

(11) Billes, F.; Várady, B. *Spectrochim. Acta, Part A* **2008**, *70*, 729−734.

(12) Long, D. A. *The Raman Effect, A Unified Treatment of the Theory of Raman Scattering by Molecules*; John Wiley and Sons, Ltd.: West Sussex, England, 2002.

(13) Jarzecki, A. A.; Spiro, T. G. *J. Phys. Chem. A* **2005**, *109*, 421−430.

(14) Zhang, Y.-H.; Chen, D.-M.; He, T.; Liu, F.-C. *Spectrochim. Acta, Part A* **2003**, *59*, 87−101.

(15) Lieber, C. A.; Mahadevan-Jansen, A. *Appl. Spectrosc.* **2003**, *57*, 1363−1367.

(16) McCreery, R. L. *Raman Spectroscopy for Chemical Analysis*; Wiley-Interscience: New York, 2000; pp 251−291.

(17) Wojdyr, M. *J. Appl. Crystallogr.* **2010**, *43*, 1126−1128.

(18) *Avogadro: an open-source molecular builder and visualization tool*, version 1.0.0; http://avogadro.openmolecules.net (accessed January 19, 2012).

(19) Schrödinger, L. L. C. *Maestro*, version 9.1; New York, 2010.

(20) Rosa, A.; Ricciardi, G.; Baerends, E. J.; Romeo, A.; Scolaro, L. M. *J. Phys. Chem. A* **2003**, *107*, 11468−11482.

(21) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(22) Oliveira, K. M. G.; Valente-Mesquita, V. L.; Botelho, M. M.; Sawyer, L.; Ferreira, S. T.; Polikarpov, I. *Eur. J. Biochem.* **2001**, *268*, 477−483.

(23) Brownlow, S.; Cabral, J. H. M.; Cooper, R.; Flower, D. R.; Yewdall, S. J.; Polikarpov, I.; North, A. C. T.; Sawyer, L. *Structure* **1997**, *5*, 481−495.

(24) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Petterson, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA* **2009**, *15*, 1219−1230.

(25) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785−2791.

(26) Lecerof, D.; Fodje, M.; Hansson, A.; Hansson, M.; Al-Karadaghi, S. *J. Mol. Biol.* **2000**, *297*, 221−232.

(27) Horning, T. L.; Fujita, E.; Fajer, J. *J. Am. Chem. Soc.* **1986**, *108*, 323−325.

(28) Geno, M. K.; Halpern, J. *J. Am. Chem. Soc.* **1987**, *109*, 1238−1240.

(29) Jentzen, W.; Song, X.; Shelnutt, J. *J. Phys. Chem. B* **1997**, *101*, 1684−1699.

(30) Moustakas, D. T.; Lang, P. T.; Pegg, S.; Petterson, E.; Kuntz, I. D.; Brooijmans, N.; Rizzo, R. C. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 601−619.

(31) Forli, S.; Botta, M. *J. Chem. Inf. Model.* **2007**, *47*, 1481−1492.

(32) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res.* **2004**, *32*, W665−W667.

(33) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704−721.

(34) Li, X.-Y.; Czernuszewicz, R. S.; Kincaid, J. R.; Su, Y. O.; Spiro, T. G. *J. Phys. Chem.* **1990**, *94*, 31−47.

(35) Rush, T. S.; Kozlowski, P. M.; Piffat, C. A.; Kumble, R.; Zgierski, M. Z.; Spiro, T. G. *J. Phys. Chem. B* **2000**, *104*, 5020−5034.

(36) Frisch, M. J.; et al. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.

(37) Cremer, D.; Kraka, E. *Curr. Org. Chem.* **2010**, *14*, 1524−1560.

(38) Hizhnyakov, V.; Tehver, I. *Phys. Status Solidi B* **1967**, *21*, 755−768.

(39) Chan, C. K.; Page, J. B. *J. Chem. Phys.* **1983**, *79*, 5234−5250.

(40) Lee, S.-Y.; Heller, E. J. *J. Chem. Phys.* **1979**, *71*, 4777−4788.

(41) Heller, E. J.; Sundberg, R. L.; Tannor, D. *J. Phys. Chem.* **1982**, *86*, 1822−1833.

(42) Golub, G.; Kahan, W. *J. Soc. Ind. Appl. Math., Ser. B* **1965**, *2*, 205−224.

(43) Saini, G. S. S. *Spectrochim. Acta, Part A* **2006**, *64*, 981−986.

(44) Bouř, P.; Záruba, K.; Urbanová, M.; Setnička, V.; Matějka, P.; Fiedler, Z.; Král, V.; Volka, K. *Chirality* **2000**, *12*, 191−198.

(45) Kozlowski, P. M.; Rush, T. S.; Jarzecki, A. A.; Zgierski, M. Z.; Chase, B. *J. Phys. Chem. A* **1999**, *103*, 1357−1366.

(46) Parthasarathi, N.; Hansen, C.; Yamaguchi, S.; Spiro, T. G. *J. Am. Chem. Soc.* **1987**, *109*, 3865−3871.

(47) Petterson, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.