
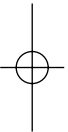



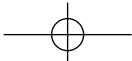
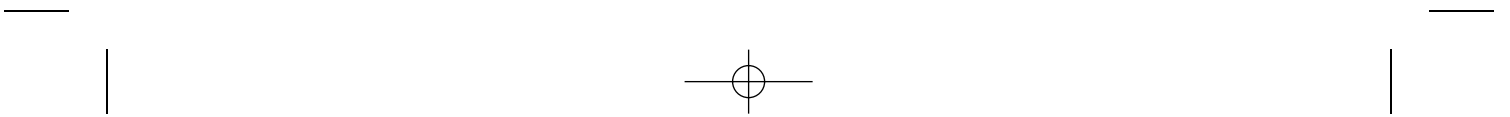
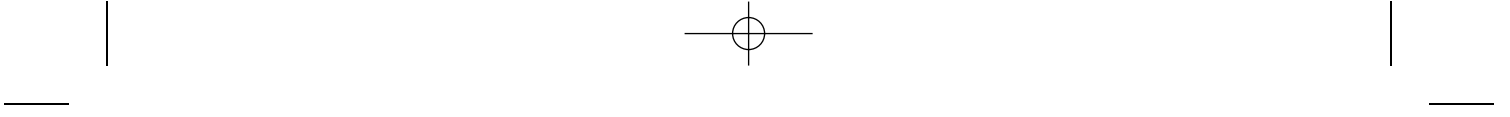


Part One
Quantum Mechanics Methods
for Molecular Biophysics and
Electronic Structure
Calculations of Biomolecules



Handbook of Molecular Biophysics. Methods and Applications. Edited by Henrik G. Bohr.
Copyright © 2009 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40702-6





1 Quantum Molecular Biological Methods Using Density Functional Theory

Karl James Jalkanen, Sándor Suhai, and Henrik G. Bohr

1.1	Introduction	9
1.2	Introduction to Density Functional Theory methods	22
1.3	Calculating Vibrational Absorption and Vibrational Circular Dichroism Spectra	28
1.3.1	Vibrational Absorption (VA) or Infrared (IR) Absorption and Vibrational Circular Dichroism	28
1.4	Calculating Raman Scattering and Raman Optical Activity Spectra	30
1.4.1	Raman Scattering	30
1.4.2	Raman Optical Activity	32
1.5	Classifying Conformational States and Structures based on VA, VCD, Raman, and ROA Spectra	35
1.5.1	Experimental Vibrational Frequencies for <i>N</i> -Methylacetamide, NALANMA, Peptides, and Proteins	35
1.5.2	Results for the DFT Calculations	35
1.5.2.1	L-Alanine	35
1.5.2.2	<i>N</i> -Acetyl L-Alanine <i>N'</i> -Methylamide	39
1.5.2.3	L-Alanyl L-Alanine	43
1.5.2.4	L-Histidine	49
1.5.2.5	<i>N</i> -Acetyl L-Histidine <i>N'</i> -Methylamide	51
1.5.2.6	<i>N</i> -Acetyl L-Alanine L-Proline L-Tyrosine <i>N'</i> -Methylamide	53
1.5.3	Future Perspectives and Applications for DFT-based Simulations	55
1.5.4	Future Perspectives of the Multidisciplinary Fields of Molecular Biophysics, Molecular Biology, Systems Biology, and Molecular Medicine and Molecular Psychiatry	56
	Acknowledgments	60
	References	60

1.1 Introduction

Toward the end of last century, when classical molecular dynamics (MD) became a very sought after technique for studying structure and conformations of large biomolecules, it was not an appropriate quest to employ and utilize quantum mechanical methods for studying dynamics of biomolecules (Warshel and Levitt, 1976; van Gunsteren *et al.*, 1983; Kollman, 1992). Initially, the potential energy surfaces were determined by fitting to experimental observed spectra (Colwell and Percival, 1983a,b), but later, also by parameterization/fitting to a combination of both experimental and theoretical data (Brooks *et al.*, 1983; Weiner *et al.*, 1984, 1986; Floris *et al.*, 1992; Brooks and Case, 1993; Cornell *et al.*, 1995; Dudek and Ponder, 1995; Maple *et al.*, 1998; MacKerell *et al.*, 1998). Owing to a limited amount of experimental data many assumptions and approximations needed to be made. These assumptions and approximations gained credence and support by being able to reproduce the (limited amount of) data. This has also created subsequent problems, as many of the assumptions and approximations involved throwing away/neglecting fundamental physics,

which has subsequently been shown to be important and is essential to predict some experimental observables. In many cases, the acceptance and widespread use of simplified models and theories have prevented researchers from further developing and extending the models and theories, as it has been hard to get funding to further develop and extend methods when the general consensus in the “application” field is that the methods are already of a sufficient level to warrant general use in applications. This has subsequently immensely slowed down the development of methods, as the funding has focused on applications to the detriment of further developing the methods. The black box approach to theory becomes prevalent and then researchers are limited to what is available in the packages like Gaussian, Amsterdam Density Functional (ADF), Dalton, QChem, Turbomole, Cambridge Analytical Derivatives Package (CADPAC), Spanish Initiative for Electronic Simulations with Thousands of Atoms (SIESTA), Vienna Ab-initio Simulation Package (VASP), and so on. The development and refinement of the methods is then left to a small number of researchers who all have their so-called development versions of these codes/packages. More recently, there has

been a call for more open-source codes, such as ABINIT, and for the National Science Foundation USA (NSF), Australian Research Council (ARC), National Institutes of Health (NIH), European Union (EU), and the National Health and Medical Research Council Australia (NHMRC) to continue to fund the development of methods for use in quantum nanobiology and molecular biology and for a core of young researchers to be supported to learn these methods by summer schools, North Atlantic Treaty Organization (NATO) workshops, and so on. This is clearly a high priority if the field is to continue to thrive and flourish and for the methods to continually be extended and developed, as problems are identified by the so-called users of the academic and commercial codes. The time lag between identifying the problems and implementing the changes/new developments needs to be reduced.

Basically, the overall goal of utilizing wave function and density functional theory based quantum mechanical methods for biomolecules is to derive from first principles functional and spectroscopic data. Here one calculates the structures, electric and magnetic properties, and vibrational, nuclear magnetic resonance (NMR), electron spin resonance (ESR), and electronic spectra of isolated molecules, aggregates, and molecular complexes in various environments where measurements have been made (nonpolar solvents, embedded within membranes, in solvents of varying polarity and ionic strengths, and finally in mixed environments, the most complex being proteins and glycoproteins that have part outside the membrane, part within the membrane, and part inside the membrane). Here one realizes that the structure, function, and properties

of these complex machines are lost when the machine (molecular complex) is broken down into its various parts (individual species). Here the combination of theory and experiment has a lot to offer, as either on its own is less useful.

Within the current set of assumptions and approximations, many physically measurable properties are predicted to be “zero,” which clearly shows that there are occasions and phenomena which cannot be treated by making the simplifying assumptions and approximations. The most commonly used ones being the so-called Born–Oppenheimer (BO) approximation and the harmonic approximation (both mechanical (for vibrational frequencies) and electronic (for infrared/vibrational intensities)). Actually, the concept of the potential energy surface which one parameterizes to generate a molecular mechanics force field is based on the validity of the BO approximation, which as we shall see later breaks down for some very interesting phenomena, the so-called non-BO properties. Many of these properties involve electric currents, magnetic fields, and angular momenta. The development/parameterization of classical mechanical force fields (and other molecular property surfaces which depend only on the nuclear positions of the atoms) is a field within itself, and the interested are encouraged to read speciality papers on this (Maple *et al.*, 1998; MacKerell *et al.*, 1998). The point here is that even though these molecular mechanics force fields are commonly used in molecular biology, they should be used only in specific cases where the assumptions and approximations made are valid. Clearly, for the case of quantum molecular biology, where a biological process is initiated by the absorption of a photon

of light, followed by subsequent electron and energy transfer, one must go beyond the use of molecular mechanics methods. Here, one can turn to either semiempirical-based methods or first principles/*ab initio*-based methods.

The field of developing and parameterizing molecular mechanical force fields (and property surfaces) has itself benefited immensely from both wave function quantum mechanics (WFQM)- and density functional theory (DFT)-based methods, as has the field of developing semiempirical-based methods, which are based on WFQM (Klopman, 1964a,b; Dewar *et al.*, 1985; Stewart, 1989, 2007) and DFT (Elstner *et al.*, 1998; Bohr *et al.*, 1999; Frimand and Jalkanen, 2002). In addition, as more accurate and complete experimental data becomes available, these data can be used both in the parameterization and testing/verifying stages. This is especially important for more exotic molecules and molecular complexes and assemblies (nanoparticles and protein–ligand complexes to name two), which have not been included in the data sets used for parameterizing and/or testing. In many cases, for example in non *ab initio* DFT methods, this new data provides a wealth of data to test the newly developed force fields, semiempirical methods, and new exchange and correlation (XC) functionals.

Since the beginning of this century, it has seemed much more appropriate to study, mainly, the electronic structure of large molecules with the help of quantum mechanical methods, with WFQM, DFT, and Green's functions being the three most commonly used. This is because the electronic structures and properties are necessary for deriving the spectroscopic data and

for understanding experimental observables (electric and magnetic properties) for the structures of molecules and molecular complexes/aggregates (Farid, 2002; Onida, Reining and Rubio, 2002; Amos, Jalkanen and Stephens, 1988; Amos *et al.*, 1987; Amos, 1982, 1986, 1984, 1987; Buckingham, 1967). Additionally, the electronic structures and properties are important for understanding the reactivity and functionality of biomolecules and biomolecular complexes (Deslongchamps, 1983; Hanesian, 1983). Hence one must extend the methods of classical MD simulations, where only one calculates the energy and gradients, to include the calculation of the various electronic and magnetic properties at each step of the MD simulation. Here, one has various ways to proceed. One can parameterize molecular property surfaces, similar to potential energy surfaces (force fields). For some properties, the property may depend on the momenta of the particles in addition to the positions (Buckingham, 1967; Amos, 1982, 1987; Buckingham, Fowler and Galwas, 1987; Lee, Colwell and Handy, 1994; Ioannnou, Colwell and Amos, 1997; Hornicek, Kapralova and Bour, 2008). The properties (observables) that depend on the momenta, in addition to the positions, are the so-called non-BO property surfaces. In some cases, there are equivalent expressions in terms of positions, momenta, and torque, which in theory should give the same results, but in practice for a variety of reasons do not (Stephens *et al.*, 1990; Amos, Jalkanen and Stephens, 1988).

Hence, the methodology and techniques which have been developed to parameterize potential energy functions can and have been extended to parameterize molecular property surfaces

(Amos, 1987). By calculating the properties at each step of the MD simulation, in addition to the energy, one can then simulate various measured properties and other experimental observables that depend on these molecular properties and how they change as a function of both positions/displacements and momenta. In addition, one would like to perturb the system and observe the system “evolve” back to equilibrium, if indeed this occurs. In some cases, the perturbation may induce conformational changes and even chemical reactions. For example, the absorption of ultraviolet (UV) radiation by DNA or RNA often involves the formation of either cyclobutane or oxetane rings, damaging the DNA or RNA. One can determine the correlated changes that occur between various quantities as the molecules and molecular complexes not only change their positions but also change their linear and angular momenta. Many of these quantities also change and are correlated with each other (this occurs due to solvation, hydration, desolvation and ligand binding, and the presence of spectator molecules like the chaperon proteins). Here, one has the ability to do correct averaging (in principle), which is harder to do when one calculates only a finite number of local minima.

Depending on the barriers between the local minima, experimental conditions such as temperature, pressure, and presence of chaperon proteins, and/or other species, which may reduce the barriers, the species and conformers present may vary (Ertl, 2008). One assumes in many cases that the system is in thermal equilibrium and that small fluctuations do not perturb the system. In some cases, the complete phase space available to the molecule

or molecular complex is sampled during the experiment (assuming the barriers are low enough). Here, one assumes that the system is in thermal equilibrium and that small fluctuations do not perturb the system, and so we do not need to take them, that is, the fluctuations in temperature (pressure), into account. In other cases, one may be trapped in a local minimum as the other local minima and even the global minimum are not accessible due to larger barriers. Hence one can measure the spectra and properties of a metastable state and think that one is in the global minimum due to the stability of the spectra and hence structure which give the unique and high resolution spectra. When one then perturbs the system and gives it enough energy to surmount the large barrier, the properties can then change drastically, which is thought to be the case for instance of the prion proteins. In this case, the global minimum of the individual molecules and the complexes (dimers, trimers, etc.) are not the same. For example, on dimerization fibrils form. The vibrational circular dichroism (VCD) and Raman optical activity (ROA) spectra of fibrils have been shown to reflect a large change from the nonfibril state. These spectroscopic techniques can hence be used to monitor such processes.

In some cases, one measures a superposition of the spectra of species present weighted by their Boltzmann populations, and in other cases, one actually measures an average spectrum, which may not even represent a real species/structure, but a sort of average structure. Hence, it is important to know both what one is measuring and one is simulating, and how to connect the two experiments, one in solution

and one *in silico*. One would like to determine the structure of the species present and also conformers in the case of flexible molecules from just vibrational spectral measurements, similar to what one now does with the X-ray and NMR methods. Many proteins and other molecules, molecular complexes, and molecular assemblies have escaped X-ray and NMR structure determination and it would be remarkable if vibrational structure determination were developed. To date, one can determine the percentage of secondary structure elements present and the side chain torsion angles for specific residues, but not the complete three-dimensional structure of a protein. What one would really like is to determine the side chain and backbone torsion angles for each residue. One promising extension of normal vibrational spectroscopy includes isotopic substitution at each residue in the protein to shift the modes and hence make them separate from those for the other $N - 1$ residues in the protein. This allows one in many cases to assign the backbone and side chain angles for only this specific residue, but not for the other $N - 1$ residues in the protein, which makes this impractical as a general method.

Finally, in addition to the ground state potential energy and property surfaces, one is interested in various excited state energy and property surfaces, and of course, the transition rates and probabilities to jump from one surface to another and back, not only in the absence of radiation but also in the presence of coherent and incoherent electromagnetic radiation (Onida, Reining and Rubio, 2002; Farid, 2002; Farid, 1999). There are many cases in biological systems where a photon (quanta) of light is necessary

to initiate a/the process, hence the term *quantum nanobiology*. One example is the repair of damaged DNA (due to UV radiation) in the presence of the photolyase protein and blue light (Jalkanen *et al.*, 2006a; Bohr, Jalkanen and Malik, 2005).

In this review, we present some of our contributions to the field of quantum nanobiology and also some main contributions from other researchers to this newly developing field and the outstanding problems that are yet to be addressed to make these simulations transparent and routine. This is done modestly by complete *ab initio* calculations of small biomolecules in various environments so as to derive structures, electronic and magnetic properties, and the various spectra of these molecules, aggregates, and molecular complexes. Hence, this area of research and study is a work in progress and each year more and more work is contributing to the field's development and evolution. The development of multidimensional infrared (IR) and Raman experiments, the analogs of two- and three-dimensional NMR, have recently been proposed (Cho, 2008; Scott, 2003; Jung, 2000; Mukamel, 1995), both of which in theory give more structural information than the simple one-dimensional spectra. What has limited the development and subsequent use of these techniques has been the development of a rigorous theory. The theories used to date involve many assumptions that have not been adequately tested, verified, and substantiated (Cho, 2008). Hence there is a need for the development of a rigorous high-level first principles theory and implementation, which can be used to test and benchmark the approximate classical, semiempirical, DFT, and *ab initio* methods. This is similar to the development of commercial VCD

and ROA instruments. Until rigorous theories were developed, the commercial companies were unwilling to invest in commercializing instrumentation for which the analysis of the spectra was not considered rigorous and transparent. In addition, pattern recognition methods should be developed which can be used to go directly from the measured spectra to structure or possible structures and not to have to do a complete MD, BOMD (Born Oppenheimer Molecular Dynamics) or even non-BOMD simulation for each system studied. Hence the term *vibrational structure determination* like X-ray and NMR structure determination.

In the following, we describe and summarize several applications of quantum mechanics to small biomolecules. We review our own and other groups' restricted Hartree–Fock (RHF) and DFT studies for hydrated L-alanine (LA) (Jalkanen, Bohr and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001, 2008a; Thar, Zhan and Kirchner, 2008), L-alanyl-L-alanine (LALA) (Knapp-Mohammady *et al.*, 1999; Jalkanen *et al.*, 2003), *N*-acetyl L-alanine *N'*-methylamide (NALANMA) (Jalkanen and Suhai, 1996; Deng *et al.*, 1996; Han *et al.*, 1998; Bohr *et al.*, 1999; Pohl *et al.*, 2007; Jalkanen *et al.*, 2008a), L-histidine (LH) (Bohr and Jalkanen, 2005; Jalkanen *et al.*, 2006a; Deplazes *et al.*, 2008), *N*-acetyl L-histidine *N'*-methylamide (NALHNMA) (Bohr and Jalkanen, 2005; Jalkanen *et al.*, 2006a), L-tryptophan (LW) (Jalkanen *et al.*, 2006a; Bombasaro, Rodriguez and Enriz, 2005), and capped L-alanine L-proline L-tyrosine (APY) [*N*-acetyl L-alanine L-proline L-tyrosine *N'*-methylamide (NALALPLYNMA) (Nardi, 1999; Nardi *et al.*, 2000; Jalkanen *et al.*, 2009c)] and examine the effects of

water hydration on the structures, the vibrational frequencies, vibrational absorption (VA), VCD, Raman scattering, and ROA intensities. The large changes that hydration has on the structures, relative stability of conformers, and on the VA, VCD, Raman, and ROA spectra observed experimentally are qualitatively reproduced by our DFT calculations when one adequately treats the environment, which to date has been neglected in the literature where one has mostly tried to correlate features in the vibrational spectra of proteins and peptides with known X-ray or NMR structures (Zhu *et al.*, 2005; Barron *et al.*, 2000). The goal has been to decompose the VA, VCD, Raman, and ROA spectra into their principle components, so that these principle components are the characteristic spectra of either local minima or parts of the structures of local minima, which together make up the measured spectra. This is the premise in the attempts to deconvolute or partition the measured spectra into their principle components (Kubelka, Pancoska and Keiderling, 1999; Baello, Pancoska and Keiderling, 1997; Baumruk, Pancoska and Keiderling, 1996; Keiderling, 1996).

In the past, theoreticians have neglected to properly take into account the effects of the environment on stabilizing structures that were thought to be stable without environmental influences (Jalkanen, Bohr and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998; Han *et al.*, 1998; Han *et al.*, 2000; Elstner *et al.*, 2000, 2001). Many of the routinely used molecular mechanics force fields have parameterized in local minima in the absence of solvent, where these local minima are only stable in the presence of the solvent (Jalkanen *et al.*, 2008a; Han *et al.*, 1998). Hence one must be

careful in using molecular mechanics force fields; cases still exist where the various force fields give different results. To be fair though, this is also the case with many of the routinely used generalized gradient approximation (GGA) and hybrid exchange correlation functionals. The methods, which have been shown to be systematically correct, are the very expensive and rigorous *ab initio* DFT and coupled-cluster methods of the Bartlett group. The methods have been used to benchmark the more approximate and computationally faster methods: GGAs, hybrid XC functionals and, of course, the SCC-DFTB (Self Consistent Charge-Density Functional Tight Binding) and PM6 (Parametric Method 6) semiempirical DFT and WFQM methods.

In addition to the work on hydration of amino acids, peptides, and proteins via vibrational spectroscopy, there is a large body of related work on gas hydrates, both experimental and theoretical (van Klaveren *et al.*, 2001). The formation of gas hydrates in natural gas pipelines initiated a lot of fundamental work on characterizing and understanding the process of hydrate formation (Hammerschmidt, 1934). Furthermore, calorimetric methods have been used to determine the compositions, heat capacities, and enthalpies of dissociation for a number of hydrates (Handa, 1986a,b). MD simulation of N₂ and O₂ hydrates have been used to interpret the double maximum in large clathrate cages and single maxima in small clathrate cages for N₂ and double maximum in both cages for O₂ (Horikawa *et al.*, 1997). Here, one is able to use MD simulations to interpret the experimentally observed spectra of gases isolated in various cages of water molecules. Similar types of experiments and MD simulations can be

used to understand the hydration of amino acids, peptides, and proteins.

Neutron and X-ray diffraction experiments have shown that the water molecules encapsulating peptides and proteins are different from bulk water (Sauter *et al.*, 2001). The question of the encapsulation of polar and nonpolar groups in amino acids, peptides, and proteins and the effects of hydration on the vibrational spectra, both frequencies and VA, VCD, Raman, and ROA intensities are very much related. To date, very little work has been done on analyzing the changes of VA, VCD, Raman, and ROA intensities due to the environment. This is due to the fact that large changes are observed because of conformational changes induced by the changing environment. Here, the two effects may not be able to be distinguished. Indeed, the work of Han *et al.* (1998) on NALANMA has shown that the species of interest may not be the isolated molecule, as it is in nonpolar solvents, in the gas phase or at low temperature in inert gas matrices, but the molecule and its bound water/solvent molecules, a molecular complex. The topology of the potential energy surface of the molecule appears to be very solvent/environment dependent, a fact that has not been appreciated in the past. Much of the complexity that appears in the gas phase appears to go away in strongly interacting polar media like water. But then the question is how does one determine the species of interest/study? This is especially important and relevant when one wishes to treat the systems in layers using various levels of theory, that is, different models for each layers.

The assumption that the properties of the molecular complex are a simple superposition or sum of the properties of

the individual molecules (parts) may be in gross error. Hence the continued development of fields of complex systems and systems biology. Chaos theory is another example of a theory being developed which describes phenomena which cannot be explained using the standard model.

Finally, the treatment of solvent effects on crystallization and predicting the morphology of the crystals formed has recently been addressed by Gale and Rohl (2007). They have extended the COSMO (Conductor-Like Screening Model) model to treat solute molecules in solution to deal with solute molecules in a mother liquor at the crystal surface. Here they calculate both the surface energy and attachment energy in the presence of implicit solvent (Gay and Rohl, 1995). Previous work calculated these energies from the gas phase, which, though important, are not consistent with the media from which many crystals are grown. The molecules in the mother liquor can either adsorb on various crystal faces or desorb and go back into solution. MD simulations of this process are very time consuming and Gale and Rohl have developed a method that is highly efficient and uses classical potentials. The method can be extended to *ab initio* and DFT methods, but to date this has not been done.

In addition to changes in the vibrational spectra, the thermal conductivity of gas hydrates changes; it is much lower than that of ice and has been assigned due to coupling between the guest molecules and the water framework (Tse *et al.*, 2001; Inoue, Tanaka and Nakanishi, 1996). Changes in the vibrational spectra were also simulated and are further evidence for the coupling of the modes of the crystalline water cages and

the embedded guest molecules. The low frequency modes have been observed in the incoherent inelastic neutron scattering (IINS) spectra, at frequencies which are very hard to observe in the IR and Raman spectra (Gutt *et al.*, 2002). Hence it is very useful to measure the IINS spectra for molecular hydrates and complexes, in addition to the VA, VCD, Raman, and ROA spectra (Oyama *et al.*, 2003; Subramanian and Sloan, 2002).

The IINS spectra for amides have also recently been analyzed with MD simulations (Itoh *et al.*, 2003). Early simulations of the IINS spectra for *N*-methylacetamide (NMA) using simple models have been shown to be in error (Fillaux *et al.*, 1993), even though they were consistent with the experimental data (Fillaux, 1999, 2004; Kearley *et al.*, 1994, 2001; Fillaux *et al.*, 1998; Kearley, 1986, 1995; Kearley, Tomkinson and Penfold, 1987). The models for the IINS reassigned the so-called amide A N–H stretch mode in NMA to be approximately half the value found in the isolated state in nonpolar solvents and in low temperature matrices. In the solid state, the NH hydrogen is hydrogen bonded with another C=O oxygen of an adjacent NMA molecule. The model used was a double-well potential and the model fit consistently all of the IINS data, so the work was accepted and caused quite a controversy in the conventional vibrational spectroscopy field, both experimental and theoretical. Subsequent work using a more rigorous *ab initio* model has shown that the initial assignment was right and the proposed new model, although able to reproduce the experimental data, was in error (Kearley *et al.*, 1994). This was due to the approximations and assumptions being made that are not strictly valid.

This emphasizes the fact that just because a simple model is able to reproduce a limited amount of experimental data does not mean it is correct. Hence the use of Ockham's razor does not always apply (it states that essentially one should use and accept the simplest possible explanation and/or model and/or theory). In this case, the more sophisticated model was necessary to get the physics correct. In some cases, one is happy to use a nonphysical model, as is quite often used in classical molecular mechanics and MD simulations that use harmonic potentials for bond stretches. But if one wants to make predictions and understand the physics of the phenomena, then one should strive to use models and theories which include the basic physics of the problem. Hence the term *ab initio* or *first principles theories*. But in many cases even the so-called first principles or *ab initio* theories involve approximations, the one most commonly made, but not explicitly stated being the BO approximation. There are many phenomena that cannot be described within the BO approximation, for example, the electronic contribution to the atomic axial tensors (AATs), which are needed to predict/simulate the VCD intensities.

Hence it is very important that the model is used to simulate as much of the available experimental data (Serdyuk, Zaccari and Zaccari, 2007). MD simulations have recently been extended to grow methyl hydrate sII from a supersaturated aqueous methane solution (Vatamanu and Kusalik, 2008). The next application is to develop inhibitors, as hydrate formation in natural gas pipelines is very problematic and much research has been undertaken to understand and to prevent it from occurring, including the use of

proteins (Zeng *et al.*, 2003; Moon, Taylor and Rodger, 2003). Here the two fields merge. The multidisciplinary nature of many problems is finally being appreciated. Here researchers can really gain from working together and getting fresh new ideas and perspectives on their problems from researchers in other fields. In many cases, having to rigorously defend one's methods and approaches, and inherent assumptions, results in one's finding some of the weaknesses that do not get noticed when one just turns the crank. It is good every once in a while to analyze the gears and mechanism that go along with the crank.

Matrix isolation (Whittle, Dows and Pimentel, 1955; Pimentel, 1975), X-ray diffraction, and solid-state vibrational studies have been undertaken to try to determine relationships between structural parameters (both intermolecular, intramolecular, and intermolecular hydrogen bonding) and vibrational frequencies (VA, VCD, Raman, and ROA intensities) (Charles and Lee, 1965; Pimentel and Charles, 1963; Pimentel and Sederholm, 1956). Early Raman studies on lysozyme, ribonuclease, and α -chymotrypsin documented that vibrational studies could also contribute information to help understand the complexities of biomolecules in aqueous solution (Lord and Yu, 1970a,b).

In addition, one has also constructed software systems such as artificial neural networks to solve the inverse scattering problem of retrieving structural information of the biomolecule from spectroscopic data, that is, vibrational frequencies as well as VA and VCD intensities of isolated NALANMA (Bohr *et al.*, 2001; Ramnarayan, Bohr and Jalkanen, 2008).

Two possible routes have been considered to predict the structure of our test

molecule from the spectroscopic data. One route is to use DFT to calculate all the possible structures/conformers and their corresponding frequencies, VA, VCD, Raman and ROA intensities and then compare them to the experimental data. The other is to train expert systems, for example, neural networks on a large combination of calculated and experimental spectra with known structures, either determined experimentally from X-ray or neutron diffraction or NMR. Then correlations and/or component spectra are used to infer or extrapolate new results, that is, to predict the structures, fold classes, and percentage of secondary structural elements present from only the measured VA, VCD, Raman, and ROA spectra.

In case of large biomolecules, one can determine whether there are correlations between the best-predicted structural details from spectroscopic data and the data connected to secondary structure stability and the structures predicted from neural network programs, which are based on only sequence and structural correlations. The main problem that prevented a complete protein structure to be determined from the combined VA, VCD, Raman, and ROA spectra has been to determine which spectroscopic data are the most important determiners of secondary structures, so that such information can be used to predict secondary structural elements, and then subsequently how these secondary elements fold up to form their tertiary and finally even quaternary structure. One assumes that the information is contained in the combined VA, VCD, Raman, and ROA spectra: frequencies, line widths, and integrated intensities; polarization properties of the intensities; and finally the frequency dependence of the Raman

and ROA intensities. To date the methods are only able to predict percentages of various secondary structural elements, but not to assign the secondary structural elements to each and every residue in the protein.

Finally, in addition to the secondary structural element (ϕ_i and ψ_i angles) for each residue, one needs to determine the side chain torsion angles $\{\tau_{ij}\}$ and any disulfide or salt bridges connecting two nonadjacent residues (giving extra stability to this structure). In many cases, the protein folds up to its “native state” and is then posttranscriptionally modified/cleaved by an enzyme to achieve its functional state. Insulin is formed from the larger proinsulin by such a process. Hence, in many cases, the functionally active state of a protein does not correspond to any one gene or protein, but to a set of proteins/genes working in a concerted manner. Hence the term *systems biology*. One cannot really understand biological function by studying individual molecules or even individual molecules in their native environment. One must know how the molecules interact with other molecules to do so. In the case of insulin, the ligand binds to the transmembrane insulin receptor extracellularly, which triggers dimerization and then a subsequent activation of intracellular tyrosine kinase activity, the so-called signal transduction process. This is as complex as its get, and one would like to be able to use spectroscopy to follow each and every step of this very complex process.

What has made this problem more difficult for biomolecules is the absorption of radiation by the environment, water being a strong absorber in the IR, and the strong interaction between the water molecules and the biomolecules

via hydrogen bonding. Hence one must consider in their native environment, the biomolecules that interact with other biomolecules, counterions, cations, anions, and finally prosthetic groups and bound ligands, which in many cases are necessary for the biomolecules to function. On ligand binding, for so-called induced fit binding, both the protein and ligand can undergo conformational changes (Keskin *et al.*, 2008). Many of the current docking algorithms assume a lock- and key-type interaction, and hence fail to predict the binding of ligands via induced fit, and hence attempts have been made to improve and extend the current docking models (Böhm and Klebe, 1996; Davis and Teague, 1999; Cavasotto and Abagyan, 2004; Oliveira *et al.*, 2006).

The use of VA, VCD, Raman, and ROA spectroscopy are ideal tools to monitor such conformational changes in both the ligand and the protein on ligand binding. In addition, VCD and optical rotation have been shown to be sensitive enough to be able to observe the induced chirality in achiral solvent molecules and ligands on binding or even interacting with a chiral environment (Losada and Xu, 2007; Jalkanen *et al.*, 2008a; Mukhopadhyay *et al.*, 2007; Losada, Tran and Xu, 2008). The sesquiterpene molecules have multiple chiral centres and in many cases flexible side chains, which though achiral, have induced chirality due to the interactions with chiral cyclic system (Jalkanen *et al.*, 2008c; Kulcitki *et al.*, 2005; Weyerstahl *et al.*, 1995; Ayafor *et al.*, 1994; Newman, 1972).

The larger goal is to find/develop/derive ways to determine structural minima, not only for the individual molecules but also when they interact to form ligand–protein complexes. In

addition, in many cases there are other participants in the stable and functional biological unit, a metal center, bound solvent molecules, prosthetic groups like the porphyrin ring in myoglobin and hemoglobin, and finally, cations or anions (Sauter *et al.*, 2001). For these minimum energy stable complexes, VA, VCD, Raman, and ROA intensities are calculated by DFT in order to produce training data, that is, sets of spectroscopic data correlated with $\phi - \psi$ angles and other structural determiners for the network. Other methods such as X-ray crystallography and NMR have only been utilized to determine the native states of proteins. The VA, VCD, Raman, and ROA spectra for proteins and peptides with known structures should also be added and be part of the training set. By using DFT methods, the VA, VCD, Raman, and ROA spectra of the so-called denatured states of proteins and peptides can also be added to the training set. Many of the so-called random coil and denatured states have been shown to have a lot of secondary structure in the CD, VCD, and ROA spectra. All that one can say for sure is that these structures are different from the so-called minimum energy structures under physiological conditions. In some cases, these structures may even be global minima and the structures which are biologically active may actually be in metastable states. For example, the prion proteins may be lower in energy in the so-called denatured state, thereby making attempts to refold the prion proteins very difficult. In addition, it has recently been proposed that the functional states of some proteins are only present when the ligand is bound and the so-called native isolated protein is not one conformer (global deep minimum energy state), but

an ensemble of structures (Keskin *et al.*, 2008; Sauter *et al.*, 2001). Only on ligand binding or in complex with another protein, does a single unique functional conformer arise. This could account for the failure of X-ray and NMR techniques in determining the unbound state(s) and structure(s) of this class of proteins.

One of the main problems is to determine the structures from the VA, VCD, Raman, and ROA spectra of the denatured states of proteins. Keiderling and coworkers have utilized neural network methodology to find correlations of electronic absorption (EA), electronic circular dichroism (ECD), VA, and VCD spectra with the native states of proteins by utilizing the known NMR and X-ray crystallographic structures (Pancoska, Yasui and Keiderling, 1989, 1991; Pancoska, Blazek and Keiderling, 1992; Pancoska *et al.*, 1994, 1995; Pancoska, Janota and Keiderling, 1996; Baumruk, Pancoska and Keiderling, 1996; Baello, Pancoska and Keiderling, 1997; Pancoska, Kubelka and Keiderling, 1999; Kubelka, Pancoska and Keiderling, 1999). Our work compliments their work in providing correlations of VA, VCD, Raman, and ROA spectra and the higher energy denatured states of peptides and proteins. These denatured states can be produced under various experimental conditions, that is, in aqueous solution, for example, under various pHs and salt conditions and by the presence of urea and other denaturing conditions.

In that sense one should be able to predict higher level intermediate energy states during folding processes of biomolecules with the help of expert software systems such as artificial neural networks and principle component analyses that have to be trained on known data in order to predict structural data novel

to the networks. Once they are trained on known sets of intermediate energy states, that is, the set of structural data we present to the network they should be able to extrapolate to new relations between spectrum and structure. The advantage of utilizing software systems such as neural network techniques and principle component analyses for the inverse scattering problem of deriving structural information from scattering data is that they go hand in hand with experiments and DFT calculations in the sense that one of the tools can support the other when it fails or is impractical. The big endeavor is to produce detailed structural data of proteins and peptides, both in the isolated state and in complex with each other and other proteins, and either measure the experimental VA, VCD, Raman, and ROA spectra and/or simulate them with DFT calculations. In many cases the spectra and structures of proteins can be assembled from the principle components, that is, small peptides that constitute the whole protein.

In many cases, the diverse conformational spectrum of small peptides is greatly reduced when they form part of a protein. Hence one may not need to determine the structures and spectra for all possible conformers of small peptides, which itself is a very challenging problem. But like the protein not sampling all possible conformational states, peptides may and probably do not sample all possible states either. This has also been confirmed via MD simulations. Hence the many groups and researchers working on methods to sample the complete conformational space for proteins and peptides, though interesting from an academic point of view, may not really be necessary. This may be necessary to do once for each amino acid and

then for sequence of dipeptide, tripeptide, and so on, but is really a waste of time and resources to do again and again for each protein and/or for each new study/simulation. What is required is to assemble/collect the results from various laboratories and annotate and verify the data before making it available to the research community. This is similar to what has been done for both experimental X-ray and NMR structures. What is needed is a similar database for theoretically determined structures. In addition, what is important for the molecular biophysicists, molecular biologists, and quantum nanobiologists to understand are the biophysical forces and interactions that are responsible for not only stabilizing the conformers, structures, and assemblies but also the biophysical processes and events which are responsible for function, and dynamics. In the opinion of these authors, too much emphasis has been put on static structures and not enough on dynamics and function. In many cases, the dynamical fluctuations and function go hand in hand. Vibrational spectroscopies are ideal techniques to study and investigate both dynamics and function.

If biomolecules and biomolecular complexes only sample a small part of the total available phase and conformational space possible, then this is what one would like to simulate, and not large portions of phase space which are not sampled and/or relevant for the experiments being studied. Part of this question can be addressed with, for example, neural network and principle component analyses tools determining the dihedral angles in the protein that are actually populated from the VA, VCD, Raman, and ROA data for each subunit. If one is able to achieve this goal, then the

combination of VA, VCD, Raman, and ROA spectroscopy will take a seat along with X-ray and neutron diffraction and NMR spectroscopies as standard tools for structure determination of proteins and peptides.

A library of various conformers for each amino acid along with its corresponding properties as a function of various environments (aqueous, nonpolar membrane, or mixed) would be very beneficial, similar to the current protein data base (PDB) for X-ray and NMR-determined structures and the Cambridge crystallographic database for small molecules and ligands. Of course, along with amino acids, its corresponding capped amino acid, and sequences of dipeptides, tripeptides, and so on, data of dipeptides, tripeptides, and so on should also be collected/assembled and most certainly annotated. Many experiments and theoretical simulations have not been able to be reproduced (a fundamental part of good science) due to the lack of being adequately annotated/documentated. A simple case is of whether one used distilled water or distilled *and* deionized water *and* whether one carefully controlled the pH with a buffer solution and then, which buffer, HEPES (N-2-Hydroxyethyl-piperazine-N¹-2-Ethanesulfonic Acid), phosphate, and so on. Some effects due to different buffering solutions have even been observed. The regrettable fact is that, in many experiments, everything seems to matter, and hence needs to be properly documented and controlled.

This large (and ever increasing) amount of data could then be used either in the parameterization and/or testing of new molecular mechanics and semiempirical methods, and even newly

developed XC functionals. It is important to continually test and retest and, if necessary, reparameterize and further develop and extend methods as problems and/or disagreements arise and are found. A classical example has been the so-called standard hybrid XC functional, the B3LYP, now being replaced by a variety of more accurate and general XC functionals, one such being the B3PW91.

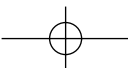
In the next sections, we present our theoretical approach to simulate the VA, VCD, Raman, and ROA spectra of proteins and peptides. First, we present an introduction to DFT methods. Then a presentation is given on how one calculates/simulates VA and VCD spectra followed by the same for Raman and ROA spectra. The last section covers the use of VA/VCD/Raman/ROA experimental and theoretical spectroscopy in molecular biophysics, molecular biology, and quantum nanobiology.

1.2 Introduction to Density Functional Theory methods

DFT is in principle an *ab initio* method that allows one to calculate the structures and properties of single isolated molecules and various aggregated complexes (Hohenberg and Kohn, 1964; Mermin, 1965; Kohn and Sham, 1965; Pople, 1999; Kohn, 1999; Bartlett *et al.*, 2005; Bartlett, Lotrich and Schwiegert, 2005). It is more complete than Hartree–Fock methodology in that it includes electron correlation and is better than other correlated methods because it requires less computational resources (disk, memory, and CPU time) than otherwise, for the same computational complexity. DFT formulates the energy of the system

in terms of the electron density and electron current density instead of the wave function. The trade-off is that one must determine the various terms in the Hamiltonian for the total energy of the system, the kinetic energy (nuclear and electron) and Coulomb energy (nuclear–nuclear, nuclear–electron, and electron–electron) in terms of the electron density, current density, and perhaps the gradients and higher order spatial derivatives of these so-called observables. Since these quantities are themselves functions of functions, they are called functionals. Various local and nonlocal electron density functionals and local electron current density functionals have been determined and tested in their ability to determine structures, relative energies, binding energies, dipole moments, polarizabilities, hyperpolarizabilities, dipole moment derivatives, NMR shielding tensors, and various other molecular and aggregate properties (Ma and Brueckner, 1968; Rajagopal and Callaway, 1973; Perdew and Levy, 1983; Sham and Schlüter, 1985; Lannoo, Schlüter and Sham, 1985; Vignale and Rasolt, 1987, 1988; Godby, Schlüter and Sham, 1988; Becke, 1988a, 1993a; Lee, Yang and Parr, 1988; Levine and Allan, 1989; Papai *et al.*, 1990; Perdew and Wang, 1992; Malkin, Malkina and Salahub, 1993a,b; Komornicki and Fitzgerald, 1993; Colwell *et al.*, 1993; Lee and Sosa, 1994; Johnson and Frisch, 1994; Schreckenback and Ziegler, 1995; Malkin *et al.*, 1995; Wilson, Amos and Handy, 1999; Wolff, 2005; Kongsted and Ruud, 2008).

The first Hohenberg–Kohn theorem showed that the electron density determines the energy and hence reformulated the basic equation to solve as one in which one has to determine the electron



density rather than the wave function (Hohenberg and Kohn, 1964). The energy function can be written as

$$E_v[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] \quad (1)$$

$$= \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \quad (2)$$

where

$$F_{HK}[\rho] = T[\rho] + V_{ee}[\rho] \quad (3)$$

and $v(\mathbf{r})$ is the external potential.

The second Hohenberg–Kohn theorem provides the energy variational principle, which enables one to find the density that minimizes this energy functional (Kohn and Sham, 1965). The problem is we do not know the functional $F_{HK}[\rho]$ and many functionals have been developed that try to address this problem. By introducing the Kohn–Sham orbitals, Kohn and Sham were able to address the problem of not being able to determine the electron kinetic energy functional in terms of only the density. At this point, they could have also used the Kohn–Sham orbitals to determine the exchange energy functional by simply substituting now the Kohn–Sham orbitals within a Slater determinant. But by doing this, they would have introduced a nonlocal exchange term and the simplicity and robustness of DFT would not have been apparent. So instead of taking this route, which would have been to search for the correlation energy functional also in terms of the Kohn–Sham orbitals, Kohn and Sham explored using the local density approximation (LDA) for both the exchange and correlation energy functionals. For problems in solid-state physics, LDA works well and has been used for decades now. Owing to problems with scaling in WFQM, quantum chemists turned

to DFT as an alternative to the WFQM semiempirical based methods. For some problems LDA did fine, but for complex organic and bioorganic molecules, especially those involving hydrogen bonding, one requires higher accuracy than LDA can give. This led to the development of the so-called GGA functionals, where in addition to the density, the exchange and correlation energy functionals are functions of the gradient of the density, in addition to the density (Perdew, Burke and Ernzerhoh, 1996, 1997). Still higher accuracy was required, and to get this higher accuracy one went to the nonlocal Hartree–Fock form for the exchange energy functional, substituting the Kohn–Sham orbitals into the expressions instead of the Hartree–Fock orbitals. This led to the so-called hybrid XC functionals.

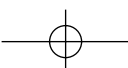
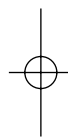
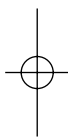
Commonly used hybrid functionals, the Becke 3LYP (B3LYP) (Becke, 1988a, 1993a; Lee, Yang and Parr, 1988; Perdew and Wang, 1992),

$$E[\rho(x, y, z), \nabla(\rho(x, y, z))] \\ = E_{RB} + E_{HF} - E_{LYP} \quad (4)$$

and Becke 3PW91 (B3PW91)

$$E[\rho(x, y, z), \nabla(\rho(x, y, z))] \\ = E_{RB} + E_{HF} - E_{PW91} \quad (5)$$

hybrid exchange correlation functionals have been implemented in CADPAC, Dalton, Turbomole, and Gaussian 94, 98 and 2003 (Stephens *et al.*, 1994; Lee and Sosa, 1994; Becke, 1993a; Lee, Yang and Parr, 1988). B3LYP and B3PW91 level analytical Hessian and atomic polar tensors (APT) calculations have also been implemented in Gaussian 94, 98, 03, CADPAC, and Dalton. More recently, the AAT have been



implemented in the ADF suite of programs utilizing Slater-type orbitals (STOs) (Nicu *et al.*, 2008). Previous implementations had all been with the Gaussian-type orbitals (GTOs), which have been used by quantum chemists, since these allow one to analytically evaluate many integrals, in contrast to the use of numerical integrations to evaluate many of the integrals when using STOs. But with the use of numerical integration to evaluate many of the exchange and correlation integrals required for various XC functionals, the necessity and advantage of using strictly GTOs no longer exists, and hence the return to the use of STOs.

The B3LYP and B3PW91 level force fields have been shown to be more accurate than RHF level Hessians, which must be scaled to get good agreement with both experimental frequencies and VA and VCD intensities (Jalkanen and Stephens, 1991; Stephens *et al.*, 1994; Jalkanen and Suhai, 1996; Jalkanen *et al.*, 2008a,b; Deplazes *et al.*, 2008). The nature of the normal modes have been shown to depend very much on the scaling scheme one chooses to scale the Hessian (Jalkanen *et al.*, 1987, 2008b; Jalkanen and Stephens, 1991). The advantage of the B3LYP and B3PW91 hybrid exchange correlation functionals is that the Hessians appear to be accurate enough to predict the VA and VCD intensities when coupled with accurate APT and distributed origin (DO) gauge and gauge invariant atomic orbital (GIAO) AAT without scaling. The number of molecules for which the B3LYP and B3PW91 Hessians, APT and AAT, have been calculated and used to predict and help interpret the VA and VCD spectra has been quite extensive now.

This level of field has become the standard where either B3LYP or B3PW91 are able to predict the correct structure (conformer) and relative energies. The good agreement shown to date has included only a small number of functional groups and the comparison has been with measurements of the VA and VCD spectra of molecules in nonpolar solvents. Recently, systematic errors with B3LYP have been found (Matsuda, Wilson and Xiong, 2006), and hence the need for a more systematic approach to develop XC functionals. The *ab initio* DFT method of Bartlett, which uses optimized effective potentials (OEPs) for both exchange and correlation (XC), appears to be the most promising (Bartlett *et al.*, 2005; Bartlett, Lotrich and Schwegert, 2005). The work of Bartlett built on previous works with OEPS is covered briefly and cited in the classical quantum chemistry text by Levine (2000).

Even though much more expensive than the GGAs like PBE, OEPs give the correct answer for the correct reasons (correct physics and chemistry), and hence the additional effort and cost is justified. Therefore, the OEP XC functionals of the Bartlett group have become the new gold standard and, all the more, simple, approximate, and robust XC functionals should be tested and benchmarked against the Bartlett OEPs. With the work of Bartlett and coworkers, DFT has now been put on as a rigorous level as WFQM, with a systematic way to improve both the XC functionals achievable, though at much higher cost and complexity than originally formulated first by Thomas, then Thomas and Fermi, and finally by Hohenberg, Kohn, and Sham. The merging and integration of the works by Kohn and Pople and many others in DFT and WFQM

appears to have finally been achieved. Now one can use DFT and WFQM and semiempirical methods based on each almost interchangeably, depending on the focus of one's work. Each method, as well as Greens functions-based methods, has its advantages and disadvantages, especially with respect to the treatment of excited electronic states and excited state potential energy surfaces, which are required for photon-induced photochemical reactions. Many reactions in molecular biology and molecular biophysics are photon initiated, so there is plenty of work to be done to further develop and refine DFT-based methods for electronic excited state MD. In this realm, WFQM and Greens functions-based methods are more accurate and should be used to test and benchmark the time-dependent density functional theory (TD-DFT)-based methods (Onida, Reining and Rubio, 2002; Reining *et al.*, 2002; Ismail-Beigi and Louie, 2003; Zhukov, Chulkov and Echenique, 2004).

In this review, we summarize previous works and present some new results on some small peptides (LALA, NALANMA, and NALHNMA) and two amino acids (LA and LH). LA and LH are two clear examples of a biological molecules whose properties and structures in water and gas phase are very different. We have found that without the water being explicitly or implicitly accounted for, the gas phase or isolated molecule calculations are of questionable use in analyzing the measurements on these molecules in aqueous solution, though they have found recent application in mass spectroscopic studies (Ipolyi *et al.*, 2006). The zwitterionic structure of LA, the predominant species in aqueous solution, is not stable in the gas phase according

to our 6-31G* B3LYP *ab initio* calculations (Jalkanen, Bohr and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998). Therefore we have explicitly included solvent, in this case water, to stabilize this structure and also to be able to compare to the measurements of the VA, VCD, Raman, and ROA spectra of LA in aqueous solution (Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001, 2008a). The VA and VCD spectra of LA in aqueous solution were initially measured by the groups of M. Diem at Hunter College in NY, NY (Diem, 1988) and L. Nafie at Syracuse University in Syracuse, NY (Freedman *et al.*, 1988). Raman and ROA measurements and calculations on LA have also been reported by Barron and coworkers (Barron *et al.*, 1991a). At our *ab initio* 6-31G* B3LYP DFT-optimized geometries we have calculated the vibrational frequencies, VA and VCD intensities. The VA and VCD spectra were calculated with the 6-31G* B3LYP force field (Hessian) and APT and the 6-31G** RHF DO gauge AAT.

For LALA, three zwitterionic structures were able to be stabilized with four, six, and seven water molecules, respectively (Jalkanen, Bohr and Suhai, 1997; Bohr and Jalkanen, 1998). The zwitterionic structure was not stable relative to proton transfer of one of the protons of the NH_3^+ group to the C=O group of the adjacent amide group, whereas the proton of the N-H group was concurrently transferred to the closest oxygen of the CO_2^- group. Subsequently completely solvated (hydrated) structures of the LALA zwitterion were extracted from MD simulations and subsequently optimized at the B3LYP level of theory (Knapp-Mohammady *et al.*, 1999). Here

only the water molecules directly hydrogen bonded with the polar groups were kept and the number of waters varied from 10 to 14, depending on the conformation of the LALA zwitterion. In a subsequent study, we used one of the structures and predicted the VA, VCD, Raman, and ROA spectra (Jalkanen *et al.*, 2003). Here we present our new results for the LALA plus 11 water clusters embedded within the polarizable continuum model (PCM). We present the VA, VCD, Raman, and ROA spectra for the LALA zwitterion with 11 water molecules. As one can see the water absorption masks many of the LALA modes. Note that in addition to masking the LALA modes, the water modes couple with the LALA modes and therefore the water molecules in the first solvation shell surrounding the LALA zwitterion are different than those in bulk water. This has been documented by X-ray and neutron diffraction studies as well as from VCD signals due to chiral water molecules (the chirality is due to the chiral environment of the LALA molecule) (Sauter *et al.*, 2001; Losada and Xu, 2007; Losada, Tran and Xu, 2008).

The goal there was and here is to determine the structure (a conformer) of the zwitterionic form of LALA in aqueous solution from the experimental vibrational spectra, VA, VCD, Raman, and ROA, if possible. As shown by our calculations on the neutral species, there are many possible conformers for this molecule and the structure(s) of this molecule in water has(have) not yet been determined, though the VA, VCD, Raman, and ROA spectra have been measured. The goal of our work is to help answer the question of the conformation and species of LALA in aqueous solution and to develop

and document a methodology which will work for larger peptides and even proteins in both native and nonnative conditions.

We also present here optimized structures of NALANMA with four waters starting from our 6-31G* B3LYP optimized structures. The relative energies of these complexes are compared with the isolated molecule values. In addition we explore the effect of changing the hybrid exchange correlation functional, as recently the B3LYP hybrid exchange correlation functional has been called into question. Hence we have replaced the LYP correlation functional with the PW91 correlation functional and have optimized the structure of NALANMA with the four bound waters embedded within the PCM and conductor-like polarizable continuum model (CPCM) continuum solvent models with the B3PW91 hybrid exchange correlation functional (Jalkanen *et al.*, 2008a).

The goal here has been to model biomolecules by explicitly adding water molecules to provide calculations which can be used to critically evaluate solvent models and specific models developed for water (liquid simulations) (Jorgensen, Madura and Swenson, 1984; Jorgensen and Swenson, 1985a,b; Jorgensen and Tirado-Rives, 1988; Jorgensen, Maxwell and Tirado-Rives, 1996). The H-bonding properties as exemplified by some of the simple water models are clearly incorrect, and we feel that conclusions based on these models can be critically evaluated utilizing the better models of water (Berendsen *et al.*, 1981). In addition, the properties for the water potentials that are used in parameterization may not necessarily be those one is interested in. For models which do not include effects due

to polarization, there is no change in the electronic charge density on hydrogen bonding, which is known to occur. Hence whenever one uses a classical potential, it should be benchmarked for the properties of interest. In many cases, a more refined model must be used for more sophisticated and elaborate experiments, for example, the vibrational and EA and CD spectra. In addition, since many models using parameters reduce the error between measured and simulated results, care must be taken to ensure that the parameters do not appear to correct the physics and chemistry that is *lacking* in the model. For example, to use parameters to “make” harmonic frequencies anharmonic. By doing so, other real anharmonic effects are lost.

Various models have been developed for implicitly and explicitly taking into account water at various levels (Onsager, 1936; Richmond, 1984; Wesson and Eisenberg, 1992; Schiffer *et al.*, 1992; Fortunelli and Tomasi, 1994; Schmidt and Fine, 1994; Tomasi and Persico, 1994; Fortunelli, 1996; Cossi *et al.*, 1996; Tawa *et al.*, 1996; Ösapay *et al.*, 1996; Craw *et al.*, 1996; Rauhut *et al.*, 1996; Mok, Nwumann and Handy, 1996; Jalkanen and Suhai, 1996, 1998; Jorgensen, Maxwell and Tirado-Rives, 1996; Tajkhorshid, Han *et al.*, 1998; Knapp-Mohammady *et al.*, 1999; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001, 2003, 2006a; Tomasi, Mennucci and Cammi, 2005; Mantz *et al.*, 2006). At the molecular mechanics level, the force field can be parameterized against experimental data measured on the molecule in the aqueous solution. The force field is then an effective force field in that it is not useful for doing calculations on the molecule in the gas phase or

for other solvents, the optimized potentials for liquid simulation (OPLS) force field being the most specific of this type (Jorgensen, Madura and Swenson, 1984; Jorgensen and Swenson, 1985a,b; Jorgensen and Tirado-Rives, 1988; Jorgensen, Maxwell and Tirado-Rives, 1996). Another approach to the solvent and/or hydration problem is to use a distance-dependent dielectric or some other perturbation to the gas phase potential energy surface and/or interactions to take into account the effect of water without actually adding explicit waters. The goal here is to save computational time and resources because the addition of explicit water molecules adds to the length of the calculation and the multiple minimum questions become exponentially worse when one tries to find the “global minimum” of the molecule solvated by water molecules. Finally the question of temperature, and hence entropic effects are very important when modelling hydrated systems, as there are many possible orientations of the water molecules encapsulating and/or hydrating the molecule.

The hydrated structures presented here for LA, NALANMA, LALA, LH, and NALHNMA can be used to test the various water models before one uses them in expensive molecular dynamic simulations on proteins and nucleic acids. The work is a part of our collaborative work involving groups at the German Cancer Research Center, the Technical University of Denmark, Helsinki University of Technology, Curtin University of Technology, and the University of Bremen to model proteins and nucleic acids along with various ligands in the presence of water and denaturing environments. In the next sections, we present the

methodology to simulate the vibrational spectrum, not only the frequencies but also the intensities, followed by the use of the spectroscopic methods in molecular biology, molecular biophysics, and quantum nanobiology.

1.3 Calculating Vibrational Absorption and Vibrational Circular Dichroism Spectra

1.3.1

Vibrational Absorption (VA) or Infrared (IR) Absorption and Vibrational Circular Dichroism

Vibrational absorption and CD spectra are related to molecular dipole and rotational strengths via,

$$\varepsilon(\bar{\nu}) = \frac{8\pi^3 N_A}{3000hc(2.303)} \sum_i \bar{\nu} D_i f_i(\bar{\nu}_i, \bar{\nu}) \quad (6)$$

$$\Delta\varepsilon(\bar{\nu}) = \frac{32\pi^3 N}{3000hc(2.303)} \sum_i \bar{\nu} R_i f_i(\bar{\nu}_i, \bar{\nu})$$

where ε and $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$ are molar extinction and differential extinction coefficients, respectively, D_i and R_i are the dipole and rotational strengths of the i th transition of wavenumbers $\bar{\nu}_i$ in cm^{-1} , and $f(\bar{\nu}_i, \bar{\nu})$ is a normalized line-shape function and N_A is Avogadro's number. For a fundamental ($0 \rightarrow 1$) transition involving the i th normal mode within the harmonic approximation

$$D_i = \left(\frac{\hbar}{2\omega_i} \right) \sum_{\beta} \left\{ \sum_{\lambda\alpha} S_{\lambda\alpha,i} P_{\alpha\beta}^{\lambda} \right\} \times \left\{ \sum_{\lambda'\alpha'} S_{\lambda'\alpha',i} P_{\alpha'\beta}^{\lambda'} \right\} \quad (7)$$

$$R_i = \hbar^2 \text{Im} \sum_{\beta} \left\{ \sum_{\lambda\alpha} S_{\lambda\alpha,i} P_{\alpha\beta}^{\lambda} \right\} \times \left\{ \sum_{\lambda'\alpha'} S_{\lambda'\alpha',i} M_{\alpha'\beta}^{\lambda'} \right\}$$

where $\hbar\omega_i$ is the energy of the i th normal mode, the $S_{\lambda\alpha,i}$ matrix interrelates normal coordinates Q_i to Cartesian displacement coordinates $X_{\lambda\alpha}$, where λ specifies a nucleus and $\alpha = x, y$, or z :

$$X_{\lambda\alpha} = \sum_i S_{\lambda\alpha,i} Q_i \quad (8)$$

$P_{\alpha\beta}^{\lambda}$ and $M_{\alpha\beta}^{\lambda}$ ($\alpha, \beta = x, y, z$) are the APT and AAT of nucleus λ . $P_{\alpha\beta}^{\lambda}$ is defined by

$$P_{\alpha\beta}^{\lambda} = \left\{ \frac{\partial}{\partial X_{\lambda\alpha}} \langle \psi_G(\mathbf{R}) | (\mu_{\text{el}})_{\beta} | \psi_G(\mathbf{R}) \rangle \right\}_{\mathbf{R}_o} \quad (9)$$

$$= 2 \left\langle \left(\frac{\partial \psi_G(\mathbf{R})}{\partial X_{\lambda\alpha}} \right)_{\mathbf{R}_o} \left| (\mu_{\text{el}}^e)_{\beta} \right| \psi_G(\mathbf{R}_o) \right\rangle + Z_{\lambda} e \delta_{\alpha\beta}$$

where $\psi_G(\mathbf{R})$ is the electronic wavefunction of the ground state G , \mathbf{R} specifies nuclear coordinates, \mathbf{R}_o specifies the equilibrium geometry, μ_{el} is the electric dipole moment operator, $\bar{\mu}_{\text{el}}^e = -e \sum_i \mathbf{r}_i$ is the electronic contribution to $\bar{\mu}_{\text{el}}$, and $Z_{\lambda} e$ is the charge on nucleus λ and $M_{\alpha\beta}^{\lambda}$ is given by

$$M_{\alpha\beta}^{\lambda} = I_{\alpha\beta}^{\lambda} + \frac{i}{4\hbar c} \sum_{\gamma} \varepsilon_{\alpha\beta\gamma} R_{\lambda\gamma}^o (Z_{\lambda} e)$$

$$I_{\alpha\beta}^{\lambda} = \left\langle \left(\frac{\partial \psi_G(\mathbf{R})}{\partial X_{\lambda\alpha}} \right)_{\mathbf{R}_e} \left| \left(\frac{\partial \psi_G(\mathbf{R}_e, B_{\beta})}{\partial B_{\beta}} \right)_{B_{\beta}=0} \right. \right\rangle \quad (10)$$

where $\psi_G(\mathbf{R}_o, B_{\beta})$ is the ground state electronic wavefunction in the equilibrium structure \mathbf{R}_e in the presence of the

perturbation $-(\mu_{\text{mag}}^e)_\beta B_\beta$, where μ_{mag}^e is the electronic contribution to the magnetic dipole moment operator. $M_{\alpha\beta}^\lambda$ is origin dependent. Its origin dependence is given by

$$(M_{\alpha\beta}^\lambda)^0 = (M_{\alpha\beta}^\lambda)^{O'} + \frac{i}{4\hbar c} \sum_{\gamma\delta} \varepsilon_{\beta\gamma\delta} Y_\gamma^\lambda P_{\delta\alpha}^\lambda \quad (11)$$

where \mathbf{Y}^λ is the vector from O to O' for the tensor of nucleus λ . Equation (11) permits alternative gauges in the calculation of the set of $(M_{\alpha\beta}^\lambda)^0$ tensors. If $\mathbf{Y}^\lambda = 0$, and hence $O = O'$, for all λ the gauge is termed the *common origin (CO) gauge*. If $\mathbf{Y}^\lambda = \mathbf{R}_{\lambda}^o$, so that in the calculation of $(M_{\alpha\beta}^\lambda)^0$ O' is placed at the equilibrium position of nucleus λ , the gauge is termed the *DO gauge* (Stephens, 1985, 1987; Buckingham, Fowler and Galwas, 1987; Amos *et al.*, 1987; Amos, Jalkanen and Stephens, 1988).

The problem with the expressions for the APT and AAT presented initially is that they involved the wavefunction and wavefunction derivatives (usually calculated with Coupled Hartree–Fock theory). These expressions were implemented in CADPAC to calculate the APT at the self-consistent field (SCF) and MP2 levels and the AAT at the SCF level. The first rigorous VA and VCD spectral simulations were performed using the Gaussian and CADPAC programs (Stephens and Lowe, 1985; Amos *et al.*, 1987; Amos, Jalkanen and Stephens, 1988, 1991). The scaled quantum mechanical (SQM) force fields (Lowe *et al.*, 1986; Jalkanen *et al.*, 1987, 1988; Kawiecki *et al.*, 1991) were initially scaled using the scaling methods of Blom and Altona (1976) and Pulay *et al.* (1979). Later more accurate MP2 force fields appeared to be accurate enough, when combined with MP2 APT

and RHF/SCF AAT tensors (Stephens *et al.*, 1993). This gave qualitative accuracy in general, and in some cases even quantitative accuracy, though there were cases which were problematic, one being *trans*-1,2-dicyanocyclopropane, which is now still considered problematic and a good test case for any new method (Jalkanen *et al.*, 2008b). Subsequently, the AAT have been implemented with GIAOs at both the SCF and multiconfiguration self-consistent field (MCSCF) levels of theory within the Dalton suite of programs (Bak *et al.*, 1993, 1994, 1995).

The advantages of DFT are numerous, but the main one is to extend rigorous methodology to the calculation of properties of large biological molecules. DFT seems to provide a way to do that. But the expressions must be reformulated as expressions amenable for DFT, where one has the density expressed in terms of the so-called Kohn–Sham orbitals and not in terms of rigorous wave functions. The Kohn–Sham orbitals are only used to generate the density, and one must be very careful in applying and using the orbitals and eigenvalues outside the scope for which they were intended.

The calculation of the DFT AATs has been implemented within the coupled perturbed Kohn–Sham formalism using GTOs in CADPAC, Dalton, and Gaussian (Cheesman *et al.*, 1996) and using STOs in the ADF program (Nicu *et al.*, 2008; Wolff, 2005; te Velde *et al.*, 2001). For many cases, SCF AAT appear to be of sufficient accuracy, but with the prevalence of implementations of DFT, DFT methodology appears to be taking over. A comparison of rotational strengths for molecules with both RHF and DFT AAT is still important, as in many cases one can use a different method for the tensors required for the VA and VCD (and

also Raman and ROA) intensities, than one requires for the geometry optimization and subsequent Hessian and normal mode calculation. The implementation of VA and VCD spectral simulation in Gaussian, Dalton, CADPAC, and ADF all allow one to accurately predict the VA and VCD spectra for molecules in nonpolar solvents, in inert low temperature noble gas and N_2 matrices, and in membranes where there is no strong interaction between the solute molecule and the solvent/environment.

Recently, various continuum models have been implemented in Gaussian, Dalton, and ADF that allow not only geometry optimizations but also analytical derivative methods for the calculation of the Hessian, APT and AAT, required to simulate the VA and VCD spectra. The continuum solvent models have required continuous refinements and extensions as one has started to use them to treat biological systems, where many of the assumptions inherent in the development of the models break down. One of these is that there are no specific directional effects which must be treated. Hydrogen bonding effects have been shown to be inadequately treated, and hybrid models have been developed to treat these systems (Jalkanen and Suhai, 1996, 1998; Jalkanen, Bohr and Suhai, 1997; Han *et al.*, 1998; Tajkhorshid, Jalkanen and Suhai, 1998; Knapp-Mohammady *et al.*, 1999; Jalkanen *et al.*, 2008a). Especially when one attempts to treat hydrogen-bonded systems with implicit solvent models one must be very careful. One cannot treat an explicit directional and specific interaction between the charged NH_3^+ and CO_2^- groups in an amino acid, peptide and protein with

solvent/water molecules, cations, or anions with an implicit continuum solvent model. Here we have introduced our embedded molecular (hydrogen-bonded) complex model (Jalkanen and Suhai, 1996; Jalkanen, Bohr and Suhai, 1997; Han *et al.*, 1998; Tajkhorshid, Jalkanen and Suhai, 1998; Knapp-Mohammady *et al.*, 1999; Jalkanen *et al.*, 2006a, 2008a; Jürgensen and Jalkanen, 2006). Here one needs to identify the important solvent molecules, cations, and anions which are (thought to be) important in stabilizing the structure/conformation/functional state of the molecule/molecular complex/assembly/aggregate. This complex/assembly/aggregate can be embedded within a solvent cavity and/or protein active site and/or micelle, where the cavity surface is treated implicitly by using a lower level of theory (perhaps using a molecular mechanics force field). A complete review of these methods is beyond the scope of this work on the use of vibrational spectroscopic methods in quantum nanobiology, molecular biology, and molecular biophysics.

1.4 Calculating Raman Scattering and Raman Optical Activity Spectra

1.4.1

Raman Scattering

If one wishes to simulate the Raman scattering intensities, one is required to calculate the electric dipole–electric dipole polarizability derivatives. Here one can use either the static values, as was done in most early works, or one can take into account the frequency response, as is now routine. In most conventional Raman scattering

experiments one uses either 488, 532, or 1064 nm source, the later for molecules for which fluorescence is a problem.

The Raman intensity is proportional to the Raman scattering activity, which is given for the j th normal mode Q_j by

$$I_j^{\text{Ram}} = g_j(45\bar{\alpha}_j^2 + 7\bar{\beta}_j^2) \quad (12)$$

where g_j is the degeneracy of the j th transition, and the electric dipole–electric dipole polarizability tensor derivative invariants $\bar{\alpha}_j^2$ and $\bar{\beta}_j^2$ are given by

$$\begin{aligned} \bar{\alpha}_j^2 &= \frac{1}{9} \left(S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} + S_{\lambda\alpha,j}\alpha_{yy}^{\lambda\alpha} + S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha} \right)^2 \\ & \quad (13) \end{aligned}$$

and

$$\begin{aligned} \bar{\beta}_j^2 &= \frac{1}{9} \left\{ (S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{yy}^{\lambda\alpha})^2 \right. \\ & \quad + (S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha})^2 \\ & \quad + (S_{\lambda\alpha,j}\alpha_{yy}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha})^2 \\ & \quad + 6[(S_{\lambda\alpha,j}\alpha_{xy}^{\lambda\alpha})^2 + (S_{\lambda\alpha,j}\alpha_{yz}^{\lambda\alpha})^2 \\ & \quad \left. + (S_{\lambda\alpha,j}\alpha_{xz}^{\lambda\alpha})^2 \right\} \quad (14) \end{aligned}$$

respectively. The $S_{\lambda\alpha,j}$ matrix relates the normal coordinates Q_j to the Cartesian displacement coordinates $X_{\lambda\alpha}$, where λ specifies a nucleus and $\alpha = x, y, \text{ or } z$ (Diem, 1993): as previously defined. In addition to normal and Cartesian coordinates there are symmetry coordinates (Cotton, 1990). The advantage of using symmetry coordinates is that the Hessian will block diagonalize, that is, the Hessian will be reduced to independent sum matrices. Before the advent of supercomputers, symmetry coordinates were used for all symmetrical molecules. They are still used as in the IR and Raman modes

for molecules with a center of inversion are mutually exclusive. That is, modes which are active in the Raman and inactive in the IR and vice versa. Here one can make use of character tables (Cotton, 1990). In addition, some modes are not seen in either the Raman or IR, and here the use of IINS can help. The selection rules for IR, Raman and IINS are all different, so the three experiments are complementary and all necessary if one wishes to get complete data for all of the vibrational modes.

The normal frequencies are the eigenvalues Λ of the Hessian matrix H (the second derivative of the energy with respect to nuclear displacements, evaluated at the equilibrium geometry):

$$H_{\lambda\alpha,\lambda'\alpha'} = \left(\frac{\partial^2 E(\mathbf{R})}{\partial X_{\lambda\alpha} \partial X_{\lambda'\alpha'}} \right)_{\mathbf{R}=\mathbf{R}_0} \quad (15)$$

and

$$C^{-1}HC = \Lambda \quad (16)$$

C is the eigenvector matrix, which defines the transformation matrix S which is given by

$$S = M^{-1/2}C \quad (17)$$

where M is the mass matrix.

Therefore, through the matrix S (obtained by diagonalizing the Hessian matrix) and the polarizability derivatives (Electric Dipole–Electric Dipole Polarizability Derivatives EDEDPD), $\alpha_{\beta\gamma}^{\lambda\alpha}$, one can obtain the Raman scattering activity I_j^{Ram} of any normal mode Q_j . The first efficient method for determining the EDEDPD was implemented utilizing finite field perturbation theory by Komornicki and McIver (1979). Subsequently the analytical EDEDPD were derived and implemented in CADPAC (Amos, 1986,

1987) and Gaussian (Frish *et al.*, 1986) at the Hartree–Fock level of theory, and subsequently at the DFT level. The ED-EDPD have also been implemented at the DFT level and used to simulate the Raman spectra without any solvent treatment (Johnson and Florian, 1995; Stirling, 1996; Collier, Magdo and Klots, 1999; Halls and Schlegel, 1999a; van Caillie and Amos, 2000; Jalkanen *et al.*, 2006b) and subsequently by treating the solvent with explicit solvent molecules, with continuum solvent models, and finally using hybrid models which combine the two (Jürgensen and Jalkanen, 2006; Jalkanen *et al.*, 2006a, 2008a; Deplazes *et al.*, 2008). In contrast to solvent subtraction which is common in IR/VA spectroscopy, this is not always done with Raman measurements. Investigations of the performance of local, gradient-corrected, and hybrid density functional models in predicting Raman intensities have shown that the hybrid functionals are very accurate when the solvent effects are not important (Halls and Schlegel, 1999b; Jalkanen *et al.*, 2006b).

Hence the question of whether and how to perform a solvent subtraction for Raman measurements is a question that we and others have started to address (Jalkanen *et al.*, 2001, 2003, 2008a; Han *et al.*, 1998; Tajkhorshid, Jalkanen and Suhai, 1998).

1.4.2

Raman Optical Activity

Finally if one wishes to simulate the ROA intensities, one is required to calculate the electric dipole–magnetic dipole polarizability derivatives (ED-MDPD) and the electric dipole–electric quadrupole polarizability derivative (ED-EQPD) (Buckingham, 1967; Atkins and

Barron, 1969; Barron and Buckingham, 1971, 2001; Fischer and Hache, 2005). The calculation of the EDMDP, G' , have been implemented in CADPAC (Amos, 1982). The EDMDP can be evaluated as the second derivative of the energy with respect to a static electric field and a time varying magnetic field. Hence the EDMDPD is a third derivative with respect to the nuclear displacement.

$$G'_{\beta\gamma}{}^{\lambda\alpha} = \left(\frac{\partial^3 W_G(\mathbf{R}, F_\beta, \dot{B}_\gamma)}{\partial X_{\lambda\alpha} \partial F_\beta \partial \dot{B}_\gamma} \right)_{\substack{\mathbf{R}=\mathbf{R}_0, F_\beta=0, \\ \dot{B}_\gamma=0}} = \left(\frac{\partial G'_{\beta\gamma}(\mathbf{R})}{\partial X_{\lambda\alpha}} \right)_{\mathbf{R}=\mathbf{R}_0} \quad (18)$$

The electric dipole–electric dipole, electric dipole–magnetic dipole, and electric dipole–electric quadrupole polarizability tensors can also be calculated at the frequency of the incident light using SCF linear response theory. London atomic orbitals have been employed, hence the results are gauge invariant (Helgaker *et al.*, 1994). In this work, the electric dipole–magnetic dipole orbitals are gauge independent when using a finite basis set due to the use of gauge-invariant atomic orbitals as implemented in Gaussian 03. Previously we used conventional basis sets which were not gauge invariant, similar to the work of Polavarapu and coworkers (Polavarapu, 1990; Barron *et al.*, 1991a,b; Deng *et al.*, 1996) and our previous work on NALANMA (Han *et al.*, 1998), where that level of theory had been shown to give predicted ROA spectra in good agreement with the experimentally observed spectra. We are thus confident that the new level of theory can be used to answer the structural questions posed here, realizing that the agreement in absolute intensities could be improved by using much larger basis sets, as we

shall do for the NALANMA4WC, but not γ_j^2 is given by the LA20WC.

$$\begin{aligned} \gamma_j^2 = & \frac{1}{2} \left\{ (S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{\gamma\gamma}^{\lambda\alpha})(S_{\lambda\alpha,j}G_{xx}^{\prime\lambda\alpha} - S_{\lambda\alpha,j}G_{\gamma\gamma}^{\prime\lambda\alpha}) + (S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha}) \right. \\ & \times (S_{\lambda\alpha,j}G_{xx}^{\prime\lambda\alpha} - S_{\lambda\alpha,j}G_{zz}^{\prime\lambda\alpha}) + (S_{\lambda\alpha,j}\alpha_{\gamma\gamma}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha})(S_{\lambda\alpha,j}G_{\gamma\gamma}^{\prime\lambda\alpha} - S_{\lambda\alpha,j}G_{zz}^{\prime\lambda\alpha}) \\ & + 3 \left[(S_{\lambda\alpha,j}\alpha_{xy}^{\lambda\alpha})(S_{\lambda\alpha,j}G_{xy}^{\prime\lambda\alpha} + S_{\lambda\alpha,j}G_{yx}^{\prime\lambda\alpha}) + (S_{\lambda\alpha,j}\alpha_{yz}^{\lambda\alpha})(S_{\lambda\alpha,j}G_{yz}^{\prime\lambda\alpha} + S_{\lambda\alpha,j}G_{zy}^{\prime\lambda\alpha}) \right. \\ & \left. \left. + (S_{\lambda\alpha,j}\alpha_{xz}^{\lambda\alpha})(S_{\lambda\alpha,j}G_{xz}^{\prime\lambda\alpha} + S_{\lambda\alpha,j}G_{zx}^{\prime\lambda\alpha}) \right] \right\} \quad (21) \end{aligned}$$

The EDEQP, A , can be evaluated as the second derivative of the energy with respect to a static electric field and a static electric field gradient. Hence the EDEQPD is a third derivative, the third derivative being with respect to the nuclear displacement.

$$\begin{aligned} A_{\beta,\gamma\delta}^{\lambda\alpha} &= \left(\frac{\partial^3 W_G(\mathbf{R}, F_\beta, F'_{\gamma\delta})}{\partial X_{\lambda\alpha} \partial F_\beta \partial F'_{\gamma\delta}} \right)_{\substack{\mathbf{R}=\mathbf{R}_0, F_\beta=0, \\ F'_{\gamma\delta}=0}} \\ &= \left(\frac{\partial A_{\beta,\gamma\delta}(\mathbf{R})}{\partial X_{\lambda\alpha}} \right)_{\mathbf{R}=\mathbf{R}_0} \quad (19) \end{aligned}$$

The Cartesian polarizability derivatives are required to calculate the ROA spectra. The quantities required have been derived by Barron and Buckingham (1971) and are $\bar{\alpha}_j \bar{G}'_j$, γ_j^2 , and δ_j^2 (Barron, 1982, 2004; Buckingham, 1967; Barron, Bogaard and Buckingham, 1973; Hecht and Barron, 1990). The quantities are calculated by combining the various polarizability derivatives with the normal mode vectors in the following equations.

$\bar{\alpha}_j \bar{G}'_j$ is given by

$$\begin{aligned} \bar{\alpha}_j \bar{G}'_j &= \frac{1}{9} \left(S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} + S_{\lambda\alpha,j}\alpha_{\gamma\gamma}^{\lambda\alpha} + S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha} \right) \\ &\times \left(S_{\lambda\alpha,j}G_{xx}^{\prime\lambda\alpha} + S_{\lambda\alpha,j}G_{\gamma\gamma}^{\prime\lambda\alpha} + S_{\lambda\alpha,j}G_{zz}^{\prime\lambda\alpha} \right) \quad (20) \end{aligned}$$

and δ_j^2 is given by

$$\begin{aligned} \delta_j^2 = & \frac{\omega}{2} \left\{ (S_{\lambda\alpha,j}\alpha_{\gamma\gamma}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha})S_{\lambda\alpha,j}A_{z,xy}^{\lambda\alpha} \right. \\ & + (S_{\lambda\alpha,j}\alpha_{xx}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha})S_{\lambda\alpha,j}A_{\gamma,zx}^{\lambda\alpha} \\ & + (S_{\lambda\alpha,j}\alpha_{zz}^{\lambda\alpha} - S_{\lambda\alpha,j}\alpha_{\gamma\gamma}^{\lambda\alpha})S_{\lambda\alpha,j}A_{x,\gamma z}^{\lambda\alpha} \\ & + S_{\lambda\alpha,j}\alpha_{xy}^{\lambda\alpha}(S_{\lambda\alpha,j}A_{\gamma,\gamma z}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{z,\gamma\gamma}^{\lambda\alpha}) \\ & + S_{\lambda\alpha,j}A_{z,xx}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{x,zz}^{\lambda\alpha} \\ & + S_{\lambda\alpha,j}\alpha_{xz}^{\lambda\alpha}(S_{\lambda\alpha,j}A_{\gamma,zz}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{z,\gamma\gamma}^{\lambda\alpha}) \\ & + S_{\lambda\alpha,j}A_{x,\gamma\gamma}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{\gamma,xx}^{\lambda\alpha} \\ & + S_{\lambda\alpha,j}\alpha_{\gamma z}^{\lambda\alpha}(S_{\lambda\alpha,j}A_{z,zx}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{x,zz}^{\lambda\alpha}) \\ & \left. + S_{\lambda\alpha,j}A_{x,\gamma\gamma}^{\lambda\alpha} - S_{\lambda\alpha,j}A_{\gamma,\gamma x}^{\lambda\alpha} \right\} \quad (22) \end{aligned}$$

The equations relating these tensor derivatives to the ROA spectra are given by Barron (1982, 2004); Hecht, Barron and Hug (1989), and Barron *et al.* (1994). These references provide a good introduction to the theory and application of ROA spectroscopy. Here we shall give the relevant equations for completeness. We follow very closely the notation of Barron from his Faraday Discussions review article (Barron *et al.*, 1994).

The commonly reported quantity from ROA measurements is the dimensionless circular intensity differential (CID) defined as

$$\Delta_{\alpha} = \left(\frac{I_{\alpha}^R - I_{\alpha}^L}{I_{\alpha}^R + I_{\alpha}^L} \right) \quad (23)$$

where I_{α}^R and I_{α}^L are the scattered intensities with linear α polarization in right- and left-circularly polarized incident light. In terms of the quantities, EDEDPD, EDMDPD and EDEQPD, the CIDs for forward, backward, and polarized and depolarized right-angle scattering from an isotropic sample for incident laser light are (Barron *et al.*, 1994):

$$\begin{aligned} \Delta(\text{forward}) &= \left(\frac{8[45\alpha G' + \beta(G')^2 - \beta(A)^2]}{2c[45\alpha^2 + 7\beta(\alpha)^2]} \right) \end{aligned} \quad (24)$$

$$\begin{aligned} \Delta(\text{backward})_{\text{ICP}} &= \left(\frac{48[\beta(G')^2 + (1/3)\beta(A)^2]}{2c[45\alpha^2 + 7\beta(\alpha)^2]} \right) \end{aligned} \quad (25)$$

$$\begin{aligned} \Delta(90^{\circ}, \text{polarized}) &= \left(\frac{2[45\alpha G' + 7\beta(G')^2 + \beta(A)^2]}{c[45\alpha^2 + 7\beta(\alpha)^2]} \right) \end{aligned} \quad (26)$$

$$\begin{aligned} \Delta(90^{\circ}, \text{depolarized}) &= \left(\frac{12[\beta(G')^2 - (1/3)\beta(A)^2]}{6c\beta(\alpha)^2} \right) \end{aligned} \quad (27)$$

$$\begin{aligned} \Delta(\text{backward})_{\text{DCP}_1} &= \left(\frac{48[\beta(G')^2 + (1/3)\beta(A)^2]}{2c[6\beta(\alpha)^2]} \right) \end{aligned} \quad (28)$$

where c is the speed of light and the isotropic invariants are defined as

$$\alpha = (1/3)\alpha_{\alpha\alpha} \quad (29)$$

and

$$G' = (1/3)G'_{\alpha\alpha} \quad (30)$$

and the anisotropic invariants as

$$\beta(\alpha)^2 = (1/2)(3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}) \quad (31)$$

$$\beta(G')^2 = (1/2)(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}) \quad (32)$$

$$\beta(G')^2 = (1/2)(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}) \quad (33)$$

$$\beta(A)^2 = (1/2)\omega\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta} \quad (34)$$

Using these expressions, we can calculate the ROA spectra within the harmonic approximation for the vibrational frequencies and with the static field limit for the EDEDPD, EDMDPD, and EDEQPD. In addition, one can take into account the frequency dependence of the Raman and ROA intensities. A new formalism for the Raman and ROA intensities in the near resonance regime has recently been presented by Nafie, though not yet implemented (Nafie, 2008). Finally the analytical derivatives of the G' and A tensors have recently been reported by Cheeseman and Frisch (2008); Ruud (2008); Liegeois and Champagne (2008); Liegeois (2008). Hence the efficiency of calculation of Raman and ROA spectra is now the same as that for VA and VCD spectra. The availability of commercial Raman and ROA and VCD and VA spectrometers and programs to simulate the spectra has been a major breakthrough in the use of these spectroscopies in quantum nanobiology, molecular biology, and molecular biophysics. In the last section, we now present some interesting and pioneering work on the use of VA, VCD, Raman, and ROA spectroscopy in quantum nanobiology, molecular biology, and molecular biophysics.

1.5 Classifying Conformational States and Structures based on VA, VCD, Raman, and ROA Spectra

1.5.1 Experimental Vibrational Frequencies for N-Methylacetamide, NALANMA, Peptides, and Proteins

The amide modes have been determined experimentally from the simple amides formamide, acetamide, and NMA. NMA has been used as a model since its two methyl groups can be assumed to represent the two C_{α} carbons in proteins. The amide A mode is mainly a N–H stretch, amide I is mainly a C=O stretch, amide II a N–H bend coupled with C–N stretching, amide III N-bending and C–N stretching (only in simple amides as it has been shown to couple with the CH deformations on the adjacent C_{α} carbons with C–HR groups [nonglycine] (Diem, 1993; Jalkanen *et al.*, 2003), amide IV OCN bending, amide V in-plane N–H bending, amide VI out-of-plane N–H bending, and finally amide VII skeletal torsion (Jung, 2000; Jalkanen and Suhai, 1996). In addition to the amide modes for NMA being identified from both experimental and theoretical studies, these have also been identified in a number of peptides and proteins and the amide frequencies have been correlated with secondary structural elements (Jung, 2000). The mode most often correlated with secondary structure is the amide I mode (called *amide I'* when measured in D_2O). The amide I frequencies need to be measured in D_2O because the amide I mode overlaps with the H_2O bending modes. Hence running the spectra in D_2O when one wants to determine the C=O stretching frequency.

Alternatively, one can make up a very concentrated solution and measure the spectra for very thin pathlengths.

In addition, by performing ^{13}C isotopic substitution for the C' carbon in specific residues, one can shift the frequency of the amide I band for the specific residue. To date, the IR spectra of proteins and their corresponding amide I spectrum have been used to only determine the percent of a specific secondary structural element present, and one has not been able to determine the specific backbone angles for the specific residue and the corresponding side chain angles. This information for each residue would uniquely identify the structure of the protein and is the goal of the combined use of VA, VCD, Raman, and ROA spectroscopy, as has been achieved for both X-ray and neutron diffraction and NMR spectroscopies. In the next section, we review our and other groups theoretical vibrational spectra simulations for amino acids and peptides and present our new work on LALA and NALHNMA. The complete work on the later two molecules will be published in a special issue of Theoretical Chemical Accounts which focuses on molecular biophysics and bioinformatics, the Suhai Festschrift, which will appear in 2009. Hence for more details on the work on LALA and NALHNMA refer to the manuscripts in this special issue (Jalkanen *et al.*, 2009a,b).

1.5.2 Results for the DFT Calculations

1.5.2.1 L-Alanine

In Table 1.1, we present results for the LA zwitterionic species, which is the dominant species found in aqueous solution at neutral pH (Jalkanen, Bohr

Tab. 1.1 L-Alanine zwitterion structures.

Coordinate	I ^a	I ^a	I ^a	II ^a	III ^b	IV ^c	V ^d
	6-31G*	6-31G**	++G(2d,2p)	6-31G*	6-31G*	6-31G*	6-31G*
r(N1-C2)	1.5115	1.5105	1.5129	1.5182	1.5142	1.5123	1.481
r(C2-C3)	1.5639	1.5638	1.5593	1.5673	1.5545	1.5516	1.534
r(C3-O4)	1.2893	1.2899	1.2854	1.2748	1.2615	1.2647	1.277
r(C3-O5)	1.2299	1.2297	1.2211	1.2421	1.2608	1.2601	1.247
r(C2-C6)	1.5267	1.5263	1.5229	1.5202	1.5227	1.5220	1.526
r(C2-H7)	1.0947	1.0945	1.0889	1.0940	1.0923	1.0925	1.096
r(N1-H8)	1.0466	1.0460	1.0344	1.0565	1.0536	1.0484	1.054
r(N1-H9)	1.0477	1.0472	1.0346	1.0559	1.0454	1.0464	1.057
r(N1-H10)	1.0340	1.0322	1.0250	1.0334	1.0529	1.0583	1.057
r(C6-H11)	1.0960	1.0950	1.0907	1.0940	1.0933	1.0938	1.094
r(C6-H12)	1.0916	1.0905	1.0866	1.0932	1.0942	1.0940	1.091
r(C6-H13)	1.0966	1.0954	1.0913	1.0937	1.0960	1.0955	1.093
θ(N1-C2-C3)	111.19	111.28	110.94	105.04	106.82	106.41	110.55
θ(C2-N3-O4)	116.14	116.18	115.80	117.31	115.31	115.83	119.06
θ(C2-N3-O5)	115.09	115.11	115.72	114.41	117.93	118.10	115.68
θ(N1-C2-C6)	109.96	110.05	109.95	110.18	109.95	109.85	111.05
θ(C3-C2-C6)	111.60	111.52	112.10	114.17	114.76	114.57	110.71
θ(N1-C2-H7)	106.19	106.32	106.13	106.06	105.79	105.78	106.84
θ(C3-C2-H7)	107.47	107.25	107.21	109.48	107.73	108.67	107.94
θ(C2-N1-H8)	109.05	108.83	109.82	108.64	112.36	111.69	111.54
θ(C2-N1-H9)	109.83	109.65	110.22	109.31	108.93	108.99	108.38
θ(H8-N1-H9)	106.25	106.27	105.95	106.22	109.40	109.14	110.87
θ(C2-N1-H10)	110.63	110.44	111.30	111.72	110.97	110.98	109.87
θ(H8-N1-H10)	110.27	110.54	109.88	110.57	109.60	110.36	107.78
θ(C2-C6-H11)	110.70	110.62	110.89	110.52	110.93	110.77	111.21
θ(C2-C6-H12)	108.27	108.21	108.42	109.44	109.85	109.84	109.80
θ(H11-C6-H12)	108.28	108.28	108.21	107.20	108.10	108.01	108.29
θ(C2-C6-H13)	111.52	111.52	111.56	111.36	110.91	110.87	110.88
θ(H11-C6-H13)	108.81	108.84	108.91	109.69	108.91	109.22	108.46
τ(N1-C2-C3-O4)	-0.53	0.08	-3.03	-101.08	94.86	86.84	-14.21
τ(N1-C2-C3-O5)	-179.57	-179.83	176.59	73.20	-81.69	-89.60	166.29
τ(C3-C2-N1-H8)	59.20	58.85	60.96	48.12	80.56	72.44	64.06
τ(C3-C2-N1-H9)	-56.85	-57.00	-55.40	-67.39	-40.83	-38.23	-58.27
τ(C3-C2-N1-H10)	-179.36	-179.64	-177.12	170.36	-156.34	-163.96	-176.50
τ(H11-C6-C2-H7)	177.17	177.27	177.02	-176.93	178.19	177.51	-175.98
τ(H12-C6-C2-H7)	-64.29	-64.26	-64.32	-59.09	-62.36	-63.27	-56.14
τ(H13-C6-C2-H7)	55.87	56.00	55.44	60.89	57.04	56.08	63.25

^aZwitterion stabilized by four water molecules.^bZwitterion stabilized by nine water molecules.^cZwitterion stabilized by nine water molecules + solvent model.^dZwitterion stabilized by 20 water molecules + CPCM model, B3PW91.

and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001). At the RHF level of theory the zwitterion was found to be stable (Barron *et al.*, 1991a), but one of the N-H bond lengths was found to be much longer than the other two. This resulted in the vibrational stretch frequency associated with N-H to be significantly lower in energy

than that for the other two N–H bonds. If the three NH bonds were all the same, the three modes would be thought to be approximately the same and vibrational frequencies would be determined by both the local symmetry (thought to be C_{3v}) and the perturbation due to the local environment. Herein lies the additional information which can in theory be extracted from the experimental VA, VCD, Raman, and ROA spectra. Subsequently Yu and coworkers tried to improve the model for the LA zwitterion by embedding the zwitterion with a continuum solvent treatment, using the Onsager continuum model (Onsager, 1936), optimizing the structure and subsequently simulating the Raman and ROA spectra (Yu *et al.*, 1995). At the B3LYP DFT level of theory, we found that the zwitterionic structure was not stable in the isolated state (Jalkanen, Bohr and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998).

We present three 6-31G* B3LYP zwitterionic structures for LA, two in the presence of four water molecules (Tajkhorshid, Jalkanen and Suhai, 1998) and one in the presence of nine water molecules (Frimand *et al.*, 2000). We also present the 6-31G** and 6-311++G(2d,2p) optimized geometries for one of the zwitterionic structures. As one can see from the basis set effect, as one adds more polarization functions and diffuse functions, the bond length changes are usually less than 0.004 Å. The valence angle changes are usually less than 0.5°. Finally we also present the LA zwitterion completely solvated by 20 water molecules and embedded with the PCM continuum solvent model (Jalkanen *et al.*, 2008).

Initially we solvated the zwitterion by adding water molecules to hydrogen bond with each of the polar groups,

and subsequently extracted the positions from classical MD simulations (Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001). In addition to the normal isotopomers of LA, we have additionally looked at the various deuteriated derivatives theoretically and compared to the experimental work of Diem *et al.* (1982); Abdali *et al.* (2002). We have recently determined the positions of the hydrogens and completely solvated the LA zwitterion with BO MD simulations (Degtyarenko *et al.*, 2007) and subsequently have simulated the VA, VCD, Raman, and ROA spectra for the completely solvated zwitterion and embedded within the COSMO and PCM continuum models (Jalkanen *et al.*, 2008a). Finally we have compared the zwitterionic structure derived from molecular dynamic simulations with that found in the crystal (Degtyarenko *et al.*, 2008). Hence the zwitterionic structure of LA has been thoroughly studied. The experimental VA, VCD, Raman, and ROA spectra have been able to be reproduced by treating the first solvation shell of water molecules explicitly and embedding the solvated zwitterionic structure within the continuum solvent models so as to treat the effects due to bulk water molecules. In Figure 1.1, we present the zwitterionic structure of the LAZ hydrated by 20 water molecules and this hydrated complex embedded within the continuum solvent model. As one can see in Figure 1.1 the 4 water molecules in the original LAZ plus 4 water molecules for the F' structure are very similar to 4 of the water molecules in the LAZ plus 20 water molecules. This can explain why the LAZ plus four water molecule plus Onsager model did quite well in reproducing the VA, VCD, Raman, and ROA spectra of the LAZ

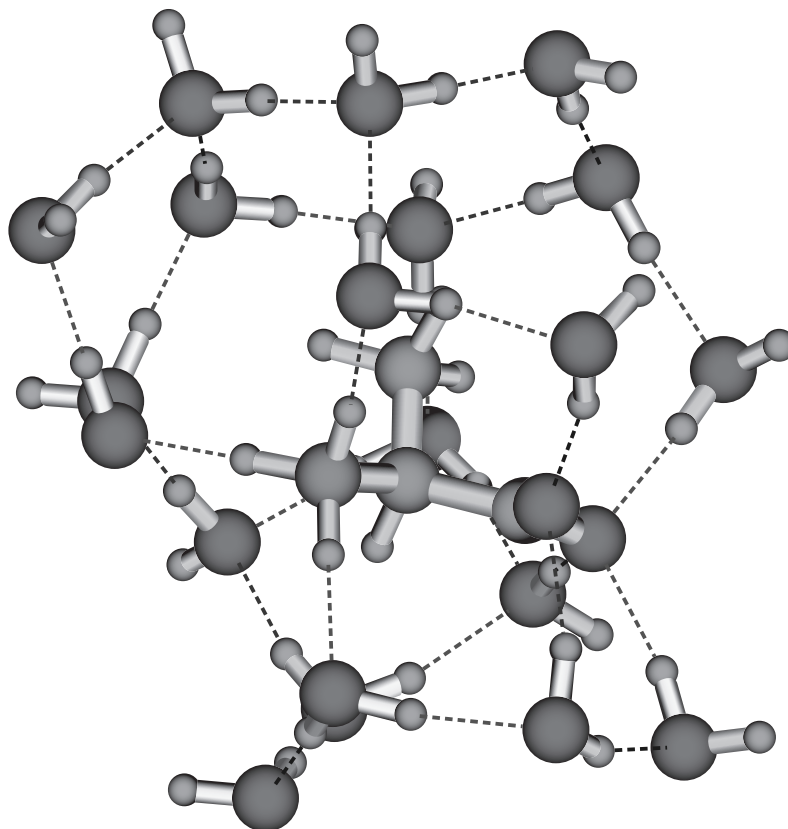


Fig. 1.1 L-Alanine zwitterion stabilized by 20 water molecules. B3PW91/6-31G* plus CPCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

(Tajkhorshid, Jalkanen and Suhai, 1998; Jalkanen *et al.*, 2001). Previously we have presented the experimental VA, VCD, Raman, and ROA spectra and the corresponding simulated VA, VCD, Raman, and ROA spectra (Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*, 2001, 2008a). We showed that the LAZ plus 20 water molecule model gives better agreement with the experimental spectra, but that our minimal model is a good starting point. Our minimal model being the LAZ plus four water molecule complex embedded in one of the continuum solvent models

(Onsager, PCM, CPCM, or Cosmo) (Tomasi, Mennucci and Cammi, 2005).

When one measures the VA, VCD, Raman, and/or ROA spectra, one has the problem of determining the spectra and spectra of the solvent. In many cases one assumes that one can subtract the spectra of the bulk solvent to get the spectra due to the isolated solute. This only applies when the solvent is inert and does not strongly interact with the solute. This is clearly not the case for amino acids and water. As stated above, the water molecules are necessary to stabilize the zwitterionic species. The relative energy

of the neutral nonionic species is lower than that of the zwitterionic species, which is not even a stable minima. So one must rethink about how one could in principle subtract out the spectra of the so-called solvent, when the solute and solvent strongly interact. Of course, not all of the solvent molecules strongly interact via hydrogen bonding with the LA zwitterion. Only the first solvation shell waters are thought to be important. In our preliminary works, we included only four and then nine water molecules, but in our last work we included a complete solvation shell, which included 20 water molecules. In addition to the structure of the LA molecule being perturbed by the aqueous environment, the individual water molecules which are hydrogen bonded with the LA zwitterion are perturbed from both isolated water molecules and water molecules in bulk water. In the case of methyl lactate, water molecule(s) bound to polar groups have been shown to possess a VCD signature (Losada and Xu, 2007). Hence on the timescale of the VCD measurement, the water molecules strongly bound to the polar chiral molecule show chirality. Hence chirality can be due to weak hydrogen bonding, in addition to strong covalent bonding. With an increase in temperature the VCD spectra will go away, whereas a drop in temperature will make other water molecules less mobile and hence increase the VCD and also the ROA spectra. Hence by cooling down the aqueous solution of biomolecules, one should be able to learn more about the character of the hydration shells of biomolecules. From these temperature-dependent spectra one should be able to identify various water molecules in the first solvation shell, the strength of their binding, and their exchange rates and

mobilities. This should also be the case for ligands bound in the binding pockets (active sites) of proteins.

In Tables 1.2a and 1.2b we present the calculated frequencies and VA and VCD intensities for the zwitterionic structures of L-alanine-d₀. We also give the experimental results reported by Diem and coworkers. In Table 1.2a we also give the calculated vibrational frequencies and VA intensities recently reported by Barron and coworkers. For a complete review of our works on LA, see our latest publication in the Stephens Festschrift in 2008 (Jalkanen *et al.*, 2008a).

1.5.2.2 N-Acetyl L-Alanine N'-Methylamide

In addition to the individual amino acids, one is interested in the amino acid residue as they appear in peptides and proteins. To model the amino acid in a protein, we have capped the LA amino acid with an N-acetyl group on the N-terminus and a N'-methylamide on the C-terminus. Some researchers have capped the N-terminus with the N-formyl group and the C-terminus with the simple amide group. In our work the methyl groups approximate the C_α carbon atoms in the adjacent residues, whereas in the simpler model, hydrogen atoms approximate the C_α carbons. With the speed of computers there is no longer a need to use the simpler model and hence we model our amino acid residues, LA/LH, with the NALANMA and NALHNMA molecules.

In Table 1.3, we present the relative energies of isolated NALANMA and with four bound water molecules. The values of ϕ and ψ (measures of secondary structure in proteins) are also given. The starting structures for the bound water optimizations were the 6-31G* B3LYP

Tab. 1.2a L-Alanine- d_0 zwitterion vibrational frequencies and VA intensities.

$\nu(\text{cm}^{-1})^a$	$\nu(\text{cm}^{-1})^b$ I	A_i^c	$\nu(\text{cm}^{-1})^b$ II	A_i^c	$\nu(\text{cm}^{-1})^b$ III	A_i^c	$\nu(\text{cm}^{-1})^d$	A_i^d
3080	3317	529.5	3327	497.5	3153	14.8	3809	65.3
3060	3170	3.6	3155	12.0	3132	32.2	3727	53.9
3020	3120	15.9	3132	19.4	3124	11.6	3333	3.7
3003	3105	336.6	3099	10.8	3114	96.1	3295	20.7
2993	3088	327.4	3057	29.0	3063	36.6	3258	32.2
2962	3077	502.0	2956	683.3	3009	63.9	3204	31.9
2949	3053	33.9	2940	691.6	2991	72.3	2813	581.0
1645	1770	324.3	1777	77.9	1775	60.3	1996	442.0
1625	1745	82.9	1745	181.4	1749	37.8	1822	42.1
1607	1741	56.8	1719	221.6	1696	544.0	1790	24.6
1498	1634	91.1	1647	78.2	1662	185.0	1641	11.9
1459	1534	2.2	1525	4.7	1526	21.5	1639	3.9
1459	1525	5.2	1513	6.5	1522	3.1	1568	47.0
1410	1433	41.3	1444	44.1	1454	92.1	1523	388.0
1375	1407	7.0	1415	5.8	1438	98.7	1501	69.9
1351	1377	34.6	1405	63.9	1428	8.9	1447	143.0
1301	1333	186.1	1342	157.5	1370	111.5	1395	238.0
1220	1281	37.5	1297	14.2	1298	22.9	1314	18.0
1145	1200	56.0	1199	20.2	1238	36.6	1205	41.8
1110	1124	16.9	1115	13.1	1121	16.9	1164	38.7
1001	1058	7.9	1075	7.0	1076	5.2	1080	72.9
995	1026	7.6	1023	1.5	1038	10.1	1060	22.6
922	901	15.3	896	17.9	921	17.7	942	99.2
850	833	43.4	840	28.9	850	34.5	881	34.1
775	771	9.8	799	13.4	804	17.9	838	16.2
640	620	4.2	658	6.8	672	8.9	681	7.9
527	602	3.7	639	1.9	658	0.7	562	34.9
477	521	16.7	533	2.3	549	9.8	413	3.0
399	432	15.7	403	28.0	400	11.3	352	111.0
296	368	39.4	352	45.4	335	34.9	303	30.4
283	287	8.1	279	3.7	283	14.1	275	17.2
219	262	4.2	231	4.2	232	3.6	255	3.3
184	178	6.9	170	3.4	171	2.4	54	3.5

^aM. Diem and coworkers, *J. Am. Chem. Soc.* **104** (1982) 3329.

^b6-31G* B3LYP frequencies.

^c6-31G* B3LYP APT.

^dBarron *et al.* (1991a).

optimized geometries. To each of these structures, four waters molecules were added by the Insight program (Biosym Technologies, San Diego, CA USA). The details of these calculations and the VA, VCD, Raman, and ROA spectra for this molecule have been presented (Jalkanen and Suhai, 1996; Han *et al.*, 1998; Jalkanen *et al.*, 2008a). Here we review the most important results. Note that recent developments by Cao *et al.* have reduced the linear birefringence in VCD measurements, which has increased significantly the sensitivity of modern

Tab. 1.2b L-Alanine-d₀ zwitterion vibrational frequencies and VCD intensities.

$\nu(\text{cm}^{-1})^a$	$10^5(\Delta\epsilon)^a$	$\nu(\text{cm}^{-1})^b$ I	R_i^c	$\nu(\text{cm}^{-1})^b$ II	R_i^c	$\nu(\text{cm}^{-1})^b$ III	R_i^c
3080	–	3317	17.1	3327	28.0	3153	–0.1
3060	–	3170	–1.4	3155	–2.3	3132	–11.5
3020	–	3120	1.7	3132	–4.1	3123	9.2
3003	–	3105	–76.2	3099	–17.6	3063	0.1
2993	–	3088	29.5	3057	–1.7	2293	–1.3
2962	–	3078	15.7	2957	–55.9	2230	0.5
2949	–	3053	–11.6	2940	21.6	2162	1.8
1623	–1.1	1770	–53.8	1777	–77.3	1698	42.9
1625	–	1745	–0.6	1746	162.9	1525	–5.4
1607	–	1741	26.4	1719	–65.2	1523	6.4
1498	–	1635	–6.4	1647	–5.5	1453	21.5
1459	–	1534	–4.6	1525	–2.5	1432	–56.5
1459	–	1525	–5.1	1513	14.9	1406	1.3
1410	–	1433	32.3	1444	26.6	1350	32.8
1375	–	1407	19.3	1415	–35.5	1275	8.6
1351	–27	1377	–71.3	1405	28.6	1252	9.6
1301	16	1333	–13.5	1341	–122.0	1246	–55.8
1220	5.3	1281	60.1	1298	18.9	1188	–4.1
1145	–4.2	1200	–24.8	1198	37.8	1122	–9.8
1110	–4.7	1124	–0.7	1115	12.3	1077	10.9
1001	–	1058	0.4	1074	15.9	941	–1.4
995	–	1026	6.6	1022	–7.1	904	–45.6
922	–	901	–23.4	896	–20.8	900	67.0
850	–	833	21.1	839	16.6	835	–25.2
775	–	771	23.6	799	–5.2	786	–8.3
640	–	620	–1.0	658	–18.2	657	0.7
527	–	602	10.1	639	7.7	540	–10.9
477	–	521	–8.5	533	–7.7	471	1.5
399	–	432	14.9	403	43.7	383	16.5
296	–	368	–26.3	352	51.4	321	45.4
283	–	287	39.9	279	–23.7	281	–34.2
219	–	262	–22.9	231	–25.9	231	3.5
184	–	179	–9.3	170	–35.5	169	–14.8

^aM. Diem, *J. Am. Chem. Soc.* **118** (1988) 6967–6970.

^b6-31G* B3LYP frequencies.

^c6-31G* B3LYP APT and 6-31G** RHF DO gauge AAT.

VCD instrumentation (Cao, Dukor and Nafie, 2008). This increased sensitivity is very important for amino acid, peptide, and protein VCD measurements in aqueous solution.

In Figure 1.2, we present the structure of the NALANMA plus four water molecular complex (Jalkanen *et al.*,

2008a). As one can see the structures and energetics of this molecule are greatly affected by the solvent, consistent with large changes in the VCD spectra of this molecule when one changes solvent from carbon tetrachloride to water. Note that the C_7^{eq} and C_5^{ext} conformers both converge to the same structure,

Tab. 1.3 *N*-Acetyl-L-alanine *N'*-methylamide conformational energies.

Conformer	ϕ^a	ψ^a	Energy ^a (kcal mol ⁻¹)	ϕ^b	ψ^b	Energy ^b (kcal mol ⁻¹)
C_7^{eq}	-82	72	0.000	-94	128	0.000
C_5^{ext}	-157	165	1.433	-94	128	0.000
C_7^{ax}	74	-60	2.612	59	-122	4.132
β_2	-136	23	3.181	-151	116	1.886
α_L	68	25	5.817	61	52	2.754
α_R	-60	-40	5.652	-82	-44	2.465
α_D	57	-133	6.467	67	-111	3.715
α'	-169	-38	6.853	-153	-92	15.140
Cryr	-84	155	-	-98	112	5.864

^aIsolated NALANMA, 6-31G* B3LYP relative energies.

^bNALANMA with 4 bound waters, 6-31G* B3LYP relative energies.

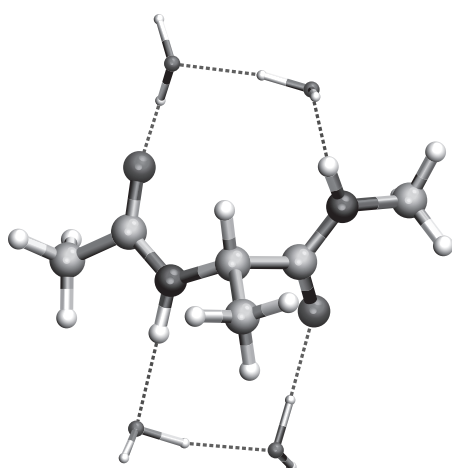


Fig. 1.2 *N*-Acetyl L-alanine *N'*-methylamide stabilized by four water molecules. B3PW91/6-31G* plus CPCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

which is the lowest energy structure of NALANMA with four bound waters found till date. This structure has also been used to reanalyze the NMR data for this molecule (Poon *et al.*, 2000; Weise and Weisshaar, 2003). The conclusions from both the NMR and VA, VCD, Raman, and ROA works is that NALANMA plus four water molecule complex is the entity (conformer) of interest, and not the isolated NALANMA molecule (Jalkanen *et al.*, 2008a). The structure

of this molecule in aqueous solution had been studied by a variety of both experimental and theoretical methods, but had alluded all previous researchers. The problem had been that most of the previous researchers had failed to take into account the effects of the first solvation water molecules, which are responsible for stabilizing a structure which is not stable on the gas phase isolated-state potential energy surface. This work has been called the gold

standard for vibrational spectroscopic structure determination for small peptides. The work on this molecule also documented the importance of measuring not only the VA and VCD or Raman and ROA but also all four spectra in addition to NMR. By measuring all these spectra and getting a structure or set of structures which are consistent with all of the measured spectra, one is able to get a self-consistent picture. This has also been advocated by Serdyuk, Zaccai and Zaccai (2007).

1.5.2.3 L-Alanyl L-Alanine

NALANMA is called the *alanine dipeptide* due to the formation of two peptide bonds by the capping. For biochemists dipeptides have only one peptide bond and keep the N-terminus and C-terminus groups. Hence to model dipeptides, one must now once again treat the N-terminus and C-terminus groups. In aqueous solution at pH 7 once again one has the neutral ionic species, the zwitterionic form of L-alanyl L-alanine.

In Tables 1.4a and 1.4b, we present the structures of zwitterionic structures stabilized by 4, 6, 7, 11, and 17 waters determined, respectively. Here one can see the effect of adding progressively more water molecules. As one can see the more water molecules one adds, the closer the LALA molecule comes to what it would be in aqueous solution. The more water molecules one adds, the more one is able to complete the first solvation shell. The polar groups are enough to hold a small number of water molecule in place, but when one starts to solvate the nonpolar groups one is faced with a problem that during the geometry optimization the water molecules tend to clump together, and this is even without any

kinetic energy (i.e. zero temperature). Hence one is faced with the problem of completing the first solvation shell completely, as we have done for LA where we found we needed 20 water molecules to fully encapsulate the LA zwitterion. In a future, study we will completely solvate the LALA zwitterion, but in this work we have only presented and refer to our structures stabilized by 4, 6, 7, 11 and 17 water molecules. In Figure 1.3, we present the LALA zwitterion stabilized by 11 water molecules and the PCM continuum solvent model with the B3PW91 hybrid XC functional and the 6-311G(2d,p) basis set. The 11 water molecules are directly hydrogen bonded to the polar groups, whereas the PCM continuum model is added to treat the effects due to the bulk water. This is similar to our previous works and also the methods recently advocated by Thar, Zhan and Kirchner (2008). In Figure 1.4, we present the simulate VA and VCD spectra for the LALA zwitterion stabilized by 11 waters at the B3PW91/6-311G(2d,p) level of theory including the PCM continuum model, and in Figure 1.5 the corresponding Raman and ROA spectra. In Figure 1.6, we present the LALA zwitterion stabilized by 17 waters and PCM continuum model with the B3PW91 hybrid XC functional and the 6-31G* basis set. In a future work, we will present all of our work on the LALA zwitterion (Jalkanen *et al.*, 2009b).

We also present two zwitterionic structures stabilized by the Onsager model implemented in Gaussian (94, 98, and 03) with and without explicit waters. In addition to the Onsager continuum model, we have optimized the structures of a collection of LALA plus N water molecule complexes, similar to our work for NALANMA with both the PCM and

Tab. 1.4a L-Alanyl-L-alanine zwitterion structures: bond lengths and dihedral angles.

Coordinate	I ^a	I ^b	I ^c	II ^d	III ^e	IV ^f	V ^g
R(C2N1)	1.5232	1.5090	1.5065	1.5060	1.5086	1.4774	1.4854
R(C3C2)	1.5638	1.5532	1.5531	1.5516	1.5460	1.5297	1.5332
R(O4C3)	1.2384	1.2348	1.2383	1.2224	1.2311	1.2372	1.2375
R(N5C3)	1.3411	1.3437	1.3384	1.3651	1.3524	1.3297	1.3404
R(C6C2)	1.5299	1.5216	1.5219	1.5224	1.5288	1.5237	1.5256
R(H7C2)	1.0930	1.0955	1.0957	1.0967	1.0917	1.0946	1.1004
R(H8C6)	1.0954	1.0929	1.0935	1.0953	1.0944	1.0924	1.0955
R(H9C6)	1.0938	1.0928	1.0925	1.0908	1.0943	1.0894	1.0921
R(H10C6)	1.0948	1.0950	1.0953	1.0952	1.0964	1.0921	1.0949
R(H11N1)	1.0289	1.0359	1.0340	1.0340	1.1524	1.0493	1.0515
R(H12N1)	1.0487	1.0566	1.0547	1.0386	1.0227	1.0417	1.0532
R(H23N1)	1.0276	1.0381	1.0457	1.0554	1.0462	1.0494	1.0473
R(C13N5)	1.4726	1.4754	1.4771	1.4862	1.4731	1.4491	1.4550
R(H22N5)	1.0216	1.0191	1.0197	1.0187	1.0168	1.0234	1.0240
R(C14C13)	1.5777	1.5590	1.5607	1.5568	1.5590	1.5392	1.5443
R(O15C14)	1.2572	1.2773	1.2757	1.2750	1.2814	1.2500	1.2680
R(O16C14)	1.2617	1.2465	1.2449	1.2476	1.2457	1.2591	1.2614
R(C17C13)	1.5340	1.5265	1.5262	1.5251	1.5284	1.5288	1.5288
R(H18C13)	1.0974	1.0951	1.0948	1.0921	1.0898	1.0883	1.0951
R(H19C17)	1.0972	1.0936	1.0938	1.0941	1.0955	1.0932	1.0938
R(H20C17)	1.0933	1.0947	1.0950	1.0967	1.0950	1.0923	1.0948
R(H21C17)	1.0946	1.0944	1.0942	1.0953	1.0947	1.0901	1.9841
τ (O4C3C2N1)	-7.35	-123.42	-126.36	-146.63	109.83	-19.01	12.38
ψ_1 (N5C3C2N1)	172.37	57.50	54.50	34.38	-67.77	162.77	-168.59
ω (C13N5C3C2)	178.91	177.48	179.06	162.30	155.23	175.15	173.89
ϕ_2 (C14C13N5C3)	-169.65	-162.49	-164.92	-135.88	-84.87	-134.60	-72.35
ψ_2 (O15C14C13N5)	-2.30	67.31	67.56	85.50	123.81	-16.03	-30.38
ψ_2' (O16C14C13N5)	178.40	-108.61	-107.74	-88.70	-55.18	164.41	150.76
τ (H8C6C2H7)	178.23	179.03	177.67	175.82	177.19	-177.97	-179.94
τ (H9C6C2H7)	-62.17	-61.73	-63.13	-65.66	-62.35	-57.55	-58.89
τ (H10C6C2H7)	56.84	57.11	56.08	54.65	56.58	61.74	59.44
τ (C3C2N1H11)	127.44	-173.41	-174.97	-174.94	-47.16	178.40	171.26
τ (C3C2N1H12)	8.00	59.80	57.88	57.92	-167.63	60.66	49.88
τ (C3C2N1H23)	-110.70	-62.10	-63.68	-63.68	73.09	-61.85	-69.26
τ (H19C17C13H18)	177.98	-177.54	-176.62	-177.83	180.00	-176.93	178.19
τ (H20C17C13H18)	57.35	60.61	61.26	59.99	58.94	62.89	57.34
τ (H21C17C13H18)	-63.59	-59.17	-58.66	-59.33	-61.10	-56.27	-61.40

^aZwitterion stabilized by Onsager solvent model at B3LYP/6-31G* level of theory.^bZwitterion stabilized by Onsager solvent model and seven water molecules at B3LYP/6-31G* level of theory.^cZwitterion stabilized by seven water molecules at B3LYP/6-31G* level of theory.^dZwitterion stabilized by six water molecules at B3LYP/6-31G* level of theory.^eZwitterion stabilized by four water molecules at B3LYP/6-31G* level of theory.^fZwitterion stabilized by PCM continuum model and 11 water molecules at B3PW91/6-311G(2d,p) level of theory.^gZwitterion stabilized by PCM continuum model and 17 water molecules at B3PW91/6-31G* level of theory.

Tab. 1.4b L-Alanyl-L-alanine zwitterion structures: valence angles.

Coordinate	I ^a	II ^b	III ^c	IV ^d	V ^e	VI ^f	VII ^g
A(N1C2C3)	104.60	108.46	108.62	111.34	107.58	108.47	106.70
A(C2C3O4)	117.62	122.01	121.46	119.87	119.28	119.48	118.36
A(C2C3N5)	118.06	113.69	113.97	115.82	115.75	115.10	118.56
A(O4C3N5)	124.32	124.30	124.56	124.30	124.92	125.40	123.08
A(N1C2C6)	110.20	111.07	111.01	110.29	111.26	110.36	109.94
A(C3C2C6)	113.96	113.57	113.14	111.90	115.40	111.57	115.45
A(N1C2H7)	106.97	105.17	105.37	105.40	106.62	107.11	107.13
A(C3C2H7)	110.19	107.92	108.05	106.97	104.87	109.27	107.83
A(C6C2H7)	110.56	110.27	110.32	110.69	110.57	109.95	109.46
A(C2C6H8)	111.12	111.80	111.89	110.92	111.81	110.79	112.81
A(C2C6H9)	109.73	109.25	109.33	109.21	109.84	109.76	110.50
A(H8C6H9)	108.18	107.80	107.68	107.68	108.46	109.00	108.03
A(C2C6H10)	111.08	110.38	110.35	110.86	110.70	110.09	109.60
A(H8H6H10)	108.89	109.32	109.03	108.95	108.11	108.72	108.19
A(H9H6H10)	107.73	108.18	108.47	109.16	107.80	108.45	107.53
A(C2N1H11)	113.36	110.71	110.89	110.80	112.75	110.30	111.06
A(C2N1H12)	104.73	111.06	110.52	110.18	110.16	111.04	111.78
A(C2N1N23)	113.73	108.64	108.71	109.57	110.58	110.68	114.04
A(11N1H12)	109.61	113.31	113.79	114.03	107.80	106.44	108.53
A(11N1H23)	106.50	102.08	101.94	101.56	107.40	108.24	105.84
A(12N1H23)	108.83	110.64	110.59	110.36	107.98	110.03	105.18
A(C3N5C13)	123.93	125.32	125.71	122.12	122.25	122.52	118.95
A(N5C13C14)	109.30	104.48	103.88	103.54	108.17	111.98	111.66
A(C13C14O15)	116.15	114.42	114.25	115.12	118.40	118.14	118.58
A(C13C14O16)	115.02	118.54	118.09	116.93	116.18	114.70	116.49
A(O15C14O16)	128.83	126.90	127.47	127.65	125.41	127.15	124.93
A(N5C13C17)	111.91	111.91	111.66	111.44	110.49	112.01	111.66
A(C14C13C17)	110.77	113.70	113.67	113.73	111.73	110.52	110.35
A(N5C13H18)	108.21	107.24	107.43	106.98	107.05	106.77	107.82
A(C14C13H18)	108.11	108.86	109.19	109.85	108.40	106.94	106.71
A(C17C13H18)	108.42	110.31	110.63	110.87	110.84	108.35	108.43
A(C13C17H19)	110.36	110.28	110.08	110.12	110.64	111.10	112.77
A(C13C17H20)	111.33	110.40	110.61	110.57	110.92	110.13	109.96
A(H19C17H20)	108.56	110.07	110.27	110.32	108.97	108.45	108.20
A(C13C17H21)	108.19	109.90	109.85	110.28	109.84	109.35	109.26
A(H19C17H21)	108.54	107.52	107.31	107.52	107.81	109.24	108.24
A(H20C17H21)	108.56	108.61	108.64	107.95	108.58	108.52	108.31
A(C3N5H22)	124.46	118.18	118.21	117.63	117.85	119.37	121.34
A(C13N5H22)	111.31	116.38	116.00	114.00	114.98	118.07	118.49

^aZwitterion stabilized by Onsager solvent model at B3LYP/6-31G* level of theory.^bZwitterion stabilized by Onsager solvent model and seven water molecules at B3LYP/6-31G* level of theory.^cZwitterion stabilized by seven water molecules at B3LYP/6-31G* level of theory.^dZwitterion stabilized by six water molecules at B3LYP/6-31G* level of theory.^eZwitterion stabilized by four water molecules at B3LYP/6-31G* level of theory.^fZwitterion stabilized by PCM continuum model and 11 water molecules at B3PW91/6-311G(2d,p) level of theory.^gZwitterion stabilized by PCM continuum model and 17 water molecules at B3PW91/6-31G* level of theory.

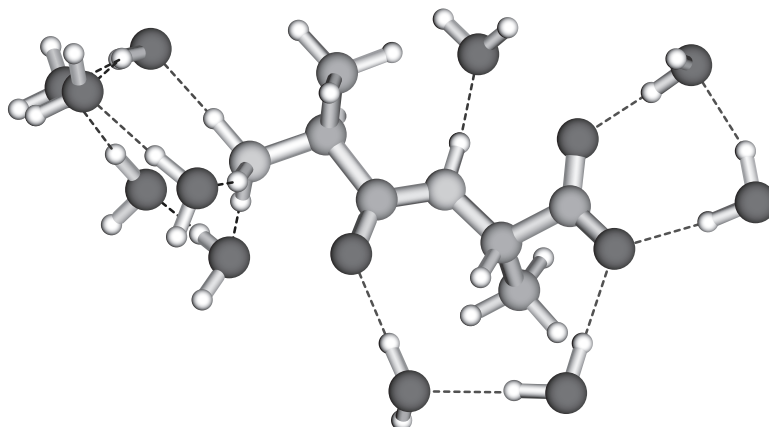


Fig. 1.3 L-alanyl-L-alanine zwitterion stabilized by 11 water molecules. B3PW91/6-311G2dp plus PCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

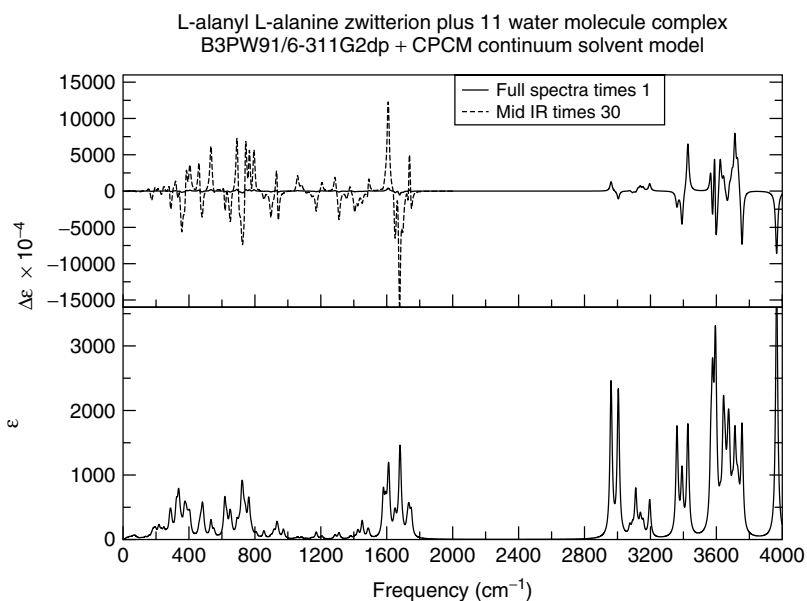


Fig. 1.4 L-Alanyl-L-alanine zwitterion stabilized by 11 water molecules. B3PW91/6-311G2dp plus PCM continuum solvent model for VA and VCD spectra.

CPCM continuum solvent models as implemented in Gaussian. The structures were extracted from classical MD simulations of the LALA zwitterion in a box of water molecules (Knapp-Mohammady

et al., 1999; Jalkanen *et al.*, 2003, 2009b). Only the waters which were directly hydrogen bonded with the polar groups were kept initially for the full DFT level calculations. This was the protocol which

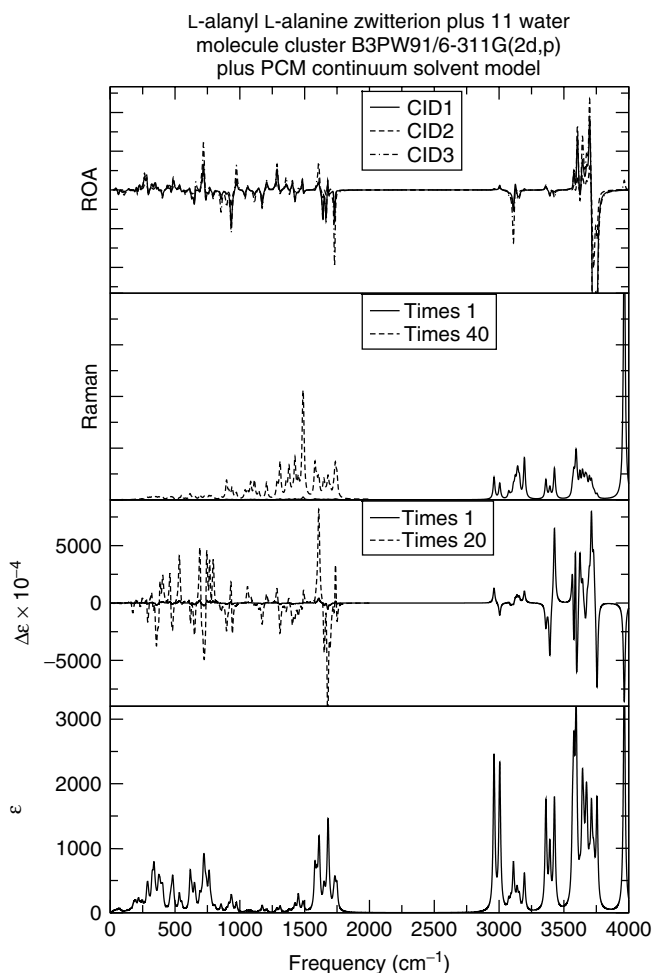


Fig. 1.5 L-Alanyl-L-alanine zwitterion stabilized by 11 water molecules. B3PW91/6-311G2dp plus PCM continuum solvent model for VA, VCD, Raman, and ROA spectra.

worked for LA and NALANMA, but it failed for the LALA zwitterion initially, as mentioned above. Since there is a competition for the hydrogen-bonded partners, some of the water molecules which were initially hydrogen bonded with the polar groups of the LALA zwitterion were pulled away, and hence polar groups of the LALA zwitterion ended up with no hydrogen bonding partners. Our simple

solution to this problem was to explicitly add another water molecule to the polar group which now lacked a hydrogen-bonded water partner. The more general problem, which we have used for LA, is to optimize the structure of the solute plus a complete solvation shell. This is of course a much more stable entity than just the solute with the water molecules directly hydrogen bonded to the polar groups of

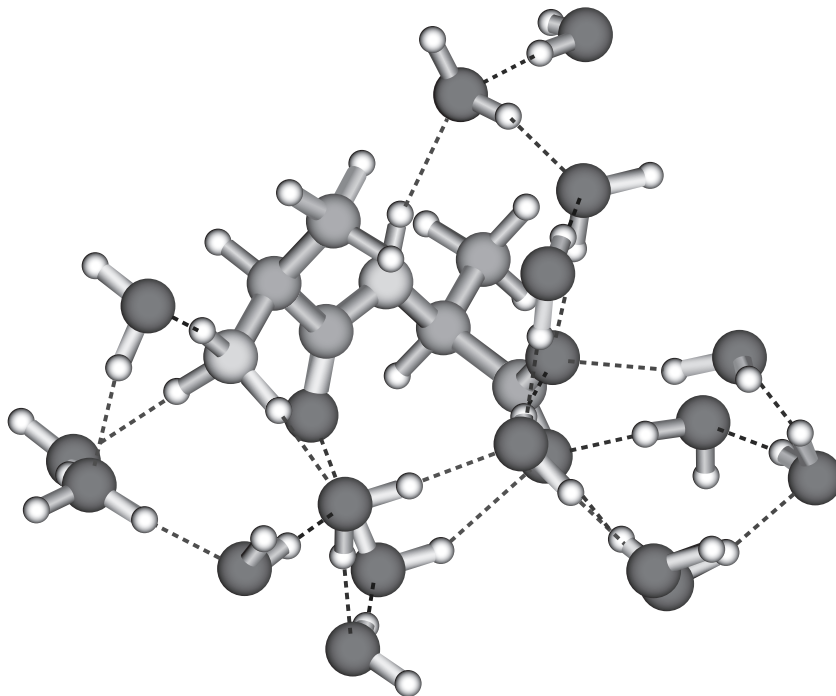


Fig. 1.6 L-Alanyl-L-alanine zwitterion stabilized by 17 water molecules. B3PW91/6-31G* plus PCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

the solute. In many cases the hydrogen bonding partners are necessary to orient and stabilize the water molecules that are hydrogen bonded to the polar groups (Degtyarenko *et al.*, 2007, 2008; Jalkanen *et al.*, 2008a).

Our goal is to try to interpret the measured VA, VCD, Raman, and ROA spectra and to be able to determine the structure or structures present under the condition of the experiment. If one cannot determine the conformer and species for simple peptides, one could not be expected to do so for polypeptides like Leu- and Met-enkephalin or proteins. If one wishes to get more than simple secondary structural elements present, which is the backbone and side chain torsion angles for all residues and the

protonation state of all polar residues, including the N- and C-termini, then one must extend the methods of vibrational spectroscopy to many dimensions, as was done for NMR, which finally allowed one to get a complete 3D structure. In many cases, in addition to the amino acid, peptide, and protein, one must also include and determine the position of certain water molecules, cations, anions, ligands, and prosthetic groups which are responsible for stabilizing the structure of the molecular complex.

To determine the structure of the LALA zwitterion in aqueous solution requires a conformational search for the lowest energy zwitterionic structure of LALA. The results here only document the feasibility of explicitly adding water to the

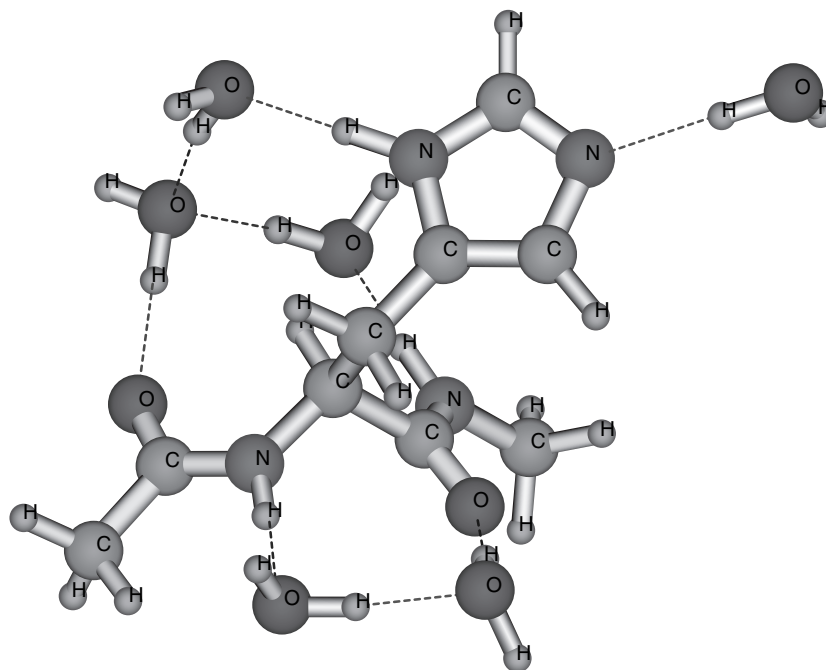


Fig. 1.7 *N*-Acetyl *L*-histidine *N'*-methylamide stabilized by six water molecules. B3PW91/6-31G* plus PCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

calculation of the zwitterionic structures of dipeptides and the calculation of the VA and VCD spectra. Our goal here is to be able to determine the structure(s) of the zwitterionic state of LALA in aqueous solution. Our preliminary results show that the VCD spectra is very sensitive to the structure of the zwitterion of LALA and one should be able to distinguish which conformer(s) is(are) present if one knew what the predicted VCD spectra for each conformer was. We are pursuing the determination of the other possible structures of the zwitterion of LALA and will report the complete results in a future publication (Jalkanen *et al.*, 2009b).

In Table 1.5, we present our calculated vibrational frequencies, VA and VCD intensities, for three of the structures presented in Tables 1.4a and 1.4b.

1.5.2.4 *L*-Histidine

In addition to LA, NALANMA, and LALA, one is of course interested in other amino acids. One of the most interesting and functionally important amino acids is LH. Recently we have reported a combined experimental and theoretical study of LH in aqueous solution (Deplazes *et al.*, 2008). For complete details on this work one is referred to that paper. Here we only summarize the results. To determine the structure of LH in aqueous solution one must first stabilize the zwitterionic part of the amino acid. Here we can use the structure of the LA zwitterion, which we determined in our earlier works (Jalkanen, Bohr and Suhai, 1997; Tajkhorshid, Jalkanen and Suhai, 1998; Frimand *et al.*, 2000; Jalkanen *et al.*,

Tab. 1.5 L-Alanyl-L-alanine zwitterion VA and VCD spectra.

$\nu(\text{cm}^{-1})^a$	A_i^b III	R_i^c	$\nu(\text{cm}^{-1})^a$	A_i^b II	R_i^c	$\nu(\text{cm}^{-1})^a$	A_i^b I	R_i^c
1745	95.12	-249.30	1793	172.50	-106.04	1764	37.59	-100.23
1711	269.18	182.81	1769	198.20	116.73	1742	441.64	8.05
1701	284.78	-418.71	1711	641.68	157.82	1718	413.77	304.51
1678	187.10	193.42	1680	67.76	-105.94	1696	73.09	-64.26
1625	265.53	388.67	1615	148.81	10.31	1635	27.49	30.81
1588	190.59	94.32	1586	200.83	70.66	1625	162.49	-67.03
1532	2.92	3.80	1526	19.72	2.27	1530	3.88	12.57
1527	23.00	-25.74	1524	13.84	-15.51	1523	10.48	13.03
1522	2.33	-6.22	1518	7.16	-1.29	1523	5.72	-4.90
1520	2.67	3.40	1516	7.73	6.35	1516	15.82	3.13
1447	9.43	15.76	1440	24.90	2.70	1453	25.63	1.29
1439	40.84	11.22	1432	117.64	24.21	1431	102.61	11.16
1416	24.09	-12.95	1418	113.12	52.53	1419	45.70	51.86
1395	232.67	-161.32	1414	88.42	-132.80	1409	67.20	97.46
1372	9.22	-14.27	1393	11.50	73.31	1404	88.05	-201.64
1359	2.81	0.36	1384	44.91	-109.25	1372	12.27	21.19
1329	43.83	-129.21	1303	107.24	-175.64	1334	83.33	-69.71
1309	27.94	74.51	1275	28.14	-21.82	1292	68.84	22.65
1240	15.70	-23.91	1266	188.86	201.30	1287	15.75	18.11
1177	114.51	4.39	1203	42.40	-69.49	1210	21.24	11.89
1162	15.55	43.66	1150	4.67	-34.19	1146	13.12	1.91
1124	8.32	-40.32	1137	24.69	49.74	1141	8.25	14.75
1110	23.95	-6.83	1112	26.83	-17.96	1119	15.40	-9.43
1088	11.17	11.86	1070	3.69	-3.71	1072	1.98	-9.97
1088	11.17	11.86	1070	3.69	-3.71	1072	1.98	-9.97
1061	10.17	6.54	1049	11.13	17.96	1067	27.22	3.84
1018	7.54	-9.04	1031	10.09	-6.41	1035	2.63	-11.54
942	8.10	-10.99	949	14.38	-45.67	953	14.15	-28.20
925	8.47	-29.98	930	7.62	-11.18	924	1.78	-8.94
883	8.91	-10.74	883	13.21	11.28	883	8.71	10.84
863	52.91	-33.16	854	30.93	-5.04	853	20.83	16.04
800	49.65	34.95	804	15.15	30.85	791	12.67	37.58
797	14.39	-3.63	771	6.69	22.72	771	6.78	-28.23
750	62.17	-10.17	708	79.82	21.54	726	39.60	9.79
693	14.89	9.25	676	20.33	0.12	665	18.49	4.71
629	1.56	-21.74	581	12.38	-42.27	601	8.46	-24.13
586	3.27	-9.87	550	3.88	-49.09	543	17.07	3.19
535	9.27	-9.73	548	11.19	19.08	531	7.81	-37.80
505	11.96	32.14	468	11.67	25.32	486	6.19	7.61
394	34.00	16.21	447	9.32	36.02	443	32.06	46.97
385	13.66	23.12	368	4.97	5.29	366	8.57	21.29
342	6.62	-49.92	361	4.56	32.73	361	11.85	14.20
303	3.18	13.31	325	10.25	-37.13	311	4.21	-27.76
276	27.88	37.18	291	19.00	61.69	292	13.35	41.52
253	4.52	15.10	251	0.32	-6.70	254	2.37	-9.71
245	1.13	6.52	236	14.52	25.11	246	6.49	-11.81
222	10.18	-51.44	230	14.44	-59.15	235	5.07	7.62

Tab. 1.5 (Continued)

$\nu(\text{cm}^{-1})^a$	A_i^b III	R_i^c	$\nu(\text{cm}^{-1})^a$	A_i^b II	R_i^c	$\nu(\text{cm}^{-1})^a$	A_i^b I	R_i^c
193	27.45	-46.99	189	10.04	1.79	197	26.41	0.31
172	9.89	-15.30	174	12.99	-8.18	170	26.67	-53.71
165	2.71	26.63	134	4.57	-25.58	157	6.95	33.04
142	6.99	18.80	105	5.51	15.96	124	5.04	-44.78
125	2.36	-9.23	83	6.30	-13.56	109	0.44	2.06

^a6-31G* B3LYP frequencies.

^b6-31G* B3LYP APT.

^c6-31G* B3LYP APT and 6-31G** RHF DO gauge AAT.

2001) and mutate the methyl group to an imidazole group. Here we have two polar groups, the ring N-H group and the N atom. Following our previous work, we have added two water molecules to hydrogen bond with each group. This is our minimal explicit solvent model. The question we wished to address was what are the two side chain angles, χ_1 and χ_2 . Hence we have optimized the structure for the various possible conformers. For the $C_\alpha-C_\beta$ bond, we have a sp^3-sp^3 bond, and hence one would expect three possibilities, gauche plus, trans, and gauche minus. For the $C_\beta-C_\gamma$ bond, we have a sp^3-sp^2 bond, and hence one would expect various possibilities depending on the interactions between the various substituents. Here one could expect cis and trans, in addition to gauche plus and gauche minus. There is the possibility of the N-H group hydrogen bonding with the C=O group of the backbone or with one of the water oxygen atoms. In the presence of explicit water molecules, the backbone C=O oxygen atom would be screened from the ring NH hydrogen atom. Hence one would expect different results for the gas phase isolated state, for implicit solvent models (Onsager, PCM,

CPCM, and/or COSMO), for explicit solvent models, and finally for our hybrid model, where we embed the LH + N water molecule complexes within the continuum solvent models. From our comparison with the VA, VCD, Raman, and ROA spectra measurements we concluded that it is important to include explicit water molecules and additionally that the gauche minus gauche minus conformer is the most likely one to be present, though we were not able to definitely rule out the presence of the gauche minus gauche plus.

We have previously presented the structures for the LH zwitterion calculated at the B3LYP/6-31G* plus PCM level of theory and the interested readers can refer to those works (Jalkanen *et al.*, 2006a; Deplazes *et al.*, 2008). In the next subsection, we present our new work on capped LH.

1.5.2.5 N-Acetyl L-Histidine N'-Methylamide

Similar to capping LA to produce capped LA, one can cap LH to produce capped LH. Here one has the backbone ϕ and ψ angles in addition to the χ_1 and χ_2 side chain angles. In Tables 1.6a,

Tab. 1.6a N-Acetyl-L-histidine N'-methylamide conformational energies.

Conformer	ϕ	ψ	χ_1	χ_2	Energy Hartrees	
gm gm	-162.26	173.64	-127.55	61.85	-720.613603718	Isolated ^a
gm gm	-124.91	132.35	-58.49	-78.23	-720.641205272	PCM ^b
gm gm	-119.14	128.43	-61.29	-87.41	-720.641751542	CPCM ^c
gm gp	-106.07	149.36	-68.19	57.63	-720.603271434	Isolated ^a
gm gp	-115.85	130.42	-61.20	85.91	-720.641232605	PCM ^b
gm gp	-116.36	131.30	-62.09	85.19	-720.641609495	CPCM ^c
gp gm	-151.60	177.47	67.74	-59.58	-720.605035221	Isolated ^a
gp gm	-142.19	161.99	62.14	-89.03	-720.640392649	PCM ^b
gp gm(ts)	-144.49	163.55	64.23	-86.38	-720.640865462	CPCM ^c
gp gm	-143.48	163.70	63.45	-86.64	-720.640832768	CPCM ^c
gp gp	-82.88	67.07	49.25	69.20	-720.618655982	Isolated ^a
gp gp	-135.38	126.74	43.41	52.26	-720.639519222	PCM ^b
gp gp	-141.30	129.55	79.86	49.23	-720.640240001	CPCM ^c
t gm	-82.84	67.74	-168.94	-70.77	-720.613246215	Isolated ^a
t gm	-132.57	118.03	179.93	-87.97	-720.641942617	PCM ^b
t gm	-135.20	119.68	179.83	-87.82	-720.641957463	CPCM ^c
t gp	-162.26	173.64	-127.54	61.85	-720.613603651	Isolated ^a
t gp (ts)	-126.44	125.05	-178.17	73.90	-720.641944678	PCM ^b
t gp	-130.52	125.55	-178.28	73.65	-720.641923448	PCM ^b
t gp	-142.74	130.04	-179.65	72.59	-720.642396472	CPCM ^c
t gp(6-31g**)	-134.57	126.85	-177.99	73.36	-720.667446267	CPCM ^c

^aNALHNMA 6-31G* B3PW91 relative energies.^bNALHNMA 6-31G* B3PW91 + PCM relative energies.^cNALHNMA 6-31G* B3PW91 + CPCM relative energies.

1.6b, and 1.6c, we present the structures for NALHNMA, NALHNMA plus four water molecule complex, and NALHNMA plus six water molecule complex at the B3PW91/6-31G*, B3PW91/6-31G* plus PCM continuum solvent model and B3PW91/6-31G* plus CPCM continuum solvent model level of theory (Jalkanen *et al.*, 2009a). As one can see in Table 1.6a for the gauche minus gauche minus conformer, structures which are stable within the continuum solvent model are not stable without it. Here the continuum solvent model has effectively taken into account the screening due to the solvent. In the absence of both explicit

solvent molecules or the continuum solvent model, the histidine side chain rotates to form a strong hydrogen bond between the ring N-H group and the N-acetyl group C=O group. Hence the importance of taking into account solvent effects. The VA, VCD, Raman, and ROA spectra for all species will be presented in a future work (Jalkanen *et al.*, 2009a). In Figure 1.7, we present one of the conformers of NALHNMA stabilized by the presence of six explicit water molecules (NALHNMA hydrated complex) which has been embedded with the PCM continuum model that treats the effects due to the bulk water molecules.

Tab. 1.6b *N*-Acetyl-L-histidine *N'*-methylamide plus four water molecule complex conformational energies.

Conformer	ϕ	ψ	χ_1	χ_2	Energy Hartrees	
gm gm	-111.17	123.21	-52.37	-50.94	-1026.22240779	Isolated ^a
gm gm	-100.52	126.20	-64.60	-88.19	-1026.25855737	PCM ^b
gm gm	-100.43	126.23	-63.78	-87.51	-1026.25912256	CPCM ^c
gm gp	-106.58	120.16	-61.32	145.72	-1026.21352819	Isolated ^a
gm gp	-97.22	124.77	-65.36	93.20	-1026.25611672	PCM ^b
gm gp	-99.94	126.74	-74.53	93.88	-1026.25703059	CPCM ^c
gp gm	-130.40	155.34	83.26	-58.20	-1026.21719542	Isolated ^a
gp gm	-108.08	153.79	86.22	-67.84	-1026.25795559	PCM ^b
gp gm	-108.42	153.92	85.79	-67.44	-1026.25846269	CPCM ^c
gp gp	-151.34	115.42	42.30	54.74	-1026.21296490	Isolated ^a
gp gp	-118.97	135.83	59.70	93.17	-1026.25224287	PCM ^b
gp gp	-144.22	144.79	58.21	94.61	-1026.25444179	CPCM ^c
gp gp	-146.21	143.91	58.24	95.18	-1026.25454373	CPCM ^c
t gm	-104.15	114.33	176.24	-77.11	-1026.21907233	Isolated ^a
t gm	-98.87	127.28	-174.74	-92.38	-1026.25819396	PCM ^b
t gm	-98.65	126.85	-174.47	-92.48	-1026.25917833	CPCM ^c
t gp	-130.97	119.11	-163.46	66.23	-1026.22157806	Isolated ^a
t gp	-96.83	123.70	-179.36	82.00	-1026.25900659	PCM ^b
t gp	-96.82	123.67	-179.38	82.02	-1026.25896453	PCM ^b
t gp	-96.96	123.60	-179.40	82.00	-1026.25948994	CPCM ^c
t gp	-96.97	123.58	-179.41	81.01	-1026.25944361	CPCM ^c

^aNALHNMA plus four water molecule complex 6-31G* B3PW91 relative energies.^bNALHNMA plus four water molecule complex 6-31G* B3PW91 + PCM relative energies.^cNALHNMA plus four water molecule complex 6-31G* B3PW91 + CPCM relative energies.

1.5.2.6 *N*-Acetyl L-Alanine L-Proline L-Tyrosine *N'*-Methylamide

In Figure 1.8, we present one of the 14 structures we have found to date for capped L-alanine L-proline L-tyrosine, *N*-acetyl L-alanine L-proline L-tyrosine *N'*-methylamide at the B3PW91 DFT level of theory. The structures have been extracted from classical molecular dynamics simulations (Nardi *et al.*, 2000; Nardi, 1999; Nardi, Worth and Wade, 1997) and the waters have been removed and the structures have been optimized with the COSMO/PCM hybrid continuum solvation model as implemented

in Gaussian 03 (Jalkanen *et al.*, 2009c). Here we have not added explicit water molecules but have only used the continuum solvent treatment to see how well one can do with these larger peptides without actually including the explicit water molecules. In the case of the Raman and ROA spectral simulations, the inclusion of the explicit water molecules increased the cost of the Raman and ROA simulations, but with the new analytical methods developed by Ruud, the tensors required for the Raman and ROA spectral simulations are not longer rate limiting. Here one can calculate the

Tab. 1.6c N-Acetyl-L-histidine N'-methylamide plus six water molecule complex conformational energies.

Conformer	ϕ	ψ	χ_1	χ_2	Energy Hartrees	
gm gm	-103.93	115.93	-33.39	-65.51	-1179.02483018	Isolated ^a
gm gm	-107.63	115.49	-51.33	-66.53	-1179.06684481	PCM ^b
gm gm	-110.35	114.40	-51.87	-70.85	-1179.06740734	CPCM ^c
gm gm	-110.41	114.37	-51.89	-70.88	-1179.06740482	CPCM ^c
gm gp	-118.80	117.55	-52.05	111.03	-1179.01745486	Isolated ^a
gm gp					-1179.06371029	PCM ^b
gm gp	-118.55	116.37	-49.90	122.50	-1179.06482351	CPCM ^c
gp gm	-154.56	155.55	38.22	-101.72	-1179.01236234	Isolated ^a
gp gm	-144.70	148.25	25.69	-94.96	-1179.01307645	isolated ^a
gp gm	-116.83	149.29	61.84	-100.02	-1179.05500717	PCM ^b
gp gm	-116.44	152.83	63.15	-98.10	-1179.0524477	CPCM ^c
gp gp	-104.27	142.30	42.07	76.01	-1179.01765151	Isolated ^a
gp gp	-112.30	141.75	49.26	76.04	-1179.05968412	PCM ^b
gp gp	-112.28	141.68	49.23	76.14	-1179.06008150	CPCM ^c
t gm	-107.25	124.15	-174.10	-97.90	-1179.02992116	isolated ^a
t gm	-106.07	123.41	-179.19	-104.51	-1179.06675368	PCM ^b
t gm	-106.57	123.11	-179.52	-103.79	-1179.06717558	CPCM ^c
t gp	-113.35	122.95	-172.82	77.92	-1179.01858560	Isolated ^a
t gp	-110.05	116.25	-177.35	78.33	-1179.06840828	PCM ^b
t gp	-110.07	116.27	-177.35	78.36	-1179.06888670	CPCM ^c

^aNALHNMA plus six water molecule complex 6-31G* B3PW91 relative energies.

^bNALHNMA plus six water molecule complex 6-31G* B3PW91 + PCM relative energies.

^cNALHNMA plus six water molecule complex 6-31G* B3PW91 + CPCM relative energies.

Hessian, VA, and VCD spectra in one calculation and the tensors required for Raman and ROA in a separate calculation, and then combine the results, that is, the Hessian, normal modes (eigenvalues and eigenvectors) from the first calculation with the EDEDPD, EDMDPD and the EDEQPD from the second, to get the most accurate Raman and ROA intensities. This was advocated earlier for the Raman intensities by Amos (1986), as the basis sets and level of theory to get accurate geometries and Hessians appear to be different than that required for accurate EDEDP, EDMDP, EDEQP and the EDEDPD, EDMDPD and the

EDEQPD. But what one must at least be able to get is the structure/conformer which is present in aqueous solution. In our geometry optimizations with the waters removed, the structures stabilized by the waters were not stable (Jalkanen *et al.*, 2009c). This is consistent with our previous work on the alanine dipeptide, showing that this phenomenon for alanine peptide may not be an exceptional case, but the general case. Hence we advocate the use of our hybrid model where one includes the first solvation shell water molecules to treat the effects due to these waters and then embed this hydrated molecule within a continuum

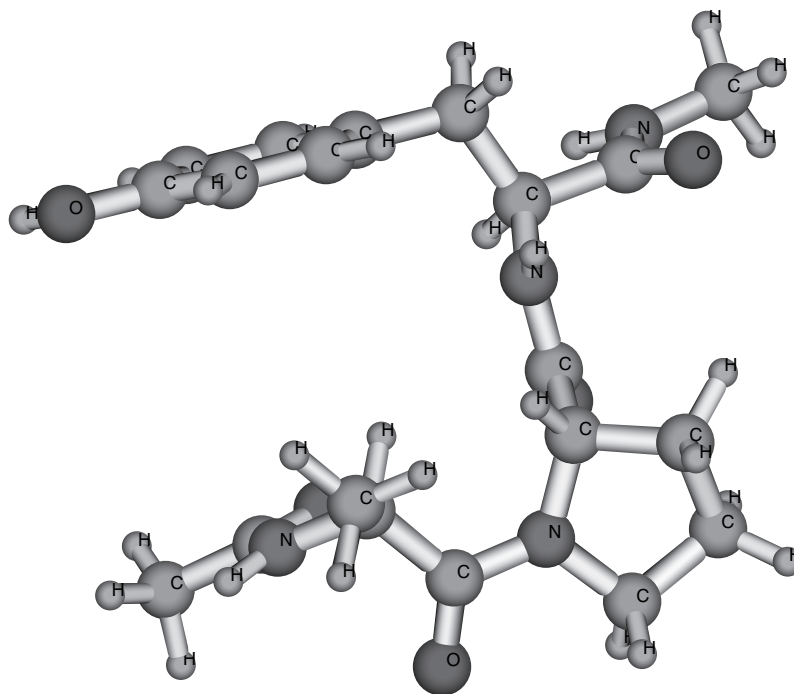


Fig. 1.8 *N*-Acetyl L-alanine L-proline L-tyrosine *N'*-methylamide molecule. B3PW91/6-31G** plus PCM continuum solvent model optimized geometry. (Please find a color version of this figure on the color plates.)

model to treat the effects due to bulk water (Jalkanen *et al.*, 2008a; Deplazes *et al.*, 2008; Frimand *et al.*, 2000; Han *et al.*, 1998; Tajkhorshid, Jalkanen and Suhai, 1998; Jalkanen, Bohr and Suhai, 1997; Jalkanen and Suhai, 1996). In some cases one may even have to include a second layer of water molecule. But the largest and most important effects appear to be due to only the first solvation/hydration layer, so this should be included minimally.

1.5.3

Future Perspectives and Applications for DFT-based Simulations

In this work, we have presented and discussed on the use of DFT simulations

and VA, VCD, Raman, and ROA spectroscopy to investigate the conformers of the ligands and proteins and the changes which occur to both on binding. There are of course relevant other approaches to the same problem (Bruno *et al.*, 1997). In addition, the binding of insulin, insulin mimetics, and other ligands to the insulin receptor is an important problem. The insulin receptor is transmembrane receptor with insulin binding to the extracellular part. Recent work has tried to characterize a second ligand binding site of the insulin receptor (Hao, Whittaker and Whittaker, 2006). In addition to the binding of insulin to the extracellular domain, the insulin receptor has an intracellular (cytoplasmic) tyrosine kinase domain, the role of which has also

been heavily investigated (Chou, 1990). Here we feel that VA, VCD, Raman, and ROA studies can contribute a lot to a better understanding of the structural changes which occur and to the mechanism/mode of action. Finally, in the last subsection of this work, we extend the multidisciplinary approach to more general problems in molecular medicine and in society in general.

1.5.4

Future Perspectives of the Multidisciplinary Fields of Molecular Biophysics, Molecular Biology, Systems Biology, and Molecular Medicine and Molecular Psychiatry

The future of chemistry, physics and mathematics appears to be secure and healthy through the application and use of the fundamental knowledge gained in these fields in the new fields of molecular biophysics, bionanotechnology and quantum nanobiology (Goodsell, 2002). In addition, the new field of molecular psychiatry has developed to the point where some behavioral disorders have been shown to be correlated with specificity genotypes (Kim-Cohen *et al.*, 2006). Monoamine oxidase A (MAOA) plays a role in regulating behavior by selectively degrading serotonin, norepinephrine, and dopamine following reuptake in the synaptic cleft. If MAOA is underexpressed, children who have been maltreated have been shown to develop conduct disorder, violent behavior and antisocial personalities in adulthood. On the other hand, children who overexpress MAOA do not appear to be affected. These results are very controversial and have not been fully accepted by everyone. Clearly more work needs to be

done in this area. The possibility of preventing and/or hindering the antisocial behaviours, conduct disorders, and violent reactions and behaviours by either gene therapy, drugs, and counselling, or a combination of the above would go a long way to solve this problem, which formerly was thought to be only due to poor upbringing and socialization. With the ever increasing prison populations in the United States and China, clearly, any alternatives are well worth exploring. In addition, many of those currently in our prisons and criminal justice system need to be tested for genetic problems that may be partly responsible for behaviours and actions inherent in the genetic makeup.

Hence, in addition to determining tendencies for diabetes, high blood pressure, heart disease, and cancer, there now appear to be genetic markers that are able to predict tendencies for antisocial and combative behavioral patterns when the individuals are psychologically, emotionally, and physically stressed (maltreated, sexually molested, physically abused, and mobbed) as children or young adults. The increasing incidences of violent behavior like the Columbine and Virginia Tech shootings may be able to be prevented by genetically testing young students who have been and continue to be mobbed and socially ostracized. Rather than demonize these young people and their behavior, more research needs to be undertaken to understand the root causes of such behavior and the triggering mechanisms. Many veterans and survivors of the Korean, Vietnam, Gulf 1 and 2, Afghanistan, and Balkan wars have shown similar behavior after being exposed to inhumane treatment and/or seeing atrocities. The increase of extremism appears to be an ever increasing problem and there does not appear

to be a lack of researchers trying to shed some light on this international problem (Gambetta, 2005; Abbas, 2005). In addition to psychological, part of the problem may indeed be biochemical. No one is saying that this is the full answer, but this field of research is well worth pursuing. Just like biomolecule interacting in various manners in various environments, it appears that many behavior patterns are the results of a complex set of physical, psychological, societal, economic, religious, and biochemical interactions. Ulcers were thought to be caused by stress, but subsequently were found to be caused by microorganisms, which when the immune response was weakened by other factors, were able to have adverse side effects that did not appear without the physical, mental, and psychological stress.

But these multidisciplinary and multifaceted fields require the researchers and scientists to be on top of many areas of endeavor. The pyramid principle of knowledge never made itself more apparent than in these problems. The idea of social workers, police and military officers, and our criminal justice worker, judges, and politicians are competent to understand these problems is really a fallacy. The systems are broken and this is shown by the record numbers of people being incarcerated in both the United States and China. When having to deal with complex problems, the simple black and white, with us or against us model is clearly not working. The world is not those who are with us and those who are against us. We have learned a lot from complex biological systems to see that the human body is a very complex machine and that when one try to fix one system with a drug, in many cases this causes adverse side effects, which in many cases

are even worse than the symptoms one was/is trying to treat. The real root cause of the problem and/or disease is not even known or dealt with. This is the same in many societies, where one tries to treat and/or deal with the symptoms of some social, religious, economic, societal, and/or political problem, and one does not deal with the root cause or causes. We must be smarter and use the scientific method, logic, reasoning, and philosophy, to get to the root causes of the problems we are facing in society, be it behavioral, environmental, economic, and/or with the public and individual health.

Tackling the multifaceted and multidisciplinary problems has changed the paradigm for many departments, funding agencies, government agencies, educational and training institutions, and publishers. The cross-fertilization of knowledge and techniques has never been more fruitful. What has prevented this interaction in the past has been the artificial barriers between what was mathematics, physics, chemistry, biology, medicine, political science, psychology, sociology, anthropology, religious studies, criminal justice, economics, and environmental sciences. We now realize that these barriers were/are really artificial, and, in many cases, these are responsible for some our global problems, like air and water pollution, chemical poisoning of our environment, high increase of allergies, global warming, world hunger, epidemics, terrorism, and the economic crises being experienced by many countries. By introducing new technologies, governmental policies, economic systems, chemicals, and even alternative energy sources, one has bettered mankind and the quality of life in one area, but has had long-term detrimental effects in other areas. The idea that free trade and

market economies would be the panacea for increasing the standard of living for everyone on the planet has been a big myth. The trade unions (organized labor), environmentalists, and religious leaders have not been included at the World Trade Organization (WTO) and G8 summit meeting. There has not been a consensus and trust from large segments of society. This has been due to the lack of experts and ideas from many of the major players: organized labor, environmentalists, religious leaders, and native indigenous nations which have been occupied, but who have almost no political, economic, and military strength (Tibetans, Tamil Tigers, Native Americans, Mexican Americans displaced by the land grab by the United States from Mexico: now the states of California, Nevada and Utah, and parts of the states of Wyoming, Texas, Arizona, Colorado, and New Mexico). Many of the 10 million so-called illegal aliens in the United States are a direct result of the failed policies and treaties resulting from the American–Mexican war. A settlement by politicians in Washington DC and Mexico City did not deal with the border problem. This was similar to the problem during the cold war between East and West, but in the Cold War the United States and the West welcomed Eastern Europe with open arms. Many so-called Americans are not so receptive now to accept the children and grandchildren of the Mexican Americans who had their land and resources stolen from them during the American–Mexican war. An open door policy similar to the expanded EU is the only fair, moral, ethical, and social solution of the American-Mexican border. Rather than build another Berlin Wall to keep the Mexican Americans out, the President Barack Obama should tear

down the Bush/Republican Berlin Wall. Ronald Regan had it right when he told the Russians to tear down the Berlin wall and now many Mexican Americans and Americans are totally right when they told Bush and the Republicans and now tell Obama and the Democrats to open up the Mexican–American border. It will be a win win for both countries. Of course the elite in both countries who are responsible for the depressed economic conditions in both countries want a status quo. America can be like Finland (bilingual Swedish and Finnish), Switzerland (trilingual German, French, and Italian), and Belgium (bilingual French and Dutch). Note that the people of these three super economies and socially just countries also learn and are fluent in English. It is high time that Mexico and the United States (and also Canada) open their borders and expand like what EU has. If the EU can add 12 new countries and hopefully in the near future Turkey, then the United States, Canada, and Mexico can open their borders and have not only free trade, but free movement of people, thoughts, ideas, religions, resources, and economic wealth and capital. North Americans should all be Americans and not Mexicans, Canadians, Cubans and US-Americans, just like Europeans are no longer Germans, French, Spanish, Portuguese, Finnish, Swedish, and so on. We need to unite and be one, and not continue to separate and deny our brothers and sisters their birthrights and rights as fellow inhabitants of this planet.

The aim of this chapter has been a multidisciplinary and generalist approach to the problems in quantum nanobiology and how they can and have been tackled and approached by various researchers

in different fields of endeavor: mathematicians, physicists, chemists, and biologists. Clearly in the future, the next generation of scientists, researchers, and even nontechnology working citizens must be better educated with respect to nature and how we and our machines, devices, vehicles, and energy sources interact with the environment and the world we live in, plant and animal life, as well as our most valuable, but taken for granted resources, pure and clean drinking water. The fate of future generations depends on the next generation being made better aware of the known and possible consequences of newly developed and being developed technologies, so that we can not only maintain our high standard of living in the Western societies but also expand it to the largest two developing nations India and China without them both having the same adverse side effects on their native environments and political systems as we have had. The goal and aim to learn from one's mistakes and errors and from those of others means that we must first admit when we have made errors, fully document them in the open literature, and not take draconian measures against honest mistakes that were made due to lack of knowledge. This will only keep the errors hidden and under the surface and prevent the root causes from being determined. In many cases, the errors and/or failures were due to the systems/approaches which were being used and the inherent assumptions. The more open the discussions and more accessible the data is made, the more honest, caring, and open researchers will tend to be. The faster the problems can be recognized, the quicker the solutions can be made. The withdrawal of the wonder drug Vioxx is a classical example of the

new and open system in the field of pharmaceuticals. As soon as problems were recognized the drug was pulled which has allowed top researchers in the whole world to start working on determining and understanding the reasons which caused the adverse side effects, with the hope and goal that the next generation of drug should be even safer than the current generation. Merck is clearly a world leader in not only developing new and better drugs but also in the way to proceed as soon as one recognizes and/or is informed of possible adverse side effects due to one's new drug or technology (small nanoparticles, genetically modified organisms causing allergic reactions for some people).

Not just corporate houses, but each and every scientist, manager, member of the board of directors, and also the Wall Street investment houses need to be more concerned about the long-term sustainability of the world's environment and global standard of living and development instead of focusing only on the short-term bottom line, even if stock options and year end bonuses might suffer and/or be reduced. We owe this to the next generation which will follow us and have to deal with the consequences of our actions, or in many cases, lack of actions, or responses to known and/or possible adverse side effects. We can in many cases no longer claim ignorance. The goal is to rid the world of ignorance and educate not only scientists but also industrialists, politicians, investment bankers, venture capitalists, and of course the citizens and noncitizens globally. A very challenging task, but working with top universities and publishers like John Wiley & Sons, Inc. surely doable, if we put our hearts, souls, and minds to the great task at hand!

Acknowledgments

K.J. Jalkanen (KJJ) would like to thank the Deutscher Akademischer Austausch Dienst (DAAD) [German Academic Exchange Service], the Deutsches Krebsforschungszentrum (DKFZ) [German Cancer Research Center] in Heidelberg and the EU for making his five year stay at the German Cancer Research Center and the Steinbeiss Transfer Centrum possible. In addition, he would like to thank the Finnish Academy of Science for sponsoring many visits to Professor Risto Nieminen's Laboratory of Physics Laboratory in Helsinki University of Technology in Espoo, Finland. We would like to thank the Danish Foundation for Fundamental Research (www.dg.dk) for funding the establishment of the Quantum Protein (QuP) Centre in the Department of Physics at the Technical University of Denmark where KJJ has been an associate professor of biophysics before moving to the Nanochemistry Research Institute at Curtin University of Technology where he is currently on unpaid sabbatical leave from QuP and being funded by the Western Australian Premier's Research Fellowship program (grant awarded to Professor J.D. Gale, Western Australian Premier's Research Fellow). KJJ would also like to acknowledge and thank iVEC (The hub of advanced computing in Western Australia) in Perth, APAC (Australian Partnership for Advanced Computing) in Canberra, and the ARC for providing computational, people and financial resources which made parts of this work possible during KJJ's three-year sabbatical stay in Perth at Curtin University of Technology in the Nanochemistry Research Institute (NRI)

in the Department of Applied Chemistry. Finally KJJ would like to thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and the DAAD for financial support for his five months stay in the laboratory of Professor Thomas Frauenheim at the Bremen Center for Computational Material Science at the University of Bremen.

References

- Abbas, H. (2005) *Pakistan's Drift into Extremism: Allah, the Army, and America's War on Terror*, M.E. Sharpe.
- Abdali, S., Jalkanen, K.J., Bohr, H., Suhai, S., and Nieminen, R.M. (2002) *Chem. Phys.*, **282**, 219.
- Amos, R.D. (1982) *Chem. Phys. Lett.*, **87**, 23.
- Amos, R.D. (1984) *Chem. Phys. Lett.*, **108**, 185.
- Amos, R.D. (1986) *Chem. Phys. Lett.*, **124**, 376.
- Amos, R.D. (1987) Molecular property derivatives, in *Ab Initio Methods in Quantum Chemistry*, John Wiley and Sons Ltd, pp. 99–153.
- Amos, R.D., Handy, N.C., Jalkanen, K.J., and Stephens, P.J. (1987) *Chem. Phys. Lett.*, **133**, 21.
- Amos, R.D., Jalkanen, K.J., and Stephens, P.J. (1988) *J. Phys. Chem.*, **92**, 5571.
- Atkins, P.W., and Barron, L.D. (1969) *Mol. Phys.*, **16**, 453.
- Ayafor, J.F., Tchuedem, M.H.K., Nyasse, B., Tillequin, F., and Anke, H. (1994) *Pure Appl. Chem.*, **66**, 2327.
- Baello, B.I., Pancoska, P., and Keiderling, T.A. (1997) *Anal. Biochem.*, **250**, 212.
- Bak, K.L., Devlin, F.J., Ashvar, C.S., Taylor, P.R., Frisch, M.J., and Stephens, P.J. (1995) *J. Phys. Chem.*, **99**, 14918.
- Bak, K.L., Jørgensen, P., Helgaker, T., Ruud, K., and Jensen, H.J.Aa. (1993) *J. Chem. Phys.*, **98**, 8873.
- Bak, K.L., Jørgensen, P., Helgaker, T., Ruud, K., and Jensen, H.J.Aa. (1994) *J. Chem. Phys.*, **100**, 6621.
- Barron, L.D. (1982) *Molecular Light Scattering and Optical Activity*, 1st edn, Cambridge University Press.

- Barron, L.D. (2004) *Molecular Light Scattering and Optical Activity*, 2nd edn, Cambridge University Press.
- Barron, L.D., Bogaard, M.P., and Buckingham, A.D. (1973) *J. Am. Chem. Soc.*, **95**, 603.
- Barron, L.D., and Buckingham, A.D. (1971) *Mol. Phys.*, **20**, 1111.
- Barron, L.D., and Buckingham, A.D. (2001) *Acc. Chem. Res.*, **34**, 781.
- Barron, L.D., Ford, S.J., Bell, A.F., Wilson, G., Hecht, L., and Cooper, A. (1994) *Faraday Discuss.*, **99**, 217.
- Barron, L.D., Gargaro, A.R., Hecht, L., and Polavarapu, P.L. (1991a) *Spectrochim. Acta, Part A*, **47**, 1001.
- Barron, L.D., Gargaro, A.R., Hecht, L., and Polavarapu, P.L. (1991b) *Spectrochim. Acta, Part A*, **48A**, 261.
- Barron, L.D., Hecht, L., Blanch, E.W., and Bell, A.F. (2000) *Prog. Biophys. Mol. Biol.*, **73**, 1.
- Bartlett, R.J., Grabowski, I., Hirata, S., and Ivanov, S. (2005) *J. Chem. Phys.*, **122**, 034104.
- Bartlett, R.J., Lotrich, V.F., and Schwiegert, I.V. (2005) *J. Chem. Phys.*, **123**, 062205.
- Baumruk, V., Pancoska, P., and Keiderling, T.A. (1996) *J. Mol. Biol.*, **259**, 774.
- Becke, A.D. (1988a) *Phys. Rev., A*, **38**, 3098.
- Becke, A.D. (1988b) *J. Chem. Phys.*, **88**, 1053.
- Becke, A.D. (1993a) *J. Chem. Phys.*, **98**, 1372.
- Becke, A.D. (1993b) *J. Chem. Phys.*, **98**, 5648.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., and Hermans, J. (1981) Interaction models for water in relation to protein hydration, in *Intermolecular Forces*, Reidel, Dordrecht, Holland, pp. 331–42.
- Blom, C.E., and Altona, C. (1976) *Mol. Phys.*, **31**, 1377.
- Böhm, H.-J., and Klebe, G. (1996) *Angew. Chem. Int.*, **35**, 2588.
- Bohr, H., Frimand, K., Jalkanen, K.J., Nieminen, R.M., and Suhai, S. (2001) *Phys. Rev. E*, **64**, 021905.
- Bohr, H., and Jalkanen, K. (1998) *Condens. Matter Theor.*, **13**, 95.
- Bohr, H.G., and Jalkanen, K.J. (2005) *Condens. Matter Theor.*, **20**, 375.
- Bohr, H.G., Jalkanen, K.J., Frimand, K., Elstner, M., and Suhai, S. (1999) *Chem. Phys.*, **246**, 13.
- Bohr, H., Jalkanen, K.J., and Malik, F.B. (2005) *Mod. Phys. Lett.*, **19**, 473.
- Bombasaro, J.A., Rodriguez, A.M., and Enriz, R.D. (2005) *J. Mol. Struct.: Theochem.*, **724**, 173.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983) *J. Comp. Chem.*, **4**, 187.
- Brooks III, C.L., and Case, D.A. (1993) *Chem. Rev.*, **93**, 2487.
- Bruno, I.J., Cole, J.C., Lommerse, J.P.M., Rowland, R.S., Taylor, R., and Verdonk, M.L. (1997) *J. Comput. Aided Mol. Des.*, **11**, 525.
- Buckingham, A.D. (1967) *Adv. Chem. Phys.*, **12**, 107.
- Buckingham, A.D., Fowler, P.W., and Galwas, P.A. (1987) *J. Chem. Phys.*, **112**, 1.
- Cao, X., Dukor, R.K., and Nafie, L.A. (2008) *Theor. Chem. Acc.*, **119**, 69.
- Cavasotto, C.N., and Abagyan, R.A. (2004) *J. Mol. Biol.*, **337**, 209.
- Charles, S.W., and Lee, K.O. (1965) *Trans. Faraday Soc.*, **61**, 2081.
- Cheeseman, J.R., and Frisch, M.J. (2008) in *First International Conference on Vibrational Optical Activity*, **24**, (eds E. Blanch, and S. Abdali), Manchester Interdisciplinary Biocentre, University of Manchester Press, Manchester.
- Cheesman, J.R., Frisch, M.J., Devlin, F.J., and Stephens, P.J. (1996) *Chem. Phys. Lett.*, **252**, 211.
- Cho, M. (2008) *Chem. Rev.*, **108**, 1331.
- Chou, C.K. (1990) *Clin. Biochem.*, **23**, 37.
- Collier, W.B., Magdo, I., and Klotz, T.D. (1999) *J. Chem. Phys.*, **110**, 5710.
- Colwell, S.M., Murray, C.W., Handy, N.C., and Amos, R.D. (1993) *Chem. Phys. Lett.*, **210**, 261.
- Colwell, S.M., and Percival, I.C. (1983a) *Chem. Phys.*, **75**, 215.
- Colwell, S.M., and Percival, I.C. (1983b) *Chem. Phys.*, **75**, 225.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz Jr, K.M., Fergusson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179.
- Cossi, M., Barone, V., Cammi, R., and Tomasi, J. (1996) *Chem. Phys. Lett.*, **255**, 327.
- Cotton, F.A. (1990) *Chemical Applications of Group Theory*, 3rd edn, John Wiley & Sons, New York.
- Craw, J.S., Guest, J.M., Cooper, M.D., Burton, N.A., and Hillier, I.H. (1996) *J. Phys. Chem.*, **100**, 6304.

- Davis, A.M., and Teague, S.J. (1999) *Angew. Chem. Int. Ed.*, **38**, 736.
- Degtyarenko, I.M., Jalkanen, K.J., Gurtovenko, A.A., and Nieminen, R.M. (2007) *J. Phys. Chem. B*, **111**, 4227.
- Degtyarenko, I.M., Jalkanen, K.J., Gurtovenko, A.A., and Nieminen, R.M. (2008) *J. Comp. Theor. Nanosci.*, **5**, 277.
- Deng, Z., Polavarapu, P.L., Ford, S.J., Hecht, L., Barron, L.D., Ewig, C.S., and Jalkanen, K.J. (1996) *J. Phys. Chem.*, **100**, 2025.
- Deplazes, E., van Bronswijk, B., Zhu, F., Barron, L.D., Ma, S., Nafie, L.A., and Jalkanen, K.J. (2008) *Theor. Chem. Acc.*, **119**, 155.
- Deslongchamps, P. (1983) *Stereoelectronic Effects in Organic Chemistry*, Volume 1 of Organic Chemistry Series, Pergamon Press, Potts Point.
- Dewar, J.S., Zoebisch, E., Healy, E.F., and Stewart, J.J.P. (1985) *J. Am. Chem. Soc.*, **107**, 3902.
- Diem, M. (1988) *J. Am. Chem. Soc.*, **110**, 6967.
- Diem, M. (1993) *Introduction to Modern Vibrational Spectroscopy*, John Wiley and Sons Ltd, New York.
- Diem, M., Polavarapu, P.L., Oboodi, M., and Nafie, L.A. (1982) *J. Am. Chem. Soc.*, **104**, 3329.
- Dudek, M.J., and Ponder, J.W. (1995) *J. Comp. Chem.*, **16**, 791.
- Elstner, M., Jalkanen, K.J., Knapp-Mohammady, M., Frauenheim, Th., and Suhai, S. (2000) *Chem. Phys.*, **256**, 15.
- Elstner, M., Jalkanen, K.J., Knapp-Mohammady, M., Frauenheim, Th., and Suhai, S. (2001) *Chem. Phys.*, **263**, 203.
- Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, Th., Suhai, S., and Seifert, G. (1998) *Phys. Rev. B*, **58**, 7260.
- Ertl, G. (2008) *Angew. Chem. Int. Ed.*, **47**, 2.
- Farid, B. (1999) *Electron Correlation in the Solid State, Chapter Ground and Low-lying Excited States of Interacting Electron Systems; A Survey and Some Critical Analyses*, World Scientific, Imperial College, pp. 103–262.
- Farid, B. (2002) *Philos. Mag. B*, **82**, 1413.
- Fillaux, F. (1999) *J. Mol. Struct.*, **511–512**, 35.
- Fillaux, F. (2004) Vibrational spectroscopy and quantum localization, in *Energy Localisation and Transfer*, Volume 22 of Advanced Series in Nonlinear Dynamics, World Scientific, pp. 73–148.
- Fillaux, F., Fontaine, J.P., Baron, M.-H., Kearley, G.J., and Tomkinson, J. (1993) *Chem. Phys.*, **176**, 249.
- Fillaux, F., Nicolai, B., Baron, M.H., Lautié, A., Tomkinson, J., and Kearley, G.J. (1998) *Ber. Bunsenges. Phys. Chem.*, **102**, 384.
- Fischer, P., and Hache, F. (2005) *Chirality*, **17**, 421.
- Floris, F., Persico, M., Tani, A., and Tomasi, J. (1992) *Chem. Phys. Lett.*, **199**, 518.
- Fortunelli, A. (1996) *Chem. Phys. Lett.*, **248**, 50.
- Fortunelli, A., and Tomasi, J. (1994) *Chem. Phys. Lett.*, **231**, 34.
- Freedman, T.B., Chernovitz, A.C., Zuk, W.M., Paterlini, M.G., and Nafie, L.A. (1988) *J. Am. Chem. Soc.*, **110**, 6970.
- Frimand, K., and Jalkanen, K.J. (2002) *Chem. Phys.*, **279**, 161.
- Frimand, K., Jalkanen, K.J., Bohr, H.G., and Suhai, S. (2000) *Chem. Phys.*, **255**, 165.
- Frish, M.J., Yamaguchi, Y., Gaw, J.F., Schaefer, H.F., and Binkley III, J.S. (1986) *J. Chem. Phys.*, **84**, 531.
- Gale, J.D., and Rohl, A.L. (2007) *Mol. Sim.*, **33**, 1237.
- Gambetta, D. (ed.) (2005) *Making Sense of Suicide Missions*, Oxford University Press.
- Gay, D.H., and Rohl, A.L. (1995) *J. Chem. Soc., Faraday Trans.*, **91**, 925.
- Godby, R.W., Schlüter, M., and Sham, L.J. (1988) *Phys. Rev. B*, **37**, 10159.
- Goodsell, D.S. (2002) The future of bionanotechnology, in *Biotechnology: Lessons from Nature*, John Wiley & Sons, Inc., Berlin, pp. 295–311.
- Gutt, C., Baumert, J., Press, W., Tse, J.S., and Janssen, J.S. (2002) *J. Chem. Phys.*, **116**, 3795.
- Halls, M.D., and Schlegel, H.B. (1999a) *J. Chem. Phys.*, **111**, 8819.
- Halls, M.D., and Schlegel, H.B. (1999b) *J. Chem. Phys.*, **109**, 10587.
- Hammerschmidt, E.G. (1934) *Ind. Eng. Chem.*, **26**, 851.
- Han, W.-G., Jalkanen, K.J., Elstner, M., and Suhai, S. (1998) *J. Phys. Chem. B*, **102**, 2587.
- Han, W.-G., Jalkanen, K.J., Elstner, M., Frauenheim, Th., and Suhai, S. (2000) *Int. J. Quantum Chem.*, **78**, 459.
- Handa, Y.P. (1986a) *J. Chem. Thermodyn.*, **18**, 891.
- Handa, Y.P. (1986b) *J. Chem. Thermodyn.*, **18**, 915.

- Hanessian, S. (1983) *Total Synthesis of Natural Products: The 'Chiron' Approach*, Volume 3 of Organic Chemistry Series, Pergamon Press, Potts Point.
- Hao, C., Whittaker, L., and Whittaker, J. (2006) *Biochem. Biophys. Res. Commun.*, **347**, 334.
- Hecht, L., and Barron, L.D. (1990) *Appl. Spectrosc.*, **44**, 483.
- Hecht, L., Barron, L.D., and Hug, W. (1989) *Chem. Phys. Lett.*, **158**, 341.
- Helgaker, T., Ruud, K., Bak, K.L., Jørgensen, P., and Olsen, J. (1994) *Faraday Discuss.*, **99**, 165.
- Hohenberg, P., and Kohn, W. (1964) *Phys. Rev.*, **136**, 864.
- Horikawa, S., Itoh, H., Tabata, J., Kawamura, K., and Hondoh, T. (1997) *J. Phys. Chem. B*, **101**, 6290.
- Hornicek, J., Kapralova, P., and Bour, P. (2008) *J. Chem. Phys.*, **127**, 084502.
- Inoue, R., Tanaka, H., and Nakanishi, K. (1996) *J. Chem. Phys.*, **104**, 9569.
- Ioannou, A.G., Colwell, S.M., and Amos, R.D. (1997) *Chem. Phys. Lett.*, **278**, 278.
- Ipolyi, I., Cicman, P., Deniff, S., Mtejcik, V., Mach, P., Urban, J., Scheier, P., Mark, T.D., and Matejcik, S. (2006) *Int. J. Mass Spectrom.*, **252**, 228.
- Ismail-Beigi, S., and Louie, S.G. (2003) *Phys. Rev. Lett.*, **90**, 076401.
- Itoh, H., Chazallon, B., Schober, H., Kawamura, K., and Kuhs, W.F. (2003) *Can. J. Phys.*, **81**, 493.
- Jalkanen, K.J., Bohr, H.G., Suhai, S. (1997) in *Proceedings of the International Symposium on Theoretical and Computational Methods in Genome Research*, (S. Suhai), Plenum Press, New York, pp. 255–77.
- Jalkanen, K.J., Degtyarenko, I.M., Nieminen, R.M., Cao, X., Nafie, L.A., Zhu, F., and Barron, L.D. (2008a) *Theor. Chem. Acc.*, **119**, 191.
- Jalkanen, K.J., Gale, J.D., Jalkanen, G.J., McIntosh, D., El Azhary, A., and Jensen, G.M. (2008b) *Theor. Chem. Acc.*, **119**, 211.
- Jalkanen, K.J., Gale, J.D., Lassen, P.R., Hemmingsen, L., Rodarte, A., Degtyarenko, I.M., Nieminen, R.M., Christensen, S.B., Knapp-Mohammady, M., and Suhai, S. (2008c) *Theor. Chem. Acc.*, **119**, 177.
- Jalkanen, K.J., Gale, J.D., Rohl, A.L., Zhu, F., Barron, L.D., Ma, S., and Nafie, L.A. (2009a) *Theor. Chem. Acc.* (Suhai Festschrift) manuscript in preparation.
- Jalkanen, K.J., Jürgensen, V.W., Claussen, A., Jensen, G.M., Rahim, A., Wade, R.C., Nardi, F., Jung, C., Nieminen, R.M., Degtyarenko, I.M., Herrmann, F., Knapp-Mohammady, M., Niehaus, T., Frimand, K., and Suhai, S. (2006a) *Int. J. Quantum Chem.*, **106**, 1160.
- Jalkanen, K.J., Jürgensen, V.W., and Degtyarenko, I.M. (2006b) *Adv. Quantum Chem.*, **50**, 91.
- Jalkanen, K.J., Knapp-Mohammady, M., Suhai, S., Nardi, F., Wade, R.C., Bohr, J., Bohr, H.G., Rohl, A.L., and Gale, J.D. (2009b) *Theor. Chem. Acc.* (Suhai Festschrift) manuscript in preparation.
- Jalkanen, K.J., Nardi, F., Wade, R.C., Suhai, S., Rohl, A.L., Gale, J.D., and Jung, C. (2009c) *Theor. Chem. Acc.* (Suhai Festschrift) manuscript in preparation.
- Jalkanen, K.J., Nieminen, R.M., Frimand, K., Bohr, J., Bohr, H., Wade, R.C., Tajkhorshid, E., and Suhai, S. (2001) *Chem. Phys.*, **265**, 125.
- Jalkanen, K.J., Nieminen, R.M., Knapp-Mohammady, M., and Suhai, S. (2003) *Int. J. Quantum Chem.*, **92**, 239.
- Jalkanen, K.J., and Stephens, P.J. (1991) *J. Phys. Chem.*, **95**, 5446.
- Jalkanen, K.J., Stephens, P.J., Amos, R.D., and Handy, N.C. (1987) *J. Am. Chem. Soc.*, **109**, 7193.
- Jalkanen, K., Stephens, P.J., Amos, R., and Handy, N. (1988) *J. Am. Chem. Soc.*, **110**, 2012.
- Jalkanen, K.J., and Suhai, S. (1996) *Chem. Phys.*, **208**, 81.
- Johnson, B.G., and Florian, J. (1995) *Chem. Phys. Lett.*, **247**, 120.
- Johnson, B.G., and Frisch, M. (1994) *J. Chem. Phys.*, **100**, 7429.
- Jorgensen, W.L., Madura, J.D., and Swenson, C.J. (1984) *J. Am. Chem. Soc.*, **106**, 6638.
- Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996) *J. Am. Chem. Soc.*, **118**, 11225.
- Jorgensen, W.L., and Swenson, C.J. (1985a) *J. Am. Chem. Soc.*, **107**, 569.
- Jorgensen, W.L., and Swenson, C.J. (1985b) *J. Am. Chem. Soc.*, **107**, 1489.
- Jorgensen, W.L., and Tirado-Rives, J. (1988) *J. Am. Chem. Soc.*, **110**, 1657.
- Jung, C. (2000) *J. Mol. Recognit.*, **13**, 325.
- Jürgensen, V.W., and Jalkanen, K.J. (2006) *Phys. Biol.*, **3**, S63.

- Kawiecki, R.W., Devlin, F.J., Stephens, P.J., and Amos, R.D. (1991) *J. Phys. Chem.*, **95**, 9817.
- Kearley, G.J. (1986) *J. Chem. Soc., Faraday Trans. 2*, **82**, 41.
- Kearley, G.J. (1995) *Nucl. Instrum. Methods Phys. Res. A*, **354**, 53.
- Kearley, G.J., Fillaux, F., Baron, M.H., Bennington, S., and Tomkinson, J. (1994) *Science*, **264**, 1285.
- Kearley, G.J., Johnson, M.R., Plazanet, M., and Suard, E. (2001) *J. Chem. Phys.*, **115**, 2614.
- Kearley, G.J., Tomkinson, J., and Penfold, J. (1987) *Z. Phys. B*, **69**, 63.
- Keiderling, T.A. (1996) Protein structural studies using vibrational circular dichroism spectroscopy, in *Spectroscopic Methods For Determining Protein Structure in Solution*, Chapter 8, VCH Publishers, Inc., Weinheim, pp. 163–89.
- Keskin, O., Gursoy, A., Ma, B., and Nussinov, R. (2008) *Chem. Rev.*, **108**, 1225.
- Kim-Cohen, J., Caspi, A., Taylor, A., Williams, B., Newcombe, R., Craig, I.W., and Moffitt, T.E. (2006) *Mol. Psychiatry*, **11**, 903.
- Klopman, G. (1964a) *J. Am. Chem. Soc.*, **86**, 1463.
- Klopman, G. (1964b) *J. Am. Chem. Soc.*, **86**, 4550.
- Knapp-Mohammady, M., Jalkanen, K.J., Nardi, F., Wade, R.C., and Suhai, S. (1999) *Chem. Phys.*, **240**, 63.
- Kohn, W. (1999) *Rev. Mod. Phys.*, **71**, 1253.
- Kohn, W., and Sham, L.J. (1965) *Phys. Rev.*, **140**, 1133.
- Kollman, P.A. (1992) *Curr. Opin. Struct. Biol.*, **2**, 765.
- Komornicki, A., and Fitzgerald, G. (1993) *J. Chem. Phys.*, **98**, 1398.
- Komornicki, A., and McIver, J.W. (1979) *J. Chem. Phys.*, **70**, 2014.
- Kongsted, J., and Ruud, K. (2008) *Chem. Phys. Lett.*, **451**, 226.
- Kubelka, J., Pancoska, P., and Keiderling, T.A. (1999) *Appl. Spectrosc.*, **53**, 666.
- Kulcitki, V., Ungur, N., Gavagnin, M., Carbone, M., and Cimino, G. (2005) *Eur. J. Org. Chem.*, **2005** (9), 1816–22.
- Lannoo, M., Schlüter, M., and Sham, L.J. (1985) *Phys. Rev. B*, **32**, 3890.
- Lee, A.M., Colwell, S.M., and Handy, N.C. (1994) *Chem. Phys. Lett.*, **229**, 225.
- Lee, C., and Sosa, C. (1994) *J. Chem. Phys.*, **100**, 9018.
- Lee, C., Yang, W., and Parr, R.G. (1988) *Phys. Rev. B*, **37**, 785.
- Levine, I.N. (2000) *Quantum Chemistry*, 5th edn, Prentice Hall, Upper Saddle River.
- Levine, Z.H., and Allan, D.C. (1989) *Phys. Rev. Lett.*, **63**, 1719.
- Liegeois, V. (2008) Ph.D. thesis Elaboration of Quantum Chemistry Methods for Predicting and Interpreting Vibrational Raman Optical Activity: Applications to Helical Structures, Facultes Universitaires Notre-Dame de la Paix, Namur.
- Liegeois, V., and Champagne, B. (2008) in *First International Conference on Vibrational Optical Activity*, **25**, (eds E. Blanch, and S. Abdali), Manchester Interdisciplinary Biocentre, University of Manchester Press, Manchester.
- Lord, R.C., and Yu, N.-T. (1970a) *J. Mol. Biol.*, **50**, 509.
- Lord, R.C., and Yu, N.-T. (1970b) *J. Mol. Biol.*, **51**, 203.
- Losada, M., Tran, H., and Xu, Y. (2008) *J. Chem. Phys.*, **128**, 014508.
- Losada, M., and Xu, Y. (2007) *Phys. Chem. Chem. Phys.*, **9**, 3127.
- Lowe, M.A., Alper, J.S., Kawiecki, R.W., and Stephens, P.J. (1986) *J. Phys. Chem.*, **90**, 41.
- Ma, S.-K., and Brueckner, K.A. (1968) *Phys. Rev.*, **165**, 18.
- MacKerell, A.D.J., Bashford, D., Bellott, M., Dunbrack, J.R.L., Evanseck, J.D., Field, M.J., Fisher, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kucyera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem. B*, **102**, 3586.
- Malkin, V.G., Malkina, O.L., Eriksson, L.A., and Salahub, D.R. (1995) The calculation of NMR and ESR spectroscopy parameters using density functional theory, *Modern Density Functional Theory: A Tool for Chemistry*, Volume 2 of Theoretical and Computational Chemistry, Elsevier Science B.V., pp. 273–347.
- Malkin, V.G., Malkina, O.L., and Salahub, D.R. (1993a) *Chem. Phys. Lett.*, **204**, 80.
- Malkin, V.G., Malkina, O.L., and Salahub, D.R. (1993b) *Chem. Phys. Lett.*, **204**, 87.

- Mantz, Y.A., Gerard, H., Iftimie, R., and Martyna, G.J. (2006) *J. Phys. Chem. B*, **110**, 13523.
- Maple, J.R., Hwang, M.-J., Jalkanen, K.J., Stockfisch, T.P., and Hagler, A.T. (1998) *J. Comp. Chem.*, **19**, 430.
- Matsuda, S.P.T., Wilson, W.K., and Xiong, Q. (2006) *Org. Biomol. Chem.*, **4**, 530.
- Mermin, N. (1965) *Phys. Rev.*, **137**, A1441.
- Mok, D.K., Nwumann, R., and Handy, N.C. (1996) *J. Phys. Chem.*, **100**, 6225.
- Moon, C., Taylor, P.C., and Rodger, P.M. (2003) *Can. J. Chem.*, **81**, 451.
- Mukamel, S. (1995) *Principles of Nonlinear Optical Spectroscopy*, Oxford University Press, Oxford.
- Mukhopadhyay, P., Zuber, G., Wipf, P., and Beratan, D.N. (2007) *Angew. Chem. Int. Ed.*, **46**, 6450.
- Nafie, L.A. (2008) *Theor. Chem. Acc.*, **119**, 39.
- Nardi, F. (1999) Ph.D. thesis Molecular Dynamics and NMR Spectroscopic Studies of Local Interactions of Small Peptides in Aqueous Solution Involving Aromatic Residues, Universite Joseph Fourier, Grenoble.
- Nardi, F., Kemmink, J., Sattler, M., and Wade, R.C. (2000) *J. Biomol. NMR*, **17**, 63.
- Nardi, F., Worth, G.A., and Wade, R.C. (1997) *Fold. Des.*, **2**, S62.
- Newman, A.A. (ed.) (1972) *Chemistry of Terpenes and Terpenoids*, Academic Press Inc., London.
- Nicu, V.P., Neugebauer, J., Wolff, S.K., and Baerends, E.J. (2008) *Theor. Chem. Acc.*, **119**, 245.
- Oliveira, F.G., Sant'Anna, C.M.R., Caffarena, E.R., Dardenne, L.E., and Barreiro, E.J. (2006) *Bioorg. Med. Chem.*, **14**, 6001.
- Onida, G., Reining, L., and Rubio, A. (2002) *Rev. Mod. Phys.*, **74**, 601.
- Onsager, L. (1936) *J. Am. Chem. Soc.*, **58**, 1486.
- Ösapay, K., Young, W.S., Bradford, D., Brooks, C.L., and Case, D.A. (1996) *J. Phys. Chem.*, **100**, 2698.
- Oyama, H., Ebinuma, T., Shimada, W., Takeya, S., Nagao, J., Uchida, T., and Narita, H. (2003) *Can. J. Phys.*, **81**, 485.
- Pancoska, P., Bitto, E., Janota, V., and Keiderling, T.A. (1994) *Faraday Discuss. Chem. Soc.*, **99**, 287.
- Pancoska, P., Bitto, E., Janota, V., Urbanova, M., Gupta, V.P., and Keiderling, T.A. (1995) *Protein Sci.*, **4**, 1384.
- Pancoska, P., Blazek, M., and Keiderling, T.A. (1992) *Biochemistry*, **31**, 10250.
- Pancoska, P., Janota, V., and Keiderling, T.A. (1996) *Appl. Spectrosc.*, **50**, 658.
- Pancoska, P., Kubelka, J., and Keiderling, T.A. (1999) *Appl. Spectrosc.*, **53**, 655.
- Pancoska, P., Yasui, S.C., and Keiderling, T.A. (1989) *Biochemistry*, **28**, 5917.
- Pancoska, P., Yasui, S.C., and Keiderling, T.A. (1991) *Biochemistry*, **30**, 5089.
- Papai, I., St-Alant, A., Ushio, J., and Salahub, D. (1990) *Int. J. Quantum Chem. Quantum Chem. Symp.*, **24**, 29.
- Perdew, J.P., Burke, K., and Ernzerhof, M. (1996) *Phys. Rev. Lett.*, **77**, 3865.
- Perdew, J.P., Burke, K., and Ernzerhof, M. (1997) *Phys. Rev. Lett.*, **78**, 1396.
- Perdew, J.P., and Levy, M. (1983) *Phys. Rev. Lett.*, **51**, 1884.
- Perdew, J.P., and Wang, Y. (1992) *Phys. Rev. B*, **45**, 13244.
- Pimentel, G.C. (1975) *Angew. Chem. Int.*, **14**, 199.
- Pimentel, G.C., and Charles, S.W. (1963) *Pure Appl. Chem.*, **7**, 111.
- Pimentel, G.C., and Sederholm, C.H. (1956) *J. Chem. Phys.*, **24**, 639.
- Pohl, G., Perczel, A., Vass, E., Magyarfalvi, G., and Tarczay, G. (2007) *Phys. Chem. Chem. Phys.*, **9**, 4698.
- Polavarapu, P.L. (1990) *J. Phys. Chem.*, **94**, 8106.
- Poon, C.-D., Samulski, E.T., Weise, C.F., and Weisshaar, J.C. (2000) *J. Am. Chem. Soc.*, **122**, 5642.
- Pople, J.A. (1999) *Rev. Mod. Phys.*, **71**, 1267.
- Pulay, P., Fogarasi, G., Pang, F., and Boggs, J.E. (1979) *J. Am. Chem. Soc.*, **101**, 2550.
- Rajagopal, A.K., and Callaway, J. (1973) *Phys. Rev. B*, **7**, 1912.
- Ramnarayan, K., Bohr, H.G., and Jalkanen, K.J. (2008) *Theor. Chem. Acc.*, **119**, 265.
- Rauhut, G., Puyear, S., Wolinski, K., and Pulay, P. (1996) *J. Phys. Chem.*, **100**, 6310.
- Reining, L., Olevano, V., Rubio, A., and Onida, G. (2002) *Phys. Rev. Lett.*, **88**, 066404.
- Richmond, T.J. (1984) *J. Mol. Biol.*, **178**, 63.
- Ruud, K. (2008) in *First International Conference on Vibrational Optical Activity*, 28, (eds E. Blanch, and S. Abdali), Manchester Interdisciplinary Biocentre, University of Manchester Press, Manchester.

- Sauter, C., Otalora, F., Gavira, J.-A., Vidal, O., Giege, R., and Garcia-Ruiz, J.M. (2001) *Acta Crystallogr. D*, **57**, 1119.
- Schiffer, C.A., Caldwell, J.W., Stroud, R.M., and Kollman, P.A. (1992) *Protein Sci.*, **1**, 396.
- Schmidt, A.B., and Fine, R.M. (1994) *Mol. Sim.*, **13**, 347.
- Schreckenback, G., and Ziegler, T. (1995) *J. Phys. Chem.*, **99**, 606.
- Scott, A. (2003) *Nonlinear Science*, 2nd edn, Oxford University Press, Oxford.
- Serdyuk, I.N., Zaccai, N.R., and Zaccai, J. (2007) *Methods in Molecular Biophysics: Structure, Dynamics, Function*, Cambridge University Press, Cambridge.
- Sham, L.J., and Schlüter, M. (1985) *Phys. Rev.*, **B**, **32**, 3883.
- Stephens, P.J. (1985) *J. Phys. Chem.*, **89**, 748.
- Stephens, P.J. (1987) *J. Phys. Chem.*, **91**, 1712.
- Stephens, P.J., Chabalowski, C., Devlin, F., and Jalkanen, K.J. (1993) *Chem. Phys. Lett.*, **225**, 247.
- Stephens, P.J., Devlin, F.J., Ashvar, C.S., Chabalowski, C.F., and Frisch, M.J. (1994) *Faraday Discuss.*, **99**, 103.
- Stephens, P.J., Devlin, F.J., Chabalowski, C., and Frisch, M.J. (1994) *J. Phys. Chem.*, **98**, 11623.
- Stephens, P.J., Jalkanen, K.J., Amos, R.D., Lazzaretti, P., and Zanasi, R. (1990) *J. Phys. Chem.*, **94**, 1811.
- Stephens, P.J., and Lowe, M.A. (1985) *Ann. Rev. Phys. Chem.*, **36**, 213.
- Stewart, J.J.P. (1989) *J. Comp. Chem.*, **10**, 209.
- Stewart, J.J.P. (2007) *J. Mol. Model.*, **13**, 1173.
- Stirling, A. (1996) *J. Chem. Phys.*, **104**, 1254.
- Subramanian, S., and Sloan Jr, E.D. (2002) *J. Phys. Chem. B*, **106**, 4348.
- Tajkhorshid, E., Jalkanen, K.J., and Suhai, S. (1998) *J. Phys. Chem. B*, **102**, 5899.
- Tawa, G.J., Martin, R.L., Pratt, L.R., and Russo, T.V. (1996) *J. Phys. Chem.*, **100**, 1515.
- te Velde, G., Bickelhaupt, F.M., Baerends, E.J., Guerra, C.F., van Gisbergen, S.J.A., Snijders, J.G., and Ziegler, T. (2001) *J. Comp. Chem.*, **22**, 931.
- Thar, J., Zhan, S., and Kirchner, B. (2008) *J. Phys. Chem. B*, **112**, 1456.
- Tomasi, J., Mennucci, B., and Cammi, R. (2005) *Chem. Rev.*, **105**, 2999.
- Tomasi, J., and Persico, M. (1994) *Chem. Rev.*, **94**, 2027.
- Tse, J.S., Shpakov, V.P., Belosludov, V.R., Trouw, F., Handa, Y.P., and Press, W. (2001) *Europhys. Lett.*, **54**, 354.
- van Caillie, C., and Amos, R.D. (2000) *Phys. Chem. Chem. Phys.*, **2**, 2123.
- van Gunsteren, W.F., Berendsen, H.J.C., Hermans, J., Hol, W.G.J., and Postma, J.P.M. (1983) *Proc. Nail. Acad. Sci.*, **80**, 4315.
- van Klaveren, E.P., Michels, J.P.J., Schouten, J.A., Klug, D.D., and Tse, J.S. (2001) *J. Chem. Phys.*, **114**, 5745.
- Vatamanu, J., and Kusalik, P.G. (2008) *J. Phys. Chem. B*, **112**, 2399.
- Vignale, G., and Rasolt, M. (1987) *Phys. Rev. Lett.*, **59**, 2360.
- Vignale, G., and Rasolt, M. (1988) *Phys. Rev.*, **B**, **37**, 10685.
- Vignale, G., Rasolt, M., and Geldart, D.J.W. (1988) *Phys. Rev.*, **B**, **37**, 2502.
- Warshel, A., and Levitt, M. (1976) *J. Mol. Biol.*, **103**, 227.
- Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., and Weiner Jr, P. (1984) *J. Am. Chem. Soc.*, **106**, 765.
- Weiner, S.J., Kollman, P.A., Nguyen, D.T., and Case, D.A. (1986) *J. Comp. Chem.*, **7**, 230.
- Weise, C.F., and Weisshaar, J.C. (2003) *J. Phys. Chem. B*, **107**, 3265.
- Wesson, L., and Eisenberg, D. (1992) *Protein Sci.*, **1**, 227.
- Weyerstahl, P., Schwieger, R., Schwöpe, I., and Hashem, Md.A. (1995) *Liebigs Ann.*, (7), 1389–92.
- Whittle, E., Dows, D.A., and Pimentel, G.C. (1955) *J. Chem. Phys.*, **22**, 1943.
- Wilson, P.J., Amos, R.D., and Handy, N.C. (1999) *Chem. Phys. Lett.*, **312**, 475.
- Wolff, S.K. (2005) *Int. J. Quantum Chem.*, **104**, 645.
- Yu, G.-S., Freedman, T.B., Nafie, L.A., Deng, T., and Polavarapu, P.L. (1995) *J. Phys. Chem.*, **99**, 835.
- Zeng, H., Wilson, L.D., Walker, V.K., and Ripmeester, J.A. (2003) *Can. J. Phys.*, **81**, 17.
- Zhu, F., Isaacs, N.W., Hecht, L., and Barron, L.D. (2005) *Structure*, **13**, 1409.
- Zhukov, V.P., Chulkov, E.V., and Echenique, P.M. (2004) *Phys. Rev. Lett.*, **93**, 096401.