Use of Vibrational Spectroscopy to Study Protein and DNA Structure, Hydration, and Binding of Biomolecules: A Combined Theoretical and Experimental Approach

K. J. JALKANEN,[1] V. WÜRTZ JÜRGENSEN,[1] A. CLAUSSEN,[1] A. RAHIM,[1] G. M. JENSEN,[1] R. C. WADE,[2] F. NARDI,[2] C. JUNG,[3] I. M. DEGTYARENKO,[4] R. M. NIEMINEN,[4] F. HERRMANN,[5] M. KNAPP-MOHAMMADY,[5] T. A. NIEHAUS,[5] K. FRIMAND,[5] S. SUHAI[5]

[1]*Quantum Protein (QuP) Center, Department of Physics, Technical University of Denmark, Building 309, DK-2800 Kgs. Lyngby, Denmark*

[2]*European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany*

[3]*Max-Delbrück-Center for Molecular Medicine, Robert-Roessle-Strasse 10, 13125 Berlin, Germany*

[4]*Laboratory of Physics, Helsinki University of Technology, P. O. Box 1100, FIN-02015 HUT, Finland*

[5]*Department of Molecular Biophysics, German Cancer Research Center, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany*

Received 16 July 2005; accepted 7 September 2005
Published online 5 December 2005 in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/qua.20863

ABSTRACT: We report on our work with vibrational absorption, vibrational circular dichroism, Raman scattering, Raman optical activity, and surface-enhanced Raman spectroscopy to study protein and DNA structure, hydration, and the binding of

Correspondence to: K. J. Jalkanen; current address: Nanochemistry Research Institute, Department of Applied Chemistry, Curtin University of Technology, P. O. Box U1987, Perth, Western Australia, 6845; e-mail: K.Jalkanen@curtin.edu.au

G. M. Jensen is currently at Gilead Sciences, 650 Cliffside Drive, San Dimas, CA 91733, USA.

R. C. Wade is currently at EML Research gGmbH, Schloss-Wolfbrunnenweg 33, D-69118 Heidelberg, Germany.

F. Nardi is currently at Drug Design, Chemical Science, Sanofi-Aventis, 13, Quai Jules Guesde, Friedel 115, 94403 Vitry-sur-Seine cedex, France.

T. A. Niehaus is currently at the Department of Theoretical Physics, University of Paderborn, D-33098 Paderborn, Germany.

Contract grant sponsor: Danish National Research Foundation.
Contract grant sponsor: Technical University of Denmark.
Contract grant sponsor: Danish Center for High Performance Computing (Sun Cluster) at DTU.
Contract grant sponsor: German Cancer Research Center.
Contract grant sponsor: Max Delbrueck Center for Molecular Medicine.
Contract grant sponsor: European Molecular Biology Laboratory.
Contract grant sponsor: Helsinki University of Technology.
Contract grant sponsor: Nanochemistry Research Institute of Curtin University of Technology.

ligands, drugs, pesticides, or herbicides via a combined theoretical and experimental approach. The systems we have studied systematically are the amino acids (L-alanine, L-tryptophan, and L-histidine), peptides (*N*-acetyl L-alanine *N*′-methyl amide, *N*-acetyl L-tryptophan *N*′-methyl amide, *N*-acetyl L-histidine *N*′-methyl amide, L-alanyl L-alanine, tri-L-serine, *N*-acetyl L-alanine L-proline L-tyrosine *N*′-methyl amide, Leu-enkephalin, cyclo-(gly-L-pro)$_3$, *N*-acetyl (L-alanine)$_n$ *N*′-methyl amide), 3-methyl indole, and a variety of small molecules (dichlobenil and 2,6-dochlorobenzamide) of relevance to the protein systems under study. We have used molecular mechanics, the SCC-DFTB, SCC-DFTB+disp, RHF, MP2, and DFT methodologies for the modeling studies with the goal of interpreting the experimentally measured vibrational spectra for these molecules to the greatest extent possible and to use this combined approach to understand the structure, function, and electronic properties of these molecules in their various environments. The application of these spectroscopies to biophysical and environmental assays is expanding, and therefore a thorough understanding of the phenomenon from a rigorous theoretical basis is required. In addition, we give some exciting and new preliminary results which allow us to extend our methods to even larger and more complex systems. The work presented here is the current state of the art to this ever and fast changing field of theoretical spectroscopic interpretation and use of VA, VCD, Raman, ROA, EA, and ECD spectroscopies. Int J Quantum Chem 106: 1160–1198, 2006

Key words: vibrational circular dichroism; Raman optical activity; hydration; density functional theory; Raman scattering

Introduction

Recently the use of vibrational spectroscopy has undergone a revival in its application and development. With the development of femtosecond lasers, one can perform time-resolved spectroscopy to probe the lifetimes of the various intermediates during a chemical reaction, protein folding, or other biochemical processes. Two forms of vibrational spectroscopy, vibrational circular dichroism (VCD), and Raman optical activity (ROA), are especially relevant for configurational and conformational studies, as these two forms of spectroscopy can be used to distinguish between two enantiomers [1, 2]. Recently there has come onto the market improved FTIR instrumentation, e.g., from Bruker AG, Ettlingen, Germany; Thermo Electron Corporation (formerly Thermo Nicolet), Madison, Wisconsin; and ABB Analytical Instrumentation Bomem, Quebec, Canada. In addition there is now commercially available ROA instrumentation. To help with the interpretation of the vibrational absorption (VA), VCD, Raman scattering, and ROA spectra, one needs to be able to reliably simulate these spectra [3, 4]. Since the intensities of these spectra require the changes of various tensor quantities that depend on both the electron density and electron currents with respect to nuclear displacements and nuclear velocities, the calculations are rather involved [5–10]. Rigorous expressions have been derived for all these quantities and, in addition, the various formalisms have been implemented in many commercial and academic codes at various levels, with respect to both the treatment of electron exchange and electron correlation and the perturbing effects: the solvent, temperature, strong static and dynamic electric and magnetic fields, counter-ions and metal atoms (e.g., metalloproteins) and metal surfaces (surface-enhanced Raman and ROA). Research in some cases has been focused on optimization of accuracy, while in other cases the research has been focused on computational efficiency (method development). In other cases, where the desire has been to interpret the experimental data of proteins, very large and diverse calculations of lower accuracy have been performed. In this case, one is usually satisfied with reproducing the qualitative features in the spectra. Ultimately one would like to have methods that are both computationally efficient and accurate for all types of interactions, including the relatively weak van der Waals (dispersion) forces (interactions) and hydrogen bonding. The currently developed generalized gradient approximation (GGA) and hybrid exchange-correlation (XC) functionals used and implemented in most commercial and academic density functional theory (DFT) codes are not adequate to treat systematically the weak van der Waals forces and interactions, charge transfer, and excited

electronic states. In many cases, the vibrational spectra of van der Waals complexes and long-lived excited states can add greatly to the understanding of these species and states, if the experimental spectra can be fully interpreted. To date most of the interpretations have not been based on rigorous high-level treatment of the complete system, including the effects of temperature, environment, and the perturbing radiation used to measure the spectra but, on simpler model systems and in many cases, no explicit treatment of the solvent. For these systems and states, the effects due to the perturbing radiation, temperature and environment may be very important because the radiation and the coupling with temperature and environment are not necessarily a small perturbation to the isolated system as is assumed to be the case where there is only one conformation or configuration of the system and bonds are stable by $\geq$50 kcal/mole. Clearly this is not the case in the biological systems discussed in this report. In addition, one must address the problem of systematic errors that may appear with any method. Scott and Radom have recently reported scale factors for Hartree–Fock, Møller–Plesset, quadratic configuration interaction, density functional theory (DFT), and some semi-empirical methods [11]. In this study, we can obtain better agreement with the experimental spectra by scaling, but then one loses the predictive ability of the methods, especially when one works on systems where the scale factors have not yet been derived. In general, one needs to have a good set of vibrational frequencies that have been carefully assigned, by for example, isotopic substitution. In addition, there is the question of whether to perform a uniform scaling, or to scale-specific force constants with different scale factors. In the field of molecular mechanics, the scaling is usually performed with simple valence coordinates, while vibrational spectroscopists prefer to use local vibrational coordinates, the same coordinates they use to do their assignments. With the advent of visualization programs such as Molden and GaussView, one no longer undertakes a so-called vibrational analysis, transforming the Cartesian Hessian to local vibrational coordinates, and giving the so-called potential energy distributions (PED). Since this report is focused on using the currently available methods to better understand, interpret and use the results of experimental vibrational studies, we have focused this review on presenting how well one can currently do by choosing one of the best hybrid DFT methods, the Becke 3LYP (B3LYP). We have chosen not to scale the force constants, but to present the raw spectra. This is useful in that one then sees the systematic errors. These are useful to be able to see, since in many cases they can be related to the physics of the approximations and enable one to see this. Hence systematic underestimation of bond lengths results in bonds that are too strong and bond stretch frequencies that are too large. We think this is very important to see, as it may lead to other errors from the methods, which one would lose sight of if the frequencies were made to agree better than they actually do. Here one would also lose sight of the difference between harmonic calculated frequencies and anharmonic experimentally measured frequencies.

In addition to the results at the B3LYP level of theory, we have presented results from the best GGA method, PW91. In this study, one does not have to introduce the nonlocal Hartree–Fock operator. Another approach is to use the optimized effective potential (OEP) method to derive local exchange and correlation functionals. The nonlocal Hartree–Fock exchange operator is used to develop orbital-based local OEPs for the exchange functional. MP2, CSD, and coupled cluster methods can be used to develop orbital-based local OEPs for the correlation functionals [12].

Although this exciting new work has great promise, these new OEP exchange and correlation functionals are quite involved and have not yet been implemented in all quantum chemistry and condensed matter codes, and certainly not for the analytical gradients, Hessians, APT, AAT, ED-EDPD, EDMDPD, and EDEQPD which are required to simulate the VA, VCD, Raman, and ROA spectra. In addition, the treatment of the environment via continuum models would also be necessary to adequately evaluate these OEP exchange and correlation functionals for biomodeling of vibrational spectroscopy. Hence the theoretical state of this field is ever changing and our work here only represents a snapshot, which unfortunately requires a fast shutter to catch the fast changing image we have attempted to present in this work.

Methods and Materials

In the modeling studies we report in this work Gaussian, Cambridge Analytical Derivatives Package (CADPAC), and self-consistent charge density functional tight binding (SCC-DFTB) have been used to perform the ab initio, DFT, and semi-em-

pirical DFT-based calculations. To simulate the VA, VCD, Raman scattering, and ROA spectra, one requires (i) an optimized geometry determined by minimizing the electronic energy until the forces are zero, and at that optimized geometry; (ii) a Hessian for which 3N-6 of the eigenvalues are positive and six are close to zero (corresponding to the three translations and three rotations); (iii) the atomic polar tensors (APT); (iv) the atomic axial tensors (AAT); (v) the electric dipole–electric dipole polarizability derivatives (EDEDPD); (vi) the electric dipole–electric quadrupole polarizability derivatives (EDEQPD); and finally (vii) the electric dipole–magnetic dipole polarizability derivatives (EDMDPD). In this study, we are working within the harmonic approximation for the vibrational frequencies and, in addition, the Born–Oppenheimer (BO) approximation for the electric dipole moment, the electric dipole–electric dipole polarizability, and the electric dipole–electric quadrupole polarizability. With respect to nuclear displacements for tensor quantities, these derivatives can be calculated either by finite difference or by analytical derivative techniques based on either coupled Hartree–Fock theory, linear response theory, or the RPA [13]. For the AAT and EDMDPD, we must go beyond the BO approximation and calculate the derivatives of the magnetic dipole moment with respect to the nuclear velocity (momentum). This is because within the BO approximation, the electronic component of the magnetic moment of the molecule is zero. Hence if one tries to calculate these tensor derivatives with respect to nuclear displacements, one obtains zero. However, these tensor quantities have been calculated by finite difference methods, but now one needs to evaluate the overlap integral between the derivative of the electronic wave function with respect to the nuclear displacement, $(\partial\phi/\partial X_{\lambda\alpha})$, that is, one of the terms neglected and assumed to be zero in the BO approximation: $\nabla^2(\phi\chi) = \phi\nabla^2\chi$ $(=\phi\partial^2\chi/\partial X_{\lambda\alpha}X_{\lambda'\alpha'})$, and the derivative of the electronic wave function with respect to the magnet field perturbation. The two terms which have been neglected in the BO approximation are $\chi\nabla^2\phi$ $(=\chi\partial^2\phi/\partial X_{\lambda\alpha}X_{\lambda'\alpha'})$ and $\nabla\chi\nabla\phi$ $(=\partial\chi/\partial X_{\lambda\alpha}\partial\phi/\partial X_{\lambda'\alpha'})$. In addition, one requires the derivative of the electronic wave function with respect to the magnetic field, that is, $\partial\phi/\partial H_\beta$ and hence the AAT is $\langle\partial\phi/\partial X_{\lambda\alpha}|\partial\phi/\partial H_\beta\rangle$. Here ϕ is $\phi(R, r)$ in the equation $H(R, r)\phi(R, r) = E(R)\phi(R, r)$, R are the nuclear coordinates, and the r are the electron coordinates. To evaluate $\partial\phi/\partial X_{\lambda\alpha}$, one needs to add the nuclear displacement perturbation, $H' = (\partial H/\partial X_{\lambda\alpha})X_{\lambda\alpha}$. To evaluate $\partial\phi/\partial H_\beta$, one needs to add the magnetic field perturbation, $H' = -(\mu_{mag})_\beta H_\beta$, where μ_{mag} is the magnetic dipole moment and H is the magnetic field. Both of these perturbations can be calculated by either finite difference techniques, coupled Hartree–Fock (CHF) theory, linear response theory, or the random phase approximation (RPA). CHF theory allows us to calculate the derivative of the wave function directly, while the RPA allows us to evaluate the sum over excited states expression. These two forms are formally equivalent, but in practice their convergence with respect to the basis set limit is different. In addition, the RPA formulation converges much faster by using special customized basis sets as implemented in the program SYSMO by Lazzeretti and Zanasi and their colleagues [14]. These expressions can also be calculated using linear response theory. These expressions have been extended to density functional theory, but here rather than calculating the derivatives with respect to the Hartree–Fock orbitals, one does with respect to the Kohn–Sham orbitals. Hence in practice, the expressions that involve knowing the wave function and its derivatives with respect to various perturbations (e.g., nuclear displacement, static and dynamic electric and magnetic fields, nuclear magnetic moments), allow us to evaluate the APT, AAT, EDEDPD, EDEQPD, and EDMDPD at the DFT level, with expressions that are formally very similar to those derived for Hartree–Fock theory. At the DFT level, one can use gauge-invariant atomic orbitals to represent the Kohn–Sham orbitals, as one has done when representing the Hartree–Fock orbitals [15, 16]. One arrives at expressions that, again, are very similar to those one has for Hartree–Fock theory, the difference being that one now needs to know how the exchange-correlation functional term varies with respect to the various applied perturbations. In addition, for those quantities that depend on the frequency of the dynamic electric or magnetic field, one must also take the frequency into account or calculate the analogous static field limit. This is what has been done for both the electric dipole–electric dipole polarizability and the electric dipole–magnetic dipole polarizability. When the frequency of the applied electric or magnetic field is resonant with either an allowed electric dipole or a magnetic dipole allowed transition in the molecule, the expressions need to be modified to take this into account. For the far from resonance condition, the static field limits give expressions and derivatives with respect to nuclear displacements, which can be

used to simulate the corresponding Raman and ROA intensities. The agreement between the simulated Raman and ROA spectra and the experimentally measured spectra is only qualitative when one uses the static limit for the polarizability derivatives. If one wishes to achieve quantitative accuracy for the Raman and ROA spectra, one must use the frequency dependent expressions for the polarizability derivatives. To do this one must use time dependent-DFT (TD-DFT) to calculate the polarizability derivatives. In addition, when the frequency of the laser is resonant with an electronic transition in the molecule, the expressions must be modified to take this into account. This results in the expressions for the resonance Raman and resonance ROA spectra. Here we see a large enhancement of the signal relative to the non resonant condition.

Applications

L-ALANINE (LA)

L-Alanine is the simplest of the chiral amino acids. It has two neutral forms, the zwitterionic form (LAZ) with the ammonium, $-NH_3^+$, and carboxylate, $-CO_2^-$, groups and the nonionic neutral form (LAN) with amine, $-NH_2$, and carboxylic acid (or carboxyl), -COOH, groups. Diem [17] reported the VA and VCD spectra of LAZ. Our initial attempts to determine the structure of the LAZ at the DFT level with the B3LYP hybrid XC functional and the 6-31G* basis set resulted in a proton being transferred from the positively charged ammonium group, $-NH_3^+$, to the negatively charged carboxylate group, $-CO_2^-$, [18, 19]. We found a stable LAZ with a network of water molecules stabilizing and bridging the two charged groups. Here we were able to identify the conformer of the dominant species present in aqueous solution. Note that the conformation of our LAZ was different from that found by Barron et al. [20] at the restricted Hartree–Fock (RHF) level and by Yu et al. [21] at the RHF Onsager continuum level of theory. Here without the explicit water molecules, one finds the OH conformer to be the only stable one. To stabilize the HHO, HOO, and HHOO conformers of the LAZ (see Ref. [19], Fig. 2 for notation) one required four explicit water molecules. In addition to the explicit water molecules that model the waters in the first hydration shell that are directly hydrogen bonded with the LAZ, we treated the effects due to the bulk water by embedding the LAZ/water complex within a spherical cavity, the Onsager continuum model. This results in better agreement with the experimental VA, VCD, and Raman spectra than the model that only includes the explicit water molecules, or the isolated LAZ reported at the RHF level and the LAZ reported at the RHF Onsager continuum level that do not include explicit water molecules. Finally, we increased the number of water molecules in the hydration shell surrounding and stabilizing the LAZ structure to see the effects of such additional water molecules [22]. For the largest LAZ/water complex (LAZ + nine explicit water molecules) also embedded within the Onsager continuum model, we have simulated the Raman and ROA spectra at the B3LYP hybrid XC DFT level of theory and compared them with the measured spectra of Barron et al. [20] and found good agreement [23]. In addition to measuring the VCD spectra of the LAZ for the most common isotope LA-d_0, Diem et al. [24, 25] have measured the VCD spectra of the LAZ for LA-C*-d_1, LA-Cd_3, LA-C-d_4 in both H_2O and D_2O, which results in LA-d_0 (LA-N-h_3 in H_2O) becoming LA-d_3 (LA-N-d_3 in D_2O), since the ammonium group hydrogens are exchangeable with the deuterium atoms of the solvent, D_2O. Using these data, Diem was able to parameterize his empirical force field. For empirical force field development, the frequencies from other isotopomers are invaluable. Another important application of isotopomers of peptides in various solutions is to resolve features and (or) bands in regions that overlap with other features. The importance is seen in the fact that the reproduction of spectral changes due to isotopic substitutions provides accurate tests of detailed spectroscopically determined force fields, and, in addition, of ab initio, DFT and semi-empirical based DFT force fields [26–28]. The VA and VCD spectra have been calculated for various isotopomers of L-alanine to a degree of detail suitable for comparison with the experiments of Diem [24]. With isotopic substitutions several approximations in the DFT methodology can be assessed, such as the harmonic approximation. In the case of the LAZ, the harmonic approximation appears to be inadequate to treat the CH stretch region [24]. Hence an anharmonic analysis of the VA and VCD spectra of the LAZ is being undertaken [29]. Most useful are the isotope effects on vibrational spectra that can be used for checking the validity of the BO approximation.

TABLE I

L-Alanyl L-alanine zwitterion structures: bond lengths, valence angles, and torsion angles.

Bond length (Å)	B3LYP Onsager	PW91 PCM	Valence angles (degrees)	B3LYP Onsager	PW91 PCM	Torsion angles (degrees)	B3LYP Onsager	PW91 PCM
R(C5N1)	1.495	1.499	A(N1C5H11)	105.14	105.42	τ(O12C11C5N1)	15.77	20.26
R(C11C5)	1.544	1.544	A(C5C11O12)	118.53	118.56	ψ_1(N15C13C11N1)	−163.57	−158.93
R(O12C11)	1.237	1.253	A(C5C11N13)	118.64	118.36	τ(N13C11C5C7)	−43.34	−37.94
R(N13C11)	1.346	1.346	A(O12C11N13)	122.83	123.08	τ(N13C11C5H6)	81.47	86.61
R(C7C5)	1.531	1.529	A(N1C5C7)	109.35	109.63	ω(C15N13C11C5)	163.79	165.18
R(H6C5)	1.096	1.110	A(C11C5C7)	114.86	115.19	ϕ_2(C21C15N13C11)	−64.77	−69.69
R(H10C7)	1.095	1.105	A(N1C5H6)	107.37	107.14	τ(C17C15N13C11)	171.59	166.36
R(H8C7)	1.093	1.102	A(C11C5H6)	109.25	108.43	τ(H16C15N13C11)	51.11	46.01
R(H9C7)	1.095	1.104	A(H6C5C7)	110.51	110.63	τ(O23C21C15N13)	−27.35	−23.47
R(H2N1)	1.035	1.037	A(C5C7H10)	110.46	111.06	τ(O22C21C15N13)	152.36	156.36
R(H3N1)	1.050	1.052	A(C5C7H8)	111.23	110.87	τ(H10C7C5H6)	178.14	179.42
R(N4N1)	1.051	1.061	A(H10C7H8)	107.86	108.29	τ(H8C7C5H6)	−62.11	−60.14
R(C15N13)	1.462	1.466	A(C5C7H9)	110.33	109.89	τ(H9C7C5H6)	57.71	58.77
R(H14N13)	1.023	1.035	A(H10C7H9)	108.87	108.97	τ(C11C5N1H2)	179.11	176.15
R(C21C15)	1.553	1.558	A(H8C7H9)	107.99	107.68	τ(C11C5N1H3)	−61.76	−64.79
R(O23C21)	1.257	1.272	A(C5N1H2)	114.27	113.21	τ(C11C5N1H4)	58.40	55.05
R(O22C21)	1.275	1.275	A(C5N1H3)	113.86	113.38	τ(H20C17C13H16)	178.97	179.81
R(C17C15)	1.534	1.534	A(C5N1H4)	109.87	110.90	τ(H18C17C13H16)	−60.41	−59.56
R(H16C15)	1.092	1.106	A(H2N1H3)	103.86	104.68	τ(H19C17C13H16)	59.42	59.89
R(H20C17)	1.094	1.103	A(H2N1H4)	107.35	107.63			
R(H18C17)	1.094	1.102	A(H3N1H4)	107.14	106.56			
R(H19C17)	1.094	1.104	A(C1N13C15)	117.61	118.25			
			A(N13C15C21)	109.66	109.27			
			A(C15C21O23)	118.24	117.91			
			A(C15C21O22)	117.19	117.28			
			A(O23C21O22)	124.56	124.81			
			A(N13C15C17)	111.10	111.47			
			A(C21C15C17)	111.44	111.70			
			A(N13C15H16)	108.19	108.08			
			A(C21C15H16)	106.60	106.74			
			A(C17H15H16)	109.70	109.41			
			A(C15C17H20)	111.25	111.50			
			A(C15C17H18)	109.91	109.55			
			A(H20C17H18)	108.84	108.91			
			A(C15C17H19)	110.04	109.87			
			A(H20C17H19)	107.91	108.18			
			A(H18C17H19)	108.83	108.71			
			A(C11N13H14)	121.40	121.11			
			A(C15N13H14)	117.11	117.59			

L-ALANINE L-ALANINE (LALA)

The L-alanyl L-alanine dipeptide is formed by the joining of two L-alanines, forming the peptide bond and giving off water. Again, two net neutral species are possible, the zwitterionic form (LALAZ) with the ammonium, -NH_3^+, and carboxylate, -CO_2^-, groups and a nonionic neutral form (LALAN) with the N-terminal amine, -NH_2, and C-terminal carboxylic acid (or carboxyl), -COOH, groups. We have explored both species. The zwitterionic species is present in aqueous solution at neutral pH, while the nonionic species is present in the gas phase as well as in nonpolar solvents.

We have found it necessary to solvate the N- and C-termini with explicit water molecules in order to

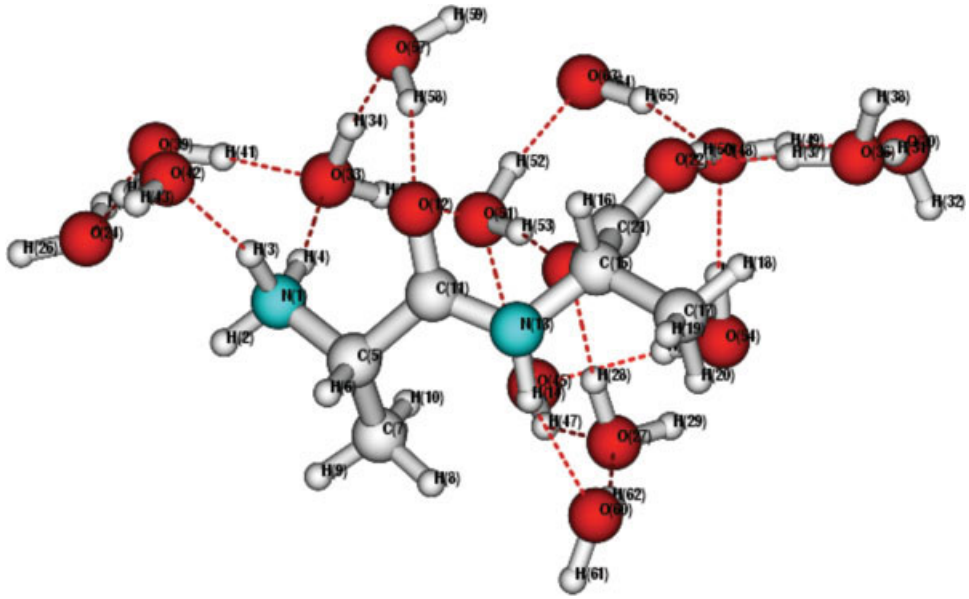

FIGURE 1. L-Alanyl L-alanine solvated with 14 water molecules, PCM PW91/aug-cc-vPDZ.

stabilize the LALAZ [30]. In addition, we have used one of the structures we have found from our modeling studies, the LALAZ + 14 water molecule complex embedded within the Onsager continuum model, to simulate the VA, VCD, Raman, and ROA spectra of LALAZW and its various isotopic species [31] and compared them with the VA and VCD spectral data available from Diem et al. [32] and the Raman and ROA spectral data from Ford et al. [33]. Previously we have investigated the effect of isotopic substitution at the two C^{α} carbons theoretically, and compared our spectral simulations at the B3LYP + Onsager level of theory to the isotopic data of Diem. We present our new results at the PW91 + PCM level of theory. The Onsager model is the simplest of the continuum models and assumes

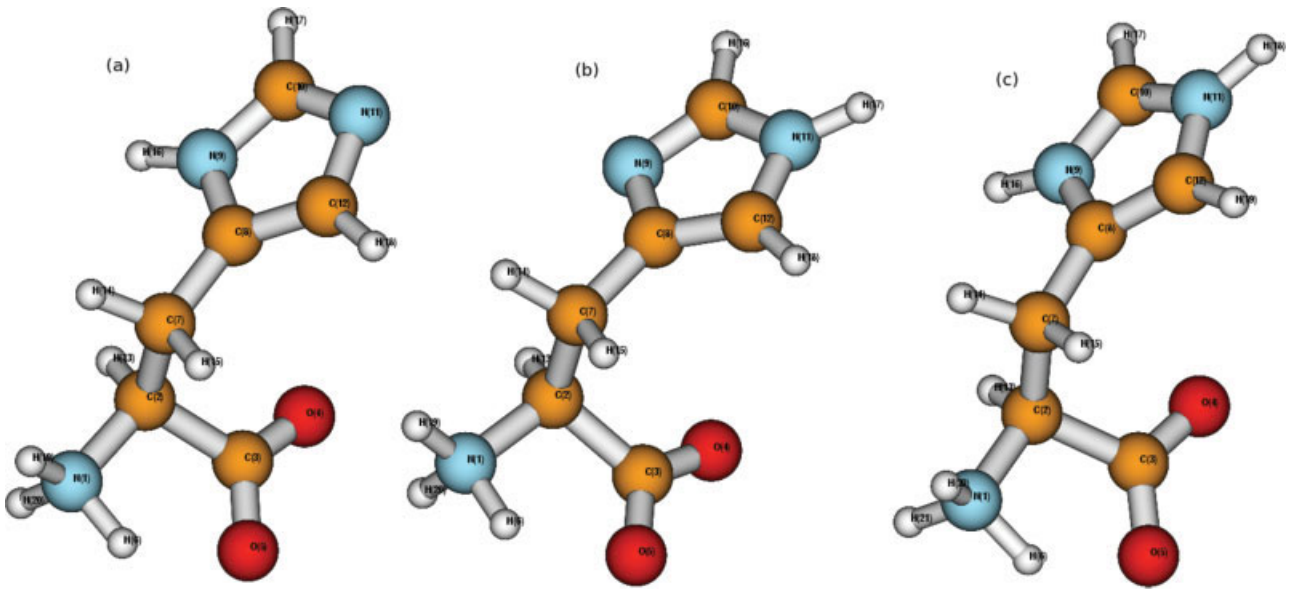

FIGURE 2. HSD (a), HSE (b), and HSP (c) structures of L-histidine.

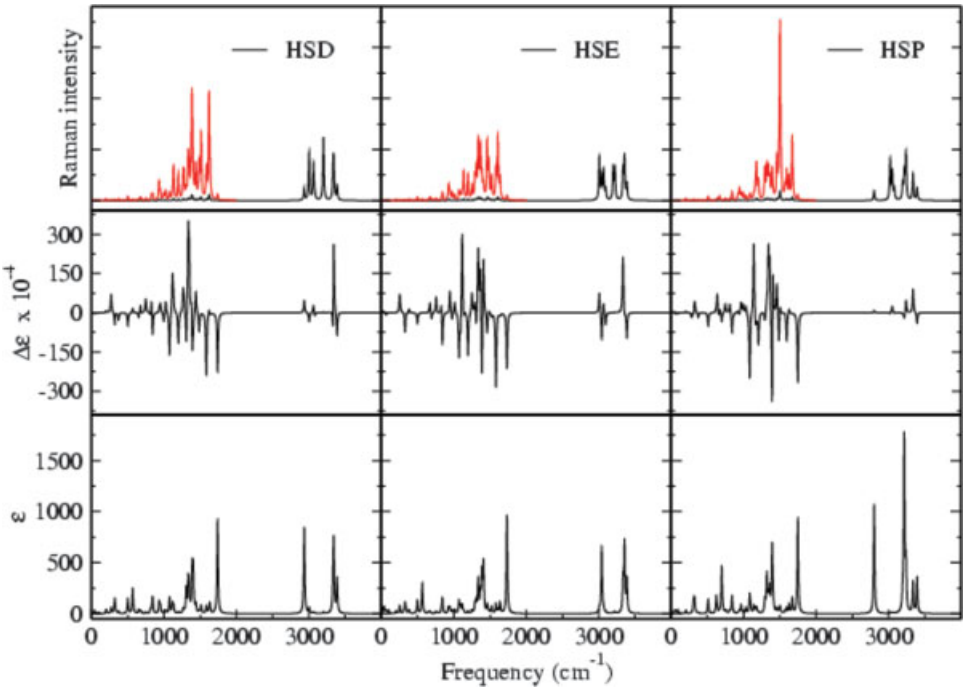

FIGURE 3. VA (bottom), VCD (middle), and Raman (top) spectras of protonated L-histidine molecules.

that the cavity within the solvent is spherically shaped, and hence one gives only the sphere radius and the dielectric constant. For the PCM model, the cavity is determined by the shape and hence conformer of the molecule. Hence it is a more realistic model; thus, it is expected that this model will simulate the effects due to the bulk solvent better, in our case water, than the Onsager model. In addition to extending the continuum model used, we have used the larger aug-cc-pVDZ basis set.

In this study, we present only the results of our geometry optimization and will report the complete set of vibrational spectra, VA, VCD, Raman, and ROA, in a future work. Table I compares our new PW91/aug-cc-pVDZ + PCM structure with our previously reported B3LYP/6-31G* + Onsager structure, and in Figure 1 we show the atom numbering. The root mean square (rms) differences between the two methods are presented in Table I: 0.009 Å for bond lengths, 0.433 degrees for valence angles and 3.434 degrees for torsion angles.

L-HISTIDINE (LH)

L-Histidine is one of the most important amino acids due, inter alia, to having a pK_a value that can change substantially with its environment. We have investigated three possible states of LH, shown in Figure 2, with two species in which the proton is on each of the two nitrogens (HSD and HSE) and then the species in which both nitrogens are protonated

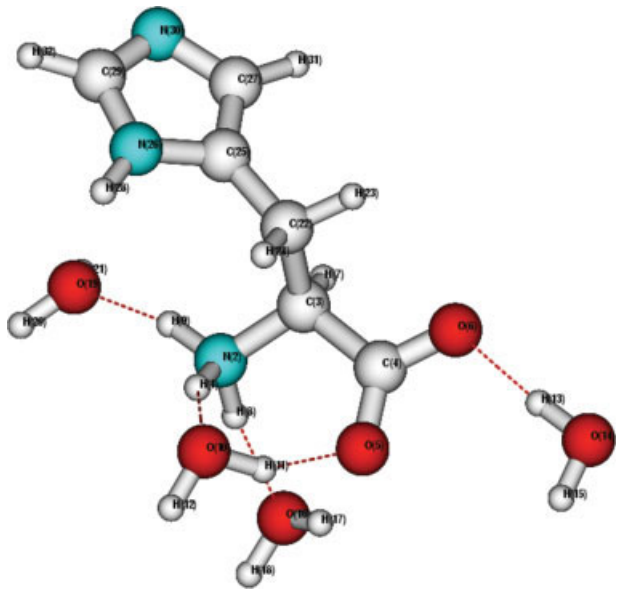

FIGURE 4. L-Histidine and four water molecules, PCM B3LYP/aug-cc-vPDZ minimum energy structure.

TABLE II

NALHNMA HSE B3LYP/6-31G conformational energies, initial structure parameters (columns 1–4), and optimized structure (columns 5–8).**

ϕ	ψ	χ_1	χ_2	ϕ	ψ	χ_1	χ_2	E (Hartrees)	E (kcal/mole)
−80	80	180	90	−83.2	74.9	−166.2	99.1	−720.90211897	3.66
−80	80	60	90	−82.1	55.9	53.5	−68.6	−720.90795537	0.00
−80	80	−60	90	−83.5	69.0	−50.2	67.9	−720.90595888	1.25
−80	80	180	−90	−83.2	74.9	−166.2	99.1	−720.90211897	3.66
−80	80	60	−90	−82.1	55.9	53.5	−68.6	−720.90795537	0.00
−80	80	−60	−90	−84.2	71.1	−56.3	−56.9	−720.90074532	4.52
180	180	180	90	−163.8	−176.6	−147.3	−62.1	−720.90750915	0.28
180	180	60	90	−160.1	−154.1	69.4	89.6	−720.90091772	4.42
180	180	−60	90	−83.5	69.0	−50.2	67.9	−720.90595892	1.25
180	180	180	−90	−163.8	−175.7	−148.6	−62.9	−720.90746703	0.31
180	180	60	−90	−168.1	157.6	60.3	−72.8	−720.89668433	7.07
180	180	−60	−90	−127.9	154.9	−72.5	−79.9	−720.89669974	6.88

(HSP) [34]. In addition, we present the VA, VCD, and Raman spectra for the lowest energy state of each species. These are shown in Figure 3, and as one can see, the vibrational spectra can be used to determine the protonation state and position of the proton for LH. This is the kind of information we are trying to determine from our modeling studies.

Previously only small parts of the VA, VCD, Raman, and ROA spectra have been fully used and appreciated. Because of the large electric dipole transition moment for the amide I mode, in many cases the concentration used for optimum VCD spectra is too dilute for studies in other regions. Also, due to the weak electric dipole transition moments for some of the other modes, one has a hard time getting a large enough concentration or path length to get the near optimal optical density (OD) of 0.6 required for the best *signal to noise* and minimal artifact for the VCD spectra. The amide group is not inherently chiral, and the VCD spectrum is therefore due to its chiral environment. The modes associated with the C^{α} carbons are inherently better because of the chiral nature of this chromophore.

Figure 4 presents the structure of LH with the ammonium and carboxylate groups solvated, and hence stabilized. In addition, the PCM continuum model has been used to treat the effects due to the bulk water molecules.

N-ACETYL L-HISTIDINE *N*′-METHYL AMIDE (NALHNMA)

To model the state of the LH in a protein, one can model the LH dipeptide, NALHNMA. In this study, we have optimized the structures of the three possible states of the NALHNMA (HSD, HSE, and HSP). In addition to seeing the spectra of the LH residue, one can also see how some of the modes change when the residue becomes part of the protein (embedded in the protein environment). Table II presents the backbone torsion angles, ϕ and ψ, and the sidechain torsion angles, χ_1 and χ_2, for the C_7^{eq} and C_5^{ext} conformers (the two lowest-energy conformers for NALANMA, where the sidechain is the methyl group) for the HSE.

Note that for the HSE structure, not all the expected local minima were stable minima, and some actually relaxed to other structures. Since the side chain is much larger than a methyl group, this is not surprising. In addition, the χ_2 torsion angle is about a bond between sp^3 and sp^2 hybridized atoms. Hence the torsion profile about this bond is more complex than that about a bond connecting two sp^3 hybridized atoms. In many cases, the torsion profile has been found to be very complex, even changing for different levels of theory. For the NALHNMA HSE species, we found the C_7^{eq} gpgm (gauche + gauche −) species to be lowest in energy. Table III presents the conformational energies for the NALHNMA HSD species.

For the HSD species of NALHNMA, the C_7^{eq} gpgp conformer is the lowest-energy structure. To date we have investigated only the side change angle dependence of the two lowest-energy conformers for the backbone angles. But it is possible that due to side chain interactions, the global minimum could be from among the other conformers,

TABLE III

NALHNMA HSD B3LYP/6-31G conformational energies, initial structure parameters (columns 1–4), and optimized structure (columns 5–8).**

ϕ	ψ	χ_1	χ_2	ϕ	ψ	χ_1	χ_2	E (Hartrees)	E (kcal/mole)
−80	80	180	90	−83.2	79.2	−165.7	90.6	−720.89797558	8.86
−80	80	60	90	−82.7	67.4	49.4	69.3	−720.91209163	0.00
−80	80	−60	90	−84.4	70.3	−58.2	167.7	−720.89843520	8.57
−80	80	180	−90	−83.3	69.5	−169.4	−70.8	−720.90724422	3.04
−80	80	60	−90	−83.0	65.3	43.2	−122.5	−720.89914477	8.12
−80	80	−60	−90	−84.5	71.6	−48.7	−67.2	−720.90046467	7.30
180	180	180	90	−162.5	173.5	−127.2	62.3	−720.90695509	3.22
180	180	60	90	−160.1	169.5	57.8	70.3	−720.89643416	9.83
180	180	60	120	−154.4	131.5	43.7	51.0	−720.90123488	6.81
180	180	−60	90	−106.9	149.4	−67.8	58.3	−720.89644061	9.82
180	180	180	−90	−160.6	158.0	−178.2	−106.3	−720.89915709	8.12
180	180	60	−90	−152.4	177.8	68.1	−59.2	−720.89826411	8.68
180	180	−60	−90	−131.0	155.4	−60.4	−20.5	−720.90265933	5.92

which are higher in energy for NALANMA, where the side chain is the methyl group. In addition, it is possible that yet a new conformer is available that is not stable without the side chain interaction. In addition, we have the HSP species of NALHNMA. Table IV presents the conformational energies for the NALHNMA HSP species.

In addition, we have simulated the effect of the water by adding explicit water molecules that are hydrogen bonded to the amide group carbonyl oxygen and NH hydrogen. We have initially started with the lowest-energy conformer for the backbone (the C_7^{eq} conformer from our NALHNMA studies above) and added the water molecules. The optimized structure is given in Figure 5. We can see that there is a hydrogen bonded network of water molecules stabilizing a structure that is not possible without an explicit water model, reinforcing our paradigm that to reliably and accurately model small peptides and amino acids in an aqueous environment one must include the first solvation shell water molecules that strongly interact with the polar groups of the parent species, and in many cases stabilize structures that are stable

TABLE IV

NALHNMA HSP B3LYP/6-31G conformational energies, initial structure parameters (columns 1–4), and optimized structure (columns 5–8).**

ϕ	ψ	χ_1	χ_2	ϕ	ψ	χ_1	χ_2	E (Hartrees)	E (kcal/mole)
−80	80	180	90	−84.0	67.8	−169.1	109.9	−721.29501011	9.87
−80	80	60	90	−83.1	75.9	60.2	54.7	−720.91209163	0.13
−80	80	−60	90	−81.8	88.7	−62.6	40.0	−720.89843520	12.12
−80	80	180	−90	−83.5	60.0	−171.6	−63.3	−720.90724422	0.47
−80	80	60	−90	−83.9	71.3	55.8	−126.0	−720.89914477	11.83
−80	80	−60	−90	−129.5	138.4	−66.7	174.1	−720.90046467	10.77
180	180	180	90	−150.7	160.0	−92.0	32.7	−720.90695509	0.00
180	180	60	90	−161.0	118.9	47.6	42.1	−720.89643416	3.02
180	180	−60	90	−138.3	160.4	−57.9	72.0	−720.89644061	11.15
180	180	180	−90	−164.6	142.1	−164.9	−43.3	−720.89915709	13.95
180	180	60	−90	172.7	166.1	41.6	−81.9	−720.89826411	6.50
180	180	−60	−90	−129.5	138.4	−66.8	174.2	−720.90265933	10.77

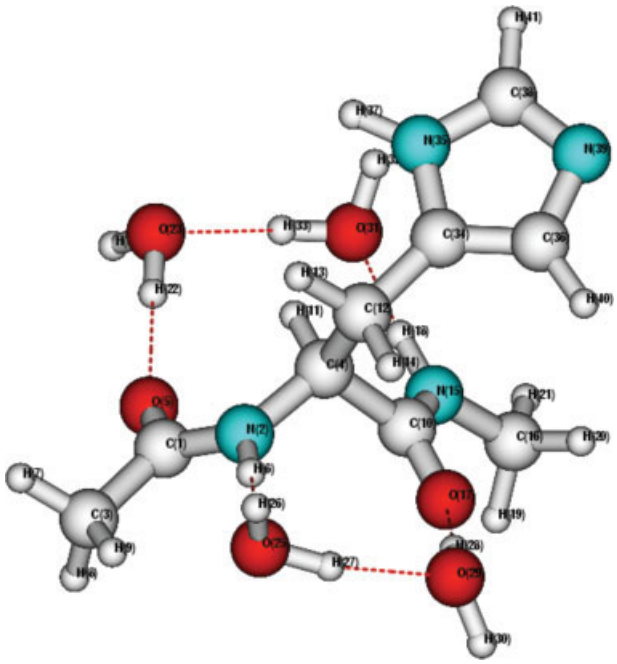

FIGURE 5. *N*-Acetyl L-histidine *N*′-methyl amide and four water molecules, PCM B3LYP/aug-cc-pVDZ optimized structure.

only because of these specific interactions. This has been confirmed by nuclear magnetic resonance (NMR) studies.

L-TRYPTOPHAN (LW)

L-Tryptophan is the amino acid with the indole side chain. Previously much work has been done on only indole and 3-methyl indole. The amino acid has two forms, depending on the environment: the so-called zwitterionic form (LWZ) and the neutral nonionic form (LWN). In addition, the cationic and anionic forms may be present, depending on the pH of the environment. Since LW can also be in the D form, one can have crystals of only LW or DW or the racemic mixture, DLW. Recently the crystal structure of DLW has been reported [35]. For LWZ, there can be four orientations of the backbone, HO, HOO, HHO, or HHOO. In addition, the side chain indole has various orientations, based on the χ_1 and χ_2 side chain angles. Since the bond between the C^{α} and C^{β} atoms is between two sp^3 hybridized atoms, there are three possible local conformers: g^+, g^-, and t. The bond between the C^{β} and C^{γ} atoms is between sp^3 and sp^2 hybridized atoms, and the number and locations of the local minimum are more complex. In this work, we present two optimized structures of LWZ determined at the PBE1 and B3LYP DFT levels of theory with the cc-pVDZ basis set and four optimized structures of LWZ determined at the PBE, PBE1, PW91, and B3LYP DFT levels of theory with the aug-cc-pVDZ basis set. The polarized continuum implicit solvent model was used to treat the effects due to the aqueous environment. Some of the more important geometrical parameters for the optimized structures are given in Table V. At the PBE and PW91 levels the zwitterionic structure was not stable with the cc-pVDZ basis set, but was with the larger aug-cc-pVDZ basis set. This is similar to our previous work on LA and LALA where the zwitterionic structures were not stable with the Onsager continuum model [18, 19, 30].

Using DFT GGA XC or hybrid XC functionals and implicit continuum models, in some cases one can obtain a stable zwitterionic structure and in other cases not. Hence it is important to be very careful with these types of calculations. An alternative approach is to add explicit water molecules to stabilize the zwitterionic structure and then use the continuum models to treat the effects due to the bulk water molecules [4]. Even though the LWZ was stable at the PBE1 and B3LYP DFT + PCM levels of theory with the cc-pVDZ basis set, the structure is not a good representation of the LWZ in aqueous solution. The evidence for this is the bond lengths of the three hydrogens attached to the ammonium group, NH_3^+. The generally accepted picture of the ammonium group is that the three N-H bond lengths are the same and that the group has local C_{3v} symmetry. The B3LYP structure has two N-H bond lengths of 1.031 and 1.033 Å and one of 1.065 Å, while the PBE1 structure has two N-H bond lengths of 1.029 and 1.031 Å and one of 1.075 Å, both with the cc-pVDZ basis set. Corresponding to these lengthened N-H bond lengths and weaker bonds, the N-H stretching frequencies are lower in frequency, being 2863 cm^{-1} for B3LYP and 2749 cm^{-1} for PBE1. The C-O bond lengths are 1.269 Å and 1.245 Å for B3LYP and 1.266 Å and 1.239 Å for PBE1, respectively, with the cc-pVDZ basis set. The corresponding C-O stretch frequencies and 1365 cm^{-1} and 1707 cm^{-1} for the B3LYP and 1411 cm^{-1} and 1756 cm^{-1} for PBE1, respectively, with the cc-pVDZ basis set. These are the most intense bands in the mid-IR spectra and are easily identified in the VA spectra. The lower-frequency band is assigned to the symmetric CO_2^- group stretch frequency and the higher-frequency band is assigned to the asym-

TABLE V

L-Tryptophan geometrical parameters.

Bond length (Å)	B3LYP/ cc-pVDZ	B3LYP/aug-cc-pVDZ	PBE1/ cc-pVDZ	PBE1/aug-cc-pVDZ	PBE/aug-cc-pVDZ	PW91/aug-cc-pVDZ
R(N1C2)	1.379	1.379	1.371	1.372	1.383	1.382
R(N1C8)	1.376	1.379	1.369	1.372	1.383	1.382
R(N1H16)	1.025	1.024	1.024	1.023	1.033	1.031
R(C2C3)	1.381	1.379	1.378	1.376	1.389	1.388
R(C3C9)	1.446	1.445	1.440	1.439	1.446	1.445
R(C9C4)	1.411	1.411	1.407	1.407	1.416	1.415
R(C4C5)	1.393	1.394	1.389	1.390	1.401	1.400
R(C5C6)	1.415	1.416	1.411	1.412	1.420	1.419
R(C6C7)	1.394	1.395	1.389	1.390	1.402	1.401
R(C7C8)	1.404	1.403	1.400	1.400	1.408	1.406
R(C3C10)	1.508	1.499	1.500	1.492	1.499	1.498
R(C10C11)	1.539	1.548	1.530	1.536	1.549	1.548
R(C11C12)	1.569	1.554	1.559	1.547	1.560	1.558
R(C12O13)	1.269	1.268	1.266	1.263	1.281	1.280
R(C12O14)	1.245	1.251	1.239	1.246	1.259	1.258
R(C11N15)	1.506	1.508	1.490	1.494	1.509	1.509
R(N15H25)	1.031	1.031	1.029	1.029	1.038	1.037
R(N15H26)	1.033	1.031	1.031	1.030	1.039	1.037
R(N15H27)	1.065	1.039	1.075	1.039	1.058	1.056

metric CO_2^- group stretch frequency, respectively. These modes can be used to determine the protonation state of the C-terminal residue. At neutral pH in aqueous solution, the state is believed to be the carboxylate (anionic) species. For the ammonium group, with C_{3v} symmetry, the group would have the same group modes as the methyl CH_3 group. As reported above, since one of the N-H bond lengths is much longer than the other two, the C_{3v} local symmetry is lost. Hence the modes will be perturbed from those found for a symmetrical NH_3 group or CH_3 group.

However, this picture changes when one increases the basis set by adding diffuse functions, e.g., the aug-cc-pVDZ basis set. Now the bond length which was longest of the three, the N15-H27 bond, is shorter and closer in length with the other two bonds, as is assumed for the NH_3^+ ammonium group. Hence the continuum model is not as bad as one would assume by using the smaller basis set, which lacks the diffuse functions. For the aug-cc-pVDZ basis set the frequencies of the three NH_3^+ stretch modes are now predicted to be 3392, 3348, and 3220 cm^{-1} with the DFT PBE1 XC functional. Note that the lowest frequency mode is now much higher in frequency and closer to what is expected for this mode. The asymmetric CO stretch mode is now calculated to occur at 1670 cm^{-1}, while the symmetric CO stretch mode is calculated to occur at 1396 cm^{-1}. Note that the lower-frequency mode also has a component due to the symmetric NH_3^+ deformation.

N-ACETYL L-TRYPTOPHAN *N*′-METHYL AMIDE (NALWNMA) AND 3-METHYLINDOLE (3MI)

3-Methylindole and *N*-acetyl L-tryptophan *N*′-methyl amide have been used as models for the L-tryptophan residue in proteins. For the study of vibrational spectroscopy of the L-tryptophan residue itself, the larger *N*-acetyl L-tryptophan *N*′-methyl amide moiety is currently being used in a manner parallel with the other amino acid studies noted above.

A separate investigation has involved the role of oxidized tryptophan in a biological setting—specifically cation and neutral tryptophan radicals that are emerging, after tyrosine, as important biological mediators of redox reactions (see, e.g., Himo and Eriksson [36]). Specifically, we have focused on the tryptophan radical formed in the heme-containing protein cytochrome c peroxidase (CCP) in collaboration with the Goodin laboratory at the Scripps Research Institute and the Bunte laboratory at the

U. S. Army Research Laboratories. This protein reacts with hydrogen peroxide to form a two-electron oxidized intermediate. This intermediate, in turn, accepts electrons sequentially from two ferrous cytochromes c. One oxidizing equivalent is stored as an oxyferryl (Fe(IV)═O) heme, while the other is stored as a radical species on the tryptophan 191 sidechain. This work builds on earlier site-directed mutagenesis experiments and an elegant isotopic enrichment/magnetic resonance experiments performed by the Goodin and Hoffman laboratories [37–41]. While the evidence favored assignment as a cation radical, rather than as a neutral radical, theoretical estimates of spin densities to match with those from the experimental work [41] were sought. Hybrid DFT methods of reasonable sophistication and proven accuracy in the modeling of ground-state properties of up to medium-size molecules were employed [42–44]. Specifically the well-established B3LYP functional and a TZ/2P basis set [45] was used to determine spin densities and partial charges for 3-methylindole. Spin densities calculated for the cation radical matched very well with those for CCP and for other known indole cation radicals [45]. This made sense in light of the structure of the CCP protein, which has a so-called histidine-aspartate-tryptophan triad where the histidine is the porphyrin ligand, and the tryptophan is adjacent to this histidine, and both amino acids are available to interact with a nearby aspartate sidechain. The ostensibly anionic aspartate can afford electrostatic stabilization to a cation at the location of the oxidizing tryptophan, and indeed when this tryptophan is modified by site-directed mutagenesis to hydrogen in CCP (tryptophan 191 mutation to glycine), a cationic binding cavity is created [46–49]. Thus, as noted in CCP, tryptophan-191 is preferentially oxidized to a cation radical to house the second oxidizing equivalent noted above. Another similar protein, ascorbate peroxidase (APX), which has a very similar molecular structure to CCP, including the histidine-aspartate-tryptophan triad, nevertheless does not exhibit an indole radical upon reaction with hydrogen peroxide. Instead, a porphyrin cation radical is formed [50, 51]. This difference in location of the stable second oxidation equivalent may be the cause of the different substrate preferences of the two peroxidases: large proteins for CCP (e.g., cytochrome c) and small molecules for APX (e.g., ascorbate). In many other plant peroxidases, this is accomplished by replacement of the triad tryptophan with another residue, such as phenylalanine, which is not so easily oxidized. Clearly, the presence of the aspartate alone in this triad is not sufficient to dictate oxidation of tryptophan in this setting. A modeling study was carried out of the oxidation of tryptophan using the x-ray crystal structures of CCP and APX, atomic partial charges determined as noted above, and an electrostatic/molecular dynamics program to model the protein and solvent [52]. Calculations of the redox potential of the indole to indole cation radical oxidation on tryptophan-191 of CCP and the analogous tryptophan-179 of APX reveal that the redox potential of the APX indole oxidation is ~350 mV higher than the equivalent reaction in CCP. This is sufficiently higher in APX to effectively account for the absence of an observed indole cation radical in APX. A bound monovalent cation observed in the APX structure ~9 Å from the triad indole and absent in the CCP structure is partly, although not completely, responsible for the calculated increase in indole redox potential in APX relative to CCP, with other more structurally subtle factors comprising the remainder of the difference. For these calculations, the atomic partial charges used for this study were determined using the electrostatic potential methods developed by Kollman and coworkers [53, 54]. Protein modeling used the Polaris protein–dipoles Langevin–dipoles (PDLD) suite of programs, which has had good success for these types of modeling schemes [55].

The demonstrated accuracy of hybrid DFT methods noted above, including, inter alia, vibrational and Raman spectra prediction, has been extended to include open-shell systems [56]. The Raman bands of tryptophan residues are known to be of sufficient intensity to be readily identified in the spectra of proteins [57]. If the tryptophan radicals can be shown to exhibit bands of adequate intensity, and the concentration and lifetime of the radical is such that the radical bands can be observed, the predicted Raman spectra of the radicals could provide possible specific spectroscopic markers for the detection of neutral or cationic tryptophan radicals in biological systems, thus providing a complement to the data available from other methods.

We calculated vibrational absorption and Raman spectra for 3-methylindole cation radical and 3-methylindole neutral radical, with the hope that these predictions will allow for the species to be identified and specific bands assigned [58]. The geometries and harmonic vibrational frequencies of 3-methylindole, the neutral radical, and the cation radical were calculated using spin-restricted DFT with the hybrid B3LYP functional for the closed

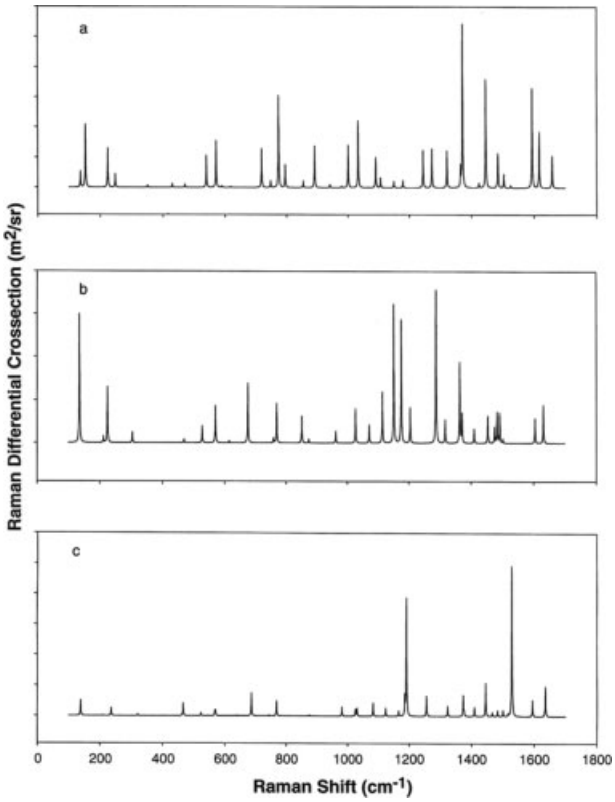

FIGURE 6. Calculated Raman spectra of (a) 3-methylindole, (b) the neutral radical of 3-methylindole, and (c) the cation radical of 3-methylindole (see Ref. [58]).

shell 3-methylindole and spin-unrestricted DFT with the B3LYP functional for the open-shell neutral and cation radicals; a TZ2P basis set was used to represent the molecular orbitals [58]. The polarizability derivatives of 3-methylindole, the neutral radical, and the cation radical were calculated at the Hartree–Fock level of theory with the 6-311++G(2d,2p) basis set [58]. The Hartree–Fock (or DFT) polarizability derivatives can be combined with the displacement vectors from the B3LYP/TZ2P Hessian to predict the Raman scattering activities. The resulting Raman spectral predictions are shown in Figure 6 for the 3-methylindole moiety and for the two oxidized radicals. Good agreement between modeled and measured spectra was observed for 3-methylindole, but no data can yet be adduced to establish the accuracy of the theoretical spectra (Raman or vibrational) for the cation radical. It is hoped that such data can be generated, e.g., for the cation radical in CCP, where only the one tryptophan among others in the protein oxidizes and where such oxidation can be controlled under well understood experimental conditions and monitored under orthogonal experimental methods, e.g., ESR and UV-Vis absorption spectroscopy. For the neutral radical, a communication has been recently published that uses our study results along with others to interpret experimental Raman measurements on the neutral tryptophan radical in DNA photolyase [59].

In the systems described, we use the 3-methylindole moiety as the model for tryptophan. When including aspects of the CCP protein, we married the DFT modeling to semi-empirical and other electrostatic modeling schemes [52]. While this has been effective in the cases noted, it is a limited approach. One might desire to incorporate more of the real protein structure, e.g., starting from the x-ray-derived structure, in DFT calculations. This approach is not trivial, however. For example, in CCP, inclusion of the aspartate moiety in calculations of the aspartate–tryptophan cation radical system leads to what is the obvious energy minimum for this system: an electron transfer from the aspartate to the tryptophan radical. (This is somewhat analogous to the cases above of two net neutral forms of our model amino acids). There is no evidence at all that this can occur in the real protein, and thus this is an artifact of incomplete inclusion of all elements of the natural system, e.g., the whole protein, the solvent, and proper understanding and modeling of the dynamics of the system and the physics of the electron transfer processes. It is possible that this tendency to create artifacts on energy minimization may be responsible for the theoretical observation of spin delocalization in a model of the tryptophan radical that includes peptide bond moieties [60]. There is no direct experimental confirmation of such delocalization at this time. Reliable inclusion of all relevant physical factors in modeling systems, or full appreciation of not doing so, is critical to the ongoing development of even more complicated theoretical systems.

N-ACETYL L-ALANINE *N*′-METHYL AMIDE (NALANMA) AND TRI-L-ALANINE

To treat the effect of the backbone atoms of the protein and to have the effects due to the ϕ and ψ backbone torsion angles, one can cap the N-terminus with the *N*-acetyl group (CH_3CO-) and the C-terminus with the *N*′-methyl amide group (-$NHCH_3$). NALANMA has been studied extensively by many groups, first as a test of empirical force fields [61], then as a test of the hybrid DFT

B3LYP XC functional [3], and finally as a test of the semi-empirical based DFT SCC-DFTB method [28] for use in reproducing the energetics and vibrational spectra of this molecule in the isolated state. The changes that occur in the VA, VCD, Raman, and ROA spectra have been used to determine the changes in structure of NALANMA that are responsible for these spectral changes, for the species present in the inert gas matrices and nonpolar solvents [3] to those found in aqueous solution [4]. Here the conformer from the modeling study, similar to the work mentioned previously to identify the conformer of the LAZ, was only present by including explicit water molecules in the study [4]. Attempts to use continuum models to treat systems that have strong hydrogen bonding interactions have noted that the vibrations are in error due to these specific hydrogen bond interactions [62]. This structure has subsequently been confirmed by nuclear magnetic resonance (NMR) studies [63]. Previous attempts to interpret the NMR signals by using structures that are a local minimum for the isolated peptide failed [64]. This has opened up a new paradigm for interpreting the spectra, VA, VCD, Raman, ROA, electronic absorption (EA), electronic circular dichroism (ECD), and NMR, for biomolecules. To understand the structure and function of the biomolecules, one must explicitly include the strongly hydrogen-bonded water molecules, as they help stabilize solute + solvent (other species present in the environment) complexes, in which the solute has conformers that are not stable without the environment [65]. This is similar to the so-called induced fit model for ligand binding of ligands to proteins, carbohydrates, or nucleic acids [66]. Here one finds that the "sequence-determines-structure" paradigm only works for the unbound (ligand free) state of the protein. To understand what happens during ligand binding, one must be able either to crystallize the protein with the ligand bound, or to monitor the spectral changes that occur during ligand binding (or turnover), and also to be able to interpret these changes. The vibrational spectra, VA, VCD, Raman and ROA, have been used to verify that this is indeed the case and to also monitor this process. Previous modeling studies that attempted to treat the solvent as only a continuum have failed. But these continuum models (or other more sophisticated models of bulk solvent) have been shown to be useful to treat the bulk solvent effects, after the solute + strongly interacting species (e.g., solvent, cation, anion, metal) complex has been identified. The key point is to be able to differentiate between the solvent and spectator ions or species and those that are a necessary and inherent part of the complex (not only its structure, but also its function and functional states). In addition, in comparative (homology) modeling of sequences, one must first identify the fold class, then perform either single or multiple alignments, and finally use the structures of the aligned sequences to get the 3D structures. The hypervariable loops are what are normally the most difficult to model; here the ab initio and DFT modeling and vibrational spectroscopy can contribute by being able to identify the secondary structure of specific residues with isotopic labeling [67].

In another application of the ab initio, DFT, and experimental data for NALANMA, artificial neural networks were used to predict 3D structures of peptides on the basis of spectroscopic data. To do that, one has to build up a database of spectroscopic data that can be used as a training set. Such a set constitutes the input to a neural network that has to "learn" to infer which structures could give rise to the corresponding spectra. This is very much analogous to the problem of inverse scattering; however untractable this problem appears to be in mathematics, it is reasonably well managed with neural networks. In this specific case, neural networks are used for classifying the conformational states from the vibrational spectra of *N*-acetyl L-alanine *N*′-methyl amide. The neural networks were shown to be capable of distinguishing the different conformers of NALANMA, especially in the gas phase [68]. Another application of neural networks has been to predict the protein hydration sites from sequences [69]. Hence the combined use of high-resolution x-ray and neutron diffraction data, NMR, and vibrational spectra along with high-level ab initio, DFT, and semi-empirical-based DFT modeling that includes explicit treatment of solvent molecules may be the way to fully understand and interpret all the available data and experiments self-consistently. In addition neural networks have been trained on ab initio data with the goal of producing a polarizable force field for water, which can be used for liquid state water simulations and also solutes, ligands, drugs, proteins, carbohydrates, and nucleic acids in aqueous solution. The complex many-body effects have been dealt with by using this methodology [70].

A recent ab initio molecular dynamics study of isolated NALANMA has shown that the time for conversion of the C_7^{eq} conformer to the C_5^{ext} conformer occurs on the picosecond timescale, in con-

trast to classical molecular dynamics simulations where the conversion did not occur even after nanoseconds [71]. In addition, our work on deriving a class two force field for proteins has shown that the simple functional form from class one force fields is inadequate and that even some of the local minima preprogrammed into the force field parameters are not even stable local minima. Hence one must be very careful if one uses the class one force fields where one only uses harmonic potentials for bond stretches, valence angle bends, a three-term Fourier series for the torsional potentials, a simple point charge model for electrostatics, and the van der Waals potential for nonbonded interactions. The recently developed force fields that treat anharmonicity in bond stretching and valence angle bending, polarizable charge distributions that allow for charge transfer and cross terms are much preferred if one wishes to have more than qualitative results. In addition, explicit treatment of the solvent is very important and the solvent force field should be of comparable complexity, that is, a polarizable water model is to be preferred. MacKerrell [72] recently presented a review of empirical force fields used in modeling biological macromolecules.

Here we extend our previous work by replacing the way we treat the bulk water. Previously we used the Onsager continuum model. In this work, we replace the Onsager continuum model with the polarized continuum model (PCM). Here the cavity is not spherical, but is generated from the shape of the molecule and hence is different for different conformers. Figure 7 shows our NALANMA plus 4 water molecules optimized structure at the PW91/aug-cc-pVDZ + PCM level of theory. The backbone torsion angles ϕ and ψ are -93 and 132 degrees, respectively. Hence once again the local minimum for the dipeptide is not from among those found for the isolated state. This structure is between the C_7^{eq} and the C_5^{ext} structures found for that case. In a future work, we will present our new VA, VCD, Raman, and ROA spectra for this NALANMA–water complex. The waters are necessary to stabilize this structure.

It is also interesting to see what the effects will be due to the various ways to treat the bulk water. Previously we have shown that some treatment of the bulk water is very important. What has been a complexity to date has been how to choose the number of waters and their positions. With the Onsager continuum model, one is forced to have a spherical cavity of waters if one wishes to have a full cavity. With the PCM model, this requirement

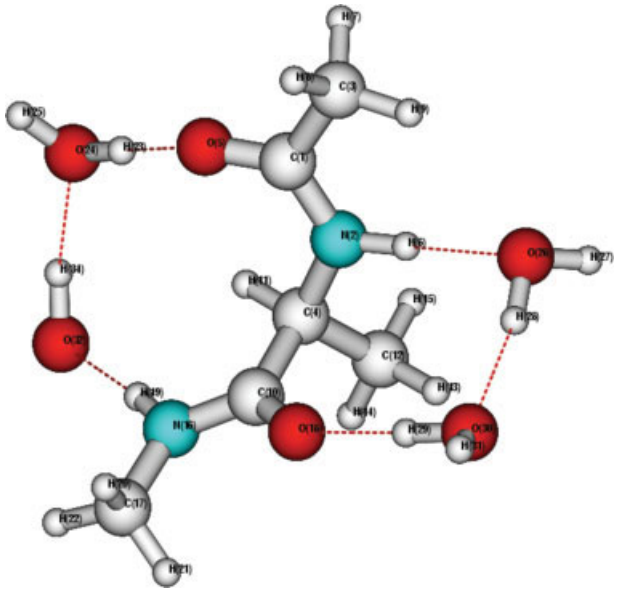

FIGURE 7. NALANMA and 4 water molecules, PCM continuum model at PW91/aug-cc-pVDZ level of theory.

is not needed. Another problem has been that the waters tend to clump together to maximize the number of hydrogen bonds. Hence, it is important to have a network which is stable for this method to work. Otherwise, the positions of the waters obtained from molecular dynamics (MD) simulations are not stable, and they will just clump together and the MD determination of their positions is for naught. Another inherent problem is with the nature of the water potential used in most MD simulations. Buckingham and Stone [73–75] have shown that an accurate potential is necessary to reproduce the structures of hydrogen-bonded water complexes, beyond the simple point charge model used in most MD simulations. Hence the solvation shell that one gets from MD simulations may be very different from that one would obtain from an ab initio MD simulation.

Tri-L-alanine (triLA) is a simple tripeptide. The infrared (IR) VCD spectra has been reported by Diem [76]. In this work the coupled oscillator model has been used to interpret the amide I (C═O stretch) and amide II and amide III (NH/CH deformation) modes. Similar to the previously reported work on LALAZ [5], there is coupling between the amide I modes and the $C^{\alpha}CH$ deformations.

In addition, Woutersen and Hamm [77] have used 2D vibrational spectroscopy to determine the ϕ and ψ backbone angles of the middle amino acid

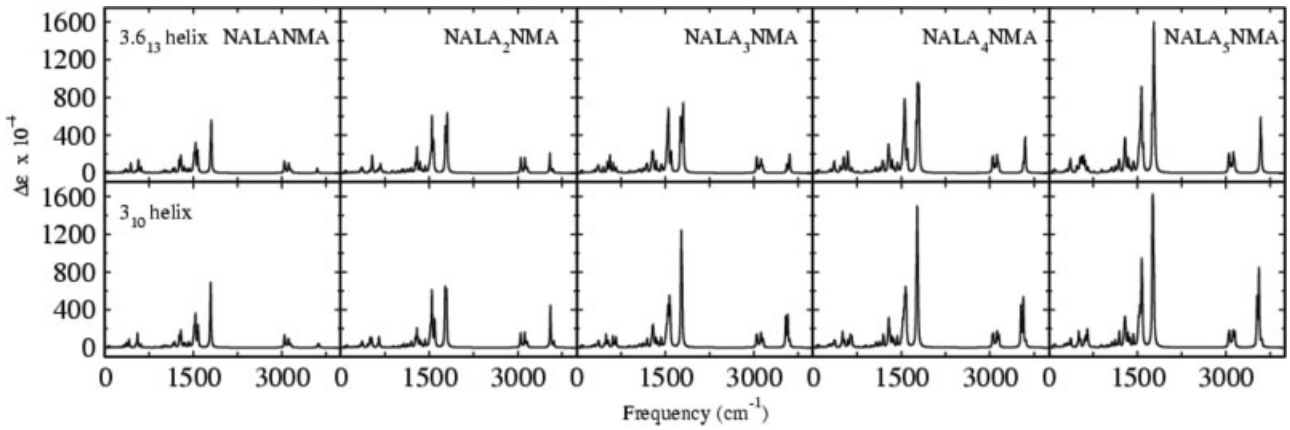

FIGURE 8. B3LYP/6-31G* VA spectra for 3.6_{13} and 3_{10} helices for NALA$_n$NMA, n = 1–5. Top set of spectra are for 3.6_{13} helix, bottom set are for 3_{10} helix. Left (first) is NALANMA, second is NALA$_2$NMA, third is NALA$_3$NMA, fourth is NALA$_4$NMA, and fifth is NALA$_5$NMA.

residue. This information can be obtained from an analysis of the amide region of the spectra. To determine the other backbone torsion angles requires additional information, which can be gained via simulations of the VA, VCD, Raman, and ROA spectra of various conformers of triLA under various conditions. Again, the problem of which conformer and species (anion, zwitterion, or cation) will depend on the environment (polar versus nonpolar) and also pH. To model these in aqueous solution at various pH requires a treatment of the aqueous environment, both water molecules strongly hydrogen bonded with the triLA and also the bulk waters. To model the backbone angles, one can once again cap the N- and C-termini to get the pentapeptide, *N*-acetyl (L-alanine)$_3$ *N*′-methyl amide. This allows us to probe the amide modes without the interfering *N*- and C-termini which have to be dealt with using explicit water models. In addition to *N*-acetyl (L-alanine)3 *N*′-methyl amide, we present the VA and VCD spectra for NA(LA)$_n$NMA, n = 1, 2, 4, and 5.

In Figure 8 one can see the enhancement of the amide I bands as one increases the number of residues. In Figure 9, one can see that the characteristic bisignate VCD feature observed in alpha helical proteins is already seen in the alanine dipeptide, also called, NALANMA. These calculations are important, since the α-helix is a characteristic secondary structure found in membrane proteins, the

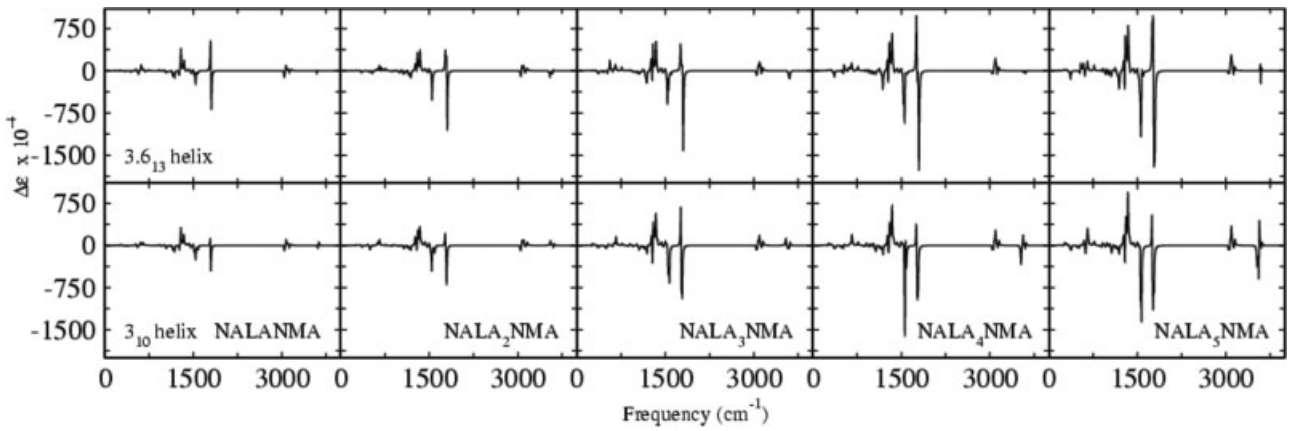

FIGURE 9. B3LYP/6-31G* VCD spectra for 3.6_{13} and 3_{10} helices for NALA$_n$NMA, n = 1–5. Top set of spectra are for 3.6_{13} helix, bottom set are for 3_{10} helix. Left (first) is NALANMA, second is NALA$_2$NMA, third is NALA$_3$NMA, fourth is NALA$_4$NMA, and fifth is NALA$_5$NMA.

TABLE VI

Relative energies of 3.6_{13} and 3_{10} helices for NALA$_n$NMA, n = 1–5, and hydrogen bond lengths at B3LYP/6-31G* level of theory.

	NALANMA	NALA$_2$NMA	NALA$_3$NMA	NALA$_4$NMA	NALA$_5$NMA
3.6_{13}	0.929 kcal/mole	3.64	4.93	5.78	6.30
D(C═O· · ·H—N)	3.193 Å	2.218	2.353	2.305	2.277
			2.328	2.240	2.203
				3.324	2.337
					2.304
3_{10}	0.0 kcal/mole	0.0	0.0	0.0	0.0
D(C═O· · ·H—N)	2.919 Å	2.050	2.132	2.102	2.089
			2.163	2.275	2.234
				2.148	2.269
					2.134

α-helical side chains being hydrophobic and interacting with the nonpolar lipid groups. In addition to the characteristic modes observed for the 3.6_{13}- and 3_{10}-helices, we are also interested in identifying characteristic features of other secondary structural elements. The relative energy of the 3.6_{13} and 3_{10} helices calculated at the B3LYP/6-31G* level of theory is presented in Table 6. Previously we have shown that neither the 3.6_{13} nor 3_{10} helix is stable with respect to unwinding at the end groups. Here we have frozen the ϕ and ψ angles at -57 and $-47°$ for the 3.6_{13} helix and at -60 and $-30°$ for the 3_{10} helix. In addition, we give the lengths of the hydrogen bonds between the oxygen of the amide group CO and the hydrogen of the amide group NH. Here we see that for all NALA$_n$NMA molecules, for n = 1–5, the 3_{10} helix is lower in energy due to the stronger hydrogen bonds.

Note that also in the NALANMA species, where there are no real hydrogen bonds, the structure with the ideal 3_{10} helix ϕ and ψ values is lower in energy than the structure with the ideal 3.6_{13} helix ϕ and ψ values. For example, the β turn is a commonly occurring secondary structural element that can connect two helices or anti parallel β sheets. But to have the turn allowed, one often requires a glycine reside within the turn.

TRI-L-SERINE (triLS)

Tri-L-serine is a tripeptide made up of three L-serine residues. The VA, VCD, and Raman spectra of triLS have been reported in the 1550–1750-cm^{-1} region by Eker et al. [78]. TriLS has three OH groups that form hydrogen bonds, in addition to the NH_3^+ ammonium and CO_2^- carboxylate groups (zwitterion) and the NH_2 amine and CO_2H carboxylic acid groups. An ab initio study of LS in the isolated state, nonzwitterionic state has also appeared [79]. In the isolated state, the conformers which have intermolecular hydrogen bonded are extra stabilized.

But when LS is in aqueous solution there is a competition between intramolecular hydrogen and intermolecular hydrogen bonding, similar to what is present when a polar ligand or drug molecule is in aqueous solution in the presence of a protein molecule. So what is important is the total energy of the system. Here we need to take not only the hydrogen bond energy, but also temperature and entropy. Figure 10 presents our PW91/aug-cc-pVDZ + PCM optimized structure. The following backbone angles are $\phi_{1(HHH)}$ = 41.4, −77.2, 162.7, ψ_1 = 154.3, χ_{11} = −178.5, χ_{12} = −73.8, ω_1 = 158.10, ϕ_2 = −66.5 ψ_2 = 157.7, χ_{21} = 175.2, χ_{22} = −72.9, ω_2 = 158.5, ϕ_3 = −91.4, $\psi_{2(OO)}$ = 10.4, −170.4, χ_{31} = −174.0, χ_{32} = −41.3. Note that the two carboxylate oxygen atoms are both hydrogen-bonded, to the hydrogen of the OH group of residues 1 and 2, while the oxygen of the second residue is hydrogen bonded to the hydrogen of the OH group of residue 3. The amide and ammonium group hydrogens are not hydrogen bonded as well as the backbone oxygen of residue 3's amide group. The hydrogen bond lengths are 1.81 Å, 1.74 Å and 1.81 Å, respectively. The two carboxylate group CO bond lengths are now 1.262 and 1.266 Å, respectively, while the three ammonium group NH bond lengths are 1.036, 1.034, and 1.035 Å, respectively.

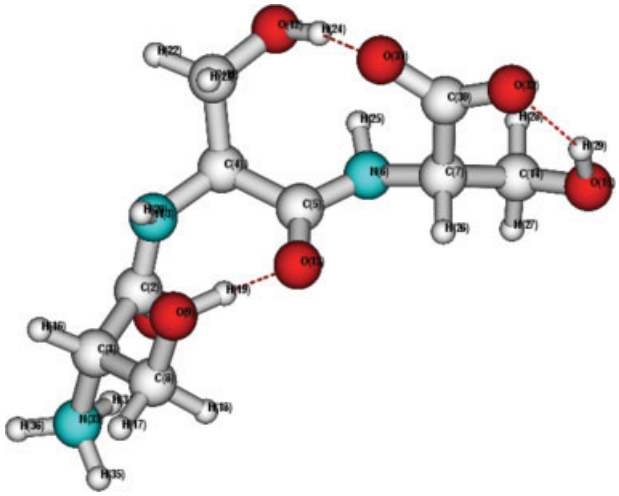

FIGURE 10. Tri-L-serine zwitterions, PCM PW91/aug-cc-pVDZ.

The three OH bond lengths of the three serine residue side chains have bond lengths of 0.984, 0.990, and 0.984 Å, respectively. All three of the OH group hydrogens are involved in H-bonding, hence are a bit longer than those not involved in hydrogen bonding. The two amide group NH bond lengths are 1.028 and 1.026 Å, respectively, while the two amide group CO bond lengths are 1.237 and 1.245 Å, respectively. The second CO bond, which is a bit longer, is involved in a hydrogen bond interaction with the hydrogen of first serine residue. Finally, the two peptide bond lengths are 1.345 and 1.344 Å, respectively. If the two end residues of a protein are serine residues, a very stable hydrogen bonding structure between the carboxylate group oxygens and serine residue OH group hydrogens can form as shown in Figure 10.

The VA and VCD spectra for this optimized structure, also calculated at the B3LYP/aug-cc-pVDZ + PCM level of theory, are presented in Figure 11. In addition to the continuum treatment of the solvation effect and hydration, we have treated the effects due to hydration and solvation by adding 22 explicit water molecules which were initially directly hydrogen bonded to the polar groups in triLS. We have also treated the effects due to the bulk water molecules by embedding the complex within a spherical solvent cavity, the so-called Onsager model. Since our experimental vibrational spectroscopic measurements were made initially at low pH, we have modeled the cationic species. To make the system neutral, we have added a solvated Cl^- anion.

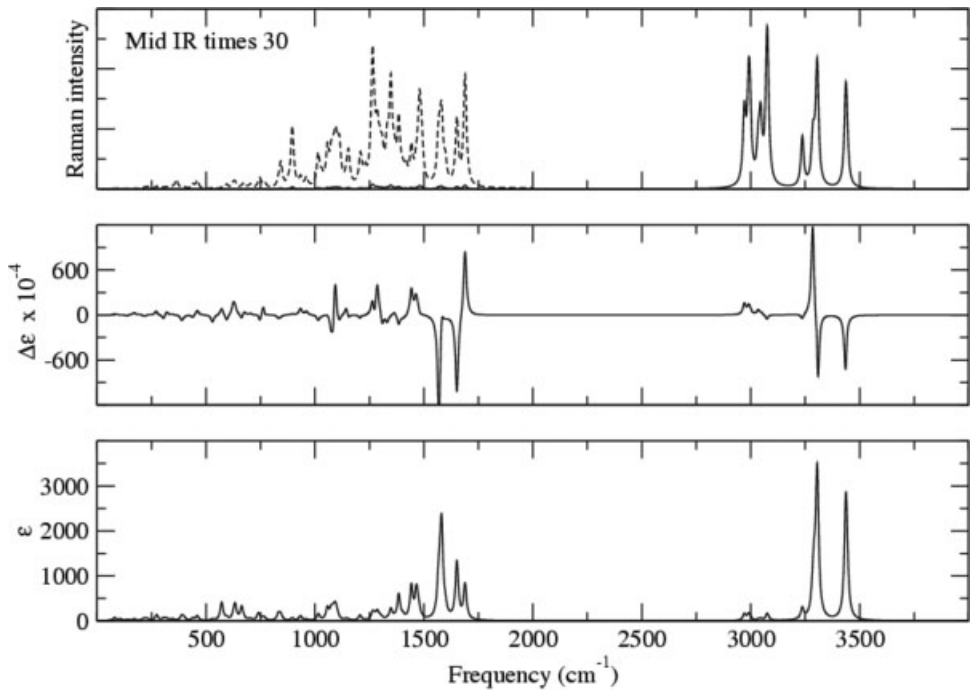

FIGURE 11. VA, VCD, and Raman spectra of tri-L-serine zwitterion, PCM B3LYP/aug-cc-pVDZ.

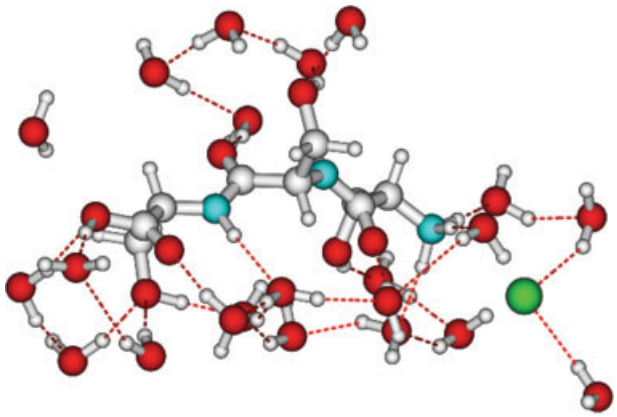

FIGURE 12. Tri-L-serine cation with Cl anion solvated and 22 water molecules, B3LYP/6-31G* level of theory.

Our optimized structures for the explicitly solvated system and the explicitly solvated system within the Onsager continuum model are given in Figures 12 and 13, respectively. In Figures 14 and 15, we present the VA, VCD, and Raman spectra for both species. Here one can see that one also has the water modes that cause real problems in certain regions, which are impossible to see in any continuum model. In addition, one can see the modes which result from the hydrogen bonded interactions between the various water molecules and also between the water molecules and the polar groups in triLS. This is a major new result from this work and will be explored in more depth in a future publication [80]. Figure 16 presents the experimentally measured VA, VCD, and Raman spectra. By comparing the VCD spectra in the region where we have made the measurement with the three theoretical simulations, we see that the model with the explicit water molecules and the Onsager continuum, our hybrid model gives the best agreement. This is consistent with our previous work on LA, NALANMA, and LALA. In addition the PCM model appears to give some of the main features, but it clearly does not do as well as the hybrid model. Hence from this work we can advocate the use of explicit water molecules to treat the effects due to the strongly interacting water molecules and the continuum solvent model to treat the effects due to the bulk water, as both are important.

APY

The tripeptide L-alanine-L-proline L-tyrosine (APY) has an interesting turn structure that has been identified by NMR spectroscopy [81]. Here one can also investigate this tripeptide via vibrational spectroscopy. The L-proline residue can be in two isomeric states, the *trans* or *cis* form. From the NMR spectra, one can determine the structure of the Ala-cisPro-Tyr peptide [81]. From molecular dynamics (MD) simulations for the Ala-transPro-Tyr peptide, we have extracted 6 conformers [82]. For these conformers we have performed geometry optimizations and at the optimized geometries, calculated the Hessians, APT, AAT, and EDEDPD that allow us to simulate the VA, VCD, and Raman spectra. In addition, we are also performing ROA simulations. The simulations allow us to determine if the structures from the molecular dynamics simulations are able to reproduce the measured VA spectra. From MD simulations for the Ala-cisPro-Tyr peptide, we have extracted two conformers. In addition, we have optimized the structure of Ala-cisPro-Tyr peptide starting from a molecular mechanics optimized structure, with the CFF91 protein force field [61, 83]. An alternative path to determining the structure of this and other molecules is to measure the experimental spectra and try to deconvolute the bands in the amide I, II, and III regions. Using the frequencies and band intensities extracted from the spectra, one is able to assign the backbone angles based on frequencies and band intensities which were extracted from the spectra of x-ray and NMR solved peptides and proteins. Here one has assumed that the spectra one extracts can then be used to generate sub-spectra which can be used in simulating spectra of unknown structures, and the structures will be

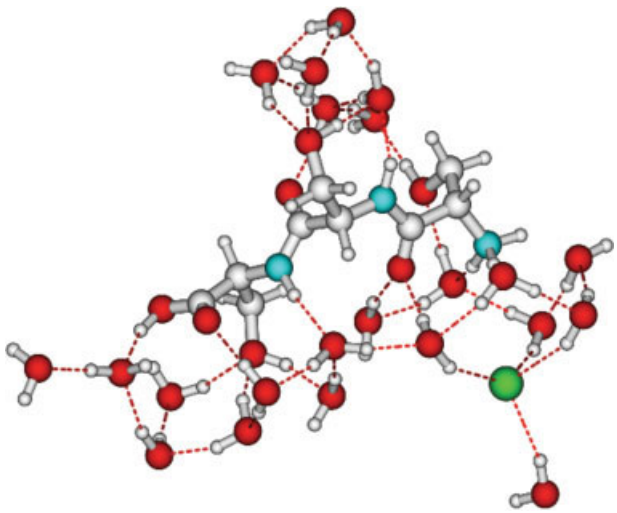

FIGURE 13. Tri-L-serine cation with Cl anion solvated with 22 water molecules, Onsager B3LYP/6-31G*.

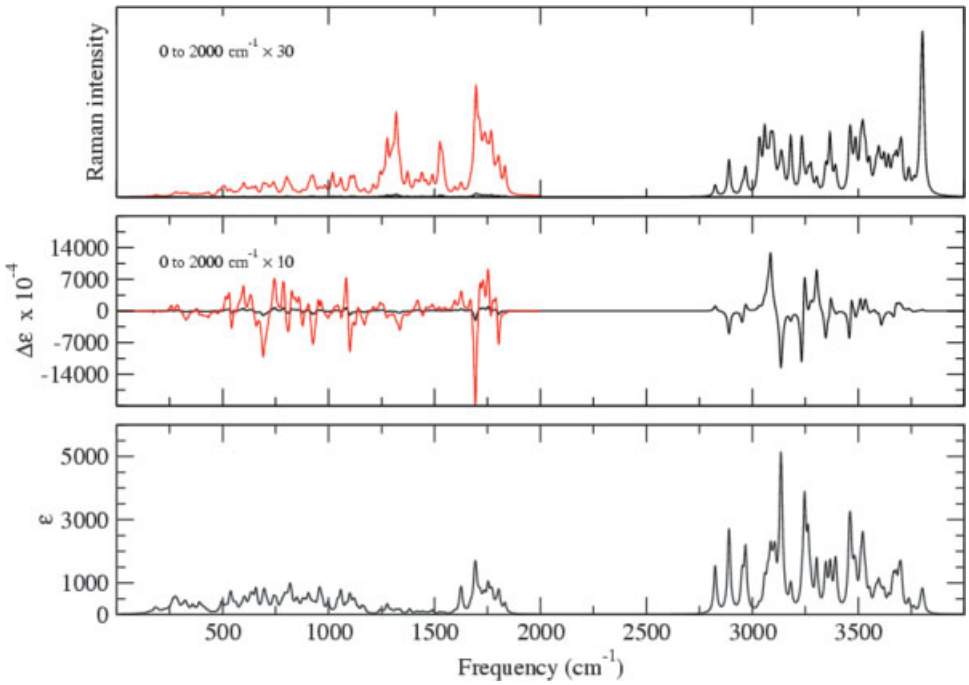

FIGURE 14. VA, VCD, and Raman spectra for tri-L-serine cation with Cl anion solvated with 22 water molecules, B3LYP 6-31G*.

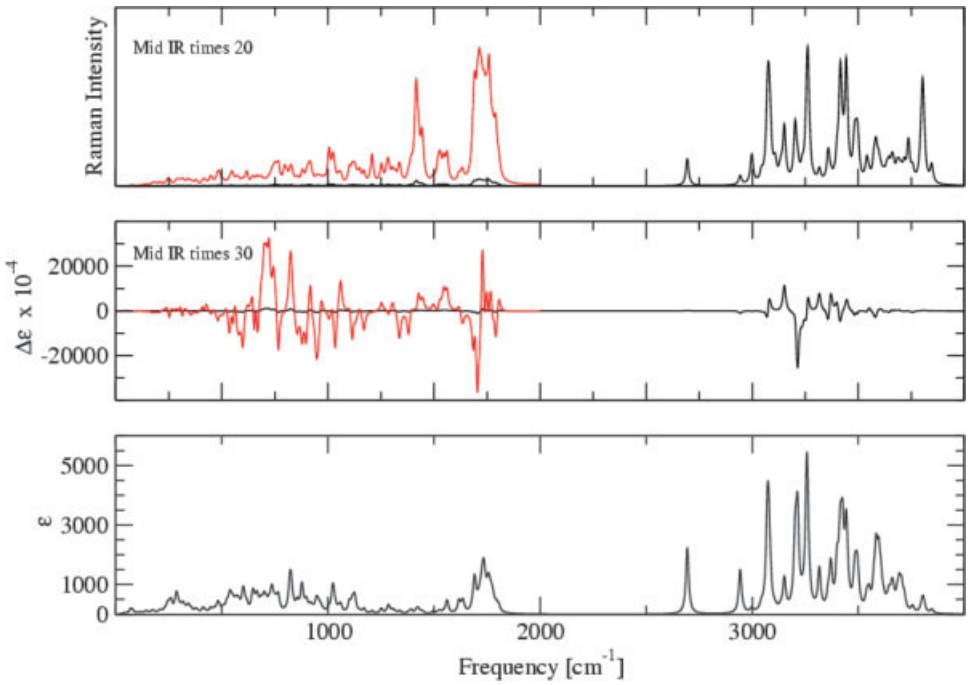

FIGURE 15. VA, VCD, and Raman spectra for tri-L-serine cation with Cl anion solvated with 22 water molecules, Onsager B3LYP 6-31G*.

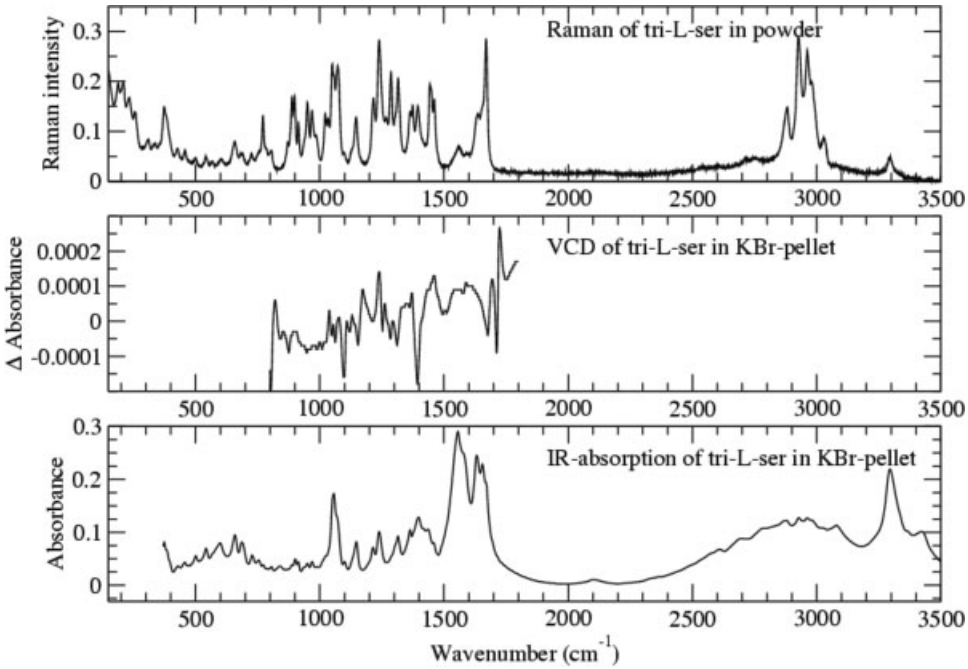

FIGURE 16. Experimental VA, VCD, and Raman spectra for tri-L-serine.

consistent with the parameters of the sub-spectra which are best able to reproduce the experimental structures. The percentage of secondary structural features has been determined, but for small peptides, where the amount of secondary structure is quite limited, the structure can be assigned.

LEU-ENKEPHALIN

Leu-enkephalin is a very interesting pentapeptide that appears to be very flexible, with its conformational states being very sensitive to its environment. To date more than one crystal structure has been reported [84]. To study a small peptide or a fragment of a protein that itself is biologically active the pentapeptide Met/Leu-enkephalin is an excellent test molecule. It is a peptide molecule with many active conformations and it can easily be studied in various solvents and as a function of pH [85]. The results of a recent molecular dynamics simulation suggest that the peptide backbone exists as a mixture of folded and unfolded forms at neutral pH, but is always unfolded at low and high pH. The folded form at neutral pH possesses a hydrogen bond between residues 2 and 5. Molecular dynamics simulations of LeuE in DMSO have also recently been performed. This work motivated us to undertake a theoretical investigation of the conformational energy surface of LeuE, initially in the isolated state, similar to our work on NALANMA. First one needed to document the level of theory (level of DFT hybrid XC functional and basis set) that is required to accurately model the interactions responsible for the stabilizing forces and interactions. Since LeuE has aromatic residues at position one and four, there exists the possibility of having the aromatic rings form stable dimer type interactions seen in the benzene dimer and in nucleic acids [86]. This molecule is working as an agonistic ligand to the opiate receptor. In various DFT studies the molecule has been simulated with respect to the structure and functionality. In the study published the energy levels and quantum states of Leu-enkephalin conformations were investigated both theoretically and experimentally both with quantum and classical mechanical methods [87]. Also functional studies were made where the probability distribution of side-chain conformations in Met- and Leu-enkephalin turned out to determine the potency and selectivity to mu- and delta-opiate receptors [88].

CYCLO-(GLY-PRO)3 (CYCLOGP3)

The cyclic hexapeptide, cyclo-(gly-L-pro)$_3$ has been shown to have interesting biochemical prop-

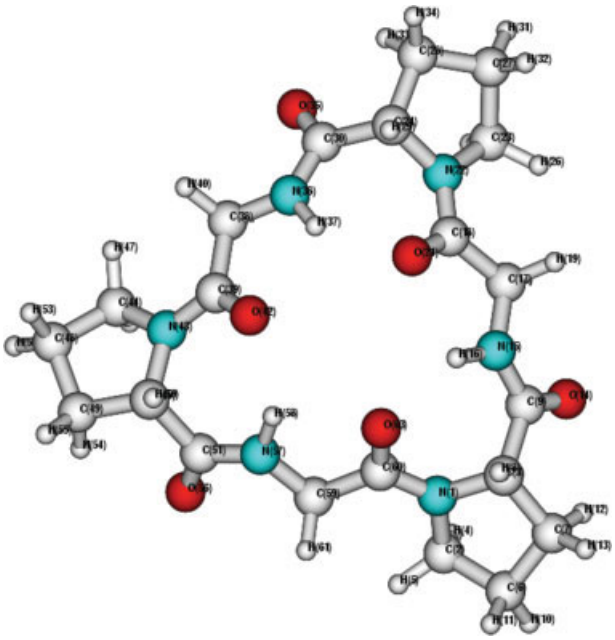

FIGURE 17. Cyclo-(gly-L-pro)$_3$, B3LYP/6-31G*.

erties. Its structure has been shown to be dependent on the presence of metal cations. The lowest energy structure we have found to date is shown in Figure 17. In this work we investigate the inherent stability of the conformers without the metal present, the so-called apo-peptide. Table 7 presents the relative energies of the six conformers we have found to date. The lowest-energy structure is the symmetric structure, which has the three amide carbonyl oxygens all pointing in one direction. As seen in Figure 17, this structure has the three electron rich oxygens in almost optimal positions to bind or coordinate to a positively charged metal ion.

N-ACETYL (L-ALANINE)$_n$ *N'*-METHYL AMIDE (NA(LA)$_n$NMA)

To investigate the relative stabilities of the conformers of the L-alanine residue in proteins and peptides as the length of the chain increases, we have modeled NA(LA)$_n$NMA, $n = 1, 2, 3, 5, 8, \ldots 20$. Here we find that various turn structures are now possible, which are not present in NALANMA. Hence it is important to increase the length of the model systems used in parameterizing molecular mechanics force fields and also in testing new force fields and also semi-empirical-based wave function and DFT-based methods. In our work we have shown that the SCC-DFTB method appears to do well in comparison with the B3LYP/6-31G* results [89, 90]. We have shown that the α (3.6_{13}) helix is not a stable conformer, and that in most cases it will convert to the 3_{10} helix. This

TABLE VII

Backbone torsion angles and relative energies for cyclo-(gly-L-pro)$_3$ B3LYP/6-31G* optimized structures.

	I	II	III	IV	V	VI	VII	VIII
ϕ_1(C80N1C3C9)	−77.4	−66.9	−50.4	−81.6	−81.5	−81.3	−73.7	−81.6
ψ_1(N1C3C9N15)	−22.6	−38.0	131.1	77.1	78.5	79.3	91.6	78.3
ω_1(C3C9N15C17)	−179.8	171.4	−161.5	−173.4	−173.2	−173.0	−156.0	−173.3
ϕ_2	−135.5	−157.4	120.3	161.5	160.4	159.4	155.4	160.6
ψ_2	138.6	145.4	−142.2	−167.5	−167.5	−167.9	−165.5	−167.3
ω_2	−25.8	−27.5	178.3	−159.0	−176.6	−176.6	−177.9	−176.5
ϕ_3	−41.1	−49.9	−69.5	−81.2	−81.7	−81.6	−81.2	−81.8
ψ_3	146.8	−171.9	−9.7	77.2	77.7	78.8	79.4	77.6
ω_3	−174.2	167.6	−161.6	−173.5	−173.4	−173.0	−173.0	−173.5
ϕ_4	−62.6	−62.5	95.4	161.5	161.2	160.1	163.0	161.3
ψ_4	138.9	153.0	153.2	−167.2	−167.2	−167.8	−160.1	−167.0
ω_4	4.9	1.6	2.8	−176.8	−176.8	−176.6	−175.5	−176.6
ϕ_5	−100.4	−59.9	−74.9	−81.4	−81.6	−81.5	−80.0	−81.7
ψ_5	37.3	−45.0	−18.6	77.2	77.9	78.5	−14.0	77.3
ω_5	−168.2	171.9	171.5	173.5	−173.4	−173.1	171.2	−173.4
ϕ_6	−70.1	171.8	−99.0	161.9	160.6	159.8	−101.4	161.2
ψ_6	−26.9	171.6	179.1	−167.2	−167.4	−167.8	−166.7	−167.3
ω_6	20.4	101.1	179.5	−177.1	−176.7	−176.6	−177.0	−176.5
E (kcal/mole)	12.39	12.65	7.75	0.02	0.00	0.03	6.79	0.01

is due to the increase in number of hydrogen bonds in the 3_{10} form. Both helices have been shown to be stabilized by either the aqueous or protein environment, but are not inherently stable structures, as they are in many of the used force fields. Hence one must be careful in utilizing molecular mechanics force fields for evaluating both the number of stable conformers and also the stabilizing interactions. Just because a structure is stable in a crystal, protein, or aqueous environment does not mean it is an inherently stable structure. Hence one must use all the information one has in parameterizing molecular mechanics and semi-empirical wave function and DFT-based methods. The balance of the various contributions in molecular force fields has been shown to be very important, especially the treatment of electrostatic charges, torsional profile, and van der Waals interactions [91]. In many cases, the parameterizations that represent these interactions are not all independent, and one must be very careful in treating parameter correlations and dependencies in the parameterization process.

In addition, with respect to the stabilizing interactions for various structures, the van der Waals interactions are responsible for the attractive forces between nonpolar species. These forces (interactions) are not treated adequately by either RHF or most DFT methods; therefore empirical and semi-empirical correction methods have been suggested [92–95]. The calculation of these weak dispersion forces has been the topic of much research over the past 30 years, both wave functional-based approaches and more recently DFT-based approaches, where the exchange correlation kernel has been the focus of time-dependent DFT studies [96–100]. The question of which data is appropriate to use to parameterize molecular mechanics force fields and then the question of transferability is also a highly researched topic [101–103].

As a further test of small drug compounds and ligands, we have also investigated the ability of DFT and the semi-empirical DFT-based method to model the VA and VCD spectra of small three-membered rings with oxygen and sulfur present and their methyl substituted derivatives.

ETHYLENE OXIDE (OXIRANE) AND PROPYLENE OXIDE (METHYL OXIRANE)

The experimental VA (IR) and Raman intensities of oxirane have been measured [104]. We have optimized the structure of oxirane at the B3LYP level and simulated the VA and VCD spectra [27, 105]. Oxirane is an interesting molecule as it is used as a monomer in the polymer industry. Hence its electronic properties are very important to understand with respect to the mechanism to optimize the conditions for the polymerization process. The lifetimes of the states, intermediates and effects of electromagnetic radiation (EMR), in the microwave, IR, and even at UV-visible wavelengths all are important. Many reactions have been shown to be very sensitive to perturbing effects, for example, catalytic surfaces, input of energy as either heat or EMR. The photolyase protein with its cofactor flavin present catalyzes the repair of damaged DNA in the presence of EMR of a specific wavelength. Hence it is very important to not only understand the density of states of the vibrational, rotational, rovibrational, electronic, and rovibronic levels, but also to understand the processes which can bring the molecule from one state to the other on the same Born–Oppenheimer potential energy surface (BO PES) and also to other excited BO PES and even when the BO model breaks down (conical intersections). The picture gets even more complicated when the system gets more complex and the system one wishes to study is in a medium in which the boundary needs to be defined, but it is not actually clear how to make the partition. Here one has various ways to make the partition: simple solute and solvent. But in some cases the structure of the solute is actually strongly perturbed by the solvent or, in the case of the photolyase, flavin, DNA, water complex, a neighboring cofactor that can absorb a photon of EMR. Here the electron density of the excited state of one of the species in the solution can induce changes in the neighboring system, which results in a charge transfer process, induced dipoles, induced magnetic moment, etc. Hence it is very important that this physics be part of the model. Clearly molecular mechanics models are not adequate to treat such processes. But how many of the molecules in the neighboring environment of the solvent, ligand, and cofactor need to be treated by a full quantum mechanical treatment? Here one would like to use vibrational spectroscopy to monitor these processes, but then one must also know how to include the important water molecules, ligands, counterions, and important residues in the geometry optimization and subsequent Hessian and tensor calculations to be able to simulate not only the frequencies, but also VA, VCD, Raman, and ROA intensities, with the VCD and ROA signals present only for

optically active species or, more importantly, optically active environments. An inherently nonchiral molecule or chromophore can have induced chirality due to its chiral environment.

ETHYLENE SULFIDE (THIIRANE) AND PROPYLENE SULFIDE (METHYL THIIRANE)

The experimental VA and Raman intensities of thiirane have been measured [104]. Thiirane differs from oxirane by the replacement of the ether oxygen in the three-membered ring by sulfur. We have optimized the structure of both thiirane and methyl thiirane at the SCC-DFTB level of theory [106]. Since the APT and AAT are not yet implemented at the SCC-DFTB level of theory, we have calculated the finite difference Hessian at the SCC-DFTB level of theory to obtain the vibrational frequencies and eigenvectors and have combined them with B3LYP/6-311++G(2*d*,2*p*) level APT and EDEDPD, which allow us to simulate the VA and Raman spectra. As seen by comparing the simulations at the mixed SCC-DFTB/B3LYP level with the simulations at the B3LYP level in Figure 18, the SCC-DFTB level of theory underestimates the frequencies; that is, there is a systematic shift to lower frequencies in the VA and Raman spectra. Note that relative intensities are very well produced, telling us that the nature of the modes is probably well reproduced, but which needs to be further documented. The VCD intensities have been shown to be more sensitive to the nature of the eigenvectors than the eigenvalues, or the VA and Raman intensities. In addition to the APT, we need the AAT. We have simulated the VCD spectra at both the DFT and SCC-DFTB levels of theory for *trans*-2,3-dideuteriothiirane, the optically active isotopomer of thiirane. In Figure 19 we can see the features of the VCD spectra are well reproduced by the SCC-DFTB Hessian, when combined with accurate high level APT and AAT.

The SCC-DFTB parameterization for sulfur has also been used to model cysteine and a series of other small sulfur-containing molecules [107]. The extension of the SCC-DFTB level of theory to second-row elements like sulfur has required an extension of the minimal basis set usually employed, that is, for first row elements. By inclusion of *d*-orbitals, the geometries and frequencies of not only hypervalent compounds improved significantly over the parameterization without the inclusion of the *d*-orbital [107]. Sulfur, below oxygen in the periodic table, also has a lone pair in many of its molecular structures, and hence also has the lone pairs that can be donated (shared) in hydrogen bonded structures. The hydrogen bond strength of the sulfur-hydrogen bond is not as strong as the oxygen-hydrogen bond, but it is stronger than the weak dispersion or van der Waals bonds, which are ever present in biomolecular systems.

AFRAMODIAL

Aframodial is an interesting molecule that has diverse medicinal properties. It has been shown to have anti-cancer, anti-malarial, and anti-fungal effects, although in the later stages of clinical testing it was found to lack an acceptable side-effect/benefit profile. For drugs that interact with protein targets, in either their efficacy, clearance, or toxicity pathways, model development is critical in the pursuit of redesign of active molecules with improved prospects. Information must be rapidly adduced to either suggest improved molecules within a class or abandonment of the given class owing to the generality of the limitations identified for a given compound. (A recent very public example is represented by the issues identified with Vioxx, and how those issues related to the whole class of COX-2 inhibitors.) The more information we can accumulate on the structure, properties, and function of these molecules and mode of action, the more rapidly improved drugs can be designed and made available, and the better our understanding will be of the molecular origin and generality of both desired and undesired effects. A classic example of this is provided by some very interesting research that has been performed on the underlying causes of the birth defects caused by women who used thalidomide during pregnancy and on the mechanism of gene damage [108, 109]. This is an example of how an understanding of the complete process is necessary to understand the adverse side effects and their molecular origin. Thalidomide, with improved understanding, is again emerging as a successful anti-cancer agent. Not only has the nature of the problem been better understood, but even an understanding the molecular defect responsible for Holt–Oram syndrome has come to light [108].

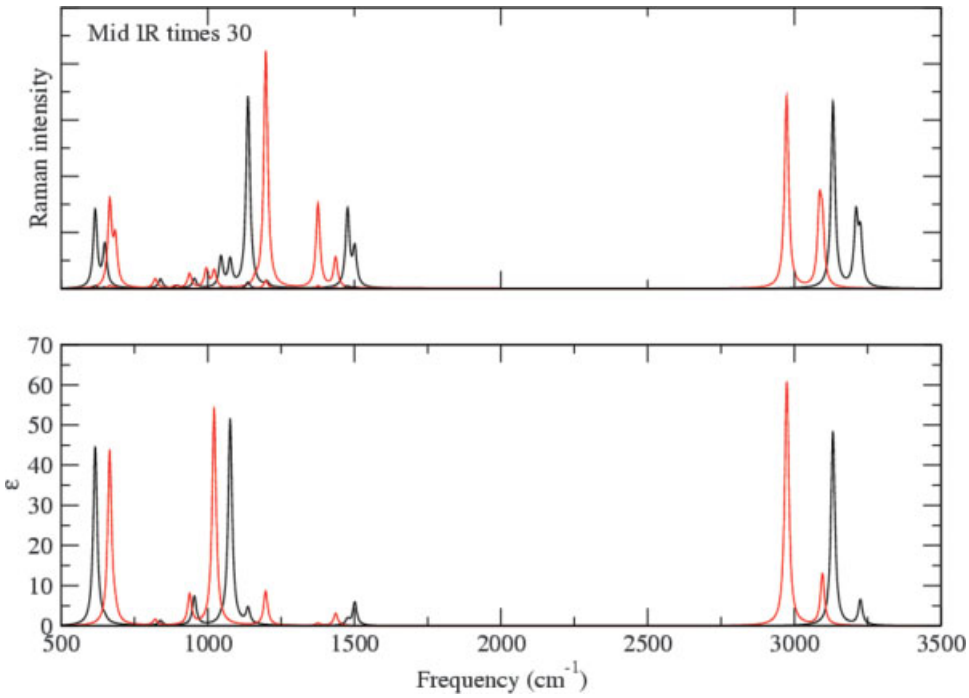

FIGURE 18. VA and Raman spectra for thiirane, B3LYP (black), and SCC-DFTB (red).

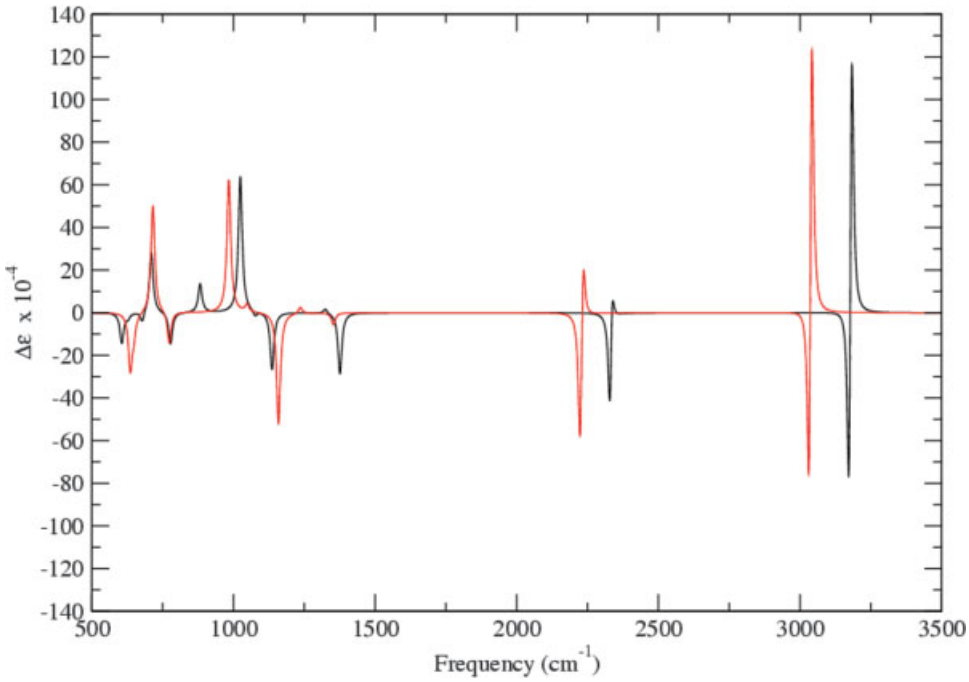

FIGURE 19. VCD spectra for *trans*-2,3-dideuteriothiirane, B3LYP (black), and SCC-DFTB (red).

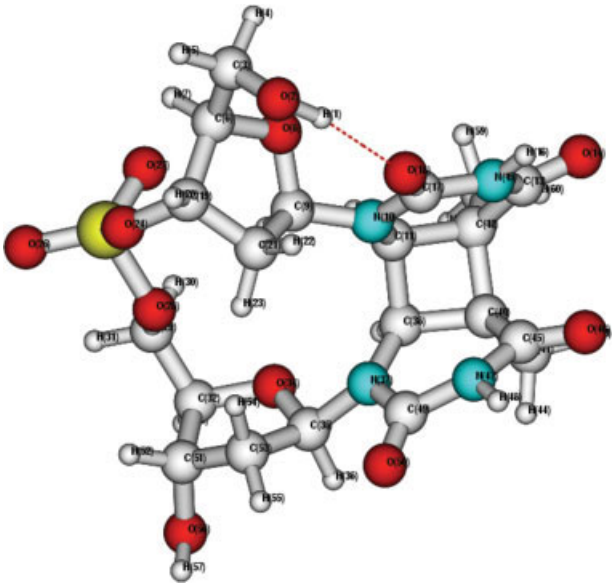

FIGURE 20. Cyclobutane ring damaged dithymine, *cis*-syn-cyclobutane pyrimidine dimer (CPD).

DITHYMINE DUPLEXES (DIMERS) AND TWO STACKED THYMINE BASES

The result of UV damage to DNA can result in the formation of a thymine duplex, that is, the formation of chemical bonds between the adjacent thymine bases. There are three possible dithymine duplexes. In Figures 20 and 21 one sees the CPD structure and its corresponding VA and VCD spectra. In Figures 22 and 23 one sees the OXD structure and its corresponding VA and VCD spectra, and finally in Figures 24 and 25 one sees the 64PPD structure and its corresponding VA and VCD spectra. Note that in Figure 20 a cyclobutane ring (CPD) connects the two thymines, in Figure 22 an oxetane ring (OXD) connects the two thymines, and in the Figure 24 only a one single bond connects the two thymine rings, the so-called 6,4-photoproduct damaged dithymine (64PPD). To distinguish the form of the resulting damaged DNA, we need to have a method to distinguish the three structures.

Vibrational spectroscopy is an obvious method, especially if we can identify regions of the IR spectra where characteristic modes will occur in one form, but not in the other two. Hence we have decided to optimize the structure of the three damaged forms of DNA, our first model being the dinucleotide with the phosphate backbone as well as the ribose rings. This is potentially a much better model than only having the two thymine rings, as we are taking into account the backbone and ribose rings [110]. An even better model would be to have two bases on either side of the damaged (mutated) residues and the aqueous environment with the cat-

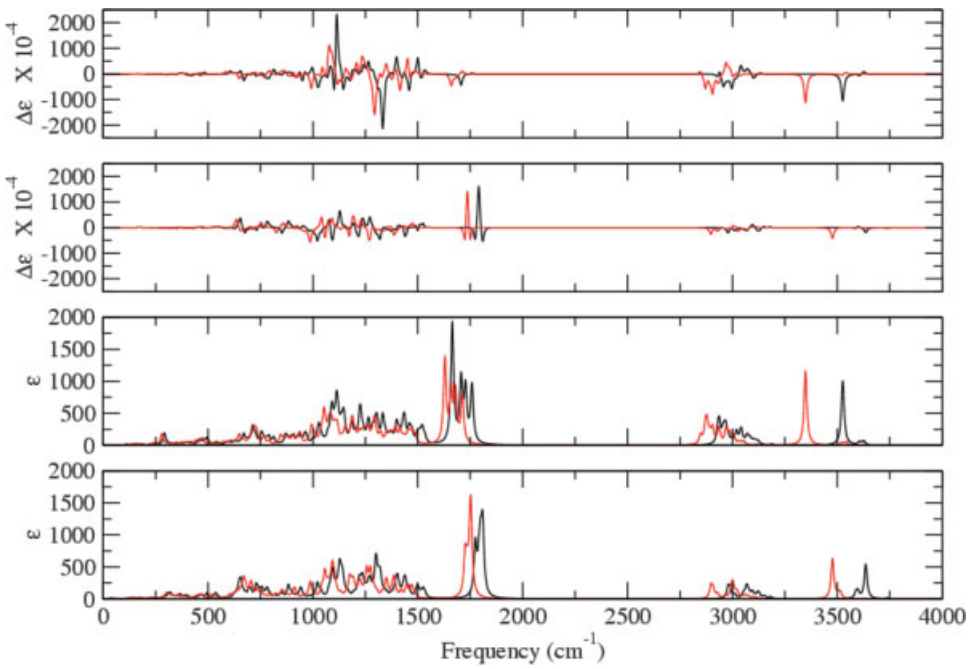

FIGURE 21. VA and VCD spectra for CPD and reduced CPD, B3LYP/6-31G* (black), and PW91/6-31G* (red).

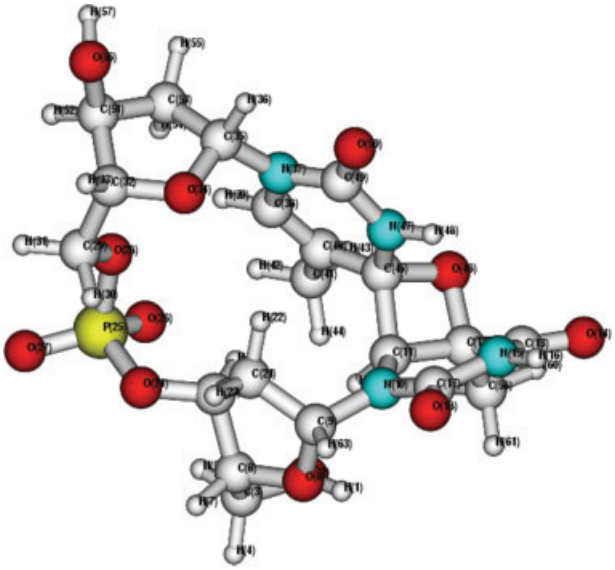

FIGURE 22. Oxetane ring damaged dithymine (OXD).

ions present. We will look at how much of the physics we can deduce from this simpler model.

Figure 21 presents the VA and VCD spectra for the dithymine CPD and reduced CPD structures at the B3LYP/6-31G* (black) and PW91/6-31G* (red) levels of theory. We can see the differences between the two levels of theory for calculating both the VA and VCD spectra of the four species. Noticeably, the carbonyl stretch, CH stretch, and NH stretch frequencies are calculated to be lower in frequency (energy) at the PW91 level of theory for all species. In spite of these differences, the characteristic features of the normal and reduced species are unique enough to characterize the two different species, especially in the VCD spectra. Also note the predicted VCD intensities are enhanced for the radical species, which should make it easier to detect, even if it is not the dominant species, due to this enhancement. Hence the VA and VCD spectra are experimental techniques which should be used to probe the mechanism of the photoreactivation repair of CPD damaged DNA.

Figure 23 presents the VA and VCD spectra for the OXD and reduced OXD structures at the B3LYP/6-31G* (black) and PW91/6-31G* (red) levels of theory.

As one can see in Figures 21 and 23, the NH and CH stretch region for the cyclobutane structure, *cis*-syn-cyclobutane pyrimidine dimer (CPD), there is a splitting of 10 cm^{-1} between the two NH stretch modes of the two thymines, while in the oxetane structure (OXD) there is a splitting of 30 cm^{-1} between the two NH stretch modes of the two thymines. Hence the splitting of these modes can be used as a characteristic feature to distinguish between the two mutated forms. In addition, there is the

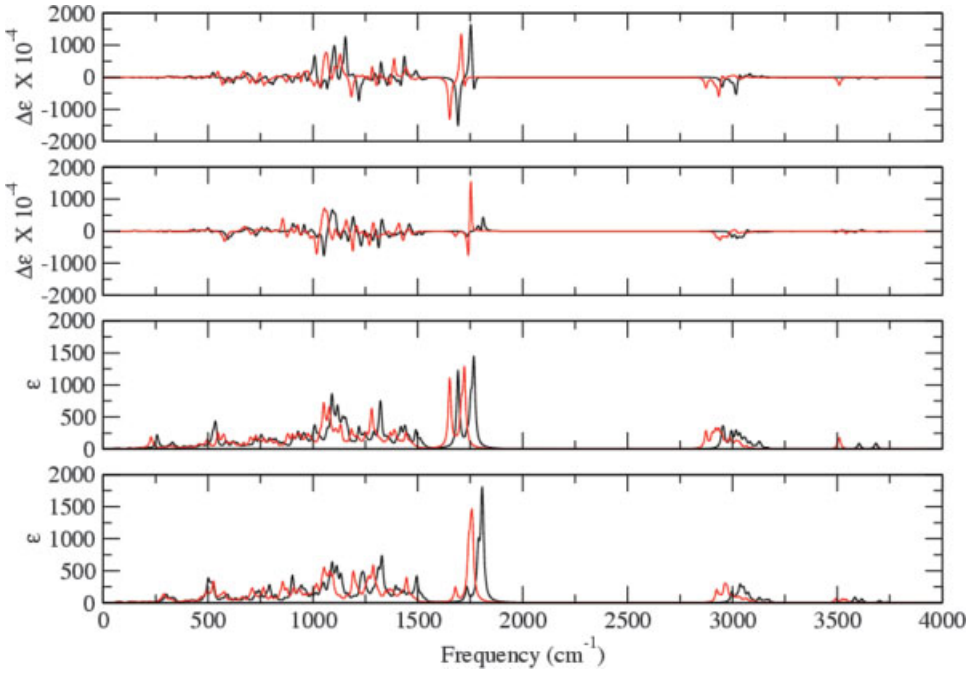

FIGURE 23. VA and VCD spectra for OXD and reduced OXD, B3LYP/6-31G* (black), and PW91/6-31G* (red).

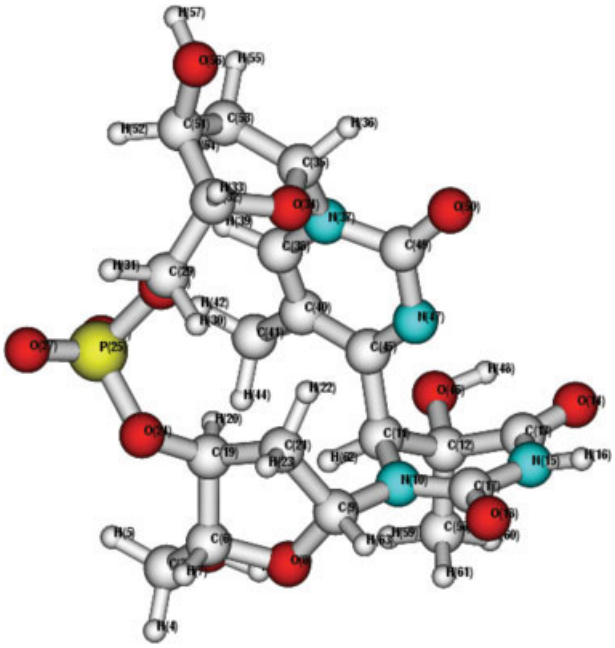

FIGURE 24. 6,4-photoproduct of UV-damaged dithymine (64PPD).

6,4-photoproduct structure (64PPD), shown in Figure 24 with the corresponding VA and VCD spectra for the 64PPD and reduced 64PPD structures in Figure 25. The CPD structure has the lowest energy of the three mutated structures at the B3LYP/6-31G* level of theory, that is, −2241.00343935 Hartrees (0.0 kcal/mole) (PW91/6-31G* − 2240.38766635 Hartrees) (B3LYP/6-31+G* −2241.0916557 Hartrees). The 64PPD structure has a relative energy of (3.8 kcal/mole), (5.0 kcal/mole), (5.0 kcal/mole), followed by the OXD structure having a relative energy of (10.4 kcal/mole), (10.0 kcal/mole), (12.4 kcal/mole), respectively. The B3LYP/6-31G* level of theory relative energy for the undamaged dithymine (UD) structure, is (−34.6 kcal/mole), [−26.7 kcal/mole] and hence the undamaged DNA is lower in energy than the three damaged forms, as it should be. The undamaged dithymine structure is shown in Figure 26, and the corresponding VA and VCD spectra are presented in Figure 27. The minimum energy required to damage the DNA is calculated at the two levels of theory to be only 34.6 kcal/mole and 26.7 kcal/mole, respectively. The repair processes should be at least this amount, however the damaging process appears to require many kcal/mole more than this to surmount the barrier (the energy of the UV radiation necessary to cause the damage). The barrier to form the three damaged forms is very high. Note that the two repair catalysts (the repair proteins CPD photolyase or 6,4 photolyase) both require input of vis-

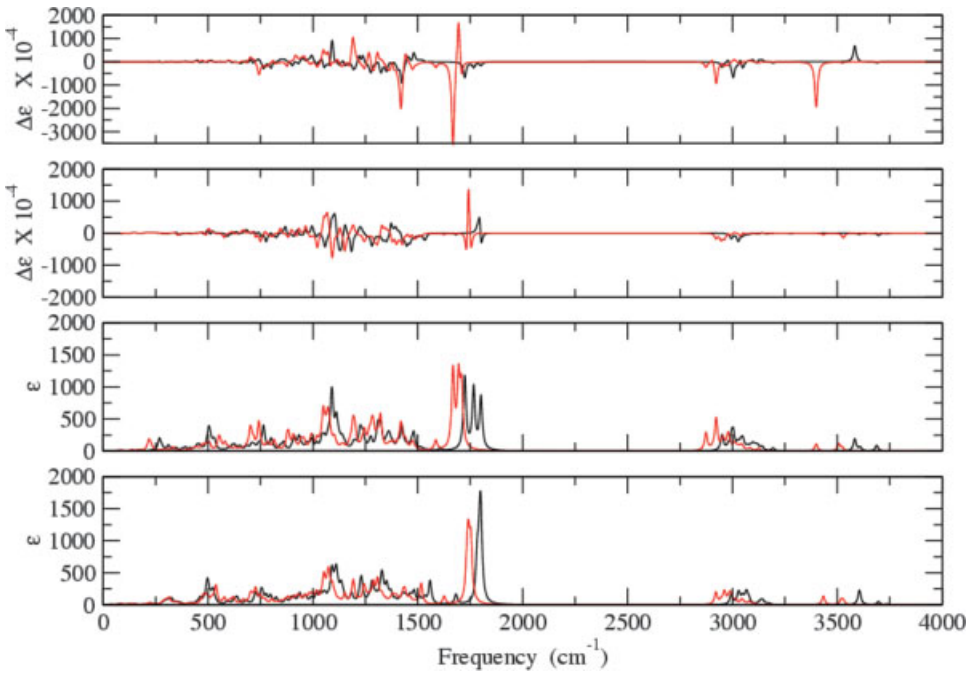

FIGURE 25. VA and VCD spectra for 6,4 photoproduct of UV-damaged dithymine (64PPD) and reduced 64PPD at B3LYP/6-31G* (black) and PW91/6-31G* (red).

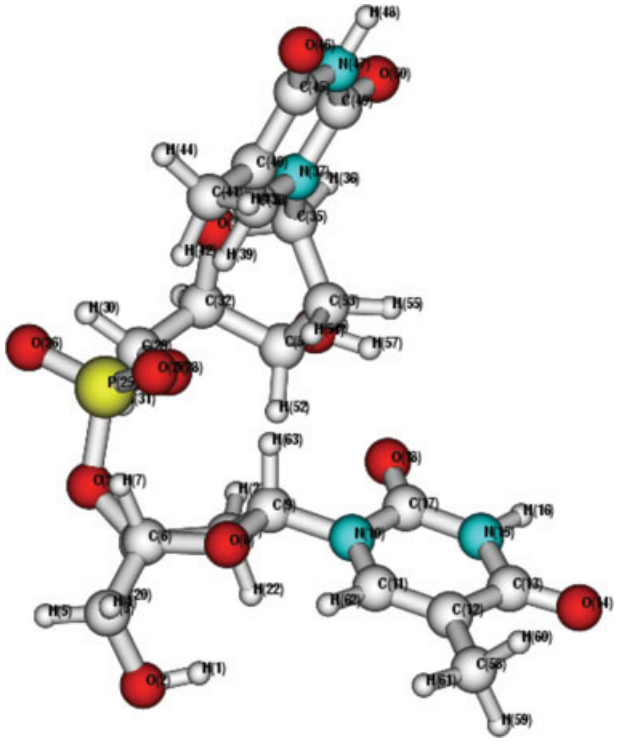

FIGURE 26. Undamaged dithymine (UD).

ible radiation to be active, hence the term photoreactivation. But the photon of energy is thought to excite the flavin, which is subsequently photoreduced to the anionic state, which then transfers an electron to the damaged structure. This is one of proposed mechanisms of the repair process. We have added an electron to the three damaged structures to see whether they are stable or decompose to the undamaged species, as has been suggested. The resulting anionic radicals are 39.3, 48.4, and 53.8 kcal/mole higher in energy than the neutral CPD structure. The CPD structure decomposes into a structure in which the C12—C40 bond in the cyclobutane ring is broken. Hence the connectivity at atoms C12 and C40 in the two thymine rings is now correct, but what remains for the CPD photolyase protein to assist is the breaking of the C11—C38 bond, which still connects the two thymine rings.

Note that the reduced UD species has largely enhanced VA and VCD intensities. In addition, there are noticeable differences in the spectra, so that the VA and VCD spectra can be used to identify and monitor the reduction process. Note also that the base stacking of the thymines in not stable at this level of theory. This is due to inadequacies in the LDA and GGA treatment of the dispersion forces which occurs with many of the current exchange-correlation functionals.

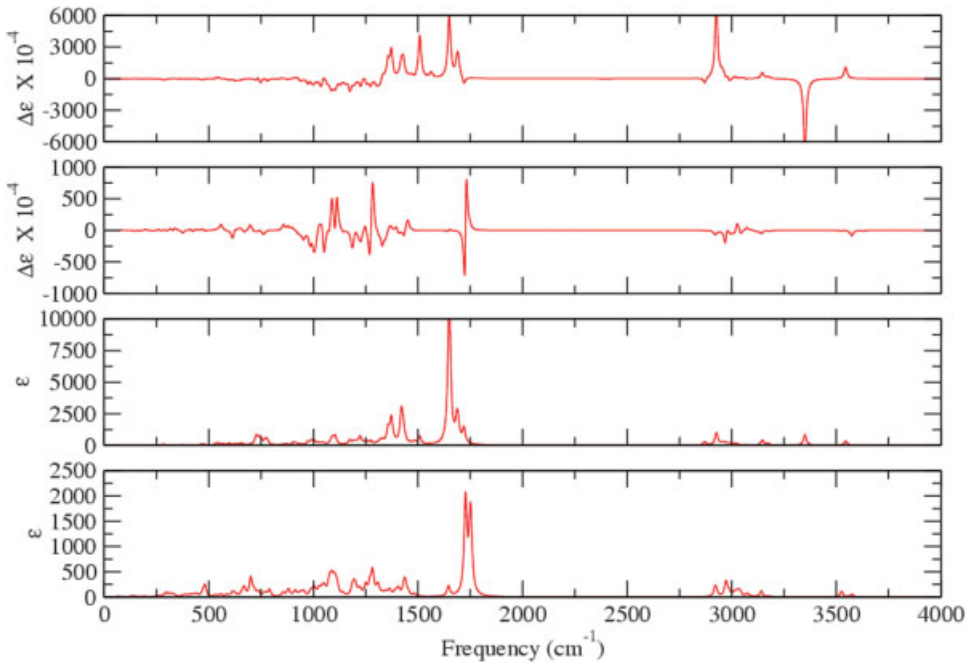

FIGURE 27. VA and VCD spectra for UD, PW91/6-31G*.

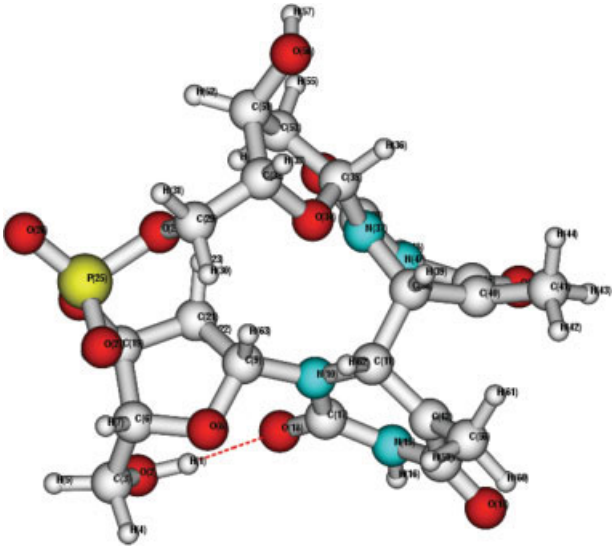

FIGURE 28. Reduced CPD state.

This problem has been solved by the recent ab initio exchange and correlation functionals developed by Bartlett and colleagues [12].

The OXD structure decomposes into a structure in which the C12O46 bond in the oxetane ring is broken. Hence the C45O46 connectivity is now correct for the thymine base, but what is still necessary is for the C11—C45 bond connecting the two thymine rings to be broken.

This is what the 6,4-photolyase protein needs to assist in. In the case of the reduced radical of the 64PPD structure, no bonds are broken, and the bonds between C45 and O46 and H48 and N47 need to be formed, the bonds between O46 and H48 and C11 and C45 need to be broken. Clearly, the 64PPD requires more assistance to repair itself than the CPD and OXD structures. The 64PPD structure is thought to be the resulting structure due to UV damage, as the oxetane structure has been claimed to be unstable with respect to degradation to it. Our oxetane structure is stable before reduction to the reduced radical, which then decomposes to another structure, not the 64PPD. Hence there may be alternative pathways and degradation products for the oxetane ring dithymine structure. Clearly, this needs to be further investigated, both theoretically and experimentally. The CPD and OXD reduced states are presented in Figures 28 and 29, respectively.

We have seen a case in which the protein only needs to assist and only when it is an inherent necessity for the decomposition of the reduced radical species to the free bases. The energy of the optimized reduced structures are much higher than the unreduced species. Hence the reduction does provide excess energy to these structures which effectively reduces the barrier for the bond breaking, and in two cases actually results in bond breaking. This is an alternative to the excited state mechanism. Note that a spectator molecule, probably the flavin, can accept the excess electron from the anionic radical species (forms) and return the two bases to their undamaged forms (if the radical anion decomposes, either one of the bases will be in a radical state or the unpaired electron will be localized elsewhere in the molecule). The thought is that one of the bases will carry the free electron. Hence we calculated the structures for the anionic radical structures of the undamaged bases to see where the free electron localizes. Clearly, the complete process is very involved, and in this work we only present our preliminary studies here at the B3LYP/6-31G* level. The mechanism of the second bond-breaking step for both the CPD and the OXD structures needs to be further investigated. In addition, the repair of the 64PPD structure requires much further work. One possibility is that the 64PPD structure is first converted back to the oxetane ring structure and then the alternative degradation pathway is then followed to form the two free bases. In the event that this is not possible, then the damaged section can be excised, which has been shown to

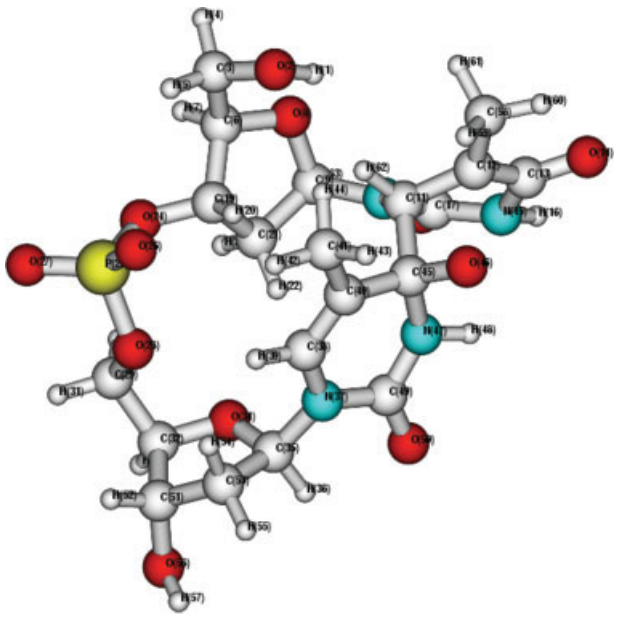

FIGURE 29. Reduced OXD state.

TABLE VIII

Singlet and triplet excitation energies and oscillator strengths, *f*.

	CPD	*f*	OXD	*f*	64P	*f*	Pure	*f*
B3LYP	4.0455 eV	0.0011	3.5162 eV	0.0	3.3039 eV	0.0007	3.8458 eV	0.0025
Singlet A	306.47 nm		352.61 nm		344.03 nm		322.39 nm	
Singlet A	41252 eV	0.0450	3.8748 eV	0.0	3.9036 eV	0.0241	4.3019 eV	0.0039
	300.55 nm		319.98 nm		317.62 nm		288.21 nm	
Singlet A	4.3380 eV	0.006	3.8994 eV	0.0	3.9391 eV	0.0509	4.4112 eV	0.0002
	285.81 nm		317.96 nm		314.75 nm		281.06 nm	
B3LYP	3.5342 eV	0.0					3.3046 eV	0.0
Triplet A	350.81 nm						375.19 nm	
Triplet A	4.0347 eV	0.0					3.3477 eV	0.0
	307.20 nm						370.36 nm	
Triplet A	4.0881 eV	0.0					3.8458 eV	0.0
	303.28 nm						322.39 nm	
PW91	2.4017 eV	0.0005	1.8745 eV	0.0	2.2977 eV	0.0008	2.3694 eV	0.0018
Singlet A	516.23 nm		661.41 nm		539.14 nm		523.88 nm	
Singlet A	2.6382 eV	0.0000	2.2428 eV	0.0	2.5470 eV	0.0014	2.8115 eV	0.0021
	469.95 nm		552.81 nm		486.69 nm		440.98 nm	
Singlet A	2.9874 eV	0.0043	2.2578 eV	0.0	2.7776 eV	0.0004	3.0430 eV	0.0
	415.03 nm		549.14 nm		446.37 nm		407.44 nm	
PW91	2.4007 eV	0.0	1.8744 eV	0.0	2.2929 eV	0.0	2.3647 eV	0.0
Triplet A	516.46 nm		661.45 nm		540.73 nm		524.30 nm	
Triplet A	2.6376 eV	0.0	2.2426 eV	0.0	2.5409 eV	0.0	2.8072 eV	0.0
	470.07 nm		552.87 nm		487.96 nm		441.66 nm	
Triplet A	2.8973 eV	0.0	2.2574 eV	0.0	2.6673 eV	0.0	3.0429 eV	0.0
	427.93 nm		549.23 nm		464.84 nm		407.45 nm	

occur. Hence the 64PPD damage is harder to repair, hence the necessity for excision.

In addition, we have presented the VA and VCD spectra for the four species that can in principle be used to monitor the presence of these species. In addition, we have also optimized the structures of the three radical anions, which are thought to be possible intermediate states and tested for their stability. These radical anions have been proposed as intermediates, and it would be nice to try to isolate these species at low temperature, where they may be stable and measure their VA and VCD spectra and, in addition, the electron spin resonance (ESR) spectra.

To investigate the possibility of excited state mechanism, we have calculated the three lowest singlet and triplet state excitation energies with the corresponding oscillator strengths, using time-dependent density functional theory (TDDFT) [111]. These are reported in Table VIII.

In addition to calculating the oscillator strengths for the three lowest singlet and triplet transition, we have also calculated the rotational strengths in both the velocity and length form, the electronic circular dichroism. Table IX presents our results.

Note that the signs of the electronic circular dichroism (ECD) transitions are even different when calculated for the three lowest-energy transitions for the B3LYP and PW91 XC functionals. Previously it has been shown that the BP86 [112, 113] exchange correlation (XC) functional combined with TDDFT give rotational strengths which agree qualitively with the experimental values [114]. Since the PW91 XC functional is of the pure GGA type, it may not treat excited states, as well as the B3LYP hybrid XC functional, which includes some Hartree–Fock exchange. Hence the B3LYP are more reliable, although the calculation of the excitation energies, oscillator strengths for the electronic spectrum (ES) and rotational strengths for the ECD are more costly. The combination of VA, VCD, ES, and ECD in addition to the ESR may be a way to get a better picture of the mechanism for CPD- and 6,4-photolyase repair of CPD, OXD, and 64PPD damaged DNA. But to be able to fully use the exper-

TABLE IX

Electronic circular dichroism (rotational strengths, velocity, and length forms) for three lowest-energy singlet and triplet transition.

	CPD	R_i	OXD	R_i	64P	R_i	Pure	R_i
B3LYP	4.0455 eV	0.3520	3.5162 eV	−0.0472	3.6039 eV	−0.2370	3.8458 eV	0.8764
Singlet A	306.47 nm	0.2467	352.61 nm	0.0031	344.03 nm	−0.4576	322.39 nm	0.8456
Singlet A	41252 eV	−19.3611	3.8748 eV	−0.3314	3.9036 eV	−7.1948	4.3019 eV	0.6424
	300.55 nm	−19.2416	319.98 nm	0.0472	317.62 nm	−8.0278	288.21 nm	0.6578
Singlet A	4.3380 eV	−0.4352	3.8994 eV	−0.0249	3.9391 eV	−4.8288	4.4112 eV	0.0516
	285.81 nm	−0.4498	317.96 nm	−0.0291	314.75 nm	−5.5271	281.06 nm	0.0518
PW91	2.4017 eV	0.1892	1.8745 eV	0.0031	2.2977 eV	−0.4338	2.3694 eV	0.4190
Singlet A	516.23 nm	0.1853	661.41 nm	0.0114	539.14 nm	−0.5535	523.28 nm	0.4388
Singlet A	2.6382 eV	0.4002	2.2428 eV	0.0011	2.5470 eV	−0.9531	2.8115 eV	−0.2934
	469.95 nm	0.3269	552.81 nm	−0.0129	486.69 nm	−0.6841	440.98 nm	−0.2365
Singlet A	2.9874 eV	2.8087	2.2578 eV	0.0339	2.7776 eV	0.1279	3.0430 eV	−0.0078
	415.03 nm	3.0070	549.14 nm	0.0154	446.37 nm	0.1494	407.44 nm	−0.0082
PW91	2.4007 eV	0.0	1.8744 eV	0.0	2.2929 eV	0.0	2.3647 eV	0.0
Triplet A	516.46 nm	0.0	661.45 nm	0.0	540.73 nm	0.0	524.30 nm	0.0
Triplet A	2.6376 eV	0.0	2.2426 eV	0.0	2.5409 eV	0.0	2.8072 eV	0.0
	470.07 nm	0.0	552.87 nm	0.0	487.96 nm	0.0	441.66 nm	0.0
Triplet A	2.8973 eV	0.0	2.2574 eV	0.0	2.6673 eV	0.0	3.0429 eV	0.0
	427.93 nm	0.0	549.23 nm	0.0	464.84 nm	0.0	407.45 nm	0.0

imental spectra, a rigorous level theory that can adequately treat the isolated case, without the protein and cofactors present is necessary. This work makes progress in this respect. In a future work, we shall also present the Raman and ROA spectra for all species where we here treat the VA and VCD spectra. To also treat the Raman and ROA spectra of all species is beyond the scope of this work. Note that preliminary results of this work have been reported previously [115].

BAM (2,6-DICHLOROBENZAMIDE)

Another application example being developed for small molecule–protein interactions involves analytical chemistry. BAM (2,6-dichlorobenzamide) is a breakdown product of the herbicide 2,6-dichlorobenzonitril, also called dichlobenil (DCB). Hence the detection of both the parent species DCB and BAM in the groundwater at low concentrations is very important. BAM has been found in the groundwater in Denmark, the United States, Sweden, and Germany. It is the source(s) of these chemicals and the processes and mechanism for their accumulation in the environment, which require increasingly sensitive and specific analytical methodologies. Hence it is very important to develop very sensitive probes for the parent molecules and degradation products. We have collected the VA and Raman scattering spectra for DCB and BAM. The goal of the collaborative project between Helsinki University of Technology (HUT), QuP, Roskilde University, GEUS, and the Statens Serum Institute in Denmark is to see whether immunochemical and spectroscopic techniques could be combined to develop a probe for this and other herbicides, pesticides, and their degradation products. The preliminary results have been published in the master's thesis of Abdoul Rahim at the Quantum Protein (QuP) Center, Lyngby, Denmark [116]. In that work, no surface-enhanced Raman spectra was observed for the BAM bound to the F(ab') fragment in the colloid solution. There could be a variety of reasons for the lack of enhancement. One possibility is that the bound BAM does not interact sufficiently with the Au particles to gain enhancement. Previously it has been observed that not all modes in the molecule are surface enhanced. Hence it is important to study and investigate, both theoretically and experimentally, the nature of the binding of BAM to the F(ab') fragment and also to the Au colloidal particles. Our initial modeling study is to simulate the VA and Raman spectra of DCB and BAM in the isolated state, followed by the use of a

continuum model of the solvent, followed by Au atoms bound to DCB and BAM, and finally the Au/DCB and Au/BAM complexes with the PCM to see which modes are enhanced by the Au atoms in their proximity. The addition of more than one Au atom and in addition the size of the Au cluster and the frequency of the incoming radiation are also important parameters (variables) that must and should be investigated. Figure 30 presents the structure of DCB.

Table X presents the geometrical parameters for DCB as determined with the B3LYP hybrid XC functional and the cc-pVDZ and aug-cc-pVDZ basis sets. Table XI presents the vibrational frequencies, VA and Raman intensities. These can be useful for assigning modes in VA and Raman spectra of DCB and in the SERS studies. The modes that involve those atoms that are strongly influenced by the gold colloid particles are those predicted to be enhanced.

TABLE X

Structural parameters for 2,6-dichlorobenzonitrile (dichlobenil) (DCB).

	cc-vPDZ	aug-cc-vPDZ
Bond lengths (Å)		
R(C1—H7)	1.0952	1.0938
R(C1—C6)	1.3975	1.3978
R(C6—H9)	1.0935	1.0922
R(C6—C5)	1.3923	1.3916
R(C5—C10)	1.7505	1.7518
R(C5—C4)	1.4126	1.4117
R(C4—C12)	1.4312	1.4301
R(C12—N13)	1.1640	1.1631
Valence angles (degrees)		
θ(7-1-6)	119.46	119.42
θ(2-1-6)	121.09	121.16
θ(1-6-5)	119.19	119.100
θ(1-6-9)	121.03	120.98
θ(6-5-4)	121.29	121.35
θ(6-5-10)	119.29	119.20
θ(5-4-12)	121.03	121.02
θ(4-12-13)	180.00	180.00

RELATED WORKS BY OTHER GROUPS

In addition to the work done in R. Nieminen's laboratory in Finland, S. Suhai's, R. C. Wade's and C. Jung's groups in Germany, the QuP in Denmark, there have been many other contributions, such as Werner Hug's laboratory in Professor Hanson's laboratory at the University of Zurich, Tim Keiderling's laboratory in Chicago, IL, Larry Nafie's laboratory in Syracuse, NY, Max Diem's laboratory in NY, NY, P. L. Polavarapu's laboratory in Memphis, TN and finally Philip J. Stephens' laboratory in LA, CA, where a substantial fraction of both the theoretical and instrument development continues to occur, in addition to the development now being pursued by Thermo Electron, Bruker, Bomen, and Biotools. In addition to the experimental developments, of course the theoretical developments have helped the fields of VCD and ROA spectroscopy flourish in recent years. A rigorous ab initio theory that was also implemented and shown to give agreement with experimental VCD and ROA spectra did not occur immediately after the initial VCD and ROA measurements. It took more than 10 years for this to occur, which really limited the interpretation of the VCD and ROA spectra. Initially, the theories were adequate to interpret and explain the VCD and ROA spectra of molecule in nonpolar solvents and where the structures were relatively rigid. Only recently, almost 10 years later has the theory been extended to be able to deal with strongly interacting solvent, conformational equilibrium, pH effects, and ionic strength. The problem was with the paradigm that one could interpret the spectra of molecules in strongly interacting solvents and environments by only studying the molecule in its isolated state.

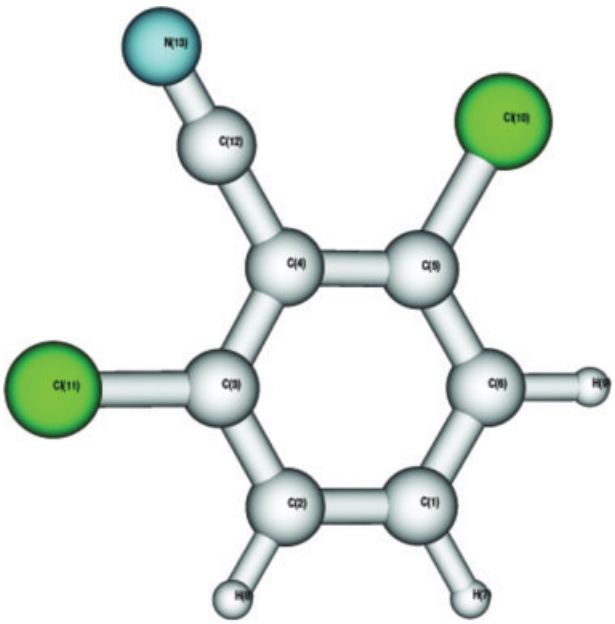

FIGURE 30. Structure of 2,6-dichlorobenzonitrile (dichlobenil) [DCB].

TABLE XI

Vibrational frequencies, VA, and Raman intensities for dichlobenil.

Symmetry	Frequency (cm^{-1})	Ai (km/mole)	Raman	Frequency (cm^{-1})	Ai (km/mole)	Raman
A1	3152	21.44	689.06	3149	25.88	803.18
B2	3150	23.09	232.37	3144	25.99	217.73
A1	3132	0.85	284.90	3130	0.31	233.96
A1	2330	97.21	1037.13	2316	149.92	1549.21
A1	1620	106.70	252.31	1614	116.52	320.42
B2	1594	62.41	33.34	1587	69.97	38.47
A1	1455	19.28	31.11	1454	16.17	29.43
B2	1450	186.85	3.47	1443	205.02	2.97
B2	1328	4.03	44.29	1320	4.06	67.75
A1	1223	0.70	165.24	1220	2.13	223.52
B2	1209	53.43	3.58	1210	59.15	5.46
B2	1159	2.29	6.91	1159	5.16	5.83
A1	1102	26.41	71.28	1099	26.59	55.88
A1	1077	1.86	47.91	1073	0.21	134.01
B1	1013	0.50	0.04	1012	0.63	0.12
A2	926	0.00	1.65	917	0.00	1.15
B1	804	48.29	4.22	798	67.99	1.35
B2	784	147.01	3.23	786	150.10	1.93
A1	777	3.19	12.02	773	2.72	19.82
B1	748	14.96	0.49	745	26.97	3.52
B2	600	0.00	1.02	597	0.07	0.50
B1	575	5.15	5.14	574	6.26	2.95
A1	546	12.40	58.49	545	10.59	90.15
A2	539	0.00	1.67	531	0.00	1.33
B1	422	8.43	5.21	419	22.36	17.46
B2	417	19.89	25.40	416	11.56	2.23
A1	360	1.93	14.81	358	1.95	17.70
B2	356	1.09	0.55	353	2.23	2.33
B1	200	0.64	1.50	208	1.09	0.90
A1	199	0.18	5.19	198	0.08	3.96
A2	196	0.00	8.04	179	0.00	7.50
B2	135	3.87	7.18	143	5.12	7.00
B1	75	0.33	0.62	76	0.37	0.25

The relative energies of the various conformers were thought to be affected, but not the number and types of stable states. Pioneering work on LA, NALANMA, and LALA was performed in the group of S. Suhai, when one of us was a visiting scientist at the German Cancer Research Center in Heidelberg, Germany. This work has been one of the cornerstones for further developments. Other cornerstones have recently been put down by the recent collaborative multidimensional IR experiments performed in the laboratories of P. Hamm in Switzerland [117], S. Woutersen in Berlin, Germany [118–120], and G. Stock in Frankfurt, Germany [121]. By performing two photon experiments one is able to determine which vibrational modes are coupled. Here, similar to modes which are due to anharmonic effects, Fermi resonance for example, a theoretical treatment beyond the harmonic level is necessary to interpret the results of these experimental studies. But once again the modeling studies have lagged a bit behind the experiments. For these new experiments to be fully used requires a rigorous ab initio model that can treat all of the effects. A recent review gives the present state of

both the experimental and theoretical aspects of nonlinear 2D IR (vibrational) spectroscopy [122]. Here current DFT theory is a good starting point, but needs to be further developed to be able to adequately and accurately treat the weak but collectively important dispersion force responsible for ligand binding and also protein, peptide, carbohydrate, and nucleic acid folding and complex formation. In addition to being able to treat the forces accurately, methods must be developed to determine the population of states, the nature of the states, and the effects due to the environment, temperature, and perturbing static and dynamic radiation, electric fields and field gradients, and magnetic fields and field gradients. Hence all these things need to be taken into account.

In addition to the work on nonlinear 2D vibrational spectroscopy, the use of the coupled oscillator model has undergone a renaissance [123]. Schweitzer-Stenner has combined the measurement of the amide I modes in the FTIR and polarized Raman with theoretical simulations by Tori and Tasumi [124] to try to determine the relative orientation of the amide groups in peptides [125–129]. Hochstrasser's group at the University of Pennsylvania has used 2D vibrational measurements to investigate the coupling of amide I modes in α-helices [130]. Recently, the femtosecond-coherent anti-Stokes Raman spectroscopy (fs-CARS) has been used to probe the ground state of porphyrin molecules [131]. In addition, Keiderling's group has worked on using ab initio and DFT-based methods to simulate the VA, VCD, Raman, and ROA spectra of small peptides to try to interpret the spectra of secondary structural elements, similar to our work, but without the treatment of explicit water molecules [132].

It is important that both experimentalists who are using ab initio, DFT, and semi-empirical based DFT codes and packages are adequately aware of the complexity of the calculations, the pit falls, and what needs to be further developed, as do the theoretically oriented scientists need to be aware of the new experiments, time-resolved experiments, so that their newly developed XC functionals, semi-empirical-based DFT methods can accurately and efficiently produce the required tensors, tensor derivatives, both static and dynamic such that they can also be calculated at each point of a QMMD simulation so that all the spectra can be calculated by Fourier methods which will give us the frequency response. The multiphoton experiments with UV-vis, IR, microwave, and radiofrequency radiation in addition to applied electric and magnetic fields will provide us with a wealth of spectral data, but for these data to be adequate, uniquely and accurately used, the theoretical methods also need to evolve, as they have done for x-ray and neutron scattering and NMR, which have allowed us to solve for the structures of proteins, peptides, carbohydrates, and nucleic acids. New and better methods of all varieties that can probe the conformational and functional states of ligands and drugs in all the states that are biologically relevant are very important, if we are to understand the mode of action of a given process at the molecular level. Such methods would be indispensable, for example, in developing better drugs and understanding the current set of drug classes in a better way. With the recent development of new techniques in vibrational spectroscopy, as well as the capabilities of the modeling, the combined use of these experiments with rigorous ab initio, DFT, and semi-empirical-based DFT methods, which have been fully tested and documented, may be able to add significantly to this understanding.

References

1. Holzwarth, G.; Hsu, E. C.; Mosher, H. S.; Faulkner, T. R.; Moscowitz, A. J Am Chem Soc 1974, 96, 251.
2. Barron, L. D.; Bougaard, M. P.; Buckingham, A. D. J Am Chem Soc 1973, 95, 603.
3. Jalkanen, K. J.; Suhai, S. Chem Phys 1996, 208, 81.
4. Han, W. G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. J Phys Chem B 1998, 102, 2587.
5. Amos, R. D. Chem Phys Lett 1984, 108, 185.
6. Stephens, P. J. J Phys Chem 1985, 89, 748.
7. Buckingham, A. D.; Fowler, P. W.; Galwas, P. A. Chem Phys 1987, 112, 1.
8. Amos, R. D. Chem Phys Lett 1986, 124, 376.
9. Amos, R. D. Chem Phys Lett 1982, 87, 23.
10. Polavarapu, P. J Phys Chem 1990, 94, 8106.
11. Scott, A. P.; Radom, L. J Phys Chem 1996, 100, 16502.
12. Lotrich, V. F.; Bartlett, R. J.; Grabowski, I. Chem Phys Lett 2005, 405, 43.
13. Amos, R. D.; Handy, N. C.; Jalkanen, K. J.; Stephens, P. J. Chem Phys Lett 1987, 13, 21.
14. Stephens, P. J.; Jalkanen, K. J.; Amos, R. D.; Lazzeretti, P.; Zanasi, R. J Phys Chem 1990, 94, 1811.
15. Bak, K. L.; Jørgensen, P.; Helgaker, T.; Ruud, K.; Jensen, H. J. A. J Chem Phys 1994, 100, 6620.
16. Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. Chem Phys Lett 1996, 252, 211.
17. Diem, M. Introduction to Modern Vibrational Spectroscopy; John Wiley & Sons: New York, 1993; p 194.
18. Jalkanen, K. J.; Bohr, H. G.; Suhai, S. In Proceedings of the

International Symposium on Theoretical and Computational Genome Research; Suhai, S., Ed.; Plenum: New York, 1997; p 255.

19. Tajkhorshid, E.; Jalkanen, K. J.; Suhai, S. J Phys Chem B 1998, 102, 5899.
20. Barron, L. D.; Gargaro, A. R.; Hecht, L.; Polavarapu, P. Spectrochim Acta 1991, 47A, 1001.
21. Yu, G. S.; Freedman, T. B.; Nafie, L. A.; Deng, Z.; Polavarapu, P. L. J Phys Chem 1995, 99, 835.
22. Frimand, K.; Bohr, H.; Jalkanen, K. J.; Suhai, S. Chem Phys 2000, 255, 165.
23. Jalkanen, K. J.; Nieminen, R. M.; Frimand, K.; Bohr, J.; Bohr, H.; Wade, R. C.; Tajkhorshid, E.; Suhai, S. Chem Phys 2001, 265, 125.
24. Diem, M.; Polavarapu, P. L.; Oboodi, M.; Nafie, L. A. J Am Chem Soc 1982, 104, 3329.
25. Diem, M. J Am Chem Soc 1988, 110, 6967.
26. Abdali, S.; Jalkanen, K. J.; Bohr, H.; Suhai, S.; Nieminen, R. M. Chem Phys 2002, 282, 219.
27. Frimand, K.; Jalkanen, K. J. Chem Phys 2002, 279, 161.
28. Bohr, H. G.; Jalkanen, Elstner, M.; Frimand, K.; Suhai, S. Chem Phys 1999, 246, 13.
29. Degtyarenko, I. M.; Degtyarenko, K. M.; Nafie, L. A.; Cao, X.; Nieminen, R. M.; Halonen, L.; Jalkanen, K. J. (in preparation).
30. Knapp-Mohammady, M.; Jalkanen, K. J.; Nardi, F.; Wade, R. C.; Suhai, S. Chem Phys 1999, 240, 63.
31. Jalkanen, K. J.; Nieminen, R. M.; Knapp-Mohammady, M.; Suhai, S. Int J Quantum Chem 2003, 92, 239.
32. Diem, M.; Oboodi, M. R.; Alva, C. Biopolymers 1984, 23, 1917.
33. Ford, S. J.; Wen, Z. Q.; Hecht, L.; Barron, L. D. Biopolymers 1994, 34, 303.
34. Degtyarenko, I. M.; Claussen, A.; Jalkanen, K. J. (in preparation).
35. Hübschle, Ch. B.; Messerschmidt, M.; Luger, P. Cryst Res Tech 2004, 39, 274.
36. Himo, F.; Eriksson, L. A. J Phys Chem 1997, 101, 9811.
37. Sivaraja, M.; Goodin, D. B.; Smith, M.; Hoffman, B. M. Science 1989, 245, 738.
38. Erman, J. E.; Vitello, L. B.; Mauro, J. M.; Kraut, J. Biochemistry 1989, 28, 7992.
39. Scholes, C. P.; Liu, Y.; Fishel, L. A.; Farnum, M. F.; Mauro, J. M.; Kraut, J. Isr J Chem 1989, 29, 85.
40. Houseman, A. L. P.; Doan, P. E.; Goodin, D. B.; Hoffman, B. M. Biochemistry 1993, 32, 4430.
41. Huyett, J. E.; Doan, P. E.; Gurbriel, R.; Houseman, A. L. P.; Sivaraja, M.; Goodin, D. B.; Hoffman, B. M. J Am Chem Soc 1995, 117, 9033, and references therein.
42. Becke, A. D. J Chem Phys 1993, 98, 5648.
43. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. J Phys Chem 1994, 98, 11623.
44. Stephens, P. J.; Devlin, F. J.; Ashvar, C. S.; Chabalowski, C. F.; Frisch, M. J. Faraday Disc 1994, 99, 103.
45. Jensen, G. M.; Goodin, D. B.; Bunte, S. W. J Phys Chem 1996, 100, 954.
46. Fitzgerald, M. M.; Churchill, M. J.; McRee, D. E.; Goodin, D. B. Biochemistry 1994, 33, 3807.
47. Fitzgerald, M. M.; Trester, M. L.; Jensen, G. M.; McRee, D. E.; Goodin, D. B. Protein Sci 1995, 4, 1844.
48. Musah, R. A.; Jensen, G. M.; Rosenfeld, R. J.; McRee, D. E.; Goodin, D. B.; Bunte, S. W. J Am Chem Soc 1997, 119, 9083.
49. Musah, R. A.; Jensen, G. M.; Bunte, S. W.; Rosenfeld, R. J.; Goodin, D. B. J Mol Biol 2002, 315, 845.
50. Patterson, W. R.; Poulos, T. L. Biochemistry 1995, 34, 4331.
51. Patterson, W. R.; Poulos, T. L.; Goodin, D B Biochemistry 1995, 34, 4342.
52. Jensen, G. M.; Bunte, S. W.; Warshel, A.; Goodin, D. B. J Phys Chem B 1998, 102, 8221.
53. Singh, U. C.; Kollman, P. A. J Comp Chem 1984, 5, 159.
54. Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. J Comp Chem 1990, 11, 431.
55. (a) Warshel, A. Computer Modeling of Chemical Reactions in Enzymes and Solutions; Wiley-Interscience: New York, 1991; (b) Lee, F. S.; Zhen-Tao, C.; Bolger, M. B.; Warshel, A. Protein Eng 1992, 5, 215; (c) Churg, A.; Warshel, A. Biochemistry 1986, 25, 1675; Cutler, R. L.; (d) Davies, A. M.; Creighton, S.; Warshel, A.; Moore, G. R.; Smith, M.; Mauk, A. G. Biochemistry 1989, 28, 3188; (e) Langen, R.; Brayer, G. D.; Berghuis, A. M.; McLendon, G.; Sherman, F.; Warshel, A. J Mol Biol 1992, 224, 589; (f) Parson, W. W.; Chu, Z. T.; Warshel, A. Biochim Biophys Acta 1990, 1017, 251; (g) Langen, R.; Jensen, G. M.; Jacob, U.; Stephens, P. J.; Warshel, A. J Biol Chem 1992, 267, 25625; (h) Jensen, G. M.; Warshel, A.; Stephens, P. J. Biochemistry 1994, 33, 10911; (i) Jensen, G. M. Ph.D. dissertation, University of Southern California, 1994; (j) Stephens, P. J.; Jollie, D. R.; Warshel, A. Chem Rev 1996, 96, 2491.
56. (a) Adamo, C.; Barone, V.; Fortunelli, A. J Chem Phys 1995, 102, 384; (b) Qin, Y.; Wheeler, R. A. J Chem Phys 1995, 102, 1689; (c) Laming, G. J.; Handy, N. C.; Amos, R. D. Mol Phys 1993, 80, 1121; (d) Eriksson, L. A.; Malkin, V. G.; Malkina, O. L.; Salahub, D. R. J Chem Phys 1993, 99, 9756; (e) Eriksson, L. A.; Malkina, O. L.; Malkin, V. G.; Salahub, D. R. J Chem Phys 1994, 100, 5066; (f) Barone, V.; Adamo, C. Chem Phys Lett 1994, 224, 432; (g) Austen, M. A.; Eriksson, L. A.; Boyd, R. J. Can J Chem 1994, 72, 695; (h) Gervasio, F. L.; Cardini, G.; Salvi, P. R.; Schettino, V. J Phys Chem A 1998, 102, 2131; (i) Pan, D.; Phillips, D. L. J Phys Chem A 1999, 103, 4737; (j) Pan, D.; Shoute, L. C. T.; Phillips, D. L. J Phys Chem A 1999, 103, 6851; (k) Halls, M. D.; Schlegel, H. B. J Chem Phys 1999, 111, 8819; (l) Pan, D.; Shoute, L. C. T.; Phillips, D. L. J Phys Chem A 2000, 104, 4140.
57. Tu, A. T. Raman Spectroscopy in Biology; John Wiley & Sons: New York, 1982.
58. Bunte, S. W.; Jensen, G. M.; McNesby, K. L.; Goodin, D. B.; Chabalowski, C. F.; Nieminen, R. M.; Suhai, S.; Jalkanen, K. J. Chem Phys 2001, 265, 13.
59. Gurudas, U.; Schelvis, J. P. M. J Am Chem Soc 2004, 126, 12788.
60. Walden, S. E.; Wheeler, R. A. J Am Chem Soc 1997, 119, 3175.
61. Maple, J.; Hwang, M.-J.; Jalkanen, K. J.; Stockfisch, T. P.; Hagler, A. T. J Comp Chem 1998, 19, 430.
62. Gontrani, L.; Mennucci, B.; Tomasi, J. J Mol Struct (Theochem) 2000, 500, 113.

63. Poon, C.-D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. J Am Chem Soc 2000, 122, 5642.
64. Deng, Z.; Polavarapu, P. L.; Ford, S. J.; Hecht, L.; Barron, L. D.; Ewig, C. S.; Jalkanen, K. J. J Phys Chem 1996, 100, 2025.
65. Schuster, P.; Wolschann, P. Monatsh Chem 1999, 130, 947.
66. Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Annu Rev Biophys Biomol Struct 2000, 29, 291.
67. Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R. A. G. D.; Decatur, S. M.; Keiderling, T. A. J Am Chem Soc 2004, 126, 2346.
68. Bohr, H. G.; Frimand, K.; Jalkanen, K. J.; Nieminen; Suhai, S. R. M. Phys Rev E 2001, 64, 021905-1-13.
69. Ehrlich, L.; Reczko, M.; Bohr, H.; Wade, R. C. Protein Eng 1998, 11, 11.
70. Cho, K.-W.; No, K. T.; Scheraga, H. A. J Mol Struct 2002, 641, 77.
71. Wie, D.; Guo, W.; Salahub, D. R. Phys Rev E 2001, 64, 011907-1-4.
72. MacKerrell, A. D., Jr. J Comp Chem 2004, 25, 1584.
73. Buckingham, A. D.; Fowler, P. W. J Chem Phys 1983, 79, 6426.
74. Buckingham, A. D. In The Hydrogen Bond. Recent theoretical and experimental advances in hydrogen bonded clusters; Xantheas, S. S., Ed.; Kluwer Academic Publishers: Dordrecht, 2000; p 1.
75. Stone, A. J. The Theory of Intermolecular Forces; Oxford University Press, 1997.
76. Diem, M. SPIE Biomol Spectrosc 1991, 1432, 28.
77. Woutersen, S.; Hamm, P. J Phys Chem B 2000, 104, 11316.
78. Eker, F.; Cao, X.; Nafie, L. A.; Schweitzer-Stenner, R. J Am Chem Soc 2002, 124, 14330.
79. Gronert, S.; O'hair, A. J. J Am Chem Soc 1995, 117, 2071.
80. Würtz Jürgensen, W.; Jalkanen, K. J. Phys Biol (accepted for publication).
81. Nardi, F.; Kemmink, J.; Sattler, M.; Wade, R. C. J Biomol NMR 2000, 17, 63.
82. Nardi, F.; Worth, G. A.; Wade, R. C. Folding Des 1997, 2, S62.
83. Jalkanen, K. J.; Nardi, F.; Wade, R. C.; Jung, C. (in preparation).
84. Hughes, J.; Smith, T. W.; Kosteriltz, H. W.; Fothergill, L. A.; Morris, H. R. Nature 1975, 258, 577.
85. Aburi, M.; Smith, P. E. Biopolymers 2002, 64, 177.
86. Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. J Chem Phys 2001, 114, 5149.
87. Jalkanen, K. J. J Phys: Condens Matter 2003, 15, S1853.
88. Nielsen, B. G.; Jensen, M. Ø.; Bohr, H. Biopolymers 2004, 71, 577.
89. Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, T.; Suhai, S. Chem Phys 2001, 263, 203.
90. Elstner, M.; Jalkanen, K. J.; Knapp-Mohammady, M.; Frauenheim, Th.; Suhai, S. Chem Phys 2000, 256, 15.
91. Palmo, K.; Mannfors, B.; Krimm, S. Chem Phys Lett 2003, 369, 367.
92. Ahlichs, R.; Penco, R.; Scoles, G. Chem Phys 1977, 19, 119.
93. Wu, Q.; Yang, W. J. Chem Phys 2002, 116, 515.
94. Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. J Chem Phys 2001, 114, 5149.
95. Cybulski, S. M.; Bledson, T. M.; Toczylowski, R. R. J Chem Phys 2002, 116, 11039.
96. Cybulski, S. M.; Chalasinki, G.; Moszynski, R. J Chem Phys 1990, 92, 4357.
97. Hesselmann, A.; Jansen, G. Chem Phys Lett 2003, 367, 778.
98. Hobza, P.; Sponer, J. J Am Chem Soc 2002, 124, 11802.
99. Hobza, P.; Selzle, H. L.; Schlag, E. W. J Phys Chem 1996, 100, 18790.
100. Wu, X.; Vargas, M. D.; Nayak, S.; Lotrich, V.; Scoles, G. J. Chem Phys 2001, 115, 8748.
101. Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. J Comp Chem 2002, 23, 1515.
102. Chen, I. J.; Yin, D.; MacKerell, A. D., Jr. J Comp Chem 2002, 23, 199.
103. Martin, M. E.; Aguilar, M. A.; Chalmet, S.; Ruiz-Lopez, M. F. Chem Phys 2002, 284, 607.
104. Spikeremann, M.; Bougeared, D.; Schrader, B. J Comp Chem 1982, 3, 354.
105. Jalkanen, K. J.; Frimand, K.; Niehaus, T.; Gale, J. D. (in preparation).
106. Jalkanen, K. J.; Degtyarenko, I. M.; Niehaus, T. A.; Frimand, K.; Suhai, S. (in preparation).
107. Niehaus, T. A.; Elstner, M.; Frauenheim, Th.; Suhai, S. J Mol Struct (Theochem) 2001, 541, 185.
108. Parman, T.; Wiley, M. J.; Wells, P. G. Nat Med 1999, 5, 582.
109. Guenzler, V. Nat Med 1999, 5, 853.
110. Wang, Y.; Gaspar, P. P.; Taylor, J.-S. J Am Chem Soc 2000, 122, 5510.
111. Bauernschmitt, R.; Ahlrichs, R. Chem Phys Lett 1996, 256, 454.
112. Becke, A. D. Phys Rev A 1998, 38, 3098.
113. (a) Perdew, J. P. Phys Rev B 1986, 33, 8822; (b) ibid 1986, 34, 7406.
114. Jorge, F. E.; Autschbach, J.; Ziegler, T. Inorg Chem 2003, 42, 8902.
115. (a) Jalkanen, K. J. World Association of Theoretical Oriented Chemists (WATOC 2005), Cape Town, South Africa; January, 2005; (b) Jalkanen, K. J. American Physical Society (APS) March Meeting, Los Angeles, CA, 2005, R.00191; (c) Bohr, H. G.; Jalkanen, K. J.; Malik, F. B. Mod Phys Lett B 2005, 19, 473.
116. Rahim, A. Towards Label Free Surface Enhanced Raman Immunoassays, Masters thesis; Quantum Protein Centre, Department of Physics, Technical University of Denmark, March 2004.
117. Hamm, P.; Woutersen, S.; Rueping, M. Helv Chim Acta 2002, 85, 3883.
118. Hamm, P.; Woutersen, S. Trends Opt Photon 2000, 43, 598.
119. Woutersen, S.; Hamm, P. J Chem Phys 2001, 114, 2727.
120. Woutersen, S.; Hamm, P. J Phys Chem B 2004, 114, 11316.
121. Woutersen, S.; Pfister, R.; Mu, Y.; Kosov, D. S.; Stock, G. J Chem Phys 2002, 117, 6833.
122. Woutersen, S.; Hamm, P. J Phys Condens Matter 2002, 14, R1035.
123. Stephens, P. J.; Jalkanen, K. J.; Kawiecki, R. W. J Am Chem Soc 1990, 112, 6518.

124. Tori, H.; Tasumi, M. J. Raman Spectrosc 1998, 29, 81.
125. Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenow, K. J Am Chem Soc 2001, 123, 9628.
126. Eker, F.; Cao, X.; Nafie, L. A.; Huang, Q.; Schweitzer-Stenner, R. J Phys Chem B 2003, 107, 358.
127. Schweitzer-Stenner, R.; Eker, F.; Perez, A.; Griebenow, K.; Cao, X.; Nafie, L. A. Biopolymers 2003, 71, 558.
128. Eker, F.; Cao, X.; Nafie, L. A.; Schweitzer-Stenner, R. J Am Chem Soc 2002, 124, 14330.
129. Schweitzer-Stenner, R. Biophys J 2002, 83, 523.
130. Fang, C.; Wang, J.; Charnley, A. K.; Barber-Armstrong, W.; Smith, A. B., III; Decatur, S. M.; Hochstrasser, R. M. Chem Phys Lett 2003, 382, 586.
131. Heid, M.; Schluecker, S.; Schmitt, U.; Chen, T.; Schweitzer-Stenner, R.; Engel, V.; Kiefer, W. J Raman Spectrosc 2001, 32, 771.
132. Bour, P.; Sopkova, J.; Bednarova, L.; Malon, P.; Keiderling, T. A. J Comp Chem 1997, 18, 646.